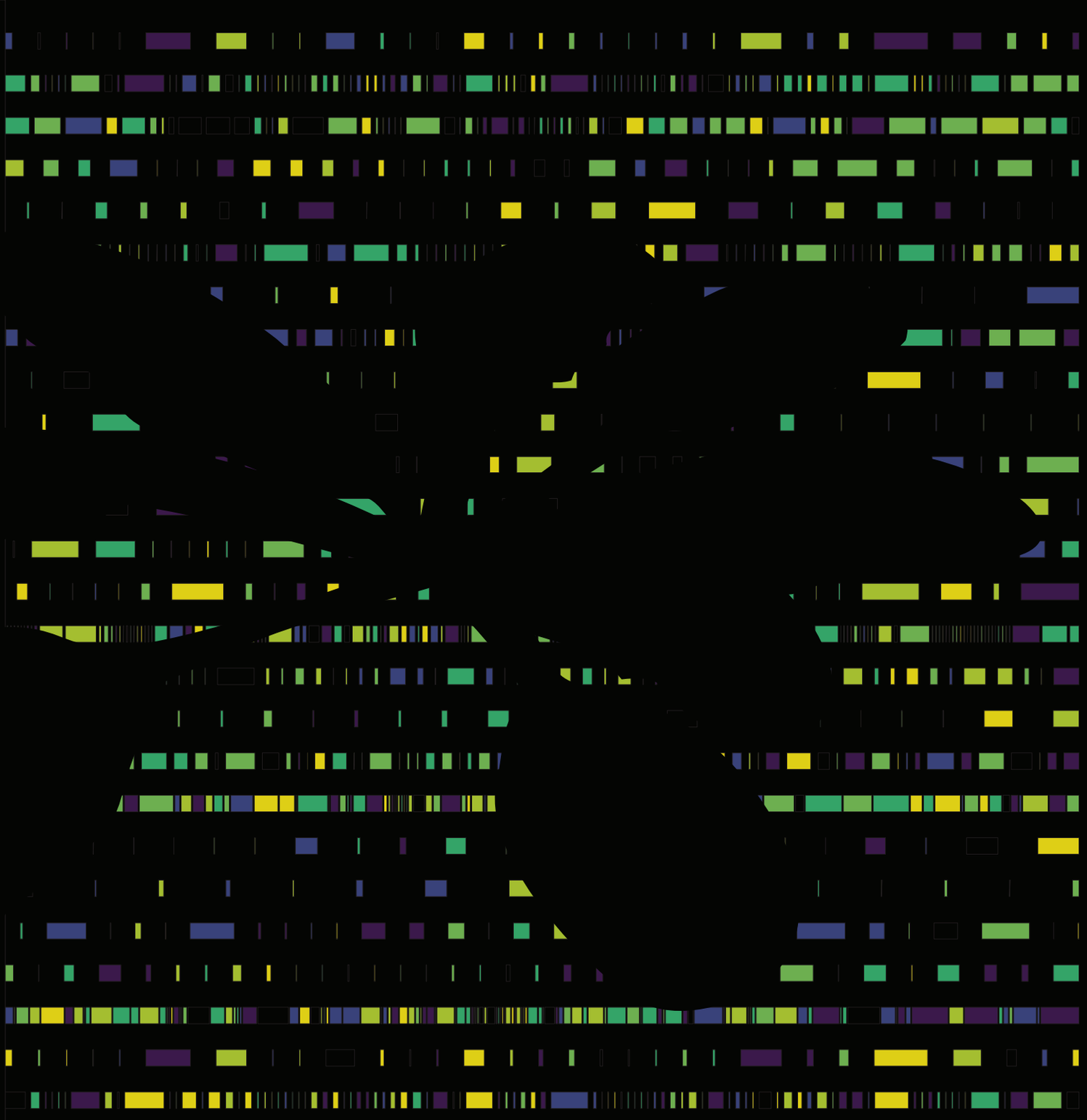


STRUCTURAL VARIATION OF A POLYMORPHIC *RPP1*-LIKE CLUSTER IN LOCAL *ARABIDOPSIS THALIANA* POPULATIONS ALTERS PHYLLOSPHERE MICROBIOME IN GORZOW WIELKOPOLSKI

PAUL RUNGE

Structural variation of a polymorphic *RPP1*-like cluster in local *Arabidopsis thaliana* populations alters phyllosphere microbiome in Gorzów Wielkopolski

Dissertation

der Mathematisch-Naturwissenschaftlichen Fakultät

der Eberhard Karls Universität Tübingen

zur Erlangung des Grades eines

Doktors der Naturwissenschaften

(Dr. rer. nat.)

vorgelegt von

Paul Runge

aus Wesel

Tübingen

2021

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

Tag der mündlichen Qualifikation:	30.06.2022
Dekan:	Prof. Dr. Thilo Stehle
1. Berichterstatter:	Prof. Dr. Eric Maik Kemen
2. Berichterstatter:	Prof. Dr. Korbininan Schneeberger

I Publications

Runge, Paul, Freddy Ventura, Eric Kemen, and Remco Stam. 2021. "Distinct Phyllosphere Microbiome of Wild Tomato Species in Central Peru upon Dysbiosis." Research Square. Research Square. https://doi.org/10.21203/rs.3.rs-648608/v1 (*in revision*)

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

Runge, Paul, Daniel Gómez-Pérez, Jane Parker, Detlef Weigel, Ruben Alcazar, Paweł Bednarek, Eric Kemen. 2021. "Chromosome-level assemblies of natural *Arabidopsis* individuals unveiled strucutral variation of the *RPP1*-like cluster on a population scale." (*in prep*)

Runge, Paul, Jane Parker, Eric Kemen .2021. "Phyllosphere microbiome of immune-compromising *Arabidopsis thaliana* mutants alters upon pathogen perturbation in nature". (*in prep*)

II Table of Content

III Table of Figures and Tables

Supplementary Figures

Supplementary Tables

IV List of abbreviations and acronyms

%	Percent
µ	Micro
cm	Centimetre
cM	Centimorgan
Col-0	Columbia
DNA	Deoxyribonucleic acid
e.g.	*Exempli gratia* (“for example”)
ETI	Effector-triggered immunity
ETS	effector-triggered susceptibility
Gw	Gorzów Wielkopolski
h	Hour
ITS	Internal transcribed spacer
Kb	Kilo base pair
La-0	natural Landsberg accession
LB	Lysogeny broth
Ler-0	*Landsberg erecta*
log	Logarithmic
MAMP	Microbe-associated molecular pattern
MAPK	Mitogen-activated protein kinase
ml	Millilitre
Mb	Mega base pair
min	Minute
MPIPZ	Max Planck Institute for Plant Breeding Research

OD	Optical density
OTU	Operational taxonomical unit
PCR	Polymerase chain reaction
PD(A)	Potato dextrose (agar)
pH	Negative decimal logarithm of H^+ concentration
PRR	Pattern-recognition receptor
PTI	PAMP-triggered immunity
rDNA	Ribosomal DNA
RNA	Ribonucleic acid
ROS	Reactive oxygen species
rpm	Rounds per minute
rRNA	Ribosomal RNA
RT	Room temperature
sec	Second
SNP	Single nucleotide polymorphism
spp	Species *pluralis*
SynCom	Synthetic community
taq	*Thermophilus aquaticus*
TSB	tryptic soy broth
UV	Ultraviolet
w/v	Weight per volume
Ws	Wassilewskija

V Summary

Plants are colonized by a myriad of microbes which affect host development, fitness, and reproduction. In recent decades, next-generation sequencing methods have enabled a substantial increase of plant microbiome studies, which have demonstrated that multiple abiotic and biotic factors influence mutualistic microbes on different plant compartments, between the phyllosphere (above-ground) and rhizosphere (below-ground) compartments. As such, plant microbiomes have been explored across a broad range of host species, including model organisms, crops and trees along with natural conditions determining multifactorial environmental traits. However, deciphering the structural impacts of host genetic factors on microbiota remain underexplored. Plants have evolved a complex network of immune signaling to cope with a range of pathogenic microbes that are known to affect microbial community assemblies. To obtain insights into the effects of host resistance genes that might indirectly affect microbial consortia through permission of resistance, I have focused on a highly polymorphic *RPP1*-like gene cluster in natural *Arabidopsis* populations of Gorzów Wielkopolski (formerly Landsberg an der Warthe) by combining population genetics with targeted-based microbial profiling over three consecutive years. 16 site-specific reference genomes were generated from individual plants combining short and long read sequencing to obtain structural insights of the *RPP1*-like cluster on chromosome 3. In addition, chromosome-level assemblies of two natural accessions of Landsberg La-0 were implemented as untreated references.

Microbial community profiling of bacteria and eukaryotes was compared between various data features such as i) sampling years, ii) geographical locations, iii) leaf compartments, and iv) nucleic acid resources. The impact of the presence or absence of *RPP1*-like copy number variations were linked to microbial diversities considering various microbiome sample features, such as Nucleic Acid (NA), Geographical Location (GC) and Leaf Compartment (LC). Further, persistent microbes were identified across Gorzów *Arabidopsis* ecotypes, which constituted a microbial core community. The present data suggest an enrichment of plant promoting *Sphingomonas* spp. in two out of three geographical locations in respect to a high *RPP1*-like copy number (*RPP1*+), while I observed an overall lower microbial diversity in *RPP1*+ compared to *RPP1*-. Overall, the phyllosphere microbiome of natural Gorzów *Arabidopsis* can be impacted by environmental cues, as well as host genetic factors such as the presence or absence of polymorphism of the *RPP1*-like cluster. While molecular functions of *RPP1*-like genes remain largely unknown, further investigations are required to identify potential microbes that can be correlated to natural observations.

VI Zusammenfassung

Pflanzen sind von einer enormen Anzahl und Vielfalt an Mikroben besiedelt, welche Einfluss auf die Entwicklung, Fitness und Vermehrung des Wirts haben können. Dabei werden die gesamten pflanzen-assoziierten Mikroben als „Mikrobiom“ bezeichnet. In den letzten Jahrzenten, haben verbesserte Technologien zur DNA-Sequenzierung („*Next-Generation Sequencing*“) einen beträchtlichen Anstieg an Mikrobiom-Studien verzeichnet, welche zur Identifikation von abiotischen und biotischen Einflussfaktoren auf mikrobielle Gemeinschaften in verschiedenen Nischen der Pflanze geführt haben. Das Mikrobiom der Pflanze, oft unterteilt in Phyllosphäre (oberirische Pflanzenteile) und Rhizosphäre (unterirdische Teile), wurde bereits an vielen natürlichen Spezies, einschließlich Model Organismen, Nutzpflanzen und Bäumen untersucht. Während Umweltfaktoren einen beachtlichen Anteil an Varianz innerhalb der mikrobiellen Gemeinschaft erklären, stellt der Einfluss des Wirtes weiterhin eine Herausforderung dar. Pflanzen beherbergen ein komplexes Immunsystem zur Abwehr einer Vielzahl an Krankheitserregern, welche einen fundamentalen Einfluss auf die strukturelle Organisation des Mikrobiom haben können. Um Erkenntnisse über den direkten oder indirekten Einfluss der Wirtspflanze auf das Mikrobiom der Phyllosphäre zu erlangen, habe ich den Einfluss von Resistenzgenen untersucht. In der vorliegenden Arbeit wurde die Anwesenheit des polymorphen *RPP1-like* Resistenzgen-Cluster in natürlichen *Arabidopsis thaliana* Populationen der polnischen Region Gorzów Wielkopolski (früher bekannt als Landsberg an der Warthe) auf die mikrobielle Gemeinschaft untersucht. Dazu wurden populationsgenetische Ansätze mit *Amplicon*-basierter DNA- und cDNA-Sequenzierung kombiniert. Durch die Genom-Sequenzierung von 16 repräsentativen Gorzów *Arabidopsis* Individuen, identifizierte ich strukturelle Variationen innerhalb des *RPP1*-like Resistenzgen-Cluster auf Populations-Ebene. Des Weiteren wurden zwei natürliche Landsberg Linien (La-0) auf Chromosomen-Level sequenziert, welche als Referenzgenome einbezogen wurden. Zusammengefasst bietet dieser Datensatz großes Potential für die Untersuchung von Genom-Analysen auf Populations-Ebene.

Die Untersuchung des phyllosphären Mikrobioms auf natürlichen *Arabidopsis* Populationen in Gorzów Wielkopolski ergab, dass die Faktoren Jahr, Standort und Blatt-Kompartiment den größten Einfluss auf die mikrobiellen Gemeinschaft haben. Zur Bestimmung des *RPP1*-like Cluster innerhalb der natürlichen Populationen, etablierte ich einen Assay zur Genotypisierung, welcher die *RPP1*-like Kopien-Anzahl im Genom bestimmt. Durch die Korrelation der *RPP1*-like Kopien

im Genom mit der mikrobiellen Diversität konnte ich eine geringere Bakterien-Vielfalt und Proben-Variabilität in *RPP1*+ Haplotypen feststellen. Dadurch konnte der Einfluss der Wirtspflanze auf das Mikrobiom der Phyllosphäre festgestellt werden. Zudem wurde eine erhöhte Abundanz des nützlichen Bakteriums *Sphingomonas* in *RPP1*+ Haplotypen festgestellt. Zusammengefasst identifizierte ich mehrere Faktoren, wie Jahr, Standort, Blatt-Kompartiment und Wirtspflanze als Einflussfaktoren auf die strukturelle Organisation des phyllosphären Mikrobioms in natürlichen *Arabidopsis* Populationen aus der Region Gorzów Wielkopolski. Während die molekulare Funktion des *RPP1*-like Clusters weiterhin unbekannt ist, können weitere Experimente unter kontrollierten Bedingungen zur Identifikation beitragen.

Introduction

1 Introduction

Within the last few decades, the emergence of large-scale sequencing techniques and constant development of bioinformatic tools have amplified our understanding of complex microbial communities, such as with the culture-independent characterization of microbiota in humans, animals and plants. The following subchapters 1.1 – 1.3 were written by myself and display partial- or full-citation of the research publication of Runge et al. (2021).

1.1 Plant microbiome

"Terrestrial plants host distinct microbial communities on various plant organs, generally divided between characterized as above-ground (phyllosphere) and below-ground (rhizosphere) compartments. Many studies have been published that examine the phyllosphere and rhizosphere microbiota (Castrillo et al., 2017; Latz et al., 2021a; Stopnisek & Shade, 2021). Both plant compartments display unique and overlapping microbial pools. However, less is known about functional traits that shape microbial community structures in plants."

1.2 Driving factors of phyllosphere microbiota

"The phyllosphere is inhabited by a tremendous variety of diverse microorganisms in nature, such as bacteria, archaea, fungi, algae, viruses and protists (nematodes, protozoa) (Agler et al. 2016). Thereby, the phyllosphere is dominated by leaf surfaces featuring an oligotrophic environment supporting microbe-microbe interactions (Schlechter, Miebach, and Remus-Emsermann 2019). Bacteria are the most dominating microbial kingdom on leaves with around 10^4 – 10^5 bacterial cells mm^{-2} (Remus-Emsermann and Schlechter 2018). Less is known about yeasts, fungi and protists on the phyllosphere, although there are an increasing number of studies profiling eukaryotic microbes across various host species (Horton et al. 2014; Agler et al. 2016; Z. Cheng et al. 2020; Yao et al. 2019). Colonizers of the phyllosphere originate from various sources, such as soil, air, rain and insects (horizontal transmission), or through pollen or seeds (vertical transmission) (Chaudhry et al. 2020; Finkel et al. 2019). Phyllospheric microbes are divided into subgroups depending on if they are colonizers of the surface (epiphytes | phylloplane) and/or the cytosolic compartment (endophytes | endosphere). Epiphytes have to continuously cope with microhabitat conditions, such as light exposure (ultraviolet), high temperatures, and sparse nutrient and water availability (Firrincieli et al. 2020; Gomes et al. 2018; J. Kumar et al. 2016; Pincebourde

and Casas 2019; Remus-Emsermann and Schlechter 2018). In contrast, endophytes are living in a more sustainable micro-environment and are likely to interact mutualistically with host cells that support host-microbe interactions (Brader et al. 2017; Dudeja et al. 2021; Latz et al. 2021; Strobel 2018; C. Xiong et al. 2021). Contradictory research findings also support the hypothesis that host genetics affect bacterial compositions of epiphytes on trees (Yao et al. 2020; Mina et al. 2020). Besides abiotic stresses, biotic factors incorporating beneficial and pathogenic microbes influence phyllosphere microbiota assemblies. In this context, the plant immune system indirectly affects microbial consortia by restricting microbial proliferation to secure host health (Chaudhry et al. 2020; Horton et al. 2014; T. Chen et al. 2020). Moreover, hormone cross-talks between abiotic and biotic stress responses have been shown to play a key role in the adaptation of plant species (Bai et al. 2018; Berens et al. 2019; N. Iqbal et al. 2013; Z. Iqbal et al. 2021)."

1.3 Host species affecting microbial assemblies

"In recent decades, many approaches have been established to study binary host-microbe interactions (Bartoli and Roux 2017; Xiao et al. 2017). While those studies have been enormously fruitful for discovering functional mechanisms, next generation sequencing methods enable the characterization of complex host-microbe interaction shaping microbiomes (Bálint et al. 2013; Latz et al. 2021; Sapkota et al. 2015; Shakir et al. 2021). Besides the integration of environmental factors, host-microbe and microbe-microbe interaction studies are required to understand microbial community compositions. However, our knowledge about plant genotypes controlling microbiota remains incomplete. Genome-wide association studies (GWAS) have been used to identify certain host factors that impact microbial consortia (Horton et al. 2014; Beilsmith et al. 2019). Depletion of symbiosis genes of the wild legume *Lotus japonicus*, such as *RAM1*, *NFR5*, *SYMRK* and *CCaMK,* has shown an effect on the structuring of bacterial and fungal root-associated communities (Thiergart et al. 2019). How microbial communities adopt to their host plant is an important scientific question. (Batstone et al. 2020) elaborate an evolutionary approach on *Medicago truncatula* to study the adoption of beneficial bacteria to encounter genotype-by-genotype (GxG) interactions over one year. Their findings revealed that bacterial symbionts rapidly adapt to local host genotypes.

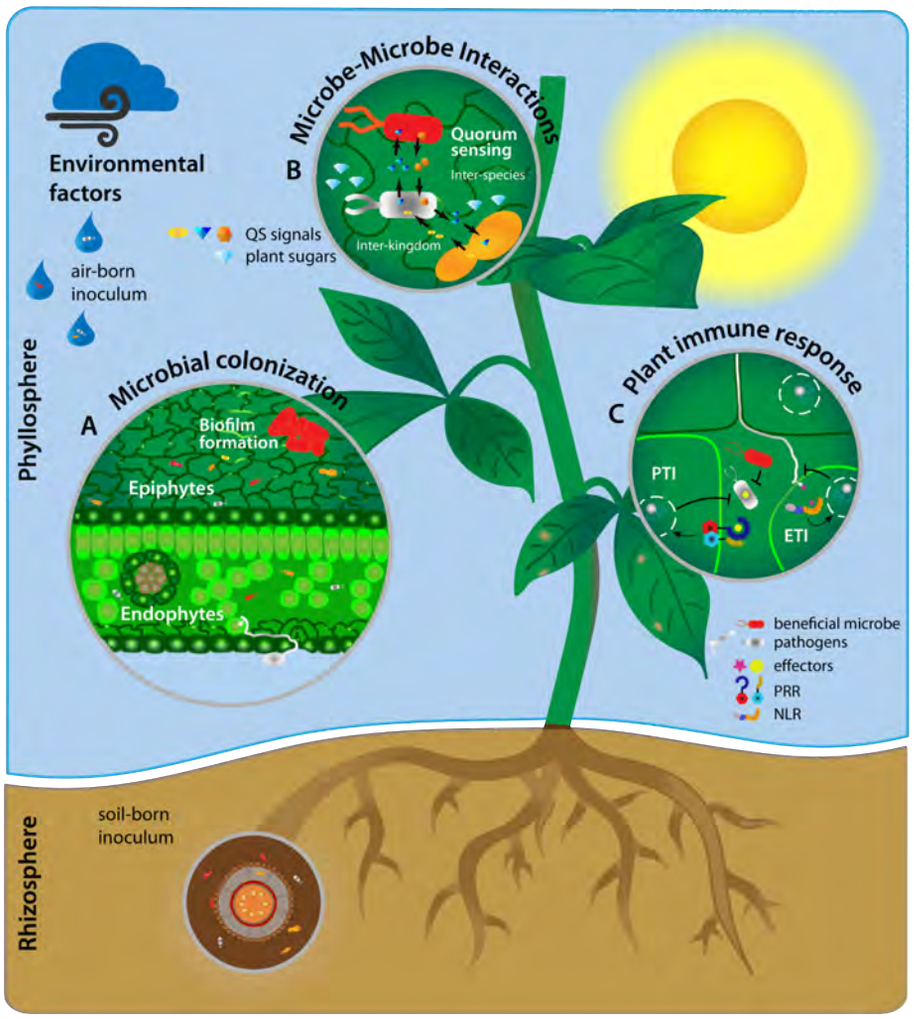


Fig. 1: Microbial colonization of the aerial part of the plant (phyllosphere), as well as the below-ground part (rhizosphere). (A) The microbial colonization on the leaf takes place on the leaf surface (epiphytes) from air-borne and soil-borne inoculum, and within the leaf (endophytes). Microbial colonization can lead to exogenous intraspecies biofilm formation on the leaf surface. (B) Microbe–microbe interactions occur inter-species and inter-kingdom, referred to as quorum sensing. Quorum-sensing molecules impact microbial recognition and biofilm formation on leaves. (C) Pathogenic microbes colonize host plants by means of their virulence. The genetic make-up of both the host and pathogen contribute to disease progression. However, other microbes in the host phyllosphere can influence this plant–pathogen interaction by either facilitation or antagonism. (D) Non-host-adapted pathogens are involved in PAMP-triggered immunity (PTI) and recognized via pattern-recognition receptors (PRRs). Host-adapted microbes are recognized via nucleotide-binding leucine-rich repeat receptors (NLRs), summarized in effector-triggered immunity (ETI).

Further bacteria community applications on multiple tomato species, such as *S. lycopersicum* and *S. pimpinellifolium* revealed robust habitats and genotype-specific selection on the phyllosphere over time. In addition, stabilized bacterial communities are vigorous against bacterial invasion of the start inoculum (Gong and Xin 2021; Morella et al. 2020).

Current research suggests a key role of host species in establishing microbial communities on various plant organs. Since pathogenic microbes have been identified as major factors in shaping of microbial consortia (Agler et al. 2016), host "resistotypes" might be of particular interest to examine genotype effects on microbiota."

1.4 Role of the plant immune system in shaping the leaf microbiome

The section was written by myself and includes a direct citation of the published review of Chaudhry et al., (2020).

"The plant innate immunity system comprises a large repertoire of plasma membrane-localized (surface) and intracellular receptors which recognize microbial or modified host molecular signatures in order to retain plant health and secure plant propagation. Surface immune receptors (often referred to pattern recognition receptors, or PRRs) are members of a diverse family of ligand-binding proteins that sense microbial, environmental, developmental, and nutritional cues (Saijo, Loo, and Yasuda 2018; Y. T. Cheng, Zhang, and He 2019). In terms of shaping microbial communities, it is the PRRs activity that are thought to gate microbial entry into leaf tissues, and effectively ward off colonization by host non-adapted strains (Boutrot and Zipfel 2017). The intracellular receptor panels (consisting mostly of nucleotide-binding/leucine-rich repeat (NLR) proteins), are similarly diverse and are selected as triggers of strain-specific resistance to host-adapted pathogens (Eitas and Dangl 2010; J. D. G. Jones, Vance, and Dangl 2016; C.-H. Wu et al. 2017; Burdett et al. 2019; Van de Weyer et al. 2019).

The activation of plant immune responses by mobilizing a network of defence and stress hormone pathways has been extensively characterized in binary plant-pathogen interactions (Noman, Aqeel, and Lou 2019; T. Zhao et al. 2019). Little is known about the impact of plant immunity signalling networks on host-microbe interactions in leaf microbial communities (see Fig. 1C). High-throughput DNA and RNA sequencing of leaf samples from natural environments have enabled examination of complex microbial communities in plant-specific niches in time and space (Agler et al. 2016). Analysis of microbial metadata and their integration with experimental testing should provide a clearer picture of the role of plant immunity signalling in shaping leaf microbial community structure and, in turn, how resident microbes influence host immunity. In this section, I consider the present evidence that abiotic and biotic stress responses modulate microbial consortia on leaves and discuss the consequences for plant fitness. It is becoming clear that microbial

community structure throughout a plant host's lifecycle is dynamic and modulated by the innate immune system, which itself is affected by environmental changes."

1.4.1 The role of pattern-triggered immunity in shaping leaf microbiota

Most microorganisms on plant leaves are non-pathogenic. However, a broad range of microbes are able to prime innate plant immunity to counter subsequent pathogen attacks (Ritpitakphong et al. 2016; Vogel et al. 2016). Many microbes are recognized by terrestrial plants through their microbial-associated molecular patterns (MAMPs) initiating pattern-triggered immune (PTI) responses. PTI is an induced and often low-level but broadly effective resistance response involving phytohormone signalling, secretion of antimicrobial compounds, generation of reactive oxygen species (ROS) and MAP kinase cascades, and stomatal closure (Bigeard, Colcombet, and Hirt 2015; Bi and Zhou 2017). Notably, the phytohormone ethylene is required for ROS production in PTI in *Arabidopsis* resistance to *Pseudomonas syringae* bacteria and rice resistance to rice blast fungus *Magnapothe oryzae* (Mersmann et al. 2010; Guan et al. 2015; Helliwell, Wang, and Yang 2016; C. Yang et al. 2017). In *Arabidopsis*, an *ethylene-insensitive2* (*ein2*) mutant displays an altered bacterial leaf community compared to wild-type plants, suggesting that ethylene signalling is important for modulating leaf microbiota (Bodenhausen et al. 2014; Nascimento, Rossi, and Glick 2018).

A recent study by (T. Chen et al. 2020) provided experimental evidence that PTI signalling controls the diversity of endophytic leaf microbiota in microorganism-rich environments. An *Arabidopsis* quadruple mutant (*min7 fls2 efr cerk1* (*mfec*)) defective in PTI and the MIN7 vesicle trafficking pathway (affecting aqueous apoplastic micro-environment), and a *constitutively activated cell death1* (*cad1*) mutant, show altered endophytic bacterial leaf diversity (T. Chen et al. 2020). In particular, the relative abundance of the bacterial phyla *Firmicutes* was significantly reduced, whereas *Proteobacteria* became dominant bacterial community members in the mutant plants. The presence of the PTI components *MIN7* and *CAD1* across major plant lineages suggests a number of common pathways might govern endophytic microbial proliferation of certain taxa in leaves.

Further research has revealed the importance of resident *Pseudomonas* sp. (*Proteobacteria*) in protecting *Arabidopsis* against infection by a fungal necrotrophic pathogen, *Botrytis cinerea* (Ritpitakphong et al. 2016). Notably, prominent bacterial clades from soil microbiota such as filamentous *Actinobacteria* (*Strepotmycetes* sp.) are able to activate plant biosynthesis of salicylic acid and promote leaf defence responses against fungal pathogens (Vergnes et al. 2020). These studies

emphasize that both commensal and pathogenic microbes are priming PTI as a barrier to colonization of the leaf compartment by host non- or poorly adapted pathogens. Nevertheless, these host-microbe interactions were examined mostly under controlled laboratory conditions.

1.4.2 Leaf effector-triggered immunity as a potential microbial gateway

Strain-specific resistance, known as effector-triggered immunity (ETI), is often mediated by intracellular NLR receptors which recognize certain pathogen-delivered virulence factors (effectors) to induce immunity (Monteiro and Nishimura 2018; Feehan et al. 2020; Seong et al. 2020). Pathogen effector-activated NLRs accelerate and amplify many PTI responses, often resulting in host localized cell death (a hypersensitive response) and rapid pathogen containment (Peng, van Wersch, and Zhang 2018). Expressed NLR genes in roots are observed in dicot plant species like the legume Lotus (Lai and Eulgem 2018). This is in contrast to tested *Brassicaceae* species including *Arabidopsis thaliana* and the crop oilseed rape (*Brassica napus*), which favour NLR expression in the phyllosphere (Munch et al. 2018).

Diverse microbial communities in leaves can be controlled directly through pathogen colonization on the host or indirectly by host-microbe interactions involving the innate immunity network (Agler et al. 2016). As such, pathogenic microbes can act as highly interconnected community members (so called "hub microbes") that dominate microbial community assemblies. For example, the causal agent of white rust on *Arabidopsis, Albugo* sp., appears to act as a hub which alters epiphytic and endophytic bacterial colonization of leaves (Agler et al. 2016; Ruhe et al. 2016). Perturbations of microbial communities by host-adapted biotrophic pathogens such as *Albugo* and *Hyaloperonospora arabidopsidis* (*Hpa*) reduces microbial diversity within leaf habitats and stabilizes microbial communities among wild plants (Karasov et al. 2019). Hence, microbial diversity can be used as an indicator of microbial community imbalance (T. Chen et al. 2020).

Whether ETI reactions directly lead to defence priming is not well-studied, although in *Arabidopsis* one important ETI branch does lead to a reinforcement and spread of pathogen resistance (so-called basal immunity) in leaf tissues (Lapin, Bhandari, and Parker 2020). A recent study by (Levy et al. 2018) analysed over 3,800 genomes of plant-associated (pathogenic and non-pathogenic) bacteria. The analysis identified plant-mimicking protein domains (named as PREPARADOS) that carry non-canonical 'embedded' nucleotide-binding leucine-rich repeat (NLR) domains. An increasing number of NLR-fused domains are related to authentic effector targets. REPARADOS are highly abundant in the bacterial families *Bacteroides* and *Xanthomonadaceae* (Frank 2019).

These findings indicate potential interactions between commensal and/or pathogenic bacteria with intracellular receptors in host plants. Additional studies are needed to test this hypothesis and dissect functional relationships between NLR panels and the leaf microbiota.

1.4.3 Stability of microbial consortia against pathogen perturbation

The plant and its associated microbiota is not a static environment but is altered by numerous factors including host genotype, environmental fluctuations, surrounding macro- and microorganisms, and geographical location and associated local variables such as climate (Laforest-Lapointe, Messier, and Kembel 2016; Poudel et al. 2016; Wagner et al. 2016; Singh et al. 2018). The stability of a leaf microbial community is measured by the ability to maintain a stable equilibrium state (homeostasis) under biotic or abiotic perturbations (Thébault and Fontaine 2010). Generally, higher community complexities reflect a more stable community structure (Mougi and Kondoh 2012). Stable microbial communities or consortia have a greater ability to resist perturbation (Ives, Klug, and Gross 2000; Luo et al. 2019; Morella et al. 2020). Studies using culture-independent DNA sequencing revealed similar microbial community patterns in successive year samplings (Copeland et al. 2015). In the phyllosphere, microbial communities can often undergo drastic changes and establish a distinctive and less diverse community (Manching, Balint-Kurti, and Stapleton 2014). Different computational and experiment-based approaches have been used to capture microbial community homeostasis or deviations over time. Computational microbial network analysis and mining of core microbes are valuable in understanding the factors underlying microbial resilience to controlled perturbations (Astudillo-García et al. 2017; Lemanceau et al. 2017). Much less is known about the dynamics and stability of leaf microbiomes in the field since there is a lack of high resolution experimental data linked to plant disease and health with respect to time, space, and environmental scale. In recent studies, leaf diseases were linked to disruption of microbial community network stability, resulting in ecosystem dysfunction (Kerdraon et al. 2019; Luo et al. 2019; Leopold and Busby 2020). Understanding how a microbial community cures under conditions of environmental stress is crucial to harness its potential in probiotic applications against aggressive plant pathogens and to track plant-associated human pathogen outbreaks.

1.4.4 Does immunity-priming affect microbial leaf communities?

Various abiotic and biotic factors impact dynamic changes on microbial leaf communities as depicted in the modes of microbial colonization, microbe-microbe and microbe-host interactions (see Fig.1 and Tab.1). Nevertheless, fundamental mechanisms of microbial community assembly remain poorly understood. One major goal of current microbiome research is to understand how microbial consortia in nature secure plant protection during pathogen perturbation. Effects of immunity priming (IP) through abiotic (applied chemical compounds) and biotic (biocontrol agents) stimuli seem to play an important role in managing abiotic stress tolerance and disease resistance (A. Kumar and Verma 2018). IP has been described as a "*positive cost-benefit balance in times of stress*" (Martinez-Medina et al. 2016). IP induction involves the phytohormones salicylic acid (SA), jasmonic acid (JA), and pipecolic acid-derived signalling molecules that are known to mediate systemic acquired resistance, as well as the non-protein amino acid defence primer β-aminobutyric acid (BABA) (Martinez-Medina et al. 2016). BABA is found naturally in *Arabidopsis* experiencing abiotic stress (high salinity) and biotic stress and induces broad-spectrum pathogen resistance (Thevenet et al. 2017; Buswell et al. 2018). Another interesting IP compound, (R)-b-homoserine (RBH), primes ethylene and JA-pathways and is effective against necrotrophic pathogens such as *Botrytis cinerea* in tomato and *Plectosphaerella cucumerina* (Buswell et al. 2018). Brassinoteroids (BR) have also been discussed as factors in an IP mechanism that balances the trade-off between immunity and growth (Yu, Zhao, and He 2018). These findings highlight the potential utility of chemical compounds for IP. They also prompt studies of how IP impacts leaf microbial diversity under conditions of abiotic and biotic stress. Effects of biocontrol agent (BCA) application on crops such as potato against hemi-biotrophic (*Phytophthora infestans*), and grapevine against necrotrophic (*Botrytis cinerea*) fungi have been studied extensively *in vitro* (Bailly and Weisskopf 2017; De Vrieze et al. 2018; Bruisson et al. 2019). In contrast, applying *Pseudomonas syringae* pathovar tomato (*Pst*) to *Arabidopsis* roots attracted *Bacillus subtilis* and led to IP upon *Pst* infection (Rudrappa, Biedrzycki, and Bais 2008; Vannier, Agler, and Hacquard 2019). The ecological impact of BCAs on the leaf microbiome while controlling disease resistance remains an open research question. Current reports emphasize a link between certain bacterial taxa (*Bacillus*, *Pantoea*, *Sphingomonas*, *Pseudomonas* and *Trichoderma*) affecting microbial diversity (Zhang et al. 2008; Bruisson et al. 2019; Ulrich et al. 2020) and IP induction on leaves (Cawoy et al. 2014; Ritpitakphong et al. 2016; C. Qin et al. 2019). In particular, highly diverse leaf communities are negatively correlated with pathogen invasion and colonization and *vice versa* (Purahong et al. 2018;

C. Qin et al. 2019). Other reports describe difficulties encountered in the application of biocontrol agents such as *Bacillus subtilis (Bs)*, which did not alter the microbial leaf community under rainy field conditions (F. Wei et al. 2016). Therefore, use of BCAs under natural conditions might be challenging and require further analysis. However, BCAs and IP-inducing compounds can potentially be used to monitor disease control to improve crop yield and production in new biological breeding strategies. There is clearly a need to increase efforts in this research field to explore the effects and underlying mechanisms of abiotic and biotic stress on immunity-priming and how they are transmitted to microbial leaf communities."

1.5 NLR gene cluster formation

The evolution of the innate immune system relies on receptor diversities correlating with pathogen recognition specificities to secure plant health. Plants harbour an enormous repertoire of immune receptors, such as intracellular nucleotide binding leucine-rich repeat (NLR) immune receptors, conferring resistance to various infectious diseases (Barragan and Weigel 2021). The abundant NLR receptor family is known for heterogenicity showing vast sequence and structural polymorphisms. Although gene cluster formations are rare in eukaryotes, NLRs tend to accumulate in clusters (J. M. Lee and Sonnhammer 2003; R. R. Q. Lee and Chae 2020). The diversity of NLRs and haplotypes, present in resistance gene clusters on a species or population level, might be unresolvable when focusing on a single reference genome (Stam, Silva-Arias, and Tellier 2019; Stam, Scheikl, and Tellier 2016; Noel et al. 1999; Kuang et al. 2004; Christopoulou et al. 2015). A recent study examined the NLR-ome of 64 natural accessions of *Arabidopsis thaliana* using long read-based resistance gene enrichment sequencing (REN-Seq) uncovered 75 novel domain architectures (Van de Weyer et al. 2019). Hence, distinct accessions revealed high genetic diversity within species. While NLRs display tendencies to occur in gene clusters, their functional significance remains largely unknown (S. van Wersch and Li 2019). Within the model organism *Arabidopsis thaliana*, many NLR gene clusters have been identified across various accessions such as *RPP1*, RPP5, RPP8, B3 and B5 clusters (Yi and Richards 2007; Bevan et al. 1998; MacQueen et al. 2019; Alcázar et al. 2014; Holub 2001; Meyers et al. 2003; R. R. Q. Lee and Chae 2020). Interestingly, NLR gene clusters vary highly in copy numbers, consisting of at least two genes (e.g. head-to-head pair RPS4-RRS1) and have been identified with up to 11 NLRs, such as the B5 cluster (Holub 2001; Narusaka et al. 2009). In addition, one of the most variable NLR gene clusters in *Arabidopsis thaliana* showing eight gene copies, such as (~77 Kb) and *RPP1* (~ 87 Kb) cluster (Alcázar et al. 2014; Meyers et al. 2003; R. R. Q. Lee and Chae 2020).

The diversity of NLRs and haplotypes present in resistance gene cluster are postulated to be hotspots where diversification and generation of new R-genes occur at high acceleration rates (Jiao and Schneeberger 2020; S. van Wersch and Li 2019; R. R. Q. Lee and Chae 2020). Although intra-species variation of those highly polymorphic gene clusters has been conducted, further studies are required to obtain intra-population structures of NLR clusters with high resolution sequencing approaches.

1.6 NLR structures and signal transduction

Plant NLRs share similar architectures comprising three core domains: a C-terminal super-structure forming repeat (SSFR) domain [typically leucine-rich repeats (LLR)], a nucleotide-binding site (NBS), and a N-terminal coiled-coiled (CNL's) or Toll/interleukin-1 receptor (TIR) domain (TNL's) (W. Song et al. 2021). Interestingly, a genome-wide survey of NLR repertoires between 38 model organisms spanning major kingdoms revealed the occurrence of core NLR domains in eubacteria and archaebacteria, as well as in green algae chloroplasts suggesting structural and functional convergence to distinguish self from non-self (implying immune responses) (Jacob, Vernaldi, and Maekawa 2013; Shao et al. 2019; Ortiz and Dodds 2018). However, there is a high degree of structural variability, including in C-terminal domains (WRKY DNA-binding domain), integrated domains (IDs) and additional types of N-terminal domains, such as kinases and α/β hydrolases (Sarris et al. 2016; Cesari et al. 2014; Andolfo et al. 2019; Le Roux C Jauneau A Camborde L Trémousaygue D Kraut A Zhou B Levaillant M Adachi H Yoshioka H Raffaele S Berthomé R Couté Y Parker JE Deslandes L. 2015).

In general, NLR proteins discern cytoplasmic pathogen invasion (Jeffery L. Dangl, Horvath, and Staskawicz 2013). In this respect, several modes of action have been described for single or paired NLRs activating local and systemic defences (Sun et al. 2021). In addition, plant NLRs are involved in either direct or indirect effector recognition signalling. Direct interactions of plant NLRs, especially with the LRR domain, have been observed in various studies including Pi-Ta/AvrPi-Ta (*Magnaporthe oryzae*) in rice, Sr50/AvrSr50 in (*Puccinia graminis*) in wheat and multiple MLAs recognising powdery mildew (*Blumeria graminis* f. sp. *hordei*) in barley (Dodds et al. 2006; Saur et al. 2019; J. Chen et al. 2017; Krasileva, Dahlbeck, and Staskawicz 2010). Within the model system *Arabidopsis thaliana*, a well-known pathosystem involves the TNL receptor *RESITANCE TO PERONOSPORA PARASITICA 1* (*RPP1*) and the biotrophic oomycete *Hyaloperonospora arabidopsidis* (*Hpa*), which is recognized by the effector ATR1. A recent study revealed that co-

expression of a natural variant of the TNL *RPP1* with its matching *Hpa* effector ATR1 leads to NLR conformational activation by tetrameric *RPP1* oligomeric structures, forming a holoenzyme for NAD+ hydrolysis (Ma et al. 2020). Interestingly, similar tetrameric complex formations were observed for the TNL ROQ1 in *Nicotiana benthamiana*, recognising the effector XopQ1 from Xanthomonas (Martin et al. 2020). These examples display direct recognition of secreted effector molecules by NLRs activating downstream immune signalling cascades. However, there are numerous NLR-Avr perceptions where physical interactions have not been observed, leading to the assumption of indirect recognition.

First indications of indirect effector recognition were described by Van der Biezen and Jones (1998). Specifically, in the case of the kinase *Pto* of *Pseudomonas syringae* interacting with the NLR *Prf* in tomato (Van der Biezen and Jones 1998). While the kinase activity is not involved in AvrPto-triggered immunity, *Prf* mediates a strong immune response leading to indirect pathogen detection (J. L. Dangl and Jones 2001; S. van Wersch et al. 2020). Indirect pathogen recognitions by NLRs are hypothesized by the guard and decoy model (Cesari et al. 2014). Hence, the guard model postulates how multiple effectors could be tackled by single NLRs (also named as guardees), leading to immune responses against a broad diversity of plant pathogens (J. L. Dangl and Jones 2001). In addition, it was suggested that guardees are indispensable for the virulence of pathogen effector lacking the R protein. Classical guardees have been described in *Arabidopsis* (*RIN4*, *PBS1*) and tomato (*RCR3*, *Pto*) (van der Hoorn and Kamoun 2008; J. D. Jones and Dangl 2006). However, recent studies support the notion that various effectors have multiple host targets (also named as decoy), and that guardee proteins are frequently dispensable for effector virulence in absence of the R protein (van der Hoorn and Kamoun 2008; Kourelis et al. 2020). In addition, the guard model is in conflict with the presence/absence scenario (polymorphism) of R genes leading to opposing natural selection forces in plant populations (van der Hoorn and Kamoun 2008).

Within the decoy model, a duplicated NLR pair form a heterocomplex receptor (often in head-to-head configuration), where one member acts as an effector target (decoy) to divert pathogen effectors from its main target, while the other induces defence resistance. Thus, the decoy (sensor) represses immune signalling activates of the transducer until pathogen perception (Williams et al. 2014; Césari et al. 2014; van der Hoorn and Kamoun 2008). Famous examples include NLR-pairs in *Arabidopsis* (*RPS4-RRS1*) and rice (*RGA4-RGA5*, *Pikp*-1-*Pikp*-2) (Huh et al. 2017; Guo, Wang, and Jones 2021; Césari et al. 2014; Zdrzałek et al. 2020).

Independent of numerous NLR signal perceptions, downstream responses of ETI have been described through multiple signalling nodes, such as *NON-RACE SPECIFIC RESISTANCE 1* (*NDR1*), a plasma membrane-anchored integrin-like protein involved in CNL signalling, or lipase-like proteins including *ENHANCED DISEASE SUSCEPTIBILITY 1* (*EDS1*) and *SENESCENCE-ASSOCIATED CARBOXYLESTERASE 101* (*SAG101*) or *PEPTIDYL ARGININE DEIMINASE 4* (*PAD4*), predominantly relevant in TNL signalling (Feys et al. 2005; Century, Holub, and Staskawicz 1995; Parker et al. 1996). Downstream immune responses are not limited through these signalling nodes as shown by *RPP8* (CNL) that acts independently of *NDR1* (Aarts et al. 1998). Further, chaperone proteins are required for NLR-mediated plant immune activation. For example, *HSP90*, *SGT1b*, and *RAR1* (Takahashi et al. 2003; Hubert et al. 2003; Liu et al. 2004; S. van Wersch et al. 2020). NLR activation leads to cellular changes, such as Ca^{2+} influx, reactive oxygen species (ROS) production, activation of mitogen-activated protein (MAP) kinases, salicylic acid (SA) accumulation, altered endomembrane trafficking, transcriptional reprogramming, and hypersensitive reaction (HR) cell death (Lolle, Stevens, and Coaker 2020; Cui, Tsuda, and Parker 2015).

Apart from NLR-mediated pathogen recognition, environmental factors have been suggested to effect immune response outputs. While high temperature and humidity have been associated with reduced disease resistance, low temperatures and high light are related to increased resistance (Y. T. Cheng, Zhang, and He 2019; Panchal et al. 2016; Jambunathan 2001; Yi Wang et al. 2009; S. H. Kim et al. 2010; Huang et al. 2010; Y. S. Kim et al. 2017; Mühlenbock et al. 2008). For instance, nuclear accumulation of SNC1 and RPP4 is inhibited by high temperatures repressing their immune responses (Y. Zhu, Qian, and Hua 2010; Mang et al. 2012).

In general, plant NLRs are steered under tense negative control to avoid self-damage in the absence of pathogens. As such, loss-of-function in negative regulators or gain-of-functions in immune receptor proteins lead regularly to autoimmunity. Phenotypic observations of autoimmune mutants show dwarfism, spontaneous cell death, or lethality (R. van Wersch, Li, and Zhang 2016; Y. Wu et al. 2020; Richard and Takken 2017; Liang, Tong, and Li 2020). While the contribution of NLR activation to autoimmunity has been discussed in various studies, a comprehensive analysis of autoimmunity on molecular bases is lacking.

1.7 *Arabidopsis thaliana* accession Landsberg

Landsberg erecta (Ler-0) is the second-widest studied *Arabidopsis thaliana* accession after Columbia (Col-0). The original Landsberg population (La-0 or La-1) was sampled among more than 150 natural accessions by Friedrich Laibach between 1930 and 1950. Laibach discovered the potential of *Arabidopsis* as a model organism for genetic, developmental and physiological analyses (Laibach 1943). In the early 1950's, Gyorgy P. Rédei recognized the new plant model system and transferred four natural *Arabidopsis* accessions, including Landsberg and Graz, Limburg and Estland from Europe to the University of Missouri (Columbia) where he started his own laboratory (George P. Rédei 1992).

Rédei decided to choose Landsberg as his model line. In 1957, he used Landsberg for a mutagenesis experiment using X-ray irradiation where he discovered the *Landsberg erecta* (Ler-0) mutant (George P. Rédei 1992; G. P. Rédei 1962; Langridge 1994). Rédei realized through his mutagenesis screen that the original Landsberg population was not a homogenous line, rather a mixture of different lines. Therefore, Rédei selected a single plant from his untreated Landsberg seed stock and named his new accession Columbia (Col-0) based on his location. While his original seed stocks originate from Europe it seems a peculiarity that he named it according to the place of his lab.

Nevertheless, the mutant Ler-0 found its way back to Europe by Willem Feenstra in 1959, who used it to study the growth habitat, and it became a standard line in the Department of Genetics at University of Groningen (Netherlands) (Feenstra 1964). Feenstra introduced a mutant induction program, which was continued by Jaap van der Veen and Maarten Koornneef (Innerebner, Knief, and Vorholt 2011).

In 1975, Rédei highlighted the importance of *Arabidopsis* in an article titled '*Arabidopsis* as a genetic tool' similar to Laibach in 1943. In respect of the people mentioned above, *Arabidopsis* got its breakthrough in science as a model system in the 1980s (Koornneef et al. 1983; Somerville and Ogren 1980). A third important article about *Arabidopsis* was published in the journal 'Science' in 1985 (Meyerowitz and Pruitt 1985). This incredible timeframe of 40 years from the first suggestion to the general acceptance of *Arabidopsis* as a plant model system displays the importance of Laibach's-, and later on Rédei's scientific perspective.

The distribution of various parental Landsberg seed stocks across laboratories in Europe prompted an increasing interest in the European *Arabidopsis* standard accession. While most likely all seeds originate from the same Ler-0 seed stock, some of the Landsberg accessions have been renamed

into Ler-1 and Ler-2 (Zapata et al. 2016). In the early 1990's, *Landsberg erecta* was the dominated *Arabidopsis* accession. However, this dramatically changed when Columbia (Col-0) was chosen as a natural accession for the whole genome sequencing project of *Arabidopsis* (*Arabidopsis* Genome Initiative 2000). Since then, Columbia was selected as the standard accession.

Nevertheless, the first whole genome investigations of Landsberg were published together with Columbia in 2000 (*Arabidopsis* Genome Initiative 2000). Large-insert bacterial artificial chromosomes (BAC) clones were used for the creation of physical maps and were integrated into genetic maps. The comparison of Columbia and Landsberg revealed a great number of single nucleotide polymorphism (SNPs) and insertion-deletions (InDels) (*Arabidopsis* Genome Initiative 2000).

The improvement of sequencing techniques led to the first *de-novo* sequencing of Landsberg using Illumina short reads with a resolution of N50: 198 kb by Korbinian Schneeberger in 2011 (Schneeberger et al. 2011). The potential of long-read sequencing techniques released by Pacific Bioscience's (PacBio) led again to an improved de-novo assembly of Landsberg in 2015. Landsberg was fully assembled in 38 contigs with an N50 of 11.2 Mb (Berlin et al. 2015). The first chromosome-level assembly of Landsberg was published by (Zapata et al. 2016), combining the advantages of long-read *de-novo* assembly and reference-based short read error correction to generate a high-quality genome including genetic maps. Further annotation improvements and description of genetic features, such as structural arrangements and sequence variations led to the currently-best available Landsberg genome resource (Jiao and Schneeberger 2020).

In contrast to Columbia, *Landsberg erecta* is a mutant variant of the original Landsberg accessions La-0 or La-1. Reference genomes the natural Landsberg accessions are still lacking and might be of interest to determine the impact of mutagenesis treatment by Rédei more than 60 years ago.

1.8 Function of *RPP1*-like cluster in *Arabidopsis*

In various *Arabidopsis* accessions, *RPP1* loci have been identified as conferring strain-specific resistance to *Hpa*. The *RPP1*-like cluster on chromosome 3 in *Ler-0* represents multiple genes that encode predominantly *RPP1*-like genes (R1 – R8), belonging to NLRs. While the molecular functions of *RPP1*-like genes (R1 – R8) remain unknown, they have been linked to hybrid incompatibilities (HI) between genetically distinct populations leading to reproductive isolation (Atanasov et al. 2014). The offspring between Ler-0 (Poland) and Kashmir2 (Kas2, central Asia) causes immune-related HI, due to an autoimmune response caused by a genetic interaction of

multiple independent *RPP1*-like genes (favoured R2, R3, R4 and R8) (Ler-0), and non-NLR alleles of Strubbeling-family receptor-like kinases (SRF3, Kas2) (Atanasov et al. 2014). Backcrossing of SUPPRESOR OF LER/KAS2 INCOMPATIBILITY (sulki) mutants into the parental line Ler/Kas2 revealed ten intragenic mutations in all three major NLR domains (TIR, NBS, LRR) of *RPP1*-like R8, and a suppression of dwarfism and cell death responses at low temperatures (14 – 16 °C). In addition, partial suppression was observed for one mutation in the TIR-domain of *RPP1*-like R3 (Atanasov et al. 2014).

Interestingly *RPP1*-like R8 is a homolog of DANGEROUS MIX 2H (DM2H) in genetic hotspots of *Arabidopsis* accessions Uk-1 and Bla-1 (Chae et al. 2014), and shows homology to At3g44670 in Columbia (Col-0) (Alcázar et al. 2014). Immune-related HI where both interacting alleles encode for TIR-NLRs (TNL) have been described for *Arabidopsis* accessions of Umkirch (Germany) Uk-1 and Uk-3 (Tran et al. 2017; Bomblies et al. 2007). As such, TNL association of the DM1 loci from Uk-1 and DM2d from Uk-3 form an autoimmune signalling complex. Thereby, DM1 has been identified as a primary signal transducer, whereas DM2d triggers activation of DM1 by complex formation (Tran et al. 2017). Interestingly, pathoassays with a local Gorzów *Hpa* strain (*Hpa Gw*) on Ler-0/Kas2 hybrids lead to the QTL mapping of RPP7 as the causal locus (Atanasov et al. 2014), suggesting that known *Hpa* strains are currently not affected by *RPP1*-like genes in pathogen perception.

Although NLR activation and downstream signalling mechanisms are extensively studied, the extent to which this layer of protection against pathogens shapes plant microbial communities is hardly understood. Further, intra-species inventories on NLRs revealed a tremendous receptor diversity among accessions.

1.9 Thesis aims

This thesis aims to identify microbial colonizers that affect microbial assemblies in the phyllosphere, in relation to host-microbe and microbe-microbe interactions. I choose to focus on the ecological relevance of a polymorphic NLR cluster on the phyllosphere assemblage in natural populations of *Arabidopsis thaliana*. I aimed to verify the ecological effect of the *RPP1*-like gene cluster on the phyllosphere community within natural *Arabidopsis* populations of Gorzów Wielkopolski (former Landsberg an der Warthe). To do so, structural variations of the *RPP1*-like cluster were identified using whole genome sequencing of individual plants including natural *Arabidopsis* accessions of Landsberg (La-0). Further, the impact of the *RPP1*-like cluster on the phyllosphere microbiome was determined using an amplicon-based sequencing approach of bacterial 16S rRNA genes and eukaryotic 18S rRNA genes, simultaneously in cDNA and DNA profiles. This study reinforces the relevance of host genetic features, such as resistance genes affecting microbial consortia in the phyllosphere of *Arabidopsis*.

Results

2 Results

In this study, I aimed to identify several factors, such as geographical location, sampling year and host genetic factors impacting the phyllosphere community. To do so, I performed a comprehensive study on the local *Arabidopsis thaliana* population in the area of Gorzów Wielkopolski (in history known as Landsberg an der Warthe, Poland) over three consecutive years. We collected wild *Arabidopsis* samples in the centre of Gorzów Wielkopolski on the Cemetery "Cmentarz Świętokrzyski" (site 100 - 400), as well as north-east (500, Różanki) and south-west (600, Łupowo) of the city (see S-Fig. 1).

2.1 Whole genome sequencing of individual Gorzów *Arabidopsis* isolates

I established a Gorzów Plant Collection (GCP) representing single *Arabidopsis thaliana* lines from sampling sites 100 – 600, collected over three consecutive years (2016 -2018). To obtain insights in the genetic background of Gorzów *Arabidopsis*, I performed whole genome sequencing of sixteen individual representative *Arabidopsis* plants, spanning all sampled sampling sites (100 – 600). In addition, I sequenced two genetically distinct wild-type accessions. La-0 [1] CS1298 and La-0 [2] CS76538, derived from Landsberg an der Warthe. PacBio long-reads (22.2 – 139.8x), were used to perform *de novo* assemblies from various geographical sites in the area of Gorzów Wielkopolski. Contigs featuring N50 values from 0.36 Mb to 42.78 Mb and L50 values from 3 to 177 indicating a range of assembly qualities (see S-Tab. 1).

I arranged contigs according to the *Ler-0* reference genome (Goel et al., 2019). Contigs were concatenated to scaffold files representing pseudo-molecules named as chromosome 1 to 5 (chr 1-5). Illumina short-reads (53.7 – 171.0 x) were used for sequencing correction and gap filling. Thus, 73.66 - 95.31 % of paired reads were mapped to the corresponding *Arabidopsis* genomes.

Comparing our genome assemblies with the reference *Ler-0* (N50: 9.4 Mb, N50: 5, No. of scaffolds: 49) revealed seven high quality genome assemblies, such as 100K, 100AA, 300AD, 400D, 400L, 400Y and La-0 [2] (97 - 204 scaffolds). Further genome assemblies were generated showing <500 scaffolds (100H, 100S, 100AD, 600H, 600I, La-0 [1]) and >500 scaffolds (100D, 100Y, 200E, 300A, 500R). The total genome length varied from 114.95 to 144.20 Mb, compared to 118.5 Mb of the reference.

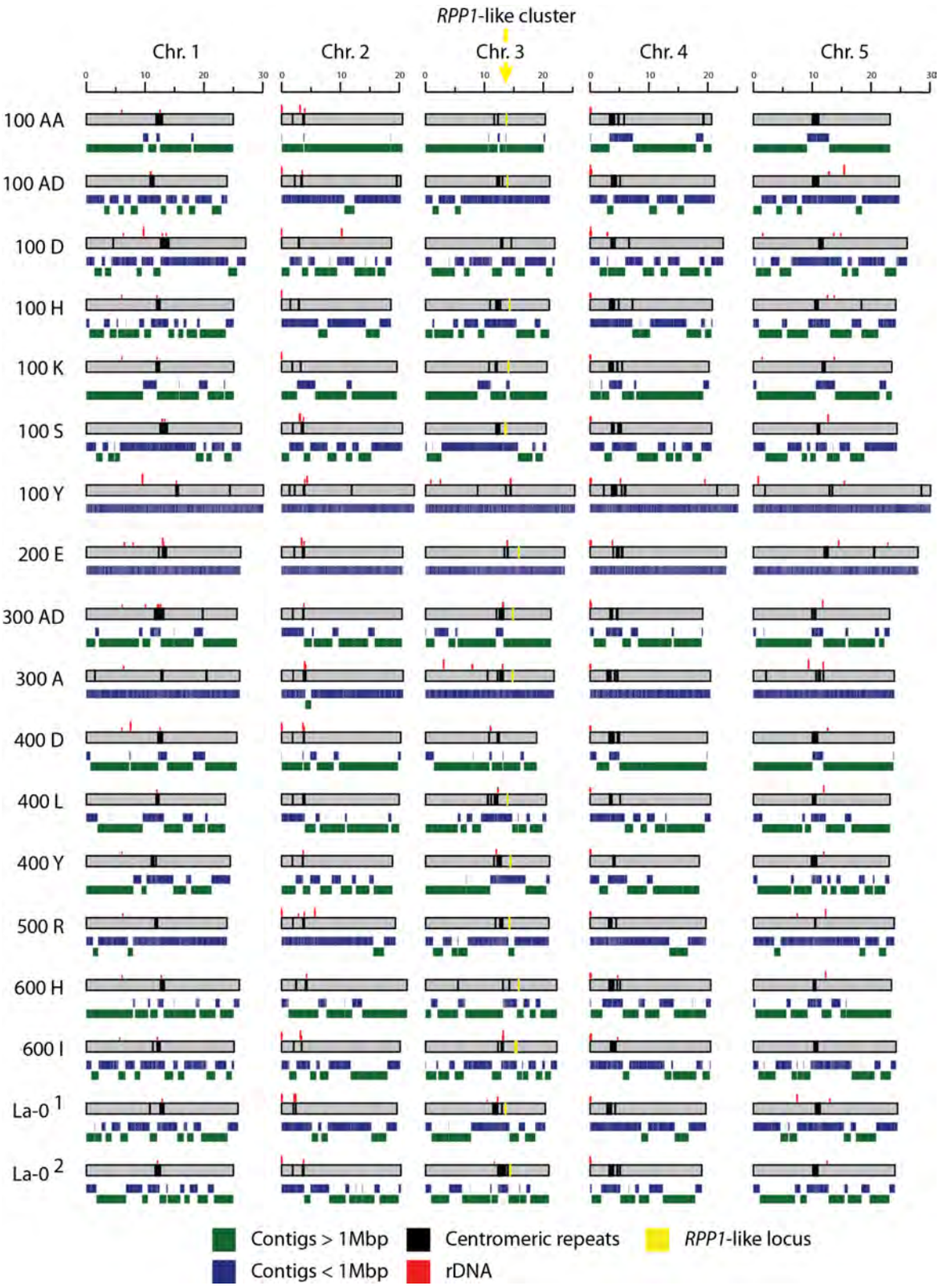


Fig. 2: Contig maps of 18 Gorzów *Arabidopsis* genomes including two La-0 references. Genome IDs representing individual plants, labelled according to their sampling sites (100-600). Reference genomes are displayed as La-0 [1] (CS1298) and La-0 [2] (CS76538). The grey bars represent each of the chromosomes, whereas centromeric repeats based on CEN180 are shown as black blocks. The *RPP1*-like cluster on chromosome 3, flanked by AT3G44600 (LB) and AT3G44690 (RB), is labelled by yellow marks. Contigs > 1 Mb are displayed as green bars, whereas contigs < 1 Mb are shown in blue. The locations of rDNA genes are marked with red ticks.

In summary, I annotated 27,240 - 34,601 protein-coding genes in each genome assembly, which is in the range of the gene annotations in the reference genome Columbia-0 (Col-0, 27,416 proteins) (Berardini et al. 2015). In addition, transposable elements were annotated in a range of 58.6 – 80.9 Kb.

2.1.1 PAN-genome analysis of a local *Arabidopsis* populations

I compared each genome assembly including Gorzów lines (100 - 600), as well as two natural accessions La-0 [1] (CS1298) and La-0 [2] (CS76538) of Landsberg an der Warthe (Gorzów Wielkopolski, Poland), against the reference genome *Landsberg erecta* (Ler-0, 2019, v2) to determine structural and sequential variation among genomes (see Fig. 3). Synteny maps were calculated for each genome on pseudochromosome-level. As an example, Fig. 3shows natural accessions La-0 [1] and La-0 [2] against Ler-0. Synteny maps of each Gorzów genomes are included in S-Fig. 2. The maps reveal a large overlap between genomes, showing 97.6 - 111.5 Mb syntenic regions against the reference. Whereas, non-syntenic regions varied from 2.2 - 10.4 Mb (see. S-Tab. 5). In total, 105.2 - 115 Mb could be aligned against the reference, leading to 4.2 - 34.1 Mb genome specific sequences.

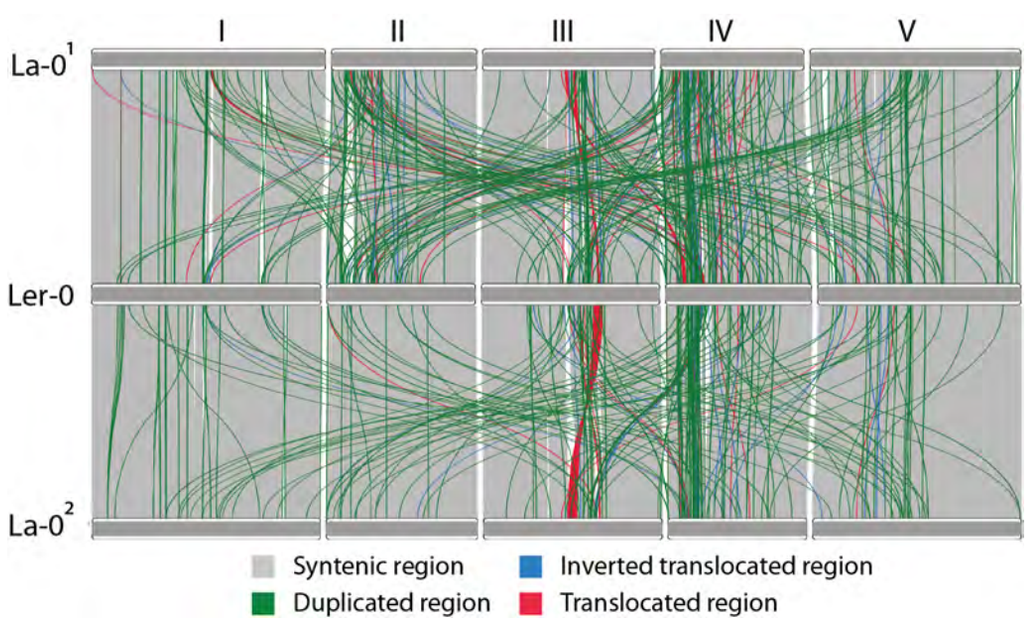


Fig. 3: Synteny maps of La-0¹ (CS1298) and La-0² (CS76538) against Ler-0 on pseudo-chromosome-level. Syntenic and non-syntenic regions were identified using SYRI and are colored as described in the figure legend.

Interestingly, syntenic regions were dominated by highly diverged regions representing 27.1 - 817.9 Kb of the genomes (see Fig. 4 and S-Tab. 6). I identified 5.0 - 184.9 Kb deletions and 0.7 - 7.7 Kb insertions. Single nucleotide polymorphisms (SNPs) were identified on 6.8 - 51.6 x 10^3 positions. In addition, I identified sequence divergence in syntenic regions varying in copy-gain- (1.7 - 32.6 Kb) and copy-loss variation (0.8 - 87.7 Kb).

Translocated regions displayed overall similar sequence divergence compared to collinear regions. However, variation across genomes was generally higher. As such, I specified 7.7 Kb - 1.75 Mb highly diverged regions. Sequence variation belonging to deletions and insertions varying between 0.8 - 49.5 Kb and 0.3 - 3.5 Kb. SNPs in translocated regions were identified on 2.3 - 26.3 x 10^3 positions, which is lower compared to collinear sequences. Copy-gain variations were two-fold higher in average, ranking from 0.0 - 88.8 Kb. Copy-loss variation was comparable to syntenic regions and varied from 0.0 - 114.9 Kb in translocated regions.

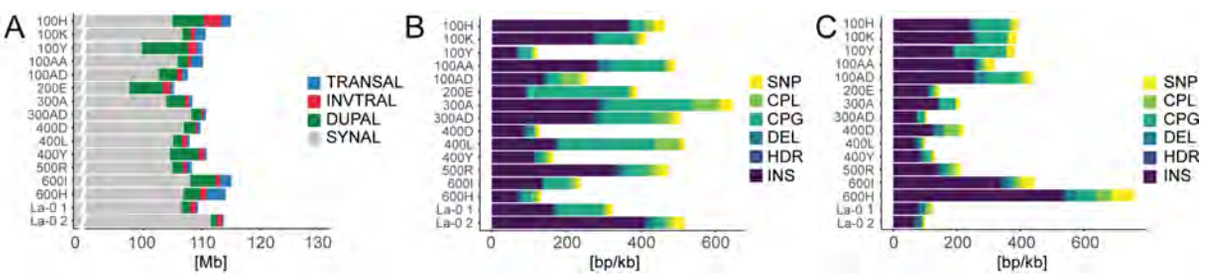


Fig. 4: Structural and sequence differences between Gorzów genomes against Ler-0. (A) Structural differences of alignments are displayed as syntenic (SYNAL), duplicated (DUPAL), inverted translocated (INVTRAL) and translocated regions (TRANSAL). Local sequence variations are described in syntenic (B) and rearranged regions (C) and displayed per kb. Variations are grouped into single nucleotide polymorphism (SNP), copy loss in query (CPL), copy gain in query (CPG), deletion in query (DEL), highly diverged regions (HDR) and insertion in query (INS).

Taken together, these results demonstrate low intra-species variations between Gorzów ecotypes. In addition, I verified that La-0 [2] (CS76538) shows the highest similarity in syntenic regions in relation to the reference.

2.1.2 Resolving *RPP1*-like gene cluster structure

I next investigated structural variation of the *RPP1*-like resistance gene loci on chromosome 3 in natural Gorzów ecotypes (100-600) and Landsberg (La-0) accessions. For simplicity, I will refer to genomes containing the Ler-0 *RPP1*-like alleles as RPP1+, while lacking the *RPP1*-like alleles as RPP1-. In the current analysis, Columbia and Landsberg served as references accounting for RPP1- (Col-0) and RPP1+ (Ler-0) haplotypes. Genomic sequences of potential *RPP1*-like loci, flanked by border genes AT3G44600 and AT3G44690, were extracted and used to calculate genomic distances (see 4.10). Notably, the sequence length of references sequences varied from 69.37 Kb in

Columbia to 158.01 Kb in Landsberg. Principle component analysis was performed on genomic distances of *RPP1*-like cluster sequences (see Fig. 5B). Interestingly, most *RPP1*-like sequences from Gorzów *Arabidopsis* ecotypes cluster to the reference sequences of Columbia (Col-0, *RPP1*-) and Landsberg (Ler-0, *RPP1*+).

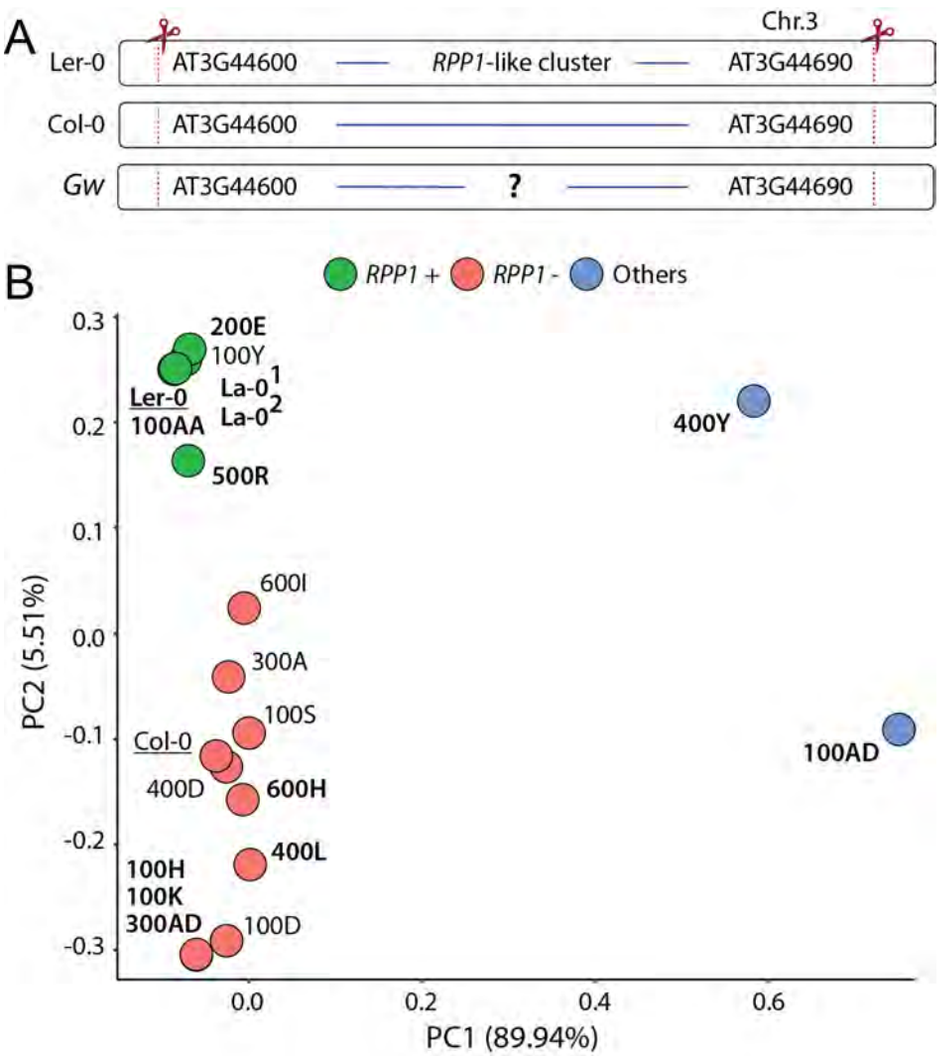


Fig. 5: Genetic distance of *RPP1*-like loci in Gorzów *Arabidopsis*, natural Landsberg and reference accessions. A) Genomic sequence extraction of *RPP1*-like loci on chromosome 3. B) Principle component analysis displaying genomic distances of extracted *RPP1*-like cluster sequences. Each datapoint represent one *RPP1*-like cluster sequence. K-means clustering using the R package 'ggfortify', revealed genomic distances between three groups: *RPP1*+ (green), *RPP1*– (red) and Others (blue). Underscored labels show reference sequences of Columbia and Landsberg. Bold labels display Gorzów lines used for further structural variation analysis.

Notably, *RPP1*-like loci of La-0[1] and La-0[2] were highly associated with Ler-0, indicating major sequence collinearity between standard laboratory and natural Landsberg accessions. In addition, *Gw* 100AA, *Gw* 100Y, *Gw* 200E and *Gw* 500R are affiliated to *RPP1*+ in the PCA. On the other hand, the majority of *RPP1*-like loci have been associated to *RPP1*–. Interestingly, *RPP1*– depicts

a higher distribution in the two-dimensional ordination, accounting for higher sequence divergence to the standard Columbia. However, *RPP1*+ and *RPP1*– alleles are predominantly separated on the PC2, showing 5.51 % explained variance. Further, an outgroup (named as 'Others') has been separated on the PC1 (89.94 %), including *Gw* 100AD and *Gw* 400Y. Taken together, it seems that the outgroup show less similarity to *RPP1*-like loci in Landsberg and Columbia based on genomic distances. To obtain further insights in the gene structure of *RPP1*-like loci in Gorzów and Landsberg, *ab initio* gene model annotations were used to identify *RPP1*-like genes, based on *RPP1*-like protein sequences [Genbank: FJ446580.1] published by (Alcázar et al. 2014). To do so, contiguous sequences of Landsberg (Ler-0, La-0) and Gorzów ecotypes (*Gw*) were included in the following analysis.

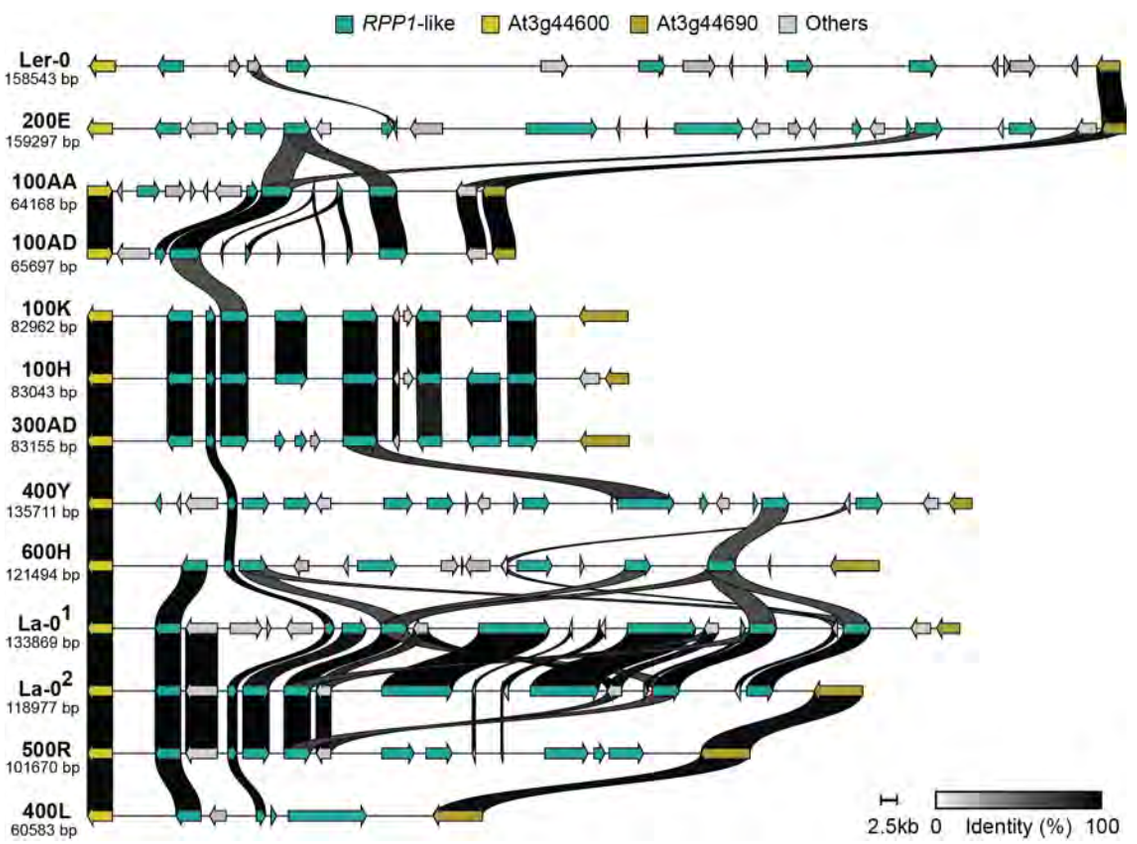


Fig. 6: Synteny among *RPP1*-like loci of Gorzów and reference genomes. Resistance gene loci were selected by flanking border genes - AT3G44600 (LB) and AT3G44690 (RB) coloured in yellow arrows. *RPP1*-like gene annotations including *RPP1*-like R1 - R8 are displayed in cyan. Other genes are labelled as grey bars. Sequence similarities are based on protein sequences and range from 75-100% displayed as colour gradient between arrows.

Global alignments were performed between all protein-coding regions to identify homologs within the *RPP1*-like loci (see Fig. 6). Interestingly, *RPP1*-like loci displayed heterogeneous sequencing lengths of 60 to 158 Kb. These are comparable to the reference sequences in Columbia and Landsberg, as shown above. Within the *RPP1*-like loci, I identified multiple *RPP1*-like protein annotations. Nevertheless, due to high similarity of full length protein sequences, I was not able to

identify single *RPP1*-like genes (R1 - R8). The *RPP1*-like locus of the reference Ler-0 showed similarities to six out of eight *RPP1*-like genes. In Gorzów ecotypes, I identified 4 - 13 *RPP1*-like coding regions. Interestingly, eight *RPP1*-like coding regions were annotated in the natural Landsberg accessions La-0 [1] and La-0 [2], as well as in *Gw* 100H and *Gw* 100K matching our expectation of eight *RPP1*-like genes within the reference locus.

Taken together, a principle component analysis on genomic distances displayed high similarities between *RPP1*-like loci of natural Landsberg (La-0) accessions and Ler-0. In addition, *RPP1*-like loci in natural Gorzów *Arabidopsis* showed two cluster formation, which were predominantly associated with *RPP1*+ and *RPP1*– haplotypes, including Landsberg and Columbia as references. Gene model annotations were used to identify *RPP1*-like coding regions within the extracted sequences. Sequence length and *RPP1*-like copy numbers varied highly between individual Gorzów genomes, collected in three natural populations of Gorzów Wielkopolski.

2.2 *RPP1*-like genotyping of wild *Arabidopsis* populations

The genotyping of local *Arabidopsis* populations around Gorzów Wielkopolski was conducted using a droplet digital PCR (ddPCR) approach targeting *RPP1* (-like) genes (see 4.8). To do so, 332 individual *Arabidopsis* plants from six sampling sites (100 – 600) were analysed in three consecutive years (see Fig. 7A). *RPP1* (-like) copy numbers showing a binomial distribution in our data set were obtained, correlating with an average *RPP1*-like CNV of 8 copies for Col-0 (negative control, *RPP1*-) and 16 copies for Ler-0 (positive control, *RPP1*+). To estimate the number of clusters, a gap statistic clustering (K-means) was performed (see Fig. 7B). These results displayed high densities of *RPP1*- (absence of the *RPP1*-like cluster) and *RPP1*+ (presence of the *RPP1*-like cluster) in natural Gorzów *Arabidopsis*. Since I identified a binomial distribution of *RPP1* (-like) CNVs around *RPP1*- and *RPP1*+, I suggest a gene gain and gene loss of *RPP1* (-like) genes in natural Gw *Arabidopsis*. I further investigated how *RPP1*-like gene copies are represented within each geographical location.

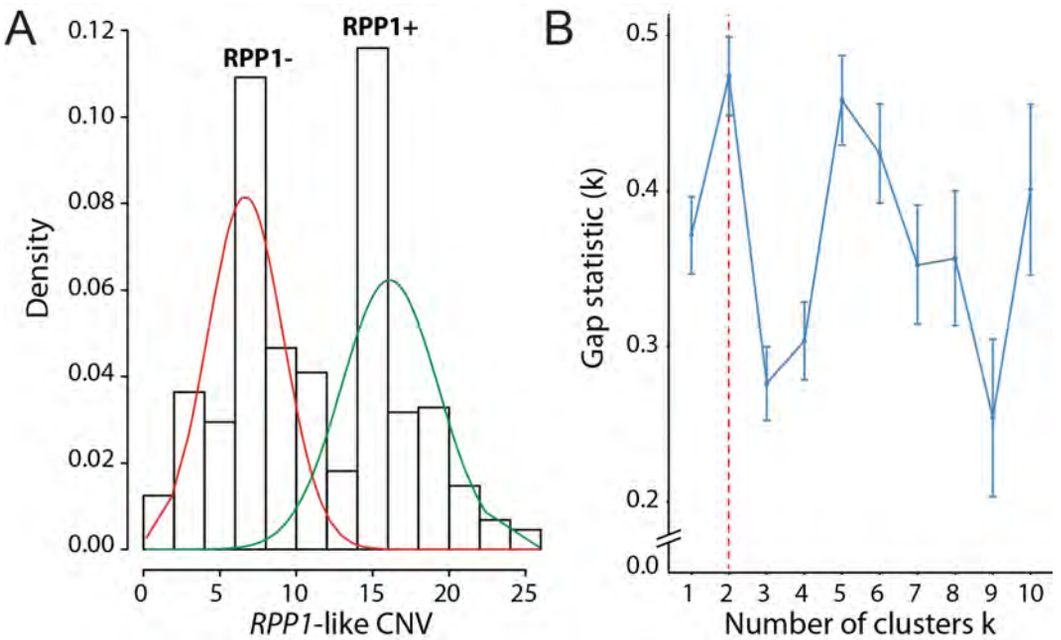


Fig. 7: *RPP1* (-like) copy number variation in Gorzów *Arabidopsis*. A) Binomial distribution of *RPP1*-like CNVs showing high densities of *RPP1*- and *RPP1*+. B) Optimal K-means clustering of *RPP1*-like CNVs was calculated using gap statistics in R (packages factoextra 'fviz_nbclust'). Red dot line displays the optimal number of clusters.

Distribution of *RPP1*-like CNVs and their frequency per location are plotted as histograms (see Fig. 8). Interestingly, central Gorzów Wielkopolski sites (100 – 400) display a diverse repertoire of *RPP1*-like CNVs among sites.

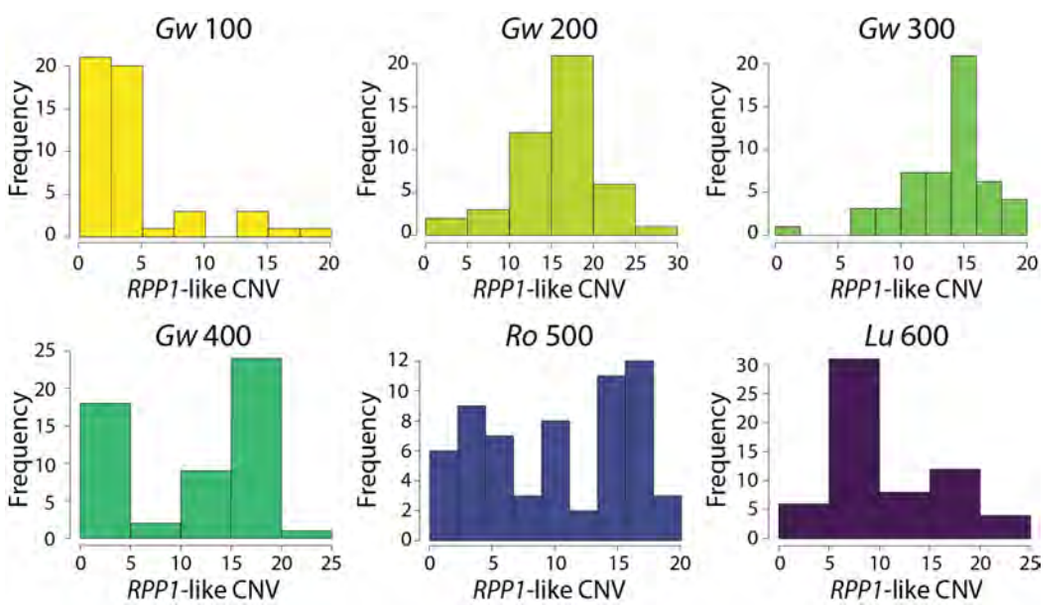


Fig. 8: Distribution of *RPP1*-like CNVs in local *Arabidopsis* populations over three consecutive years. Sampling sites are displayed as 100 -400 (Gw), 500 (Ro) and 600 (Lu).

While site 100 predominantly shows the *RPP1*-, site 200 and 300 frequently show *RPP1*+. Further, site 400 shows *RPP1*- and *RPP1*+ in similar ratios. In addition, sampling sites outside of Gorzów Wielkopolski showed either an equal distribution of *RPP1* haplotypes (site 500) or else preferentially *RPP1*- haplotypes (site 600). Our data suggest a complex *RPP1*-like gene cluster displaying divergence in CNVs among local *Arabidopsis* populations.

2.3 Phyllosphere microbiota of Gorzów *Arabidopsis thaliana*

This study aims to investigate how active phyllosphere microbiota vary in wild *Arabidopsis* populations in the area of Gorzów Wielkopolski. To do so, we collected surface-colonizing microbes (epiphytes | phylloplane) and cytosolic microbes (endophytes | endosphere) from 180 individual *Arabidopsis* rosettes over three consecutive years (2016 – 2018). To obtain insights into the active phyllosphere microbiome, a simultaneous extraction of DNA and RNA was performed on each sample and conducted an amplicon sequencing approach (hereafter described as DNA- and cDNA-seq) focusing on bacterial 16S V4/V5 regions and eukaryotic 18S rRNA V8/V9 regions. A total 1440 sequencing libraries, were sequenced on four MiSeq runs. The whole sample set was fully randomized over the process of data generation. In addition, an internal 18S rRNA spike-in was included for further downstream analysis. Further, a DNA sequencing approach was performed targeting bacterial 16S (V4/V5) rRNA, eukaryotic 18S (V8/V9) rRNA and fungal ITS2 regions to compare our results with a standard DNA amplicon sequencing approach (see Supplementary chapter 7.5).

2.3.1 Microbial richness and composition differs between local *Arabidopsis* population

Initially, I estimated the microbial richness of epiphytic and endophytic samples across all sampling locations (100 - 600). Overall, similar patterns were observed in microbial richness between DNA and RNA samples, suggesting that the active microbiome based on RNA amplicons reflects microbial profiles of classical DNA sequencing. While I identified comparable microbial diversities of *Arabidopsis* between closely related sampling sites in Gorzów Wielkopolski (100 - 400), more distinct sites around the city show higher diversities in general (500 - 600) in epiphytic and endophytic samples (see Fig. 9). In detail, bacterial richness of epiphytic DNA samples revealed higher richness in Ro (500) compared to Gw2 (200) and Gw3 (300) (see Fig. 9A). In line, bacterial richness of endophytic DNA samples from Gw1 – Gw4 (100 - 400) were significantly lower compared to Ro (500) and Lu (600).

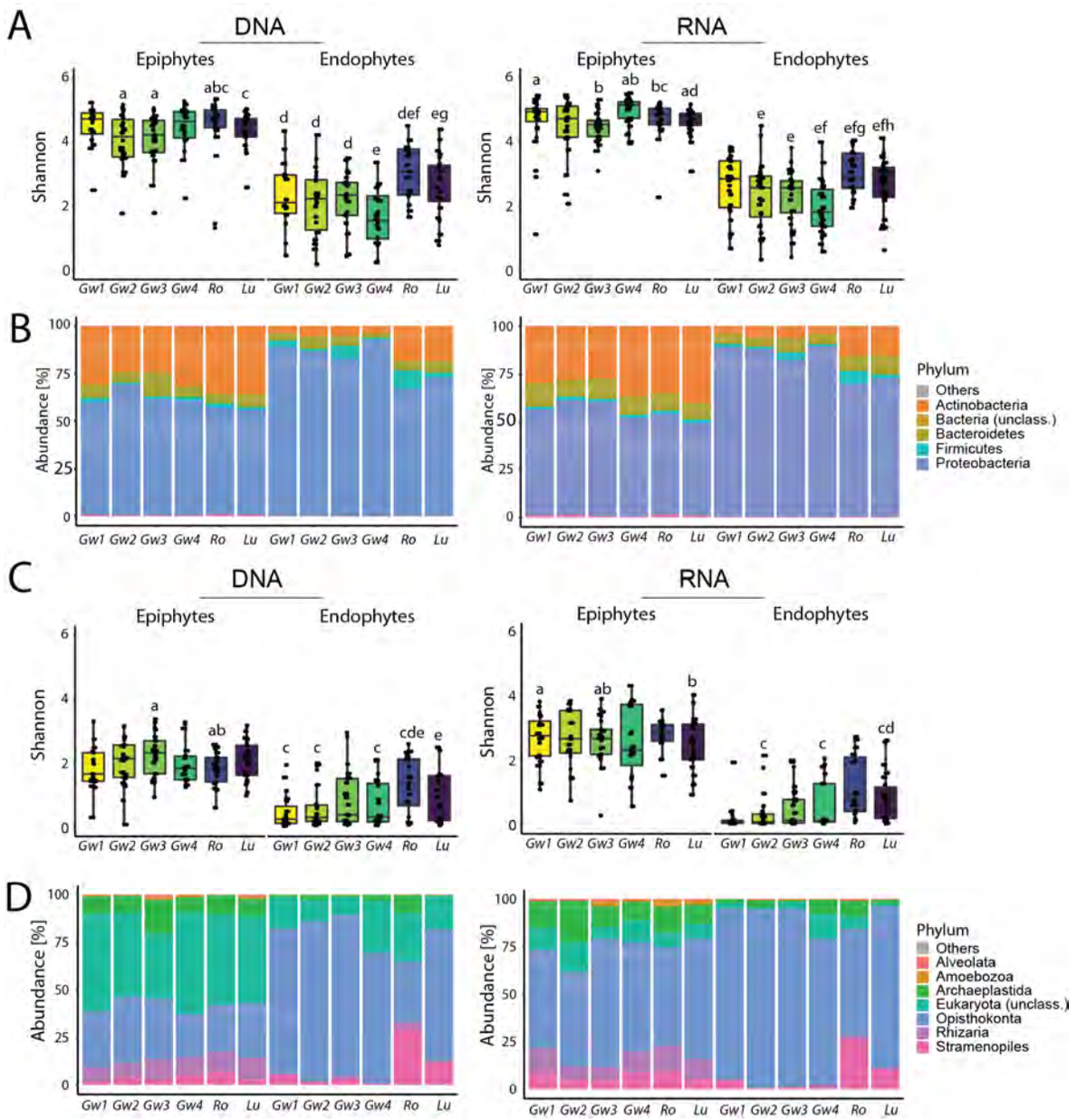


Fig. 9: Microbial richness and taxonomic assignment in phyllosphere microbiota across sampling sites. Samples are grouped by Nucleic Acid x Compartment x Sites. A) Bacterial richness displayed as Shannon index. Statistics are calculated using Kruskal-Wallis test, $p < 0.05$. B) Taxonomic plots showing relative abundances of bacterial phyla. Phyla showing < 1 % relative abundance were summarized as Others. C) Eukaryotic richness displayed as Shannon index. D) Taxonomic plots showing relative abundances of eukaryotic phyla.

Microbial profiling of the active microbiome revealed higher bacterial richness of epiphytes in Ro (500) and Lu (600) compared to Gw1 and Gw3. Interestingly, site Gw4 (400) showed higher bacterial richness of epiphytes was identified in Gw4 (400) compared to all other sites. Similar to DNA samples, active endophytes from Gw2 to Gw4 (200 - 400) showed lower bacterial richness compared to Ro and Lu (500 - 600). To obtain insights in the microbial community composition on the phyllosphere, I performed a taxonomic assessment on phylum level (see Fig. 9B).

I performed a statistical analysis using pairwise-Wilcoxon tests on Nucleic Acid x Compartment x Site against each phylum. Proteobacteria (54.90 – 92.97 % [DNA], 48.26 – 88.58 % [RNA]), Actinobacteria (3.80 – 35.35 % [DNA], 3.85 – 40.11 % [RNA]), Bacteroides (2.16 – 12.23 % [DNA], 4.25 – 12.19 % [RNA]) and Firmicutes (0.21 – 9.61 % [DNA], 0.22 – 6.83 % [RNA]) are the most abundant bacterial phyla in our data set. Low abundant microbes are represented by unclassified bacteria (0.05-1.21% [DNA], 0.16 – 0.38 % [RNA]). By comparing the eukaryotic richness of active and dormant microbes, I determined higher richness in epiphytic RNA samples. This might be explainable by the nucleic acid input concentration of RNA samples, which was at least twice as much as DNA samples. The nucleic acid concentration can impact PCR amplification upon library preparation and affects the detection level of low abundant microbes. Nevertheless, epiphytic samples displayed comparable richness between sampling sites in DNA and RNA samples. Further, the eukaryotic richness of endophytic samples was significantly lower ($p < 0.05$) in Gorzów Wielkopolski (site 100-400) compared to Różanki and Łupowo (500-600). Taxonomic profiles on phylum level revealed a similar pattern between RNA and DNA samples, considering leaf compartments and sampling sites. However, DNA samples displayed higher relative abundances of unclassified eukaryotes (kingdom level), especially in epiphytic samples.

Opisthokonta (22.16 – 84.58 % [DNA], 49.67 – 94.32 % [RNA]), Archaeplastida (0.76 – 9.45 % [DNA], 1.02 – 21.81 % [RNA]), Stramenopiles (0.33 – 29.60 % [DNA], 0.33 – 26.27 % [RNA]) and Rhizaria (0.23-11.26 % [DNA], 0.02 – 13.36 % [RNA]) are the most abundant eukaryotic phyla of *Arabidopsis* phyllosphere microbiota across all sampling sites. Further, Amoebozoa (0.09 – 1.35 % [DNA], 0.35 – 3.35 % [RNA]) and Alveolata (0.09 – 1.34 % [DNA], 0.39 – 1.40 % [RNA]) could be identified, preferentially in epiphytic samples. Unclassified microbes belonging to the kingdom Eukaryota were identified mainly in epiphytic DNA samples (9.47 -54.05 % [DNA], 0.90 – 15.33 % [RNA]).

While I identified that microbial richness remains rather stable across Gorzów Wielkopolski sampling sites (100 - 400), samples taken in Różanki (500) and Łupowo (600) showed higher microbial richness.

2.3.2 Microbial diversity remains stable across sampling years and sites

In this chapter, I aimed to identify metadata traits affecting microbial community compositions. I performed a multivariate analysis calculating Bray-Curtis dissimilarities. Thus, samples were grouped by metadata features, such as nucleic acid x year, compartment or sampling site. OTU-tables have been rarefied according to bacterial DNA (1001 reads) and RNA (1029 reads), as well as eukaryotic DNA (15743 reads) and RNA (15159 reads). Notably, after rarefaction I removed spike-in reads for the current analysis. Initially, I determined year-to-year variation in our dataset. Thus, bacterial diversity was significantly different ($p < 0.017$ [DNA], $p=0.001$ [RNA]) accross three consecutive years (2016 to 2018). While eukaryotic diversities remained similar for DNA samples, RNA samples were significantly different between sampling years ($p < 0.008$). I further determined sample variations across leaf compartments, previously described as epiphytes and endophytes. The microbial diversity of bacteria and eukaryotes differ consistently between epiphytes and endophytes ($p < 0.001$ [DNA, RNA]).

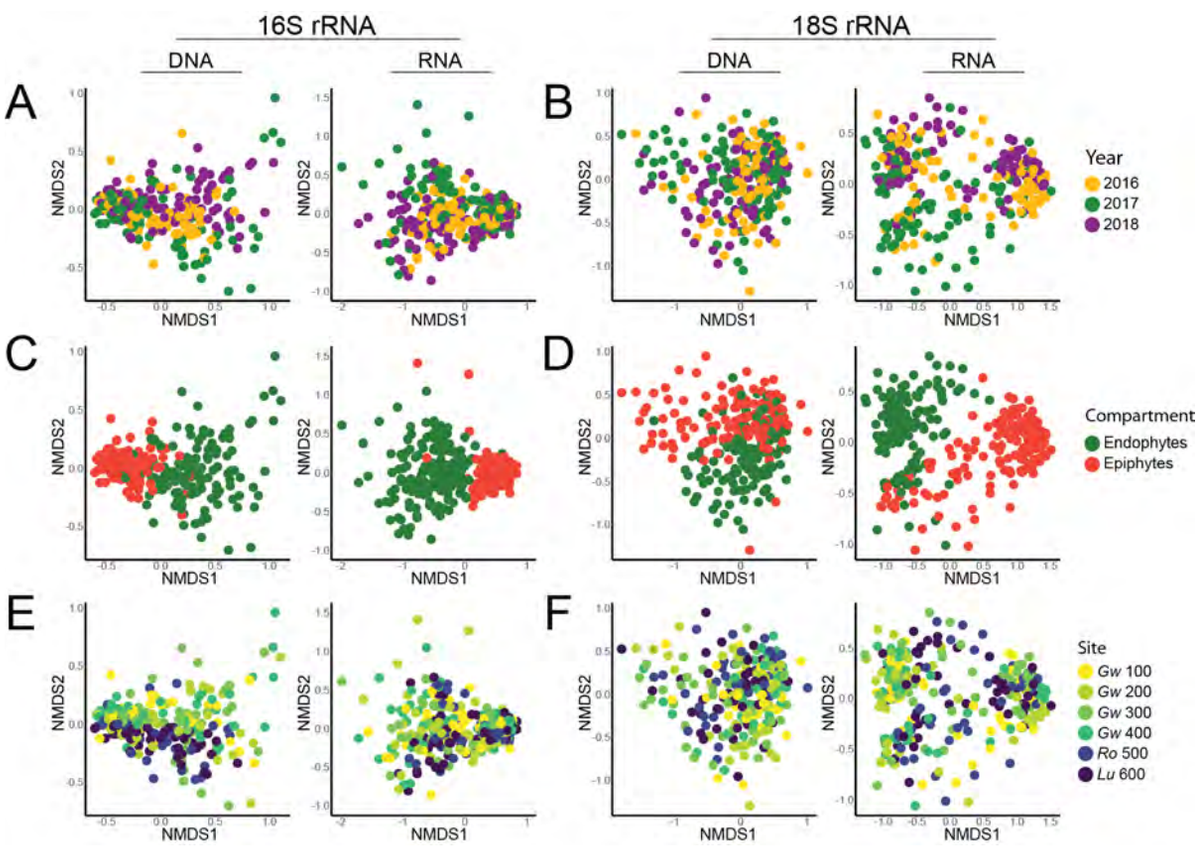


Fig. 10: Multivariate analysis of phyllosphere microbiome samples grouped by year, compartment and sampling sites. Non-metric multidimensional scaling (NMDS) plot on Bray-Curtis dissimilarities representing NMDS1 and NMDS2 of the 2D ordination. Single dots representing microbiomes of individual samples. NMDS plots showing bacterial 16S rRNA amplicons (A,C,E) and eukaryotic 18S rRNA amplicons (B,D,F). Color-coding displays sample groups of sampling years (A, B), leaf compartments (C, D) and sampling sites (E, F).

To capture the impact of geographical locations in the area of Gorzów Wielkopolski, I grouped samples according to sampling sites. Hence, bacterial diversities were consistently different between all sampling sites across the three years ($p < 0.05$ [DNA, RNA]). Interestingly, eukaryotic diversities were significantly different between Gorzów sites (100-400) and Różanki (500) ($p < 0.006$ [DNA], $p < 0.01$ [RNA]). In addition, Gw3 (300, $p=0.021$) and Różanki (500, $p=0.003$) were significantly different from Łupowo (600) in RNA samples. These results suggest that sampling years and leaf compartment are major factors in shaping microbial communities. However, I also observed a large effect on bacterial diversities between local *Arabidopsis* populations. In addition, eukaryotic diversities remained stable across central Gorzów Wielkoposlki sampling sites (small geographical distances). Further, sampling sites in the surround area of Gorzów Wielkopolski (Różanki and Łupowo) displayed significantly different eukaryotic diversities.

I performed a beta-dispersion analysis on Bray-Curtis dissimilarities (see S-Fig. 5 - 6) to validate sample-to-sample variability for bacteria and eukaryotes across multiple sample groups (year, compartment, site), which were described above. Considering the year-to-year variability, samples collected in 2016 showed a lower distance to centroid for bacterial (see S-Fig. 5A) and eukaryotic diversities (S-Fig. 6A) compared to 2017 ($p < 0.05$ [18S DNA], $p < 0.001$ [16S RNA]) and 2018 ($p < 0.05$ [18S DNA], $p < 0.01$ [16S RNA]). Notably, the year-to-year effect on sample compositions was not significant in bacterial DNA-seq and eukaryotic cDNA-seq suggesting certain inconsistency between both sequencing approaches. Further, the highest sample-to-sample variation was observed between leaf compartments (see S-Fig. 5 - 6B). While bacterial epiphytes unveiled lower sample variations contrasting to endophytes ($p < 0.001$ [DNA, RNA]), the opposite trend was observed for eukaryotes ($p < 0.001$ [RNA]) showing lower sample variation within the endophytic fraction. Nevertheless, sample variation in eukaryotic were higher compared to bacterial profiles. Interestingly, sample-to-sample variation were comparable between different sampling sites using cDNA-seq. In contrast, DNA-seq revealed a lower sample variability of outlier sites (Różanki, Łupowo) to particular sampling sites in Gorzów Wielkopolski. Interestingly, the sampling site Gw4 (400) displayed the most disparate distance to centroid to outlier sites ($p < 0.05$: Ro, Lu [16S DNA], Ro, Lu [18S DNA]), as well as to *RPP1*-like hotspot sites ($p < 0.05$: Gw1/2/3 [16S DNA], Gw3 [18S DNA]) in the surrounding area.

Those findings lead to the hypothesis that bacterial communities are more stable across environmental features (year, sampling site) and leaf compartments. In general, higher fluctuations were observed for eukaryotes. However, sample variability was predominantly determined to leaf compartments, followed by site-to-site and year-to-year variation, depending on the sequencing approach.

2.3.3 Persistent microbial core communities across geographical locations

Although the majority of leaf-associated microbes underlie huge fluctuations in terms of occurrence and abundance throughout a plant lifecycle, a small number of microorganisms occur consistently across sampling years, sampling sites and leaf compartments. I aimed to identify persistent microbes in Gorzów *Arabidopsis* populations also known as the "core" phyllosphere microbiome. Microbial core members are defined as bacterial and eukaryotic OTUs showing an occurrence of ≥ 85 % in the whole data set. While I identified 17 bacteria and 2 eukaryotes as core members in RNA samples, 3 bacteria and 2 eukaryotes were consistently found in DNA samples (see Fig. 11). Bacterial core members were classified as Proteobacteria (13 OTUs), Actinobacteria (2 OTUs) and Bacteroides (2 OTU). Among bacterial core members, I identified leaf-colonizing taxa, such as *Sphingomonas* (Otu0000002, Otu0000066, Otu0000084), *Pseudomonas* (Otu0000003, Otu0000009), *Methylobacterium* (Otu0000005, Otu0000014) and *Variovorax* (Otu0000008) in RNA samples. Consistently between RNA and DNA, I found *Sphingomonas*, *Methylobacterium* and *Variovorax* to be present in the bacterial core community. In addition, I identified *Flavobacterium succinicans* (Otu0000011), *Devosia* (Otu0000017), *Brevundimonas* (Otu0000018), *Hymenobacter* (Otu0000020), *Rhizobium* (Otu0000022), *Rathayibacter caricis* (Otu0000025), *Salinibacterium* (Otu0000049), *Bosea* (Otu0000046) and *Aureimonas* (Otu0000071) in RNA samples. Multicellular core members were identified as the alfalfa weevil *Hypera postica* (OTU000002) and a terrestrial green algae *Chloroidium saccharophilum* (OTU000065), which were consistently found in DNA and RNA samples. Interestingly, the most abundant microbial core members across Gorzów Wielkopolski (100-400), Różanki (500) and Łupowo (600) were *Sphingomonas faeni* (Otu0000002), *Methylobacterium adhaesivum* (Otu0000005), Pseudomonas (Otu0000003), Variovorax (Otu0000008), *Hypera postica* (OTU000002) and *Chloroidium saccharophilum* (OTU000065) (see Fig. 11). These results indicate the majority of the phyllosphere microbiome are subject to fluctuation, whereas a restricted number of microbes are able to persist on *Arabidopsis* leaves across local populations.

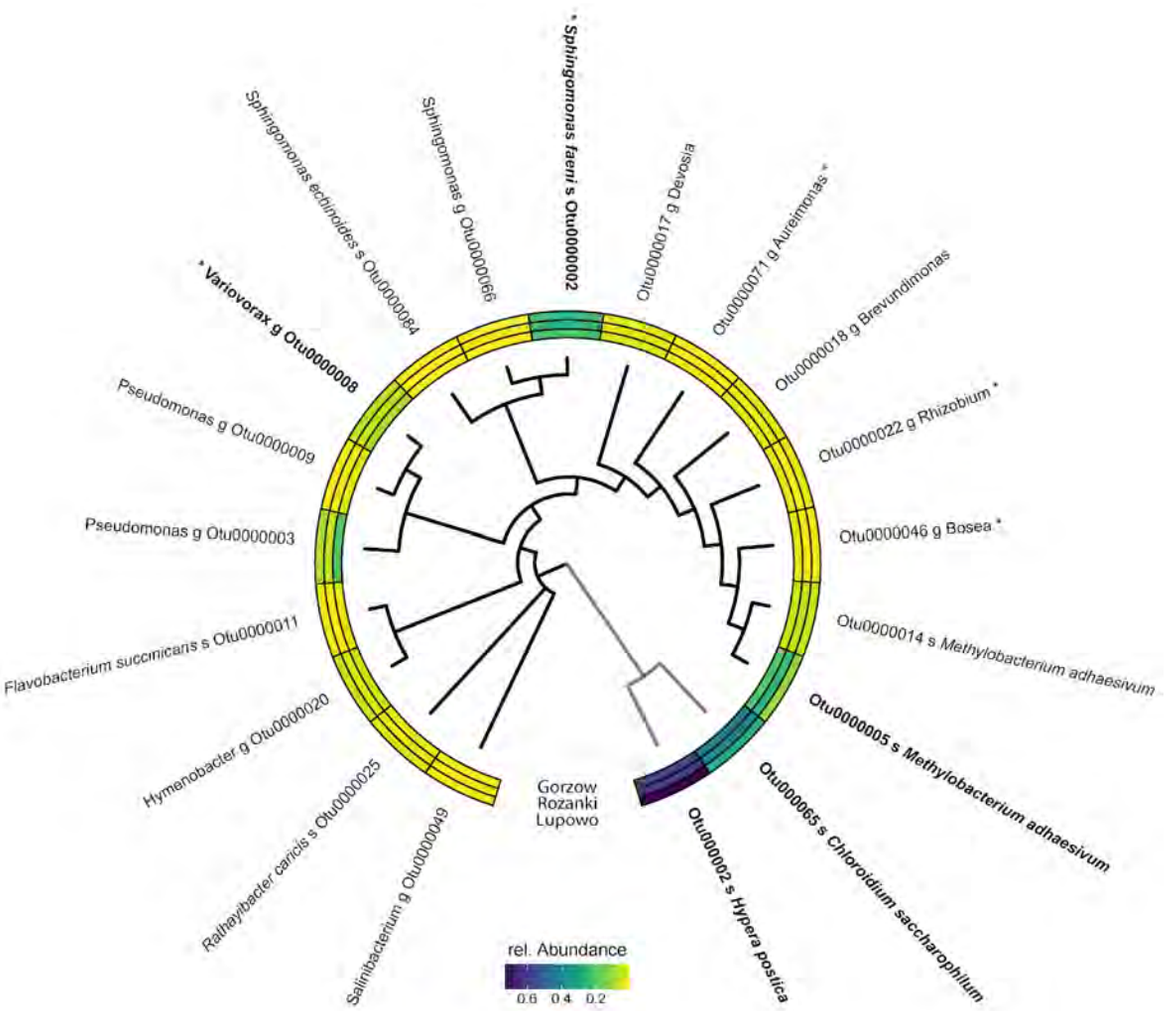


Fig. 11: Microbial core taxa of Gorzów *Arabidopsis thaliana*. Persistent microbial taxa across sampling years and sites are identified as OTUs showing ≥ 85 % occurrence of all DNA or RNA samples. Phylogenetic trees are calculated using representative sequences of OTUs. Bold description correspond to co-occurring core taxa in DNA and RNA samples.

2.3.4 Host resistance genes mediating bacterial community in the phyllosphere

In the previous chapter, I described multiple traits affecting microbial community compositions, such as sampling year, leaf compartments and geographical location. In addition, genotyping of wild *Arabidopsis thaliana* plants from all sampling sites (sites 100 – 600), in respect of *RPP1*-like gene copy numbers, revealed a complex resistance gene repertoire in the area of Gorzów Wielkopolski. In this chapter I aimed to identify the impact of host genetic factors, such as *RPP1*-like gene copies on the phyllosphere microbiome. Therefore, I grouped samples (Nucleic Acid x Compartment x *RPP1*) in *RPP1*+ (high *RPP1*-like copy number) and *RPP1*- (low *RPP1*-like copy number) according to our *RPP1* genotyping (see 5.5). Microbial richness (Shannon index) was calculated for bacteria and eukaryotes in DNA and RNA samples (see Fig. 12 A-B). Interestingly,

bacterial richness of RNA samples was significantly lower ($p < 0.05$) in *RPP1*+ compared to *RPP1*- . Although similar trends were observed for bacterial richness in DNA samples, no differences were observed for microbial richness of eukaryotes. Further, I performed a principal component analysis calculating Bray-Curtis dissimilarities on *RPP1* samples (see Fig. 12 C-D).

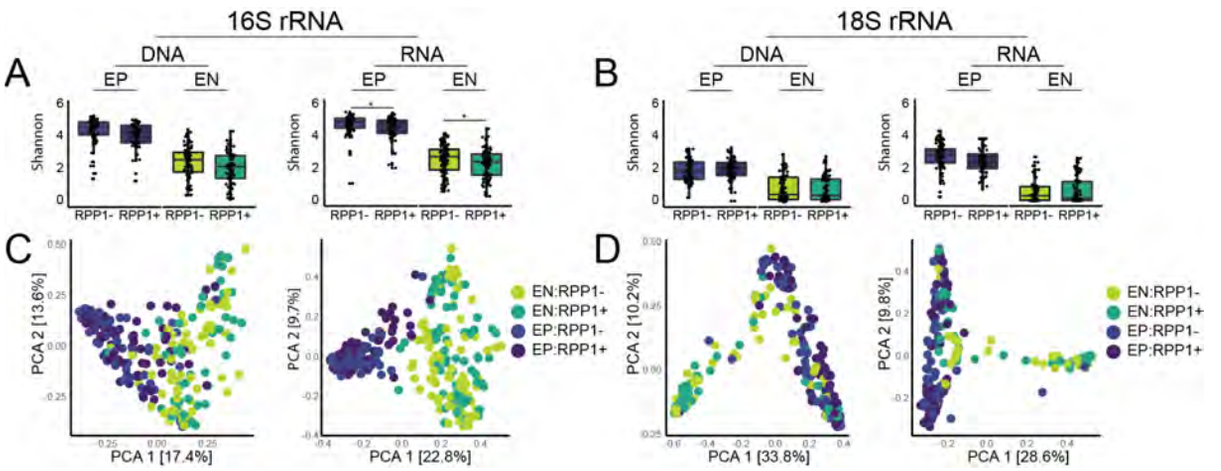


Fig. 12: Microbial diversity of phyllosphere samples grouped by NucleicAcid x Compartment x *RPP1*. Shannon indices bacteria (A) and eukaryotes (B). Principal component analysis on Bray-Curtis dissimilarities. Explained variance are calculated on the first two axes PCA 1 and PCA 2 within the ordination system. Kruskal-Wallis Test: *p < 0.05.

I was able to confirm significantly different bacterial diversities between *RPP1* sample groups (Nucleic Acid x Compartment x *RPP1*). Hence, bacterial diversity of epiphytic (Permanova, P=999; p=0.001 [DNA], p=0.007 [RNA]) and endophytic compartments (Permanova, P=999; p=0.024 [DNA], p=0.017 [RNA]) were significantly different, in respect of *RPP1*-like copy numbers (*RPP1*+ / *RPP1*-). The microbial diversity of eukaryotes remained stable in *RPP1*+ and *RPP1*- sample groups (Nucleic Acid x Compartment x *RPP1*). Further, a beta dispersion analysis calculating distance to centroid (Geographical Location x Compartment x *RPP1*) unveil lower sample variability in *RPP1*+ for bacterial profiles (see S-Fig. 7) in Łupowo ($p < 0.05$ [RNA]) independent of leaf compartments. In contrast, sample-to-sample variation was higher in *RPP1*+ for eukaryotic profiles (see S-Fig. 8) in Gorzów. This results suggest lower bacterial diversities in *RPP1*+ samples and affected bacterial diversities in respect of *RPP1*-like haplotypes in our Gorzów *Arabidopsis* dataset. In addition, sample-to-sample variability was lower in RPP1+ haplotypes of Łupowo (bacteria), while higher in *RPP1*+ of Gorzów (eukaryotes).

Since bacterial communities were affected between *RPP1* haplotypes, I aimed to identify bacterial genera that might be affected in presence or absence of *RPP1*-like gene copies in local *Arabidopsis thaliana* along geographical locations. Thus, I grouped samples according to Nucleic Acid x Compartment x Geographical Location x *RPP1*. In detail, sampling sites were summarized into

geographical locations, like Gorzów Wielkopolski (Gorzów, site 100 – 400), Różanki (site 500) and Łupowo (site 600). By calculating bacterial richness of *RPP1*-genotypes between geographical locations, I obtained significantly ($p < 0.05$ [DNA, RNA]) lower diversities of endophytic bacteria in Gorzów Wielkopolski (see Fig. 13).

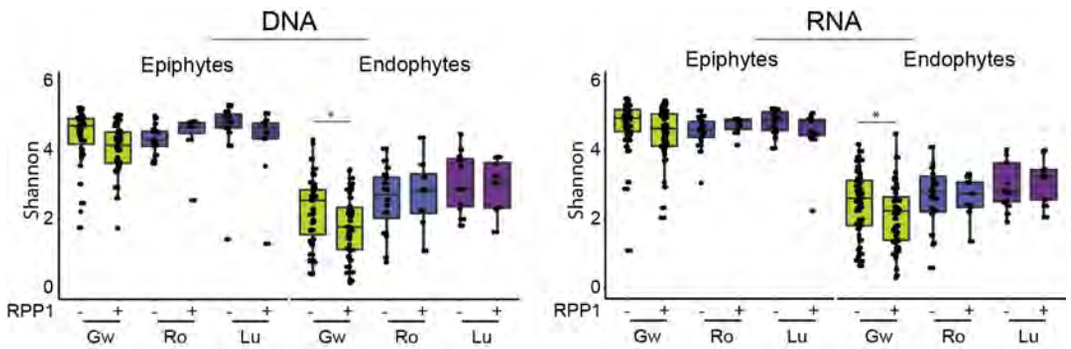
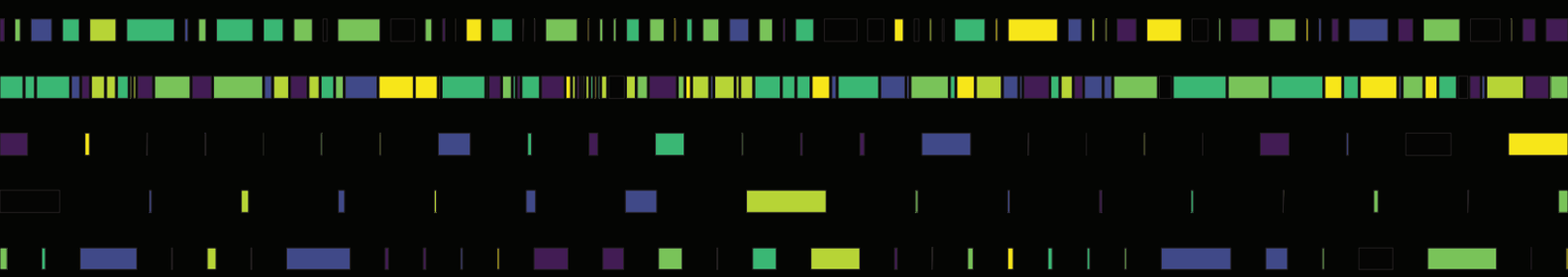


Fig. 13: Bacterial richness in relation to *RPP1*-like haplotypes among geographical location. Samples are grouped by Nucleic Acid x Compartment x Sites x *RPP1*. Sampling sites are grouped by geographical location into Gorzów Wielkopolski [Gorzów (100-400)], Różanki (500) and Łupowo (600).

Interestingly, I identified the bacterial genus *Sphingomonas*, which was significantly higher abundant in endophytic compartments of *Arabidopsis thaliana RPP1*+ genotypes, collected Gorzów Wielkopolski (sites 100-400, Wilcoxon-Test, p=0.0081 [DNA], p=0.029 [RNA]) and Łupowo (site 600, Wilcoxon-Test, p=0.0081 [DNA], p=0.0014 [RNA]). In terms of relative abundance, samples taken from Gorzów (23.38 % *RPP1*-, 24.83 % *RPP1*+) and Łupowo (26.14 % *RPP1*-, 34.56 % *RPP1*+) displayed higher abundances of *Sphingomonas* in *RPP1*+ genotypes. In contrast to that I observed the opposite case in Różanki (33.73 % *RPP1*-, 20.79 % *RPP1*+). We identified 99 OTUs belonging to *Sphingomonas* in our data set. Thus, 14 OTUs showed a relative abundance > 1% of all *Sphinomonas* OTUs, such as Otu0000002, Otu0000066, Otu0000084, Otu0000095, Otu0000139, Otu0000181, Otu0000229, Otu0000238, Otu0000262, Otu0000401, Otu0000447, Otu0000466, Otu0001387, Otu0001521). The most abundant Otu0000002 made up 80.28 - 91.13 % [DNA] or 78.82 – 89.06 % [RNA] of the *Sphingomonas* profile. Nevertheless, significant effects on relative abundance between *RPP1* genotypes were only observed by considering the majority of *Sphingomonas* OTUs. Further, Flavobacteria was found more than two times more abundant in Różanki (site 500, Wilcoxon-Test, p=0.027 [DNA]) *RPP1*- genotypes (2.14% [DNA], 0.84 % [RNA]) compared to *RPP1*+ (5.01 % [DNA], 2.17 % [RNA]). In contrast, an opposite trend was observed in Łupowo, showing higher abundances of Flavobacteria in *RPP1*- genotypes (5.52 % [DNA], 1.55 % [RNA]) compared to $RPP1^{+}$ (2.72 % [DNA], 0.66 % [RNA]). This data suggests an indirect impact of *Arabidopsis RPP1*+ haplotype from Gorzów

Wielkopolski and Łupowo on the abundance of *Sphingomonas*, as well as Flavobacteria in Różanki. *RPP1*- genotypes (2.14% [DNA], 0.84 % [RNA]) compared to *RPP1*+ (5.01 % [DNA], 2.17 % [RNA]). In contrast, an opposite trend was observed in Łupowo, showing higher abundances of Flavobacteria in *RPP1*- genotypes (5.52 % [DNA], 1.55 % [RNA]) compared to $RPP1^{+}$ (2.72 % [DNA], 0.66 % [RNA]). This data suggests an indirect impact of *Arabidopsis RPP1*+ haplotype from Gorzów Wielkopolski and Łupowo on the abundance of *Sphingomonas*, as well as *Flavobacteria* in Różanki.

Discussion



3 Discussion

Plants are concurrently facing a multilayer of ecological and biological stress factors along their developmental stages in nature (Niinemets 2010; Chaudhry et al. 2020). Crosstalk between plants and their respective environment leads to an evolutionary-driven genetic diversity, and adaptive traits (Gentzbittel et al. 2015; De Kort et al. 2021). At the same time, plants harbour complex microbial communities (collectively described as microbiota) that can affect growth, performance and health of their respective host (Babalola et al. 2020; C. Song et al. 2020). The dissection of plant-microbiome interactions are conceptually described as the "holobiont" (Cavalier-Smith 1992; Simon et al. 2019). While numerous studies have been published to understand binary plant-pathogen interactions, establishing a broad context of host factors shaping microbiota remains a challenge. Plants have evolved a complex innate immune system to cope with myriad pathogen pressures. Plant-pathogenic microbes have been identified as important "hub microbes" showing high degrees in complex microbial networks (Agler et al. 2016; Gao et al. 2021). The indirect relation of pathogen-recognizing immune receptors and the impact of microbial consortia might be a target to identify host-microbiome interaction (S. Yang et al. 2010). Here, I performed whole genome sequencing on 16 individual *Arabidopsis* plants from local populations around Gorzów Wielkopolski and two natural Landsberg accessions (La-0) to address intra-population variation of the RPP1-like cluster. In addition, microbial profiling was conducted to identify active and dormant phyllosphere microbiota considering *RPP1*-like copy number variations among a sample set collected over three consecutive years.

3.1 Whole genome sequencing of natural Gorzów ecotypes unveils intra-population variation

During the last decade, species-wide inventories of *Arabidopsis* helped to define genetic and phenotypic diversity among natural accessions across various regions of the world (Horton et al. 2012; Weigel and Mott 2009). Several studies suggest a global-scale population structure and genetic isolation by distance of *Arabidopsis* accessions (Sharbel, Haubold, and Mitchell-Olds 2000; Nordborg et al. 2005; Ostrowski et al. 2006; Beck, Schmuths, and Schaal 2008). Within the C-S-R (competitive C, stress tolerant S, ruderal R) theory, stress-tolerant plants evolve under high stress and low disturbance, whereas ruderal plants are affiliated to low stress and high disturbance (Grime 1977). In this respect, fundamental intra-species variation of *Arabidopsis* has been allocated to

climate adaptations (positive correlation to stress tolerance), such as temperature affecting flowering time in an annual cycle and leaf dry weight (Grime 1977; May, Warner, and Wingler 2017). In addition to stress reactions, local adaptation and co-variation have been shown to affect plant life history and growth rates, as well as metabolomes (Takou et al. 2019; Katz et al. 2021). In contrast to global-scale genetic pattern, population variation contributes to local adaptation (Montesinos et al. 2009). In this respect, most research on intra-species variation has been SNP-based and lacking structural genomic resolution (Gomaa et al. 2011; Alcázar et al. 2014; May, Warner, and Wingler 2017; Takou et al. 2019; Exposito-Alonso et al. 2019; Sharbel, Haubold, and Mitchell-Olds 2000; Horton et al. 2012; 1001 Genomes Consortium 2016; De Coster, Weissensteiner, and Sedlazeck 2021). Here, I present a comprehensive local-scale *Arabidopsis* genome dataset of Gorzów Wielkopolski (sub-) populations collected over three consecutive years. Local *Arabidopsis* populations of Gorzów Wielkopolski represent likely the ancestral population of *Landsberg erecta* (Ler-0), which is ranked as the second most widely-used accession in the world. I have performed whole genome sequencing on 16 individual *Arabidopsis* plants. Further, natural La-0 accessions were included to compare genetic backgrounds of Gorzów *Arabidopsis* with independent lines collected in 1992 (ABRC: CS1298) and 2010 (ABRC: CS76538). The latter natural Landsberg accession was part of the 1001 Genomes Project of *Arabidopsis* (1001 Genomes Consortium 2016). The whole genome sequencing of single plant *Arabidopsis* genomes remained precarious, due to restricted quantities of plant material. However, I generated seven high quality genomes from Gorzów *Arabidopsis* on pseudo-chromosome level. In addition, nine (here described as low-quality) genomes were generated from Gorzów *Arabidopsis* individuals. Challenges in *de novo* assemblies, due to regional sequence collapse or expansion, might explain obtained variations in total genome lengths (Asalone et al. 2020). Since second filial (F2) generations were used for individual genome sequencing, haplotype - heterozygosity might be attributable to genome length variation in diploid genome assemblies (Kronenberg et al. 2018). Thus, genome assembly contiguity could be proven by implementing linkage information, such as with physical and genetic maps (Mozo et al. 1999; Giraut et al. 2011; Jiao and Schneeberger 2020).

Gorzów *Arabidopsis* genome annotations, based on *ab initio* gene models revealed a nearly complete overlap (min. 99.4%) of protein-coding genes from the golden standard Columbia. However, ambiguous protein-coding genes (up to 26.2 %) were predicted in several gene model that require manual revision of gene models. It might be also true that genome assembly continuity displayed in high contig numbers might directly affect annotation qualities (Florea et al. 2011). The alignment of RNA-sequencing data to genomic sequences could clearly improve genome

annotations, even if it might be not truly comprehensive due to low level or tissue-specific gene expressions (Salzberg 2019). Nevertheless, high quality genomes displayed a significant consensus of benchmarking universal single-copy orthologs (BUSCO) comparable to the Landsberg reference (Jiao and Schneeberger 2020).

My data suggest a prodigious collinearity between Gorzów *Arabidopsis* genomes and natural accessions La-0 against the reference genome Ler-0. High collinearity might be explainable by the hypothesis of natural re-colonization pattern of *Arabidopsis* after the glaciation epochs (François et al. 2008; 1001 Genomes Consortium 2016). While higher genetic diversities were observed in southern European populations, genetic diversities declined in eastern populations, including Ler-0 (François et al. 2008). Changing climates, such as higher frequencies in droughts and rising temperatures might also increase natural selection towards northern Europe in the future (Exposito-Alonso et al. 2019). However, various proportions of genome-specific sequences were identified among local Gorzów *Arabidopsis*. Genome diversification in Gorzów populations varied from 3.1 % to 31.0 %, which is in average comparable to (Stenøien et al. 2004) who identified an overall genome variability of 12 % within *Arabidopsis* populations in Norway. The genetic variation in a local population permits flexibility and survival during variable environmental circumstances. (Castilla et al. 2020) displayed isolation-by-distance (IBD ~ dispersal limitation, genetic drift), and to a lower degree isolation-by-environment (IBE ~ local adaptation), accounting mainly for genetic differentiation of local *Arabidopsis* populations from the Iberian Peninsula (Castilla et al. 2020).

In line with (Zapata et al. 2016), most of the non-syntenic sequences (2.2 – 10.4 Mb) originate from duplications (1.1 -8.0 Mb). Duplication events occur highly frequently in plant genomes contributing to adaptation and evolutionary novelties (Van de Peer, Maere, and Meyer 2009; Conant and Wolfe 2008; Panchy, Lehti-Shiu, and Shiu 2016). Gene duplicates, also known as paralogs, correspond to their homologous relationship and appear by numerous mechanism, such as whole genome duplication (WGD), tandem, proximal, transposed and dispersed duplications (Yupeng Wang et al. 2011; Mascagni et al. 2021; Kono et al. 2018; Lallemand et al. 2020). Interestingly, duplication modes among dicots, including *Arabidopsis,* suggest tandem and proximal duplicates are occurring with higher frequencies, while WGDs underlie purifying selection (Mascagni et al. 2021; Xie et al. 2019). Notably, many tandem duplicate gene cluster have been identified in various plants showing intra-species variation among accessions (Kono et al. 2018). In this respect, Ler-0 represents various hotspots of rearrangements that contain tandem

duplications in defence-related genes that evolve race-specific resistances, such as the *RPP1*-like locus on chromosome 3 (Zapata et al. 2016).

Large-scale rearrangements in non-syntenic sequences, such as translocation (0.2 -1.9 Mb) and inversion (0.3 -1.5 Mb) making up 0.5 to 3.4 Mb in Gorzów *Arabidopsis* genomes, which is comparable to ~3.6 Mb in Ler-0 against Col-0 (Zapata et al. 2016). Respectively, obtained intra-population variation of rearrangements were lower compared to intra-species variation based on chromosome-level assemblies of seven *Arabidopsis* accessions, representing species-wide diversity (Jiao and Schneeberger 2020).

Taken together, a comprehensive Gorzów *Arabidopsis* genome dataset on a (sub-) population scale revealed a prodigious collinearity between local individuals and natural accessions against the Landsberg reference (Ler-0). Nevertheless, genome-specific variations in Gorzów *Arabidopsis* may provide the cause of the genetic diversification in local populations. Further investigations are required to determine population variability on the base of proportion of polymorphic loci (P_P), gene diversities (H_E) and heterozygosity (Stenøien et al. 2004). Since I obtained multiple genomes from consecutive sampling years, heritable variation can be estimated. This whole genome dataset guides us towards long-read-based structural variation analysis on a population scale.

3.2 Heterogeneous *RPP1*-like gene cluster structure in local Gorzów *Arabidopsis*

The architecture of the NLR gene cluster has been identified in multiple distinct accessions, leading to the assumption of complex evolutionary tasks that generate distinct haplotypes (Van de Weyer et al. 2019; R. R. Q. Lee and Chae 2020). Notably, the NLR gene cluster have been identified as genetic rearrangement hotspots, showing high frequencies of gene duplication to quickly response to biotic stress (Jiao and Schneeberger 2020). In this study, I examined the *RPP1*-like cluster structure, also referred to the *DANGEROUS MIX* 2 (DM2) locus, among WGS of individual *Arabidopsis thaliana* plants from Gorzów Wielkopolski (former Landsberg an der Warthe, Poland) to highlight structural variation on a local scale. The *RPP1*-like cluster is remarkably variable in terms of cluster size and gene copy numbers even within natural population. Thus, the *RPP1*-like cluster loci varied highly in size between 60 Kb and 159 Kb in our sampled populations. Notably, the genomic region of Col-0, flanked by AT3G44600 and AT3G44690, displayed a similar sequence length to *RPP1*- haplotypes. In contrast to Col-0, the laboratory accession Landsberg (Ler-0) contains a ~ 100 Kb insertion within the region on chromosome 3, accounting for the *RPP1*-like locus (total length of 158 Kb). Intra-species variation has been previously described for

RPP1-like cluster in various *Arabidopsis* accessions (Goritschnig et al. 2016; Chae et al. 2014; R. R. Q. Lee and Chae 2020; Van de Weyer et al. 2019). In this study, I provide evidence for high heterogeneity within the *RPP1*-like loci on a local scale within natural *Arabidopsis* populations of Gorzów Wielkopolski based on whole genome sequences. Based on *ab initio* gene model predictions, I identified immunity-related and -unrelated genes within the *RPP1*-like cluster loci that varied pretty much between individual plants of local subpopulations. Hence, high sequence divergence was observed favoured between sub-populations, in terms of *RPP1*-like cluster organisation, number and size of immunity-related and unrelated genes, as well as general R-gene cluster length. However, in order to precisely capture all paralogs, further investigations are required to annotate single *RPP1*-like genes (R1 – R8).

A recent study by (R. R. Q. Lee and Chae 2020), using a species-wide inventory of NLR cluster on 64 *Arabidopsis* accessions, suggested that larger NLR cluster, such as DM2/*RPP1*, DM8/*RPP8* and DM8/*RPP4*/*RPP5* generally unveil the highest heterogeneity varying in cluster size and copy numbers (R. R. Q. Lee and Chae 2020). In addition, the DM2/*RPP1* cluster occurs in a presence/absence scenario in some of the *Arabidopsis* accessions (R. R. Q. Lee and Chae 2020). Although those findings were observed on an intra-species survey, my data suggest similar trends on a local population scale.

Population studies of natural *Arabidopsis* individuals from Tübingen (Germany) unveiled high homozygosity within a set of 436 SNP markers, which is mostly attributable for nearest-neighbour plants (Bomblies et al. 2010). Similarly, I observed a major collinearity between individual Gorzów genomes against the reference Ler-0. However, structural variation of the *RPP1*-like loci suggested high heterozygosity within NLR cluster, which were described as rearrangement hotspots, in natural population or even between neighbouring plants (Jiao and Schneeberger 2020).

Genetic distances of *RPP1*-like loci favored the formation of two major cluster, here named as *RPP1*+ and *RPP1*-. Notably, the reference sequence of Landsberg (Ler-0) was affiliated to *RPP1*+, while Columbia unveiled a high centrality in *RPP1*-. Considering the recorded origin of Col-0, which was collected by Friedrich Laibach in Landsberg an der Warthe (Poland), it might be assumable that the natural *Arabidopsis* population of Gorzów Wielkopolski still harbors haplotypes reflecting the genetic background of the current laboratory accessions Columbia and Landsberg in our days (Somssich 2019).

In addition to the whole genome sequencing of Gorzów *Arabidopsis* individuals, I performed a population-wide *RPP1*-like genotyping of collected phyllosphere microbiome samples to correlate

RPP1-like copy number variations to microbial community compositions. Interestingly, genotyping of 332 individual plants from six sampling sites, representing three geographical locations (Gorzów, Różanki, Łupowo), unveiled a binominal distribution of *RPP1*-like CNVs in natural *Arabidopsis* populations around Gorzów Wielkopolski. Similar to structural variation of the *RPP1*-like cluster in genomes, two major *RPP1*-like CNV clusters were correlated to Columbia (*RPP1*-, ~ 8 copies) and Landsberg (*RPP1*+, ~ 16 copies). Since a binominal distribution of *RPP1*-like CNV's was detected, there was no clear cut between *RPP1*+ and *RPP1*- haplotypes, suggesting a segregation of the *RPP1*-like loci in Gorzów Wielkopolski.

Notably, a previous study, based on SNP-genotyping already unveiled copy number variations within Gorzów *Arabidopsis* (Alcázar et al. 2014; Atanasov et al. 2014). Sampling sites in central Gorzów Wielkopolski had been described as *RPP1*-like hotspots (Alcázar et al. 2014). In line with these previously reported findings, three (200, 300, 400) out of four central Gorzów sites could be confirmed as *RPP1*-like hotspots in my screening of 332 individual plants. Genotyping of a wild *Arabidopsis* population in Różanki (north-east of Gorzów Wielkopolski) displayed a various mix of *RPP1*-like CNV's, while Łupowo (south-west of Gorzów Wielkopolski) favoured *RPP1*- haplotypes. In addition, high fluctuations of *RPP1*-like CNV's were observed between sampling years, even within sampling sites.

Taken together, whole genome sequencing and ddPCR genotyping unveiled a complex and highly diverged *RPP1*-like cluster structure in natural Gorzów *Arabidopsis* that suggest high heterozygosity attributable for cluster size and copy numbers in individual plants among sampling sites and years. *Arabidopsis* sampling of central Gorzów Wielkopolski were confirmed as *RPP1*-like hotspots in three out of four sites. This study supports that the NLR cluster, such as the *RPP1*-like locus on chromosome 3 in Gorzów / Landsberg accessions of *Arabidopsis* represents highly divergent cluster formations even on local population (max. plant distance: 20 km) or sub-population levels (plant distance: 3 m radius). An in depth analysis using multi-sequence alignments and domain structure predictions found that *RPP1*-like genes are necessary to substantiate structural variation of characteristic functional regions within the NLR cluster (Proell et al. 2008). Initial results of an evolutionary sequence analysis calculating dN/dS ratios [dN: number of nonsynonymous substitutions per non-synonymous site, dS: number of synonymous substitutions per synonymous site] (Runge et al. unpublished data provided by Daniel Gómez-Pérez) unveiled that only a few *RPP1*-like genes are under positive selection (Mondragón-Palomino et al. 2017, 2002; Howe et al. 2016; Stam, Silva-Arias, and Tellier 2019; Koenig et al. 2019). These data suggest that the *RPP1*-

like cluster is actively maintained in *Arabidopsis* populations by positive or balancing selection, while genetic drift alone is much more unlikely. In addition, our whole genome dataset could be used to obtain structural insights of additional NLR cluster in natural Gorzów *Arabidopsis* population.

3.3 Phyllosphere microbiota of natural *Arabidopsis* populations around Gorzów Wielkopolski altered by environmental cues

Microbial consortia on the phyllosphere are facing a challenging habitat, which undergoes drastically fluctuations in environmental conditions, poorly available nutrient repertoires and inter-kingdom competitions related to plant pathogenic microbes (Perreault and Laforest-Lapointe 2021; Y. T. Cheng, Zhang, and He 2019). Previous studies on phyllosphere microbiota in natural *Arabidopsis* populations have unveiled general colonization pattern (Remus-Emsermann et al. 2014; Bodenhausen, Horton, and Bergelson 2013; Agler et al. 2016). However, characterizations of the phyllosphere microbiomes of the most common *Arabidopsis* accessions have yet to be established. In this study, I present the first observation of phyllosphere microbiota in natural *Arabidopsis* from Gorzów Wielkopolski, which still harbors Landsberg-like genotypes (Alcázar et al. 2014) and might represent the ancestral population of the second-most widely studied *Arabidopsis* accession.

To assign which factors drive microbial assemblies on the *Arabidopsis* phyllosphere on a population-scale, I analyzed amplicon-based sequencing data (targeting bacteria and eukaryotes), simultaneously obtained from DNA and cDNA-sequencing of individual plants taken over a time period of three years. Here, the factor 'leaf compartment' explained a significant part of the species richness and beta diversity consistent in DNA- and cDNA-sequencing. A recent study, unveiled higher species richness of epiphytes in a culturable approach (Bruisson et al. 2019), while lower diversity in endophytes can be linked to spatial niche colonization (Bulgarelli et al. 2013), which might require tight host adaptation. In addition, plant phyllosphere compartmentation of grapevine and lotus (*L. corniculatus*) plants exhibited diverging patterns for root, shoot, leaf, flower and seed microbial communities (Zarraonaindia *et al.*, 2015, Lutap et al. (unpublished)). Similar observations were made for compartmentation of the rhizosphere (soil, root) (Hou et al. 2021).

In addition, the year and geographical location of sampled *Arabidopsis* populations significantly impacted the microbial community assemblage. This had been already suggested across agro-climatic zones in various host species, such as grapevine and tomato (Singh et al. 2018; Runge et

al. 2021; Agler et al. 2016). The effect of sampling sites in the geographical area of Gorzów Wielkopolski on the phyllosphere microbiome may be related to microbial soil inoculum available for colonization that vary in space and time among sampling sites (Zarraonaindia et al. 2015; Shade et al. 2014; Goss-Souza et al. 2017; Brown et al. 2020). In addition, local colonization patterns are affected by different microclimates (Noh et al. 2020; Al Ashhab et al. 2021; Lindström and Langenheder 2012), as well as by natural host genotype variation (Runge et al. 2021; Bálint et al. 2013; Singh et al. 2018). Recent studies also highlight the impact of host developmental stages on the microbial community progression over the lifecycle (Edwards et al. 2018; Almario et al. 2021; Berens et al. 2019; Wagner et al. 2016), which might partially account for the observed year-effects in this dataset. Overall, this highlights how phyllosphere microbiota are shaped upon constantly fluctuating environmental conditions in space and time over the growth period.

The taxonomic composition of phyllospheric microbiota of Gorzów *Arabidopsis* displayed common bacterial phyla, such as Proteobacteria, Actinobacteria, Firmicutes and Bacteroides. While bacterial taxa in epiphytic compartments were dominated by Proteobacteria and Actinobacteria, endophytic compartments were predominantly composited of Proteobacteria. In contrast to endophytes present in populations from sampling sites in Gorzów, endophytes in Ro 500 and Lu 600 showed higher abundances of Actinobacteria accompanied by a reduced abundance of Proteobacteria. Interestingly, endophytic Actinobacteria play an important role in plant growth promotion showing an enormous capacity of bioactive compounds, such as antimicrobials and antifungal compounds (Passari et al. 2017; P. Chen et al. 2019; S. Qin et al. 2011).

Overall, eukaryotes have been shown to display a lower species richness compared to bacteria (Stone, Weingarten, and Jackson 2018; Lindow and Brandl 2003). In line with these observations, DNA- and cDNA - sequencing revealed similar richness trends across sampling sites. Observation of species richness on eukaryotic epiphytes unveiled high similarities between geographical locations. Thus, Gorzów *Arabidopsis* was dominated by eukaryotic phyla, such as Opisthokonta, Archaeplastida, Stramenopiles and Rhizaria. Interestingly, the phylum Opisthokonta was dominating epiphytic and endophytic compartments in cDNA-seq, while many OTU's remained as unclassified Eukaryota in epiphytic compartments using classical DNA-seq. As such, eukaryotic species richness seems to be higher in cDNA-seq in contrast to DNA-seq. In contrast to epiphytes, eukaryotic endophytes displayed higher species richness in natural *Arabidopsis* populations of Różanki and Łupowo. In addition, those outlier sites (Różanki, Łupowo) showed higher abundances of Stramenopiles.

Taken together, while species richness of epiphytes was rather comparable between *Arabidopsis* populations, endophytic microbes displayed higher species richness in Różanki and Łupowo. Further investigations on soil properties and microbial assessments of soil inoculate might be relevant to indicate whether high species richness occurs due to environmental cues or host-species dependent on individual *Arabidopsis* populations.

3.4 Common and atypical persistent core microbiome

Various studies have attempted to identify taxa within plant microbiome that are persistent on single compartments, nutrient levels or abiotic stresses on population-scale or among populations of various hosts (Toju et al. 2018; Runge et al. 2021; Noble et al. 2020). Since, environmental cues are major factors impacting microbial communities, core members were stringently determined across the whole dataset to perceive persistent taxa (Müller et al. 2016). In this study, the microbial core community consists of 3 - 17 (DNA, cDNA) bacteria and 2 eukaryotes, which accounts for 34.6 % of bacterial to 55.1 % of eukaryotic relative abundance of total reads including high abundant OTUs. Here, the most abundant bacterial core members were taxonomically assigned as *Sphingomonas faeni* (OTU2), *Methylobacterium adhaesivum* (OTU5, 14), *Variovorax* spp. (OTU8) and *Pseudomonas* spp. (OTU3, 9). Notably, bacterial genera, like *Sphingomonas*, *Methylobacterium* and *Pseudomonas* have been identified as microbial taxa universally present on various host species (Rastogi et al. 2012; Guittar and Shade 2019; Runge et al. 2021; X. Chen et al. 2021; Agler et al. 2016; Vorholt 2012). In natural *Arabidopsis thaliana*, *Sphingomonas* and *Pseudomonas* have been identified as major taxa on the phyllosphere (Lundberg et al. 2021). Pan-genome analysis of the bacterial genus *Sphingomonas* has unveiled high genomic diversity (Y.-J. Kim et al. 2020). Genetic features of endophytic *Sphingomonas* showed plant-protective properties, as well as frequently presence of protein secretion systems hints to host-microbe interaction (Lundberg et al. 2021). Plant protection has been described for the endophytic *S. melonis* shaping disease resistance against *Pseudomonas syringae* DC3000 in *Arabidopsis* and *Burkholderia plantarii* in rice (Matsumoto et al. 2021; Innerebner, Knief, and Vorholt 2011). Here, the most abundant bacterium *Sphingomonas faeni* has been previously described as an air- and dust-borne bacterium (Busse et al. 2003). Further, *S. echinoides* has been shown to have a root growth stimulating ability via phytohormone production (Khan et al. 2014). A recent study has indicated that single strains of *Sphingomonas* and *Microbacterium* showed the highest potential to affect microbial community structure (Carlström et al., 2019). Facultative methylotrophic bacterial core members were closely assigned to *Methylobacterium adhaesivum*, which was identified as a novel species in a drinking water

distribution system (Gallego, García, and Ventosa 2006). Similar to this study, *Methylobacterium adhaesivum* has been shown to be a core member of epiphytic phyllosphere microbiota of wild tomato species in central Peru (Runge et al. 2021). Further, *M. adhaesivum* has been described as phyllosphere colonizer showing plant protective properties (Wellner, Lodders, and Kämpfer 2011; Peredo and Simmons 2018; Sy et al. 2005; Yoshida et al. 2017; Sanjenbam et al. 2020). In particular, growth promotion provided by *Methylobacterium* was measurable on biomass and fruit yield in tomato (Innerebner et al., 2011; Senthilkumar & Krishnamoorthy, 2017). Interestingly, *Methylobacterium adhaesivum* and *Rathayibacter* spp. showed higher abundances in dysbiotic tomato samples (Runge et al. 2021). In addition, *Rathayibacter* and *Flavobacterium* are commonly present on the phyllosphere of *Arabidopsis* and other host species (Dorofeeva et al. 2002; Bodenhausen, Horton, and Bergelson 2013; Sivakumar et al. 2020). While member of *Flavobacterium* have been associated with the ability to degrade complex organic compounds (Kolton et al. 2016), their functional role in microbial community assemblies remain unknown.

Furthermore, *Variovorax* was also found as a microbial core member. *Variovorax* has been identified as a key player of the root microbiome, affecting auxin degradation processes in *Arabidopsis* and tomato (Décamps & Lebreton, 2011). On the phyllosphere, *Variovorax* is involved in degradation of isoprene carbon sources, which might be produced by the plant under stress conditions (Crombie et al., 2018; Fini et al., 2017; Jardine et al., 2020).

The bacterial genus *Pseudomonas* has been widely described as a colonizer of various ecological niches, including the phyllosphere microbiome (Dong et al., 2019) and comprises common plant pathogens, as well as non-pathogenic strains (Desrut et al. 2020; Vacheron et al. 2018; Chahtane et al. 2018; Chu et al. 2019; Karasov et al. 2019). *Pseudomonas* species are strong competitors in the environment, such as *P. viridiflava* found to be a dominant pathogen on natural *Arabidopsis* and present on other host species (Goss, Kreitman, and Bergelson 2005; Karasov et al. 2018). Here, our taxonomic resolution uncovered common core OTU's belonging to the Pseudomonas genus using amplicon sequencing. While *Pseudomonas* spp. were among the highest abundant OTUs, their function on the phyllosphere microbiome remains elusive, since strain resolution could not be achieved for most OTU's.

Persistent microbial taxa that frequently occur may have a functional or adaptive role in their microbiome. Nevertheless, rare taxa have been shown to be as important in microbe-microbe interaction to remain host health (Velazquez et al. 2019). In this study, rare core taxa have been taxonomically assigned to *Hymenobacter* spp. (OTU20), *Rathayibacter caricis* (OTU25),

Salinibacterium (OTU49), *Devosia* spp. (OTU17) and *Bosea* spp.(OTU46). A previous study on various plant compartments of wild strawberry microbiomes unveiled significantly enriched *Salinibacterium* and *Hymenobacter* on flowers (N. Wei and Ashman 2018). *Devosia* spp. were found as endophytes in potato roots (Manter et al. 2010) and *Bosea* spp. have been identified as biodegrades (Shin et al. 2012).

Eukaryotic core microbes, comprising Arthropoda and green algae were represented by 18S rRNA amplicons. Abundant core microbes, represented by DNA and cDNA amplicons of the small component of eukaryotic ribosomes were *Hypera postica* (OTU2) and *Chloroidium saccharophilum (*OTU65). The alfalfa weevil *H. postica* is a species of beetle and an invasive legume pest, feeding on *Medicago*, *Vicia*, *Trifolium* and *Astragalus* species (Fabaceae) (Tuda et al. 2021; Iwase et al. 2015). The geographical distribution of *H. postica* on its native hosts has been described for western Europe, including southern Poland. Herein, *H. postica* was firstly described to occur highly abundant (>3 mio. total reads, OUT2, OTU565, OTU612) on natural *Arabidopsis* population (Tuda et al. 2021). Notably, we never observed any adults or larvae on leaf rosettes during sample collections. While taxonomy assignment accurately matched *H. postica* (OTU2), misassignments on partial 18S rRNA gene sequences or database biases cannot be completely ruled out. Interestingly, other weevils of the same family Curculionidae, such as *Ceutorhynchus atomus* and *C. contractus* have been described as major insect herbivores on natural *Arabidopsis* plants (Arany, de Jong, and van der Meijden 2005). However, further investigations are required to unveil the occurrence of *H postica* on the non-host plants *Arabidopsis thaliana*. Nevertheless, multiple studies have shown that insect eggs are able to activate plant immune responses, such as systemic acquired resistance, in order to defend themselves against other pathogens, like *Pseudomonas syringae*, *Botrytis cinerea* BMM and *Hyaloperonospora arabidopsidis* Noco2 (Hilfiker et al. 2014; Alfonso et al. 2021). Whether the highly abundant *H. postica* on *Arabidopsis* leads to SAR has to be examined. Further, a photopic green algae *Chloroidium saccharophila* was detected as a microbial core member on Gorzów *Arabidopsis* population. Notably, *Chloroidium* spp are ubiquitous in various ecosystems (Darienko et al. 2010; Stewart et al. 2021; Nelson et al. 2017) and have been detected on the phyllosphere of trees (H. Zhu et al. 2018; Ching-Su, Yu-Hsin, and Jiunn-Tzong 2012). A recent study on rhizosphere protists using network analysis unveiled a positive correlation between the presence of *C. saccharophila* and plant health in tomatoes (W. Xiong et al. 2020). The identification of persistent plant-associated microbes highlights current knowledge for the majority of the achieved taxa, while multiple microbes have been barely or never described on *Arabidopsis thaliana*.

3.5 *RPP1*-like copy number variation alters bacterial diversity in nature

The *RPP1*-like locus displays a highly polymorphic NLR cluster found in a handful of *Arabidopsis thaliana* accessions (Alcázar et al. 2009; R. R. Q. Lee and Chae 2020; Chae et al. 2014). Most studies have described structural organizations of NLR clusters on the intra-species level. Notably, previous studies on *Arabidopsis* and tomato unveiled high heterogeneity within the NLR repertoire, even on a population scale (Stam, Scheikl, and Tellier 2016; Alcázar et al. 2014). For the *RPP1*-like locus, (Alcázar et al. 2014) unveiled that natural *Arabidopsis thaliana* from Gorzów Wielkopolski harbor predominantly two *RPP1*-like haplotypes within a population (Alcázar et al. 2014). In detail, roughly 30 % of *Arabidopsis* individuals from Gorzów displayed similarities to the Ler *RPP1*-like locus, while the majority of tested individuals (~ 70%) showed no amplification of Ler *RPP1*-like genes or non-Ler *RPP1*-like genes, respectively (Alcázar et al. 2014). These data suggest predominantly two major Gw *RPP1*-like haplotypes, such as Gw+ (~ *RPP1*+) and Gw- (~*RPP1*-), which were confirmed in my *RPP1*-like genotyping for central Gorzów *Arabidopsis* population. I observed notable differences in *RPP1*-like CNV's in natural Gorzów populations. However, the presence / absence polymorphism within the *RPP1*-like cluster enabled us to study the impact of on phyllosphere microbiota in nature. While polymorphisms in NLR clusters might occur more frequently on a local-scale, this is the first study that has actually the potential to correlate microbial community assemblies in respect of NLR cluster variation in natural *Arabidopsis* populations. Here, I linked *RPP1*-like polymorphism, grouped in *RPP1*+ (high copy number) and *RPP1*- (low copy number), to microbial community compositions of the phyllosphere of natural *Arabidopsis* populations over three consecutive years. Interestingly, bacterial species richness was lower in *RPP1*+ haplotypes, consistent in epiphytic and endophytic compartments of cDNA-sequencing. Similar observations were made in DNA-seq, yet, after testing the were not statistically significant. Further observation on bacterial richness were made on sample groups, considering Nucleic Acid x Compartment x Geographical Location x *RPP1*, leading to a reduced species richness in bacterial endophytes of Gorzów Wielkopolski (DNA, cDNA) considering *RPP1*+ haplotypes. In addition, sample-to-sample variation of bacterial profiles decreased in *RPP1*+ haplotypes for leaf compartments in cDNA-seq. Those data show that bacterial richness and sample variation decreases in presence of the *RPP1*-like cluster in natural *Arabidopsis* population. Recent studies have suggested to use lower microbial diversities as a predictor of microbial dysbiosis in plant and animals (Karasov et al. 2019; Runge et al. 2021; Kriss et al. 2018; Brüssow 2020).

However, I was not able to distinguish healthy and dysbiotic samples in our dataset referring to phenotypical or microbial diversity measurements.

Further, *RPP1*+ haplotypes displayed an elevated abundance of the bacterial genus *Sphingomonas*. Therefore, I hypothesize that the presence of the *RPP1*-like cluster (*RPP1*+) suppresses the proliferation of certain bacterial species, which might indirectly result in a higher abundance of *Sphingomonas*. While NLR signaling affecting microbiome composition haven't been described in plants, animal NLR's, such as NOD1/2, NLRP12 and NLRP3 are activated by conserved MAMPs and have been suggested to affect host microbiota in the human gut (L. Chen et al. 2017; Y.-Y. Li et al. 2017). Although animal NLR's are known for mainly sensing conserved MAMPs and show functional relations to animal PRR's (Maekawa, Kufer, and Schulze-Lefert 2011). For example, NOD1/2 can sense bacterial cell-wall peptidoglycan (Strober et al. 2006; Wolf and Underhill 2018). In contrast, plant NLR's are known to recognize highly variable effector proteins in the cytoplasm (described as ETI), whereas plant PRR's are surface-localized receptor mediating PTI signaling, historically (Lu and Tsuda 2021a; J. D. Jones and Dangl 2006). While plant NLR's and PRR's showing distinct functional activation mechanism, current research suggest an intimate association of PTI-ETI signaling to potentiate immune responses against bacterial pathogens (Lu and Tsuda 2021b; Naveed et al. 2020; Yuan, Ngou, et al. 2021; Ngou et al. 2021; Yuan, Jiang, et al. 2021). As an example, ETI triggered by oestradiol-inducible treatments of AvrRpp4 (recognized by *RPP4*) lead to an accumulation of PTI-related proteins, such as BAK1, SOBIR1, BIK1, RBOHD and MPK3, but not for other PTI-related proteins, such as CERK1, FLS2. MPK4 and MPK6 (Ngou et al. 2021). Note, the plant NADPH oxidase RBOHD, responsible for ROS production during PTI and ETI, has been identified as a crucial host factor for homeostasis in phyllosphere microbiota (Pfeilmeier et al. 2021). Moreover, plant NLR's have been widely described in host resistance, while evidence for NLR non-host resistance (NHR) emerged from *Arabidopsis WRR4* and maize *Rxo1* (Redkar et al. 2021; Borhan et al. 2008; Cui, Tsuda, and Parker 2015; H.-A. Lee et al. 2017; Cevik et al. 2019). Thus, *Rxo1* shows resistance to pathogenic and non-host bacteria (B. Y. Zhao et al. 2004).

In contrast to bacteria richness, no differences were observed in eukaryotic profiles in respect of *RPP1*-like haplotypes. However, I observed a slight shift in eukaryotic richness in *RPP1*+ haplotypes using cDNA-seq suggesting a consistent advantage of cDNA-seq compared to DNA-seq to resolve the impact of host-genotype effect in natural populations.

Overall, these results highlight the fact that the host-genotype, considering *RPP1*-like haplotypes, affect the phyllosphere microbiome in natural *Arabidopsis* populations, besides other effects, such as leaf compartments and environmental cues. Nevertheless, the molecular function of *RPP1*-like genes in natural Gorzów *Arabidopsis*, as well as in natural and laboratory Landsberg accessions (Ler-0, La-0) still remains unknown. In the future, correlations between the microbiome composition and *RPP1*-like haplotypes will have the potential to identify key microbes that will allow us to validate our findings *in planta* under controlled conditions. Therefore, highly associated microbes will be tested on Landsberg (Ler-0, *RPP1*+) and Ler-0 ΔDM2 (*RPP1*-, deletion of the *RPP1*-like cluster in Ler-0 background).

3.6 Conclusions

The current study aimed to determine the *RPP1*-like cluster structure and whether the presence of the *RPP1*-like genes impacts the phyllosphere microbiota in local Gorzów *Arabidopsis* populations (see Fig. 14A). Here, I supplied a comprehensive dataset comprising whole genome sequences of 16 single *Arabidopsis thaliana* genomes from natural sampling sites in the area of Gorzów Wielkopolski (former Landsberg an der Warthe). Furthermore, two natural *Arabidopsis* accessions from Landsberg (La-0) displayed high collinear regions in relation to the reference Ler-0. In respect of the *RPP1*-like loci, high heterogeneity was observed in terms of sequencing length and gene copy variation (see Fig. 14D). *RPP1*-like genotyping revealed two major *RPP1*-like haplotypes, described as *RPP1*- and *RPP1*+ across natural *Arabidopsis* populations (see Fig. 14B). Hence, three out of four central Gorzów sites (200, 300, 400) were confirmed as *RPP1*-like hotspot, while outlier sites displayed a complex mixture of *RPP1*-like haplotypes. By studying the composition of phyllospheric microbes using simultaneously DNA- and cDNA - amplicon sequencing has unveiled environmental cues, such as sampling year and sampling sites, as well as leaf compartments affecting microbial community assemblies (see Fig. 14C). Natural variation of *RPP1*-like copy numbers allowed correlations between *RPP1*-like haplotypes and resulted in lower bacterial richness, as well as lower sample-to-sample variability in *RPP1*+. While environmental cues were leading factors in shaping phyllosphere microbiota, natural NLR polymorphism, displayed on *RPP1*-like haplotypes altered bacterial communities in wild Gorzów *Arabidopsis* on a local-scale. In addition, I determined persistent plant-associated microbes across the whole dataset and I resolved the presence of common taxa, as well as uncommon ones that had not yet been described on *Arabidopsis thaliana*. These results support the currently established knowledge about sampling timepoints and geographical locations impacting microbial consortia. In addition, I overserved a host species effect on the diversity of microbial community compositions, mainly affected in bacteria and protist profiles. In terms of diversity, fungal community compositions remained rather stable across data features, like sampling year, geographical locations and host species.

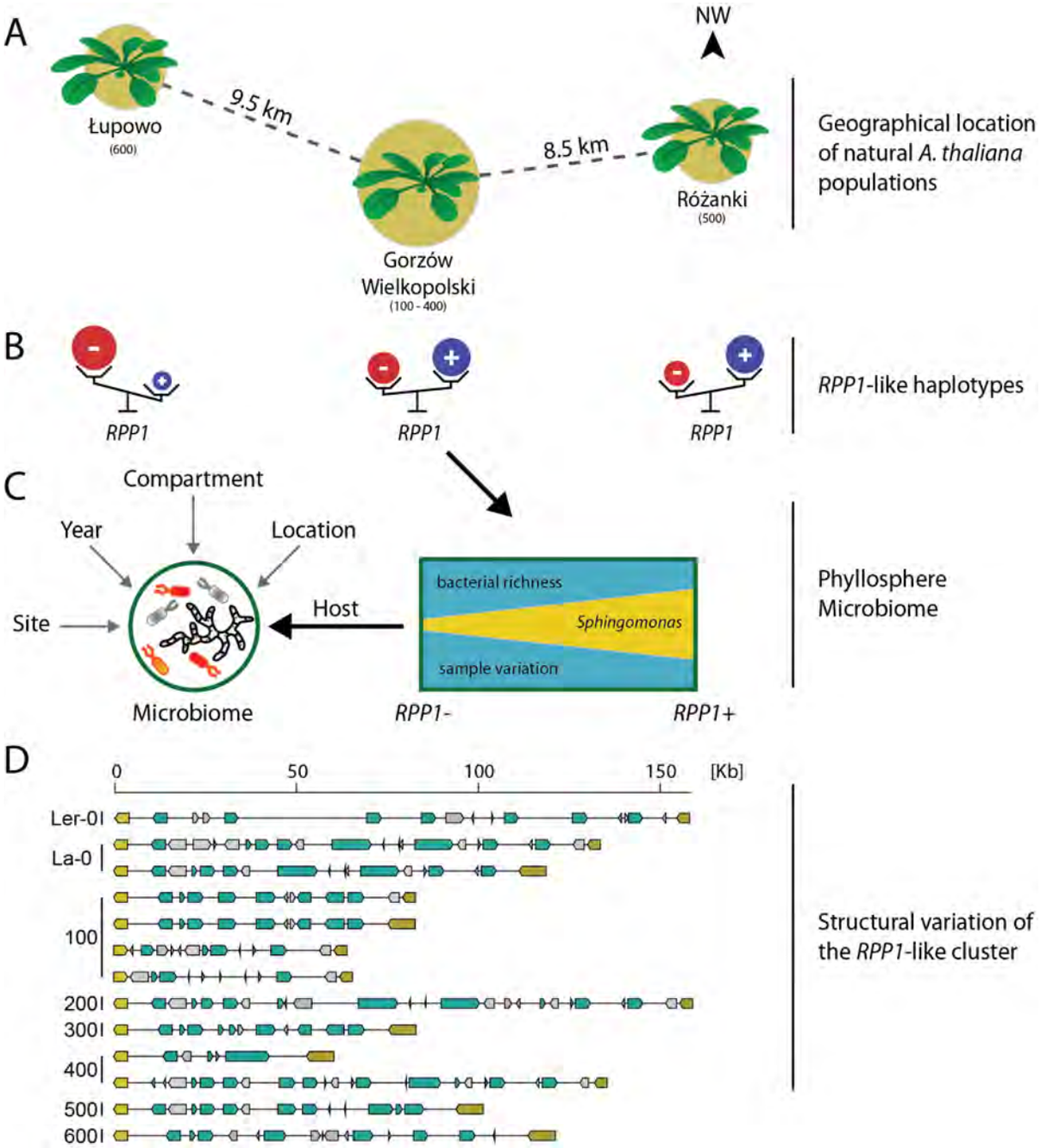


Fig. 14: Overview of major conclusions. A) Geographical location of natural *Arabidopsis* population around Gorzów Wielkopolski. B) Distribution of *RPP1*-like haplotypes simplified as *RPP1*- and *RPP1*+. C) Ecological factors impacting phyllosphere microbiota. D) Structural variation of *RPP1*-like cluster in Ler-0, La-0 and Gw individuals.

3.7 Future perspectives

Scientific results obtained in my thesis rises a plethora of new research questions and objectives. The whole genome sequencing of single *Arabidopsis* plants from Gorzów Wielkopolski revealed heterogenous *RPP1*-like gene cluster structures. The *RRP1*-like genotyping of individual plants gave a first indication of copy number variations among the whole samples set.

From this study, whole genome sequences of various individual *Arabidopsis* plants whole genome sequencing of Gorzów *Arabidopsis* plants call for further investigations about the *RPP1*-like gene cluster organization among *Arabidopsis* subpopulations of Gorzów Wielkopolski, Różanki and Łupowo. However, to study *RPP1*-like gene cluster structures and the impact of single *RPP1*-like genes on the phyllosphere microbiome, it would be useful to apply a metagenomic approach on the endophytic sample set (see Fig. 15A). On one hand, plant reads could be used for a mapping approach of the *RPP1*-like gene cluster against the reference Ler-0 and our Gorzów *Arabidopsis* genome collection to verify gene cluster organisation and single nucleotide polymorphism in non-synonymous amino acids on a population-scale. On the other hand, sequencing reads of endophytic microbes could be used to extent the microbiome study to an untargeted approach. To further understand the complex colonization and succession of the phyllosphere microbiome in respect to the plant, metagenomic data could be used for a microbial genome wide association study (microbial-GWAS), implementing further genomic features to identify host-genetic factors on whole genomes (Deng et al. 2021; Clouse and Wagner 2021; Brown et al. 2020; F. Li et al. 2019). Thus, evidence of positive selection pressure on whole genome information on a local-scale can be used to identify

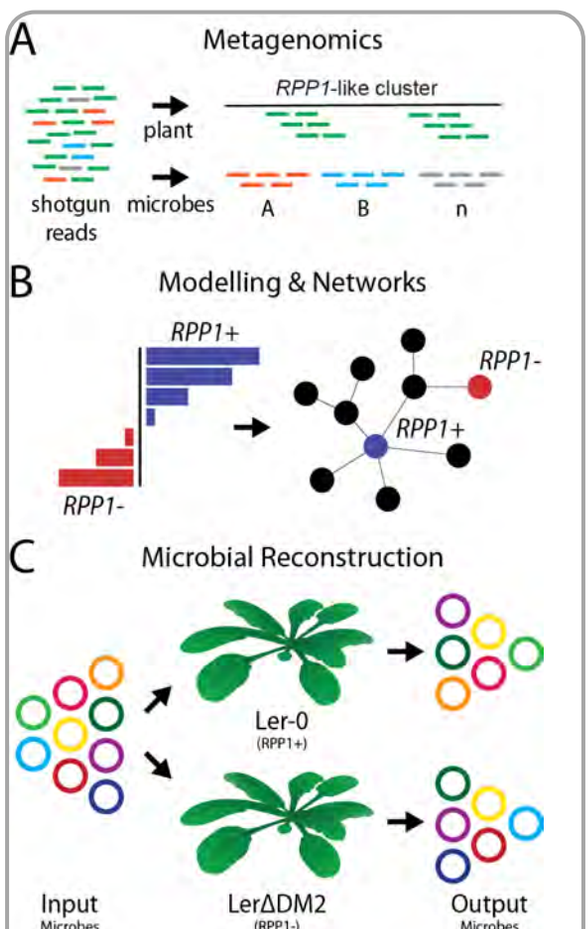
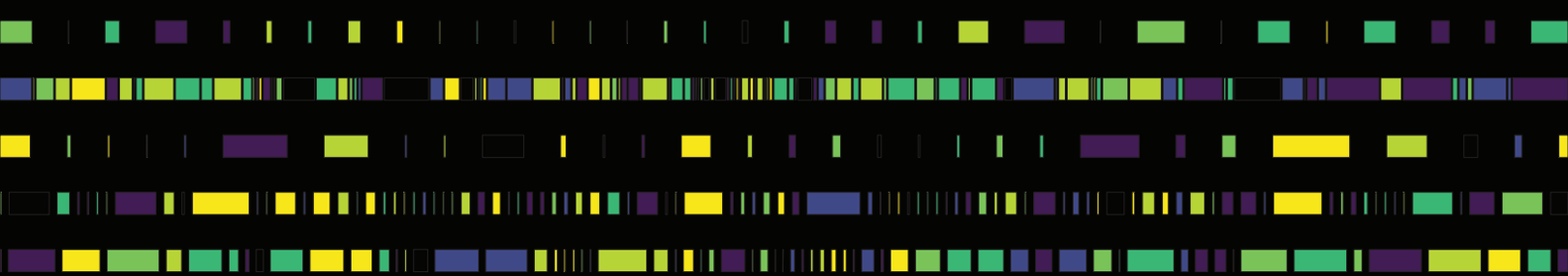


Fig. 15: Future perspectives to verify the impact of *RPP1*-like haplotypes on the phyllosphere microbiota. A) Metagenomics of endophytic compartments will be used to obtain structural variation of the *RPP1*-like cluster. B) Linear discriminant analysis and co-occurrence networks will be used to identify *RPP1*-like-associated microbes. C) Microbial reconstruction experiments using Ler-0 (*RPP1+*) and LerΔDM2 (*RPP1-*) are suggested to confirm obtained conclusions under controlled conditions.

candidate genes or families that have the potential to unveil host-microbe interactions to undergo functional relation of host genetics and microbial compositions (Mondragón-Palomino et al. 2002; Dyachkova, Chekalin, and Danilenko 2019).

In this project, I observed higher bacterial diversities in the absence of the *RPP1*-like gene cluster in Gorzów *Arabidopsis*. Hence, the cause might be directly- or indirectly - related to the presence of *RPP1*-like genes. Therefore, modelling approaches and co-occurrence network calculations can be used to identify *RPP1*-like-related microbes (see Fig. 15B). In planta experiments using a Ler-0 Δ*RPP1*-like mutant (Deletion of the *RPP1*-like gene cluster in Ler-0) or Col-0 *RPP1*-like (R1-R8) complementation lines could be useful to perform functional analysis of this NLR cluster under the use of highly diverse synthetic microbial communities, so called SynCom's (see Fig. 15C). SynCom's are ideal to narrow down functional relations of the *RPP1*-like haplotypes, since we established a profound microbial Gorzów Culture Collection (GCC), representing several hundreds of bacteria, yeasts and fungi. In addition, co-occurrence network analysis and supervised linear learning models, based on support-vector machine algorithm are considered for the future to limit experimental efforts. I hypothesize that potential resistance genes, such as *RPP1*-like genes might play a role in phyllosphere microbiome dynamics by implementing ETI signalling pathways. However, the molecular function of the *RPP1*-like genes on chromosome 3 remains unknown. Further research is needed to obtain insights in the signalling events mediated by *RPP1*-like genes in Gorzów and Landsberg *Arabidopsis*.

Material & Methods



4 Material and Methods

4.1 Sampling sites and local sample processing

Natural *Arabidopsis thaliana* populations were collected on six sites within the region of Gorzów Wielkopolski (former Landsberg an der Warthe) in three consecutive years (2016 – 2018). The geographical location of sampling sites were determined by Rubén Alcázar and were summarized in Tab. 1. *Arabidopsis* leaf rosettes (10 samples per site) were collected in April of each sampling year representing 180 individual plant samples. From each plant rosette, leaf-surface colonizing microbes (epiphytes) and cytoplasmic microbes (endophytes) were separated along the sampling site according to (Agler et al. 2016). Leaf compartments were directly frozen on dry ice and stored at -80˚C for microbiome studies.

Tab. 1: Geographical location of local *Arabidopsis* population around Gorzów Wielkopolski.

Sampling Sites	Location	Latitude	Longitude
100	Gorzów Wielkopolski	52.73493 N	15.25088 E
200	Gorzów Wielkopolski	52.73542 N	15.25115 E
300	Gorzów Wielkopolski	52.73654 N	15.25004 E
400	Gorzów Wielkopolski	52.79430 N	15.32715 E
500	Różanki	52.79430 N	15.32715 E
600 #1	Łupowo	52.70058 N	15.12207 E
600 #2	Łupowo	52.70059 N	15.12220 E

4.2 Gorzów plant collection

During our samplings, we established a Gorzów plant collection (GPC) comprising 145 single *Arabidopsis* lines, collected according to sampling sites from 2016 to 2018. *Arabidopsis* plants were transplanted to the greenhouse for seed propagation. From each parental line, offspring generations (F1, F2) were generated for further genotyping and whole genome sequencing. A full list of the Gorzów plant collection is displayed in Tab. 2.

Tab. 2: Description of the Gorzów plant collection (GPC).

Sites	GPC ID	Sampling Year
100 - 600	A - J	2016
100 - 600	K - S	2017
100 - 600	T - AC	2018

4.3 Simultaneous DNA and RNA extraction

Homogenization of leave and microbial samples was done with the Precellys Evolution bead beater (Bertin Coorp.) with Cryolys at 6300rpm [2x30s, 15s break] at -20˚C. The epiphytic and endophytic fraction was homogenized with a zirconium bead combination [0.1mm, 0.5mm 2.8mm]. The simultaneous extraction of high quality DNA and RNA from epiphytic and endophytic samples has been obtained with an established method, based on the AllPrep DNA/RNA Kit (Qiagen, Germany). The protocol was started according to the manufacturer's protocol with the adding of 600µl Lysis Buffer. The supernatant was transferred to a QIAshredder column, to remove plant cell debris and optimize the homogenization procedure. Further, the simultaneous extraction of DNA / RNA was conducted according to manual recommendation, including a DNAse I on column digestion. RNA samples were treated additionally with Turbo™ DNAse (Thermo Fisher Scientific, USA). The complete sample set was fully randomized on 96-well plates. Finally, nucleic acids were purified using an in-house bead clean-up protocol (provided by D. S. Lundberg, MPI Tübingen) to gain sample quality. Quality check was performed using QuantiFluor® dsDNA and RNA System. Measurements were taken on a plate reader.

4.4 Amplicon sequencing spike-in

I established an internal control for amplicon sequencing libraries using 18S rRNA from zebrafish (*Danio rerio*). Wild-type zebrafish embryos (ABTL, 5h old) were provided by Aristides Arrenberg (Werner Reichardt Centre for Integrative Neuroscience, Tübingen, Germany). Approximately 100 embryos were snap frozen and homogenized with the Precellys Evolution bead beater (Bertin Coorp.) with Cryolys at 6500rpm [2x30s, 15s break] at -20˚C using zirconium beads [0.1mm, 0.5mm 2.8mm]. Genomic DNA was isolated with our custom phenol-chloroform isoamyl alcohol protocol, described by (Agler et al. 2016). Ribosomal RNA gene primers were designed (see below), based on the *Danio rerio* GRCz11 reference genome (Genbank: CM002889.2, RefSeq: NC_007116.7). Amplicons were tacked with a T7 promotor (fw) and poly-A-tail (rv) and 18S rDNA amplicons were generated using PCR [5x Q5 reaction buffer, 5x High GC Enhancer, each 2.5 mM dNTP's, each 10 µM Oligonucleotides, 100 ng gDNA, 1 U NEB Q5® High-Fidelity DNA Polymerase, nuclease-free water] cycling [95 °C 2 min; 35 times 95 °C 30 s, 65 °C 45 °C, 72 °C 60 s; 72 °C 2 min]. 18S rRNA gene amplicons were purified using NucleoSpin® Gel and PCR Clean-up mini kit (Macherey-Nagel, Germany).

An *in-vitro* transcription with PCR products was performed using MEGAscript T7 transcription Kit (Thermo Scientific Fisher, USA) according to manufacturer recommendations, to generate RNA spike-in control. Deoxynucleic acids were degraded using Turbo™ DNase (Thermo Fisher Scientific, USA). Ribonucleic acids were purified using NucleoSpin® RNA Clean-up (Macherey-Nagel, Germany). RNA quality check was performed using RNA native agarose gel electrophoresis, NanoDrop (Thermo Fisher Scientific, USA), QuantiFluor®RNA system (Promega, USA) and Agilent 2100 Bioanalyzer system (Agilent, USA).

Tab. 3: In-vitro transcription oligonucleotides

ID	Name	Sequence 5' – 3'
nPR116	T7prom-18SDr F	TAATACGACTCACTATAGGGTACCTGGTTGATCCTGCCAGTAA
nPR117	18S Dr R2 30(T)_AcII	AACGTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTAAAAACG GGCGGTCGGC

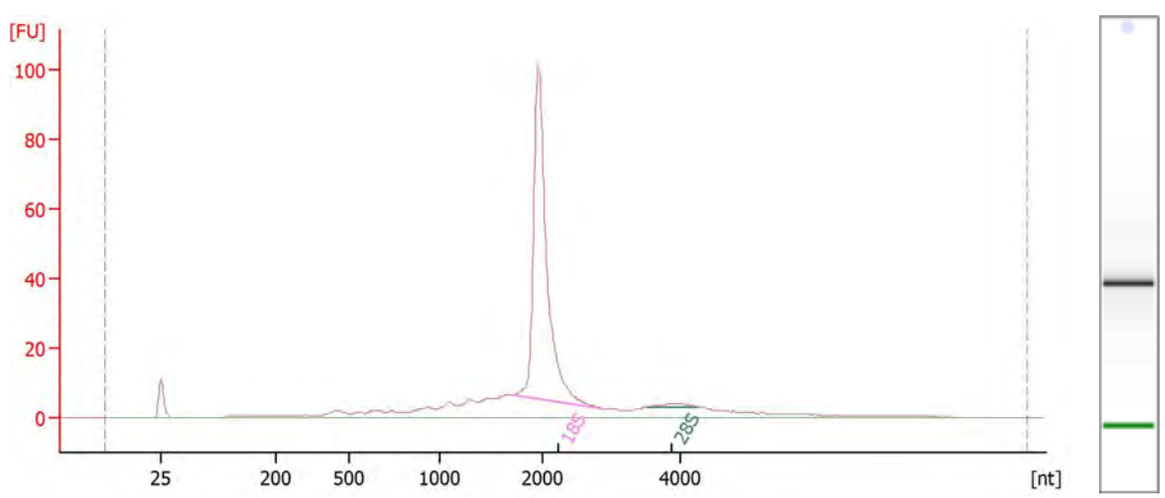


Fig. 16: Quality assessment of amplicon-sequencing spike-in control. Electropherogram of 18S rRNA *spike-in* on a high sensitivity RNA microfluidics-based detection system (Agilent 2100 bioanalyzer) shows one clear peak (A), as well as bioanalyzer image shows one major band (B) indicating good quality RNA. Concentration: 18S rRNA *Danio rerio* ~ 375.8 ng/µl (#181024-2)

4.5 Amplicon library preparation invoking spike-in control

Custom amplicon sequencing libraries were prepared using simultaneously extracted DNA and RNA samples. Ribonucleic acid samples were transcribed into complementary DNA (cDNA) using SuperScript IV™ Reverse Transcriptase (Thermo Scientific Fisher, USA). The reaction was set up as recommended by the manufacturer. Oligonucleotides are displayed in Tab. 4. The input nucleic acid was normalized to an equal amount and a ratio of nucleic acids (DNA, RNA) between epiphytic

and endophytic samples were set to 1:100. In addition to each template nucleic acid, I added an internal control (spike-in) to each sequencing library. Thus, 25 pg spike-in was added to each DNA (18S rDNA) and cDNA (18S rcDNA) sequencing library. From that point forward, custom libraries were prepared as published by (M. Agler et al., 2016). Bacterial amplicons targeting 16S rRNA genes (referred to rDNA) V4/V5 regions and protist amplicons were targeting 18S rDNA V8/V9 regions. Sequencing of the sequencing library pools (total 1440 libraries) was performed at the Genome Center of the Max Planck Institute for Developmental Biology Tübingen, Germany (https://www.eb.tuebingen.mpg.de/ infrastructure/genome-center/) using an Illumina MiSeq platform (MiSeq V3 kits). Basecall BCL files were converted into fastq files using bcl2fastq.

Tab. 4: Oligonucleotides for complementary DNA

ID	Name	Sequence 5' – 3'	Reference
T052	B806R	GGACTACHVGGGTWTCTAAT	(Bodenhausen et al., 2013)
T027	1193R	ACGTCATCCCCACCTTCC	(Bodenhausen et al., 2013)
G002	R1200	CCCGTGTTGAGTCAAATTAAGC	(Hadziavdic et al., 2014)
G004	R1797	TGATCCTTCTGCAGGTTCACCTAC	(Hadziavdic et al., 2014)
T025	799F	AACMGGATTAGATACCCKG	(Bodenhausen et al., 2013)
G003	F1422	ATAACAGGTCTGTGATGCCC	(Hadziavdic et al., 2014)
5M30-F		AGATCAGGGGCTCAGCTAACGCGTGAACAC	(Agler et al. 2016)
CI1BV5		TTTTGGCAGGGCGTACTAAACCCACTTACT	(Agler et al. 2016)
GC006		TGATGTATTCAACGAGTTCACACCTTGGCCGACAG	Mari et al., unpublished
GC008		TCTAAATGATAAGGTTTAGTGGACTTCTCGCGACG	Mari et al., unpublished

4.6 Amplicon quality processing, clustering and classification

Raw sequencing data were processed using the mothur pipeline (v. 1.42.2). Paired-reads were quality filtered with screen.seqs (parameter= minoverlap=5, maxambig=0, maxhomop=10, minlength=100, maxlength=600) and demultiplexed according to their 12 bp barcode-indices. Chimeric sequences were removed using uchime and sequences were grouped into operational taxonomic units (OTUs) (Edgar et al. 2011). Taxonomic assignment was performed based on reference databases for bacterial 16S rRNA genes (Greengenes gg_13_8_99) and protists (PR2, v. 4.11.0) (Guillou et al. 2013; DeSantis et al. 2006). Reference databases were completed with the full phage genomes of PhiX (sequence and taxonomy files), an internal Illumina sequencing standard and spike-in sequence (18S rRNA full length gene of *Danio rerio*). Final OTU-tables were converted into biom-files and further processed using Qiime2 (Bolyen et al. 2019) and in-house R scripts. R packages such as qiime2R, phyloseq and microbiome were implemented to calculate microbial diversities. A beta-dispersion analysis was conducted on Bray-Curtis dissimilarities using

rarefied relative abundance OTU-tables to calculate sample-to-sample variability. The multivariate homogeneity of group dispersions analysis was conducted within the R package Vegan using 'betadisper'.

4.7 Identification of microbial core community

Persistent core microbes across sampling sites and years were calculated using CORE-function in qiime2. Core microbes, represented by operational taxonomic units, had to be present in ≥ 85% of all samples to be counted, applied to all amplicons. Multi sequence alignments using Clustal Omega (v.1.2.4) were conducted based on representative sequences of each core OTU. Multi Sequence alignments (ClustalW-format) of single amplicons were used to calculate rooted phylogenetic trees using iqtree (parameter: iqtree -s clustalo_output -st DNA -m TEST -bb 1000 -alrt 1000).

4.8 Genotyping determining *RPP1*-like CNV in *Gw Arabidopsis*

Genotyping of *RPP1*-like genes in *Arabidopsis thaliana* was established using a droplet digital PCR (ddPCR) approach. Thus, copy number variations (CNVs) of *RPP1*-like genes (target) were calculated, based on UBQ5 (At3G62250, reference) genes. Oligonucleotides and Taq-Man probes were designed for target and reference genes (see Tab. 4). Each duplexed ddPCR reaction was performed in triplicates and included unique oligonucleotides for target and reference genes. The ddPCR reactions [20 µl] were set up as follow: 2 x ddPCR™ Supermix for Probes (Bio-Rad, Ref: 1863023), each 900nM *RPP1*-like Fw / Rv, 250nM FAM-Taq-Man probe, each 900nM UBQ5 Fw / Rv, 250nM Hex-Taq-Man probe, 0.5 ng gDNA / HindIII, NFW. Droplets were generated using a QX200 Droplet Generator (Bio-Rad, Ref: 1864002) and transferred to 96-well PCR plates for thermocycling [98 °C for 600 s; 40 cycles of 94 °C (ramp 2 °C/s) for 30 s, 60 °C for 20 s and 72 °C for 15 s, 98 °C 600 s]. Quantitative measurements were conducted with a QX200 Droplet Reader (Bio-Rad, Ref: 1864003) and analyzed using Quanta Soft (v. 1.0) and in-house R scripts. FAM-detected droplets of *RPP1*-like copies were normalized to Hex-detected droplets of UBQ5 copies to perform an absolute quantification.

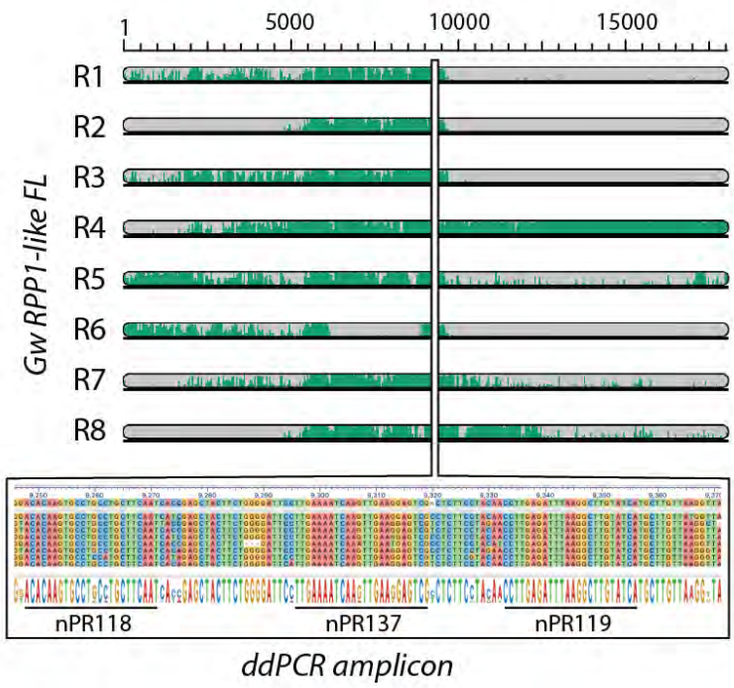


Fig. 17: Oligonucleotide and probe design of *RPP1*-like genotyping using droplet digital PCR.

Tab. 5: Oligonucleotides for ddPCR genotyping in Gorzów *Arabidopsis* populations.

ID	Sequence 5' -> 3'	Description
nPR118	ACA CAA GTG CCT SCM TGC TTC AAT	*RPP1*-like Fw primer
nPR119	TGA TAC AAG CCT TAA ATC TCA AGG	*RPP1*-like Rv primer
nPR137	FAM-TTGAAAATCAAGTTGAAGGAGTCG-BHQ1	*RPP1*-like probe
nPR124	GAAGGCGAAGATCCAAGACAAGGAA	UBQ5 Fw primer
nPR125	GGAGGACGAGATGAAGCG	UBQ5 Rv primer
nPR136	Hex-AACAGCTTGAAGACGGCCG-BHQ1	UBQ5 probe

4.9 Whole genome sequencing, assembly and annotation

The isolation of high-molecular weight DNA (HMW-gDNA) from *Arabidopsis* bud collections of individual plants (up to 10 cycles) was performed using an optimized Phenol-Chloroform-Isoamyl alcohol extraction protocol, described by (Schwessinger and Rathjen 2017) or followed manufacturer recommendations using Plant Midi II Kits (Macherey-Nagel, Germany). Nucleic acids were further purified using an in-house magnetic bead clean protocol, provided by D. S. Lundberg (MPI Tübingen). Quality controls of HMW-gDNA was conducted using NanoDrop

(Thermo Fisher Scientifics, USA), QuantiFluor® dsDNA system (Promega, USA) and Bioanalyzer 2100 system (Agilent Scientific Instruments, USA).

Long read whole genome sequencing was performed by the Max Planck-Genome-centre Cologne, Germany (https://mpgc.mpipz.mpg.de/home/). Each *Arabidopsis* genome was sequenced on two SMRT- cells and raw data were delivered via FTP-server. Illumina short reads (150PE) were generated for each *Arabidopsis* genome by Novogene (Norwich, UK).

Genome assemblies using long reads were performed with Canu (v. 1.8) (Koren et al. 2017). A reference-based scaffolding was conducted with RagTag (v.2.0.1) (Alonge et al. 2019). The assembly of Ler-0 (accession: GCA_900660825) published by (Goel et al. 2019) was used as a reference. Obtained scaffolds were polished with Illumina short reads using Pilon (v. 1.23) to reduce sequencing errors (Walker et al. 2014). Quality checks of assemblies and scaffolded genomes were conducted using QUAST (v. 5.02) (Gurevich et al. 2013).

Coding genes were annotated implementing *ab initio* gene predictions, such as Augustus (v. 3.2.3), GlimmerHMM (v. 3.0.4) and SNAP (Majoros, Pertea, and Salzberg 2004; Leskovec and Sosič 2016; Stanke and Morgenstern 2005). Reference protein sequences from Araport11 annotation (Columbia-0) were used for protein sequence alignments within each genome annotation using Exonerate (v. 2.2.0) (C. Y. Cheng et al. 2017; Slater and Birney 2005). Gene predictions and annotations were concatenated and integrated into consensus gene models based on Evidence Modeler (v. 1.1.1) (Haas et al. 2008). Noncoding genes were annotated using tRNAscan-SE (v.2.0) incorporating Infernal (v1.1) (Chan et al. 2021; Nawrocki and Eddy 2013). In addition, ULTRA was used to locate short tandem repeats (Olson and Wheeler 2018). Transposable elements were annotated using RepeatMasker (v. 4.1.2-p1) (Smit 1993). Resistance genes were further annotated using RGAugury (v.) (P. Li et al. 2016). Centromeric regions were identified by searching (blastn, hmmer) for the consensus sequence of 178-bp tandem repeats published by (Naish et al. 2021). *Arabidopsis* gene models were used for further downstream analysis.

4.10 Pan-genome analysis

Synteny maps and genomic rearrangements were analysed using SYRI (Synteny and Rearrangement Identifier) (Goel et al. 2019). Downstream analysis was performed with in-house R scripts implementing packages, such as karyoploteR, rtracklayer, RIdeogram and Gviz (Gel and Serra 2017; Lawrence, Gentleman, and Carey 2009; Hao et al. 2020; Hahne and Ivanek 2016).

RPP1-like resistance gene loci were extracted using border gene locations of AT3G44600 (left border) and AT3G44690 (right border) in gene model annotations. The *RPP1*-like loci in the reference *Landsberg erecta* (Ler-0, 2019, v2) was extracted using blastn to locate border genes. Genomic sequences of *RPP1*-like loci were extracted using bedtools (v. 2.26.0) getfasta function (Quinlan and Hall 2010). Genomes showing both border genes were considered for further downstream analysis. Genomic distances of *RPP1*-like loci were calculated by genome skimming using Skmer (v. 3.0.2) (Sarmashghi et al. 2019). Visualization and statistics of genomic distance matrices were conducted in R (v using R packages 'ggfortify' and 'vegan' (Tang, Horikoshi, and Li 2016; "Vegan: Community Ecology Package" n.d.).

Gorzów gene models and reference genomic feature files were used to extract gene annotations within the *RPP1*-like loci. Genomic feature files were converted into genbank files using seqret (Rice, Longden, and Bleasby 2000) and visualized with the gene cluster comparison tool Clinker (Gilchrist and Chooi 2021). Protein translations within *RPP1*-like loci were extracted from Clinker global alignments (all vs. all). Thus, border proteins and *RPP1*-like proteins were identified using blastp from Diamond against a self-made database including AT3G44600, AT3G44690 (TAIR10) and *RPP1*-like R1-R8 [ACJ64856.1] proteins) (Buchfink, Xie, and Huson 2014).

4.11 Statistical analysis

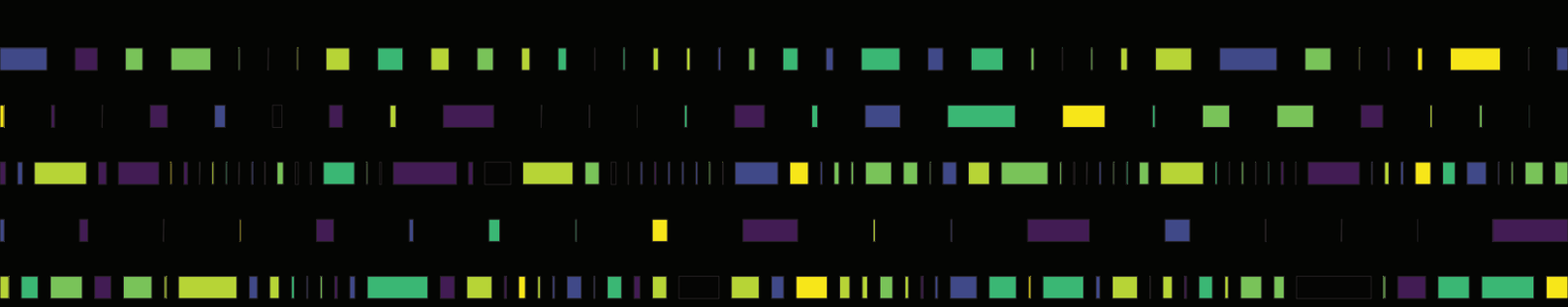
Statistics were performed with either customized scripts using R (pairwise-Wilcoxon-Test, PERMANOVA) or qiime2-implemented functions (Kruskal-Wallis-Test, PERMANOVA).

4.12 Data and code availability

Supporting data of the current work are available within the paper and appending Supplementary Information files. The datasets generated and analyzed within this study are available from the corresponding author upon request. Raw sequencing data of phyllosphere microbiota and *Arabidopsis* whole genome sequencing, including assemblies and annotations are accessible in the European Nucleotide Archive.

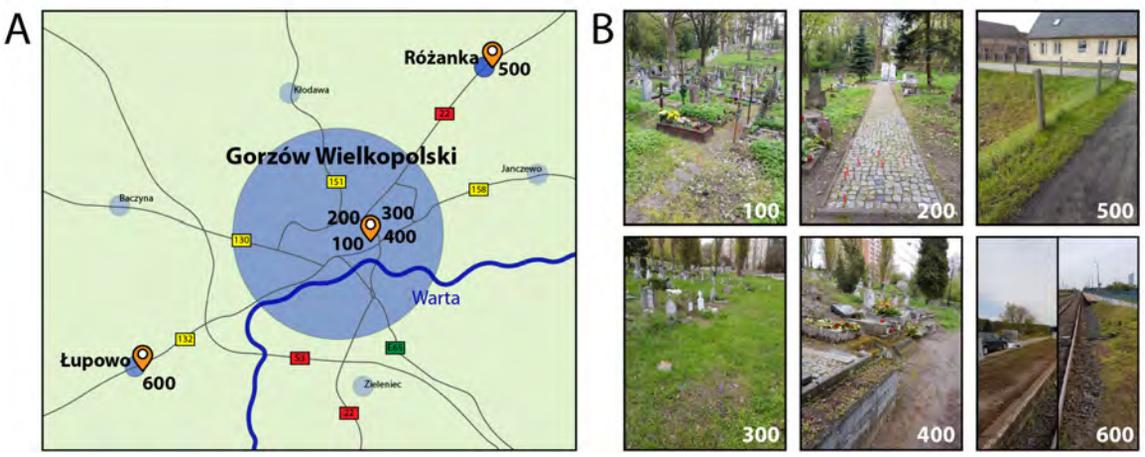
Custom scripts are freely accessible at github/RUNGEP.

Supplementary Data



5 Supplementary Data

5.1 *Arabidopsis* sampling sites around Gorzów Wielkopolski



S-Fig. 1: Geographical location of natural Gorzów *Arabidopsis* sampling sites. Map (A) and representative views (B) of sampling sites. Sites are located on the cemetery Świętokrzyski in central Gorzów Wielkopolski (100-400), as well as north east in Różanki (500) and south west in Łupowo (600).

5.2 Whole genome sequencing data

S-Tab. 1: Whole genome sequencing overview using long-read PacBio (LR) and short-read (SR) Illumina sequencing. The majority of LR sequences were generated on two SMRT-cells per genome. PacBio coverage was calculated on high-quality sequences within the trimming phase of Canu (v.1.8) assemblies. Illumina SR (PE150) coverage was calculated on an average *Arabidopsis thaliana* genome size of 135Mbp. Long-reads were assembled into contigs using Canu and scaffolded on the reference genome *Landsberg erecta* (Ler-0) using RagTag (v. 2.0.1). Illumina reads were used for sequencing correction and gap filling using samtools (v. 1.10) and pilon (v. 1.24).

Accession			PacBio			Illumina paired-end		
Plant line	Site	Stock number	Reads	Mean Length	Depth	Pairs	Length	Depth
100D	100	sPR0373	449,327	13,745	45.7	24,507,954	150	54.5
100H	100	sPR0376	303,761	15,563	35.0	25,139,262	150	55.9
100K	100	sPR0379	342,441	14,066	35.7	24,224,082	150	53.8
100S	100	sPR0383	544,571	12,789	51.6	25,735,510	150	57.2
100Y	100	sPR0738	1,863,495	10,128	139.8	24,152,102	150	53.7
100AA	100	sPR0740	237,057	20,761	36.5	24,280,336	150	54.0
100AD	100	sPR0743	556,238	12,161	50.1	24,008,698	150	53.4
200E	200	sPR0431	457,517	13,751	46.6	24,959,970	150	55.5
300A	300	sPR0478	432,222	13,122	42.0	24,492,390	150	54.4
300AD	300	sPR0754	361,911	14,007	37.6	26,244,262	150	58.3
400D	400	sPR0537	362,509	13,626	36.6	29,747,746	150	66.1
400L	400	sPR0544	445,445	10,784	35.6	25,386,064	150	56.4
400Y	400	sPR0760	306,026	9,781	22.2	26,436,620	150	58.7
500R	500	sPR0613	447,794	10,617	35.2	25,358,458	150	56.4
600H	600	sPR0683	267,864	17,407	34.5	24,408,276	150	54.2
600I	600	sPR0684	349,821	14,039	36.4	24,446,250	150	54.3
La-0a	Ref	CS1298	379,694	12,532	35.2	76,960,134	150	171.0
La-0b	Ref	CS76538	347,204	13,404	34.5	74,380,684	150	165.3

5.3 Contig assembly statistics

S-Tab. 2: PacBio long reads of one to three SMRT- cells were assembled into contigs using Canu (v. 1.8). Contig assembly statistics were calculated using QUAST (v. 5.0.2).

Plant line	Seq ID	Contigs	Total bases	GC (%)	N50	N90	L50	L90	Max
100D	3812A2	4,382	239,759,699	37.2	171,211	28,700	203	1,398	3,204,991
100H	3812B2	758	138,444,090	36.2	773,525	369,308	44	108	5,307,894
100K	3812F2	263	126,243,422	36.1	3,943,680	1,734,872	9	21	11,174,918
100S	3812D2	1,824	172,859,825	36.8	432,715	85,845	98	302	2,785,117
100Y	3812G2	4,528	233,943,511	36.7	85,201	36,933	750	1,818	612,465
100AA	3812M2	222	126,902,589	36.21	7,949,486	2,634,134	6	13	14,343,189
100AD	3812N2	1,153	152,166,702	36.25	451,280	187,394	99	232	1,994,504
200E	3812E2	3,128	181,534,898	36.4	85,894	43,514	616	1,357	551,899
300A	3812I2	1,497	149,310,330	36.2	249,438	107,365	177	394	1,001,840
300AD	3812J2	484	131,981,105	36.34	2,402,987	900,762	16	37	9,118,441
400D	4415A2	661	132,400,675	36.6	4,280,141	2,083,359	10	20	11,640,849
400L	4415B2	596	130,298,647	36.2	1,360,043	597,147	26	61	7,452,123
400Y	3812P1	334	121,850,295	36.1	1,580,099	769,787	23	51	5,905,003
500R	4415E2	1,388	135,248,263	36.5	361,680	146,590	104	243	1,883,544
600H	3812K2	1,456	168,867,214	36.1	1,502,553	94,775	27	140	6,545,858
600I	3812L2	1,771	171,027,729	36.2	441,123	69,636	85	325	3,815,722
La-0a	4415C2	1,312	142,230,605	36.86	673,011	238,670	60	148	3,890,436
La-0b	4415D3	1,293	141,589,923	36.7	1,055,862	409,679	36	85	5,434,911

5.4 Polished scaffolds statistics

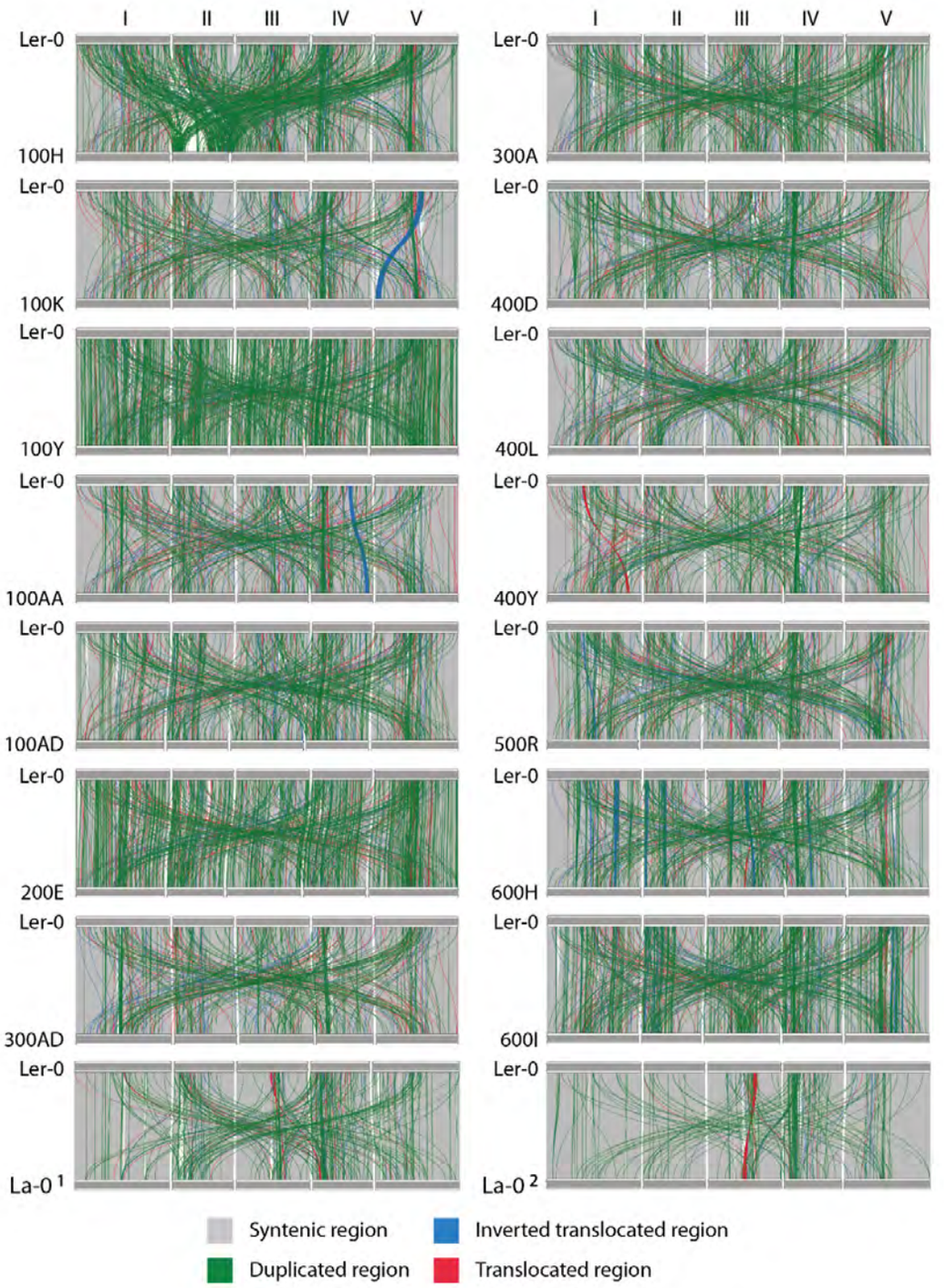
S-Tab. 3: The completeness of Gorzów *Arabidopsis* genomes were validated using Benchmark Universal Single-Copy Orthologs (BUSCO, v. 5.2.2) of the lineage dataset "embryophyta_odb10" including the reference genome *Landsberg erecta* (Ler-0, 2019, v2) published by (Jiao and Schneeberger 2020).

Plant line	Seq ID	Busco				
		Complete	[%]	Fragmented	Missing	Total
100D	3812A2	1,568	97.1	3	43	1,614
100H	3812B2	1,594	98.8	3	17	1,614
100K	3812F2	1,602	99.3	3	9	1,614
100S	3812D2	1,576	97.6	5	33	1,614
100Y	3812G2	1,506	93.3	4	104	1,614
100AA	3812M2	1,599	99.1	3	12	1,614
100AD	3812N2	1,559	96.6	3	52	1,614
200E	3812E2	1,462	90.6	18	134	1,614
300A	3812I2	1,568	97.1	9	37	1,614
300AD	3812J2	1,598	99.0	4	12	1,614
400D	4415A2	1,591	98.6	3	20	1,614
400L	4415B2	1,579	97.8	3	32	1,614
400Y	3812P1	1,547	95.8	21	46	1,614
500R	4415E2	1,565	97.0	7	42	1,614
600H	3812K2	1,587	98.3	6	21	1,614
600I	3812L2	1,574	97.5	3	37	1,614
La-0a	4415C2	1,581	98.0	5	28	1,614
La-0b	4415D3	1,593	98.7	4	17	1,614
Ler-0	Ref	1,603	99.3	3	8	1,614

S-Tab. 4: Reference-based scaffolding of contigs was conducted using chromosome-level assemblies of *Arabidopsis thaliana* ecotype *Landsberg erecta* (Ler-0, 2019, v2) published by (Jiao and Schneeberger 2020). Scaffolds were generated using RagTaq (v. 2.0.1) and polished including Illumina short reads using pilon (v. 1.24). Scaffold statistics of pseudomolecules (split into scaffolds) were calculated using QUAST (v. 5.0.2).

Plant Line	Seq ID	Pseudomolecules	Scaffolds	N's per 100 kbp	Total length	N50	NG50	N75	NG75	L50	LG50	L75	LG75	Max.
100D	3812A2	5	555	42.4	126,196,764	723,028	779,775	327,456	360,748	50	45	117	100	3,205,132
100H	3812B2	5	257	20.6	118,617,912	995,606	995,606	520,729	520,729	34	34	78	78	5,307,756
100K	3812F2	5	97	7.4	117,980,688	4,468,310	3,943,704	1,478,587	1,458,477	8	9	20	21	10,920,042
100S	3812D2	5	357	25.8	121,697,851	659,245	665,853	339,631	353,279	54	51	118	111	2,785,117
100Y	3812G2	5	1,622	110.8	144,195,761	125,696	147,235	75,353	100,863	360	266	729	508	612,530
100AA	3812M2	5	107	7.6	118,567,429	7,935,770	7,935,770	3,152,733	3,152,733	6	6	12	12	14,298,175
100AD	3812N2	5	405	32.6	120,075,074	512,989	516,737	276,559	280,857	75	73	155	150	1,994,914
200E	3812E2	5	1,706	123.9	131,421,521	106,540	115,632	63,212	74,049	394	336	795	654	551,899
300A	3812I2	5	707	56.3	122,276,386	285,935	301,354	163,697	176,019	131	125	271	255	1,002,046
300AD	3812J2	5	163	13.0	118,791,981	2,458,501	2,458,501	1,220,899	1,220,899	15	15	32	32	9,118,995
400D	4415A2	5	128	9.6	117,449,313	4,279,672	4,279,672	2,438,895	2,438,895	9	9	18	18	10,560,973
400L	4415B2	5	178	13.6	116,441,951	1,521,071	1,423,705	767,193	720,841	22	23	49	51	7,452,123
400Y	3812P1	5	173	13.8	114,953,803	1,623,323	1,620,258	821,742	779,617	22	23	47	50	5,381,675
500R	4415E2	5	577	46.4	116,619,118	408,615	400,716	213,419	201,804	82	84	179	186	1,883,776
600H	3812K2	5	262	20.4	122,843,210	2,498,146	2,542,532	1,004,887	1,084,084	17	16	37	33	6,011,160
600I	3812L2	5	450	36.2	121,440,666	752,706	812,207	357,344	390,299	45	43	104	98	3,817,048
La-0a	4415C2	5	320	18.7	119,126,468	798,208	798,208	396,503	402,712	45	45	97	96	3,890,848
La-0b	4415D3	5	204	14.7	118,350,391	1,322,577	1,322,577	784,137	784,137	28	28	57	57	5,266,805
Ler-0	Ref	5	49	18.6	118,502,335	11,168,341	11,168,341	6,074,025	6,074,025	5	5	9	9	14,816,562

5.5 Synteny maps of Gorzów genomes



S-Fig. 2: Synteny map of Gorzów genomes against the reference genome *Landsberg erecta* (Ler-0). Synteny was calculated using SYRI (Goel et al. 2019). Graphics were created using R (package Rideogram, v 0.2.2) (Hao et al. 2020).

5.6 Number and total length of rearranged regions

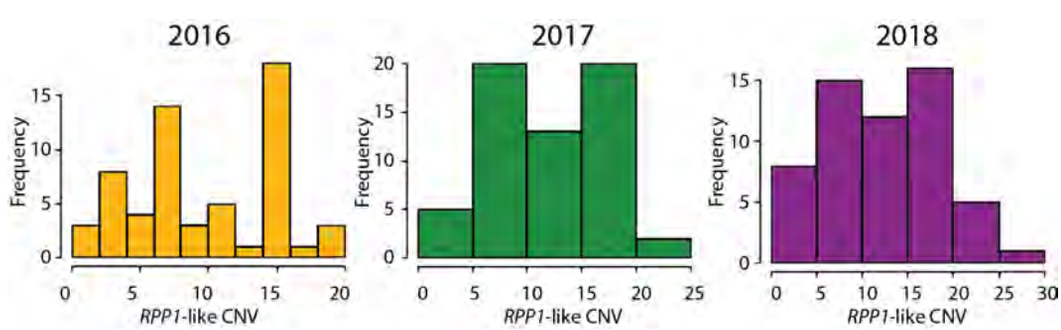
S-Tab. 5: Aligned regions in megabases of Gorzów genomes against the reference genome Ler-0 using SYRI. Annotations are described as alignment in syntenic region (SYNAL), translocated region (TRANSAL), inverted translocated region (INVTRAL) and duplicated region (DUPAL). Total displays full alignment length.

Genome	[Mb]				
	SYNAL	TRANSAL	INVTRAL	DUPAL	TOTAL
100H	105.0	1.6	3.0	5.3	114.9
100K	106.6	1.8	0.7	1.5	110.6
100Y	99.7	0.9	1.5	8.0	110.1
100AA	105.9	1.9	0.6	1.7	110.1
100AD	102.6	0.9	0.9	3.1	107.5
200E	97.6	0.5	1.3	5.7	105.2
300A	104.0	0.5	0.7	3.2	108.3
300AD	108.2	0.4	0.3	1.8	110.7
400D	106.9	0.5	0.4	2.0	109.7
400L	105.1	0.4	0.9	1.5	107.9
400Y	104.5	0.3	1.0	4.9	110.8
500R	105.0	0.9	0.7	1.8	108.3
600I	107.9	1.8	0.8	4.4	115.0
600H	106.9	3.3	1.0	2.9	114.0
La-0 1	106.5	0.3	0.9	1.6	109.2
La-0 2	111.5	0.2	0.9	1.1	113.7

S-Tab. 6: Syntenic and non-syntenic (translocated) regions of Gorzów genomes and La-0 against the reference genome Ler-0 using SYRI. Syntenic and non-syntenic regions are described as single nucleotide polymorphism (SNP), Deletion in query (DEL), highly diverged regions (HDR), Copy gain in query (CPG), Copy loss in query (CPL) and Insertion in query (INS).

Genome	Syntenic regions [Kb]						Translocated regions [Kb]					
	SNP	DEL	HDR	CPG	CPL	INS	SNP	DEL	HDR	CPG	CPL	INS
100H	31.4	111.8	817.9	32.6	43.5	2.8	15.1	13.6	231.2	2.9	38.4	1.3
100K	12.4	25.4	183.7	2.5	25.4	0.9	17.3	29.5	283.8	4.9	114.9	1.8
100Y	51.6	120.2	230.3	15.4	87.7	7.7	10.2	40.0	127.8	1.5	59.3	1.3
100AA	11.9	6.3	190.5	1.7	26.3	0.8	12.5	39.0	237.8	35.5	25.0	1.4
100AD	14.5	22.4	205.0	7.4	5.1	1.8	6.0	3.7	132.3	1.5	9.8	0.6
200E	32.4	184.9	207.6	1.8	66.8	5.5	6.2	5.7	31.2	10.4	5.6	0.6
300A	15.9	30.4	547.2	2.6	8.9	2.1	5.9	24.6	48.0	1.6	0.0	0.6
300AD	15.1	28.8	324.0	2.5	9.1	1.4	4.7	6.4	17.7	1.5	0.0	0.4
400D	6.8	5.0	27.1	4.2	20.3	0.7	6.1	2.6	72.6	0.0	1.7	0.7
400L	16.0	18.0	195.3	0.9	3.6	1.4	5.6	3.0	10.9	7.5	5.8	0.5
400Y	10.9	14.7	92.6	7.8	11.5	1.2	3.5	9.8	7.7	1.1	5.1	0.3
500R	8.5	9.9	232.9	4.9	9.7	0.8	10.2	9.0	114.2	10.8	6.4	1.1
600I	12.9	60.0	68.9	2.2	10.3	1.2	19.9	27.1	662.7	88.8	38.0	2.5
600H	12.7	11.1	64.3	3.8	6.1	1.4	26.3	49.5	1746.1	31.1	38.4	3.5
La-0 1	15.2	21.5	186.2	2.1	26.0	1.5	4.0	1.4	22.4	0.0	2.1	0.4
La-0 2	13.5	19.1	386.0	17.3	23.4	1.1	2.3	0.8	15.8	0.0	0.0	0.3

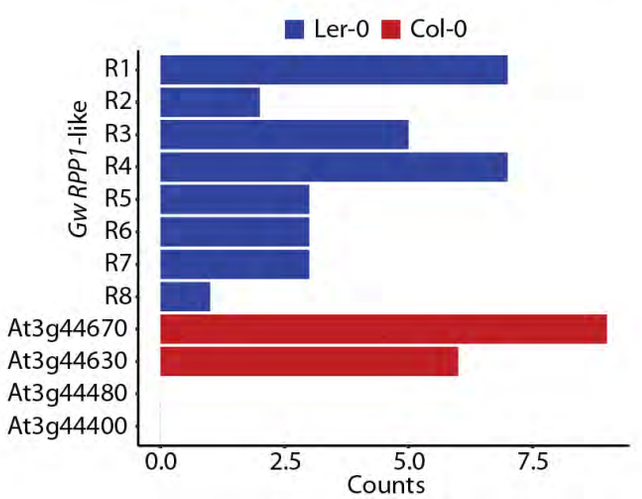
5.7 Year distribution of *RPP1*-like CNVs



S-Fig. 3: Distribution of *RPP1*-like copy numbers across sampling years, obtained by ddPCR genotyping.

5.8 Validation of *RPP1*-like genotyping

The genotyping of local Gorzów *Arabidopsis thaliana* using ddPCR was verified by sanger sequencing of *RPP1* amplicons from control lines Columbia (Col-0) and Landsberg (Ler-0). The ddPCR assay was performed on controls as described in the method section (see 4.8). After PCR cycling, ddPCR reactions were used for DNA extraction using a Phenol-Chloroform-Isoamyl alcohol protocol described by Agler et al. (2016). Obtained DNA was purified using an in-house magnetic bead clean-up protocol (provided by D. S. Lundberg, MPI Tübingen) and quantified using QuantiFluor® dsDNA system (Promega, USA). Purified *RPP1* amplicons from Col-0 and Ler-0 were cloned in the blunt-end vector pJET1.2 according to manufacturer recommendation (0.15 pmol ends ~ 6 ng [121bp amplicons]) (ThermoFisher, USA). Plasmids were transformed into *E.coli* Top10 competent cells and streaked onto selective media (LB + Amp100) ON at 37℃. High-copy plasmid DNA was extracted using NucleoSpin® Plasmid Mini kit (Macherey-Nagel, Germany). Plasmids were sent for sanger sequencing including primer of T7 promoter (nPR131).

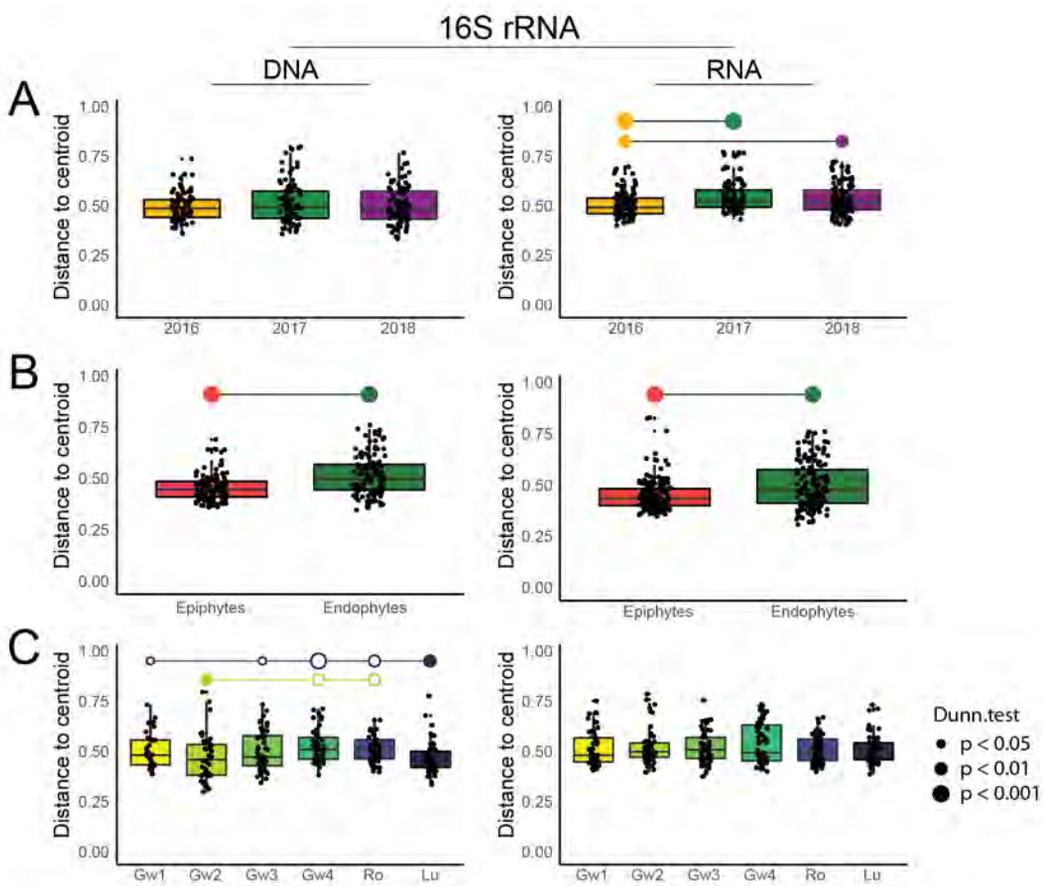


S-Fig. 4: Validation of ddPCR genotyping for Gorzów *Arabidopsis* populations. Barplot representing local blastn results of control-based (Col-0 ~ *RPP1*-, Ler-0 ~ *RPP1*+) ddPCR amplicons.

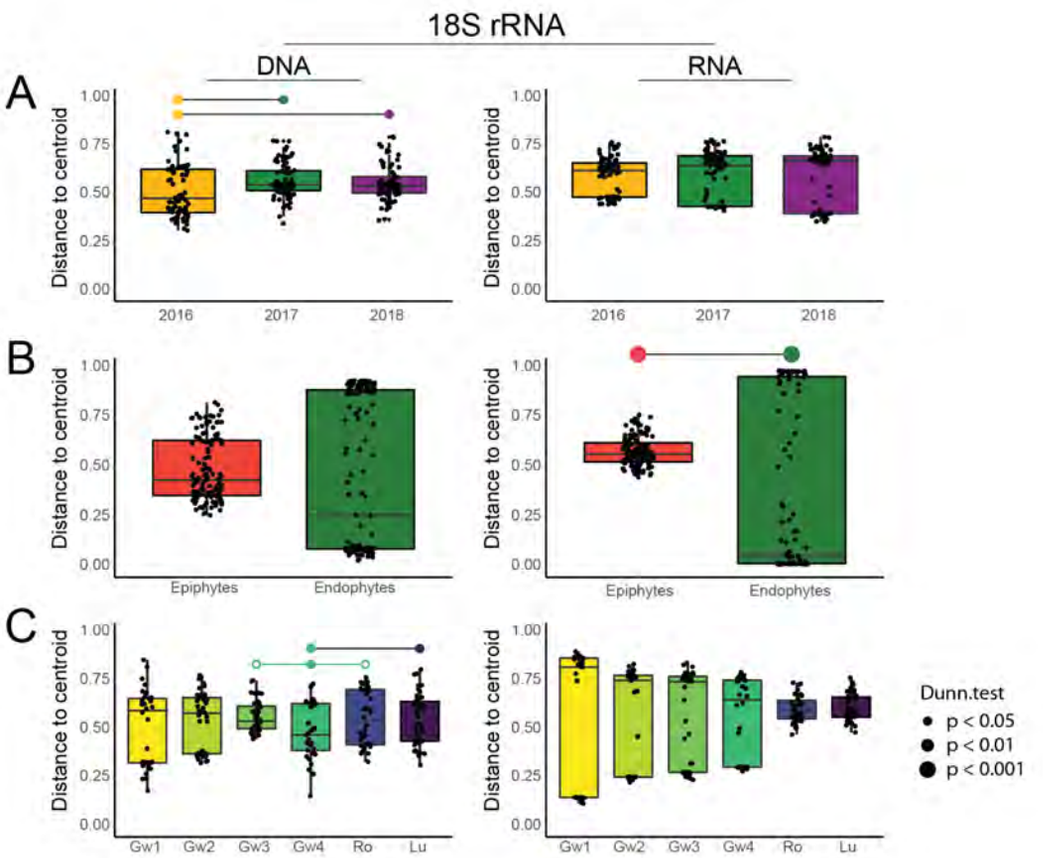
In total, I obtained sequences of 24 plasmids (pJET1.2-*RPP1*) for control lines Col-0 (*RPP1*-) and Ler-0 (*RPP1*+). Sequencing results were analysed with local blastn (ncbi-blast-2.2.30+) using a *RPP1*-specific database (makeblastdb -in ddPCR_amplicons.fasta -dbtype nucl -out ddPCR_amplicons_mydb.fa), including ddPCR amplicons of *RPP1* and *RPP1*-like genes from Col-0 and Ler-0. Blast results were considered with sequence identities > 97 %. In addition, *RPP1*-like blast hits in Col-0 were considered as false positives. Interestingly, all Col-0 amplicons are

related to *RPP1*-like genes (AT3G44630, AT3G44670). *RPP1* amplicons obtained from Ler-0 showed high similarities to *Gw RPP1*-like genes (see S-Fig. 4). Our results display a high specificity of the ddPCR genotyping approach applicable for natural Gorzów *Arabidopsis* populations.

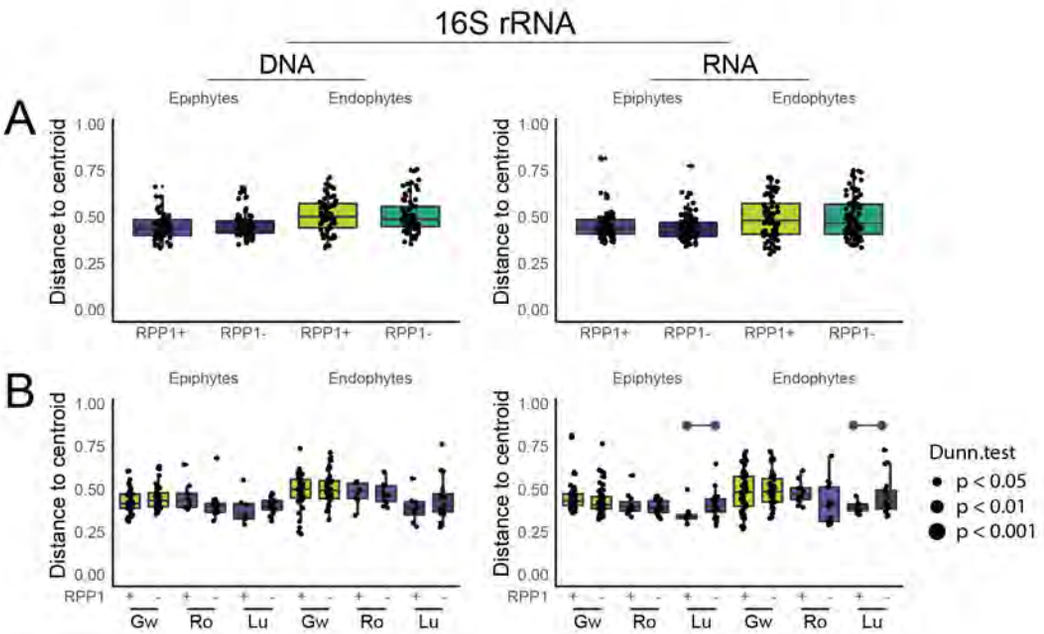
5.9 Beta-dispersion analysis on Bray-Curtis dissimilarities to compare sample-to-sample variation accounting for environmental factors and *RPP1*-like alleles



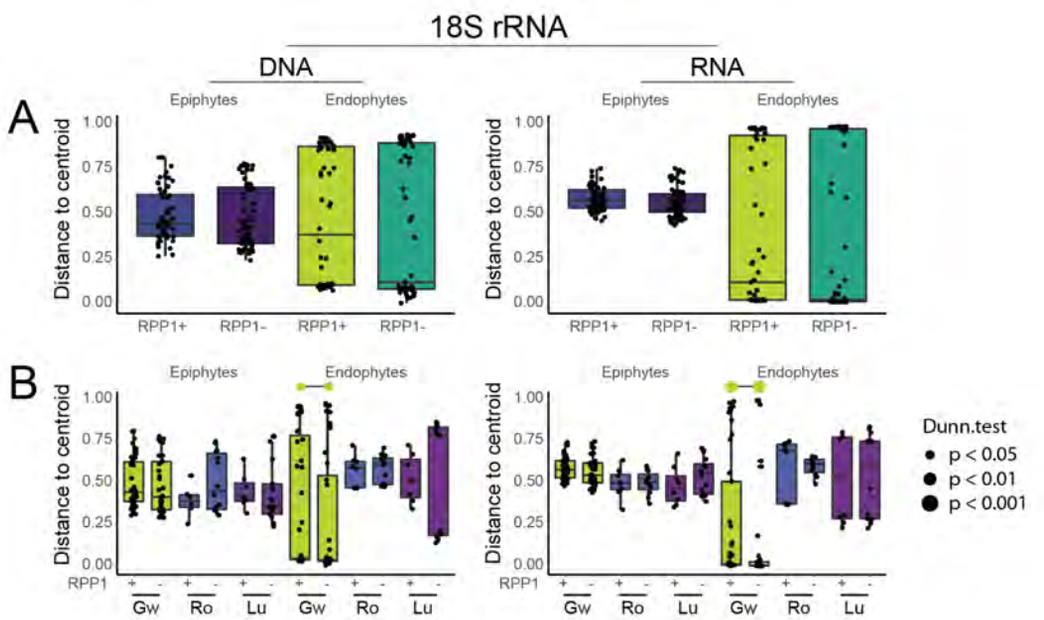
S-Fig. 5: Beta-dispersion analysis on Bray-Curtis dissimilarities to compare sample-to-sample variation on environmental features for 16S rRNA amplicons using simultaneously DNA-seq (left) and cDNA-seq (right). Samples are grouped by A) sampling year, B) leaf compartment, B) and C) sampling sites. Statistics: Dunn test. Horizontal lines accounting for statistical significance between sample groups. Circle diameter reflect p-values.



S-Fig. 6: Beta-dispersion analysis on Bray-Curtis dissimilarities to compare sample-to-sample variation on environmental features for 18S rRNA amplicons using simultaneously DNA-seq (left) and cDNA-seq (right). Samples are grouped by A) sampling year, B) leaf compartment and C) sampling sites. Statistics: Dunn test. Horizontal lines accounting for statistical significance between sample groups. Circle diameter reflect p-values.

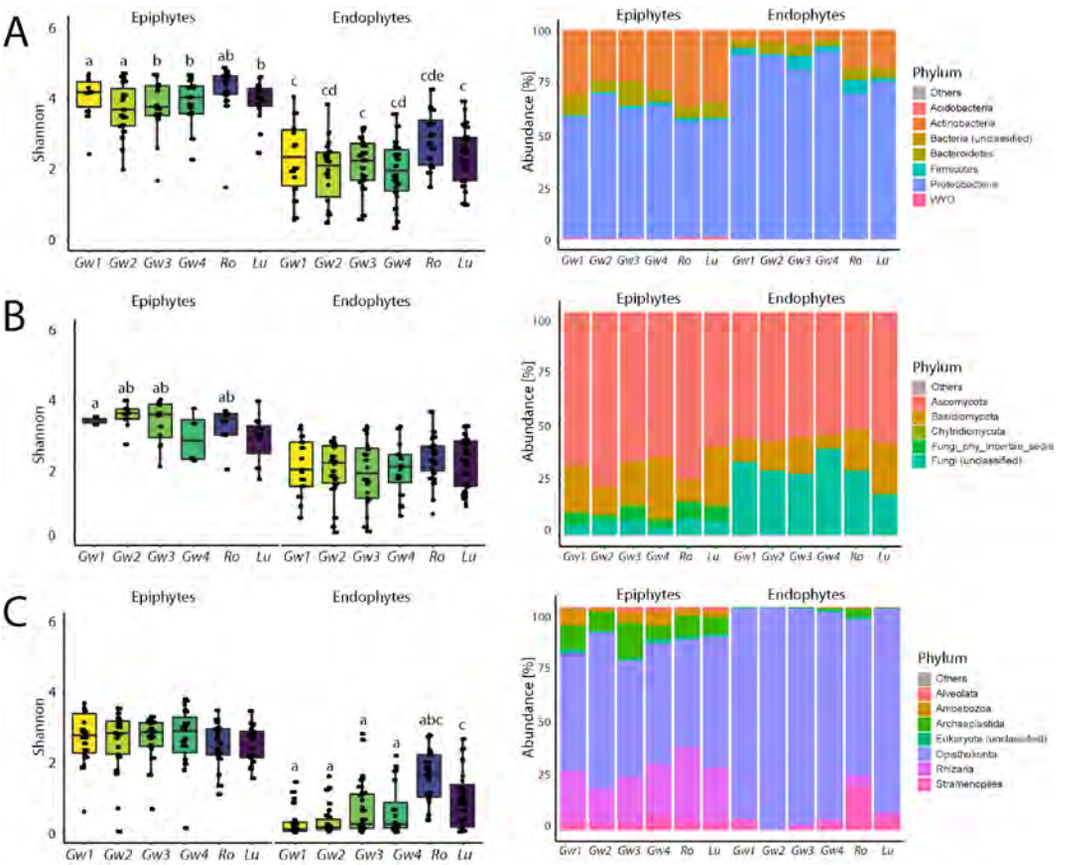


S-Fig. 7: Beta-dispersion analysis on Bray-Curtis dissimilarities to compare sample-to-sample variation in *RPP1* haplotypes for 16S rRNA amplicons using simultaneously DNA-seq (left) and cDNA-seq (right). Samples are grouped by A) Compartment x *RPP1*-like allele and B) Location x Compartment x *RPP1*-like allele. Statistics: Dunn test. Horizontal lines accounting for statistical significance between sample groups. Circle diameter reflect p-values.

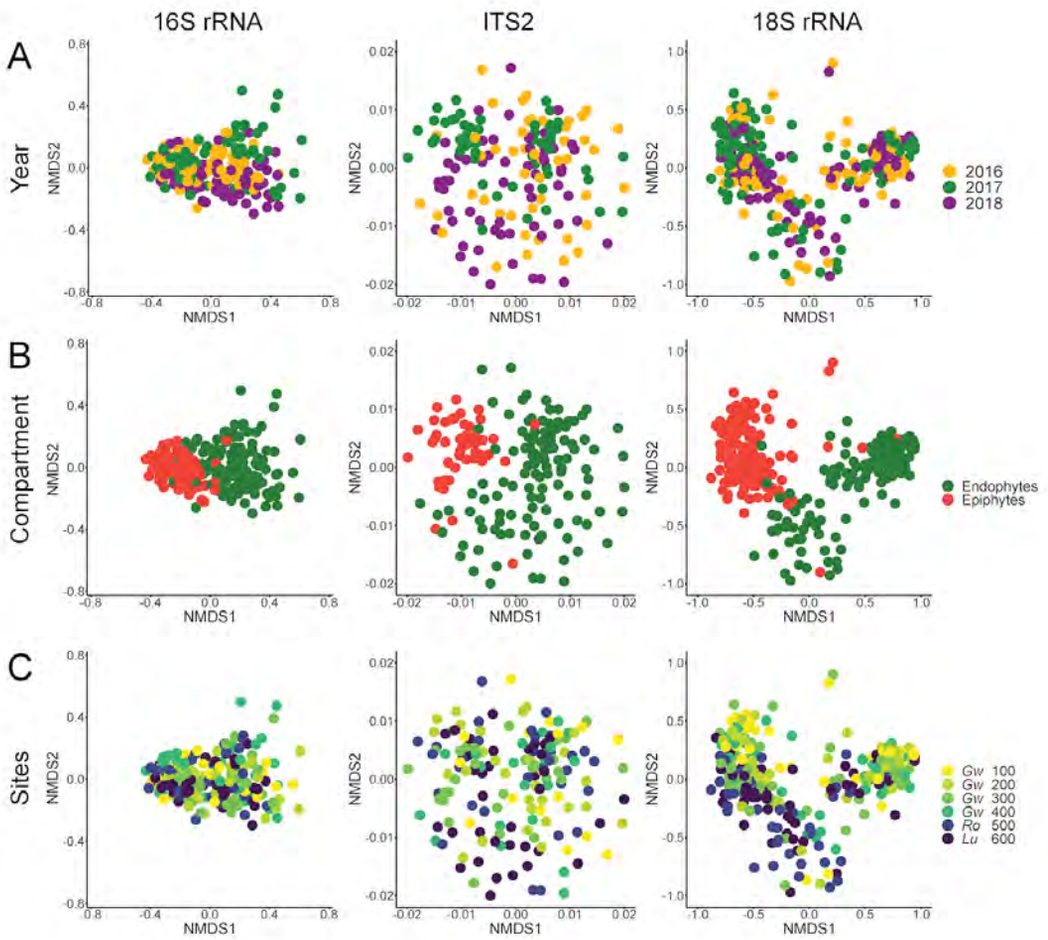


S-Fig. 8: Beta-dispersion analysis on Bray-Curtis dissimilarities to compare sample-to-sample variation in *RPP1* haplotypes for 18S rRNA amplicons using simultaneously DNA-seq (left) and cDNA-seq (right). Samples are grouped by A) Compartment x *RPP1*-like allele and B) Location x Compartment x *RPP1*-like allele. Statistics: Dunn test. Circle diameter reflect p-values.

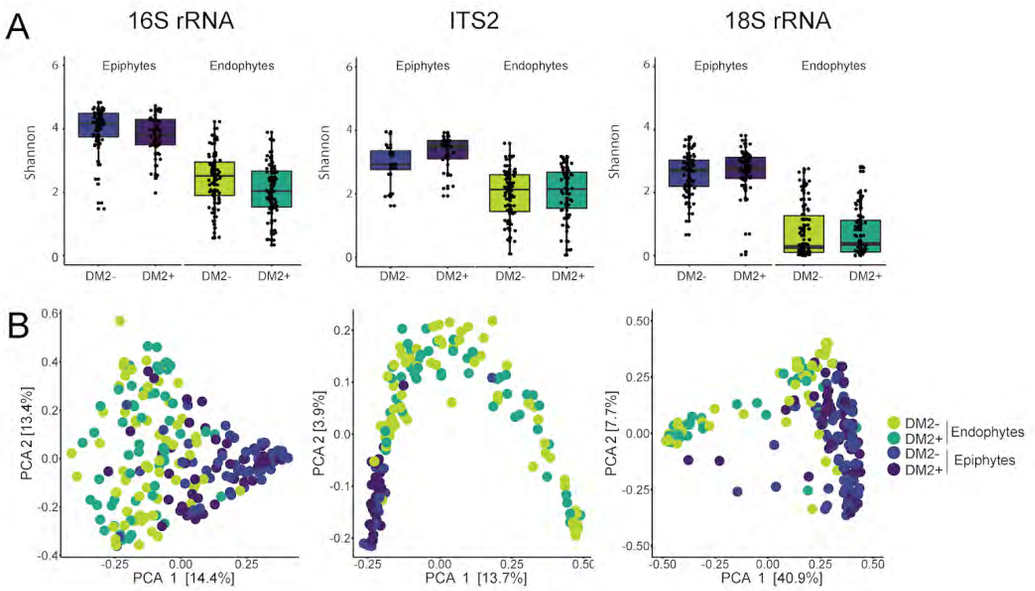
5.10 DNA amplicon sequencing



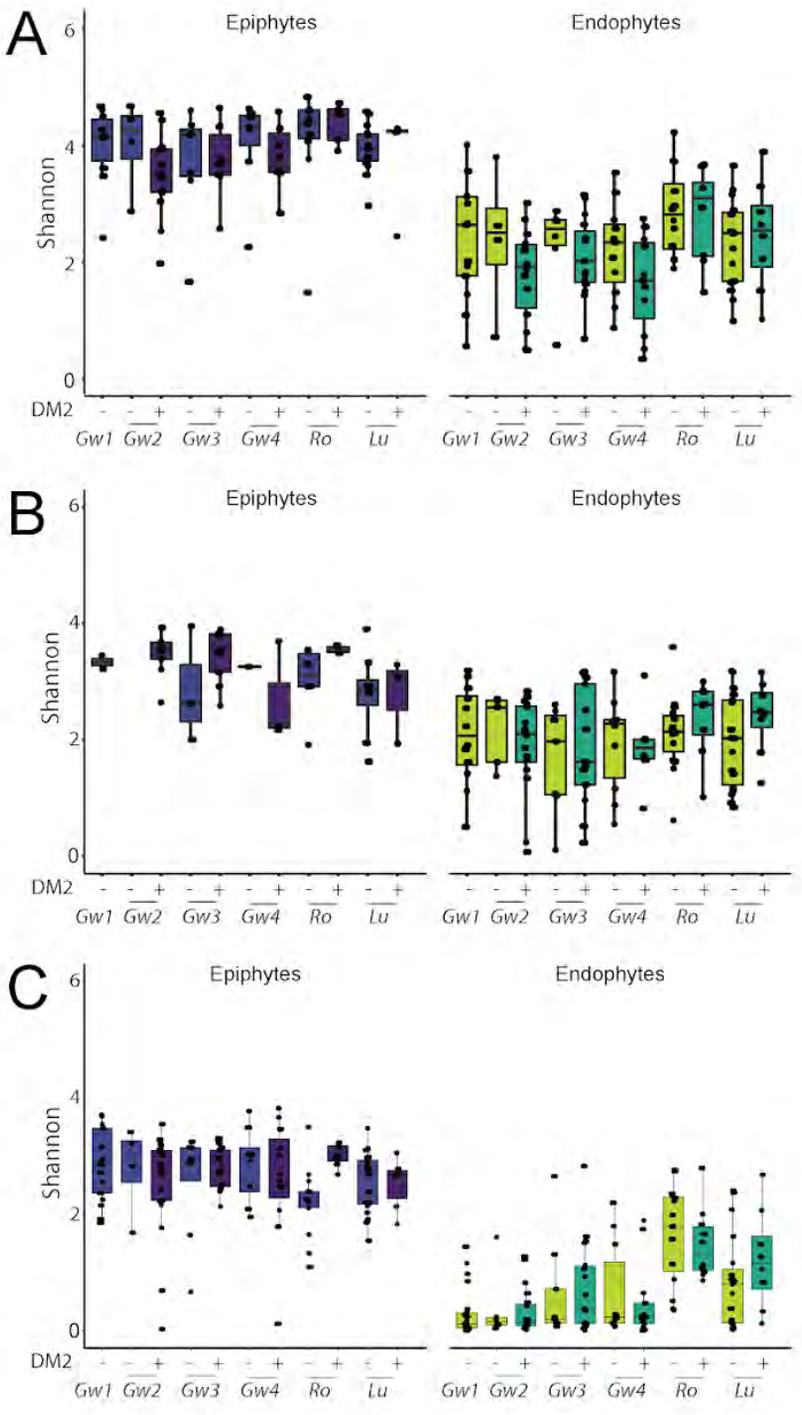
S-Fig. 9: Microbial richness and taxonomic composition of classical DNA-sequencing. A) bacteria B) fungi and C) eukaryotes. Kruskal-Wallis test $p < 0.05$.



S-Fig. 10: Multivariate analysis of (DNA) microbiome samples considering year, compartment and geographical location (sites). Non-metric multidimensional scaling plots on Bray-Curtis dissimilarities displaying NMDS1 and NMDS2 representing the first two axis of the two-dimensional ordination. Plots are from left to right: 16S rRNA (bacteria), ITS2 (fungi) and 18S rRNA (eukaryotes). Microbiome features are displayed for sampling year (A), leaf compartments (B) and geographical locations (C).

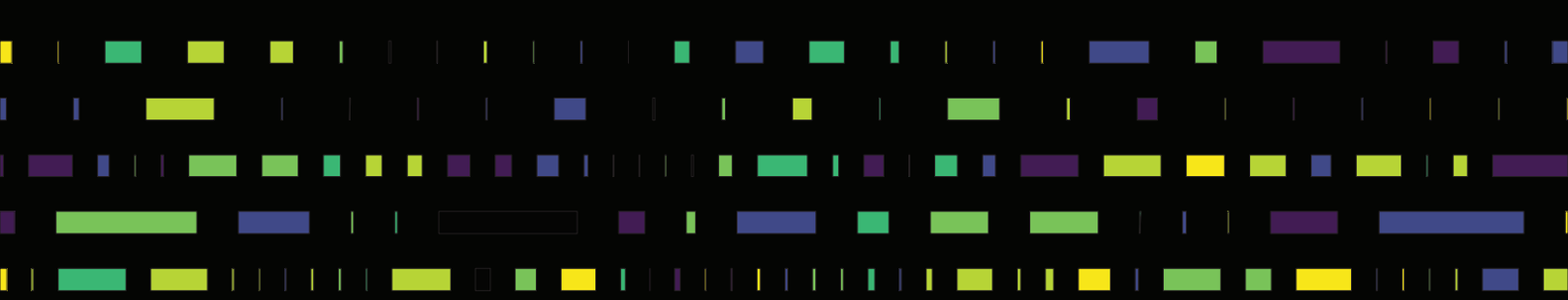


S-Fig. 11: Microbial diversities of wild *Arabidopsis* microbiome samples from six geographical locations around Gorzów Wielkopolski across three consecutive years. Considering amplicons of 16S rRNA, ITS2 and 18S rRNA (left to right). Samples are grouped by [Compartment x *RPP1*]. A) Shannon indices were calculated for sample groups. B) Principle component analysis (PCA) of taxonomic composition according to sample groups. The explained variance relies on the first two ordinations PCA1 and PCA2 within the ordination system.



S-Fig. 12: Microbial composition of sample groups [Compartment x *RPP1*] remains stable across geographical locations.

References



6 References

1001 Genomes Consortium. 2016. "1,135 Genomes Reveal the Global Pattern of Polymorphism in Arabidopsis Thaliana." *Cell* 166 (2): 481–91.

Aarts, N., M. Metz, E. Holub, B. J. Staskawicz, M. J. Daniels, and J. E. Parker. 1998. "Different Requirements for EDS1 and NDR1 by Disease Resistance Genes Define at Least Two R Gene-Mediated Signaling Pathways in Arabidopsis." *Proceedings of the National Academy of Sciences of the United States of America* 95 (17): 10306–11.

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.

Al Ashhab, Ashraf, Shiri Meshner, Rivka Alexander-Shani, Hana Dimerets, Michael Brandwein, Yael Bar-Lavan, and Gidon Winters. 2021. "Temporal and Spatial Changes in Phyllosphere Microbiome of Acacia Trees Growing in Arid Environments." *Frontiers in Microbiology* 12 (July): 656269.

Alcázar, Rubén, Ana V. García, Jane E. Parker, and Matthieu Reymond. 2009. "Incremental Steps toward Incompatibility Revealed by Arabidopsis Epistatic Interactions Modulating Salicylic Acid Pathway Activation." *Proceedings of the National Academy of Sciences of the United States of America* 106 (1): 334–39.

Alcázar, Rubén, Marcel von Reth, Jaqueline Bautor, Eunyoung Chae, Detlef Weigel, Maarten Koornneef, and Jane E. Parker. 2014. "Analysis of a Plant Complex Resistance Gene Locus Underlying Immune-Related Hybrid Incompatibility and Its Occurrence in Nature." *PLoS Genetics* 10 (12): 1–14.

Alfonso, Esteban, Elia Stahl, Gaétan Glauser, Etienne Bellani, Tom M. Raaymakers, Guido Van den Ackerveken, Jürgen Zeier, and Philippe Reymond. 2021. "Insect Eggs Trigger Systemic Acquired Resistance against a Fungal and an Oomycete Pathogen." *The New Phytologist*, September. https://doi.org/10.1111/nph.17732.

Almario, Juliana, Maryam Mahmudi, Samuel Kroll, Mathew Agler, Aleksandra Placzek, Alfredo Mari, and Eric Kemen. 2021. "The Leaf Microbiota of Arabidopsis Displays Reproducible Dynamics and Patterns throughout the Growing Season." *Research Square*. https://doi.org/10.21203/rs.3.rs-808966/v1.

Alonge, Michael, Sebastian Soyk, Srividya Ramakrishnan, Xingang Wang, Sara Goodwin, Fritz J. Sedlazeck, Zachary B. Lippman, and Michael C. Schatz. 2019. "RaGOO: Fast and Accurate Reference-Guided Scaffolding of Draft Genomes." *Genome Biology* 20 (1): 224.

Andolfo, Giuseppe, Antimo Di Donato, Pasquale Chiaiese, Antonino De Natale, Antonino Pollio, Jonathan D. G. Jones, Luigi Frusciante, and Maria Raffaella Ercolano. 2019. "Alien Domains Shaped the Modular Structure of Plant NLR Proteins." *Genome Biology and Evolution* 11 (12): 3466–77.

Arabidopsis Genome Initiative. 2000. "Analysis of the Genome Sequence of the Flowering Plant Arabidopsis Thaliana." *Nature* 408 (6814): 796–815.

Arany, A. Mosleh, T. J. de Jong, and E. van der Meijden. 2005. "Herbivory and Abiotic Factors Affect Population Dynamics of Arabidopsis Thaliana in a Sand Dune Area." *Plant Biology (Stuttgart, Germany)* 7 (5): 549–56.

Asalone, Kathryn C., Kara M. Ryan, Maryam Yamadi, Annastelle L. Cohen, William G. Farmer, Deborah J. George, Claudia Joppert, et al. 2020. "Regional Sequence Expansion or Collapse in Heterozygous Genome Assemblies." *PLoS Computational Biology* 16 (7): e1008104.

Astudillo-García, Carmen, James J. Bell, Nicole S. Webster, Bettina Glasl, Jamaluddin Jompa, Jose M. Montoya, and Michael W. Taylor. 2017. "Evaluating the Core Microbiota in Complex Communities: A Systematic Investigation." *Environmental Microbiology* 19 (4): 1450–62.

Atanasov, Kostadin E., Changxin Liu, Alexander Erban, Joachim Kopka, and Jane E. Parker. 2014. "NLR Mutations Suppressing Immune Hybrid Incompatibility and Their Effects on Disease Resistance."

Babalola, Olubukola Oluranti, Ayomide E. Fadiji, Ben J. Enagbonma, Elizabeth T. Alori, Modupe S. Ayilara, and Ayansina S. Ayangbenro. 2020. "The Nexus between Plant and Plant Microbiome: Revelation of the Networking Strategies." *Frontiers in Microbiology* 11 (September). https://doi.org/10.3389/fmicb.2020.548037.

Bai, Yuling, Christos Kissoudis, Zhe Yan, Richard G. F. Visser, and Gerard van der Linden. 2018. "Plant Behaviour under Combined Stress: Tomato Responses to Combined Salinity and Pathogen Stress." *The Plant Journal: For Cell and Molecular Biology* 93 (4): 781–93.

Bailly, Aurélien, and Laure Weisskopf. 2017. "Mining the Volatilomes of Plant-Associated Microbiota for New Biocontrol Solutions." *Frontiers in Microbiology* 8 (August): 1638.

Bálint, Miklós, Peter Tiffin, Björn Hallström, Robert B. O'Hara, Matthew S. Olson, Johnathon D. Fankhauser, Meike Piepenbring, and Imke Schmitt. 2013. "Host Genotype Shapes the Foliar Fungal Microbiome of Balsam Poplar (Populus Balsamifera)." *PloS One* 8 (1). https://doi.org/10.1371/journal.pone.0053987.

Barragan, A. Cristina, and Detlef Weigel. 2021. "Plant NLR Diversity: The Known Unknowns of Pan-NLRomes." *The Plant Cell* 33 (4): 814–31.

Bartoli, Claudia, and Fabrice Roux. 2017. "Genome-Wide Association Studies in Plant Pathosystems: Toward an Ecological Genomics Approach." *Frontiers in Plant Science* 8 (May). https://doi.org/10.3389/fpls.2017.00763.

Batstone, Rebecca T., Anna M. O'Brien, Tia L. Harrison, and Megan E. Frederickson. 2020. "Experimental Evolution Makes Microbes More Cooperative with Their Local Host Genotype." *Science* 370 (6515): 23–26.

Beck, James B., Heike Schmuths, and Barbara A. Schaal. 2008. "Native Range Genetic Variation in Arabidopsis Thaliana Is Strongly Geographically Structured and Reflects Pleistocene Glacial Dynamics." *Molecular Ecology* 17 (3): 902–15.

Beilsmith, Kathleen, Manus P. M. Thoen, Benjamin Brachi, Andrew D. Gloss, Mohammad H. Khan, and Joy Bergelson. 2019. "Genome-Wide Association Studies on the Phyllosphere Microbiome: Embracing Complexity in Host–Microbe Interactions." *The Plant Journal: For Cell and Molecular Biology* 97 (1): 164–81.

Berens, Matthias L., Katarzyna W. Wolinska, Stijn Spaepen, Jörg Ziegler, Tatsuya Nobori, Aswin Nair, Verena Krüler, et al. 2019. "Balancing Trade-Offs between Biotic and Abiotic Stress Responses through Leaf Age-Dependent Variation in Stress Hormone Cross-Talk." *Proceedings of the National Academy of Sciences of the United States of America* 116 (6): 2364–73.

Berlin, Konstantin, Sergey Koren, Chen-Shan Chin, James P. Drake, Jane M. Landolin, and Adam M. Phillippy. 2015. "Assembling Large Genomes with Single-Molecule Sequencing and Locality-Sensitive Hashing." *Nature Biotechnology* 33 (6): 623–30.

Bevan, M., I. Bancroft, E. Bent, K. Love, H. Goodman, C. Dean, R. Bergkamp, et al. 1998. "Analysis of 1.9 Mb of Contiguous Sequence from Chromosome 4 of Arabidopsis Thaliana." *Nature* 391 (6666): 485–88.

Bi, Guozhi, and Jian-Min Zhou. 2017. "MAP Kinase Signaling Pathways: A Hub of Plant-Microbe Interactions." *Cell Host & Microbe* 21 (3): 270–73.

Bigeard, Jean, Jean Colcombet, and Heribert Hirt. 2015. "Signaling Mechanisms in Pattern-Triggered Immunity (PTI)." *Molecular Plant* 8 (4): 521–39.

Bodenhausen, Natacha, Miriam Bortfeld-Miller, Martin Ackermann, and Julia A. Vorholt. 2014. "A Synthetic Community Approach Reveals Plant Genotypes Affecting the Phyllosphere Microbiota." *PLoS Genetics* 10 (4). https://doi.org/10.1371/journal.pgen.1004283.

Bodenhausen, Natacha, Matthew W. Horton, and Joy Bergelson. 2013. "Bacterial Communities Associated with the Leaves and the Roots of Arabidopsis Thaliana." *PloS One* 8 (2). https://doi.org/10.1371/journal.pone.0056329.

Bolyen, Evan, Jai Ram Rideout, Matthew R. Dillon, Nicholas A. Bokulich, Christian C. Abnet, Gabriel A. Al-Ghalith, Harriet Alexander, et al. 2019. "Reproducible, Interactive, Scalable and Extensible Microbiome Data Science Using QIIME 2." *Nature Biotechnology* 37 (8): 852–57.

Bomblies, Kirsten, Janne Lempe, Petra Epple, Norman Warthmann, Christa Lanz, Jeffery L. Dangl, and Detlef Weigel. 2007. "Autoimmune Response as a Mechanism for a Dobzhansky-Muller-Type Incompatibility Syndrome in Plants." *PLoS Biology* 5 (9): 1962–72.

Bomblies, Kirsten, Levi Yant, Roosa A. Laitinen, Sang-Tae Kim, Jesse D. Hollister, Norman Warthmann, Joffrey Fitz, and Detlef Weigel. 2010. "Local-Scale Patterns of Genetic Variability, Outcrossing, and Spatial Structure in Natural Stands of Arabidopsis Thaliana." *PLoS Genetics* 6 (3): e1000890.

Borhan, M. Hossein, Nick Gunn, Abigail Cooper, Sigrun Gulden, Mahmut Tör, S. Roger Rimmer, and Eric B. Holub. 2008. "WRR4 Encodes a TIR-NB-LRR Protein That Confers Broad-Spectrum White Rust Resistance in Arabidopsis Thaliana to Four Physiological Races of Albugo Candida." *Molecular Plant-Microbe Interactions: MPMI* 21 (6): 757–68.

Boutrot, Freddy, and Cyril Zipfel. 2017. "Function, Discovery, and Exploitation of Plant Pattern Recognition Receptors for Broad-Spectrum Disease Resistance." *Annual Review of Phytopathology* 55 (1): 257–86.

Brader, Günter, Stéphane Compant, Kathryn Vescio, Birgit Mitter, Friederike Trognitz, Li Jun Ma, and Angela Sessitsch. 2017. "Ecology and Genomic Insights into Plant-Pathogenic and Plant-Nonpathogenic Endophytes." *Annual Review of Phytopathology* 55: 61–83.

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.

Bruisson, Sébastien, Mónica Zufferey, Floriane L'Haridon, Eva Trutmann, Abhishek Anand, Agnès Dutartre, Mout De Vrieze, and Laure Weisskopf. 2019. "Endophytes and Epiphytes from the Grapevine Leaf Microbiome as Potential Biocontrol Agents against Phytopathogens." *Frontiers in Microbiology* 10 (November): 2726.

Brüssow, Harald. 2020. "Problems with the Concept of Gut Microbiota Dysbiosis." *Microbial Biotechnology* 13 (2): 423–34.

Buchfink, Benjamin, Chao Xie, and Daniel H. Huson. 2014. "Fast and Sensitive Protein Alignment Using DIAMOND." *Nature Methods* 12 (1): 59–60.

Bulgarelli, Davide, Klaus Schlaeppi, Stijn Spaepen, Emiel Ver Loren van Themaat, and Paul Schulze-Lefert. 2013. "Structure and Functions of the Bacterial Microbiota of Plants." *Annual Review of Plant Biology* 64: 807–38.

Burdett, Hayden, Adam R. Bentham, Simon J. Williams, Peter N. Dodds, Peter A. Anderson, Mark J. Banfield, and Bostjan Kobe. 2019. "The Plant 'Resistosome': Structural Insights into Immune Signaling." *Cell Host & Microbe* 26 (2): 193–201.

Busse, Hans-Jürgen, Ewald B. M. Denner, Sandra Buczolits, Mirja Salkinoja-Salonen, Antonio Bennasar, and Peter Kämpfer. 2003. "Sphingomonas Aurantiaca Sp. Nov., Sphingomonas Aerolata Sp. Nov. and Sphingomonas Faeni Sp. Nov., Air- and Dustborne and Antarctic, Orange-Pigmented, Psychrotolerant Bacteria, and Emended Description of the Genus Sphingomonas." *International Journal of Systematic and Evolutionary Microbiology* 53 (Pt 5): 1253–60.

Buswell, Will, Roland E. Schwarzenbacher, Estrella Luna, Matthew Sellwood, Beining Chen, Victor Flors, Pierre Pétriacq, and Jurriaan Ton. 2018. "Chemical Priming of Immunity without Costs to Plant Growth." *The New Phytologist* 218 (3): 1205–16.

Castilla, Antonio R., Belén Méndez-Vigo, Arnald Marcer, Joaquín Martínez-Minaya, David Conesa, F. Xavier Picó, and Carlos Alonso-Blanco. 2020. "Ecological, Genetic and Evolutionary Drivers of Regional Genetic Differentiation in Arabidopsis Thaliana." *BMC Evolutionary Biology* 20 (1): 71.

Cavalier-Smith, Tom. 1992. "Symbiosis as a Source of Evolutionary Innovation: Speciation and Morphogenesis." *Trends in Ecology & Evolution* 7 (12): 422–23.

Cawoy, Hélène, Martin Mariutto, Guillaume Henry, Christophe Fisher, Natallia Vasilyeva, Philippe Thonart, Jacques Dommes, and Marc Ongena. 2014. "Plant Defense Stimulation by Natural Isolates of Bacillus Depends on Efficient Surfactin Production." *Molecular Plant-Microbe Interactions: MPMI* 27 (2): 87–100.

Century, K. S., E. B. Holub, and B. J. Staskawicz. 1995. "NDR1, a Locus of Arabidopsis Thaliana That Is Required for Disease Resistance to Both a Bacterial and a Fungal Pathogen." *Proceedings of the National Academy of Sciences of the United States of America* 92 (14): 6597–6601.

Cesari, Stella, Maud Bernoux, Philippe Moncuquet, Thomas Kroj, and Peter N. Dodds. 2014. "A Novel Conserved Mechanism for Plant NLR Protein Pairs: The Â€œintegrated Decoyâ€ Hypothesis." *Frontiers in Plant Science* 5 (November): 606.

Césari, Stella, Hiroyuki Kanzaki, Tadashi Fujiwara, Maud Bernoux, Véronique Chalvon, Yoji Kawano, Ko Shimamoto, Peter Dodds, Ryohei Terauchi, and Thomas Kroj. 2014. "The NB-LRR Proteins RGA4 and RGA5 Interact Functionally and Physically to Confer Disease Resistance." *The EMBO Journal* 33 (17): 1941–59.

Cevik, Volkan, Freddy Boutrot, Wiebke Apel, Alexandre Robert-Seilaniantz, Oliver J. Furzer, Amey Redkar, Baptiste Castel, et al. 2019. "Transgressive Segregation Reveals Mechanisms of Arabidopsis Immunity to Brassica-Infecting Races of White Rust (Albugo Candida)." *Proceedings of the National Academy of Sciences of the United States of America* 116 (7): 2767–73.

Chae, Eunyoung, Kirsten Bomblies, Sang-Tae Kim, Darya Karelina, Maricris Zaidem, Stephan Ossowski, Carmen Martín-Pizarro, et al. 2014. "Species-Wide Genetic Incompatibility Analysis Identifies Immune Genes as Hot Spots of Deleterious Epistasis." *Cell* 159 (6): 1341–51.

Chahtane, Hicham, Thanise Nogueira Füller, Pierre-Marie Allard, Laurence Marcourt, Emerson Ferreira Queiroz, Venkatasalam Shanmugabalaji, Jacques Falquet, Jean-Luc Wolfender, and Luis Lopez-Molina. 2018. "The Plant Pathogen Pseudomonas Aeruginosa Triggers a DELLA-Dependent Seed Germination Arrest in Arabidopsis." *ELife* 7 (August). https://doi.org/10.7554/eLife.37082.

Chan, P. P., B. Y. Lin, A. J. Mak, and T. M. Lowe. 2021. "TRNAscan-SE 2.0: Improved Detection and Functional Classification of Transfer RNA Genes." *BioRxiv*. https://www.biorxiv.org/content/10.1101/614032v2.abstract.

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

Chen, Jiapeng, Narayana M. Upadhyaya, Diana Ortiz, Jana Sperschneider, Feng Li, Clement Bouton, Susan Breen, et al. 2017. "Loss of AvrSr50 by Somatic Exchange in Stem Rust Leads to Virulence for Sr50 Resistance in Wheat." *Science (New York, N.Y.)* 358 (6370): 1607–10.

Chen, Liang, Justin E. Wilson, Mark J. Koenigsknecht, Wei-Chun Chou, Stephanie A. Montgomery, Agnieszka D. Truax, W. June Brickey, et al. 2017. "Erratum: NLRP12 Attenuates Colon Inflammation by Maintaining Colonic Microbial Diversity and Promoting Protective Commensal Bacterial Growth." *Nature Immunology* 18 (8): 951.

Chen, Pan, Chunmei Zhang, Xiuyun Ju, Youwei Xiong, Ke Xing, and Sheng Qin. 2019. "Community Composition and Metabolic Potential of Endophytic Actinobacteria from Coastal Salt Marsh Plants in Jiangsu, China." *Frontiers in Microbiology* 10 (May): 1063.

Chen, Tao, Kinya Nomura, Xiaolin Wang, Reza Sohrabi, Jin Xu, Lingya Yao, Bradley C. Paasch, et al. 2020. "A Plant Genetic Network for Preventing Dysbiosis in the Phyllosphere." *Nature* 580 (7805): 653–57.

Chen, Xiaoyulong, Wisnu Adi Wicaksono, Gabriele Berg, and Tomislav Cernava. 2021. "Bacterial Communities in the Plant Phyllosphere Harbour Distinct Responders to a Broad-Spectrum Pesticide." *The Science of the Total Environment* 751 (141799): 141799.

Cheng, Chia Yi, Vivek Krishnakumar, Agnes P. Chan, Françoise Thibaud-Nissen, Seth Schobel, and Christopher D. Town. 2017. "Araport11: A Complete Reannotation of the Arabidopsis Thaliana Reference Genome." *The Plant Journal: For Cell and Molecular Biology* 89 (4): 789–804.

Cheng, Yu Ti, Li Zhang, and Sheng Yang He. 2019. "Plant-Microbe Interactions Facing Environmental Challenge." *Cell Host & Microbe* 26 (2): 183–92.

Cheng, Zhiqiang, Shaonan Lei, Ye Li, Wei Huang, Rongqin Ma, Juan Xiong, and Ting Zhang. 2020. "Revealing the Variation and Stability of Bacterial Communities in Tomato Rhizosphere Microbiota," 1–15.

Ching-Su, L. I. N., L. I. N. Yu-Hsin, and W. U. Jiunn-Tzong. 2012. "Biodiversity of the Epiphyllous Algae in a Chamaecyparis Forest of Northern Taiwan." 2012. https://ejournal.sinica.edu.tw/bbas/content/2012/4/Bot534-08.pdf.

Christopoulou, Marilena, Sebastian Reyes-Chin Wo, Alex Kozik, Leah K. McHale, Maria-Jose Truco, Tadeusz Wroblewski, and Richard W. Michelmore. 2015. "Genome-Wide Architecture of Disease Resistance Genes in Lettuce." *G3 (Bethesda, Md.)* 5 (12): 2655–69.

Chu, Thanh Nguyen, Bao Thi Hoai Tran, Le Van Bui, and Minh Thi Thanh Hoang. 2019. "Plant Growth-Promoting Rhizobacterium Pseudomonas PS01 Induces Salt Tolerance in Arabidopsis Thaliana." *BMC Research Notes* 12 (1): 11.

Clouse, Kayla M., and Maggie R. Wagner. 2021. "Plant Genetics as a Tool for Manipulating Crop Microbiomes: Opportunities and Challenges." *Frontiers in Bioengineering and Biotechnology* 9 (May): 567548.

Conant, Gavin C., and Kenneth H. Wolfe. 2008. “Turning a Hobby into a Job: How Duplicated Genes Find New Functions.” *Nature Reviews. Genetics* 9 (12): 938–50.

Copeland, Julia K., Lijie Yuan, Mehdi Layeghifard, Pauline W. Wang, and David S. Guttman. 2015. “Seasonal Community Succession of the Phyllosphere Microbiome.” *Molecular Plant-Microbe Interactions PMI* 274 (3): 274–85.

Cui, Haitao, Kenichi Tsuda, and Jane E. Parker. 2015. “Effector-Triggered Immunity: From Pathogen Perception to Robust Defense.” *Annual Review of Plant Biology* 66 (October 2015): 487–511.

Dangl, J. L., and J. D. Jones. 2001. “Plant Pathogens and Integrated Defence Responses to Infection.” *Nature* 411 (6839): 826–33.

Dangl, Jeffery L., Diana M. Horvath, and Brian J. Staskawicz. 2013. “Pivoting the Plant Immune System from Dissection to Deployment.” *Science (New York, N.Y.)* 341 (6147): 746–51.

Darienko, Tatyana, Lydia Gustavs, Opayi Mudimu, Cecilia Rad Menendez, Rhena Schumann, Ulf Karsten, Thomas Friedl, and Thomas Pröschold. 2010. “Chloroidium, a Common Terrestrial Coccoid Green Alga Previously Assigned ToChlorella(Trebouxiophyceae, Chlorophyta).” *European Journal of Phycology* 45 (1): 79–95.

De Coster, Wouter, Matthias H. Weissensteiner, and Fritz J. Sedlazeck. 2021. “Towards Population-Scale Long-Read Sequencing.” *Nature Reviews. Genetics* 22 (9): 572–87.

De Kort, H., J. G. Prunier, S. Ducatez, O. Honnay, M. Baguette, V. M. Stevens, and S. Blanchet. 2021. “Life History, Climate and Biogeography Interactively Affect Worldwide Genetic Diversity of Plant and Animal Populations.” *Nature Communications* 12 (1): 516.

De Vrieze, Mout, Fanny Germanier, Nicolas Vuille, and Laure Weisskopf. 2018. “Combining Different Potato-Associated Pseudomonas Strains for Improved Biocontrol of Phytophthora Infestans.” *Frontiers in Microbiology* 9 (October): 2573.

Deng, Siwen, Daniel F. Caddell, Gen Xu, Lindsay Dahlen, Lorenzo Washington, Jinliang Yang, and Devin Coleman-Derr. 2021. “Genome Wide Association Study Reveals Plant Loci Controlling Heritability of the Rhizosphere Microbiome.” *The ISME Journal* 15 (11): 3181–94.

DeSantis, T. Z., P. Hugenholtz, N. Larsen, M. Rojas, E. L. Brodie, K. Keller, T. Huber, D. Dalevi, P. Hu, and G. L. Andersen. 2006. “Greengenes, a Chimera-Checked 16S RRNA Gene Database and Workbench Compatible with ARB.” *Applied and Environmental Microbiology* 72 (7): 5069–72.

Desrut, Antoine, Bouziane Moumen, Florence Thibault, Rozenn Le Hir, Pierre Coutos-Thévenot, and Cécile Vriet. 2020. “Beneficial Rhizobacteria Pseudomonas Simiae WCS417 Induce Major Transcriptional Changes in Plant Sugar Transport.” *Journal of Experimental Botany* 71 (22): 7301–15.

Dodds, Peter N., Gregory J. Lawrence, Ann-Maree Catanzariti, Trazel Teh, Ching-I A. Wang, Michael A. Ayliffe, Bostjan Kobe, and Jeffrey G. Ellis. 2006. “Direct Protein Interaction Underlies Gene-for-Gene Specificity and Coevolution of the Flax Resistance Genes and Flax Rust Avirulence Genes.” *Proceedings of the National Academy of Sciences of the United States of America* 103 (23): 8888–93.

Dorofeeva, Lubov V., Lyudmila I. Evtushenko, Valentina I. Krausova, Alexander V. Karpov, Sergey A. Subbotin, and James M. Tiedje. 2002. “Rathayibacter Caricis Sp. Nov. and Rathayibacter Festucae Sp. Nov., Isolated from the Phyllosphere of Carex Sp. and the Leaf Gall Induced by the Nematode Anguina Graminis on Festuca Rubra L., Respectively.” *International Journal of Systematic and Evolutionary Microbiology* 52 (Pt 6): 1917–23.

Dudeja, Surjit S., Pooja Suneja-Madan, Minakshi Paul, Rajat Maheswari, and Erika Kothe. 2021. “Bacterial Endophytes: Molecular Interactions with Their Hosts.” *Journal of Basic Microbiology*, no. vember 2020. https://doi.org/10.1002/jobm.202000657.

Dyachkova, Marina S., Evgeny V. Chekalin, and Valery N. Danilenko. 2019. “Positive Selection in Bifidobacterium Genes Drives Species-Specific Host-Bacteria Communication.” *Frontiers in Microbiology* 10 (October): 2374.

Edgar, Robert C., Brian J. Haas, Jose C. Clemente, Christopher Quince, and Rob Knight. 2011. “UCHIME Improves Sensitivity and Speed of Chimera Detection.” *Bioinformatics* 27 (16): 2194–2200.

Edwards, Joseph A., Christian M. Santos-Medellín, Zachary S. Liechty, Bao Nguyen, Eugene Lurie, Shane Eason, Gregory Phillips, and Venkatesan Sundaresan. 2018. “Compositional Shifts in Root-Associated Bacterial and Archaeal Microbiota Track the Plant Life Cycle in Field-Grown Rice.” *PLoS Biology* 16 (2): e2003862.

Eitas, Timothy K., and Jeffery L. Dangl. 2010. "NB-LRR Proteins: Pairs, Pieces, Perception, Partners, and Pathways." *Current Opinion in Plant Biology* 13 (4): 472–77.

Exposito-Alonso, Moises, 500 Genomes Field Experiment Team, Hernán A. Burbano, Oliver Bossdorf, Rasmus Nielsen, and Detlef Weigel. 2019. "Natural Selection on the Arabidopsis Thaliana Genome in Present and Future Climates." *Nature* 573 (7772): 126–29.

Feehan, Joanna M., Baptiste Castel, Adam R. Bentham, and Jonathan D. G. Jones. 2020. "Plant NLRs Get by with a Little Help from Their Friends." *Current Opinion in Plant Biology* 56 (Cc): 99–108.

Feenstra, W. J. 1964. "Isolation of Nutritional Mutants in Arabidopsis Thaliana." *Genetica* 35 (4): 259–69.

Feys, Bart J., Marcel Wiermer, Riyaz A. Bhat, Lisa J. Moisan, Nieves Medina-Escobar, Christina Neu, Adriana Cabral, and Jane E. Parker. 2005. "Arabidopsis SENESCENCE-ASSOCIATED GENE101 Stabilizes and Signals within an ENHANCED DISEASE SUSCEPTIBILITY1 Complex in Plant Innate Immunity." *The Plant Cell* 17 (9): 2601–13.

Finkel, Omri M., Isai Salas-González, Gabriel Castrillo, Stijn Spaepen, Theresa F. Law, Paulo José Pereira Lima Teixeira, Corbin D. Jones, and Jeffery L. Dangl. 2019. *The Effects of Soil Phosphorous Content on Microbiota Are Driven by the Plant Phosphate Starvation Response*.

Firrincieli, Andrea, Mahsa Khorasani, A. Carolin Frank, and Sharon Lafferty Doty. 2020. "Influences of Climate on Phyllosphere Endophytic Bacterial Communities of Wild Poplar." *Frontiers in Plant Science* 11 (February): 1–13.

Florea, Liliana, Alexander Souvorov, Theodore S. Kalbfleisch, and Steven L. Salzberg. 2011. "Genome Assembly Has a Major Impact on Gene Content: A Comparison of Annotation in Two Bos Taurus Assemblies." *PloS One* 6 (6): e21400.

François, Olivier, Michael G. B. Blum, Mattias Jakobsson, and Noah A. Rosenberg. 2008. "Demographic History of European Populations of Arabidopsis Thaliana." *PLoS Genetics* 4 (5): e1000075.

Frank, A. Carolin. 2019. "Molecular Host Mimicry and Manipulation in Bacterial Symbionts." *FEMS Microbiology Letters* 366 (4). https://doi.org/10.1093/femsle/fnz038.

Gallego, Virginia, Maria Teresa García, and Antonio Ventosa. 2006. "Methylobacterium Adhaesivum Sp. Nov., a Methylotrophic Bacterium Isolated from Drinking Water." *International Journal of Systematic and Evolutionary Microbiology* 56 (2): 339–42.

Gao, Min, Chao Xiong, Cheng Gao, Clement K. M. Tsui, Meng-Meng Wang, Xin Zhou, Ai-Min Zhang, and Lei Cai. 2021. "Disease-Induced Changes in Plant Microbiome Assembly and Functional Adaptation." *Microbiome* 9 (1): 187.

Gel, Bernat, and Eduard Serra. 2017. "KaryoploteR: An R/Bioconductor Package to Plot Customizable Genomes Displaying Arbitrary Data." *Bioinformatics* 33 (19): 3088–90.

Gentzbittel, Laurent, Stig U. Andersen, Cécile Ben, Martina Rickauer, Jens Stougaard, and Nevin D. Young. 2015. "Naturally Occurring Diversity Helps to Reveal Genes of Adaptive Importance in Legumes." *Frontiers in Plant Science* 6 (April): 269.

Gilchrist, Cameron L. M., and Yit-Heng Chooi. 2021. "Clinker & Clustermap.Js: Automatic Generation of Gene Cluster Comparison Figures." *Bioinformatics* , January. https://doi.org/10.1093/bioinformatics/btab007.

Giraut, Laurène, Matthieu Falque, Jan Drouaud, Lucie Pereira, Olivier C. Martin, and Christine Mézard. 2011. "Genome-Wide Crossover Distribution in Arabidopsis Thaliana Meiosis Reveals Sex-Specific Patterns along Chromosomes." *PLoS Genetics* 7 (11): e1002354.

Goel, Manish, Hequan Sun, Wen-Biao Jiao, and Korbinian Schneeberger. 2019. "SyRI: Finding Genomic Rearrangements and Local Sequence Differences from Whole-Genome Assemblies." *Genome Biology* 20 (1): 277.

Gomaa, Nasr H., Alicia Montesinos-Navarro, Carlos Alonso-Blanco, and F. Xavier Picó. 2011. "Temporal Variation in Genetic Diversity and Effective Population Size of Mediterranean and Subalpine Arabidopsis Thaliana Populations." *Molecular Ecology* 20 (17): 3540–54.

Gomes, Teresa, José Alberto Pereira, Jacinto Benhadi, Teresa Lino-Neto, and Paula Baptista. 2018. "Endophytic and Epiphytic Phyllosphere Fungal Communities Are Shaped by Different Environmental Factors in a Mediterranean Ecosystem." *Microbial Ecology* 76 (3): 668–79.

Gong, Tianyu, and Xiu Fang Xin. 2021. “Phyllosphere Microbiota: Community Dynamics and Its Interaction with Plant Hosts.” *Journal of Integrative Plant Biology* 63 (2): 297–304.

Goritschnig, S., A. D. Steinbrenner, D. J. Grunwald, and B. J. Staskawicz. 2016. “Structurally Distinct Arabidopsis Thaliana NLR Immune Receptors Recognize Tandem WY Domains of an Oomycete Effector.” *The New Phytologist* 210 (3): 984–96.

Goss, Erica M., Martin Kreitman, and Joy Bergelson. 2005. “Genetic Diversity, Recombination and Cryptic Clades in Pseudomonas Viridiflava Infecting Natural Populations of Arabidopsis Thaliana.” *Genetics* 169 (1): 21–35.

Goss-Souza, Dennis, Lucas William Mendes, Clovis Daniel Borges, Dilmar Baretta, Siu Mui Tsai, and Jorge L. M. Rodrigues. 2017. “Soil Microbial Community Dynamics and Assembly under Long-Term Land Use Change.” *FEMS Microbiology Ecology* 93 (10). https://doi.org/10.1093/femsec/fix109.

Grime, J. P. 1977. “Evidence for the Existence of Three Primary Strategies in Plants and Its Relevance to Ecological and Evolutionary Theory.” *The American Naturalist* 111 (982): 1169–94.

Guan, Rongxia, Jianbin Su, Xiangzong Meng, Sen Li, Yidong Liu, Juan Xu, and Shuqun Zhang. 2015. “Multilayered Regulation of Ethylene Induction Plays a Positive Role in Arabidopsis Resistance against Pseudomonas Syringae.” *Plant Physiology* 169 (1): 299–312.

Guillou, Laure, Dipankar Bachar, Stéphane Audic, David Bass, Cédric Berney, Lucie Bittner, Christophe Boutte, et al. 2013. “The Protist Ribosomal Reference Database (PR2): A Catalog of Unicellular Eukaryote Small Sub-Unit RRNA Sequences with Curated Taxonomy.” *Nucleic Acids Research* 41 (Database issue): D597-604.

Guittar, John, and Ashley Shade. 2019. “Microbiota on Perennial Biofuel Crops.” *Nature Communications*, 1–10.

Guo, Hailong, Shanshan Wang, and Jonathan D. G. Jones. 2021. “Autoactive Arabidopsis RPS4 Alleles Require Partner Protein RRS1-R.” *Plant Physiology* 185 (3): 761–64.

Gurevich, Alexey, Vladislav Saveliev, Nikolay Vyahhi, and Glenn Tesler. 2013. “QUAST: Quality Assessment Tool for Genome Assemblies.” *Bioinformatics* 29 (8): 1072–75.

Haas, Brian J., Steven L. Salzberg, Wei Zhu, Mihaela Pertea, Jonathan E. Allen, Joshua Orvis, Owen White, C. Robin Robin, and Jennifer R. Wortman. 2008. “Automated Eukaryotic Gene Structure Annotation Using EVidenceModeler and the Program to Assemble Spliced Alignments.” *Genome Biology* 9 (1). https://doi.org/10.1186/gb-2008-9-1-r7.

Hahne, Florian, and Robert Ivanek. 2016. “Visualizing Genomic Data Using Gviz and Bioconductor.” *Methods in Molecular Biology* 1418: 335–51.

Hao, Zhaodong, Dekang Lv, Ying Ge, Jisen Shi, Dolf Weijers, Guangchuang Yu, and Jinhui Chen. 2020. “RIdeogram: Drawing SVG Graphics to Visualize and Map Genome-Wide Data on the Idiograms.” *PeerJ. Computer Science* 6 (January): e251.

Helliwell, Emily E., Qin Wang, and Yinong Yang. 2016. “Ethylene Biosynthesis and Signaling Is Required for Rice Immune Response and Basal Resistance against Magnaporthe Oryzae Infection.” *Molecular Plant-Microbe Interactions: MPMI* 29 (11): 831–43.

Hilfiker, Olivier, Raphaël Groux, Friederike Bruessow, Karin Kiefer, Jürgen Zeier, and Philippe Reymond. 2014. “Insect Eggs Induce a Systemic Acquired Resistance in Arabidopsis.” *The Plant Journal: For Cell and Molecular Biology* 80 (6): 1085–94.

Holub, E. B. 2001. “The Arms Race Is Ancient History in Arabidopsis, the Wildflower.” *Nature Reviews. Genetics* 2 (7): 516–27.

Hoorn, Renier A. L. van der, and Sophien Kamoun. 2008. “From Guard to Decoy: A New Model for Perception of Plant Pathogen Effectors.” *The Plant Cell* 20 (8): 2009–17.

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 (May): 1–7.

Horton, Matthew W., Angela M. Hancock, Yu S. Huang, Christopher Toomajian, Susanna Atwell, Adam Auton, N. Wayan Muliyati, et al. 2012. “Genome-Wide Patterns of Genetic Variation in Worldwide Arabidopsis Thaliana Accessions from the RegMap Panel.” *Nature Genetics* 44 (2): 212–16.

Hou, Shiji, Thorsten Thiergart, Nathan Vannier, Fantin Mesny, Jörg Ziegler, Brigitte Pickel, and Stéphane Hacquard. 2021. "A Microbiota-Root-Shoot Circuit Favours Arabidopsis Growth over Defence under Suboptimal Light." *Nature Plants* 7 (8): 1078–92.

Howe, K., P. H. Schiffer, J. Zielinski, T. Wiehe, G. K. Laird, J. C. Marioni, O. Soylemez, F. Kondrashov, and M. Leptin. 2016. "Structure and Evolutionary History of a Large Family of NLR Proteins in the Zebrafish." *Open Biology* 6 (4): 1–15.

Huang, Xiaozhen, Jianyong Li, Fei Bao, Xiaoyan Zhang, and Shuhua Yang. 2010. "A Gain-of-Function Mutation in the Arabidopsis Disease Resistance Gene RPP4 Confers Sensitivity to Low Temperature." *Plant Physiology* 154 (2): 796–809.

Hubert, David A., Pablo Tornero, Youssef Belkhadir, Priti Krishna, Akira Takahashi, Ken Shirasu, and Jeffery L. Dangl. 2003. "Cytosolic HSP90 Associates with and Modulates the Arabidopsis RPM1 Disease Resistance Protein." *The EMBO Journal* 22 (21): 5679–89.

Huh, Sung Un, Volkan Cevik, Pingtao Ding, Zane Duxbury, Yan Ma, Laurence Tomlinson, Panagiotis F. Sarris, and Jonathan D. G. Jones. 2017. "Protein-Protein Interactions in the RPS4/RRS1 Immune Receptor Complex." *PLoS Pathogens* 13 (5): e1006376.

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.

Iqbal, Noushina, Asim Masood, M. Iqbal R. Khan, Mohd Asgher, Mehar Fatma, and Nafees A. Khan. 2013. "Cross-Talk between Sulfur Assimilation and Ethylene Signaling in Plants © 2013 Landes Bioscience . Do Not Distribute © 2013 Landes Bioscience . Do Not Distribute." *Plant Signaling & Behavior*, no. January: 104–12.

Iqbal, Zahra, Mohammed Shariq Iqbal, Abeer Hashem, Elsayed Fathi Abd_Allah, and Mohammad Israil Ansari. 2021. "Plant Defense Responses to Biotic Stress and Its Interplay With Fluctuating Dark/Light Conditions." *Frontiers in Plant Science* 12 (March): 1–22.

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

Iwase, S-I, K. Nakahira, M. Tuda, K. Kagoshima, and M. Takagi. 2015. "Host-Plant Dependent Population Genetics of the Invading Weevil Hypera Postica." *Bulletin of Entomological Research* 105 (1): 92–100.

Jacob, Florence, Saskia Vernaldi, and Takaki Maekawa. 2013. "Evolution and Conservation of Plant NLR Functions." *Frontiers in Immunology* 4 (SEP): 1–16.

Jambunathan, N. 2001. "A Humidity-Sensitive Arabidopsis Copine Mutant Exhibits Precocious Cell Death and Increased Disease Resistance." *The Plant Cell* 13 (10): 2225–40.

Jiao, Wen-Biao, and Korbinian Schneeberger. 2020. "Chromosome-Level Assemblies of Multiple Arabidopsis Genomes Reveal Hotspots of Rearrangements with Altered Evolutionary Dynamics." *Nature Communications* 11 (1): 989.

Jones, J. D., and J. L. Dangl. 2006. "The Plant Immune System." *Nature* 444: 323–29.

Jones, Jonathan D. G., Russell E. Vance, and Jeffery L. Dangl. 2016. "Intracellular Innate Immune Surveillance Devices in Plants and Animals." *Science (New York, N.Y.)* 354 (6316): aaf6395–aaf6395.

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): 168-179.e4.

Karasov, Talia L., Manuela Neumann, Alejandra Duque-Jaramillo, Sonja Kersten, Ilja Bezrukov, Birgit Schröppel, Efthymia Symeonidi, et al. 2019. "The Relationship between Microbial Biomass and Disease in the Arabidopsis Thaliana Phyllosphere." *BioRxiv*, 828814.

Katz, Ella, Jia-Jie Li, Benjamin Jaegle, Haim Ashkenazy, Shawn R. Abrahams, Clement Bagaza, Samuel Holden, Chris J. Pires, Ruthie Angelovici, and Daniel J. Kliebenstein. 2021. "Genetic Variation, Environment and Demography Intersect to Shape Arabidopsis Defense Metabolite Variation across Europe." *ELife* 10 (May): e67784.

Kerdraon, Lydie, Matthieu Barret, Valérie Laval, and Frédéric Suffert. 2019. "Differential Dynamics of Microbial Community Networks Help Identify Microorganisms Interacting with Residue-Borne Pathogens: The Case of Zymoseptoria Tritici in Wheat." *Microbiome* 7 (1): 125.

Khan, Abdul Latif, Muhammad Waqas, Sang-Mo Kang, Ahmed Al-Harrasi, Javid Hussain, Ahmed Al-Rawahi, Salima Al-Khiziri, et al. 2014. "Bacterial Endophyte Sphingomonas Sp. LK11 Produces Gibberellins and IAA and Promotes Tomato Plant Growth." *The Journal of Microbiology* 52 (8): 689–95.

Kim, Sang Hee, Fei Gao, Saikat Bhattacharjee, Joseph A. Adiasor, Ji Chul Nam, and Walter Gassmann. 2010. "The Arabidopsis Resistance-like Gene SNC1 Is Activated by Mutations in SRFR1 and Contributes to Resistance to the Bacterial Effector AvrRps4." *PLoS Pathogens* 6 (11): e1001172.

Kim, Yeon-Ju, Joon Young Park, Sri Renukadevi Balusamy, Yue Huo, Linh Khanh Nong, Hoa Thi Le, Deok Chun Yang, and Donghyuk Kim. 2020. "Comprehensive Genome Analysis on the Novel Species Sphingomonas Panacis DCY99T Reveals Insights into Iron Tolerance of Ginseng." *International Journal of Molecular Sciences* 21 (6): 2019.

Kim, Yong Sig, Chuanfu An, Sunchung Park, Sarah J. Gilmour, Ling Wang, Luciana Renna, Federica Brandizzi, Rebecca Grumet, and Michael F. Thomashow. 2017. "CAMTA-Mediated Regulation of Salicylic Acid Immunity Pathway Genes in Arabidopsis Exposed to Low Temperature and Pathogen Infection." *The Plant Cell* 29 (10): 2465–77.

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.

Kolton, Max, Armin Erlacher, Gabriele Berg, and Eddie Cytryn. 2016. "The Flavobacterium Genus in the Plant Holobiont: Ecological, Physiological, and Applicative Insights." In *Microbial Models: From Environmental to Industrial Sustainability*, 189–207. Singapore: Springer Singapore.

Kono, Thomas J. Y., Alex B. Brohammer, Suzanne E. McGaugh, and Candice N. Hirsch. 2018. "Tandem Duplicate Genes in Maize Are Abundant and Date to Two Distinct Periods of Time." *G3 (Bethesda, Md.)* 8 (9): 3049–58.

Koornneef, M., J. van Eden, C. J. Hanhart, P. Stam, F. J. Braaksma, and W. J. Feenstra. 1983. "Linkage Map of Arabidopsis Thaliana." *The Journal of Heredity* 74 (4): 265–72.

Koren, Sergey, Brian P. Walenz, Konstantin Berlin, Jason R. Miller, Nicholas H. Bergman, and Adam M. Phillippy. 2017. "Canu: Scalable and Accurate Long-Read Assembly via Adaptive k-Mer Weighting and Repeat Separation." *Genome Research* 27 (5): 722–36.

Kourelis, Jiorgos, Shivani Malik, Oliver Mattinson, Sonja Krauter, Parvinderdeep S. Kahlon, Judith K. Paulus, and Renier A. L. van der Hoorn. 2020. "Evolution of a Guarded Decoy Protease and Its Receptor in Solanaceous Plants." *Nature Communications* 11 (1): 4393.

Krasileva, Ksenia V., Douglas Dahlbeck, and Brian J. Staskawicz. 2010. "Activation of an Arabidopsis Resistance Protein Is Specified by the in Planta Association of Its Leucine-Rich Repeat Domain with the Cognate Oomycete Effector." *The Plant Cell* 22 (7): 2444–58.

Kriss, Michael, Keith Z. Hazleton, Nichole M. Nusbacher, Casey G. Martin, and Catherine A. Lozupone. 2018. "Low Diversity Gut Microbiota Dysbiosis: Drivers, Functional Implications and Recovery." *Current Opinion in Microbiology* 44 (August): 34–40.

Kronenberg, Zev N., Arang Rhie, Sergey Koren, Gregory T. Concepcion, Paul Peluso, Katherine M. Munson, Stefan Hiendleder, et al. 2018. "Extended Haplotype Phasing of de Novo Genome Assemblies with FALCON-Phase." *BioRxiv*. bioRxiv. https://doi.org/10.1101/327064.

Kuang, Hanhui, Sung-Sick Woo, Blake C. Meyers, Eviatar Nevo, and Richard W. Michelmore. 2004. "Multiple Genetic Processes Result in Heterogeneous Rates of Evolution within the Major Cluster Disease Resistance Genes in Lettuce." *The Plant Cell* 16 (11): 2870–94.

Kumar, Akhilesh, and Jay Prakash Verma. 2018. "Does Plant—Microbe Interaction Confer Stress Tolerance in Plants: A Review?" *Microbiological Research* 207 (March): 41–52.

Kumar, Jay, Piyoosh K. Babele, Divya Singh, and Ashok Kumar. 2016. "UV-B Radiation Stress Causes Alterations in Whole Cell Protein Profile and Expression of Certain Genes in the Rice Phyllospheric Bacterium Enterobacter Cloacae." *Frontiers in Microbiology* 7 (SEP): 1–14.

Laforest-Lapointe, Isabelle, Christian Messier, and Steven W. Kembel. 2016. "Host Species Identity, Site and Time Drive Temperate Tree Phyllosphere Bacterial Community Structure." *Microbiome* 4: 1–10.

Lai, Yan, and Thomas Eulgem. 2018. "Transcript-Level Expression Control of Plant NLR Genes." *Molecular Plant Pathology* 19 (5): 1267–81.

Laibach, F. 1943. "Arabidopsis Thaliana (L.) Heynh. Als Objekt Für Genetische Und Entwicklungsphysiologische Untersuchungen." *Bot Arch* 44: 439–455.

Lallemand, Tanguy, Martin Leduc, Claudine Landès, Carène Rizzon, and Emmanuelle Lerat. 2020. "An Overview of Duplicated Gene Detection Methods: Why the Duplication Mechanism Has to Be Accounted for in Their Choice." *Genes* 11 (9): 1046.

Langridge, J. 1994. "Arabidopsis Thaliana, a Plant Drosophila." *BioEssays: News and Reviews in Molecular, Cellular and Developmental Biology* 16 (10): 775–78.

Lapin, Dmitry, Deepak D. Bhandari, and Jane E. Parker. 2020. "Origins and Immunity Networking Functions of EDS1 Family Proteins." *Annual Review of Phytopathology* 58: 253–76.

Latz, Meike A. C., Mads Herbert Kerrn, Helle Sørensen, David B. Collinge, Birgit Jensen, James K. M. Brown, Anne Mette Madsen, and Hans Jørgen Lyngs Jørgensen. 2021. "Succession of the Fungal Endophytic Microbiome of Wheat Is Dependent on Tissue-Specific Interactions between Host Genotype and Environment." *The Science of the Total Environment* 759: 143804–143804.

Lawrence, Michael, Robert Gentleman, and Vincent Carey. 2009. "Rtracklayer: An R Package for Interfacing with Genome Browsers." *Bioinformatics (Oxford, England)* 25 (14): 1841–42.

Le Roux C Jauneau A Camborde L Trémousaygue D Kraut A Zhou B Levaillant M Adachi H Yoshioka H Raffaele S Berthomé R Couté Y Parker JE Deslandes L., Huet G. 2015. "A Receptor Pair with an Integrated Decoy Converts Pathogen Disabling of Transcription Factors to Immunity." *Cell*, 2015.

Lee, Hyun-Ah, Hye-Young Lee, Eunyoung Seo, Joohyun Lee, Saet-Byul Kim, Soohyun Oh, Eunbi Choi, Eunhye Choi, So Eui Lee, and Doil Choi. 2017. "Current Understandings of Plant Nonhost Resistance." *Molecular Plant-Microbe Interactions: MPMI* 30 (1): 5–15.

Lee, Jennifer M., and Erik L. L. Sonnhammer. 2003. "Genomic Gene Clustering Analysis of Pathways in Eukaryotes." *Genome Research* 13 (5): 875–82.

Lee, Rachelle R. Q., and Eunyoung Chae. 2020. "Variation Patterns of NLR Clusters in Arabidopsis Thaliana Genomes." *Plant Communications* 1 (4): 100089.

Lemanceau, P., M. Barret, S. Mazurier, S. Mondy, B. Pivato, T. Fort, and C. Vacher. 2017. "Plant Communication With Associated Microbiota in the Spermosphere, Rhizosphere and Phyllosphere." In *Advances in Botanical Research*, 101–33. Advances in Botanical Research. Elsevier.

Leopold, Devin R., and Posy E. Busby. 2020. "Host Genotype and Colonist Arrival Order Jointly Govern Plant Microbiome Composition and Function." *Current Biology: CB* 30 (16): 3260-3266.e5.

Leskovec, Jure, and Rok Sosič. 2016. "SNAP: A General Purpose Network Analysis and Graph Mining Library." *ACM Transactions on Intelligent Systems and Technology* 8 (1): 1–20.

Levy, Asaf, Isai Salas Gonzalez, Maximilian Mittelviefhaus, Scott Clingenpeel, Sur Herrera Paredes, Jiamin Miao, Kunru Wang, et al. 2018. "Genomic Features of Bacterial Adaptation to Plants." *Nature Genetics* 50 (1): 138–50.

Li, Fuyong, Changxi Li, Yanhong Chen, Junhong Liu, Chunyan Zhang, Barry Irving, Carolyn Fitzsimmons, Graham Plastow, and Le Luo Guan. 2019. "Host Genetics Influence the Rumen Microbiota and Heritable Rumen Microbial Features Associate with Feed Efficiency in Cattle." *Microbiome* 7 (1): 92.

Li, Pingchuan, Xiande Quan, Gaofeng Jia, Jin Xiao, Sylvie Cloutier, and Frank M. You. 2016. "RGAugury: A Pipeline for Genome-Wide Prediction of Resistance Gene Analogs (RGAs) in Plants." *BMC Genomics* 17 (1): 1–10.

Li, Yang-Yang, James A. Pearson, Chen Chao, Jian Peng, Xiaojun Zhang, Zhiguang Zhou, Yu Liu, F. Susan Wong, and Li Wen. 2017. "Nucleotide-Binding Oligomerization Domain-Containing Protein 2 (Nod2) Modulates T1DM Susceptibility by Gut Microbiota." *Journal of Autoimmunity* 82 (August): 85–95.

Liang, Wanwan, Meixuezi Tong, and Xin Li. 2020. "SUSA2 Is an F-Box Protein Required for Autoimmunity Mediated by Paired NLRs SOC3-CHS1 and SOC3-TN2." *Nature Communications* 11 (1): 5190.

Lindow, Steven E., and Maria T. Brandl. 2003. "Microbiology of the Phyllosphere." *Applied and Environmental Microbiology* 69 (4): 1875–83.

Lindström, Eva S., and Silke Langenheder. 2012. "Local and Regional Factors Influencing Bacterial Community Assembly." *Environmental Microbiology Reports* 4 (1): 1–9.

Liu, Yule, Tessa Burch-Smith, Michael Schiff, Suhua Feng, and Savithramma P. Dinesh-Kumar. 2004. "Molecular Chaperone Hsp90 Associates with Resistance Protein N and Its Signaling Proteins SGT1 and Rar1 to Modulate an Innate Immune Response in Plants." *The Journal of Biological Chemistry* 279 (3): 2101–8.

Lolle, Signe, Danielle Stevens, and Gitta Coaker. 2020. "Plant NLR-Triggered Immunity: From Receptor Activation to Downstream Signaling." *Current Opinion in Immunology* 62 (February): 99–105.

Lu, You, and Kenichi Tsuda. 2021a. "Intimate Association of PRR- and NLR-Mediated Signaling in Plant Immunity." *Molecular Plant-Microbe Interactions: MPMI* 34 (1): 3–14.

Lundberg, Derek S., Roger de Pedro Jové, Pratchaya Pramoj Na Ayutthaya, Talia L. Karasov, Or Shalev, Karin Poersch, Wei Ding, Anita Bollmann-Giolai, Ilja Bezrukov, and Detlef Weigel. 2021. "Contrasting Patterns of Microbial Dominance in the Arabidopsis Thaliana Phyllosphere." *BioRxiv*. bioRxiv. https://doi.org/10.1101/2021.04.06.438366.

Luo, Luyun, Zhuo Zhang, Pei Wang, Yongqin Han, Decai Jin, Pin Su, Xinqiu Tan, et al. 2019. "Variations in Phyllosphere Microbial Community along with the Development of Angular Leaf-Spot of Cucumber." *AMB Express* 9 (1): 76.

Ma, S., D. Lapin, L. Liu, Y. Sun, W. Song, X. Zhang, and J. Chai. 2020. "Direct Pathogen-Induced Assembly of an NLR Immune Receptor Complex to Form a Holoenzyme." *Science* 370 (6521): 3069.

MacQueen, Alice, Dacheng Tian, Wenhan Chang, Eric Holub, Martin Kreitman, and Joy Bergelson. 2019. "Population Genetics of the Highly Polymorphic RPP8 Gene Family." *Genes* 10 (9): 691.

Maekawa, T., T. A. Kufer, and P. Schulze-Lefert. 2011. "NLR Functions in Plant and Animal Immune Systems: So Far and yet so Close." *Nature Immunology* 12 (9): 817–26.

Majoros, W. H., M. Pertea, and S. L. Salzberg. 2004. "TigrScan and GlimmerHMM: Two Open Source Ab Initio Eukaryotic Gene-Finders." *Bioinformatics* 20 (16): 2878–79.

Manching, Heather C., Peter J. Balint-Kurti, and Ann E. Stapleton. 2014. "Southern Leaf Blight Disease Severity Is Correlated with Decreased Maize Leaf Epiphytic Bacterial Species Richness and the Phyllosphere Bacterial Diversity Decline Is Enhanced by Nitrogen Fertilization." *Frontiers in Plant Science* 5 (August): 403.

Mang, Hyung-Gon, Weiqiang Qian, Ying Zhu, Jun Qian, Hong-Gu Kang, Daniel F. Klessig, and Jian Hua. 2012. "Abscisic Acid Deficiency Antagonizes High-Temperature Inhibition of Disease Resistance through Enhancing Nuclear Accumulation of Resistance Proteins SNC1 and RPS4 in Arabidopsis." *The Plant Cell* 24 (3): 1271–84.

Manter, Daniel K., Jorge A. Delgado, David G. Holm, and Rachel A. Stong. 2010. "Pyrosequencing Reveals a Highly Diverse and Cultivar-Specific Bacterial Endophyte Community in Potato Roots." *Microbial Ecology* 60 (1): 157–66.

Martin, Raoul, Tiancong Qi, Haibo Zhang, Furong Liu, Miles King, Claire Toth, Eva Nogales, and Brian J. Staskawicz. 2020. "Structure of the Activated Roq1 Resistosome Directly Recognizing the Pathogen Effector XopQ." *BioRxiv*. bioRxiv. https://doi.org/10.1101/2020.08.13.246413.

Martinez-Medina, Ainhoa, Victor Flors, Martin Heil, Brigitte Mauch-Mani, Corné M. J. Pieterse, Maria J. Pozo, Jurriaan Ton, Nicole M. van Dam, and Uwe Conrath. 2016. "Recognizing Plant Defense Priming." *Trends in Plant Science* 21 (10): 818–22.

Mascagni, Flavia, Gabriele Usai, Andrea Cavallini, and Andrea Porceddu. 2021. "Structural Characterization and Duplication Modes of Pseudogenes in Plants." *Scientific Reports* 11 (1): 5292.

Matsumoto, Haruna, Xiaoyan Fan, Yue Wang, Peter Kusstatscher, Jie Duan, Sanling Wu, Sunlu Chen, et al. 2021. "Bacterial Seed Endophyte Shapes Disease Resistance in Rice." *Nature Plants* 7 (1): 60–72.

May, Rose-Lucy, Stuart Warner, and Astrid Wingler. 2017. "Classification of Intra-Specific Variation in Plant Functional Strategies Reveals Adaptation to Climate." *Annals of Botany* 119 (8): 1343–52.

Mersmann, Sophia, Gildas Bourdais, Steffen Rietz, and Silke Robatzek. 2010. "Ethylene Signaling Regulates Accumulation of the FLS2 Receptor and Is Required for the Oxidative Burst Contributing to Plant Immunity." *Plant Physiology* 154 (1): 391–400.

Meyerowitz, E. M., and R. E. Pruitt. 1985. "Arabidopsis Thaliana and Plant Molecular Genetics." *Science (New York, N.Y.)* 229 (4719): 1214–18.

Meyers, Blake C., Alexander Kozik, Alyssa Griego, Hanhui Kuang, and Richard W. Michelmore. 2003. "Genome-Wide Analysis of NBS-LRR-Encoding Genes in Arabidopsis." *The Plant Cell* 15 (4): 809–34.

Mina, Diogo, José Alberto Pereira, Teresa Lino-Neto, and Paula Baptista. 2020. "Epiphytic and Endophytic Bacteria on Olive Tree Phyllosphere: Exploring Tissue and Cultivar Effect." *Microbial Ecology* 80 (1): 145–57.

Mondragón-Palomino, Mariana, Blake C. Meyers, Richard W. Michelmore, and Brandon S. Gaut. 2002. "Patterns of Positive Selection in the Complete NBS-LRR Gene Family of Arabidopsis Thaliana." *Genome Research* 12 (9): 1305–15.

Mondragón-Palomino, Mariana, Remco Stam, Ajay John-Arputharaj, and Thomas Dresselhaus. 2017. "Diversification of Defensins and NLRs in Arabidopsis Species by Different Evolutionary Mechanisms." *BMC Evolutionary Biology* 17 (1). https://doi.org/10.1186/s12862-017-1099-4.

Monteiro, Freddy, and Marc T. Nishimura. 2018. "Structural, Functional, and Genomic Diversity of Plant NLR Proteins: An Evolved Resource for Rational Engineering of Plant Immunity." *Annual Review of Phytopathology* 56 (1): 243–67.

Montesinos, Alicia, Stephen J. Tonsor, Carlos Alonso-Blanco, and F. Xavier Picó. 2009. "Demographic and Genetic Patterns of Variation among Populations of Arabidopsis Thaliana from Contrasting Native Environments." *PloS One* 4 (9): e7213.

Morella, Norma M., Francis Cheng Hsuan Weng, Pierre M. Joubert, C. Jessica, 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.

Mougi, A., and M. Kondoh. 2012. "Diversity of Interaction Types and Ecological Community Stability." *Science (New York, N.Y.)* 337 (6092): 349–51.

Mozo, T., K. Dewar, P. Dunn, J. R. Ecker, S. Fischer, S. Kloska, H. Lehrach, et al. 1999. "A Complete BAC-Based Physical Map of the Arabidopsis Thaliana Genome." *Nature Genetics* 22 (3): 271–75.

Mühlenbock, Per, Magdalena Szechynska-Hebda, Marian Plaszczyca, Marcela Baudo, Alfonso Mateo, Philip M. Mullineaux, Jane E. Parker, Barbara Karpinska, and Stanislaw Karpinski. 2008. "Chloroplast Signaling and LESION SIMULATING DISEASE1 Regulate Crosstalk between Light Acclimation and Immunity in Arabidopsis." *The Plant Cell* 20 (9): 2339–56.

Müller, Daniel B., Christine Vogel, Yang Bai, and Julia A. Vorholt. 2016. "The Plant Microbiota: Systems-Level Insights and Perspectives." *Annual Review of Genetics* 50 (1): 211–34.

Munch, David, Vikas Gupta, Asger Bachmann, Wolfgang Busch, Simon Kelly, Terry Mun, and Stig Uggerhøj Andersena. 2018. "The Brassicaceae Family Displays Divergent, Shoot-Skewed NLR Resistance Gene Expression." *Plant Physiology* 176 (2): 1598–1609.

Naish, M., M. Alonge, P. Wlodzimierz, and A. J. Tock. 2021. "The Genetic and Epigenetic Landscape of the Arabidopsis Centromeres." *BioRxiv*. https://www.biorxiv.org/content/10.1101/2021.05.30.446350v2.abstract.

Narusaka, Mari, Ken Shirasu, Yoshiteru Noutoshi, Yasuyuki Kubo, Tomonori Shiraishi, Masaki Iwabuchi, and Yoshihiro Narusaka. 2009. "RRS1 and RPS4 Provide a Dual Resistance-Gene System against Fungal and Bacterial Pathogens." *The Plant Journal: For Cell and Molecular Biology* 60 (2): 218–26.

Nascimento, Francisco X., Márcio J. Rossi, and Bernard R. Glick. 2018. "Ethylene and 1-Aminocyclopropane-1-Carboxylate (ACC) in Plant–Bacterial Interactions." *Frontiers in Plant Science* 9 (February). https://doi.org/10.3389/fpls.2018.00114.

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.

Nawrocki, Eric P., and Sean R. Eddy. 2013. “Infernal 1.1: 100-Fold Faster RNA Homology Searches.” *Bioinformatics* 29 (22): 2933–35.

Nelson, David R., Basel Khraiwesh, Weiqi Fu, Saleh Alseekh, Ashish Jaiswal, Amphun Chaiboonchoe, Khaled M. Hazzouri, et al. 2017. “The Genome and Phenome of the Green Alga Chloroidium Sp. UTEX 3007 Reveal Adaptive Traits for Desert Acclimatization.” *ELife* 6 (June). https://doi.org/10.7554/eLife.25783.

Ngou, Bruno Pok Man, Hee-Kyung Ahn, Pingtao Ding, and Jonathan D. G. Jones. 2021. “Mutual Potentiation of Plant Immunity by Cell-Surface and Intracellular Receptors.” *Nature* 592 (7852): 110–15.

Niinemets, Ülo. 2010. “Responses of Forest Trees to Single and Multiple Environmental Stresses from Seedlings to Mature Plants: Past Stress History, Stress Interactions, Tolerance and Acclimation.” *Forest Ecology and Management* 260 (10): 1623–39.

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.

Noel, Laurent, Tracey L. Moores, Erik A. van der Biezen, Martin Parniske, Michael J. Daniels, Jane E. Parker, and Jonathan D. G. Jones. 1999. “Pronounced Intraspecific Haplotype Divergence at the RPP5 Complex Disease Resistance Locus of Arabidopsis.” *The Plant Cell* 11 (11): 2099.

Noh, Hyun-Ju, Yung Mi Lee, Chae Haeng Park, Hong Kum Lee, Jang-Cheon Cho, and Soon Gyu Hong. 2020. “Microbiome in Cladonia Squamosa Is Vertically Stratified According to Microclimatic Conditions.” *Frontiers in Microbiology* 11 (February): 268.

Noman, Ali, Muhammad Aqeel, and Yonggen Lou. 2019. “PRRs and NB-LRRs: From Signal Perception to Activation of Plant Innate Immunity.” *International Journal of Molecular Sciences* 20 (8). https://doi.org/10.3390/ijms20081882.

Nordborg, Magnus, Tina T. Hu, Yoko Ishino, Jinal Jhaveri, Christopher Toomajian, Honggang Zheng, Erica Bakker, et al. 2005. “The Pattern of Polymorphism in Arabidopsis Thaliana.” *PLoS Biology* 3 (7): e196.

Olson, Daniel, and Travis Wheeler. 2018. “ULTRA: A Model Based Tool to Detect Tandem Repeats.” *ACM-BCB ... ... : The ... ACM Conference on Bioinformatics, Computational Biology and Biomedicine. ACM Conference on Bioinformatics, Computational Biology and Biomedicine* 2018 (August): 37–46.

Ortiz, Diana, and Peter N. Dodds. 2018. “Plant NLR Origins Traced Back to Green Algae.” *Trends in Plant Science*.

Ostrowski, Marie-France, Jacques David, Sylvain Santoni, Heather Mckhann, Xavier Reboud, Valerie Le Corre, Christine Camilleri, et al. 2006. “Evidence for a Large-Scale Population Structure among Accessions of Arabidopsis Thaliana: Possible Causes and Consequences for the Distribution of Linkage Disequilibrium.” *Molecular Ecology* 15 (6): 1507–17.

Panchal, Shweta, Reejana Chitrakar, Blaine K. Thompson, Nisita Obulareddy, Debanjana Roy, W. Sealy Hambright, and Maeli Melotto. 2016. “Regulation of Stomatal Defense by Air Relative Humidity.” *Plant Physiology* 172 (3): 2021–32.

Panchy, Nicholas, Melissa Lehti-Shiu, and Shin-Han Shiu. 2016. “Evolution of Gene Duplication in Plants.” *Plant Physiology* 171 (4): 2294–2316.

Parker, Jane E., Eric B. Holub, Louise N. Frost, Anders Falk, Nick D. Gunn, and Michael J. Daniels. 1996. “Characterization of Eds1, a Mutation in Arabidopsis Suppressing Resistance to Peronospora Parasitica Specified by Several Different RPP Genes.” *The Plant Cell* 8 (11): 2033–46.

Passari, Ajit Kumar, Vineet Kumar Mishra, Garima Singh, Pratibha Singh, Brijesh Kumar, Vijai Kumar Gupta, Rupak Kumar Sarma, Ratul Saikia, Anthonia O’ Donovan, and Bhim Pratap Singh. 2017. “Insights into the Functionality of Endophytic Actinobacteria with a Focus on Their Biosynthetic Potential and Secondary Metabolites Production.” *Scientific Reports* 7 (1): 1–17.

Peng, Yujun, Rowan van Wersch, and Yuelin Zhang. 2018. “Convergent and Divergent Signaling in PAMP-Triggered Immunity and Effector-Triggered Immunity.” *Molecular Plant-Microbe Interactions: MPMI* 31 (4): 403–9.

Peredo, Elena L., and Sheri L. Simmons. 2018. “Leaf-FISH: Microscale Imaging of Bacterial Taxa on Phyllosphere.” *Frontiers in Microbiology* 8 (JAN): 1–14.

Perreault, Rosaëlle, and Isabelle Laforest-Lapointe. 2021. “Plant-Microbe Interactions in the Phyllosphere: Facing Challenges of the Anthropocene.” *The ISME Journal*, September. https://doi.org/10.1038/s41396-021-01109-3.

Pfeilmeier, Sebastian, Gabriella C. Petti, Miriam Bortfeld-Miller, Benjamin Daniel, Christopher M. Field, Shinichi Sunagawa, and Julia A. Vorholt. 2021. "The Plant NADPH Oxidase RBOHD Is Required for Microbiota Homeostasis in Leaves." *Nature Microbiology* 6 (7): 852–64.

Pincebourde, Sylvain, and Jérôme Casas. 2019. "Narrow Safety Margin in the Phyllosphere during Thermal Extremes." *Proceedings of the National Academy of Sciences of the United States of America* 116 (12): 5588–96.

Poudel, Ravin, Ari Jumpponen, Daniel Schlatter, Timothy Carl Paulitz, Brian McSpadden Gardener, Linda L. Kinkel, and Karen A. Garrett. 2016. "Microbiome Networks: A Systems Framework for Identifying Candidate Microbial Assemblages for Disease Management." *Phytopathology* 1: PHYTO-02-16-0058-FI.

Proell, Martina, Stefan J. Riedl, Jörg H. Fritz, Ana M. Rojas, and Robert Schwarzenbacher. 2008. "The Nod-like Receptor (NLR) Family: A Tale of Similarities and Differences." *PloS One* 3 (4): e2119.

Purahong, Witoon, Luigi Orrù, Irene Donati, Giorgia Perpetuini, Antonio Cellini, Antonella Lamontanara, Vania Michelotti, Gianni Tacconi, and Francesco Spinelli. 2018. "Plant Microbiome and Its Link to Plant Health: Host Species, Organs and Pseudomonas Syringae Pv. Actinidiae Infection Shaping Bacterial Phyllosphere Communities of Kiwifruit Plants." *Frontiers in Plant Science* 9 (November): 1563.

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.

Qin, Sheng, Ke Xing, Ji-Hong Jiang, Li-Hua Xu, and Wen-Jun Li. 2011. "Biodiversity, Bioactive Natural Products and Biotechnological Potential of Plant-Associated Endophytic Actinobacteria." *Applied Microbiology and Biotechnology* 89 (3): 457–73.

Quinlan, Aaron R., and Ira M. Hall. 2010. "BEDTools: A Flexible Suite of Utilities for Comparing Genomic Features." *Bioinformatics (Oxford, England)* 26 (6): 841–42.

Rastogi, Gurdeep, Adrian Sbodio, Jan J. Tech, Trevor V. Suslow, Gitta L. Coaker, and Johan H. J. Leveau. 2012. "Leaf Microbiota in an Agroecosystem: Spatiotemporal Variation in Bacterial Community Composition on Field-Grown Lettuce." *The ISME Journal* 6 (10): 1812–22.

Rédei, G. P. 1962. "Supervital Mutants of Arabidopsis." *Genetics* 47 (4): 443–60.

Rédei, George P. 1992. "A Heuristic Glance at the Past of Arabidopsis Genetics." In *Methods in Arabidopsis Research*, 1–15. WORLD SCIENTIFIC.

Redkar, Amey, Volkan Cevik, Kate Bailey, Oliver J. Furzer, Sebastian Fairhead, M. Hossein Borhan, Eric B. Holub, and Jonathan D. G. Jones. 2021. "The Arabidopsis WRR4A and WRR4B Paralogous NLR Proteins Both Confer Recognition of Multiple Albugo Candida Effectors." *BioRxiv*. bioRxiv. https://doi.org/10.1101/2021.03.29.436918.

Remus-Emsermann, Mitja N. P., Sebastian Lücker, Daniel B. Müller, Eva Potthoff, Holger Daims, and Julia A. Vorholt. 2014. "Spatial Distribution Analyses of Natural Phyllosphere-Colonizing Bacteria on Arabidopsis Thaliana Revealed by Fluorescence in Situ Hybridization." *Environmental Microbiology* 16 (7): 2329–40.

Remus-Emsermann, Mitja N. P., and Rudolf O. Schlechter. 2018. "Phyllosphere Microbiology: At the Interface between Microbial Individuals and the Plant Host." *The New Phytologist* 218 (4): 1327–33.

Rice, P., I. Longden, and A. Bleasby. 2000. "EMBOSS: The European Molecular Biology Open Software Suite." *Trends in Genetics: TIG* 16 (6): 276–77.

Richard, Manon M. S., and Frank L. W. Takken. 2017. "Plant Autoimmunity: When Good Things Go Bad." *Current Biology: CB*. Elsevier BV.

Ritpitakphong, Unyarat, Laurent Falquet, Artit Vimoltust, Antoine Berger, Jean Pierre Métraux, and Floriane L'Haridon. 2016. "The Microbiome of the Leaf Surface of Arabidopsis Protects against a Fungal Pathogen." *The New Phytologist* 210 (3): 1033–43.

Rudrappa, Thimmaraju, Meredith L. Biedrzycki, and Harsh P. Bais. 2008. "Causes and Consequences of Plant-Associated Biofilms." *FEMS Microbiology Ecology* 64 (2): 153–66.

Ruhe, Jonas, Matthew T. Agler, Aleksandra Placzek, Katharina Kramer, Iris Finkemeier, and Eric M. Kemen. 2016. "Obligate Biotroph Pathogens of the Genus Albugo Are Better Adapted to Active Host Defense Compared to Niche Competitors." *Frontiers in Plant Science* 7 (JUNE2016): 1–17.

Runge, Paul, Freddy Ventura, Eric Kemen, and Remco Stam. 2021. "Distinct Phyllosphere Microbiome of Wild Tomato Species in Central Peru upon Dysbiosis." *Research Square*. Research Square. https://doi.org/10.21203/rs.3.rs-648608/v1.

Saijo, Yusuke, Eliza Po-Iian Loo, and Shigetaka Yasuda. 2018. "Pattern Recognition Receptors and Signaling in Plant-Microbe Interactions." *The Plant Journal: For Cell and Molecular Biology* 93 (4): 592–613.

Salzberg, Steven L. 2019. "Next-Generation Genome Annotation: We Still Struggle to Get It Right." *Genome Biology* 20 (1): 92.

Sanjenbam, Pratibha, Radhika Buddidathi, Radhika Venkatesan, P. V. Shivaprasad, and Deepa Agashe. 2020. "Phenotypic Diversity of Methylobacterium Associated with Rice Landraces in North-East India." *PloS One* 15 (2): e0228550.

Sapkota, Rumakanta, Kamilla Knorr, Lise Nistrup Jørgensen, Karen A. O'Hanlon, and Mogens Nicolaisen. 2015. "Host Genotype Is an Important Determinant of the Cereal Phyllosphere Mycobiome." *The New Phytologist* 207 (4): 1134–44.

Sarmashghi, Shahab, Kristine Bohmann, M. Thomas P Gilbert, Vineet Bafna, and Siavash Mirarab. 2019. "Skmer: Assembly-Free and Alignment-Free Sample Identification Using Genome Skims." *Genome Biology* 20 (1): 34.

Sarris, Panagiotis F., Volkan Cevik, Gulay Dagdas, Jonathan D. G. Jones, and Ksenia V. Krasileva. 2016. "Comparative Analysis of Plant Immune Receptor Architectures Uncovers Host Proteins Likely Targeted by Pathogens." *BMC Biology* 14. https://doi.org/10.1186/s12915-016-0228-7.

Saur, Isabel Ml, Saskia Bauer, Barbara Kracher, Xunli Lu, Lamprinos Franzeskakis, Marion C. Müller, Björn Sabelleck, et al. 2019. "Multiple Pairs of Allelic MLA Immune Receptor-Powdery Mildew AVRA Effectors Argue for a Direct Recognition Mechanism." *ELife* 8 (February). https://doi.org/10.7554/eLife.44471.

Schlechter, Rudolf O., Moritz Miebach, and Mitja N. P. Remus-Emsermann. 2019. "Driving Factors of Epiphytic Bacterial Communities: A Review." *Journal of Advanced Research* 19: 57–65.

Schneeberger, Korbinian, Stephan Ossowski, Felix Ott, Juliane D. Klein, Xi Wang, Christa Lanz, Lisa M. Smith, et al. 2011. "Reference-Guided Assembly of Four Diverse Arabidopsis Thaliana Genomes." *Proceedings of the National Academy of Sciences of the United States of America* 108 (25): 10249–54.

Schwessinger, Benjamin, and John P. Rathjen. 2017. "Extraction of High Molecular Weight DNA from Fungal Rust Spores for Long Read Sequencing." *Methods in Molecular Biology (Clifton, N.J.)* 1659: 49–57.

Seong, Kyungyong, Eunyoung Seo, Kamil Witek, Meng Li, and Brian Staskawicz. 2020. "Evolution of NLR Resistance Genes with Noncanonical N-Terminal Domains in Wild Tomato Species." *The New Phytologist* 227 (5): 1530–43.

Shade, Ashley, Stuart E. Jones, J. Gregory Caporaso, Jo Handelsman, Rob Knight, Noah Fierer, and Jack A. Gilbert. 2014. "Conditionally Rare Taxa Disproportionately Contribute to Temporal Changes in Microbial Diversity." *MBio* 5 (4): e01371-14.

Shakir, Sara, Syed Shan e. Ali Zaidi, Franciska T. de Vries, and Shahid Mansoor. 2021. "Plant Genetic Networks Shaping Phyllosphere Microbial Community." *Trends in Genetics: TIG* 37 (4): 306–16.

Shao, Zhu-Qing, Jia-Yu Xue, Qiang Wang, Bin Wang, and Jian-Qun Chen. 2019. "Revisiting the Origin of Plant NBS-LRR Genes." *Trends in Plant Science* 24 (1): 9–12.

Sharbel, T. F., B. Haubold, and T. Mitchell-Olds. 2000. "Genetic Isolation by Distance in Arabidopsis Thaliana: Biogeography and Postglacial Colonization of Europe." *Molecular Ecology* 9 (12): 2109–18.

Shin, Dong-Hyeon, Dong-Uk Kim, Chi-Nam Seong, Hong-Gyu Song, and Jong-Ok Ka. 2012. "Genetic and Phenotypic Diversity of Carbofuran-Degrading Bacteria Isolated from Agricultural Soils." *Journal of Microbiology and Biotechnology* 22 (4): 448–56.

Simon, Jean-Christophe, Julian R. Marchesi, Christophe Mougel, and Marc-André Selosse. 2019. "Host-Microbiota Interactions: From Holobiont Theory to Analysis." *Microbiome* 7 (1): 5.

Singh, Prashant, Sylvain Santoni, Patrice This, and Jean-Pierre Péros. 2018. "Genotype-Environment Interaction Shapes the Microbial Assemblage in Grapevine's Phyllosphere and Carposphere: An NGS Approach." *Microorganisms* 6 (4). https://doi.org/10.3390/microorganisms6040096.

Sivakumar, Natesan, Ramamoorthy Sathishkumar, Gopal Selvakumar, Rajaram Shyamkumar, and Kalimuthu Arjunekumar. 2020. "Phyllospheric Microbiomes: Diversity, Ecological Significance, and Biotechnological Applications." In *Sustainable Development and Biodiversity*, 113–72. Cham: Springer International Publishing.

Slater, Guy St C., and Ewan Birney. 2005. "Automated Generation of Heuristics for Biological Sequence Comparison." *BMC Bioinformatics* 6: 1–11.

Smit, A. F. 1993. "Identification of a New, Abundant Superfamily of Mammalian LTR-Transposons." *Nucleic Acids Research* 21 (8): 1863–72.

Somerville, C. R., and W. L. Ogren. 1980. "Inhibition of Photosynthesis in Arabidopsis Mutants Lacking Leaf Glutamate Synthase Activity." *Nature* 286 (5770): 257–59.

Somssich, Marc. 2019. "A Short History of Arabidopsis Thaliana (L.) Heynh. Columbia-0." *PeerJ*. https://doi.org/10.7287/peerj.preprints.26931v5.

Song, Chunxu, Feng Zhu, Víctor J. Carrión, and Viviane Cordovez. 2020. "Beyond Plant Microbiome Composition: Exploiting Microbial Functions and Plant Traits via Integrated Approaches." *Frontiers in Bioengineering and Biotechnology* 8 (August): 896.

Song, Wen, Alexander Forderer, Dongli Yu, and Jijie Chai. 2021. "Structural Biology of Plant Defence." *The New Phytologist* 229 (2): 692–711.

Stam, Remco, Daniela Scheikl, and Aurélien Tellier. 2016. "Pooled Enrichment Sequencing Identifies Diversity and Evolutionary Pressures at NLR Resistance Genes within a Wild Tomato Population." *Genome Biology and Evolution* 8 (5): 1501–15.

Stam, Remco, Gustavo A. Silva-Arias, and Aurelien Tellier. 2019. "Subsets of NLR Genes Show Differential Signatures of Adaptation during Colonization of New Habitats." *The New Phytologist* 224 (1): 367–79.

Stanke, Mario, and Burkhard Morgenstern. 2005. "AUGUSTUS : A Web Server for Gene Prediction in Eukaryotes That Allows User-Defined Constraints" 33: 465–67.

Stenøien, Hans K., Charles B. Fenster, Anni Tonteri, and Outi Savolainen. 2004. "Genetic Variability in Natural Populations of Arabidopsis Thaliana in Northern Europe." *Molecular Ecology* 14 (1): 137–48.

Stewart, Adeline, Delphine Rioux, Fréderic Boyer, Ludovic Gielly, François Pompanon, Amélie Saillard, Wilfried Thuiller, Jean-Gabriel Valay, Eric Maréchal, and Eric Coissac. 2021. "Altitudinal Zonation of Green Algae Biodiversity in the French Alps." *Frontiers in Plant Science* 12 (June): 679428.

Stone, Bram W. G., Eric A. Weingarten, and Colin R. Jackson. 2018. "The Role of the Phyllosphere Microbiome in Plant Health and Function." *Annual Plant Reviews Online*. Wiley. https://doi.org/10.1002/9781119312994.apr0614.

Strobel, Gary. 2018. "The Emergence of Endophytic Microbes and Their Biological Promise." *Journal of Fungi* 4 (2). https://doi.org/10.3390/jof4020057.

Strober, Warren, Peter J. Murray, Atsushi Kitani, and Tomohiro Watanabe. 2006. "Signalling Pathways and Molecular Interactions of NOD1 and NOD2." *Nature Reviews. Immunology* 6 (1): 9–20.

Sun, Xinhua, Dmitry Lapin, Joanna M. Feehan, Sara C. Stolze, Katharina Kramer, Joram A. Dongus, Jakub Rzemieniewski, et al. 2021. "Pathogen Effector Recognition-Dependent Association of NRG1 with EDS1 and SAG101 in TNL Receptor Immunity." *Nature Communications* 12 (1): 3335.

Sy, Abdoulaye, Antonius C. J. Timmers, Claudia Knief, and Julia A. Vorholt. 2005. "Methylotrophic Metabolism Is Advantageous for Methylobacterium Extorquens during Colonization of Medicago Truncatula under Competitive Conditions." *Applied and Environmental Microbiology* 71 (11): 7245–52.

Takahashi, Akira, Catarina Casais, Kazuya Ichimura, and Ken Shirasu. 2003. "HSP90 Interacts with RAR1 and SGT1 and Is Essential for RPS2-Mediated Disease Resistance in Arabidopsis." *Proceedings of the National Academy of Sciences of the United States of America* 100 (20): 11777–82.

Takou, Margarita, Benedict Wieters, Stanislav Kopriva, George Coupland, Anja Linstädter, and Juliette De Meaux. 2019. "Linking Genes with Ecological Strategies in Arabidopsis Thaliana." *Journal of Experimental Botany* 70 (4): 1141–51.

Tang, Yuan, Masaaki Horikoshi, and Wenxuan Li. 2016. "Ggfortify: Unified Interface to Visualize Statistical Results of Popular R Packages." *The R Journal* 8 (2): 474.

Thébault, Elisa, and Colin Fontaine. 2010. "Stability of Ecological Communities and the Architecture of Mutualistic and Trophic Networks." *Science (New York, N.Y.)* 329 (5993): 853–56.

Thevenet, Damien, Victoria Pastor, Ivan Baccelli, Andrea Balmer, Armelle Vallat, Reinhard Neier, Gaétan Glauser, and Brigitte Mauch-Mani. 2017. "The Priming Molecule β-Aminobutyric Acid Is Naturally Present in Plants and Is Induced by Stress." *The New Phytologist* 213 (2): 552–59.

Thiergart, Thorsten, Rafal Zgadzaj, Zoltán Bozsóki, Ruben Garrido-Oter, Simona Radutoiu, and Paul Schulze-Leferta. 2019. "Lotus Japonicus Symbiosis Genes Impact Microbial Interactions." *MBio* 10 (5): 15–15.

Toju, Hirokazu, Kabir G. Peay, Masato Yamamichi, Kazuhiko Narisawa, Kei Hiruma, Ken Naito, Shinji Fukuda, et al. 2018. "Core Microbiomes for Sustainable Agroecosystems." *Nature Plants* 4 (5): 247–57.

Tran, Diep T. N., Eui-Hwan Chung, Anette Habring-Müller, Monika Demar, Rebecca Schwab, Jeffery L. Dangl, Detlef Weigel, and Eunyoung Chae. 2017. "Activation of a Plant NLR Complex through Heteromeric Association with an Autoimmune Risk Variant of Another NLR." *Current Biology: CB* 27 (8): 1148–60.

Tuda, Midori, Shun-Ichiro Iwase, Khadim Kébé, Julien Haran, Jiri Skuhrovec, Ehsan Sanaei, Naomichi Tsuji, et al. 2021. "Diversification, Selective Sweep, and Body Size in the Invasive Palearctic Alfalfa Weevil Infected with Wolbachia." *Scientific Reports* 11 (1): 9664.

Ulrich, Kristina, Regina Becker, Undine Behrendt, Michael Kube, and Andreas Ulrich. 2020. "A Comparative Analysis of Ash Leaf-Colonizing Bacterial Communities Identifies Putative Antagonists of Hymenoscyphus Fraxineus." *Frontiers in Microbiology* 11 (May): 966.

Vacheron, Jordan, Guilhem Desbrosses, Sébastien Renoud, Rosa Padilla, Vincent Walker, Daniel Muller, and Claire Prigent-Combaret. 2018. "Differential Contribution of Plant-Beneficial Functions from Pseudomonas Kilonensis F113 to Root System Architecture Alterations in Arabidopsis Thaliana and Zea Mays." *Molecular Plant-Microbe Interactions: MPMI* 31 (2): 212–23.

Van de Peer, Yves, Steven Maere, and Axel Meyer. 2009. "The Evolutionary Significance of Ancient Genome Duplications." *Nature Reviews. Genetics* 10 (10): 725–32.

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): 1260-1272.e14.

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.

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

"Vegan: Community Ecology Package." n.d. Accessed October 21, 2021. https://CRAN.R-project.org/package=vegan.

Velazquez, Eric M., Henry Nguyen, Keaton T. Heasley, Cheng H. Saechao, Lindsey M. Gil, Andrew W. L. Rogers, Brittany M. Miller, et al. 2019. "Endogenous Enterobacteriaceae Underlie Variation in Susceptibility to Salmonella Infection." *Nature Microbiology* 4 (6): 1057–64.

Vergnes, Sophie, Damien Gayrard, Marine Veyssière, Justine Toulotte, Yves Martinez, Valérie Dumont, Olivier Bouchez, Thomas Rey, and Bernard Dumas. 2020. "Phyllosphere Colonization by a Soil Streptomyces Sp. Promotes Plant Defense Responses Against Fungal Infection." *Molecular Plant-Microbe Interactions: MPMI* 33 (2): 223–34.

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.

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.

Walker, Bruce J., Thomas Abeel, Terrance Shea, Margaret Priest, Amr Abouelliel, Sharadha Sakthikumar, Christina A. Cuomo, et al. 2014. "Pilon: An Integrated Tool for Comprehensive Microbial Variant Detection and Genome Assembly Improvement." *PloS One* 9 (11). https://doi.org/10.1371/journal.pone.0112963.

Wang, Yi, Zhilong Bao, Ying Zhu, and Jian Hua. 2009. "Analysis of Temperature Modulation of Plant Defense against Biotrophic Microbes." *Molecular Plant-Microbe Interactions: MPMI* 22 (5): 498–506.

Wang, Yupeng, Xiyin Wang, Haibao Tang, Xu Tan, Stephen P. Ficklin, F. Alex Feltus, and Andrew H. Paterson. 2011. "Modes of Gene Duplication Contribute Differently to Genetic Novelty and Redundancy, but Show Parallels across Divergent Angiosperms." *PloS One* 6 (12): e28150.

Wei, Fang, Brian Fanella, Liang Guo, and Xuemin Wang. 2016. "Membrane Glycerolipidome of Soybean Root Hairs and Its Response to Nitrogen and Phosphate Availability." *Scientific Reports* 6 (November): 36172.

Wei, Na, and Tia-Lynn Ashman. 2018. "The Effects of Host Species and Sexual Dimorphism Differ among Root, Leaf and Flower Microbiomes of Wild Strawberries in Situ." *Scientific Reports* 8 (1): 5195.

Weigel, Detlef, and Richard Mott. 2009. "The 1001 Genomes Project for Arabidopsis Thaliana." *Genome Biology* 10 (5): 107.

Wellner, S., N. Lodders, and P. Kämpfer. 2011. "Diversity and Biogeography of Selected Phyllosphere Bacteria with Special Emphasis on Methylobacterium Spp." *Systematic and Applied Microbiology* 34 (8): 621–30.

Wersch, Rowan van, Xin Li, and Yuelin Zhang. 2016. "Mighty Dwarfs: Arabidopsis Autoimmune Mutants and Their Usages in Genetic Dissection of Plant Immunity." *Frontiers in Plant Science* 7 (November): 1717.

Wersch, Solveig van, and Xin Li. 2019. "Stronger When Together: Clustering of Plant NLR Disease Resistance Genes." *Trends in Plant Science* 24 (8): 688–99.

Wersch, Solveig van, Lei Tian, Ryan Hoy, and Xin Li. 2020. "Plant NLRs: The Whistleblowers of Plant Immunity." *Plant Communications* 1 (1): 100016.

Williams, Simon J., Kee Hoon Sohn, Li Wan, Maud Bernoux, Panagiotis F. Sarris, Cecile Segonzac, Thomas Ve, et al. 2014. "Structural Basis for Assembly and Function of a Heterodimeric Plant Immune Receptor." *Science (New York, N.Y.)* 344 (6181): 299–303.

Wolf, Andrea J., and David M. Underhill. 2018. "Peptidoglycan Recognition by the Innate Immune System." *Nature Reviews. Immunology* 18 (4): 243–54.

Wu, Chih-Hang, Ahmed Abd-El-Haliem, Tolga O. Bozkurt, Khaoula Belhaj, Ryohei Terauchi, Jack H. Vossen, and Sophien Kamoun. 2017. "NLR Network Mediates Immunity to Diverse Plant Pathogens." *Proceedings of the National Academy of Sciences of the United States of America* 114 (30): 8113–18.

Wu, Yujun, Yang Gao, Yanyan Zhan, Hong Kui, Hongyan Liu, Li Yan, Birgit Kemmerling, Jian-Min Zhou, Kai He, and Jia Li. 2020. "Loss of the Common Immune Coreceptor BAK1 Leads to NLR-Dependent Cell Death." *Proceedings of the National Academy of Sciences of the United States of America* 117 (43): 27044–53.

Xiao, Yingjie, Haijun Liu, Liuji Wu, Marilyn Warburton, and Jianbing Yan. 2017. "Genome-Wide Association Studies in Maize: Praise and Stargaze." *Molecular Plant* 10 (3): 359–74.

Xie, Jianbo, Ying Li, Xiaomin Liu, Yiyang Zhao, Bailian Li, Pär K. Ingvarsson, and Deqiang Zhang. 2019. "Evolutionary Origins of Pseudogenes and Their Association with Regulatory Sequences in Plants." *The Plant Cell* 31 (3): 563–78.

Xiong, Chao, Ji-Zheng He, Brajesh K. Singh, Yong-Guan Zhu, Jun-Tao Wang, Pei-Pei Li, Qin-Bing Zhang, et al. 2021. "Rare Taxa Maintain the Stability of Crop Mycobiomes and Ecosystem Functions." *Environmental Microbiology* 23 (4): 1907–24.

Xiong, Wu, Yuqi Song, Keming Yang, Yian Gu, Zhong Wei, George A. Kowalchuk, Yangchun Xu, Alexandre Jousset, Qirong Shen, and Stefan Geisen. 2020. "Rhizosphere Protists Are Key Determinants of Plant Health." *Microbiome* 8 (1). https://doi.org/10.1186/s40168-020-00799-9.

Yang, Chao, Wen Li, Jidong Cao, Fanwei Meng, Yongqi Yu, Junkai Huang, Lan Jiang, et al. 2017. "Activation of Ethylene Signaling Pathways Enhances Disease Resistance by Regulating ROS and Phytoalexin Production in Rice." *The Plant Journal: For Cell and Molecular Biology* 89 (2): 338–53.

Yang, Shengming, Fang Tang, Muqiang Gao, Hari B. Krishnan, and Hongyan Zhu. 2010. "R Gene-Controlled Host Specificity in the Legume-Rhizobia Symbiosis." *Proceedings of the National Academy of Sciences of the United States of America* 107 (43): 18735–40.

Yao, Hui, Xiang Sun, Chao He, Xing-Chun Li, and Liang-Dong Guo. 2020. "Host Identity Is More Important in Structuring Bacterial Epiphytes than Endophytes in a Tropical Mangrove Forest," no. August 2019: 1–16.

Yao, Hui, Xiang Sun, Chao He, Pulak Maitra, Xing Chun Li, and Liang Dong Guo. 2019. "Phyllosphere Epiphytic and Endophytic Fungal Community and Network Structures Differ in a Tropical Mangrove Ecosystem." *Microbiome* 7 (1): 1–15.

Yi, Hankuil, and Eric J. Richards. 2007. "A Cluster of Disease Resistance Genes in Arabidopsis Is Coordinately Regulated by Transcriptional Activation and RNA Silencing." *The Plant Cell* 19 (9): 2929–39.

Yoshida, Shigenobu, Syuntaro Hiradate, Motoo Koitabashi, Tsunashi Kamo, and Seiya Tsushima. 2017. "Phyllosphere Methylobacterium Bacteria Contain UVA-Absorbing Compounds." *Journal of Photochemistry and Photobiology. B, Biology* 167 (February): 168–75.

Yu, Mei-Hui, Zhe-Ze Zhao, and Jun-Xian He. 2018. "Brassinosteroid Signaling in Plant–Microbe Interactions." *International Journal of Molecular Sciences* 19 (12): 4091.

Yuan, Minhang, Zeyu Jiang, Guozhi Bi, Kinya Nomura, Menghui Liu, Yiping Wang, Boying Cai, Jian-Min Zhou, Sheng Yang He, and Xiu-Fang Xin. 2021. "Pattern-Recognition Receptors Are Required for NLR-Mediated Plant Immunity." *Nature* 592 (7852): 105–9.

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 (102030): 102030.

Zapata, Luis, Jia Ding, Eva-Maria Willing, Benjamin Hartwig, Daniela Bezdan, Wen-Biao Jiao, Vipul Patel, et al. 2016. "Chromosome-Level Assembly of *Arabidopsis Thaliana* L *Er* Reveals the Extent of Translocation and Inversion Polymorphisms." *Proceedings of the National Academy of Sciences* 113 (28): E4052–60.

Zarraonaindia, Iratxe, Sarah M. Owens, Pamela Weisenhorn, Kristin West, Jarrad Hampton-Marcell, Simon Lax, Nicholas A. Bokulich, et al. 2015. "The Soil Microbiome Influences Grapevine-Associated Microbiota." *MBio* 6 (2): 1–10.

Zdrzałek, Rafał, Sophien Kamoun, Ryohei Terauchi, Hiromasa Saitoh, and Mark J. Banfield. 2020. "The Rice NLR Pair Pikp-1/Pikp-2 Initiates Cell Death through Receptor Cooperation Rather than Negative Regulation." *PloS One* 15 (9): e0238616.

Zhang, Baoguo, Zhihui Bai, Daniel Hoefel, Ling Tang, Zhiguang Yang, Guoqiang Zhuang, Jianzhou Yang, and Hongxun Zhang. 2008. "Assessing the Impact of the Biological Control Agent Bacillus Thuringiensis on the Indigenous Microbial Community within the Pepper Plant Phyllosphere." *FEMS Microbiology Letters* 284 (1): 102–8.

Zhao, B. Y., E. Ardales, E. Brasset, L. E. Claflin, J. E. Leach, and S. H. Hulbert. 2004. "The Rxo1/ Rba1 Locus of Maize Controls Resistance Reactions to Pathogenic and Non-Host Bacteria." *Theoretical and Applied Genetics* 109 (1): 71–79.

Zhao, Tingting, Wenhong Liu, Zhentong Zhao, Huanhuan Yang, Yufang Bao, Dongye Zhang, Ziyu Wang, et al. 2019. "Transcriptome Profiling Reveals the Response Process of Tomato Carrying Cf-19 and Cladosporium Fulvum Interaction." *BMC Plant Biology* 19 (1): 572.

Zhu, Huan, Shuyin Li, Zhengyu Hu, and Guoxiang Liu. 2018. "Molecular Characterization of Eukaryotic Algal Communities in the Tropical Phyllosphere Based on Real-Time Sequencing of the 18S RDNA Gene." *BMC Plant Biology* 18 (1): 365.

Zhu, Ying, Weiqiang Qian, and Jian Hua. 2010. "Temperature Modulates Plant Defense Responses through NB-LRR Proteins." *PLoS Pathogens* 6 (4): e1000844.

7 Acknowledgment

At the beginning, I'd like to say many thanks to my two supervisor Eric and Jane for all the possibilities to extent my scientific expertise in many ways and for you trust during my PhD. I appreciated a lot our fruitful discussion across many projects. To be the most expensive PhD candidate of both groups seems to play off in the end. Many thanks goes again to Eric, for your understanding during many 'exceptional occurrences' that happened during the time.

In addition to my supervisors, many thanks goes also to my thesis advisory committee's in Cologne (Alga Zuccaro, Angela Hancock) and Tübingen (Thorsten Nürnberger, Oliver Bossdorf).

In line with this, I'd like to thank Korbinian Schneeberger for being part of my thesis committee.

Eine unvergessliche Zeit wäre undenkbar ohne die PALS (Angels), welche die Tätigkeiten am Institut und im Privaten untrennbar machten. Eine Hommage an die Kemen vs. Tsuda *competition*. In diesem Kontext auch ein Hoch auf Alina Kuroczik (†), welche mir tatkräftig bei nächtlichen Auswertungen zur Seite stand und welche noch heute mein Handeln beeinflusst. Danke für diese Zeit! Ein dreifaches Hoch auch an ALLE Cologne – Boyz, welche im sonntägigen Gottesdienst für göttliche Klänge sorgten.

Der Umzug nach Tübingen veränderte so einiges und 800Kg Trockeneis ermöglichten es einen kühlen Kopf zu bewahren. Many thanks goes also to the Tübingen-Startup team Juliana and Alfi, which helped a lot to establish a functional lab and social activities in the south of Germany. Further, many thanks for your patience in bioinformatical support.

In addition, I 'd like to say many thanks to Vasvi for your permanent scientific feedback and funny conversations. Big thanks also too Ariane, Daniel, Maryam and Remco Stam for our fruitful collaborative projects. Many thanks also to those of you, which were involved in numerous projects and are not named so far!

Diana, dieses Werk wäre ohne deinen Zuspruch, wohl noch nicht fertig – beschten Dank! Vielen Dank für deine Unterstützung und Prost, auf eine gemeinsame Zukunft! Zu guter Letzt, möchte ich meinen Eltern danken. Vielen Dank für euer Vertrauen über all die Jahre. Zusammen haben wir ein goldenes Dreieck geschaffen. ;)

Cheers!