

Identification, Characterization, and Agricultural Uses of Precipitation-borne and Plant-associated Bacteria

Marco Enrique Mechan-Llontop

Dissertation submitted to the faculty of the Virginia Polytechnic Institute and State University in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

In

Plant Pathology, Physiology, and Weed Science

Boris A. Vinatzer, Chair

Aureliano Bombarely

David C. Haak

John M. McDowell

Mark A. Williams

20 November 2019

Blacksburg, VA

**Keywords:** Plant microbiota, rain, sources, sink, community assembly, biopesticides, ice nucleation, disease diagnosis



Identification, Characterization, and Agricultural Uses of Precipitation-borne and

## Plant-associated Bacteria

Marco Enrique Mehan-Llontop

### Abstract

Bacteria are ubiquitously present in every ecosystem on earth. While bacterial communities that reside in specific habitats, called the microbiota, have characteristic compositions, their constituents are exchanged between habitats. To understand the assembly processes and function of a microbial community in an ecosystem, it is thus important to identify its putative sources and sinks.

The sources and sinks of the plant leaf microbiome, also called the phyllosphere microbiome, are still under debate. Here, I hypothesized that precipitation is a so far neglected source of the phyllosphere microbiome. Using 16S rRNA amplicon and metagenomic sequencing, I identified the genera *Massilia*, *Sphingomonas*, *Methylobacterium*, *Pseudomonas*, *Acidiphilium*, and *Pantoea* as members of the core rain microbiome in Blacksburg, VA. Further, I used rainwater as a bacterial inoculum to treat tomato plants. I showed that rain-borne bacteria of the genera *Chryseobacterium*, *Enterobacter*, *Pantoea*, *Paenibacillus*, *Duganella*, *Streptomyces*, *Massilia*, *Shinella*, *Janthinobacterium*, *Erwinia*, and *Hyphomicrobium* were significantly more abundant in the tomato phyllosphere 7 days post-inoculation, suggesting that these rain-borne bacteria successfully colonized the tomato phyllosphere and had a direct impact on the composition of its microbiome. These results were confirmed by comparing the phyllosphere microbiota of tomato plants grown under greenhouse conditions, and thus never exposed to rain, compared to plants grown outside under environmental conditions, including precipitation.

Since a large diversity of bacteria is associated with rain, I also hypothesized that rain-borne bacteria are well adapted to environmental stresses, similar to the stressors microbial biopesticides are exposed to in the field. I thus explored rain as a source of resilient biopesticides to control fire blight, caused by the bacterial pathogen *Erwinia amylovora*, on apple. In an *in-vitro* dual culture assay, I identified rain-borne isolates displaying broad-range inhibition against *E. amylovora* and several other plant pathogens. Two rain-borne isolates, identified as *Pantoea agglomerans* and *P. ananatis*, showed the strongest inhibition of *E. amylovora*. Further experiments showed that these two *Pantoea* isolates survive under environmental conditions and have a strong protective effect against *E. amylovora*. However, protection from disease in an orchard was inconsistent, suggesting that the timing of application and formulations must be improved for field applications. Using a UV-mutagenesis screen and whole-genome sequencing, I found that a phenazine antibiotic produced by the *P. agglomerans* isolate was the likely active molecule that inhibited *E. amylovora*.

Bacterial communities are constantly released as aerosols into the atmosphere from plant, soil, and aquatic sources. When in the atmosphere, bacteria may play crucial roles in geochemical processes, including the formation of precipitation. To understand the potential role of decaying vegetation as a source of atmospheric Ice Nucleation Particles (INPs), I analyzed a historic leaf litter sample collected in 1970 that had maintained Ice Nucleation Activity (INA) for 48 years. A culture-dependent analysis identified the bacterial species *Pantoea ananatis* and the fungal species *Mortierella alpina* to have INA and to be present in the leaf litter sample. Further, I determined that both *P. ananatis* and *M. alpina* produced heat-sensitive sub-micron INPs that may contribute to atmospheric INPs.

The development of new sequencing technologies has facilitated our understanding of microbial community composition, assembly, and function. Most research in bacterial

community composition is based on the sequencing of a single region of the 16S rRNA gene. Here, I tested the potential of culture-independent 16S rRNA sequencing of the phyllosphere microbiome for disease diagnosis. I compared the community composition of the microbiome of the aerial parts of cheddar pinks (*Dianthus gratianopolitanus*) that showed disease symptoms with the microbiome of healthy plants to identify the causative agent. However, I found that the pathogen is probably ubiquitous on cheddar pinks since it was present at similar abundance levels in symptomatic as well as healthy plants. Moreover, the low-resolution of 16S rRNA sequencing did not allow to identify the pathogen at the species or strain level.

In summary, in this thesis, I found support for the hypothesis that rain is one of the sources of the phyllosphere microbiome, that rain is a promising source of biopesticides to control plant diseases in the field, that leaf litter is a source of atmospheric INPs, and that 16S rRNA sequencing is not well suited for pathogen identification in support of plant disease diagnosis. Finally, in additional research to which I contributed but that is not included in this thesis, I found that metagenomic sequencing can identify pathogens at the species and strain level and can overcome the limitations of 16S rRNA sequencing.



# Identification, Characterization, and Agricultural Uses of Precipitation-borne and Plant-associated Bacteria

Marco Enrique Mechan-Llontop

## General Audience Abstract

Bacteria are present in nearly every ecosystem on earth. Bacterial communities that reside in a specific habitat are known as microbiota and have characteristic compositions and functions that directly impact the health of ecosystems. Microbiota associated with plants, the so-called plant microbiota, play a crucial role in plant fitness. Thus, it is important to study the assembly and diversity of plant microbiota and their impact on the ecosystem. The sources of leaf microbiota remain to be elucidated. Here, I have studied the contribution of rainfall to the bacteria that live on and in plant leaves. First, using DNA sequencing, I identified the bacteria present in rainfall in Blacksburg, VA. Then, using rain as bacterial inoculum, I found that some rain-borne bacteria, including members of the genera *Pantoea*, *Massilia*, *Janthinobacterium*, and *Enterobacter*, are efficient colonizers of tomato leaves. Either absence or low abundance of rain-borne bacteria from tomato leaves never exposed to rainfall confirmed further that bacteria in rain contribute to the assembly of plant leaf microbiota. The identification of all putative sources and sinks of leaf microbiota is important when trying to manipulate them to improve plant health and crop yield. Since I found that rainfall contains many different bacteria, I also studied the potential application of rain-borne bacteria in agriculture. The main limitations of commercial bio-pesticides are their poor survival and limited efficacy in the field. Here, I speculated that rain-borne bacteria are well adapted to environmental stressors and could represent efficient bio-pesticides under field conditions. In fact, I isolated two rain-borne

bacteria from the genus *Pantoea* that strongly inhibited *Erwinia amylovora*, the causal agent of the fire blight disease of apple, in the laboratory under controlled conditions. However, I observed inconsistent results in a 2-year field trial in an orchard. Using mutagenesis and DNA sequencing, I found the active molecule that likely inhibited *E. amylovora*, in one of the rain-borne isolates. Finally, the access to newer and cheaper sequencing technologies has recently facilitated the study of bacteria at large scale. Most research of microbiota is based on the sequencing of a single region of one gene, the 16S rRNA gene. Here, I tested the potential of 16S rRNA sequencing of leaf microbiota for disease diagnosis. However, I identified the pathogen in healthy and diseased plants, suggesting its ubiquitous presence. Further, due to the low-resolution of 16S rRNA sequencing, it was impossible to identify the pathogen at the species level. In summary, I found that rain is a source that contributes to leaf microbiota, that rain is a promising source of bio-pesticides to control plant diseases, and that 16S rRNA sequencing is not recommended as a tool to diagnose plant diseases.

## **Acknowledgements:**

Completing my Ph.D. has been a dream come true and a truly life-changing experience for me. However, it would not have been possible without the support and encouragement I have received from many people.

First, I would like to express my most sincere gratitude to my advisor Dr. Boris A. Vinatzer for trusting me, motivating me, and for all his support throughout my Ph.D. study. I am thankful for his patience, guidance, and encouragement he provided me to become a better scientist. I have been fortunate enough to have not only a great advisor but an amazing friend. I would also like to thank all the members of my committee, Dr. David Haak, Dr. John McDowell, Dr. Aureliano Bombarely and Dr. Mark Williams for their advice in finishing this thesis, and for their career advice.

I am thankful to the entire Vinatzer lab members, who I had the opportunity to work during all these years Dr. Rongman Cai, Dr. Christopher Clarke, Haijie Liu, Dr. Kevin Failor, Dr. Noam Eckshtain-Levi, Marcela Aguilera, Parul Sharma, and Shu Yang. They have all been great colleagues, collaborators, and wonderful friends.

I would like to thank the Translational Plant Science and the Plant, Pathology and Weed Science communities for creating a supportive and friendly environment for their members, especially to international students. Additionally, I would like to thank Dr. Pedro J. Chimoy Effio, my undergraduate advisor for introducing me to academic research and helping me during my application to Graduate School.

Finally, I would like to thank my family for their love, patience and support all these years we have been living apart, my wife Leyla, my son Marco Lucca and Maria Gracia. I would like to thank my parents, Marcos Mehan and Teresa Llontop for giving me love, guidance and education to be a better person. And finally, I would like to thank Mery, Jorge, and Jessica for being the best siblings. I hope that one day I can return what all of them have done for me.

## Table of contents

Acknowledgements.....	vii
-----------------------	-----

### Chapter One : Literature review: Sources and Sink of the Phyllosphere microbiota

Abstract .....	1
Keywords.....	1
Glossary.....	2
Graphical abstract.....	2
Introduction to the plant-associated microbiota.....	3
The phyllosphere microbiome.....	4
The seed microbiome: Vertical transmission of the microbiome.....	6
Horizontal transmission.....	8
The soil and root microbiomes.....	8
Microbes above the plant surfaces: air and rain associated-microbiomes.....	10
Long-distance microbial movement.....	12
Interconnection of habitats through the movement of microbes.....	14
Conclusions and future directions.....	14
References.....	16
<b>Chapter two: An investigation of rain as reservoir of tomato phyllosphere microbiota.....</b>	<b>26</b>
Abstract.....	27
Introduction.....	28
Materials and methods.....	30
Determination of the population size of tomato phyllosphere microbiota.....	30
Culture-dependent analysis of phyllosphere diversity.....	31
Rain collection.....	31
Rain as bacterial inoculum and plant treatments.....	32
Harvesting.....	32
Microbial DNA extraction.....	32
Library preparation and sequencing.....	32
Bioinformatic analysis.....	34
Results.....	35

Comparison of the bacterial population size in the phyllosphere of tomato plants exposed to rain compared to plants not exposed to rain.....	35
Characterization of rain-borne microbiota in Blacksburg, Virginia.....	36
Comparison of the rain-borne microbiota with microbiota of plants either treated with concentrated rain microbiota, sterile rain, or double-distilled water.....	37
Identification of the bacterial rain-borne taxa that are the most efficient tomato leaf colonizers under controlled laboratory conditions.....	38
Phyllosphere microbiota of tomato plants never exposed to rain.....	39
Identification of the bacterial rain-borne taxa that are the most efficient tomato leaf colonizers under controlled laboratory conditions using a metagenomic approach.....	41
Discussion.....	41
References.....	47
Tables.....	55
Figures.....	60
Supplementary figures.....	70
<b>Chapter three: Exploring rain as source of biological control agents for fire blight on apple...</b>	<b>73</b>
Abstract.....	74
Introduction.....	75
Materials and Methods.....	77
In vitro screening for inhibition of <i>E. amylovora</i> growth.....	77
Survival assay under environmental conditions.....	78
Assay of inhibitory activity against <i>E. amylovora</i> on detached apple blossoms.....	78
Disease suppression assays on attached blossoms in an orchard .....	79
UV mutagenesis screen to identify the molecular basis of growth inhibition.....	80
Genome sequencing.....	80
Genome analysis.....	81
Analysis of blossom microbiomes.....	82
Results.....	83
Inhibitory activity of rain-isolated bacteria in-vitro.....	83
Bacterial survival on appl trees.....	83
Fire blight control on detached apple blossom.....	84
Fire blight control in an apple orchard.....	85
Whole genome sequencing and genome-based identification.....	86

Prediction of biosynthetic gene clusters (BGCs) and a UV-mutagenesis screen to identify genes at the basis of the observed inhibitory activity.....	87
Discussion.....	88
Conflict of Interest.....	95
Author Contributions.....	95
Funding.....	96
Acknowledgments.....	96
Data Availability Statement.....	96
References.....	96
Figures.....	106
Tables.....	111
Supplementary tables.....	116
Supplementary figures.....	121
<b>Chapter four: Comprehensive characterization of an Aspen (<i>Populus tremuloides</i>) leaf litter sample that maintained ice nucleation activity for 48 years.....</b>	<b>123</b>
Abstract.....	124
Introduction.....	125
Materials and Methods.....	127
Description of leaf litter sample 70-S-14 and how it has been stored.....	127
Characterization of the cumulative ice nucleation spectrum of leaf litter sample 70-S-14...	128
Culture-independent microbial community analysis.....	129
Culture-dependent bacterial community analysis.....	130
INA testing of individual bacterial colonies.....	131
INA testing of individual fungal colonies.....	131
Results.....	132
Aspen leaf litter 70-S-14 contains mostly INPs of submicron size that are heat-sensitive....	132
Microbial population analysis of the Aspen leaf litter 70-S-14 sample reveals few known microbial taxa with INA.....	133
INA testing of bacterial and fungal colonies reveals presence of INA strains belonging to the bacterial species <i>P. syringae</i> and <i>Pa. ananatis</i> strains and the fungal species <i>M. alpina</i> .....	134
Discussion and Conclusions.....	136
Acknowledgements.....	139

References.....	140
Figures.....	145
<b>Chapter five: Assessing the potential of culture-independent 16S rRNA microbiome analysis in disease diagnostics: the example of <i>Dianthus gratianopolitanus</i> and <i>Robbsia andropogonis</i>.</b> .....	<b>149</b>
Abstract.....	150
Introduction.....	151
Materials and Methods.....	153
Plant samples and plant inoculations.....	153
DNA extraction.....	153
16S rRNA amplicon sequencing.....	154
Community composition analysis based on 16S rRNA sequences.....	154
Results.....	155
Analysis and comparison of the microbiome composition of cheddar pink nursery plants.....	156
Relative abundance of the confirmed pathogen <i>R. andropogonis</i> .....	157
Analysis and comparison of the diversity of cheddar pink microbiomes.....	158
The microbiome of cheddar pink plants artificially inoculated with <i>R. andropogonis</i> .....	160
Discussion.....	160
Funding .....	165
Compliance with Ethical Standards.....	165
Conflict of Interest.....	165
References.....	165
Tables.....	170
Figures.....	172

## List of Figures

Graphical abstract. Ecosystems are interconnected by cycling microorganisms. Putative pathways of microbial dispersal that originate the phyllosphere microbiota assembly.....	2
Figure 1. Bacterial population size in the phyllosphere of tomato plants A) grown in a plastic pot at the Kentland research farm compared to plants grown under laboratory conditions B) grown on the roof of a research building compared to plants grown under laboratory conditions, and C) grown under laboratory conditions treated with rain and autoclaved rain.....	60
Figure 2. Relative abundance of bacterial taxa in rainfall collected in Blacksburg, VA. at the class level (abundance >1%).....	61
Figure 3. Alpha diversity measurements for rain compared to treated plants at day 0 vs day 7. A) Rarefaction curves showing the richness of samples, and B) three measures of alpha diversity (observed OTUs, Shannon diversity index, and Simpson diversity index).....	62
Figure 4. Principal coordinates analysis (PCoA) derived from the dissimilarity matrix of Bray-Curtis is shown for all rain samples and plants treated at day 0 and day 7.....	63
Figure 5. Relative abundance of bacterial taxa of all rain samples and plants treated at day 0 and day 7 at the phylum level (abundance > 1%).....	64
Figure 6: Operational Taxonomic Unit (OTU) differential abundance. A) differences at the OUT level between plants treated with 100X-rain at the day 0 and day 7 time points, B) differences at the OUT level between rain and tomato plants treated with 100X- rain at the day 7 time point.....	65
Figure 7. Relative abundance of rain-borne bacterial taxa that either failed or succeed in colonizing the tomato phyllosphere under controlled laboratory conditions. A) <i>Acidiphilium</i> , B) <i>Flavobacterium</i> , C) <i>Duganella</i> , D) <i>Janthinobacterium</i> , E) <i>Pantoea</i> , F) <i>Massilia</i> , G) <i>Pseudomonas</i> , and H) <i>Bacillus</i> .....	66
Figure 8. Relative abundance of bacterial taxa of all tomato plants grown in a hydroponic system, tomato plants grown in soil (both in a commercial greenhouse) and tomato plants grown outside (in a roof of a research building) at the phylum level (abundance > 1%).....	67



Figure 9. Alpha diversity measurements for tomato plants grown in a hydroponic system, in soil and outside A) Rarefaction curves showing the richness of samples, and B) three measures of alpha diversity (observed OTUs, Shannon diversity index, and Simpson diversity index).....	68
Figure 10. Abundance of rain samples and tomato plants treated with 100X-rain at time point 0 and 7 days post inoculation using metagenome shotgun sequencing.....	69
Supplementary figure 1. Differences at the Operational Taxonomic Unit (OTU) level between rain and tomato plants treated with sterile- water at the day 0 time point.....	70
Supplementary figure 2. Differences at the Operational Taxonomic Unit (OTU) level between rain and tomato plants treated with Filtered rain at the day 0 time point.....	71
Supplementary figure 3. Differences at the Operational Taxonomic Unit (OTU) level between rain and tomato plants treated with 100X-rain at the day 0 time point.....	72
Figure 1. Dual culture screening assay showing the inhibitory activity of rain-isolated bacteria against <i>E. amylovora</i> represented by A) the halo inhibition where the pathogen did not growth, and B) average size of the inhibition zone induced by BAV2934 and BAV3296.....	106
Figure 2. Survival, measured as population size, of rain-isolated bacteria BAV2934 and BAV3296 compared with <i>E. amylovora</i> on detached apple branches under environmental conditions for up to 12 days. An <i>Escherichia coli</i> isolate was included as control.....	107
Figure 3. Survival of rain-isolated <i>Pantoea agglomerans</i> BAV2934 on apple blossoms in an orchard. A) Culturable bacteria, and B) Relative abundance of bacterial taxa at the genus level (16S rRNA gene) present in blossoms up to 10 days after inoculation with BAV2934.....	108
Figure 4. Antagonistic effect of rain-isolated bacteria BAV2934 and BAV3296 against <i>E. amylovora</i> on detached blossoms under laboratory conditions. A) Disease severity measured by the necrotic tissue caused by <i>E. amylovora</i> invasion on the pedicel, B) Pathogen population size on the receptacle, and C) Pathogen population size on the pedicel 7 days after pre-treatment with either rain-isolated bacteria or commercial products.....	109

Figure 5. Fire blight control on blossoms in an apple orchard in 2018. A) Disease incidence of Fire blight on ‘Golden Delicious’ and B) Disease incidence of Fire blight on ‘York’.....	110
Supplementary Figure S1. Bacterial survival rate of rain-isolated BAV2934, BAV3296 and <i>E. amylovora</i> BAV5616 on apple branches under environmental conditions in A) November 2016 and B) March 2017 under environmental conditions.....	121
Supplementary Figure S2. Identified BGC predicted to encode a phenazine compound. The red dashed box indicates the gene cluster identified by antiSMASH. The black vertical line indicates the location of the base pair mutation. JGI IMG base pair coordinates are indicated.....	122
Supplementary Figure S3. PCR detection of contig 4 in BAV2934. A) Contig 4 gene NOOGOKNH_04505, B) Contig 4 gene NOOGOKNH_04435, C) Contig 4 gene NOOGOKNH_04535, D) Contig 4 gene NOOGOKNH_04609, E) Chromosomal gene NOOGOKNH_2217, and F) Chromosomal gene NOOGOKNH_02219.....	123
Figure 1. Characterization of INA of Aspen leaf litter 70-S-14.....	145
Figure 2. Cumulative ice nucleation spectra of the Aspen leaf litter sample 70-S-14.....	146
Figure 3. Analysis of the composition of the microbial communities present in Aspen leaf litter sample 70-S-14.....	147
Figure 4. Cumulative ice nucleation spectra of the <i>M. alpina</i> strain LL118 (A) isolated from leaf litter sample 70-S-14 and of <i>Pa. ananatis</i> BAV 3057 (B) isolated from rain, but over 99% identical in its 16S rRNA sequence to the two <i>Pa. ananatis</i> isolates from sample 70-S-14.....	148
Figure 1. Relative abundance of bacterial taxa. Relative abundance of bacterial taxa at the genus level is shown for symptomatic and asymptomatic cheddar pink plants from nursery 1 during the disease outbreaks in 2015, plants from nursery 2 without a disease outbreak in 2016, and plants from both nurseries in the absence of disease in 2018.....	173

Figure 2. Alpha diversity. Three measures of alpha diversity (observed OTUs, Shannon diversity index, and Simpson diversity index) for nursery 1 during the disease outbreak in 2015, nursery 2 in 2016 (no disease), and both nurseries in 2018 (no disease).....	174
Figure 3. Rarefaction curves. A Rarefaction curves for nursery 1 during the disease outbreak in 2015, nursery 2 in 2016 (no disease), and both nurseries in 2018 (no disease). B rarefaction curves for each individual plant at nursery 1 during the disease outbreak in 2015.....	175
Figure 4. Principal Coordinates Analysis. Principal Coordinates Analysis based on weighted Unifrac distances is shown for all samples from nursery 1 during the disease outbreak in 2015, nursery 2 in 2016 in the absence of disease, and both nurseries in 2018 in the absence of disease.....	176
Supplementary Figure S1. Rarefaction curves of all individual samples drawn from raw data.....	177
Supplementary Figure S2. <i>Dianthus gratianopolitanus</i> plant with leaf spotting symptoms after artificial inoculation.....	178

## List of table

Table 1. Identity of 9 rain-isolated bacteria that showed antagonistic activity against *E. amylovora* in our initial inhibition screen

Table 2. Inhibition assay of rain-isolated bacteria against bacterial, fungal and oomycetes plant pathogens

Table 3. BAV2934 UV-generated non-synonyms mutations

Table 4. Biosynthetic gene clusters predicted in BAV2934 by antiSMASH.

Table S1. List of rain-isolated bacteria tested in the initial inhibition assay.

Table S2. Antagonistic activity of rain-isolated bacteria against plant pathogens.

Table S3. Bacterial survival assay on apple branches under environmental conditions.

Table S4. Weather conditions during field treatments at Kentland Farm, VA.

Table S5. Infection (% infected clusters) and control (% reduction compared to MgSO<sub>4</sub>) in the field.

Table 1. Plant samples analyzed in this study:

Table 2. Diversity indices computed for each individual nursery plant sample:

## **Chapter One: Literature review**

### **Sources and sinks of phyllosphere microbiota**

#### **Abstract**

Microbes colonize nearly every single environment on earth. These microbial communities, called microbiota, are not immobile entities but their members are exchanged between environments. Interconnecting pathways of microbial dispersal between environments are poorly explored and identifying which environments act more as source than as sink and which environments act more as sink than as source is difficult to discern. Plants host microbiota of relatively stable composition but of high genetic diversity that are crucial for plant resilience and health. Current research is focused on discovering the composition of plant microbiota, which biotic and abiotic factors influence their genetic and genomic composition, *i.e.*, the microbiome structure, and how to manipulate them for agricultural improvements. Knowledge of the identity of the sources and sinks of plant microbiota is still limited. However, a better understanding of these sources and sinks is crucial to successfully manipulate plant microbiota and improve plant health. Here, we review existing literature on known sources and sinks that may play a significant role in the assembly of plant microbiota, with a focus on the phyllosphere compartment. Finally, we suggest an experimental framework to gain further insight into sources and sinks of plant microbiota.

**Keywords:** Phyllosphere, microbiome, assembly, microbial dispersal

## Glossary:

Microbiota: the microorganisms that inhabit a specific environment.

Microbiome: the collection of genomes or genes of microbiota.

Microbial dispersal: Process of movement of microbes from one location to another.

Ecosystem: the organisms living in a specific environment and the physical environment in which they live.

Disturbance: an event or process that affects an ecosystem.

Driver: a factor that induces a change in an ecosystem.

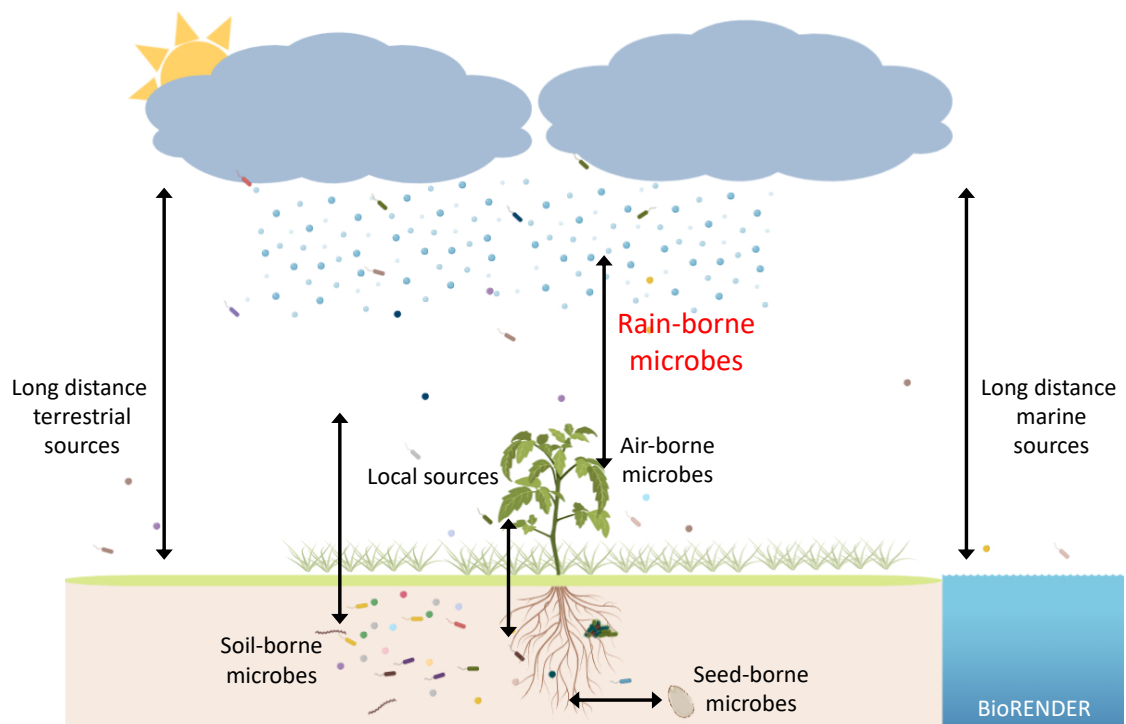
Habitat: an environment where organisms live.

Source: the habitat from where microbes originate before dispersal.

Sink: the habitat where microbes are deposited after dispersal.

Vertical transmission: transmission of microbes from parents to offspring.

Horizontal transmission: transmission of microbes between environments that exist at the same time.



**Graphical abstract:** Ecosystems are interconnected by cycling microorganisms. Putative pathways of microbial dispersal that originate the phyllosphere microbiota assembly.

## **1. Introduction to the plant-associated microbiota:**

Each plant compartment harbors structured microbial communities (Bulgarelli et al. (2012) (Links et al., 2014). However, these non-random, well-balanced plant microbiota are strongly affected by biotic and abiotic factors, including plant developmental stage (Chaparro et al., 2014), genotype (Bodenhausen et al., 2014), nutrient availability (Berg & Koskella, 2018), environmental conditions (Kadivar & Stapleton, 2003), and the soil in which plants grow (Tkacz et al., 2015).

Plants are surrounded by several habitats, each inhabited by specific microbiota, including soil, air, and many multicellular organisms, such as insects, nematodes, and other animals. These habitats are interconnected with each other by the movement of microbes from one habitat to the other, generally referred to as horizontal transmission (van Bruggen et al., 2019). Plants also inherit microbes from the previous generation present on seeds, which is referred to as vertical transmission.

Plants can discriminate between pathogenic, beneficial, and neutral microbes (Vogel et al., 2016) through the recognition of Microbe Associated Molecular Patterns (MAMPs) by diverse Pattern Recognition Receptors (PRRs) (Hacquard et al., 2017) that induce a first layer of the plant immune system (MAMP-triggered Immunity – MTI). However, plant-associated microbes either evade recognition or overcome this first barrier through the injection of effector proteins into plant cells (Jones & Dangl, 2006, Berrabah et al., 2018). The recognition of these microbial effectors induces a second layer of defense known as Effector Triggered Immunity (ETI) (Cui et al., 2015). Hacquard et al. (2017) illustrated two additional barriers: 1) competition with native microbiota, and 2) different host physical barriers. Overall these barriers are able to maintain an equilibrium in the plant microbiome, unless a pathogen overcomes them all.

Considering that microbes cycle through different habitats and that environmental disturbances regularly occur, such as changes in temperatures and humidity, the following questions arise: 1) What microbial sources drive plant microbiome assembly? 2) to what degree do foreign microbes influence the restoration of the environment at the micro and macro levels? Our knowledge about the sources

and reservoirs of microbes that determine the assembly of the plant microbiome is relatively limited. While there is strong evidence that plant roots recruit the root-associated microbiome from the surrounding soil (Fitzpatrick et al., 2018), the sources that drive the phyllosphere microbiome assembly are still under debate (Grady et al., 2019).

The purpose of this review is to 1) describe the microbial sources of the phyllosphere microbiome proposed in the published literature, 2) describe what is known about the role of the phyllosphere microbiome as source of other microbiomes, in particular, the atmosphere, soil, and seeds, 3) discuss the general importance of understanding the interconnection between habitats and the movement of microbes to better understand phyllosphere microbiome dynamics, and 4) describe research that is needed to further our understanding of the sources and sinks of the phyllosphere microbiome. However, before addressing these topics, we need to introduce the phyllosphere microbiome itself.

## **2. The phyllosphere microbiome:**

The phyllosphere (Ruinen, 1961) is the microenvironment that extends from the outside to the inside of the plant leaf (Morris, 2002, Vacher et al., 2016). This habitat is exposed to several environmental stresses, including water scarcity, osmotic stress, temperature extremes, fluctuations in humidity, UV exposure, and precipitation (Jacobs & Sundin, 2001). Despite these stresses, the phyllosphere is colonized by a dense population of microorganisms, with bacteria being the most abundant domain (Vorholt, 2012). However, because of the fluctuating environmental conditions, the phyllosphere microbiome is more ephemeral compared to the root microbiome, which is exposed to a more stable soil environment (Hacquard et al., 2015).

Moreover, besides the already listed environmental stresses, upon arrival on the plant surface, bacteria must also deal with additional challenges, such as nutrient limitation, various microclimatic conditions, and multilayered surfaces (Morris, 2002). For example, changes in the cuticle, the most



external leaf surface for microbial attachment and a protective barrier, affect not only the colonization of specific bacteria (Marcell & Beattie, 2002) but the abundance and composition of the total bacterial microbiome (Ritpitakphong et al., 2016). Also changes in the availability of water in the phyllosphere, induced by pathogen invasion, have been shown to influence the composition of the phyllosphere microbiome in *A. thaliana* (Hernandez & Lindow, 2019). Similarly, the availability of nutrients in tomato has been shown to have a direct impact on the composition of the bacterial community (Berg & Koskella, 2018). Finally, the phyllosphere microbiome composition is also influenced by biotic factors, including plant age and genotype (Wagner et al., 2016).

The result of all the listed factors is that the phyllosphere microbiome, compared to other plant microbiomes, such as the rhizosphere microbiome, seems to be composed of fewer bacterial taxa. The most abundant taxa of the phyllosphere microbiome are Proteobacteria, Actinobacteria, Bacteroidetes, and Firmicutes (Bulgarelli et al., 2013). At the genus level, *Pseudomonas*, *Sphingomonas*, *Methylobacterium*, *Bacillus*, *Massilia*, *Arthrobacter*, *Flavobacterium*, and *Pantoea* define the phyllosphere core microbiome of perennial grasses (Grady et al., 2019)). In contrary, the rhizosphere core microbiome of common beans, for example, grown in agricultural soil is dominated by the bacterial taxa *Rhizobium*, followed by *Chitinophaga*, *Dyadobacter*, *Arthrobacter*, *Streptomyces*, *Caulobacter*, *Mesorhizobium*, *Bosea*, *Devosia*, *Phenylobacterium*, *Novosphingobium*, *Sphingomonas* and *Bradyrhizobium* (Pérez-Jaramillo et al., 2019). Moreover, it has been observed that soil type and plant genotype affect the core rhizosphere microbiome composition (Pérez-Jaramillo et al., 2019, Simonin et al., 2019).

Next, we will describe the potential sources that may play a role in the assembly of the phyllosphere microbiome and how these sources are interconnected by dispersal mechanisms.

### 3. The seed microbiome: Vertical transmission of the microbiome:

Although seeds harbor a microbiome that is less diverse than that of the phyllosphere, the seed microbiome plays a crucial role in seed germination and protection against pathogens (Barret et al., 2015). Seeds may acquire their microbiota through 1) non-vascular or xylem tissues in the maternal plant, 2) the stigma of the maternal plant, and 3) contamination from the surrounding environment (Maude, 1996). Thus, the composition of seed-associated microbiota can be expected to differ between seed compartments: the embryo, the endosperm, and the seed coat (Shade et al., 2017) (Eyre et al., 2019).

Seed endophytes are considered to be directly transmitted vertically from parents to offspring (Johnston-Monje & Raizada, 2011) (Hardoim et al., 2012). Abundance and endophyte diversity differ between genotypes within the same plant species (Walitang et al., 2018). It has been hypothesized that an extreme microbial bottleneck occurs in individual seeds, with none or a single endophyte colonizer cell per seed (Newcombe et al., 2018). Although the authors in effect found none or just one colonizer per seed, either a bacterial or fungal isolate, isolation was limited to a culture-dependent approach on PDA medium. However, Lundberg et al. (2012), in agreement with this bottleneck, found no endophytic bacteria on surface-disinfected *A. thaliana* seeds using culture-dependent and culture-independent approaches. However, in a different study, a large diversity of seed-associated bacterial endophytes including *Sinorhizobium*, *Micrococcus*, *Sphingomonas*, *Rhizobium*, *Acidovorax*, *Variovorax*, *Methylobacterium*, *Bacillus*, and *Staphylococcus* were isolated from surface-sterilized *A. thaliana* seeds (Truyens et al., 2013). These opposite results may be the result of the different sterilization protocols used in the two studies. In the first study, seeds were surface-sterilized with ethanol and NaClO solution for over 10 minutes, while seeds were rinsed for only one minute in the second study.

In a different study, *Paenibacillus*, *Acidovorax*, and *Pantoea* isolates were found exclusively in the endophytic compartment of seeds of rice plants, while *Bacillus*, *Curtobacterium*,

*Methylobacterium*, *Sphingomonas*, *Xanthomonas*, and *Micrococcus* were found both inside and on the seed surface (Mano et al., 2006). Similar results were obtained by Hardoim et al. (2012), where endophytic *Pseudomonas*, *Stenotrophomonas*, *Ochrobactrum*, *Sphingomonas*, *Flavobacterium*, *Paenibacillus*, *Agromyces*, *Curtobacterium*, *Frigoribacterium*, *Microbacterium*, *Mycobacterium*, and *Plantibacter* were associated with rice seeds. Studies in maize revealed *Clostridium*, *Paenibacillus*, *Enterobacter*, *Methylobacteria*, *Pantoea*, and *Pseudomonas* as the seed-associated core microbiome (Johnston-Monje & Raizada, 2011). Endophytic *Bacillus*, *Staphylococcus*, *Paenibacillus*, *Micrococcus*, and *Methylobacterium* have been isolated from bean seeds (López-López et al., 2010).

Culture-independent approaches instead found a much more diverse endophytic bacterial population. For examples, the core microbiome of *Brassica napus* was found to consist of members of Alphaproteobacteria, Betaproteobacteria, Gammaproteobacteria, Bacilli, and Actinobacteria (Rybakova et al., 2017). Additionally, culture-independent approaches revealed the diversity of endophytic seed populations to be affected by seed maturation (Mano et al., 2006) and abiotic conditions (Truyens et al., 2013).

Although the seed coat as an external layer is constantly in contact with the surrounding environment, it carries a characteristic microbiome with a protective function against pathogens (Cottyn et al., 2009, Links et al., 2014). For example, *Coryneform*, *Pantoea*, *Pseudomonas* *Bacillus*, *Burkholderia*, *Enterobacter*, *Paenibacillus*, *Staphylococcus*, and *Xanthomonas* are associated with the rice seed surface (Cottyn et al., 2009) (Mano et al., 2006) (Midha et al., 2016). *Pantoea*, *Massilia*, *Porphyrobacter*, *Pseudomonas*, *Pyrenophora*, *Sphingobium*, *Sphingomonas*, *Telluria*, and *Xanthomonas* have been observed on both *Triticum* and *Brassica* seed surfaces (Links et al., 2014).

Might the seed microbiome be a source of the phyllosphere microbiome? Several studies found that seed microbial communities may in fact represent the first microbial inoculum for the plant microbiome (López-López et al., 2010, Hardoim et al., 2012, Mitter et al., 2016). Further, some members of the seed-associated microbiome may even impact the surrounding soil (Hardoim et al.,

2012). However, as roots develop, the soil may have a stronger impact on the plant microbiome than the seed microbiome (Kristin & Miranda, 2013). Interestingly though, Mano et al. (2007) found that endophytic seed microbial communities are more similar to phyllosphere-associated microbial communities than to the soil microbiome.

#### **4. Horizontal transmission:**

##### **4.1. The soil and root microbiomes:**

Soil is considered the largest reservoir of microbial diversity on earth (Fierer et al., 2012). This microbial diversity may be capable of influencing biogeochemical processes and has a direct impact on soil characteristics. Soil structure creates many micro-environments that have distinct microbial compositions (Fierer, 2017) adapted to a nutrient-limited environment. As in every other ecosystem, abiotic and biotic disturbances limit microbial growth and survival (Bahram et al., 2018). Microbes in soil may adapt to unfavorable conditions by forming spores or by entering a dormancy stage. Although most of the microbial diversity in soil remains to be described (Ramirez et al., 2014), the major bacterial taxa in soil include Acidobacteria, followed by Actinobacteria, Bacteroidetes, Proteobacteria, and Verrucomicrobia (Ramirez et al., 2014). The soil microbiota also plays a crucial role in plant growth and health (Weller et al., 2002) (Mendes et al., 2011) (Schlatter et al., 2017).

The rhizosphere, the interface region between the soil and the plant root, hosts complex and dynamic microbiota derived from the surrounding soil. Plant roots influence directly the rhizospheric compartment by extracellular depositions (Zhalnina et al., 2018), which facilitate microbial colonization (Hu et al., 2018). Compared with soil, fewer bacterial taxa dominate the rhizosphere, suggesting a selection process that drives the microbiome assembly (Bulgarelli et al., 2012, Lundberg et al., 2012). The rhizosphere microbiome composition further depends on soil type, plant age, genotype, and root morphology (Walters et al., 2018). Recruited microbes are key players improving

plant fitness (Kwak et al., 2018). Proteobacteria followed by Actinobacteria, Bacteroidetes and Firmicutes are enriched in the rhizosphere compartment compared to bulk soil (Kwak et al., 2018).

A subset of the surrounding microbes can migrate and inhabit the endophytic root compartment (Gaiero et al., 2013). It has been shown that, in addition to root exudates signaling (Zhalnina et al., 2018) and the plant immune system (Lebeis et al., 2015), bacterial genetics (Cole et al., 2017) determine internal root colonization. Proteobacteria followed by Actinobacteria are the dominant endophytic phyla suggesting that these bacterial taxa may intimately interact with plants. Several bacterial families are significantly enriched in the endophytic compartment compared with the surrounding soil (Lundberg et al., 2012).

There is strong evidence that soil represents a starting inoculum for the assembly of the root microbiome (Bulgarelli et al., 2012). However, microbial colonization is modulated by the plant immune system and soil nutrient composition (Lebeis et al., 2015) (Castrillo et al., 2017).

Is soil also a source of microorganisms that drive the microbiome assembly of the above-ground part of plants? Are at least some soil-associated microbes also adapted to survive and colonize the phyllosphere compartment? Some research has suggested that soil is a major source that drives the assembly of the phyllosphere microbiome. For example, Grady et al. (2019) observed an overlap of bacterial taxa in soil and the phyllosphere of perennial grasses. Similar results have been found in sugarcane (Hamonts et al., 2018), grapevine (Zarraonaindia et al., 2015), mustard (Wagner et al., 2016), and *A. thaliana* (Bai et al., 2015) suggesting that phyllosphere microbiota originate from soil. Although soil and leaves share many bacterial communities, their microbiome structure is different suggesting an evolutionary adaptation to colonize these habitats.

However, all this evidence has been limited by sequencing only the 16S rRNA gene, which low resolution does not allow to identify bacteria at the species level and within species (Mechan-Llontop et al., 2019). Due to bacterial diversity within species, this approach makes it impossible to conclude if the same bacterial strains found in soil are also present in leaves or if one or more strains

of one species are found in soil but a different set of strains of the same species are found in the phyllosphere. Further, since field-grown plants were used in these studies, additional putative sources, such as aerosols, may include the same species as soil, and therefore, also constitute one of the possible sources of the phyllosphere microbiome.

Microbes from the surrounding soil are attracted and repelled by root exudates. In association with the rhizospheric compartment, microbes compete for nutrients and space while playing a role in plant fitness. Further, some recruited microbes may have co-evolved within the plant host to colonize the endophytic root compartment in a mutualistic interaction. How the above-ground plant microbiota assemble is still unclear.

#### **4.2. Microbes above the plant surfaces: air and rain associated-microbiomes:**

The Earth's surface releases a large number of microbial cells into the atmosphere as aerosol (Morris et al., 2014). While ubiquitous in the lower atmosphere (Haas et al., 2013), clouds (Kourtev et al., 2011), and the higher atmosphere (Burrows et al., 2009), metabolically active air-borne microbes are subjected to diverse environmental stresses, including acidic pH, the presence of toxic compounds, high UV light exposure, and extreme temperatures. (Amato et al., 2007).

Despite these adverse conditions, some bacteria can survive long periods of time and travel long distances in the atmosphere before being deposited to the Earth surface through precipitation (Burrows et al., 2009). Some research suggests that one role that bacteria may play in the atmosphere is to facilitate the formation of precipitation. It has been hypothesized that air-borne bacteria may produce ice nucleating particles (INPs) as a mechanism to return to land surfaces (Morris et al., 2014). Ice Nucleation Active (INA) bacteria present in the atmosphere belong to the genera *Pseudomonas*, *Pantoea*, *Xanthomonas*, *Erwinia*, *Stenotrophomonas*, and *Lysinibacillus*, and there are at least two different mechanisms used by these bacteria to catalyze ice formation (Failor et al., 2017).

However, bacterial diversity in the atmosphere is not limited to INA bacteria, but includes a large number of diverse beneficial and pathogenic bacteria (Griffin, 2007) (Polymenakou, 2012) that move locally and globally (Smith et al., 2013) (Weil et al., 2017). Further, it has been observed that bacteria in the atmosphere are not randomly dispersed (Cáliz et al., 2018, Aho et al., 2019, Aho et al., 2019). At the phylum level Proteobacteria, Cyanobacteria, Bacteroidetes, Actinobacteria, Firmicutes, and Acidobacteria are ubiquitous and, at lower taxonomic ranks Alphaproteobacteria, Betaproteobacteria, and Gammaproteobacteria are the dominant bacterial groups (Cáliz et al., 2018, Aho et al., 2019). A core microbiome has been identified that includes *Sphingomonas*, *Methylobacterium*, *Massilia*, *Pseudomonas*, *Polaromonas*, *Acidiphilium*, *Ramlibacter*, *Mucilaginibacter*, *Hymenobacter*, and *Noviherbaspirillum* (Cáliz et al., 2018, Aho et al., 2019), suggesting that these genera may include strains that are specifically adapted to the environmental stresses typically found in the atmosphere (Weil et al., 2017).

Although air-borne microbes originate from the soil, plants, water sources, and other surfaces on earth (Amato et al., 2017), plants are likely the major source of air-borne bacteria (Šantl-Temkiv et al., 2018). As we have described above, bacteria in the atmosphere and the phyllosphere are exposed to similar environmental stresses, suggesting the ability of these bacteria to be adapted to colonize both environments. For example, bacteria in the phyllosphere are UV-tolerant, non-spore-forming and heterotroph and similar characteristics are observed in metabolically active bacteria in the atmosphere. However, these bacterial features differ from those found below ground (Hirano & Upper, 2000).

Is it possible that wet deposition of bacteria, during rain fall, and dry deposition of bacteria, in the absence of rain fall, are important sources of the phyllosphere microbiome? On the other hand, what is the relative importance of plants as sources of bacteria in the atmosphere considering that bacteria in the phyllosphere and atmosphere are exposed to similar stresses but soil bacteria are not? It was found that bacteria are removed from leaves during precipitation (Lindemann & Upper, 1985), but less attention has been given to the idea that rain-associated bacteria colonize plant surfaces.

Morris et al. (2008) provided the first evidence in support of the hypothesis that precipitation is a source of the phyllosphere microbiome when studying the movement of *Pseudomonas syringae* between the atmospheric and the phyllosphere compartment. This result was later confirmed by (Monteil et al., 2016) by identifying that leaf-associated *P. syringae* pathogenic strains and precipitation-associated strains belong to the same population. Further, Maignien et al. (2014), showed that air-borne bacteria represent an important inoculum for the phyllosphere microbiome assembly in *A. thaliana* plants. Similar, Ottesen et al. (2016) found that the phyllosphere microbiome of tomato originates from airborne deposition. However, plant selection processes may determine what bacteria, even between closely related taxa, successfully establish in the phyllosphere.

Besides the work of Cindy Morris and her team studying the dispersal of *P. syringae* through precipitation and over long distances using culture-dependent approaches, most studies have been limited to the use of a microbial profile approach by sequencing the 16S rRNA gene. Similar to what we pointed out above about testing the hypothesis that soil is a possible source and sink of the phyllosphere microbiome, it will be necessary to use strain-level resolution metagenomics to precisely determine the link between the atmosphere and the phyllosphere microbiome.

## **5. Long-distance microbial movement:**

Microbes are key components of global ecological processes (Graham et al., 2016). Air-borne microbes do not originate only from local sources but global dispersal of microbes through the atmosphere has been extensively described (Cáliz et al., 2018). A high diversity and abundance of fungi, bacteria, and archaea are transported over long distances and deposited from the atmosphere (Cáliz et al., 2018). However, this microbial dispersal is strongly influenced by air mass pathways and seasonal patterns (Cáliz et al., 2018). An important implication of microbial dispersal over long distances is the dissemination of pathogenic agents that can affect human, animal, and plant health (Brown & Hovmøller, 2002), since entire microbial communities are being dispersed from one place



to another (Weil et al., 2017). Multi-year studies have made it possible to identify the sources of air-borne communities throughout the seasons. In the Mediterranean area, freshwater, cropland, and urban biomes have been found to be the sources of air-borne bacteria in summer; in contrast, marine and forest biomes are the sources in winter (Cáliz et al., 2018). This suggests that microbial origin and trajectory can be traced for epidemiological and ecological purposes. However, this study was limited to the identification of microbes at the genus level. Further, collections were performed only at the sink location but not in the putative source areas.

Is it possible that long-distance microbial dispersal can alter sensitive sink environments including plant microbiomes? Can we predict future plant disease outbreaks following microbial dispersal mechanisms? Highly virulent fungal plant pathogens including *Triphragmium ulmariae* (Smith et al., 2003), *Phakospora pachyrizi* (Krupa et al., 2006), *Puccinia graminis* (Stokstad, 2007), *Puccinia striiformis* (Hovmøller et al., 2008), and *Fusarium graminearum* (Prussin et al., 2015) disperse as spores over long distances through the atmosphere. Further, viable bacterial plant pathogens have been found to move with African dust to the Caribbean islands (Griffin et al., 2001). *Erwinia carotovora* and *P. syringae* strains have been found to originate and to disperse through several compartments of the water cycle (Morris et al., 2008).

It may be possible to predict the potential movement of particular pathogens through the atmosphere since bacterial signatures have been correlated with source, season, and chemical composition of precipitation over time (Cáliz et al., 2018). However, a more comprehensive analysis at the strain level should be used. Additionally, microbial collections at the putative source locations would help us to confirm microbial dispersal over long distances.

## **6. Interconnection of habitats through the movement of microbes:**

Similar to plants, every single multicellular organism harbors a microbiome (Gilbert et al., 2012). The One Health concept suggests that all organisms are interconnected to each other and within the environment through the dispersal of microbes (van Bruggen et al., 2019).

Microbial dispersal influences gene flow (Whittaker & Ryneerson, 2017) and horizontal gene transfer (Aminov, 2011). Some microbiota are more resilient to disturbances than others (Todman et al., 2018). However, how much does a habitat change after disturbance? For example, if we consider precipitation, which carries a high diversity of microbes, as a disturbance process for the phyllosphere ecosystem (Lindemann & Upper, 1985), how much of this foreign diversity establishes on leaves before a new equilibrium is restored? Our current understanding of how plants recognize and assemble this large diversity of microbial foreigners is still limited. Since microbial hubs (Aglar et al., 2016) can also be affected by the disturbance, it may be possible to observe a new re-arrangement in the community structure after the habitat is perturbed. As we have discussed above, pathogenic agents are cycling globally and are deposited after traveling long distances. Pathogens can be suppressed in a habitat with a well-established microbiome structure (Mendes et al., 2011) but, how is bacterial diversity affected by the introduction of a pathogenic agent during disturbances, like precipitation?

The study of sources and drivers that shape the microbiome structure of the different plant compartments have been focused on the one-way source and inferring that additional environmental forces may influence the assembly. We believe that to truly understand microbial dispersal and assembly a macro-level perspective must be used, whereby the microbiomes of several interconnected habitats are investigated over time and at the strain level.

## **7. Conclusions and future directions:**

Every single habitat and multicellular organisms harbor a microbiome. These microbial communities are cycling from one habitat to another not only locally but globally.

Since bacterial dispersal may directly influence plant health, we believe it is mandatory to identify all potential sources that drive microbial assembly on plants. We suggest that to better understand the mechanisms that drive the plant microbiome assembly new tools and a macrolevel perspective must be used. Due to high genetic diversity within one bacterial species, the low resolution of the either 16S or 18S rRNA sequencing make it impossible to precisely determine if microbes found in two different habitats belong to the same population. This limitation makes it difficult to conclude that they derive from the same initial source. We suggest that the use of metagenomic sequencing and strain-level microbial identification approaches together will help to determine the mode of transmission and dispersal mechanisms that facilitate microbiome assembly.

We also suggest that an interconnected-level perspective must be considered when studying microbiome assembly and bacterial dispersal so that we can quantify the contribution of several sources of inoculum to a particular sink. Additionally, we suggest that continued monitoring of putative microbial sources and sinks will help us to reveal the movement of microbial clones or members of the same microbial population from one habitat to another and in which direction they move.

Further, the development of a reductionist approach with synthetic communities (SynCom) from culturable soil and air/rain microbes, for example, could reveal their relative fitness when inoculated on leaves. Moreover, since some genera are present in soil, the atmosphere, and on leaf surfaces, mixtures of soil-isolated and atmosphere-isolated strains of the same genus could be co-inoculated on plant leaves. Over time, their relative abundance could then be determined using metagenomic sequencing to discover if it is the soil-isolated strains, the rain/air-isolated strains, or both that colonize plant leaves.

In conclusion, we suggest the implementation of strain-level microbial identification applied to all possible sources and sinks of the phyllosphere microbiome over time to understand dispersal dynamics in all possible directions across plant compartments and the habitats that surround plants.

## 8. References:

- Agler MT, Ruhe J, Kroll S, *et al.*, 2016. Microbial Hub Taxa Link Host and Abiotic Factors to Plant Microbiome Variation. *PLOS Biology* **14**, e1002352.
- Aho K, Weber CF, Christner BC, *et al.*, 2019. Spatiotemporal patterns of microbial composition and diversity in precipitation. *Ecological Monographs* **0**.
- Amato P, Joly M, Besaury L, *et al.*, 2017. Active microorganisms thrive among extremely diverse communities in cloud water. *PloS one* **12**, e0182869.
- Aminov R, 2011. Horizontal gene exchange in environmental microbiota. *Frontiers in microbiology* **2**.
- Bahram M, Hildebrand F, Forslund SK, *et al.*, 2018. Structure and function of the global topsoil microbiome. *Nature* **560**, 233-7.
- Bai Y, Müller DB, Srinivas G, *et al.*, 2015. Functional overlap of the Arabidopsis leaf and root microbiota. *Nature* **528**, 364.
- Barret M, Briand M, Bonneau S, *et al.*, 2015. Emergence shapes the structure of the seed microbiota. *Applied and Environmental Microbiology* **81**, 1257.
- Berg M, Koskella B, 2018. Nutrient- and dose-dependent microbiome-mediated protection against a plant pathogen. *Current Biology* **28**, 2487-92.e3.
- Berrabah F, Ratet P, Gourion B, 2018. Legume Nodules: Massive Infection in the Absence of Defense Induction. *Molecular Plant-Microbe Interactions* **32**, 35-44.
- Bodenhause N, Bortfeld-Miller M, Ackermann M, Vorholt JA, 2014. A synthetic community approach reveals plant genotypes affecting the phyllosphere microbiota. *PLOS Genetics* **10**, e1004283.
- Brown JKM, Hovmøller MS, 2002. Aerial dispersal of pathogens on the global and continental scales and its impact on plant disease. *Science* **297**, 537.

- Bulgarelli D, Rott M, Schlaeppi K, *et al.*, 2012. Revealing structure and assembly cues for Arabidopsis root-inhabiting bacterial microbiota. *Nature* **488**, 91.
- Bulgarelli D, Schlaeppi K, Spaepen S, Van Themaat EVL, Schulze-Lefert P, 2013. Structure and functions of the bacterial microbiota of plants. *Annual Review of Plant Biology* **64**, 807-38.
- Burrows SM, Elbert W, Lawrence MG, Pöschl U, 2009. Bacteria in the global atmosphere – Part 1: Review and synthesis of literature data for different ecosystems. *Atmospheric Chemistry and Physics* **9**, 9263-80.
- Cáliz J, Triadó-Margarit X, Camarero L, Casamayor EO, 2018. A long-term survey unveils strong seasonal patterns in the airborne microbiome coupled to general and regional atmospheric circulations. *Proceedings of the National Academy of Sciences* **115**, 12229.
- Castrillo G, Teixeira PJPL, Paredes SH, *et al.*, 2017. Root microbiota drive direct integration of phosphate stress and immunity. *Nature* **543**, 513.
- Chaparro JM, Badri DV, Vivanco JM, 2014. Rhizosphere microbiome assemblage is affected by plant development. *The ISME journal* **8**, 790-803.
- Cole BJ, Feltcher ME, Waters RJ, *et al.*, 2017. Genome-wide identification of bacterial plant colonization genes. *PLOS Biology* **15**, e2002860.
- Cottyn B, Debode J, Regalado E, Mew TW, Swings J, 2009. Phenotypic and genetic diversity of rice seed-associated bacteria and their role in pathogenicity and biological control. *Journal of Applied Microbiology* **107**, 885-97.
- Cui H, Tsuda K, Parker JE, 2015. Effector-Triggered Immunity: From pathogen perception to robust defense. *Annual Review of Plant Biology* **66**, 487-511.
- Eyre AW, Wang M, Oh Y, Dean RA, 2019. Identification and characterization of the core rice seed microbiome. *Phytobiomes Journal* **3**, 148-57.

- Failor KC, Schmale III DG, Vinatzer BA, Monteil CL, 2017. Ice nucleation active bacteria in precipitation are genetically diverse and nucleate ice by employing different mechanisms. *The ISME journal* **11**, 2740.
- Fierer N, 2017. Embracing the unknown: disentangling the complexities of the soil microbiome. *Nature Reviews Microbiology* **15**, 579.
- Fierer N, Leff JW, Adams BJ, *et al.*, 2012. Cross-biome metagenomic analyses of soil microbial communities and their functional attributes. *Proceedings of the National Academy of Sciences* **109**, 21390.
- Fitzpatrick CR, Copeland J, Wang PW, Guttman DS, Kotanen PM, Johnson MTJ, 2018. Assembly and ecological function of the root microbiome across angiosperm plant species. *Proceedings of the National Academy of Sciences* **115**, E1157.
- Gaiero JR, McCall CA, Thompson KA, Day NJ, Best AS, Dunfield KE, 2013. Inside the root microbiome: Bacterial root endophytes and plant growth promotion. *American Journal of Botany* **100**, 1738-50.
- Gilbert SF, Sapp J, Tauber AI, 2012. A symbiotic view of life: We have never been individuals. *The Quarterly Review of Biology* **87**, 325-41.
- Grady KL, Sorensen JW, Stopnisek N, Guittar J, Shade A, 2019. Assembly and seasonality of core phyllosphere microbiota on perennial biofuel crops. *Nature communications* **10**, 4135-.
- Graham EB, Knelman JE, Schindlbacher A, *et al.*, 2016. Microbes as engines of ecosystem function: When does community structure enhance predictions of ecosystem processes? *Frontiers in microbiology* **7**.
- Griffin DW, 2007. Atmospheric movement of microorganisms in clouds of desert dust and implications for human health. *Clin Microbiol Rev* **20**, 459-77, table of contents.
- Griffin DW, Garrison VH, Herman JR, Shinn EA, 2001. African desert dust in the Caribbean atmosphere: Microbiology and public health. *Aerobiologia* **17**, 203-13.

- Haas D, Galler H, Luxner J, *et al.*, 2013. The concentrations of culturable microorganisms in relation to particulate matter in urban air. *Atmospheric Environment* **65**, 215-22.
- Hacquard S, Garrido-Oter R, González A, *et al.*, 2015. Microbiota and Host Nutrition across Plant and Animal Kingdoms. *Cell Host & Microbe* **17**, 603-16.
- Hacquard S, Spaepen S, Garrido-Oter R, Schulze-Lefert P, 2017. Interplay between innate immunity and the plant microbiota. *Annu Rev Phytopathol* **55**, 565-89.
- Hamonts K, Trivedi P, Garg A, *et al.*, 2018. Field study reveals core plant microbiota and relative importance of their drivers. *Environmental microbiology* **20**, 124-40.
- Hardoim PR, Hardoim CCP, Van Overbeek LS, Van Elsas JD, 2012. Dynamics of Seed-Borne Rice Endophytes on Early Plant Growth Stages. *PloS one* **7**, e30438.
- Hernandez MN, Lindow SE, 2019. *Pseudomonas syringae* increases water availability in leaf microenvironments via production of hygroscopic syringafactin. *Applied and Environmental Microbiology* **85**, e01014-19.
- Hirano SS, Upper CD, 2000. Bacteria in the leaf ecosystem with emphasis on *Pseudomonas syringae*-a pathogen, ice nucleus, and epiphyte. *Microbiology and molecular biology reviews : MMBR* **64**, 624-53.
- Hovmøller MS, Yahyaoui AH, Milus EA, Justesen AF, 2008. Rapid global spread of two aggressive strains of a wheat rust fungus. *Molecular Ecology* **17**, 3818-26.
- Hu L, Robert CaM, Cadot S, *et al.*, 2018. Root exudate metabolites drive plant-soil feedbacks on growth and defense by shaping the rhizosphere microbiota. *Nature communications* **9**, 2738.
- Jacobs JL, Sundin GW, 2001. Effect of solar UV-B radiation on a phyllosphere bacterial community. *Applied and Environmental Microbiology* **67**, 5488.
- Johnston-Monje D, Raizada MN, 2011. Conservation and diversity of seed associated endophytes in *Zea* across boundaries of evolution, ethnography and ecology. *PloS one* **6**, e20396.
- Jones JDG, Dangl JL, 2006. The plant immune system. *Nature* **444**, 323-9.

- Kadivar H, Stapleton AE, 2003. Ultraviolet radiation alters maize phyllosphere bacterial diversity. *Microbial Ecology* **45**, 353-61.
- Kourtev PS, Hill KA, Shepson PB, Konopka A, 2011. Atmospheric cloud water contains a diverse bacterial community. *Atmospheric Environment* **45**, 5399-405.
- Kristin A, Miranda H, 2013. The root microbiota—a fingerprint in the soil? *Plant and Soil* **370**, 671-86.
- Krupa S, Bowersox V, Claybrooke R, *et al.*, 2006. Introduction of Asian soybean rust *Urediniospores* into the Midwestern United States—A case study. *Plant Dis* **90**, 1254-9.
- Kwak M-J, Kong HG, Choi K, *et al.*, 2018. Rhizosphere microbiome structure alters to enable wilt resistance in tomato. *Nature Biotechnology* **36**, 1100.
- Lebeis SL, Paredes SH, Lundberg DS, *et al.*, 2015. Salicylic acid modulates colonization of the root microbiome by specific bacterial taxa. *Science* **349**, 860.
- Lindemann J, Upper CD, 1985. Aerial dispersal of epiphytic bacteria over bean plants. *Applied and Environmental Microbiology* **50**, 1229-32.
- Links MG, Demeke T, Gräfenhan T, Hill JE, Hemmingsen SM, Dumonceaux TJ, 2014. Simultaneous profiling of seed-associated bacteria and fungi reveals antagonistic interactions between microorganisms within a shared epiphytic microbiome on *Triticum* and *Brassica* seeds. *New Phytologist* **202**, 542-53.
- López-López A, Rogel MA, Ormeño-Orrillo E, Martínez-Romero J, Martínez-Romero E, 2010. *Phaseolus vulgaris* seed-borne endophytic community with novel bacterial species such as *Rhizobium endophyticum* sp. nov. *Systematic and Applied Microbiology* **33**, 322-7.
- Lundberg DS, Lebeis SL, Paredes SH, *et al.*, 2012. Defining the core *Arabidopsis thaliana* root microbiome. *Nature* **488**, 86-90.



- Maignien L, Deforce EA, Chafee ME, Eren AM, Simmons SL, 2014. Ecological succession and stochastic variation in the assembly of *Arabidopsis thaliana* phyllosphere communities. *mBio* **5**, e00682-13.
- Mano H, Tanaka F, Nakamura C, Kaga H, Morisaki H, 2007. Culturable endophytic bacterial flora of the maturing leaves and roots of rice plants (*Oryza sativa*) cultivated in a paddy field. *Microbes and Environments* **22**, 175-85.
- Mano H, Tanaka F, Watanabe A, Kaga H, Okunishi S, Morisaki H, 2006. Culturable surface and endophytic bacterial flora of the maturing seeds of rice plants (*Oryza sativa*) cultivated in a paddy field. *Microbes and Environments* **21**, 86-100.
- Marcell LM, Beattie GA, 2002. Effect of leaf surface waxes on leaf colonization by *Pantoea agglomerans* and *Clavibacter michiganensis*. *Molecular Plant-Microbe Interactions* **15**, 1236-44.
- Maude RB, 1996. *Seedborne diseases and their control: principles and practice*. CAB International.
- Mechan-Llontop ME, Tian L, Bernal-Galeano V, *et al.*, 2019. Assessing the potential of culture-independent 16S rRNA microbiome analysis in disease diagnostics: the example of *Dianthus gratianopolitanus* and *Robbsia andropogonis*. *European Journal of Plant Pathology*.
- Mendes R, Kruijt M, De Bruijn I, *et al.*, 2011. Deciphering the rhizosphere microbiome for disease-suppressive bacteria. *Science* **332**, 1097.
- Midha S, Bansal K, Sharma S, *et al.*, 2016. Genomic resource of rice seed associated bacteria. *Frontiers in microbiology* **6**.
- Mitter B, Pfaffenbichler N, Sessitsch A, 2016. Plant-microbe partnerships in 2020. *Microbial biotechnology* **9**, 635-40.
- Monteil CL, Yahara K, Studholme DJ, *et al.*, 2016. Population-genomic insights into emergence, crop adaptation and dissemination of *Pseudomonas syringae* pathogens. *Microbial genomics* **2**, e000089-e.

- Morris CE, 2002. Phyllosphere.
- Morris CE, Conen F, Alex Huffman J, Phillips V, Pöschl U, Sands DC, 2014. Bioprecipitation: a feedback cycle linking Earth history, ecosystem dynamics and land use through biological ice nucleators in the atmosphere. *Global Change Biology* **20**, 341-51.
- Morris CE, Sands DC, Vinatzer BA, *et al.*, 2008. The life history of the plant pathogen *Pseudomonas syringae* is linked to the water cycle. *The ISME journal* **2**, 321-34.
- Newcombe G, Harding A, Ridout M, Busby PE, 2018. A hypothetical bottleneck in the plant microbiome. *Frontiers in microbiology* **9**, 1645-.
- Ottesen AR, Gorham S, Reed E, *et al.*, 2016. Using a control to better understand phyllosphere microbiota. *PloS one* **11**, e0163482.
- Pérez-Jaramillo JE, De Hollander M, Ramírez CA, Mendes R, Raaijmakers JM, Carrión VJ, 2019. Deciphering rhizosphere microbiome assembly of wild and modern common bean (*Phaseolus vulgaris*) in native and agricultural soils from Colombia. *Microbiome* **7**, 114.
- Polymenakou PN, 2012. Atmosphere: A source of pathogenic or beneficial microbes? *Atmosphere* **3**.
- Prussin AJ, Marr LC, Schmale DG, Stoll R, Ross SD, 2015. Experimental validation of a long-distance transport model for plant pathogens: Application to *Fusarium graminearum*. *Agricultural and Forest Meteorology* **203**, 118-30.
- Ramirez KS, Leff JW, Barberán A, *et al.*, 2014. Biogeographic patterns in below-ground diversity in New York City's Central Park are similar to those observed globally. *Proceedings of the Royal Society B: Biological Sciences* **281**, 20141988.
- Ritpitakphong U, Falquet L, Vimoltust A, Berger A, Métraux J-P, L'haridon F, 2016. The microbiome of the leaf surface of *Arabidopsis* protects against a fungal pathogen. *New Phytologist* **210**, 1033-43.
- Ruinen J, 1961. The phyllosphere: I. An ecologically neglected milieu *Plant and Soil* **15**, 81-109.

- Rybakova D, Mancinelli R, Wikström M, *et al.*, 2017. The structure of the *Brassica napus* seed microbiome is cultivar-dependent and affects the interactions of symbionts and pathogens. *Microbiome* **5**, 104.
- Šantl-Temkiv T, Gosewinkel U, Starnawski P, Lever M, Finster K, 2018. Aeolian dispersal of bacteria in southwest Greenland: their sources, abundance, diversity and physiological states. *FEMS Microbiol Ecol* **94**.
- Schlatter D, Kinkel L, Thomashow L, Weller D, Paulitz T, 2017. Disease suppressive soils: New insights from the soil microbiome. *Phytopathology* **107**, 1284-97.
- Shade A, Jacques M-A, Barret M, 2017. Ecological patterns of seed microbiome diversity, transmission, and assembly. *Current Opinion in Microbiology* **37**, 15-22.
- Simonin M, Dasilva C, Terzi V, *et al.*, 2019. Influence of plant genotype and soil on the wheat rhizosphere microbiome: identification of a core microbiome across eight African and European soils. *bioRxiv*, 777383.
- Smith DJ, Timonen HJ, Jaffe DA, *et al.*, 2013. Intercontinental Dispersal of Bacteria and Archaea by Transpacific Winds. *Applied and Environmental Microbiology* **79**, 1134.
- Smith DL, Ericson L, Burdon JJ, 2003. Epidemiological patterns at multiple spatial scales: an 11-year study of a *Triphragmium ulmariae*–*Filipendula ulmaria* metapopulation. *Journal of Ecology* **91**, 890-903.
- Stokstad E, 2007. Deadly wheat fungus threatens world's breadbaskets. *Science* **315**, 1786.
- Tkacz A, Cheema J, Chandra G, Grant A, Poole PS, 2015. Stability and succession of the rhizosphere microbiota depends upon plant type and soil composition. *The ISME journal* **9**, 2349.
- Todman LC, Fraser FC, Corstanje R, *et al.*, 2018. Evidence for functional state transitions in intensively-managed soil ecosystems. *Scientific reports* **8**, 11522.
- Truyens S, Weyens N, Cuypers A, Vangronsveld J, 2013. Changes in the population of seed bacteria of transgenerationally Cd-exposed *Arabidopsis thaliana*. *Plant Biology* **15**, 971-81.

- Vacher C, Hampe A, Porté AJ, Sauer U, Compant S, Morris CE, 2016. The Phyllosphere: Microbial Jungle at the Plant–Climate Interface. *Annual Review of Ecology, Evolution, and Systematics* **47**, 1-24.
- Van Bruggen AHC, Goss EM, Havelaar A, Van Diepeningen AD, Finckh MR, Morris JG, 2019. One Health - Cycling of diverse microbial communities as a connecting force for soil, plant, animal, human and ecosystem health. *Science of The Total Environment* **664**, 927-37.
- Vogel C, Bodenhausen N, Gruissem W, Vorholt JA, 2016. The *Arabidopsis* leaf transcriptome reveals distinct but also overlapping responses to colonization by phyllosphere commensals and pathogen infection with impact on plant health. *New Phytologist* **212**, 192-207.
- Vorholt JA, 2012. Microbial life in the phyllosphere. *Nature Reviews Microbiology* **10**, 828.
- Wagner MR, Lundberg DS, Del Rio TG, Tringe SG, Dangl JL, Mitchell-Olds T, 2016. Host genotype and age shape the leaf and root microbiomes of a wild perennial plant. *Nature communications* **7**, 12151.
- Walitang DI, Kim C-G, Kim K, Kang Y, Kim YK, Sa T, 2018. The influence of host genotype and salt stress on the seed endophytic community of salt-sensitive and salt-tolerant rice cultivars. *BMC Plant Biology* **18**, 51.
- Walters WA, Jin Z, Youngblut N, *et al.*, 2018. Large-scale replicated field study of maize rhizosphere identifies heritable microbes. *Proceedings of the National Academy of Sciences* **115**, 7368.
- Weil T, De Filippo C, Albanese D, *et al.*, 2017. Legal immigrants: invasion of alien microbial communities during winter occurring desert dust storms. *Microbiome* **5**, 32-.
- Weller DM, Raaijmakers JM, Gardener BBM, Thomashow LS, 2002. Microbial populations responsible for specific soil suppressiveness to plant pathogens. *Annu Rev Phytopathol* **40**, 309-48.

Whittaker KA, Rynearson TA, 2017. Evidence for environmental and ecological selection in a microbe with no geographic limits to gene flow. *Proceedings of the National Academy of Sciences* **114**, 2651-6.

Zarraonaindia I, Owens SM, Weisenhorn P, *et al.*, 2015. The soil microbiome influences grapevine-associated microbiota. *mBio* **6**, e02527-14.

Zhalnina K, Louie KB, Hao Z, *et al.*, 2018. Dynamic root exudate chemistry and microbial substrate preferences drive patterns in rhizosphere microbial community assembly. *Nature Microbiology* **3**, 470-80.

## **Chapter 2:**

### **An investigation of rain as reservoir of tomato phyllosphere microbiota**

Marco E. Mechan-Llontop<sup>1</sup>, Long Tian<sup>1</sup>, Parul Sharma<sup>1</sup>, David Haak<sup>1</sup>, Christopher Clarke<sup>2</sup> and  
Boris A. Vinatzer<sup>1</sup>

<sup>1</sup> School of Plant and Environmental Sciences, Virginia Tech, Blacksburg-VA, 24061

<sup>2</sup> Genetic Improvement for Fruits and Vegetables Laboratory, Beltsville Agricultural Research Center, U.S. Department of Agriculture-Agricultural Research Service, Beltsville, MD 20705

#### **\* Correspondence:**

Boris A. Vinatzer

vinatzer@vt.edu

Keywords: Rain, disturbance, sources, sink, microbiome assembly, phyllosphere, tomato

**Abstract:**

Understanding microbial dispersal pathways provides foundational knowledge about the drivers of microbiota assembly. After finding that leaves of tomato plants exposed to rain carried a higher microbial population size than leaves of tomato plants not exposed to rain, we hypothesized that rain was one of the sources of tomato phyllosphere microbiota. First, using 16S rRNA amplicon sequencing, *Massilia*, *Sphingomonas*, *Methylobacterium*, *Acidiphilium*, *Pseudomonas*, *Pantoea*, *Endobacter* and *Duganella* were identified as the most diverse genera in the core rain microbiota in Blacksburg, Virginia. After inoculating tomatoes under controlled laboratory conditions with concentrated microbiota derived from the analyzed rain events, 29 taxa belonging to the core rain microbiota and 4 additional rain-borne taxa of the genera *Enterobacter*, *Janthinobacterium*, and *Kluyvera* significantly increased in relative abundance based on six independent experiments. Using metagenomic sequencing, the most efficient rain-borne tomato colonizers were identified as *Pantoea vagans*, *Pantoea agglomerans*, and *Pseudomonas citronellosis*. Not a single taxon increased significantly in abundance when filter-sterilized rain was used to inoculate tomatoes instead. Comparing the phyllosphere microbiota of tomatoes never exposed to rain because grown in a commercial greenhouse with tomatoes grown outside and exposed to rain, we found that several taxa identified as the most efficient tomato colonizers in our inoculation experiments were absent. This result further supports our hypothesis that rain is a source of tomato phyllosphere microbiota. However, since rain harbors a much higher bacterial diversity than the phyllosphere, only a sub-set of rain-borne taxa appear to be adapted to interact with plants. To further our understanding of the assembly of phyllosphere microbiota, we suggest to use metagenomics for strain-level identification applied contemporarily to phyllosphere microbiota from all its possible sources: seed, roots, soil, air, and precipitation.

## 1. Introduction

Microbial communities associated with plants, often referred to as plant-associated microbiota or the plant microbiome, influence a remarkable number of processes of plant biology (Goh et al., 2013, Badri et al., 2013, Hacquard et al., 2015, Lu et al., 2018, Torres-Cortés et al., 2018) and plant survival (Durán et al., 2018). The phyllosphere, described here as the plant compartment that extends from the outside to the inside of the leaf (Morris, 2002, Vacher et al., 2016), harbors a high diversity of microorganisms, with bacteria being the most abundant domain (Lindow & Brandl, 2003, Vorholt, 2012) and crucial for plant health (Ritpitakphong et al., 2016, Berg & Koskella, 2018) (Durán et al., 2018)

Phyllosphere microbiota are exposed to fluctuating environmental stresses, including changes in UV exposure, temperature, water availability, osmotic stress, and humidity (Amato et al., 2007). The core bacterial taxa that can withstand these environmental stressors include Proteobacteria, Actinobacteria, Bacteroidetes, and Firmicutes (Vorholt, 2012, Bulgarelli et al., 2013). However, the structure of the phyllosphere microbiome is variable because of the changing biotic and abiotic factors (Lindemann & Upper, 1985, Copeland et al., 2015, Wagner et al., 2016).

Next-generation sequencing technologies have expanded our knowledge of the microbiome structure and composition present in the phyllosphere of several plant species (Knief et al., 2012, Williams et al., 2013, Kembel et al., 2014, Copeland et al., 2015, Grady et al., 2019), including tomato (*Solanum Lycopersicum*). The bacterial genera *Pseudomonas*, *Erwinia*, *Sphingomonas*, *Janthinobacterium*, *Curtobacterium*, *Agrobacterium*, *Stenotrophomonas*, *Aurantimonas*, *Thermomonas*, *Buchnera*, *Enterococcus*, *Rubrobacter*, *Methylobacterium*, *Deinococcus*, and *Acidovorax* have been observed to be associated with the tomato phyllosphere grown in the field (Ottesen et al., 2013, Ottesen et al., 2016, Toju et al., 2019) protecting tomato from pathogenic agents (Berg & Koskella, 2018) .

Microbes are constantly cycling across interconnected habitats maintaining healthy ecosystems (van Bruggen et al., 2019). It has been demonstrated that plants represent an important source of



microbes that are continuously released into the atmosphere as aerosol (Šantl-Temkiv et al., 2013). After an ecological filtering in the atmospheric compartment due to environmental stresses, metabolically active bacteria (Amato et al., 2007) contribute to crucial biogeochemical cycles, including the formation of precipitation (Christner et al., 2008, Morris et al., 2014, Amato et al., 2015, Failor et al., 2017). Although precipitation events are largely influenced by different sources, Proteobacteria followed by Bacteroidetes, Actinobacteria, and Firmicutes are ubiquitous and dominate the atmosphere (Peter et al., 2014, Hiraoka et al., 2017, Cáliz et al., 2018). *Sphingomonas*, *Methylobacterium*, *Massilia*, *Pseudomonas*, *Polaromonas*, *Acidiphilium*, *Ramlibacter*, *Mucilaginibacter*, *Hymenobacter*, and *Noviherbaspirillum* have been identified as core taxa of precipitation (Cáliz et al., 2018, Aho et al., 2019). The atmosphere serves as a vehicle for microbial dispersal not only locally but globally (Bovallius et al., 1978, Brown & Hovmøller, 2002, Schmale & Ross, 2015). Complete air-borne microbial communities are deposited back to earth surfaces through dry and wet depositions, representing a crucial route for the dissemination of microbes, including beneficial and pathogenic species (Polymenakou, 2012, Monteil et al., 2014)(Monteil et al., 2016).

Microbial community assembly is largely influenced by deterministic (selection) and stochastic (dispersal) processes that determine the complex structure and function of microbiomes (Powell et al., 2015, Zhou & Ning, 2017, Graham & Stegen, 2017). This assembly process has been extensively studied in roots, which recruit microbial communities from the surrounding soil (Fitzpatrick et al., 2018, Pérez-Jaramillo et al., 2019). However, sources that drive the phyllosphere microbiome assembly are still under debate. It has been suggested that soil is the major reservoir of the phyllosphere microbiome (Zarraonaindia et al., 2015, Wagner et al., 2016, Grady et al., 2019). In fact, it has been observed that the leaf microbiome reflects soil bacterial diversity at an early stage of growth but significantly differs as plants grow (Copeland et al., 2015). On the contrary, (Maignien et al., 2014, Ottesen et al., 2016) found evidence that air-borne microbes are an important source for the

phyllosphere microbial diversity. However, these studies have been limited to the use of a 16S rRNA sequencing approach, which low resolution does not allow to identify bacteria at the species level and strain level (Mechan-Llontop et al., 2019, Regalado et al., 2019). This limitation makes it impossible to determine if the presence of the same taxon in two different habitats means that the same population occupies both habitats or if two genetically distinct populations belonging to the same taxon occupy the two habitats.

Only a few studies so far have explored precipitation as source of the phyllosphere microbiota (Morris et al., 2008). We previously used whole genome sequencing and found evidence that *Pseudomonas syringae* bacteria isolated from diseased cantaloupe plants and *P. syringae* in compartments of the water cycle are members of the same population (Monteil et al., 2016). Moreover, it is well known that fungal spores are released from plants, travel long distance through the atmosphere, and are deposited back on plants with rain (Woo et al., 2018)..

After finding that the phyllosphere of tomato plants exposed to rain contains a higher bacterial population size than the phyllosphere of tomato plants not exposed to rain, we hypothesized that rain contributes bacteria to the assembly of the phyllosphere microbiota beyond pathogenic bacteria and fungi. To test this hypothesis, we 1) characterized the rain-associated microbiota using 16S rRNA and metagenomic sequencing and compared it to published 16S rRNA data of tomato phyllosphere microbiota, 2) used concentrated rain microbiota to inoculate tomato plants and determined their effect on the composition of the tomato phyllosphere microbiota, and 3) compared phyllosphere microbiota of tomato plants exposed to rain with those of tomato plants not exposed to rain.

## **2. Materials and methods**

### **2.1. Determination of the population size of tomato phyllosphere microbiota**

Tomato seeds of the cultivar ‘Rio Grande’ were germinated in autoclaved soil (60 min/ fast cycle). For growth under laboratory conditions, plants were kept for 4 weeks on shelves under 14 h of light and

10 h of dark. For exposure to outdoor conditions, 4 week-old tomato plants were transplanted into large plastic pots. Pots were then either transported to a research farm and placed on gravel near a maintained lawn (Kentland Farm, Blacksburg, VA, USA) or put on the flat roof of the 3-story Latham Hall research building at Virginia Tech (Blacksburg, VA, USA) several meters above and away from soil and vegetation.

## **2.2. Culture-dependent analysis of phyllosphere microbiota**

Leaf disks were aseptically collected with a #1 cork borer and placed in a tube containing 200  $\mu$ L of sterile 10 mM MgSO<sub>4</sub> solution and 3 glass beads. Tubes were placed in a mini bead beater (Biospec Products, Inc., Bartlesville, OK, USA) and shaken for 2 min to release bacterial cells. Serial dilutions were plated on Reasoner's 2A Agar plates supplemented with cyclohexamide to inhibit fungal growth. Plates were incubated at room temperature and colony-forming units were counted 4 days later.

## **2.2. Rain collection**

Rain was collected as described in (Failor et al., 2017). In short, autoclavable bags were wrapped in aluminum foil and autoclaved for 40 min/fast cycle. Trash cans were arranged away from structures on the roof of the Latham Hall research building. Surfaces of containers were sprayed with 75% ethanol to prevent contamination. Sterile bags were placed in the cans and the lid placed back on top until the beginning of a rainfall event, at which point they were removed. The lids of three cans were then removed but one can was kept closed during the precipitation event as a negative control. After the end of precipitation events, 1 liter (L) of sterile water was poured into the negative control can, simulating the precipitation event. After rainfall events ended, bags containing rain water were removed and placed at 4°C until processing.

Three L of rainwater were vacuum filtrated (reusable filter holders from Thermo Scientific Nalgene, USA) through a 0.2  $\mu$ m pore filter membrane (Supor<sup>®</sup> 200 PES membrane Disc Filter, PALL, USA). Filters were removed using sterile tweezers, placed into a 15 mL Eppendorf tube and stored at -80°C

until processing. The same procedure was performed for the 1 L of sterile water collected as negative control.

### **2.3. Rain as bacterial inoculum and plant treatments**

Two L of rainwater were vacuum-filtered as described above. To concentrate the bacterial microbiota present in rain 100 fold (referred to as 100X-rain from now on), the filter membrane was incubated with shaking in 20 mL of sterile water after which it was used as inoculum. The rain that passed through the filter was used as bacterial-free inoculum (filtered rain). Autoclaved double-distilled water (sterile-water) was included as negative control treatment. Four-weeks old tomato plants were placed into a 13in x 16in plastic bag and were sprayed until run off was visible with 10 mL of either 1) 100X-rain, 2) filtered rain or 3) sterile-water (four plants per treatment). Two plants per treatment were harvested two hours post-inoculation while the other two plants were kept in the bag for one day and then maintained in the lab for an additional six days.

### **2.4. Harvesting**

All plant leaves were removed and collected in a ziplock plastic bag. Sterile distilled water was added (300ml) and samples were sonicated for 10 minutes using a 1510 BRANSON sonicator (Brandsonic, Mexico)(Ottesen et al., 2013). The leaf washings were vacuum-filtered on to the same kind of 0.22  $\mu\text{m}$  pore filter described above to collect microbial cells dislocated from leaf tissue during sonication, as described above. Filter membranes were placed into a 15 ml Eppendorf tube and stored at  $-80^{\circ}\text{C}$  until processing.

### **2.5. Microbial DNA extraction**

DNA extraction from all the 0.22  $\mu\text{m}$  filter membranes was performed using the Power Water DNA isolation kit (Qiagen, USA) according to the manufacture's protocol with minor modifications. DNA concentration and quality were assessed by UV spectrophotometry (NanoDrop 1000, Thermo, USA) and visualized on a 1 % agarose gel.

### **2.6. Library preparation and sequencing**

For the 16S rRNA amplicon sequencing, the V4 hypervariable region of the 16S rRNA gene was amplified and sequenced using barcoded primers 799F (anti-chloroplast, 5'AACMGGATTAGATACCCKG3') and 1115R ("universal", 5'AGGGTTGCGCTCGTTG3') (Shade et al., 2013). A 28 cycle PCR was performed using the HotStarTaq Plus Master Mix Kit (Qiagen, USA) under the following conditions: 94°C for 3 minutes, followed by 28 cycles of 94°C for 30 seconds, 53°C for 40 seconds and 72°C for 1 minute, after which a final elongation step at 72°C for 5 minutes was performed. After amplification, PCR products obtained from the various samples were mixed in equal concentrations and purified using Agencourt Ampure beads (Agencourt Bioscience Corporation, Beverly, MA, USA). All steps from PCR to paired-end (2 × 300 bp) amplicon sequencing on the Illumina MiSeq platform were performed at Molecular Research LP (MR DNA™, Shallowater, TX, USA).

For the metagenomic sequencing using the Illumina platform, total DNA was sequenced using a 150 bp paired-end library with HiSeq 4000 of Illumina platform at the Duke University Sequencing and Genomic Technologies Shared Resource, Durham, NC, USA.

For the metagenomic sequencing using the Nanopore platform, DNA library was prepared following the '1D Native barcoding genomic DNA protocols (SQK-LSK109 and EXP-NBD104) provided by the Oxford Nanopore Technologies (ONT). In short, each DNA sample was repaired with the NEBNext Ultra II End Repair/dA-Tailing module (New England Biolabs Inc.). Repaired DNA was purified with 1.5 volumes of AMPure XP beads and eluted in 25 µl nuclease-free water. 22.5 µl of each DNA sample was individually mixed with the Blunt/TA Ligase Master Mix (New England Biolabs Inc.) and with the Native Barcode kit (ONT, Native barcoding Expansion Kit EXP-NBD104). DNA samples were purified with 1.5 volumes of AMPure XP beads and eluted with 26 µl nuclease-free water. Equimolar amounts of each barcoded sample were pooled for ligation. 15 µl Elution buffer (ONT, SQK-LSK109). Pooled samples were mixed with Adapter Mix (ONT, SQK-LSK109), NEBNext Quick Ligation Reaction Buffer (New England Biolabs Inc.) and Quick T4 DNA Ligase (New England Biolabs Inc.).

## 2.7. Bioinformatic analysis

Raw 16S rRNA paired-end sequences were processed following the analysis pipeline from the Molecular Research LP: 1) reads were joined together after q25 trimming of the ends and reoriented in the 5'-3' direction, 2) barcodes and primer sequences were removed, and 3) sequences shorter than 200bp, sequences with ambiguous base calls, and sequences with homopolymer runs exceeding 6 bp were removed. Operational taxonomic units (OTU) were assigned using the open source Quantitative Insights into Microbial Ecology (QIIME) bioinformatic pipeline (Caporaso et al., 2010), using the open-reference protocol at 97% sequence identity with UCLUST as the clustering tool and SILVA (Quast et al., 2013) as the database. All OTUs annotated as mitochondria, chloroplasts, cyanobacteria, unassigned, and OTUs with fewer than five reads were removed from the dataset. The QIIME-generated output file in the Biological Observation Matrix (BIOM) format was used for downstream data analysis and visualization in R version 3.3 using the Phyloseq 1.19.1 and ggplot2 2.2.1 packages (McMurdie & Holmes, 2013) (Wickham, 2009).

Raw 150 bp paired-end reads from Illumina were filtered to exclude short and low quality reads using Trimmomatic (version 0.38) (Bolger et al., 2014). Reads with an average quality per base below 30 and read length below 150bp were filtered out. The remaining reads were then used to perform taxonomic classification using Centrifuge (version 1.0.4) (Kim et al., 2016). All Illumina sequencing samples were analyzed with centrifuge to find the most abundant genomes in each of the samples.

The filtered raw reads from all day 7 samples were also used for assembly followed by binning. Briefly the filtered reads from all five samples for day 7 from different experiments were taken and assembled using IDBA (version 1.1.3) (Peng et al., 2012). The assembled contigs for each sample were indexed using Bowtie2 (version 2.3.5) (Langmead & Salzberg, 2012) and converted into sorted bam files using Samtools (version 1.3.1) (Li et al., 2009) which provides the information of depth coverage and is used as an input for binning. The bam file along with the file with the assembled contigs are provided as the input for Metabat (version 2.12.1) (Kang et al., 2015) for binning. The bins from Metabat were

taxonomically classified using the classify function of Sourmash (version 2.0.0) and Blast (nucleotide blast, online version) against NCBI RefSeq database.

The fast5 files containing the raw reads obtained from the MinION sequencer were base-called using Guppy (v3.3.2). The ONT workflow What is in my pot (WIMP v2019.7.9) was used for bacterial identification and a classification at the species level (Juul et al., 2015).

### **3. Results**

#### **3.1. Comparison of the bacterial population size in the phyllosphere of tomato plants exposed to rain compared to plants not exposed to rain**

Our investigation into the role of rain in shaping the phyllosphere microbiome started by comparing the bacterial community size on tomato plants grown inside under controlled conditions with that of plants grown outside exposed to environmental disturbances including rainfall. We observed that tomato plants grown for four weeks in a plastic pot at the Kentland research farm harbored larger bacterial community sizes compared to plants grown indoors during the same period (**Figure 1A**). Interestingly, even plants grown on the roof of a research building, which minimized microbial dispersal from soil and plants compared to the farm environment, had larger bacterial community sizes compared to plants grown inside (**Figure 1B**), suggesting that air-borne and rain-borne bacteria through dry and/or wet deposition respectively had colonized tomato plants grown outside.

To test the effect of rain on the bacterial population size in the tomato phyllosphere under controlled conditions, we collected rain and used it as an inoculum to treat 4-week old tomato plants that had been grown under laboratory conditions. Interestingly, 7 days after plants were treated with rain, they carried a larger bacterial population size compared to plants that had been treated with autoclaved rain (**Figure 1C**). This result directly suggested that at least some rain-borne bacteria are able to colonize plant leaves efficiently and have a direct impact on bacterial abundance in the phyllosphere.

### **3.2. Characterization of rain-borne microbiota in Blacksburg, Virginia**

In a first step towards identification of which bacterial taxa present in rainfall may efficiently colonize the tomato phyllosphere, we characterized the bacterial diversity associated with eight rainfall events during 2015 and 2016. Rainfall was collected in sterile plastic bags on the roof of the research building previously used to grow tomatoes, DNA was extracted, and the 16S rRNA gene was amplified and sequenced (**Table 1**). In total, 755,035 short reads were obtained. After 97% OTU clustering and removal of all non-bacterial and unassigned reads, a total of 398,646 reads remained and 4,887 OTUs were identified based on these reads.

The taxonomic diversity analysis revealed Proteobacteria, Actinobacteria, Bacteroidetes, Acidobacteria, and Firmicutes to be the dominant taxa at the phylum level. At the class level, Alphaproteobacteria followed by Gammaproteobacteria, Betaproteobacteria, Actinobacteria, Acidobacteria, and Bacilli were most abundant. However, there were considerable differences between samples. For example, Clamydiae at 35% relative abundance represented the most abundant taxon in the rain sample collected in August 2015, Betaproteobacteria were the most abundant taxon in the rain sample in March 2016 at 42% relative abundance, and Actinobacteria were most abundant in rain samples collected in April 2016 and December 2016 (**Figure 2**).

Even more dramatic differences between samples were observed at lower taxonomic ranks. However, a small number of bacterial genera were identified in all samples and thus can be considered the core rain microbiome at the genus level in the set of the eight collected samples: *Massilia*, *Sphingomonas*, *Methylobacterium*, *Pseudomonas*, *Acidiphilium*, *Bacillus*, *Endobacter*, *Duganella*, *Microbacterium* and *Pantoea*.

### **3.3. Comparison of the rain-borne microbiota with microbiota of plants either treated with concentrated rain microbiota, sterile rain, or double-distilled water**

In a second step towards determination of which bacterial taxa identified in the rain-borne microbiota were the most efficient colonizers of the tomato phyllosphere, we used 100-fold concentrated rain-



borne microbiota (referred to as 100X-rain from now on) as inoculum to treat tomato plants grown from sterilized seeds in autoclaved soil under controlled laboratory conditions (**Table 1**). In six independent experiments, we used the 100X-rain derived from six out of the eight collected rainfall events described above (April 2015, August 2015, March 2016, April 2016, May 2016, July 2016) to treat four tomato plants. In parallel, each time additional tomato plants were treated (1) with the corresponding bacteria-free rain-filtrate obtained from the filtration of the collected rain through 0.22  $\mu\text{m}$  filters and (2) with autoclaved double-distilled water (referred to as dd-water from now on). 16S rRNA amplicons were sequenced after amplification from microbiome DNA extracted from tomato leaves at two time points for all three treatments: two hours after treatments as soon as leaves had dried (day 0) and seven days later (day 7).

In total, 2,169,888 reads were obtained from 32 phyllosphere samples. After 97% OTU clustering and removing all non-bacterial and unassigned reads, a total of 2,071,156 reads remained in the data set and 15,547 OTUs were identified. Rarefaction curves show that most of the samples were deeply sequenced (**Figure 3A**). After subsampling depth to 7028 reads per sample, we measured the alpha diversity by the total number of observed species, and by the Shannon and Simpson diversity indexes. **Figure 3B** shows the alpha diversity measurements for rain compared to treated plants at day 0 vs day 7. Pairwise comparisons using Wilcoxon rank sum test showed a significant depletion of phyllosphere community diversity on plants treated with 100X-rain after 7 days (**Table 2**). For the Shannon index, we observed that the phyllosphere communities were depleted in richness in plants treated with filtered rain after 7 days. No significant difference was observed between day 0 and day 7 for plants treated with sterile-water.

The principal coordinates analysis (PCoA) derived from the dissimilarity matrix of Bray-Curtis (**Figure 4**), which measures the beta diversity, revealed that rain samples group together and phyllosphere samples treated with 100X-rain at day 7 are different than other treatments. This data

suggest a distinction in community structure in plants treated with 100X-rain at day 7 compared with plants treated with filtered rain and dd water at day 7.

Taxonomic diversity analysis revealed an enrichment in Proteobacteria in the tomato phyllosphere 7 days post inoculation compared to the day 0 time point regardless of treatment suggesting that the incubation conditions following inoculation favored growth of Proteobacteria independently of the presence or absence of bacteria in the inoculum. On the contrary, Actinobacteria and Firmicutes were almost completely depleted 7 days post inoculation in all plants treated with 100X-rain but not after treatments with rain-filtrate or dd-water suggesting that they may have been outcompeted by rain-borne bacteria present in the 100X-rain (**Figure 5**).

### **3.4. Identification of the bacterial rain-borne taxa that are the most efficient tomato leaf colonizers under controlled laboratory conditions**

To identify the OTUs that significantly differ in abundance between the day 0 and day 7 time points, we performed a differential abundance test using DeSeq2 (Love et al., 2014). A total of 39 OTUs of the genera *Flavobacterium*, *Pantoea*, *Duganella*, *Enterobacter*, *Janthinobacterium*, *Pseudomonas*, *Massilia*, and *Kluyvera* significantly increased in relative abundance on plants treated with 100X-rain from day 0 to day 7 (**Figure 6A**). To the contrary, 2 OTUs of the genera *Bacillus* and *Microbispora* had a significantly higher abundance at day 0 compared to day 7. Importantly, we did not find any OTU that significantly differed in abundance between day-0 and day-7 for tomato plants treated with either filtered rain or dd water, suggesting that the OTUs that increased in relative abundance from day 0 to day 7 on tomatoes after treatment with 100X-rain originated from rain and were not members of the microbiota that were present in the tomato phyllosphere before inoculation.

To further investigate the dynamics of the rain-borne microbiota, we also compared the relative abundance of OTUs present in rain with OTUs present on plants treated with 100X-rain at the day 7 time point (**Figure 6B**). We observed that OTUs from 57 bacterial taxa detected in rain were absent on plants at the day 7 time point suggesting that these taxa were unable to colonize tomato leaves. For

example, the genus *Acidiphilium* that was detected at high relative abundance in rain samples and that was also observed on plants treated with 100X-rain on day 0 was absent on plants treated with 100X-rain at day 7 (**Figure 7A**), suggesting that *Acidiphilium* failed to colonize the phyllosphere compartment. 124 OTUs instead had a significantly higher abundance on tomatoes at day 7 compared to their relative abundance in rain. Many of these OTUs belonged to the same genera that were already found to increase in abundance between day 0 and day 7 on tomatoes treated with 100X-rain (**Figure 6A**): *Flavobacterium*, *Pantoea*, *Duganella*, *Enterobacter*, *Janthinobacterium*, *Pseudomonas*, *Massilia*, and *Kluyvera*. However, 23 OTUs in additional genera (for example, *Dyella*, *Streptomyces*, or *Brevundimonas*) also significantly increased in relative abundance.

Finally, we also compared the relative abundance of OTUs in rain with that in plants treated with either 100X-rain, filtered rain, or dd-water at day 0 to identify taxa associated with laboratory grown plants but absent in rain (**Supplementary Figures 1 and 2 and 3**). These results support our previous findings since *Pantoea*, *Sphingomonas*, *Enterobacter*, *Duganella* and *Massilia* were only present in rain samples but not in the tomato phyllosphere on day 0. To the contrary, *Bacillus* OTUs were observed in high relative abundance in rain samples as well as in the tomato phyllosphere at day 0 independently of treatment and on day 7 after sterile-rain and dd-water treatments but not after 100X-rain treatments (**Figure 7H**). This suggests that *Bacillus* OTUs were present in rain and soil and/or seeds but that they could not compete with the rain-borne bacteria added with the 100X-rain treatment.

### **3.5. Phyllosphere microbiota of tomato plants never exposed to rain**

After identifying bacterial taxa present in rain that efficiently colonized tomato leaves under laboratory conditions, we hypothesized that these taxa would be missing, or at least be underrepresented, in phyllosphere microbiota of greenhouse-grown tomatoes that had never been exposed to rain. This included tomato plants grown in a hydroponic system and tomato plants grown in soil (both in a commercial greenhouse). A small number of tomato plants grown outside in pots containing autoclaved

soil on the top of the same research building where we had grown tomato plants before and where we had collected rain were analyzed for comparison purposes.

We sequence 29 phyllosphere samples collected from the hydroponic system, 18 samples from the soil system, and 3 samples from tomato plants grown on the roof. In total, we obtained 2,724,244 reads. After 97% OTU clustering and removing all non-bacterial and unassigned reads, as described previously, a total of 1,981,912 reads remained and 10,854 OTUs were identified. Taxonomic diversity analysis showed Proteobacteria, Firmicutes, Actinobacteria and Bacteroidetes were the dominant bacterial taxa (**Figure 8**). Enrichments and depletions of taxa were observed in correlation with plant age (date of collection) but not based on growing conditions.

Rarefaction curves shows that most of the samples were deeply sequenced (**Figure 9A**). All samples were subsampled to depth of 2553 reads. The alpha diversity analysis supported by the total number of observed species, and by the Shannon and Simpson diversity indexes showed differences in bacterial richness between growing systems (**Figure 9B**). Pairwise comparisons using Wilcoxon rank sum test showed a significant difference only in regard to community diversity between plants grown in hydroponic and soil systems (**Table 3**). No significant difference was observed for the Shannon index.

We then identified *Methylobacterium* (2 OTUs), *Propionibacterium* (1 OTU), *Bacillus* (1 OTU), and 1 OTU from the order Rhizobiales as the core phyllosphere microbiome in 100% of the analyzed tomato plants grown hydroponically. We identified *Bacillus* (7 OTUs), *Staphylococcus* (3 OTUs), *Methylobacterium* (2 OTUs), *Microbacterium* (2 OTUs), *Massilia* (1 OTU), *Pseudomonas* (1 OTU), *Raoultella* (1 OTU), *Propionibacterium* (1 OTU), *Curtobacterium* (1 OTU), and *Brevibacterium* (1 OTU) as the core phyllosphere microbiome of tomato plants grown in soil. Although fewer tomato plants were grown outside (n=3), we identified *Sphingomonas* (22 OTUs), *Methylobacterium* (18 OTUs), *Rhizobium* (12 OTUs), *Massilia* (11 OTUs), *Acidiphilium* (9 OTUs), *Pantoea* (7 OTUs), *Bacillus* (6 OTUs), *Pseudomonas* (6 OTUs), *Mycobacterium* (6 OTUs),

*Brevundimonas* (4 OTUs), *Bradyrhizobium* (4 OTUs), *Roseomonas* (3 OTUs), *Microbacterium* (3 OTUs), *Bryocella* (3 OTUs), *Novosphingobium* (3 OTUs) and *Lysinibacillus* (2 OTUs) as the core phyllosphere microbiome of the tomato plants (**Table 4**). Therefore, a large number of OTUs of the genera that had been found to effectively colonize tomato plants in the lab were also found in the phyllosphere microbiomes of tomato plants grown on the roof (exposed to rain and air) but not in the phyllosphere microbiomes of tomato plants grown in the greenhouse (and not exposed to rain), suggesting that rain is the source of at least some of the members of the core tomato phyllosphere microbiota.

### **3.6. Identification of the bacterial rain-borne taxa that are the most efficient tomato leaf colonizers under controlled laboratory conditions using a metagenomic approach**

Since the 16S rRNA amplicon sequencing approach is limited to the genus level by its low resolution, we decided to use metagenome shotgun sequencing to re-sequence the microbiota associated with rain samples and plants treated with 100X-rain at the day 0 and day 7 time points to identify which rain-borne bacteria were the most efficient colonizers at the species level.

The metagenome shotgun sequencing approach generated 260,035,170 short- reads. After quality control 7,543,305 reads remained of which 98.26% were identified as bacterial reads. We ranked the species present in rain by their relative average abundance in tomato leaves 7 days after being treated with 100X-rain (**Figure 10**). This allowed us to determine that, for example, the rain-borne species *Pantoea vagans* and *Pantoea agglomerans* are the most effective tomato phyllosphere colonizers in the genus *Pantoea*. Our metagenomic analysis thus supports our 16S rRNA results that rain harbors taxa that successfully colonize the tomato phyllosphere.

## **4. Discussion**

Ecosystems are interconnected by the cycling of microorganisms (van Bruggen et al., 2019). Identifying all potential sources that contribute to the assembly of microbiota in an ecosystem are

important to understand how microbiota composition changes over time and the potential to manipulate them. The sources of the highly diverse plant microbiota are still under debate (Ottesen et al., 2016, Grady et al., 2019).

Here we investigated rain as a potential source of bacteria that compose tomato phyllosphere microbiota. We developed the hypothesis that rain may be a source of the tomato phyllosphere microbiota because rain is well known to carry spores of fungal plant pathogens to leaf surfaces, for example (Monteil et al., 2014), and we had found that *P. syringae* bacteria isolated from rain and plants belong to the same population (Monteil et al., 2016). Also, previous research had suggested that air-borne bacteria may represent an important source of microbes in the bean phyllosphere (Lindemann 1985), the *A. thaliana* phyllosphere when grown under greenhouse conditions (Maignien et al., 2014), and tomato plants grown in the field (Ottesen et al., 2016). In agreement with these studies, here we observed that tomato plants exposed to environmental conditions, including rain, harbor a larger total bacterial community size (even when grown away from native soil and other plants) compared to plants grown indoors under controlled conditions (**Figure 1A and 1B**). Moreover, we found that laboratory-grown tomato plants treated with rainwater harbored a larger bacterial community size compared to plants treated with autoclaved rainwater (**Figure 1C**). Taken together, these results suggested that some rain-borne bacteria may successfully colonize the tomato phyllosphere and may possibly be an important source of the phyllosphere microbiome.

However, contrary to our hypothesis, evidence had also been found that soil is the main source of microbes of the phyllosphere microbiome of *Arabidopsis thaliana* (Bulgarelli, 2012, Bai et al 2015), grapevine (Zarraonaindia et al., 2015), mustard (Wagner et al., 2016), sugarcane (Hamonts et al, 2018), and perennial grasses (Grady et al., 2019). One limitation of these studies however is that conclusions were based on similarities and correlations between soil and phyllosphere microbiota without determining if there was a directed dispersal from soil to leaves. Therefore, we wanted to use an

experimental approach using rain microbiota as inoculum of tomato plants grown under controlled conditions to directly verify directed dispersal from rain to plants.

To do this, we needed to characterize rain microbiota first. In agreement with (Aho et al., 2019), we found Proteobacteria, Actinobacteria, Bacteroidetes, Acidobacteria and Firmicutes as the dominant bacterial phyla and found that at lower taxonomic ranks bacterial diversity and abundance changes with each rainfall event (**Figure 2**), probably related to season and air mass sources (Cáliz et al., 2018, Aho et al., 2019). However, even if relative abundance fluctuated, we identified the genera *Massilia*, *Sphingomonas*, *Methylobacterium*, *Pseudomonas*, *Acidiphilium*, and *Pantoea* as the core rain-borne bacteria in Blacksburg, VA. Interestingly, these same bacterial genera had been observed not only in different locations in North America (Aho et al., 2019) but also in Europe (Vařtilingom et al., 2012, Cáliz et al., 2018), suggesting common microbial sources and an evolutionary adaptation of these taxa to survive in the atmospheric environment. Since the genera of the core rain microbiome include genera commonly found in plant microbiota as well as soil microbiota, the composition of rain microbiota in this study could not resolve whether soil or rain were the more important sources of phyllosphere microbiota.

After inoculating tomato plants with 100X-rain and using filtered rain and sterile-water as controls, we observed an enrichment in Proteobacteria and depletion of Actinobacteria and Firmicutes in the phyllosphere microbiota going from day 0 to day 7 (**Figure 3**). We also observed a depletion in bacterial richness and diversity of the tomato phyllosphere at day 7, regardless of the inoculum (**Figure 4**). However, the beta diversity analysis showed that tomato plants inoculated with 100X-rain harbor distinct phyllosphere microbiota at day 7 compared to plants treated with either filtered rain or sterile-water (**Figure 5**). Our results thus suggest that rain does not only wash off part of the bacterial community on the phyllosphere but may also be a source of bacterial emigrants to establish a new community (Lindemann et al., 1982, Lindemann & Upper, 1985, Hirano & Upper, 1991).

After using 100X-rain as bacterial inoculum to treat plants under laboratory conditions, we found *Chryseobacterium*, *Enterobacter*, *Pantoea*, *Paenibacillus*, *Duganella*, *Streptomyces*, *Massilia*, *Shinella*, *Janthinobacterium*, *Erwinia*, and *Hyphomicrobium* significantly more abundant on the tomato phyllosphere 7 days post inoculation compared to the rain microbiota (**Figure 6A-B**). Although OTUs of the genera *Pseudomonas*, *Flavobacterium*, *Brevundimonas*, *Herbaspirillum*, *Limnobacter*, *Dyella*, *Pandoraea*, *Methylobacterium*, and *Tuberibacillus* were significantly more abundant in the tomato phyllosphere 7 days after being inoculated with 100X-rain, OTUs of these genera were also observed in tomato plants treated with sterile-water and filtered rain at time point 0 (**Supplementary Figure 1**), suggesting that some OTUs of these genera were part of the native tomato microbiota delivered most likely by seeds (Bergna et al., 2018) since sterile soil was used in this study.

*Pseudomonas*, *Erwinia*, *Sphingomonas*, *Janthinobacterium*, *Curtobacterium*, *Buchnera*, *Enterococcus*, *Rubrobacter*, *Methylobacterium*, *Deinococcus* and *Acidovorax* are bacterial genera commonly associated to the phyllosphere of tomato plants grown in the field (Ottesen et al., 2013, Ottesen et al., 2016, Toju et al., 2019). These bacterial genera were also detected in our rain samples, suggesting that rain was most likely the source of those tomato-associated bacteria in the field.

*Pantoea* have been commonly found associated with tomato seeds (Morella et al., 2019) and many other plant species (Links et al., 2014, Barret et al., 2015). However, in our analysis the genus *Pantoea* was present in high abundance in rain but absent in tomato plants treated with dd-water and most plants treated with filtered rain (**Figure 7**), suggesting that the *Pantoea* OTUs we observed on tomato leaves 7 days after being inoculated with 100X-rain most likely originated from rain. However, in one of our treatments, we observed that *Pantoea* reached high relative abundance (25.5%) after being inoculated with filtered rain at day 7. Since non-sterilized commercially available seeds were used in our study we believe that seeds used in that particular experiment carried *Pantoea* that proliferate under humid conditions after inoculation.



*Bacillus* was observed at higher relative abundance in almost all tomato plants at day 0 (**Figure 7**), suggesting that *Bacillus* associated with the tomato phyllosphere originated either from non-sterile seeds or survived soil sterilization. It was interesting to observe that after treatment with 100X-rain, *Bacillus* OTUs decreased in relative abundance at day 7 suggesting that rain-borne bacteria possibly outcompeted *Bacillus* OTUs.

Because of the limited resolution provided by sequencing a small region of the conserved 16S rRNA gene (Mechan-Llontop et al., 2019), we believe that to correctly determine bacterial dispersal pathways, a more stringent approach must be used. Therefore, we also used metagenome shotgun sequencing to analyze our rain samples and the metagenome of tomato plants treated with 100X-rain at time points 0 and 7 days post inoculation. We identified several bacterial species that in the majority of experiments increased in abundance from day 0 to day 7 including several members of genera identified as effective colonizers using the 16S rRNA approach (**Figure 10**). Our results support the hypothesis that rain precipitation is an ecologically relevant process that shapes the tomato phyllosphere microbiota under laboratory conditions. Studying the effect of rain in community assembly in the field will be essential to expand our understanding on community diversity, dynamics, and function (Konopka et al., 2015).

To discriminate bacteria that originate from rain sources and other putative sources, including soil, we decided to compare the microbiota structure of tomato plants grown in a greenhouse that were never exposed to rain. At higher taxonomic level Proteobacteria, Firmicutes, Actinobacteria and Bacteroidetes were the dominant bacterial taxa (**Figure 8**). We observed that changes in the microbiota were associated mainly with plant age (Wagner et al., 2016). Our core microbiome analysis revealed low bacterial diversity consistently present in tomato plants grown hydroponically under greenhouse conditions. We identified only 2 OTUs from the genus *Methylobacterium*, 1 OTU from *Propionibacterium*, 1 OTU from the genus *Bacillus*, and 1 OTU from the order Rhizobiales. While *Methylobacterium* is a common phyllosphere-associated bacterial genus (Knief et al., 2010), the

presence of *Propionibacterium* may be correlated with human activity indoors (Miletto & Lindow, 2015, Adams et al., 2015). Similarly, while *Bacillus* species are associated as part of the seed microbiota, it is likely that the *Bacillus* OTU observed on plants grown hydroponically was due to the use of a commercial *Bacillus*-based biopesticide (Punja et al., 2016). Also bacterial diversity might have been affected by chemical pesticides used in the hydroponic system to prevent plant diseases in the greenhouse.

However, we observed higher bacterial diversity in the core phyllosphere microbiome of tomato plants grown in soil under greenhouse conditions collected at the same time as samples from the hydroponic system. We identified 7 *Bacillus* OTUs, 3 *Staphylococcus* OTUs, 2 *Methylobacterium* OTUs, 2 *Microbacterium* OTUs, 1 *Massilia* OTU, 1 *Pseudomonas* OTU, 1 *Raoultella* 1 OTU, 1 *Propionibacterium* OTU, 1 *Curtobacterium* OTU, and 1 *Brevibacterium* 1 OTU. Some of these bacterial genera have been described associated with the tomato phyllosphere (Ottesen et al., 2013, Berg & Koskella, 2018, Toju et al., 2019, Ottesen et al., 2016). However, it is likely that some OTUs originated from soil (Wei et al., 2019) and human (Adams et al., 2015, Miletto & Lindow, 2015) sources. Since phyllosphere collections were performed in an organic growing system, chemical pesticides were not included in disease control, but a *Bacillus*-based biopesticide product was commonly used.

Importantly, the core phyllosphere microbiota of tomato plants grown in sterile soil and placed outside under environmental conditions, including rain precipitation, was more diverse compared with the core microbiota of plants grown indoors. We identified 22 *Sphingomonas* OTUs, 18 *Methylobacterium* OTUs, 12 *Rhizobium* OTUs, 11 *Massilia* OTUs, 9 *Acidiphilium* 9 OTUs, 7 *Pantoea* OTUs, 6 *Bacillus* OTUs, 6 *Pseudomonas* OTUs, 6 *Mycobacterium* OTUs, 4 *Brevundimonas* OTUs, 4 *Bradyrhizobium* OTUs, 3 *Roseomonas* OTUs, 3 *Microbacterium* OTUs, 3 *Bryocella* OTUs, 3 *Novosphingobium* OTUs and 2 *Lysinibacillus* OTUs. These results suggest that rain-borne and air-borne microbes influence the tomato phyllosphere microbiome assembly and composition.

In summary, we investigated the role of rain as a source of microbes contributing to the assembly of tomato phyllosphere microbiota using a controlled experimental system. We identified, at the genus and species level, bacteria that successfully colonize the tomato phyllosphere and may represent key players in community assembly. We also found that the taxa we identified as the most effective colonizers of tomato leaves are under-represented in greenhouse-grown tomatoes not exposed to rain. Taken together, these results suggest that rain harbors effective colonizers of the tomato phyllosphere and may represent a so far neglected source of phyllosphere microbiota. We suggest to further explore additional sources and sinks of the tomato phyllosphere microbiome besides rain, such as seed, roots, soil, and air, together over multiple time points, to quantify bacterial dispersal between all these environments in all possible directions. To better understand bacterial source-sink dynamics, we propose to do this using metagenomic sequencing for strain-level identification to precisely identify bacterial dispersal pathways. Finally, we also propose to use synthetic communities composed of genera that include both, soil-borne and rain-borne strains, as inoculum of leaves of tomato plants grown in the field, to determine if the soil-borne or rain-borne members of these taxa are the more effective tomato leaf colonizers. In summary, we believe that combining microbial ecology with experimental microbiology approaches using strain-level metagenomics will make it possible to identify the relative contribution of all possible sources and sinks of the phyllosphere microbiome.

## 5. References

- Adams RI, Bateman AC, Bik HM, Meadow JF, 2015. Microbiota of the indoor environment: a meta-analysis. *Microbiome* **3**, 49-.
- Aho K, Weber CF, Christner BC, *et al.*, 2019. Spatiotemporal patterns of microbial composition and diversity in precipitation. *Ecological Monographs*.
- Amato P, Joly M, Schaupp C, *et al.*, 2015. Survival and ice nucleation activity of bacteria as aerosols in a cloud simulation chamber. *Atmos. Chem. Phys.* **15**, 6455-65.

- Amato P, Parazols M, Sancelme M, Laj P, Mailhot G, Delort AM, 2007. Microorganisms isolated from the water phase of tropospheric clouds at the Puy de Dome: major groups and growth abilities at low temperatures. *FEMS Microbiol Ecol* **59**, 242-54.
- Badri DV, Zolla G, Bakker MG, Manter DK, Vivanco JM, 2013. Potential impact of soil microbiomes on the leaf metabolome and on herbivore feeding behavior. *New Phytologist* **198**, 264-73.
- Barret M, Briand M, Bonneau S, *et al.*, 2015. Emergence Shapes the Structure of the Seed Microbiota. *Applied and Environmental Microbiology* **81**, 1257.
- Berg M, Koskella B, 2018. Nutrient- and Dose-Dependent Microbiome-Mediated Protection against a Plant Pathogen. *Current Biology* **28**, 2487-92.e3.
- Bergna A, Cernava T, Rändler M, Grosch R, Zachow C, Berg G, 2018. Tomato Seeds Preferably Transmit Plant Beneficial Endophytes. *Phytobiomes Journal* **2**, 183-93.
- Bolger AM, Lohse M, Usadel B, 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics (Oxford, England)* **30**, 2114-20.
- Bovallius A, Bucht B, Roffey R, Anäs P, 1978. Long-range air transmission of bacteria. *Applied and Environmental Microbiology* **35**, 1231-2.
- Brown JKM, Hovmöller MS, 2002. Aerial Dispersal of Pathogens on the Global and Continental Scales and Its Impact on Plant Disease. *Science* **297**, 537.
- Bulgarelli D, Schlaeppi K, Spaepen S, Van Themaat EVL, Schulze-Lefert P, 2013. Structure and Functions of the Bacterial Microbiota of Plants. *Annual Review of Plant Biology* **64**, 807-38.
- Cáliz J, Triadó-Margarit X, Camarero L, Casamayor EO, 2018. A long-term survey unveils strong seasonal patterns in the airborne microbiome coupled to general and regional atmospheric circulations. *Proceedings of the National Academy of Sciences* **115**, 12229.
- Caporaso JG, Kuczynski J, Stombaugh J, *et al.*, 2010. QIIME allows analysis of high-throughput community sequencing data. *Nat Methods* **7**, 335-6.

- Christner BC, Morris CE, Foreman CM, Cai R, Sands DC, 2008. Ubiquity of biological ice nucleators in snowfall. *Science* **319**, 1214.
- Copeland JK, Yuan L, Layeghifard M, Wang PW, Guttman DS, 2015. Seasonal community succession of the phyllosphere microbiome. *Mol Plant Microbe Interact* **28**, 274-85.
- Durán P, Thiergart T, Garrido-Oter R, *et al.*, 2018. Microbial Interkingdom Interactions in Roots Promote Arabidopsis Survival. *Cell* **175**, 973-83.e14.
- Failor KC, Schmale DG, 3rd, Vinatzer BA, Monteil CL, 2017. Ice nucleation active bacteria in precipitation are genetically diverse and nucleate ice by employing different mechanisms. *The ISME Journal* **11**, 2740-53.
- Fitzpatrick CR, Copeland J, Wang PW, Guttman DS, Kotanen PM, Johnson MTJ, 2018. Assembly and ecological function of the root microbiome across angiosperm plant species. *Proceedings of the National Academy of Sciences* **115**, E1157.
- Goh C-H, Veliz Vallejos DF, Nicotra AB, Mathesius U, 2013. The Impact of Beneficial Plant-Associated Microbes on Plant Phenotypic Plasticity. *Journal of Chemical Ecology* **39**, 826-39.
- Grady KL, Sorensen JW, Stopnisek N, Guittar J, Shade A, 2019. Assembly and seasonality of core phyllosphere microbiota on perennial biofuel crops. *Nature Communications* **10**, 4135.
- Graham BE, Stegen CJ, 2017. Dispersal-Based Microbial Community Assembly Decreases Biogeochemical Function. *Processes* **5**.
- Hacquard S, Garrido-Oter R, González A, *et al.*, 2015. Microbiota and Host Nutrition across Plant and Animal Kingdoms. *Cell Host & Microbe* **17**, 603-16.
- Hirano SS, Upper CD. Bacterial Community Dynamics. In: Andrews JH, Hirano SS, eds. *Proceedings of the Microbial Ecology of Leaves, 1991*. New York, NY: Springer New York, 271-94.

- Hiraoka S, Miyahara M, Fujii K, Machiyama A, Iwasaki W, 2017. Seasonal Analysis of Microbial Communities in Precipitation in the Greater Tokyo Area, Japan. *Frontiers in microbiology* **8**, 1506-.
- Juul S, Izquierdo F, Hurst A, *et al.*, 2015. What's in my pot? Real-time species identification on the MinION™. *bioRxiv*, 030742.
- Kang DD, Froula J, Egan R, Wang Z, 2015. MetaBAT, an efficient tool for accurately reconstructing single genomes from complex microbial communities. *PeerJ* **3**, e1165-e.
- Kembel SW, O'connor TK, Arnold HK, Hubbell SP, Wright SJ, Green JL, 2014. Relationships between phyllosphere bacterial communities and plant functional traits in a neotropical forest. *Proceedings of the National Academy of Sciences* **111**, 13715.
- Kim D, Song L, Breitwieser FP, Salzberg SL, 2016. Centrifuge: rapid and sensitive classification of metagenomic sequences. *Genome research* **26**, 1721-9.
- Knief C, Delmotte N, Chaffron S, *et al.*, 2012. Metaproteogenomic analysis of microbial communities in the phyllosphere and rhizosphere of rice. *The ISME Journal* **6**, 1378-90.
- Knief C, Ramette A, Frances L, Alonso-Blanco C, Vorholt JA, 2010. Site and plant species are important determinants of the Methylobacterium community composition in the plant phyllosphere. *The ISME Journal* **4**, 719-28.
- Konopka A, Lindemann S, Fredrickson J, 2015. Dynamics in microbial communities: unraveling mechanisms to identify principles. *The ISME Journal* **9**, 1488-95.
- Langmead B, Salzberg SL, 2012. Fast gapped-read alignment with Bowtie 2. *Nature methods* **9**, 357-9.
- Li H, Handsaker B, Wysoker A, *et al.*, 2009. The Sequence Alignment/Map format and SAMtools. *Bioinformatics (Oxford, England)* **25**, 2078-9.

- Lindemann J, Constantinidou HA, Barchet WR, Upper CD, 1982. Plants as Sources of Airborne Bacteria, Including Ice Nucleation-Active Bacteria. *Applied and Environmental Microbiology* **44**, 1059.
- Lindemann J, Upper CD, 1985. Aerial Dispersal of Epiphytic Bacteria over Bean Plants. *Applied and Environmental Microbiology* **50**, 1229-32.
- Lindow SE, Brandl MT, 2003. Microbiology of the Phyllosphere. *Applied and Environmental Microbiology* **69**, 1875.
- Links MG, Demeke T, Gräfenhan T, Hill JE, Hemmingsen SM, Dumonceaux TJ, 2014. Simultaneous profiling of seed-associated bacteria and fungi reveals antagonistic interactions between microorganisms within a shared epiphytic microbiome on Triticum and Brassica seeds. *The New phytologist* **202**, 542-53.
- Love MI, Huber W, Anders S, 2014. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biology* **15**, 550.
- Lu T, Ke M, Lavoie M, *et al.*, 2018. Rhizosphere microorganisms can influence the timing of plant flowering. *Microbiome* **6**, 231.
- Maignien L, Deforce EA, Chafee ME, Eren AM, Simmons SL, 2014. Ecological Succession and Stochastic Variation in the Assembly of Arabidopsis thaliana Phyllosphere Communities. *mBio* **5**, e00682-13.
- McMurdie PJ, Holmes S, 2013. phyloseq: An R Package for Reproducible Interactive Analysis and Graphics of Microbiome Census Data. *PLOS ONE* **8**, e61217.
- Mechan-Llontop ME, Tian L, Bernal-Galeano V, *et al.*, 2019. Assessing the potential of culture-independent 16S rRNA microbiome analysis in disease diagnostics: the example of Dianthus gratianopolitanus and Robbsia andropogonis. *European Journal of Plant Pathology*.
- Miletto M, Lindow SE, 2015. Relative and contextual contribution of different sources to the composition and abundance of indoor air bacteria in residences. *Microbiome* **3**, 61.

- Monteil CL, Bardin M, Morris CE, 2014. Features of air masses associated with the deposition of *Pseudomonas syringae* and *Botrytis cinerea* by rain and snowfall. *The ISME Journal* **8**, 2290-304.
- Monteil CL, Yahara K, Studholme DJ, *et al.*, 2016. Population-genomic insights into emergence, crop adaptation and dissemination of *Pseudomonas syringae* pathogens. *Microbial genomics* **2**, e000089-e.
- Morella NM, Zhang X, Koskella B, 2019. Tomato Seed-Associated Bacteria Confer Protection of Seedlings Against Foliar Disease Caused by *Pseudomonas syringae*. *Phytobiomes Journal* **3**, 177-90.
- Morris CE, 2002. Phyllosphere. In. *eLS*.
- Morris CE, Conen F, Alex Huffman J, Phillips V, Pöschl U, Sands DC, 2014. Bioprecipitation: a feedback cycle linking Earth history, ecosystem dynamics and land use through biological ice nucleators in the atmosphere. *Global Change Biology* **20**, 341-51.
- Morris CE, Sands DC, Vinatzer BA, *et al.*, 2008. The life history of the plant pathogen *Pseudomonas syringae* is linked to the water cycle. *Isme j* **2**, 321-34.
- Ottesen AR, González Peña A, White JR, *et al.*, 2013. Baseline survey of the anatomical microbial ecology of an important food plant: *Solanum lycopersicum* (tomato). *BMC Microbiology* **13**, 114.
- Ottesen AR, Gorham S, Reed E, *et al.*, 2016. Using a Control to Better Understand Phyllosphere Microbiota. *PLOS ONE* **11**, e0163482.
- Peng Y, Leung HCM, Yiu SM, Chin FYL, 2012. IDBA-UD: a de novo assembler for single-cell and metagenomic sequencing data with highly uneven depth. *Bioinformatics (Oxford, England)* **28**, 1420-8.



- Pérez-Jaramillo JE, De Hollander M, Ramírez CA, Mendes R, Raaijmakers JM, Carrión VJ, 2019. Deciphering rhizosphere microbiome assembly of wild and modern common bean (*Phaseolus vulgaris*) in native and agricultural soils from Colombia. *Microbiome* **7**, 114.
- Peter H, Hörtnagl P, Reche I, Sommaruga R, 2014. Bacterial diversity and composition during rain events with and without Saharan dust influence reaching a high mountain lake in the Alps. *Environmental microbiology reports* **6**, 618-24.
- Polymenakou PN, 2012. Atmosphere: A Source of Pathogenic or Beneficial Microbes? *Atmosphere* **3**.
- Powell JR, Karunaratne S, Campbell CD, Yao H, Robinson L, Singh BK, 2015. Deterministic processes vary during community assembly for ecologically dissimilar taxa. *Nature Communications* **6**, 8444.
- Punja ZK, Rodriguez G, Tirajoh A, 2016. Effects of *Bacillus subtilis* strain QST 713 and storage temperatures on post-harvest disease development on greenhouse tomatoes. *Crop Protection* **84**, 98-104.
- Quast C, Pruesse E, Yilmaz P, *et al.*, 2013. The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic Acids Res* **41**, D590-6.
- Regalado J, Lundberg DS, Deusch O, *et al.*, 2019. Combining whole genome shotgun sequencing and rDNA amplicon analyses to improve detection of microbe-microbe interaction networks in plant leaves. *bioRxiv*, 823492.
- Ritpitakphong U, Falquet L, Vimoltust A, Berger A, Métraux J-P, L'haridon F, 2016. The microbiome of the leaf surface of *Arabidopsis* protects against a fungal pathogen. *New Phytologist* **210**, 1033-43.
- Šantl-Temkiv T, Finster K, Dittmar T, *et al.*, 2013. Hailstones: A Window into the Microbial and Chemical Inventory of a Storm Cloud. *PLOS ONE* **8**, e53550.
- Schmale DG, Ross SD, 2015. Highways in the Sky: Scales of Atmospheric Transport of Plant Pathogens. *Annual Review of Phytopathology* **53**, 591-611.

- Shade A, Mcmanus PS, Handelsman J, 2013. Unexpected diversity during community succession in the apple flower microbiome. *mBio* **4**, e00602-12.
- Toju H, Okayasu K, Notaguchi M, 2019. Leaf-associated microbiomes of grafted tomato plants. *Scientific Reports* **9**, 1787.
- Torres-Cortés G, Bonneau S, Bouchez O, *et al.*, 2018. Functional Microbial Features Driving Community Assembly During Seed Germination and Emergence. *Frontiers in Plant Science* **9**.
- Vacher C, Hampe A, Porté AJ, Sauer U, Compant S, Morris CE, 2016. The Phyllosphere: Microbial Jungle at the Plant–Climate Interface. *Annual Review of Ecology, Evolution, and Systematics* **47**, 1-24.
- Vařtilingom M, Attard E, Gaiani N, *et al.*, 2012. Long-term features of cloud microbiology at the puy de Dôme (France). *Atmospheric Environment* **56**, 88-100.
- Van Bruggen AHC, Goss EM, Havelaar A, Van Diepeningen AD, Finckh MR, Morris JG, 2019. One Health - Cycling of diverse microbial communities as a connecting force for soil, plant, animal, human and ecosystem health. *Science of The Total Environment* **664**, 927-37.
- Vorholt JA, 2012. Microbial life in the phyllosphere. *Nature Reviews Microbiology* **10**, 828-40.
- Wagner MR, Lundberg DS, Del Rio TG, Tringe SG, Dangl JL, Mitchell-Olds T, 2016. Host genotype and age shape the leaf and root microbiomes of a wild perennial plant. *Nature Communications* **7**, 12151-.
- Wei Z, Gu Y, Friman V-P, *et al.*, 2019. Initial soil microbiome composition and functioning predetermine future plant health. *Science Advances* **5**, eaaw0759.
- Wickham H, 2009. *Ggplot2: Elegant graphics for data analysis*. Springer.
- Williams TR, Moyne A-L, Harris LJ, Marco ML, 2013. Season, Irrigation, Leaf Age, and *Escherichia coli* Inoculation Influence the Bacterial Diversity in the Lettuce Phyllosphere. *PLOS ONE* **8**, e68642.

- Woo C, An C, Xu S, Yi S-M, Yamamoto N, 2018. Taxonomic diversity of fungi deposited from the atmosphere. *The ISME Journal* **12**, 2051-60.
- Zarraonaindia I, Owens SM, Weisenhorn P, *et al.*, 2015. The Soil Microbiome Influences Grapevine-Associated Microbiota. *mBio* **6**, e02527-14.
- Zhou J, Ning D, 2017. Stochastic Community Assembly: Does It Matter in Microbial Ecology? *Microbiology and Molecular Biology Reviews* **81**, e00002-17.

## 6. Tables:

**Table 1.** Rain fall collections

<b>Rain event</b>	<b>Date of collection</b>	<b>Season</b>	<b>Location</b>
VTP049	Apr-15	Spring	Latham Hall
VTP050	Aug-15	Summer	Latham Hall
VTP061	Mar-16	Spring	Latham Hall
VTP062	Apr-16	Spring	Latham Hall
VTP063	May-16	Spring	Latham Hall
VTP064	Jul-16	Summer	Latham Hall
VTP065	Oct-16	Fall	Latham Hall
VTP066	Dec-16	Winter	Latham Hall

**Table 2.** Pairwise comparisons using the Wilcoxon rank-sum test to evaluate differences in alpha diversity indexes in rain samples and plants treated at day 0 and day 7

<b>Observed OTUs</b>	Filtered Rain day 0	Filtered Rain day 7	Rain	100X-Rain day 0	100X-Rain day 7	dd water day 0
Filtered Rain day7	0.0591	-	-	-	-	-
Rain	0.3595	0.0066	-	-	-	-
100X-Rain day 0	1	0.0413	1	-	-	-
100X-Rain day 7	0.0178	1	0.0035	0.0178	-	-
dd water day 0	0.7636	0.4364	0.2286	0.4364	1	-
dd water day 7	0.0213	1	0.0152	0.04	1	0.2286

<b>Shannon index</b>	Filtered Rain day 0	Filtered Rain day 7	Rain	100X-Rain day 0	100X-Rain day 7	dd water day 0
Filtered Rain day 7	0.039	-	-	-	-	-
Rain	1	0.049	-	-	-	-
100X-Rain day 0	0.379	0.053	1	-	-	-
100X-Rain day 7	0.049	1	0.079	0.079	-	-
dd water day 0	1	0.339	1	1	0.369	-
dd water day 7	0.079	1	0.079	0.079	1	0.379

**Table 3.** . Pairwise comparisons using the Wilcoxon rank-sum test to evaluate differences in alpha diversity indexes in tomato plants grown in a hydroponic system, soil and outside.

<b>Observed OTUs</b>	Hydroponics	Soil
Soil	0.036	-
Outside	0.876	0.876

<b>Shannon index</b>	Hydroponics	Soil
Soil	0.25	-
Outside	0.91	0.96

**Table 4.** Core microbiome analysis. OTUs present on tomato plants grown in a hydroponic system, tomato plants grown in soil and tomato plants grown outside.

Plant grown in Hydroponic (n = 29)		Plant grown in soil (n = 18)		Plant grown outside (n = 3)	
Genus	# OTUs	Genus	# OTUs	Genus	# OTUs
<i>Methylobacterium</i>	2	<i>Bacillus</i>	7	<i>Sphingomonas</i>	22
<i>Propionibacterium</i>	1	<i>Massilia</i>	1	<i>Methylobacterium</i>	18
<i>Bacillus</i>	1	<i>Methylobacterium</i>	2	<i>Rhizobium</i>	12
Rhizobiales OUT	1	<i>Microbacterium</i>	2	<i>Massilia</i>	11
		<i>Pseudomonas</i>	1	<i>Acidiphilium</i>	9
		<i>Raoultella</i>	1	<i>Pantoea</i>	7
		<i>Staphylococcus</i>	3	<i>Bacillus</i>	6
		<i>Propionibacterium</i>	1	<i>Pseudomonas</i>	6
		<i>Curtobacterium</i>	1	<i>Mycobacterium</i>	6
		<i>Brevibacterium</i>	1	<i>Brevundimonas</i>	4
				<i>Bradyrhizobium</i>	4
				<i>Roseomonas</i>	3
				<i>Microbacterium</i>	3
				<i>Bryocella</i>	3
				<i>Novosphingobium</i>	3
				<i>Lysinibacillus</i>	2
				<i>Gaiella</i>	2
				<i>Granulicella</i>	2
				<i>Devosia</i>	2
				<i>Endobacter</i>	2
				<i>Curtobacterium</i>	2
				<i>Nocardioides</i>	2
				<i>Brevibacterium</i>	2
				<i>Amnibacterium</i>	2
				<i>Methylocella</i>	2
				<i>Rhodococcus</i>	2
				<i>Kineococcus</i>	2
				<i>Modestobacter</i>	2
				<i>Wolbachia</i>	1
				<i>Bosea</i>	1
				<i>Exiguobacterium</i>	1
				<i>Friedmanniella</i>	1
				<i>Blastococcus</i>	1
				<i>Actinomycetospora</i>	1
				<i>Actinoplanes</i>	1

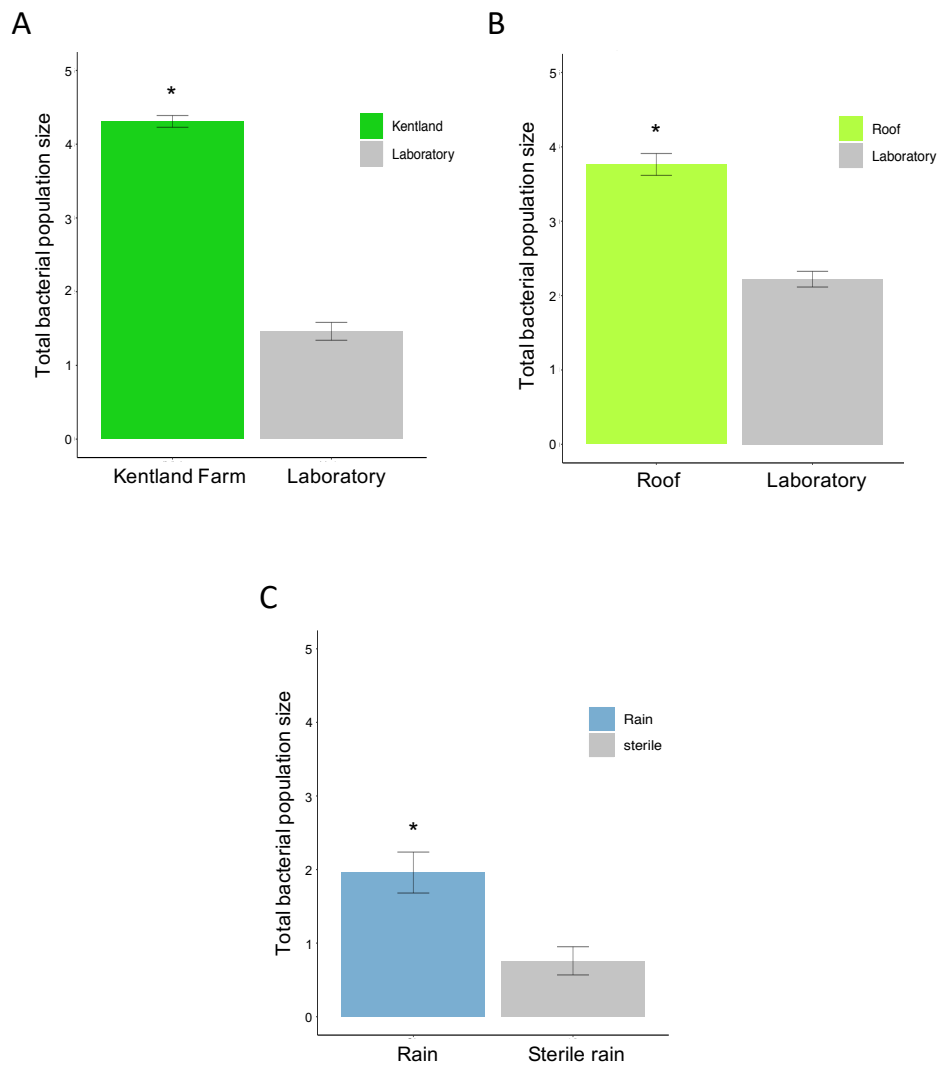
---

<i>Zymomonas</i>	1
<i>Tuberibacillus</i>	1
<i>Rubellimicrobium</i>	1
<i>Tatumella</i>	1
<i>Lysobacter</i>	1
<i>Amaricoccus</i>	1
<i>Methylophilus</i>	1
<i>Propionibacterium</i>	1
<i>Williamsia</i>	1
<i>Clostridium</i>	1
<i>Edaphobacter</i>	1
<i>Stenotrophomonas</i>	1
<i>Turcibacter</i>	1
<i>Spirosoma</i>	1
<i>Pullulanibacillus</i>	1
<i>Rhodopseudomonas</i>	1
<i>Frigoribacterium</i>	1
<i>Variovorax</i>	1
<i>Candidatus</i>	1
<i>Mesorhizobium</i>	1
<i>Aeromicrobium</i>	1
<i>Aureimonas</i>	1
<i>Streptomyces</i>	1
<i>Quadrisphaera</i>	1

---

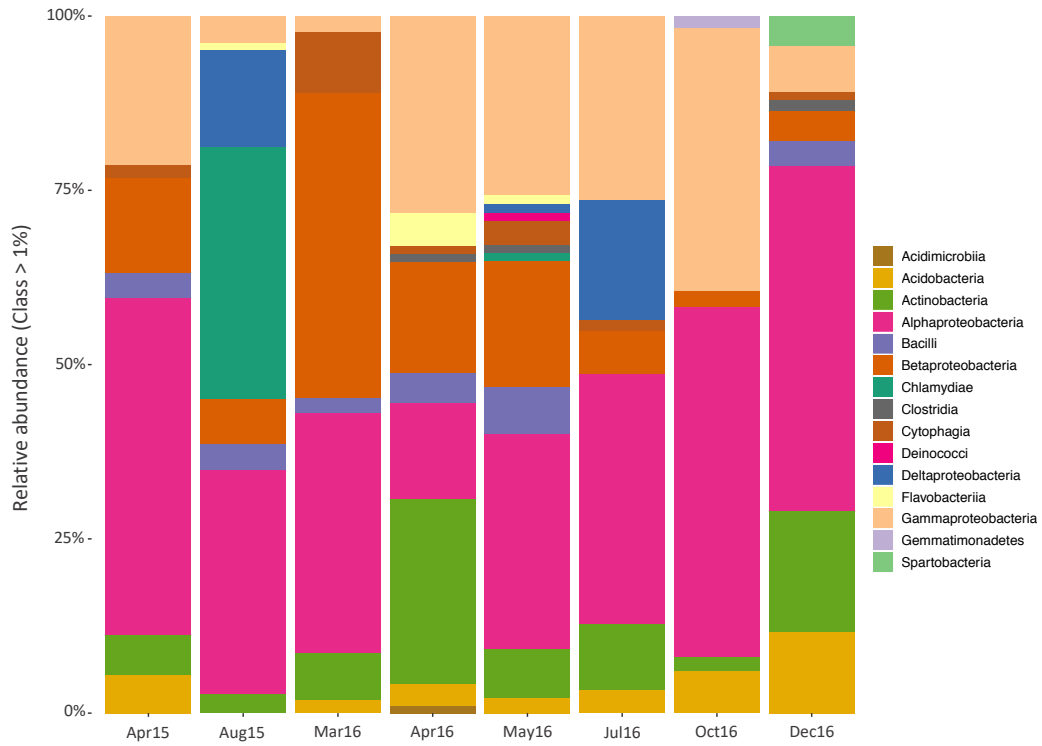
## 7. Figures

**Figure 1.** Bacterial population size in the phyllosphere of tomato plants A) grown in a plastic pot at the Kentland research farm compared to plants grown under laboratory conditions B) grown on the roof of a research building compared to plants grown under laboratory conditions, and C) grown under laboratory conditions treated with rain and autoclaved rain. T-test,  $P < 0.001$ .

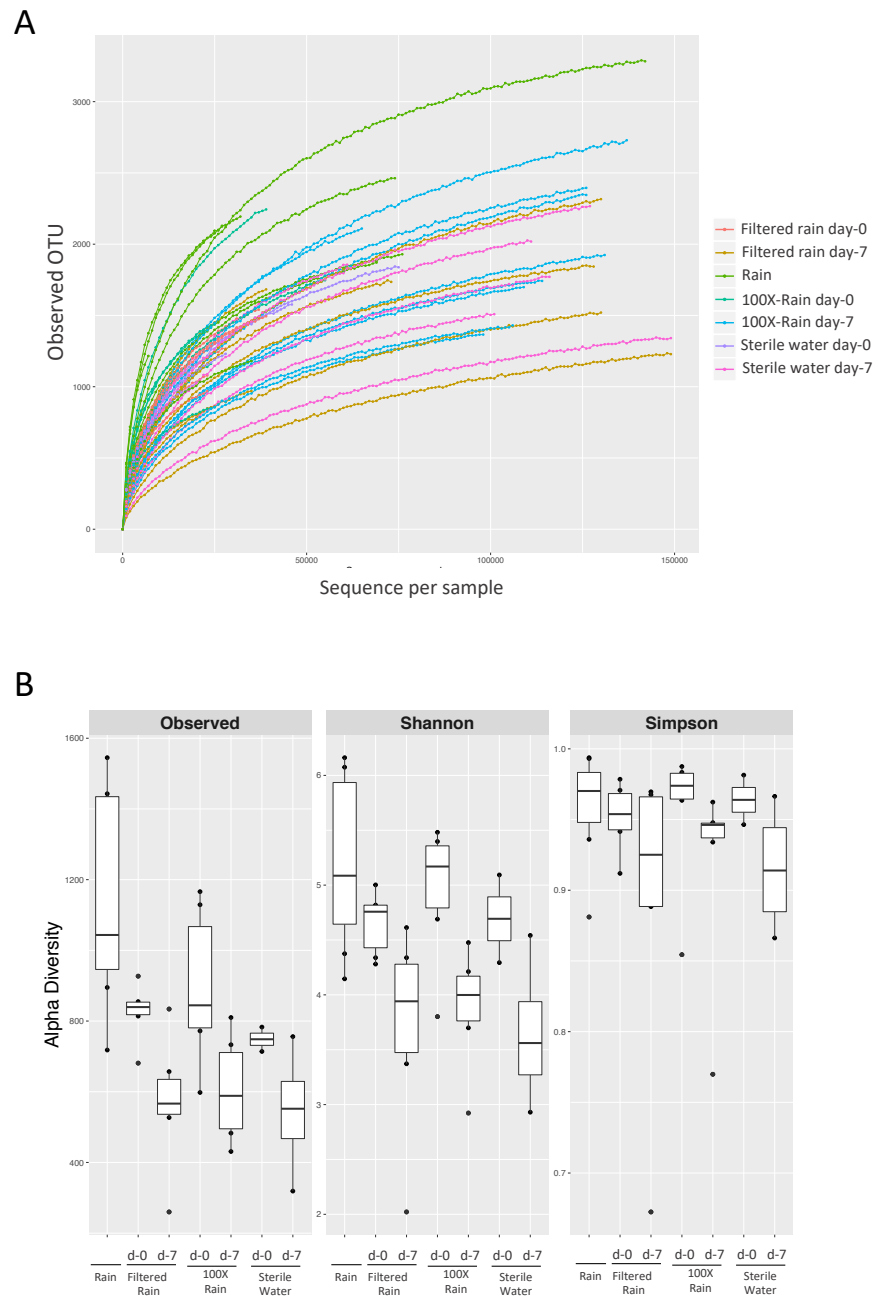




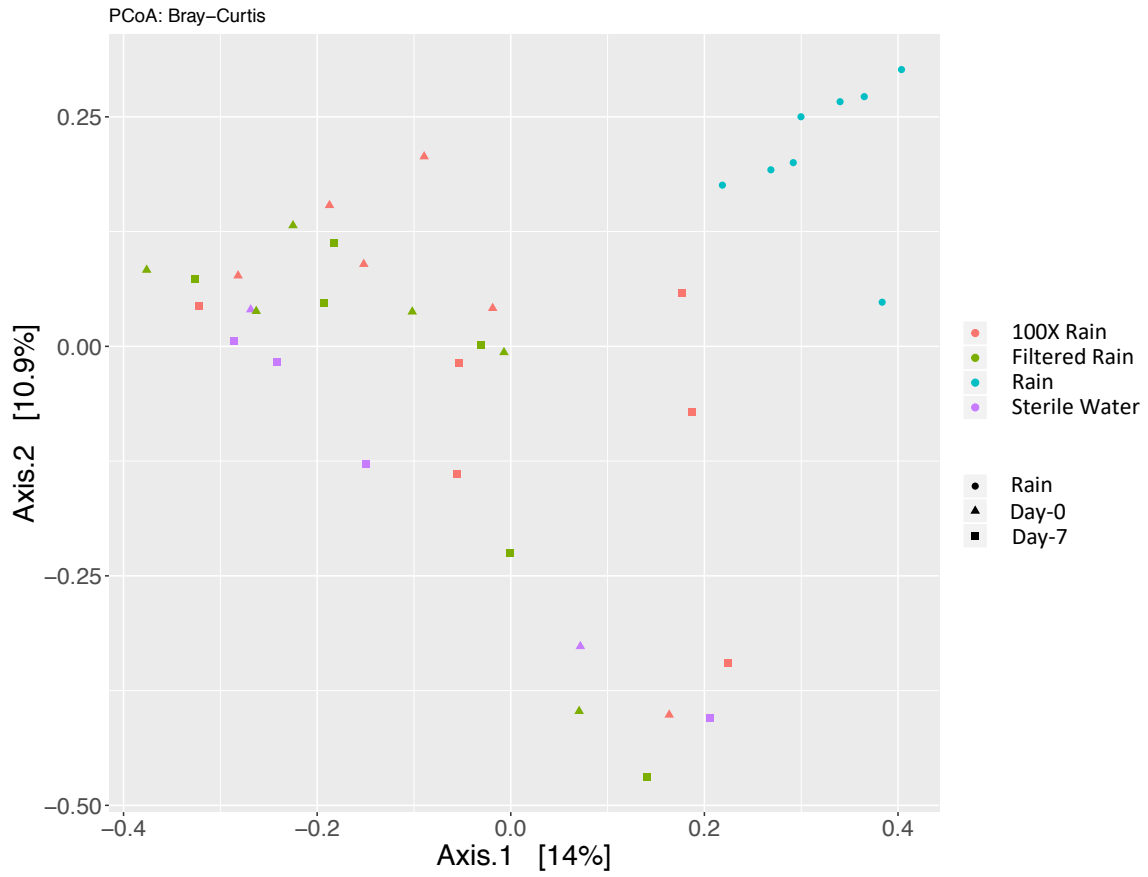
**Figure 2.** Relative abundance of bacterial taxa in rainfall collected in Blacksburg, VA. at the class level (abundance > 1%).



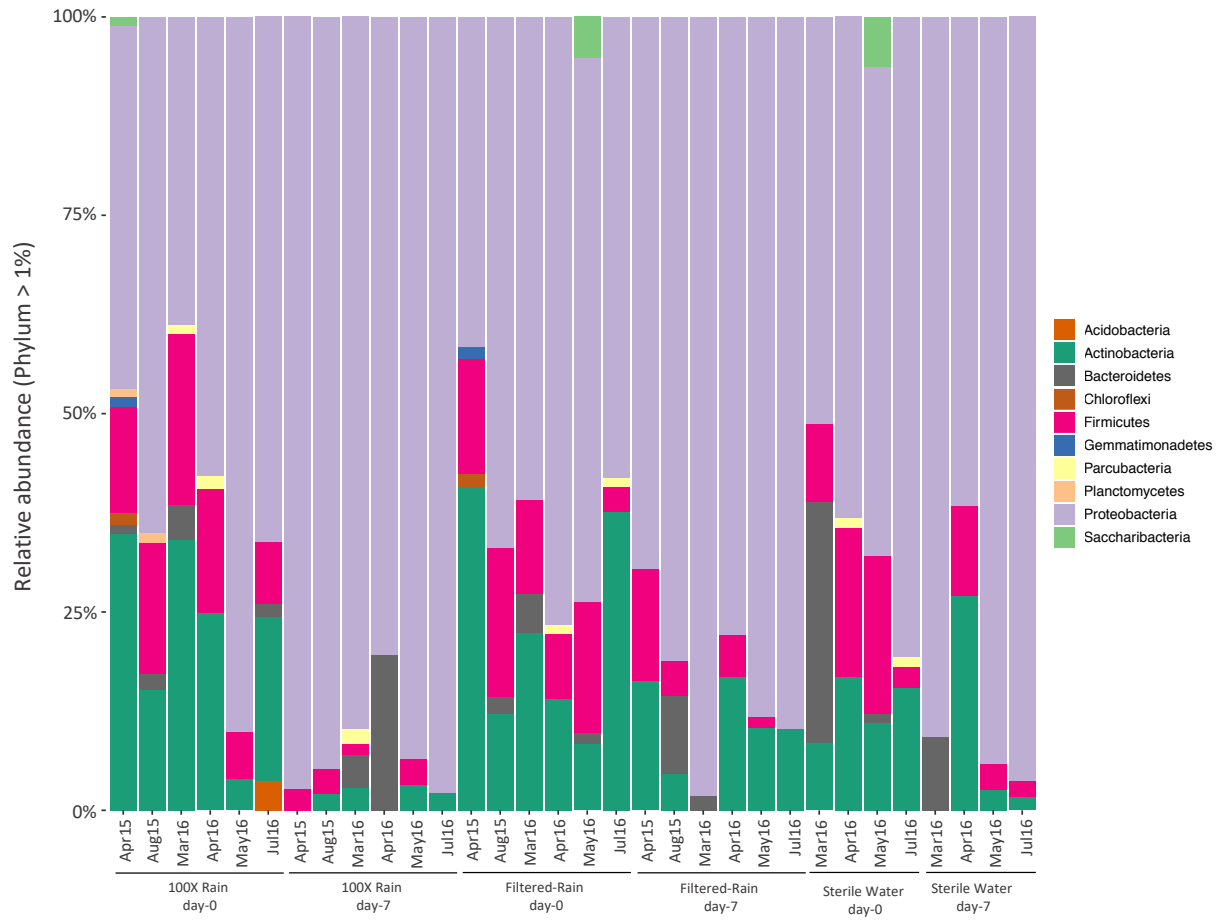
**Figure 3.** Alpha diversity measurements for rain compared to treated plants at day 0 vs day 7. A) Rarefaction curves showing the richness of samples, and B) three measures of alpha diversity (observed OTUs, Shannon diversity index, and Simpson diversity index).



**Figure 4.** Principal coordinates analysis (PCoA) derived from the dissimilarity matrix of Bray-Curtis is shown for all rain samples and plants treated at day 0 and day 7.

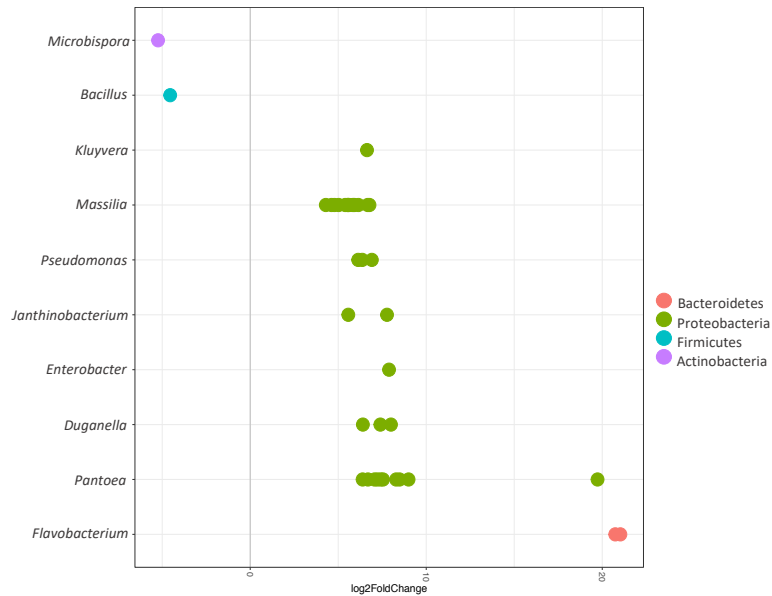


**Figure 5.** Relative abundance of bacterial taxa of all rain samples and plants treated at day 0 and day 7 at the phylum level (abundance > 1%).

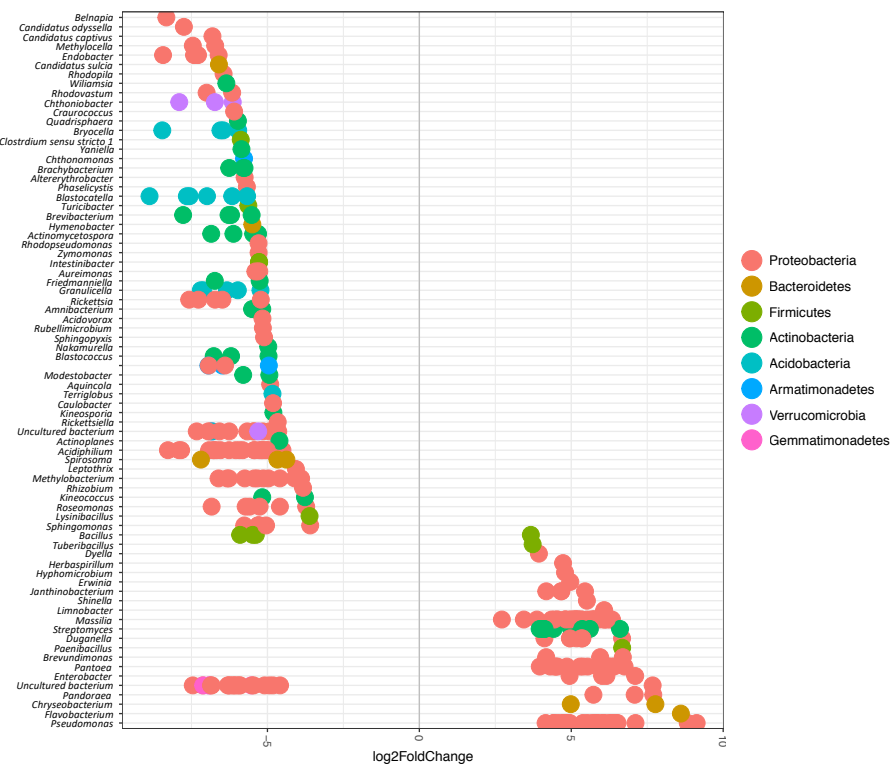


**Figure 6:** Operational Taxonomic Unit (OTU) differential abundance. A) differences at the OUT level between plants treated with 100X-rain at the day 0 and day 7 time points, B) differences at the OUT level between rain and tomato plants treated with 100X- rain at the day 7 time point.

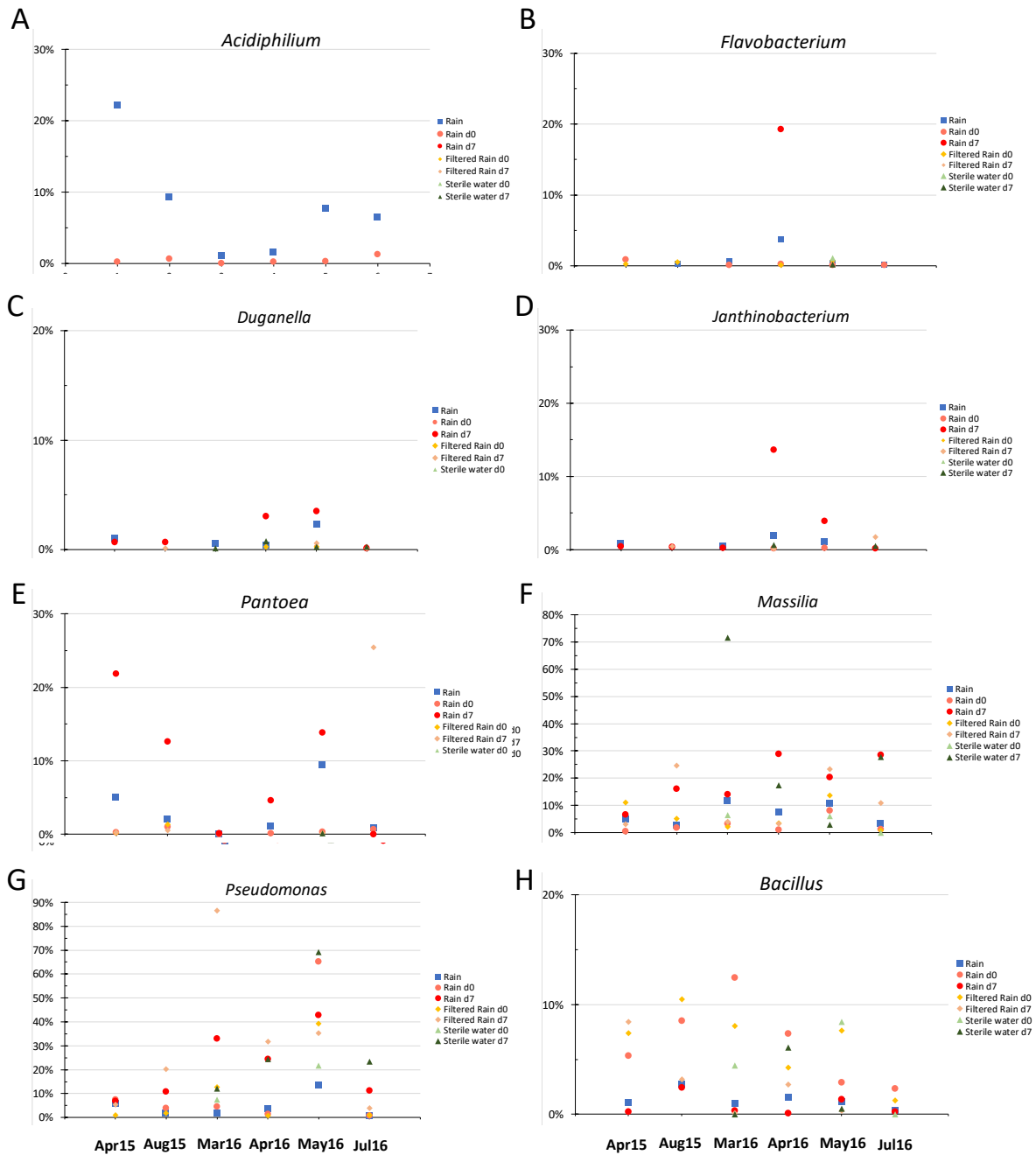
A



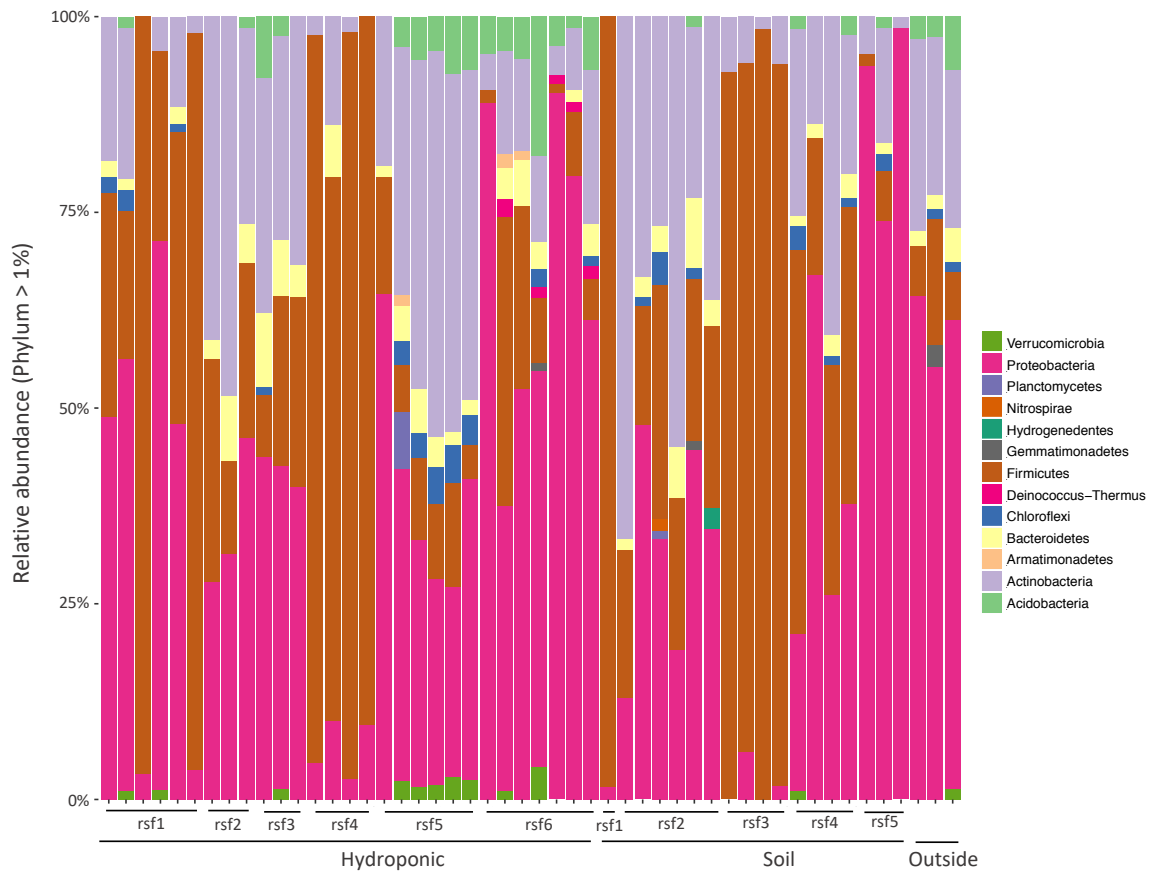
B



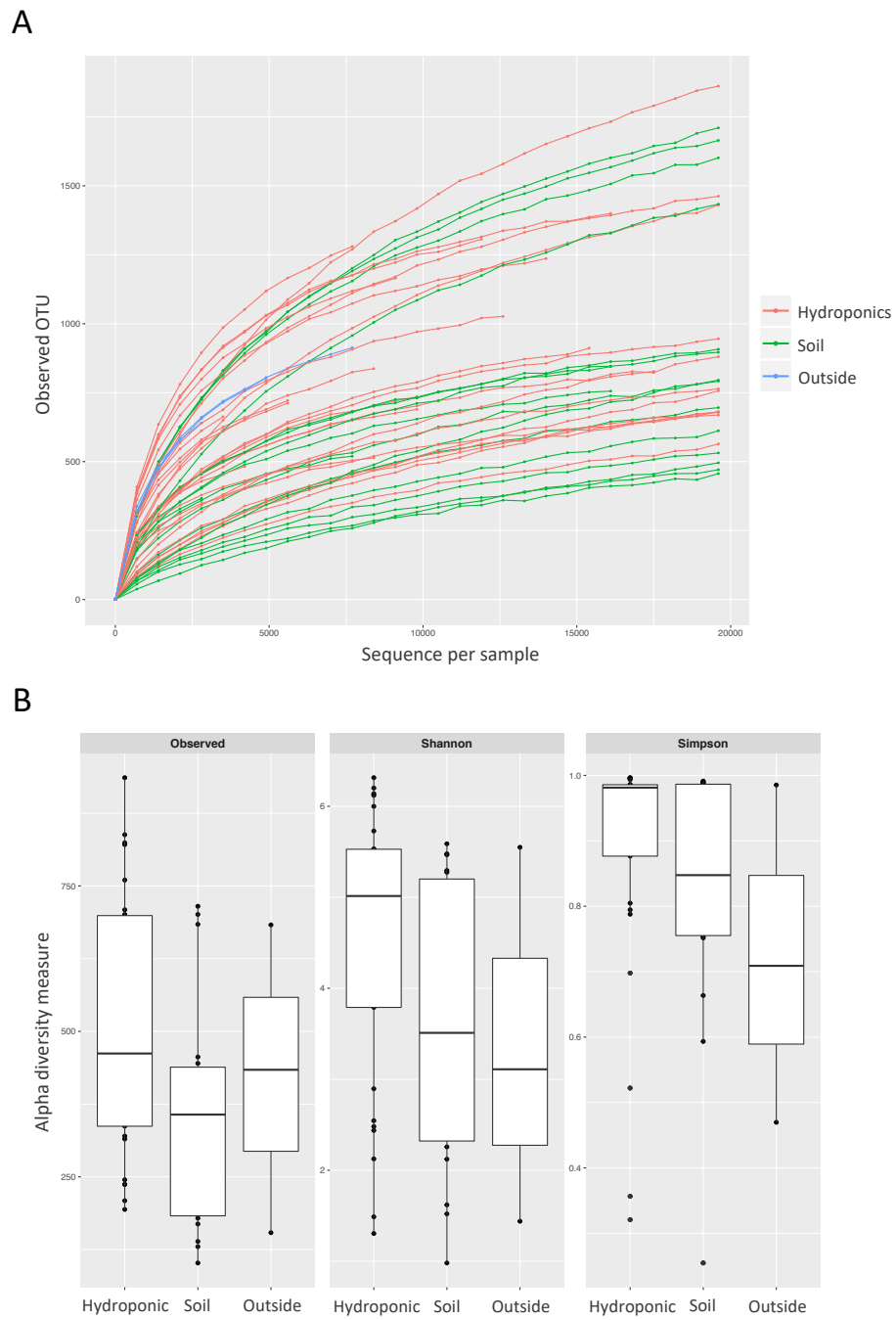
**Figure 7.** Relative abundance of rain-borne bacterial taxa that either failed or succeed in colonizing the tomato phyllosphere under controlled laboratory conditions. A) *Acidiphilium*, B) *Flavobacterium*, C) *Duganella*, D) *Janthinobacterium*, E) *Pantoea*, F) *Massilia*, G) *Pseudomonas*, and H) *Bacillus*.



**Figure 8.** Relative abundance of bacterial taxa of all tomato plants grown in a hydroponic system, tomato plants grown in soil (both in a commercial greenhouse) and tomato plants grown outside (in a roof of a research building) at the phylum level (abundance > 1%).



**Figure 9.** Alpha diversity measurements for tomato plants grown in a hydroponic system, in soil and outside A) Rarefaction curves showing the richness of samples, and B) three measures of alpha diversity (observed OTUs, Shannon diversity index, and Simpson diversity index).



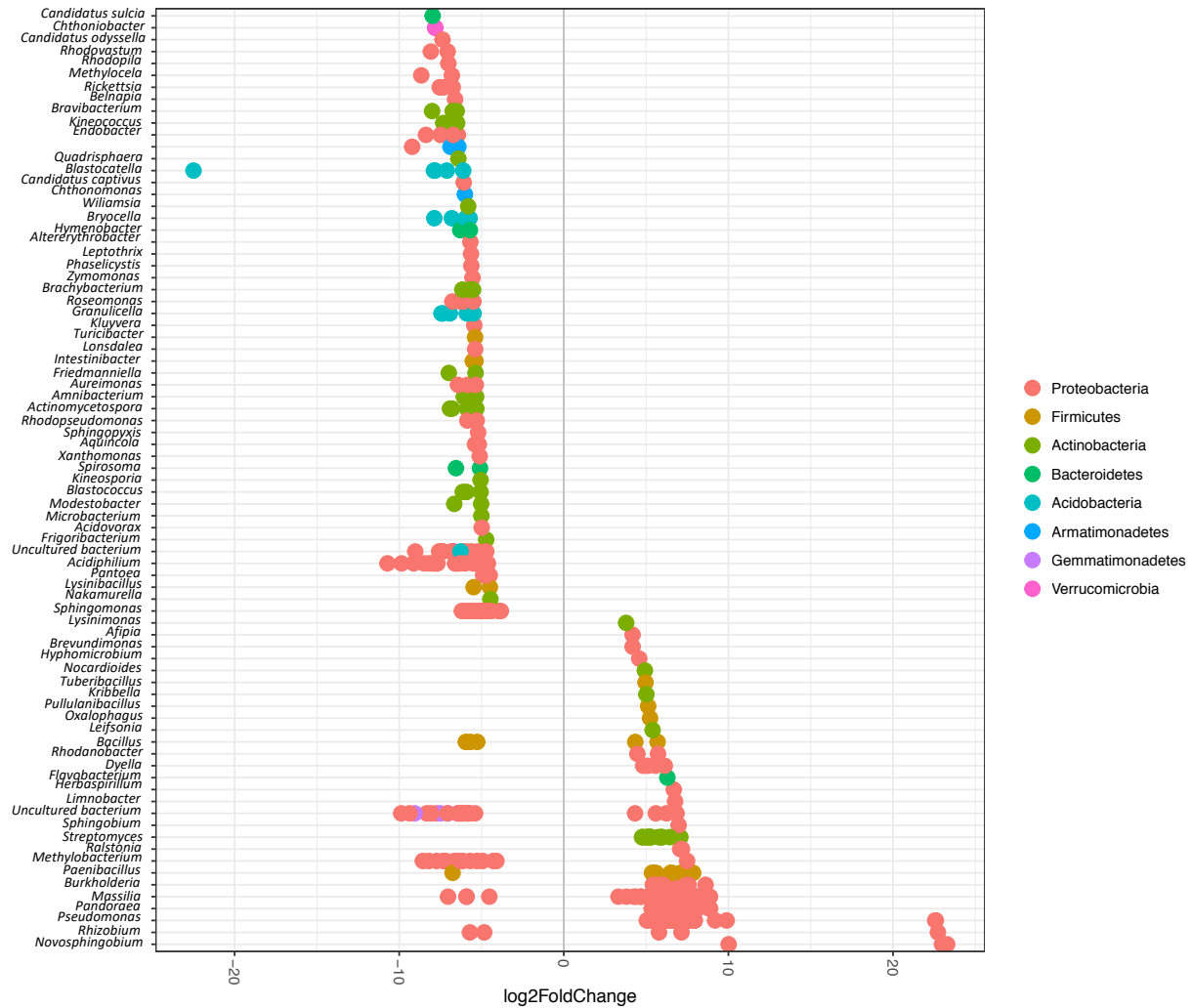


**Figure 10.** Abundance of rain samples and tomato plants treated with 100X-rain at time point 0 and 7 days post inoculation using metagenome shotgun sequencing.

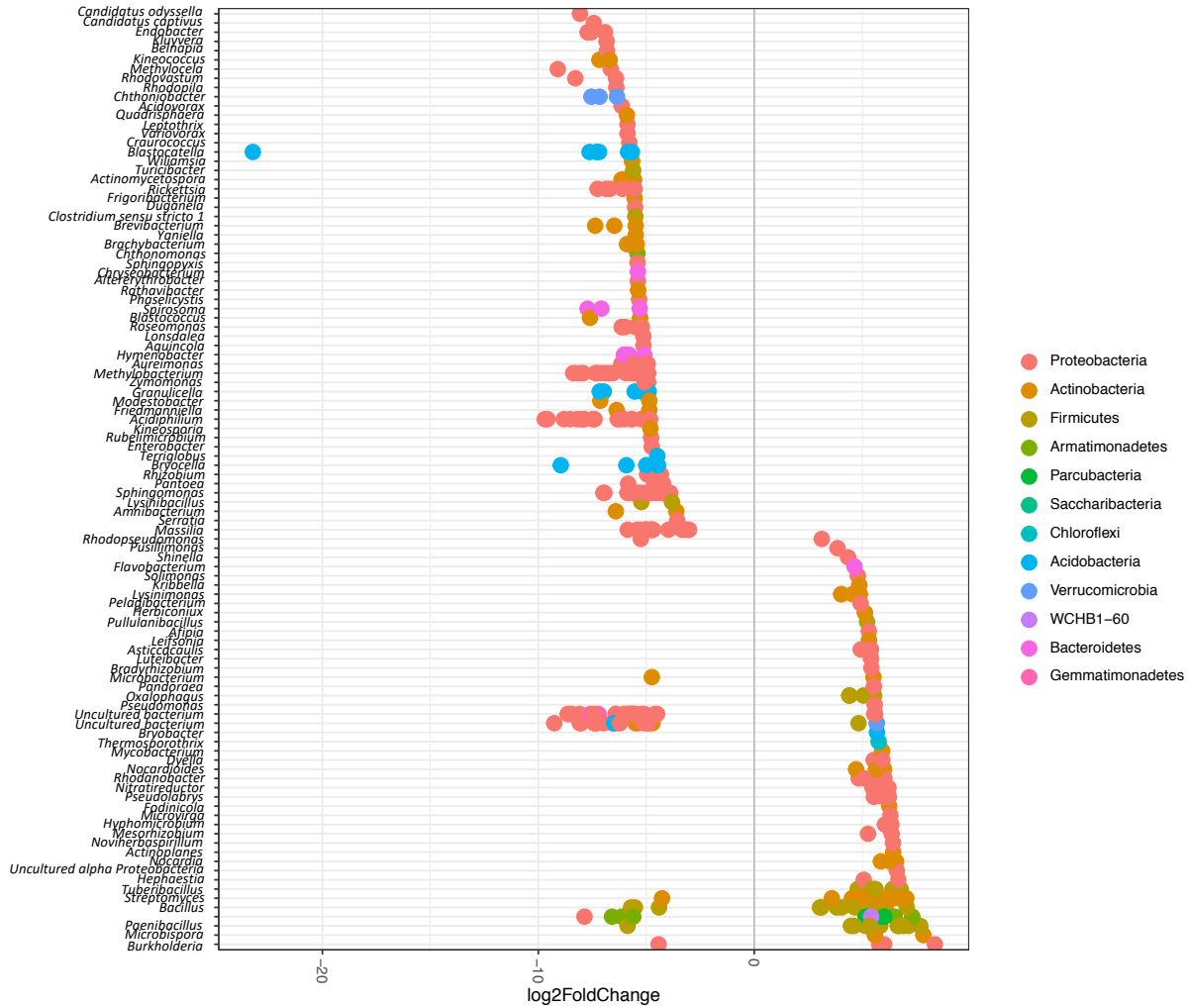
Bacterial species	Phyllosphere treated with 100X-rain at day-7					Average abundance in rain	Average abundance in day-7	Highest abundance
	Apr 15	Aug15	Nov15	Mar16	Apr16			
<i>Pantoea vagans</i>						0.088%	2.865%	8.691%
<i>Pantoea agglomerans</i>						0.025%	2.123%	4.246%
<i>Pseudomonas citronellolis</i>						0.006%	1.208%	1.975%
<i>Novosphingobium resinovorum</i>						0.004%	1.535%	1.535%
<i>Buttiauxella</i> sp. 3AFRM03						0.008%	0.786%	1.490%
<i>Erwinia gerundensis</i>						0.010%	0.753%	1.313%
<i>Pseudomonas fluorescens</i>						0.028%	0.753%	1.156%
<i>Cedecea neteri</i>						0.012%	0.533%	1.003%
<i>Massilia</i> sp. WG5						0.011%	0.455%	0.903%
<i>Pseudomonas orientalis</i>						0.004%	0.532%	0.789%
<i>Janthinobacterium</i> sp. 1_2014MBL_MicDiv						0.007%	0.765%	0.765%
<i>Massilia putida</i>						0.013%	0.627%	0.756%
<i>Janthinobacterium agaricidamnorum</i>						0.006%	0.617%	0.617%
<i>Pseudomonas azotoformans</i>						0.005%	0.311%	0.479%
<i>Flavobacterium</i> sp. HYN0086						0.055%	0.411%	0.411%
<i>Flavobacterium anhuiense</i>						0.083%	0.401%	0.401%
<i>Pseudomonas</i> sp. LG1E9						0.005%	0.269%	0.334%
<i>Flavobacterium</i> sp. HYN0056						0.043%	0.331%	0.331%
<i>Pseudomonas rhizosphaerae</i>						0.008%	0.233%	0.273%
<i>Pseudomonas putida</i>						0.021%	0.268%	0.268%
<i>Enterobacter</i> sp. SA187						0.002%	0.257%	0.257%
<i>Klebsiella michiganensis</i>						0.004%	0.255%	0.255%
<i>Pseudomonas protegens</i>						0.005%	0.249%	0.249%
<i>Pantoea ananatis</i>						0.004%	0.244%	0.244%
<i>Pseudomonas trivialis</i>						0.002%	0.243%	0.243%
<i>Enterobacter cloacae</i>						0.019%	0.154%	0.242%
<i>Brevundimonas</i> sp. GW460-12-10-14-LB2						0.007%	0.232%	0.232%
<i>Pseudomonas veronii</i>						0.001%	0.226%	0.226%
<i>Escherichia coli</i>						0.136%	0.150%	0.224%
<i>Pseudomonas syringae</i>						0.005%	0.217%	0.217%
<i>Sphingobium</i> sp. YG1						0.003%	0.149%	0.149%
<i>Hyphomicrobium</i> sp. MC1						0.007%	0.127%	0.127%
<i>Paraburkholderia fungorum</i>						0.004%	0.114%	0.114%
<i>Staphylococcus aureus</i>						0.115%	0.106%	0.106%
<i>Pseudomonas</i> sp. 31-12						0.004%	0.104%	0.104%
<i>Pseudomonas</i> sp. URMO17WK12:111						0.009%	0.103%	0.103%

## 8. Supplementary Figures

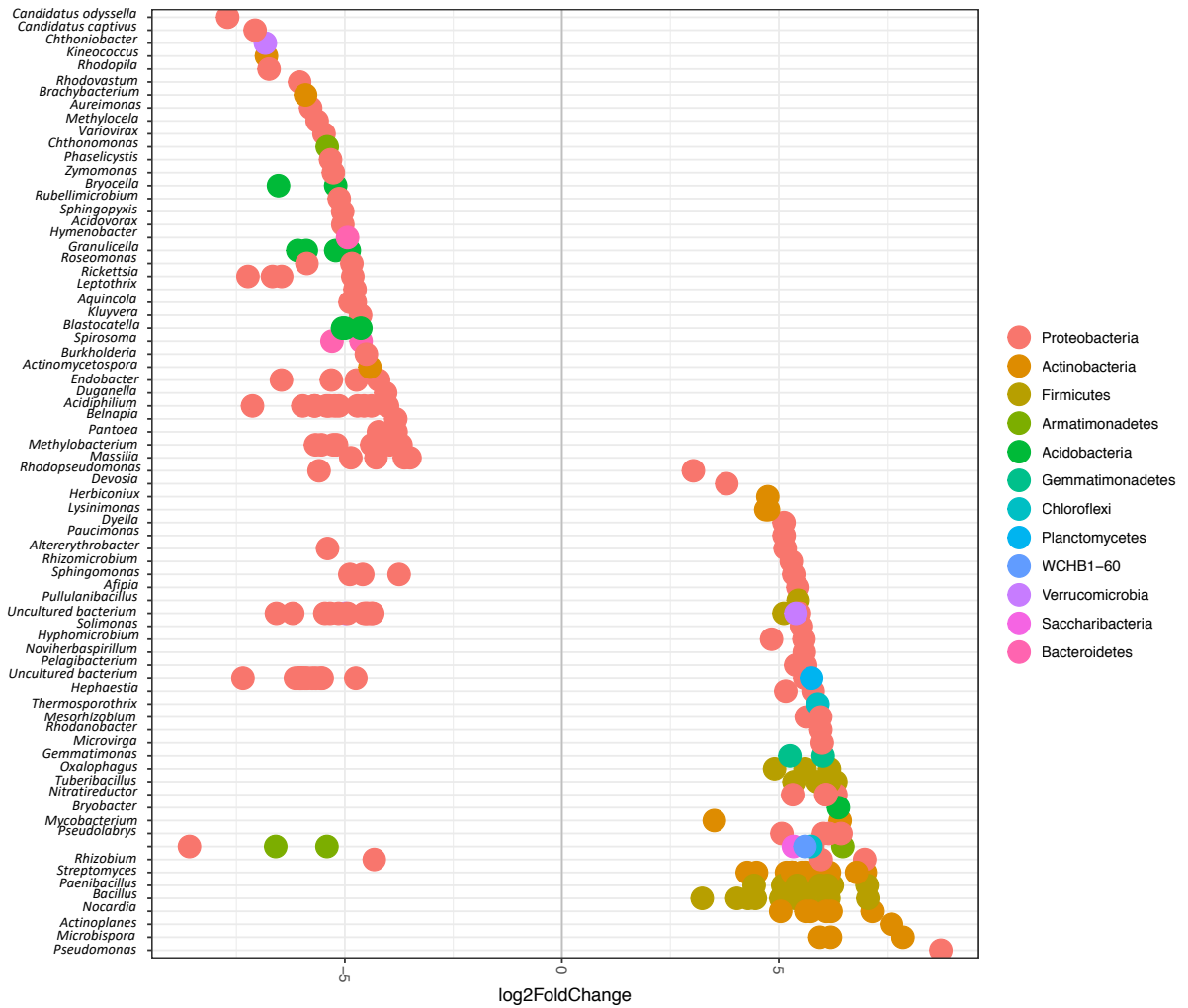
**Supplementary figure 1.** Differences at the Operational Taxonomic Unit (OTU) level between rain and tomato plants treated with sterile- water at the day 0 time point.



**Supplementary figure 3.** Differences at the Operational Taxonomic Unit (OTU) level between rain and tomato plants treated with Filtered rain at the day 0 time point.



**Supplementary figure 3.** Differences at the Operational Taxonomic Unit (OTU) level between rain and tomato plants treated with 100X-rain at the day 0 time point.



## Chapter three:

### Exploring rain as source of biological control agents for fire blight on apple

Marco E. Mechan-Llontop<sup>1</sup>, Kelly Hurley<sup>1</sup>, Long Tian<sup>1</sup>, Vivian Bernal-Galeano<sup>1</sup>, Hans Wildschutte<sup>2</sup>,  
Sasha C. Marine<sup>3</sup>, Keith S. Yoder<sup>1,4</sup> and Boris A. Vinatzer<sup>1\*</sup>

<sup>1</sup>School of Plant and Environmental Sciences, Virginia Tech, Blacksburg, VA, USA

<sup>2</sup>Department of Biological Sciences, Bowling Green State University, Bowling Green, OH, USA

<sup>3</sup>Department of Biochemistry, Virginia Tech, Blacksburg, VA, USA

<sup>4</sup>Alson H. Smith Agricultural Research and Extension Center, Virginia Tech, Winchester, VA, USA

#### \* Correspondence:

Boris A. Vinatzer

vinatzer@vt.edu

**Keywords:** rain, fire blight, biocontrol, genome-based identification, UV mutagenesis

\*\* *Submitted to the journal Frontiers in Microbiology in October 2019*

**Abstract:**

Poor survival on plants can limit the efficacy of Biological Control Agents (BCAs) in the field. Yet bacteria survive in the atmosphere, despite their exposure to high solar radiation and extreme temperatures. If conditions in the atmosphere are similar to, or more extreme than, the environmental conditions on the plant surface, then precipitation may serve as a reservoir of robust BCAs. To test this hypothesis, two hundred and fifty-four rain-borne isolates were screened for *in vitro* inhibition of *Erwinia amylovora*, the causal agent of fire blight, as well as of other plant pathogenic bacteria, fungi and oomycetes. Two isolates showed strong activity against *E. amylovora* and other plant pathogenic bacteria, while other isolates showed activity against fungal and oomycete pathogens. Survival assays suggested that the two isolates that inhibited *E. amylovora* were able to survive on apple blossoms and branches similarly to *E. amylovora*. Pathogen population size and associated fire blight symptoms were significantly reduced when detached apple blossoms were treated with the two isolates before pathogen inoculation, however, disease reduction on attached blossoms within an orchard was inconsistent. Using whole genome sequencing, the isolates were identified as *Pantoea agglomerans* and *P. ananatis*, respectively. A UV-mutagenesis screen pointed to a phenazine antibiotic D-alanylgriseoluteic acid synthesis gene cluster as being at the base of the antimicrobial activity of the *P. agglomerans* isolate. Our work reveals the potential of precipitation as an under-explored source of BCAs, whole genome sequencing as an effective approach to precisely identify BCAs, and UV-mutagenesis as a technically simple screen to investigate the genetic basis of BCAs. More field trials are needed to determine the efficacy of the identified BCAs in fire blight control.

## 1. Introduction:

There has been a growing effort in scouting for bacteria and fungi to be deployed as biological control agents (BCAs) against crop pests (Jaffuel et al., 2019) and diseases (Durán et al., 2018). The motivation behind this effort includes the emergence of fungicide resistance in plant pathogenic fungi and antibiotic resistance in plant pathogenic bacteria, as well as an increase in consumer demand for crops produced without synthetic pesticides (Dean et al., 2012; Reganold and Wachter, 2016; Sundin and Wang, 2018).

One plant disease that is challenging to control is fire blight of apple (*Malus x domestica*) and pear (*Pyrus communis*) (Norelli et al., 2003). Since this bacterial disease was first reported in the United States in the 1870s, the causal agent, *Erwinia amylovora*, has spread to Europe, Asia, and Africa and New Zealand, causing significant economic losses (CABI, 2018). The main vectors of transmission are rainfall and pollinating insects. *E. amylovora* mainly invades plants through open blossoms and wounds and infects trees systemically through the vascular system (Thomson, 1985; Koczan et al., 2011). The first sign of disease consists of droplets of bacterial ooze on the surface of infected tissues. Blossoms and young fruits are later aborted, followed by necrosis and wilting of leaves and shoots on infected branches. In the worst case, the entire tree may die (Norelli et al., 2003).

The antibiotic streptomycin sulfate is generally an efficient method to control fire blight, but antibiotic use in crop production is illegal in many countries, and antibiotic-resistant strains have emerged in several apple and pear growing regions (Loper et al., 1991; McManus and Jones, 1994; Russo et al., 2008; Förster et al., 2015; Tancos et al., 2015). Therefore, BCAs for fire blight control have been explored for many years (Ishimaru et al., 1988). Currently available commercial products include: BlightBan™ A506 (*Pseudomonas fluorescens* A506, isolated from leaves of pear trees (Wilson and Lindow, 1992)), BlightBan™ C9-1 (*Pantoea vagans* C9-1, isolated from apple stem tissue (Ishimaru et al., 1988)), Serenade Optimum™ (*Bacillus subtilis* QST713, isolated from soil), Double Nickel™ (*Bacillus amyloliquefaciens* D747, isolated from soil), Biopro™ (*Bacillus subtilis* BD170),

Bloomtime Biological™ (*Pantoea agglomerans* E325, isolated from apple blossoms (Pusey, 1999)), and Blossom Protect™ (*Aureobasidium pullulans* strains DSM 14940 and DSM 14941, isolated from leaves of apple trees in 1989, Germany). Disease suppression by these BCAs is achieved by multiple modes of action including: the production of antimicrobial compounds (Ishimaru et al., 1988; Temple et al., 2004), colonization rates higher than those of the pathogen (Wilson and Lindow, 1994), competition for nutrients (Wilson and Lindow, 1992; Lindow, 1993), induction of plant defenses (Van Wees et al., 1997; Pieterse et al., 2014; Alamri et al., 2019), or a combination of mechanisms (Neeno-Eckwall et al., 2001).

A major hurdle to the introduction of commercial BCAs is their regulation. Precise identification and thorough characterization are necessary to exclude the potential for a BCA to cause disease in plants, animals, or humans. In fact, Bloomtime Biological™ is not available in the European Union because of safety concerns in regard to the species *P. agglomerans*, which has been reported to occasionally cause human infections (Cruz et al., 2007; Büyükcam et al., 2018). Precise genome-based classification and identification methods, such as the Life Identification Number (LIN) system implemented in the LINbase web service (Tian et al., 2019), could aid in regulation of BCAs, but those methods have not been thoroughly explored within this context.

Another challenge with BCAs is that their efficacy under field conditions is more variable than synthetic pesticides since BCAs contain living organisms, whose survival on plant surfaces and/or internal tissues depends on environmental conditions (Sundin et al., 2009; Bonaterra et al., 2012). This is especially true for BCAs that are applied to aerial parts of plants, where conditions can change rapidly in regard to temperature, humidity, and solar radiation. In a previous study (Failor et al., 2017), we isolated 33,134 bacterial strains from precipitation, of which 1144 strains were found to be putative ice nucleation active strains. Later characterization determined that 551 of those strains (i) did not have ice nucleation activity and (ii) were members of plant-associated species within the genera *Bacillus*, *Pseudomonas*, and *Pantoea*. These plant-associated species are also known to include BCAs. Because



these bacterial strains were isolated from precipitation, we hypothesized that they may resist environmental stresses (solar radiation, temperature, moisture, etc.) better than bacteria isolated from soil or plant tissues. Therefore, we screened 254 of these ice nucleation inactive rain-borne isolates for inhibition of *E. amylovora* and for their ability to persist on apple trees. The strongest inhibitors of *E. amylovora* growth *in vitro* were then tested under laboratory and field conditions to evaluate survival on apple branches and blossoms and for suppression of fire blight on apple blossoms. The most promising BACs were then identified to species using genome-based methods, including the LINbase Web service (Tian et al., 2019), and the genetic basis of the biocontrol activity for one of the bacteria was explored using a combination of UV-mutagenesis and genome sequencing.

## **2. Materials and Methods:**

### *2.1 In vitro screening for inhibition of E. amylovora growth*

A dual culture assay was used to detect antimicrobial activity in 254 rain-borne bacteria (listed in **Supplementary Table 1**) against two isolates of *Erwinia amylovora* from Northern Virginia (BAV5616 and BAV5617). The assay was performed in Petri dishes containing yeast extract dextrose agar NYDA (Nutrient agar 23g l<sup>-1</sup>, dextrose 10g l<sup>-1</sup>, yeast extract 5g l<sup>-1</sup>). 100µl of *E. amylovora* suspensions were spread on each plate and up to five 10µl-droplets of suspension of rain-isolated bacteria were placed at equal distance from each other. In an initial screening, suspensions were made by adding a loop-full of rain-borne bacteria to 1ml of sterile 50mM potassium phosphate buffer (pH 7.0). In follow-up experiments, concentrations of putative BCAs were adjusted to an optical density of 0.1 at 600nm (OD<sub>600</sub>). Suspensions of increasing concentration of *E. amylovora* were spread on plates (OD<sub>600</sub> of 0.001, 0.01, 0.1, and 1, with the latter corresponding to a concentration of 3x10<sup>9</sup> CFU ml<sup>-1</sup>). Plates were incubated at 28°C for 24 hours. Antagonistic activity was quantified by the diameter of the inhibition zone forming around the rain-isolated bacteria. Rain-isolated bacteria that showed the most pronounced inhibition were also screened for activity against other economically important bacterial

(10 isolates), fungal (5 isolates) and oomycete (1 isolate) plant pathogens (**Table 2; Supplementary Table 2**) using similar protocols.

The two rain-isolated bacteria, BAV2934 and BAV3296, that showed the strongest inhibition of *E. amylovora in vitro* were chosen for further characterization. To facilitate re-isolation from plants, rifampicin-resistant mutants were selected for these strains and for *E. amylovora*. This was done by growing strains on NYDA solid medium containing 20 mg ml<sup>-1</sup> of antibiotic. Colonies that grew at this concentration were then serially transferred to media with increasing antibiotic concentration (50 mg ml<sup>-1</sup>, 100 mg ml<sup>-1</sup>, 150 mg ml<sup>-1</sup>, and 200 mg ml<sup>-1</sup>).

### 2.2 Survival assay under environmental conditions

To compare the survival of BAV2934, BAV3296, and *E. amylovora* rifampicin-resistant mutant strains (BAV5616), branches of potted 5-year old apple trees cultivar ‘Golden delicious’ were spray-inoculated to run-off at an OD<sub>600</sub> of 0.1 and placed on the roof of a three-story research building. Apple trees were exposed to the outdoor environment for the duration of the experiment. The non-pathogenic *Escherichia coli* strain DH5-Alpha was included for comparison. Bacterial population sizes were determined on the day of inoculation, and 4 days, 8 days, and 12 days after inoculation by dilution-plating on NYDA containing 200 mg ml<sup>-1</sup> rifampicin.

### 2.3 Assay of inhibitory activity against *E. amylovora* on detached apple blossoms

Blossoms of apple cultivar ‘Golden Delicious’ were detached at the first bloom stage (central blossom just opened) and maintained with the cut pedicel submerged in 10% sucrose solution. Potential BCAs were cultured on NYDA plates for 24h at 28°C, and cell suspensions were prepared in 10mM MgSO<sub>4</sub> containing 0.01% of Silwet. BlightBan™ A560 (active ingredient: *P. fluorescens* strain A560; NuFarm Americas, Burr Ridge, IL, USA) and FireWall™ 17WP (active ingredient: streptomycin sulfate; AgroSource, NJ, USA) were diluted to 0.8 g l<sup>-1</sup> water and 2 g l<sup>-1</sup> water, respectively, following manufacturer’s instructions. A rifampicin-resistant mutant strain of *E. amylovora* was cultured on

NYDA plates containing 200 mg ml<sup>-1</sup> rifampicin for 24 hours at 28°C, and a cell suspension was prepared in 10mM MgSO<sub>4</sub> and 0.01% of Silwet.

Six individual detached blossoms were pre-treated with one of the following treatments: bacterial suspension (BAV2934, BAV3296 or *E. coli* DH5-Alpha) at a concentration of 10<sup>9</sup> CFU ml<sup>-1</sup>; 10mM MgSO<sub>4</sub> (mock treatment negative control), or a commercial product (BlightBan™ A560 or FireWall™ 17WP). Detached blossoms were pre-treated by dipping the blossoms into the bacterial suspension or commercial product for 10 seconds. Blossoms were then allowed to air-dry for 2 minutes, enclosed in a planting tray, and incubated at room temperature for 48 hours. The experiment was repeated a total of four times.

Pre-treated blossoms were then inoculated with *E. amylovora* at 10<sup>7</sup> CFU ml<sup>-1</sup>, enclosed in a planting tray, and incubated at room temperature for 72 hours. Pathogen population sizes were determined in the receptacle and pedicel separately by grinding plant tissue followed by dilution-plating on NYDA supplemented with 200 mg ml<sup>-1</sup> of rifampicin. Disease severity was calculated as fraction of the length of the necrotic pedicel divided by the total pedicel length.

#### 2.4 Disease suppression assays on attached blossoms in an orchard

Field experiments were conducted in April 2018 and April 2019 in an experimental orchard at Virginia Tech's Kentland Farm (McCoy, VA, USA). Dates of treatments and weather conditions are reported in **Supplementary Table 4**. Clusters of apple blossoms of the varieties 'Golden Delicious', 'York Imperial', 'Rome Beauty' and 'Empire' at the first bloom stage were spray-inoculated to the point of run off. Inoculations were performed during morning hours to avoid warm temperatures that may affect bacterial attachment and survival. Between 20 and 30 blossom clusters were randomly pre-treated using BAV2934 (10<sup>9</sup> CFU ml<sup>-1</sup>), BAV3296 (10<sup>9</sup> CFU ml<sup>-1</sup>), FireWall™ 17WP (2 g l<sup>-1</sup>), and BlightBan™ A506 (0.8 g diluted in 1 l) as described above. 10mM MgSO<sub>4</sub> was used as mock treatment negative control. In 2018, three 'Golden Delicious' and three 'York' trees located in different rows in an orchard were inoculated as replicates, with each tree receiving all treatments. In 2019, three 'Golden

Delicious' and two 'Empire' and two 'Rome' apple trees located in different rows in an orchard were inoculated as replicates, with each tree receiving all treatments. Pre-treated blossoms were allowed to air-dry before inoculation with *E. amylovora* ( $10^7$  CFU ml<sup>-1</sup>). Disease incidence for each treatment was evaluated as percentage of symptomatic clusters out of all clusters. All data were analyzed using JMP statistical software (Statistical Analysis Systems Institute, Cary, NC, USA). We performed a one-way ANOVA analysis of variance, using a significance level of 0.05, to test the null hypothesis that treatments do not differ in controlling fire blight. Tukey-Kramer HSD was used for all pairwise comparisons in our detached and attached blossoms assays. Graphs were plotted in RStudio version 1.1.456 (RStudio, 2015).

### 2.5 UV mutagenesis screen to identify the molecular basis of growth inhibition

To identify genes necessary for the suppression of *E. amylovora*, mutants of BAV2934 were generated by UV radiation following the protocol by Barrick (2014) with small modifications. Briefly, an overnight bacterial culture was resuspended in water at an OD<sub>600</sub> of 0.1. 10µl droplets were subjected to 10,000µJ of UV radiation utilizing a UV Stratalinker (Stratagene UV Stratalinker 1800). UV treated cells were plated onto TSA medium and incubated for 24 hours at 28°C. All single colonies were tested for inhibitory activity following the plate assay as described above. Any putative bacterial colony that lost inhibitory activity against *E. amylovora* was confirmed in subsequent tests.

### 2.6 Genome sequencing

Genomic DNA was extracted using the Genomic DNA purification kit (PUREGENE - Genra Systems, USA) according to the manufacturer's protocol. The DNA concentration and quality were evaluated by UV spectrophotometry (NanoDrop 1000, Thermo, USA) and visualized on a 1% agarose gel. DNA library preparation and Paired-end (PE150) sequencing was performed on the Illumina HiSeq 2500 at the Biocomplexity Institute at Virginia Tech (Blacksburg, VA, USA).

Isolate BAV2934 was also sequenced using Oxford Nanopore Technologies' (Oxford, UK) MinION. A total of 3µg of purified bacterial DNA was treated with RNase A 100mg ml<sup>-1</sup> (1µl for 100

µl of gDNA) and the Short Read Eliminator Kit (Circulomics, MD, USA), according to the manufacturer's protocol. The purified bacterial DNA was used to prepare a sequencing library with the 1D genomic DNA ligation kit SQK-LSK109 (Oxford Nanopore Technologies, UK). Sequencing was performed with a FLO-MIN106 (R 9.4.1) flow cell for 48 hours.

### 2.7 Genome analysis

FastQC (Andrews, 2010) was used to assess sequencing reads for quality control. Reads were trimmed using Trimmomatic (Bolger et al., 2014) to remove adapter sequences and low quality reads (Q<30). *De-novo* assemblies were generated using SPAdes (Bankevich et al., 2012).

The fast5 files generated by MinION sequencing of the BAV2934 genome were base-called using guppy (version 3.3.2). A hybrid assembly using both Illumina and ONT reads was performed using Unicycler v0.4.7 (Wick et al., 2017) with default parameters. Bandage (Wick et al., 2015) was used to visualize the bacterial assembly graph. Prokka (Seemann, 2014) was used to annotate the assembled genome sequence.

Genomic DNA of five BAV2934 UV-induced mutants was pooled and sequenced using Illumina as described above. Mutations were identified by mapping reads against the BAV2934 hybrid assembly. In brief, a reference genome index was pre-built with HiSat2 (Kim et al., 2015). Illumina paired-end reads were aligned to the reference genome using SAMtools (Li et al., 2009) and VarScan (Koboldt et al., 2009) was used for calling variants selecting parameters for pooled samples (--min-coverage8 --min-var-freq 0.15 -p-value 0.05). Read coverage of each gene was determined using bbmap (Bushnell, 2014).

Assembled genome sequences in fasta format were used for strain identification at the LINbase website at linbase.org (Tian et al., 2019) using the "Identify using a genome sequence" function.

The antibiotic & Secondary Metabolite Analysis Shell [antiSMASH (Medema et al., 2011)], which is capable of identifying loci that cover the whole range of known secondary metabolite compound classes, was used to identify gene clusters in the wildtype strain BAV2934.

## 2.8 Analysis of blossom microbiomes

‘Golden Delicious’ apple blossoms were inoculated in an orchard as described above using BAV2934 ( $10^7$  CFU ml<sup>-1</sup>). Non-inoculated blossoms served as negative control. Fifteen blossoms per treatment were collected into a clean Ziploc plastic bag on day 0 (day of inoculation), five days after inoculation, and 10 days after inoculation. Blossoms were immediately transported to the laboratory and processed. Culturable bacterial population sizes were determined on TSA by dilution-plating.

For the culture-independent microbiome analysis, 300 ml of sterile distilled water was added to each Ziploc bag containing the blossoms. Samples were sonicated for 10 minutes using a 1520 BRANSON sonicator (Branson Ultrasonics, Danbury, CT, USA). The liquid was then vacuum-filtrated through a 0.2 µm pore-size filter membrane (Supor® 200 PES membrane Disc Filter, PALL Corporation, Port Washington, NY, USA). Genomic DNA extraction from filters was performed using the DNeasy PowerWater kit (Qiagen, Germantown, MD, USA) according to the manufacturer’s protocol. DNA concentration and quality were determined by NanoDrop 1000 (Thermo, Waltham, MA, USA) and visualization on a 1% agarose gel. Primers 799F (anti-chloroplast, 5’AACMGGATTAGATACCCCKG3’) and 1115R (“universal,” 5’AGGGTTGCGCTCGTTG3’) were used to amplify and sequence the V4 hypervariable region of the 16S rRNA gene. All steps from PCR to paired-end (2 × 300 bp) amplicon sequencing on the Illumina MiSeq platform were performed at Molecular Research LP (MR DNA™, Shallowater, TX, USA). QIIME version 1.9.1 was used for amplicon data analysis (Caporaso et al., 2010). Briefly, quality filtered paired-end reads (phred quality scores above 30) were joined together in a single read. Operational taxonomic units (OTUs) were picked using the open-reference pipeline at 97% sequence similarity using UCLUST as the clustering tool and the SILVA database (Quast et al., 2013). All OTUs unassigned or assigned to mitochondria, chloroplast and Cyanobacteria were removed. Data was visualized in RStudio version 1.1.456 using the Phyloseq 1.19.1 (McMurdie and Holmes, 2013) and ggplot2 2.2.1 packages.

### 3. Results:

#### 3.1 Inhibitory activity of rain-isolated bacteria in vitro

Nine out of 254 tested rain-isolated bacteria (**Supplementary Table 1**) displayed an inhibitory effect against *E. amylovora* during initial *in vitro* dual culture screening (**Figure 1A**). These bacteria had been previously identified as *Pseudomonas sp.*, *Bacillus sp.*, *Pseudomonas sp.*, *Pantoea agglomerans*, *Pantoea sp.*, *Pseudomonas sp.*, *Pseudomonas sp.*, *Pantoea ananatis.*, and *Pseudomonas sp.* by 16S rRNA sequencing (Failor et al., 2017) (**Table 1**). A more stringent assay using higher concentrations of *E. amylovora* allowed us to identify two isolates (BAV2934 and BAV3296) from the genus *Pantoea* with the highest inhibitory effect, as measured by the diameter of the inhibition zone that they induced (**Figure 1B**). BAV2934 and BAV3296 were chosen for further characterization.

We also evaluated the inhibitory effect of the nine bacteria that had shown activity against *E. amylovora* for activity against additional bacterial, fungal, and oomycete plant pathogens. BAV2934 and BAV3296 showed strong activity against the majority of the tested bacterial plant pathogens (including species of *Xanthomonas*, *Pseudomonas*, and *Ralstonia*), but not against fungal and oomycete pathogens, while the *Pseudomonas* strains BAV3226 and BAV3280 were most effective against fungal and oomycete pathogens (**Table 2; Supplementary Table 2**).

#### 3.2 Bacterial survival on apple trees

Survival, measured as bacterial population size over time, of *Pantoea* strains BAV2934 and BAV3296 and rifampicin-resistant *E. amylovora* BAV5616 was determined on detached apple branches placed on the top of our 3-story research building in three independent experiments performed on the following dates: from November 18<sup>th</sup> to 30<sup>th</sup>, 2016, from February 13<sup>th</sup> to 25<sup>th</sup>, 2017, and from March 24<sup>th</sup> to April 5<sup>th</sup>, 2017. Minimum and maximum temperatures during the three experiments ranged from -6 to 18°C, -7 to 18°C, 0 to 19°C, respectively. Apple branches in November 2016 and March 2017 were exposed to rain and light snow (details on weather conditions during the experiments are reported in **Supplementary Table 3**). Results for the experiment performed in February 2017 are shown in

**Figure 2** and results for experiments in November 2016 and March 2017 are shown in **Supplementary Figure 1**. In all three experiments, survival of BAV2934 and BAV3296 was comparable to *E. amylovora* after 12 days under environmental conditions (p-value 0.0001 ANOVA, Tukey HDS). The main difference between experiments was that in March 2017 the bacterial population size remained the same after 4 days post-inoculation, in contrast to the other two experiments in which the bacterial population size decreased 2-fold by day 4 post-inoculation.

Population size of BAV2934 was also evaluated on apple blossoms at Kentland Farm in spring 2018 at 0, 5, and 10 days post-inoculation (April 25, April 30, and May 5 of 2018). Minimum and maximum temperatures during this experiment ranged from 0 to 27°C (**Supplementary Table 4**). In this case, the total bacterial population size was analyzed in combination with the relative abundance of BAV2934, as determined by culture-independent microbiome analysis. The total culturable bacterial population size on blossoms inoculated with BAV2934 was three-fold higher at 0, 5 and, 10 days post-inoculation, compared to the non-inoculated controls (**Figure 3A**). The relative abundance of BAV2934 was 95% and 85% at 5 days and 10 days post-inoculation, respectively (**Figure 3B**).

### 3.3 Fire blight control on detached apple blossoms

Fire blight disease severity measured as a fraction of the length of necrosis along the pedicel divided the total pedicel length was significantly reduced by pre-treatments of blossoms with either BAV2934 (approximately 20% of necrotic pedicel) or BAV3296 (approximately 10% of necrotic pedicel), as compared to *E. coli* (60% of necrotic pedicel) and the MgSO<sub>4</sub> mock treatment negative control (70% of necrotic pedicel) (p-value 0.0001, ANOVA, Tukey HDS). BlightBan™ A506 (approximately 10% of necrotic pedicel) and FireWall™ (approximately 5% of necrotic pedicel) treatments were similarly effective, and both were lower than the mock treatment negative control (**Figure 4A**).

*E. amylovora* population size in the receptacle and the pedicel was also significantly reduced by pre-treatment with either BAV2934 or BAV3296, compared to *E. coli* and MgSO<sub>4</sub> (p-value 0.0001 ANOVA, Tukey HDS). In the receptacle, for example, the *E. amylovora* population was reduced by



over 2-fold on blossoms treated with BAV2934 and BAV3296, compared to blossoms treated with *E. coli* or MgSO<sub>4</sub> (p-value 0.0001 ANOVA, Tukey HDS). Similar pathogen population sizes were found in the receptacles of blossoms treated with BlightBan™. However, BlightBan™ was not as effective as FireWall™, which suppressed pathogen growth almost completely (**Figure 4B**). In the pedicel, the effect of BAV2934 or BAV3296 was similar to BlightBan™, in terms of *E. amylovora* population reduction. FireWall™ again provided the most protection (**Figure 4C**).

### 3.4 Fire blight control in an apple orchard

Disease incidence of *E. amylovora*-inoculated apple blossoms pre-treated with BAV2934 and BAV3296 was compared with disease incidence following pre-treatments with commercial products BlightBan™ A506 and FireWall™ or a mock treatment negative control (10mM MgSO<sub>4</sub>) in spring 2018 and spring 2019. Disease incidence was calculated as percentage of blossom clusters with fire blight symptoms out of the total number of clusters used in each treatment.

In 2018 on ‘Golden Delicious’, FireWall™ was the most effective treatment with only 11% of treated clusters developing disease, while the mock treatment showed 89% disease incidence (p-value 0.0028 ANOVA, Tukey HDS). No significant differences were observed between BAV2934 with 48%, BlightBan™ with 63%, and BAV3296 with 70% disease incidence, compared to the mock treatment (**Figure 5A** and **Supplementary Table 5**).

On ‘York’, the mock treatment showed 76.5% disease incidence. FireWall™ and BAV2934 had 0% and 11.5% disease incidence, respectively, both significantly lower compared to the mock treatment and compared to BlightBan™ with 55.5% disease incidence (p-value 0.0016 ANOVA, Tukey HDS). Also BAV3296 with 32.5% disease incidence showed significant control compared to the mock treatment, but not compared to BlightBan™. No significant difference was observed between BlightBan™ and the mock treatment (**Figure 5B** and **Supplementary Table 5**).

In 2019, apple blossoms of the varieties ‘Golden Delicious’, ‘Empire’ and ‘Rome’ were pre-treated as in 2018, but one of the BAV2934 UV-generated mutants that had shown no inhibitory

activity *in vitro* was also included. On ‘Golden Delicious’, although no significant differences were observed among pre-treatments (p-value 0.3152 ANOVA, Tukey HDS), FireWall™ was the best inhibitor of fire blight with 30% disease incidence. It was followed by BAV2934 with 55%, the BAV2934 mutant with 65%, BAV3296 with 70 %, and BlightBan™ with 80% disease incidence. The mock treatment gave 80% disease incidence (**Figure 5C** and **Supplementary Table 5**).

On ‘Rome’, blossoms pre-treated with BAV3296 had 15% disease incidence, while the FireWall™ and BAV2934 UV-mutant pre-treatment each had 20% disease incidence. These pre-treatments were significantly different compared to the pre-treatment with BlightBan™, which showed 65% disease incidence, and with the mock treatment, which showed 70% disease incidence (p-value 0.0083 ANOVA, Tukey HDS). The BAV2934 treatment with 35% disease incidence was not significantly lower than the mock treatment (**Figure 5D** and **Supplementary Table 5**).

On ‘Empire’, the mock treatment showed only 50% disease incidence. FireWall™ showed 37% disease incidence, followed by the BAV2934 mutant with 43%, BAV2934 with 47%, BAV3296 with 53% disease incidence. BlightBan™ had 80% disease incidence, which was higher than the mock treatment negative control (**Figure 5E** and **Supplementary Table 5**). None of the treatments were significantly different compared to mock (p-value 0.4420 ANOVA, Tukey HDS).

In summary, BAV2934 and BAV3296 reduced disease incidence compared to a mock treatment in four out of five field experiments and this reduction was significant in two of the experiments for BAV3296 and in one experiment for BAV2934. For most field experiments, the reduction in disease incidence by BAV2934 and BAV3296 was similar to BlightBan™ but not as good as FireWall™. The BAV2934 UV-mutant showed inconsistent control and was not significantly different from the BAV2934 wild-type strain in any of the trials (**Supplementary Table 5**).

### 3.5. Whole genome sequencing and genome-based identification

The genomes of BAV2934 and BAV3296 were sequenced using Illumina HiSeq technology, assembled, and identified using the “Identify using a genome sequence” function at the LINbase web service (Tian et al., 2019) (**Table**

1). BAV2934 was identified as *Pantoea agglomerans* and BAV3296 as *P. ananatis*. For BAV2934, we also generated 27 Gb of long reads using the Oxford Nanopore Technologies MinION™ sequencer and carried out a hybrid genome assembly of both the short Illumina and long MinION™ reads, which allowed us to obtain a closed genome of 4,003,977 bp and four circular contigs of 528,933 bp, 205,248 bp, 203,868 bp, and 2,968 bp, respectively. These four contigs probably represent plasmids.

### *3.6 Prediction of biosynthetic gene clusters (BGCs) and a UV-mutagenesis screen to identify genes at the basis of the observed inhibitory activity*

To identify putative genes at the basis of the inhibitory activity of BAV2934 against *E. amylovora*, the genome was annotated using Prokka (Seemann, 2014) and biosynthetic gene clusters (BGCs) were predicted using antiSMASH (Medema et al., 2011) (Table 4). Eight BGCs were identified that range from 11,391 to 59,810 bp and were predicted to produce a series of different products (Table 4). To identify if any BGC was involved in the inhibitory activity, a UV-mutant screen was performed in parallel. After UV treatment, 1099 colonies of BAV2934 were screened for loss of inhibition *in vitro* and five mutants were identified. DNA of the five mutants was pooled and sequenced on the Illumina platform. 21,742,895 reads of a total length of 7.3 Gb were obtained and mapped against the annotated BAV2934 genome to identify non-synonymous mutations that could explain the loss of the inhibitory activity. **Table 3** lists the 17 non-synonymous mutations that were found. Gene NOOGOKNH\_04505, annotated as a dimodular nonribosomal peptide synthase, was one of them. Using BLAST, it was identified as part of a phenazine antibiotic D-alanylgriseoluteic acid synthesis gene cluster in *Erwinia herbicola* Eh1087 (Giddens et al., 2002). Among the BGCs predicted in BAV2934 by antiSMASH (Table 4), a phenazine gene cluster was identified in contig 4, and although not in this BGC, the mutation was 1,843 bp downstream of the phenazine cluster (**Supplementary Figure 2**) and may contribute to production of the compound. This mutation was confirmed by PCR and Sanger sequencing in one of the five mutants; the gene could not be amplified from the other four mutants. Since, based on genome sequencing, the NOOGOKNH\_04505 gene is located on the 203,868 bp-long

contig in the BAV2934 genome assembly, the lack of amplification may have been due to plasmid loss. In fact, genome coverage data of the reads derived from the sequencing of the five mutant-pool showed that the entire contig was present in the pool only at 1/4th of the average coverage of the main chromosome and PCR with gene-specific primers of chromosomal genes and putative plasmid-encoded genes confirmed the absence of the 203,868 bp-long plasmid in these four mutants (Supplementary Figure 3).

#### 4. Discussion:

Efficacy of BCAs in the field, in particular when applied to aerial plant surfaces, is dependent on weather conditions during and after application because BCAs need to colonize and survive on plants surfaces in order to efficiently inhibit pathogens (Johnson et al., 2000). In this study, we hypothesized that precipitation may be a promising source of BCAs for aerial plant surfaces, in particular against the fire blight pathogen *E. amylovora* on apple blossoms, since bacteria that are ubiquitous in the atmosphere and precipitation (Polymenakou, 2012) can be expected to be adapted to some of the same environmental stresses to which BCAs are exposed, such as dramatic changes in temperature, humidity, and UV radiation (Lindow, 1991). Additionally, as part of a separate study (Failor et al., 2017), we had isolated bacteria from rain that were members of the genera *Pseudomonas*, *Pantoea*, and *Bacillus*, which include common plant-associated species, some of which are already used in commercial BCAs (Matyjaszczyk, 2015).

As a first step, we screened 254 rain-isolated bacteria for inhibition of *E. amylovora in vitro*. We excluded bacteria with confirmed ice nucleation activity from this screen because they could worsen frost damage when used during weather conditions favorable for frost formation (Lindow, 1983). Nine out of the 254 tested bacteria (3.5%) showed a measurable inhibitory effect against *E. amylovora*. Two isolates, identified as members of the genus *Pantoea*, maintained their efficacy against *E. amylovora* even when pathogen concentrations were increased, while the other isolates failed under

these conditions. Our use of precipitation as a source of potential BCAs is unconventional, as traditionally plant surfaces and soil are the most commonly utilized sources for BCAs isolation. Gerami et al. (2013) found that almost 50% of 120 tested epiphytic bacterial strains, isolated from blossoms, leaves and shoots of pome-fruit and stone-fruit trees, inhibited *E. amylovora in vitro* but only 4 isolates worked efficiently in both *in vitro* and *in planta* assays. Sharifazizi et al. (2017) found that 45% of 22 pear leaf-associated bacterial isolates inhibited *E. amylovora in vitro*. It is challenging to compare our success rate with rain-isolated bacteria with the success rate in these screens using epiphytic bacteria since many factors contribute to the efficiency of bacteria to inhibit pathogens *in vitro*, including: the identity of the pathogen strain, the concentration at which potential BCAs and pathogens are plated, and the type of growth medium used (Dickie and Bell, 1995; Borowicz and Saad Omer, 2000). Therefore, in the absence of a direct comparison with other isolation sources, we cannot conclude if rain harbors a higher or lower proportion of bacteria that effectively inhibit pathogens *in vitro* compared to bacteria isolated from other sources. Nonetheless, since we succeeded in identifying two *Pantoea* isolates that strongly inhibited *E. amylovora in vitro*, further tests to determine their survival rate and their efficiency *in vivo* were warranted.

Since the main motivation behind our study was the expectation that rain-isolated bacteria could persist on plant surfaces despite environmental stresses, we tested survival on both apple branches and blossoms. On apple branches, the rain-isolated strains BAV2934 and BAV3296 showed a very similar survival rate to *E. amylovora* declining approximately 100-fold after 12 days post-inoculation in three separate experiments in November, February, and March, while *E. coli* was already undetectable 4 days post-inoculation. The decline was faster in November and February probably due to colder temperatures, compared to those recorded in March (**Supplementary Table 3**). Although we did not extend survival assays beyond 12 days, the fact that survival of BAV2934 and BAV3296 was similar to survival of *E. amylovora* on apple branches gives confidence that rain-borne bacteria are able to persist on plant surfaces as we had hypothesized. This result also suggests that winter and spring

applications of these potential BCAs should be tested, to determine if such applications could reduce fire blight incidence in the following spring and summer.

When testing survival on apple blossoms, the total bacterial population size on blossoms inoculated with BAV2934 was significantly larger than the population size on non-inoculated blossoms 10 days post-inoculation (**Figure 3A**). Further, the relative abundance of BAV2934 remained stable, representing over 85% of the total population at 10 days post-inoculation (**Figure 3B**). This evidence also suggests that BAV2934 is able to robustly colonize and persist on healthy apple blossoms under field conditions. In contrast, Wei et al. (2016) found that the relative abundance of a potential BCA (*Bacillus subtilis*) in strawberry leaves was found to be depleted by 50% 8 days post-inoculation under field conditions. Another study found the population size of a potential *Lactobacillus plantarum* BCA decreased significantly 10 days after inoculation on kiwifruit, strawberry, and *Prunus* leaves, even under stable greenhouse conditions (Daranas et al., 2019). When compared to those BCAs, which were isolated from plants or soil, our rain-borne isolate BAV2934 demonstrated better survival on aerial plant surfaces. However, since differences in bacterial survival rate depend on the type of plant surface (Pujol et al., 2006; Bonaterra et al., 2007) and environmental conditions (Nucló et al., 1998), we cannot make general conclusions before testing our rain-isolated bacteria in different geographic locations.

Some BCAs have a broad spectrum of inhibition against many plant pathogens (Ishimaru et al., 1988; Mora et al., 2015; Daranas et al., 2019). Here, we observed that our initial 9 rain-isolated bacteria have the ability to suppress a wide range of bacterial, fungal, and oomycetes pathogens *in vitro* (**Supplementary Table 2**). However, the efficacy of the rain-isolated *Pantoea* and *Bacillus* strains was limited to bacterial pathogens. In contrast, the *Pseudomonas* isolates inhibited bacterial, fungal, and oomycete plant pathogens. This suggests either the production of several antimicrobial compounds (Bender et al., 1999) in these strains or the production of an antibiotic with broad spectrum activity (Ishimaru et al., 1988). The broad spectrum activity against fungi and bacteria observed in the rain-isolated *Pseudomonas* strain agrees with previous reports that *Pseudomonas* isolates produce

siderophores with antimicrobial properties in disease-suppressive soils (Kloepper et al., 1980). For example, *P. fluorescens* strain UP61 produces 3 different antibiotics, including pyrrolnitrin, pyoluteorin and 2,4-diacetylphloroglucinol, involved in the inhibition of fungal, oomycete and bacterial plant pathogen strains (La Fuente et al., 2004). Also, several *Pseudomonas syringae* strains produce syringomicins and syringopeptins that inhibit a broad spectrum of fungi and bacteria, respectively (Bensaci et al., 2011).

The effect of timing of BCA application compared to pathogen inoculation is variable. While (Wilson et al., 1992) reported a significant control of *E. amylovora* on hawthorn blossoms when biocontrol and pathogen were co-inoculated, (Wilson et al., 1992) found that *Pseudomonas fluorescens* strain A506 effectively protected pear blossoms when blossoms were inoculated with the BCA in advance, but not when blossoms were co-inoculated with the BCA and *E. amylovora*. In detached blossom assays, we decided to apply rain-isolated bacteria 2 days before we applied the pathogen to allow sufficient time for blossom colonization by our putative BCAs. In contrast, in the field experiments, BCA treatments were done only two 2 hours before *E. amylovora* inoculation because we wanted to inoculate both the BCA and the pathogen when the blossoms were most susceptible. Also, since *E. amylovora* can persist in symptomless infected plant tissue (Crepel and Maes, 2000; Weißhaupt et al., 2016), the pathogen can spread and colonize blossoms as soon as they open. Therefore, testing BCAs by applying them only 2 hours before inoculation with the pathogen may represent a more realistic scenario for the use of BCAs in agricultural practice.

In some studies, no relationship was found between *in vitro* antibiosis and *in vivo* bacterial performance (Özaktan et al., 1999; Gerami et al., 2013). Our rain-isolated bacteria were initially tested for antibiosis in a dual culture assay against *E. amylovora* (**Table 2**) with the goal of finding potential new secondary metabolites in further experiments. Then, the best inhibitors of *E. amylovora in vitro* were tested for their ability to suppress fire blight *in planta*. Published assays for testing the efficiency of BCAs in controlling fire blight on detached blossoms only rate disease incidence on whole blossoms

(Bonaterra et al., 2007; Roselló et al., 2013). We decided to measure the pathogen population size separately in the receptacle and the pedicel and to measure the necrotic portion of the blossom pedicel. Our reasoning for evaluating the pedicel separately from the receptacle was that the ability of a BCA to reduce symptoms and pathogen population size should be a measure of its ability to interfere with pathogen colonization and migration into the rest of the apple tree. Using this experimental setup, we found that strains with the best *in vitro* performance had *in vivo* performance similar to the commercial BCA BlightBan™ A506, but were not as efficient as the streptomycin product FireWall™.

We observed inconsistent results when the rain-isolated bacteria were inoculated on attached blossoms in an apple orchard (**Figure 5**). BAV2934 significantly reduced disease incidence compared to the mock treatment in only one out of five trials, while BAV3296 significantly reduced disease incidence in only two out of five trials. In addition, BAV2934 or BAV3296 only performed significantly better than the commercial BCA BlightBan™ in one trial each. The streptomycin product FireWall™ generally performed the best in field trials, significantly reducing disease incidence in three out of five trials. We expected the BAV2934 mutant, which had lost activity in the *in vitro* assays, to perform poorer than the wild-type BAV2934 on attached blossoms, but we could not observe any significant differences between the two. Some of the non-significant differences we observed between treatments may have simply been due to small sample sizes (20 to 30 clusters per treatment per trial), but such variability is generally in line with previous studies using other BCAs. For example, (Sundin et al., 2009) found that BCA applications were inefficient and highly variable in the control of fire blight in the field in Michigan, New York, and Virginia. However, it is possible that better and more consistent results could have been obtained if BAV2934, BAV3296, and BlightBan™ had been applied more in advance of pathogen inoculation, allowing BCA organisms more time to colonize the blossoms. BCA variability may also be related to inoculum preparation. (Stockwell et al., 1998) showed that lyophilized cells of *P. fluorescens* A506 and *P. agglomerans* C9-1R established better on apple blossoms than inoculum prepared from fresh bacterial cells harvested directly from solid agar



medium. Özaktan (1999) tested several formulations of *P. agglomerans* where talc-based formulations showed better control of fire blight on pear blossoms than lyophilized and whey-based treatments. Further adaptive strategies, including osmoadaptation and acidic conditions, have also been tested as means to increase bacterial survival and fire blight control in both controlled laboratory and field conditions (Bonaterra et al., 2007; Daranas et al., 2018). In our study, bacterial cells were harvested from agar medium and used immediately to inoculate apple blossoms in our laboratory and field trials. Therefore, it may be possible to improve the control efficiency of the rain-isolated bacteria by optimizing their formulation. Additional trials in other geographic locations using alternative formulations will be necessary to determine if our rain-isolated BCAs can provide more efficient fire blight control under field conditions than currently available BCA products.

It is possible that temperatures following application and during incubation affected the control demonstrated in the orchard (**Figure 5 and Supplemental Figure 2**). In 2019, control was generally least effective on the cultivar ‘Empire’ and this was true for all biocontrol organisms and by FireWall™. Post-inoculation temperatures were also coolest on this cultivar: 10.9°C during the first eight days after application and 12.6°C during the 16-day period until data collection. Control was generally better on cultivars ‘Golden Delicious’ and ‘Rome’ in 2019, and post-inoculation temperatures for these cultivars were warmer with 16.6°C during the first eight days after application and 17.1°C during the 12-day period until data collection. Despite inconsistent results across cultivars, relative percent control appears to be greater on ‘Golden Delicious’ and ‘Rome’ because fire blight infection on the MgSO<sub>4</sub> control was greater under these warmer conditions. However, varietal differences may also be a factor, as control was greater on ‘Rome’ than on ‘Golden Delicious’ although both cultivars were treated and inoculated on the same day in 2019.

*Pantoea* species have been shown to produce different antibiotic compounds (Wright et al., 2001). In this study we identified a BGC that is predicted to produce a phenazine by *P. agglomerans* isolate BAV2934, and this molecule has been shown to be required for inhibition of *E. amylovora* in

*vitro* by (Giddens et al., 2002). The locus identified by UV mutagenesis is likely to contribute to the phenazine production, since the gene is in close proximity to the predicted BGC, in the same polarity, and because Dimodular NRPS are known to contribute to assist in the production of antibiotics (Felnagle et al., 2008). Phenazines are also produced by many other bacteria including: *Pseudomonas spp.* (Thomashow and Weller, 1988), *Streptomyces spp.* (Karnetová et al., 1983), and *P. agglomerans* (Giddens et al., 2002). However, in our rain-isolated bacteria, which included fluorescent *Pseudomonas* and *Pantoea* isolates, we only identified a phenazine cluster gene in BAV2934.

Due to the extensive genetic diversity found within many bacterial species, some species can include both beneficial and pathogenic bacteria. For example, while some *P. agglomerans* strains are commercialized as BCAs, others are known plant pathogens or recognized opportunistic human pathogens, isolated from wounds (Cruz et al., 2007; Lee et al., 2010; Smits et al., 2015). *P. ananatis* also includes both plant pathogens (De Baere et al., 2004; Coutinho and Venter, 2009) and BCAs (Wu et al., 2016). Unfortunately, it is not well known for either *P. agglomerans* or *P. ananatis* if the same strains can be both beneficial and pathogenic, or if some strains are pathogenic, while others are beneficial. The taxonomy within the *Pantoea* genus is also rapidly changing, which leads to another complication in that some *Pantoea* isolates reported as members of a certain *Pantoea* species have been incorrectly identified (Rezzonico et al., 2009). The LINbase web service (Tian et al., 2019) represents a practical tool to improve precise genome-based identification of BCAs, since it allows classification and identification of bacteria at the genus-, species-, and intraspecies ranks. However, precise genome-based prediction of pathogenicity and other phenotypes for strains belonging to a certain species only becomes possible after careful phenotypic characterization of many reference strains within that species. Because this foundational work has not been done yet with either *P. agglomerans* or *P. ananatis*, LINbase was useful in identifying BAV2934 and BAV3296 as members of these species, but it was impossible to infer their safety as BCAs.

In summary, we have found that rain serves as a reservoir of robust bacteria that suppress growth of *E. amylovora* and other plant pathogens *in vitro*. Two of the bacteria we isolated were shown to control fire blight on detached blossoms, to survive similarly to *E. amylovora* on apple branches and blossoms, and, although inconsistently, to control fire blight in the field. Using a combination of genomics and UV-mutagenesis, we also determined the likely mechanism of antibiosis in one of the isolates. While it was straightforward to identify potential BCAs to the species-level using genomics, a database with more thorough phenotypic characterization of strains in regard to pathogenicity on plants and humans will be necessary to infer safety of BCAs based on their genome sequences alone. Therefore, additional field tests, and formulations and safety tests will need to be performed before making a conclusive determination of the potential of the identified rain-borne bacteria as commercial BCAs.

## **5. Conflict of Interest:**

LINbase uses the trademarks Life Identification Number<sup>®</sup> and LIN<sup>®</sup>, which are registered by This Genomic Life, Inc. BAV reports in accordance with Virginia Tech policies and procedures and his ethical obligation as researcher that he has a financial interest in This Genomic Life, Inc. Therefore, his financial interests may be affected by the research reported in this manuscript. He has disclosed those interests fully to Virginia Tech, and he has in place an approved plan for managing any potential conflicts arising from this relationship.

## **5. Author Contributions:**

MEMLL conducted most of the experiments with contributions from VBG and KH. LT contributed to the bioinformatics analysis. HW analyzed genomes for prediction of the biosynthetic gene clusters. BAV, KSY, and SCM developed the overall project. MEMLL and BAV wrote the manuscript with input from the other authors.

## 6. Funding:

This research was supported by the Virginia Agricultural Council (Project number 671) and, in part, by the National Science Foundation (DEB- 1241068 and IOS-1754721). Funding to BAV and KSY was also provided in part by the Virginia Agricultural Experiment Station and the Hatch Program of the National Institute of Food and Agriculture, US Department of Agriculture.

## 7. Acknowledgments:

The authors are grateful to Ricky Hughes for facilitating the access to the experimental orchard at Kentland Farm, Virginia Tech (Virginia, USA).

## 9. Data Availability Statement:

The genomes of the nine rain-isolated bacteria described in Table 1 for this study have been submitted to GenBank (<https://www.ncbi.nlm.nih.gov/genome>) under the submission number SUB5968955 and accession numbers will be added here when they become available.

## 10. References:

- Alamri, S.A.M., Hashem, M., Mostafa, Y.S., Nafady, N.A., and Abo-Elyousr, K.A.M. 2019. Biological control of root rot in lettuce caused by *Exserohilum rostratum* and *Fusarium oxysporum* via induction of the defense mechanism. *Biological Control* 128:76-84.
- Andrews, S. (2010). FastQC: a quality control tool for high throughput sequence data. Available online at: <http://www.bioinformatics.babraham.ac.uk/projects/fastqc/>
- Bankevich, A., Nurk, S., Antipov, D., Gurevich, A.A., Dvorkin, M., Kulikov, A.S., Lesin, V.M., Nikolenko, S.I., Pham, S., Prjibelski, A.D., Pyshkin, A.V., Sirotkin, A.V., Vyahhi, N., Tesler, G., Alekseyev, M.A., and Pevzner, P.A. 2012. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *Journal of Computational Biology* 19:455-477.

- Barrick, J.E. (2014). UV mutagenesis of bacteria. Available online at: (<http://barricklab.org/twiki/bin/view/Lab/ProtocolsUVLibrary>).
- Bender, C.L., Rangaswamy, V., and Loper, J. 1999. Polyketide production by plant-associated *Pseudomonads* Annual Review of Phytopathology 37:175-196.
- Bensaci, M.F., Gurnev, P.A., Bezrukov, S.M., and Takemoto, J.Y. 2011. Fungicidal Activities and Mechanisms of Action of *Pseudomonas syringae* pv. *syringae* Lipodepsipeptide Syringopeptins 22A and 25A. Frontiers in Microbiology 2:216-216.
- Bolger, A.M., Lohse, M., and Usadel, B. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics 30:2114-2120.
- Bonaterra, A., Cabrefiga, J., Camps, J., and Montesinos, E. 2007. Increasing survival and efficacy of a bacterial biocontrol agent of fire blight of rosaceous plants by means of osmoadaptation. FEMS Microbiology Ecology 61:185-195.
- Bonaterra, A., Badosa, E., Cabrefiga, J., Francés, J., and Montesinos, E. 2012. Prospects and limitations of microbial pesticides for control of bacterial and fungal pomefruit tree diseases. Trees (Berlin, Germany : West) 26:215-226.
- Borowicz, J.J., and Saad Omer, Z. 2000. Influence of rhizobacterial culture media on plant growth and oninhibition of fungal pathogens. BioControl 45:355-371.
- Bushnell, B. (2014). BBMap: A Fast, Accurate, Splice-Aware Aligner (United States).
- Büyükcam, A., Tuncer, Ö., Gür, D., Sancak, B., Ceyhan, M., Cengiz, A.B., and Kara, A. 2018. Clinical and microbiological characteristics of *Pantoea agglomerans* infection in children. Journal of Infection and Public Health 11:304-309.
- CABI. (2018). Invasice Species Compendium. Available online at: <https://www.cabi.org/ISC>
- Caporaso, J.G., Kuczynski, J., Stombaugh, J., Bittinger, K., Bushman, F.D., Costello, E.K., Fierer, N., Peña, A.G., Goodrich, J.K., Gordon, J.I., Huttley, G.A., Kelley, S.T., Knights, D., Koenig, J.E., Ley, R.E., Lozupone, C.A., McDonald, D., Muegge, B.D., Pirrung, M., Reeder, J., Sevinsky,

- J.R., Turnbaugh, P.J., Walters, W.A., Widmann, J., Yatsunenkov, T., Zaneveld, J., and Knight, R. 2010. QIIME allows analysis of high-throughput community sequencing data. *Nature Methods* 7:335.
- Coutinho, T.A., and Venter, S.N. 2009. *Pantoea ananatis*: an unconventional plant pathogen. *Molecular Plant Pathology* 10:325-335.
- Crepel, C., and Maes, M. 2000. Hibernation of the fire blight pathogen *Erwinia amylovora* in host plants. *Mededelingen - Faculteit Landbouwkundige en Toegepaste Biologische Wetenschappen, Universiteit Gent* 65:19-25.
- Cruz, A.T., Cazacu, A.C., and Allen, C.H. 2007. *Pantoea agglomerans*, a plant pathogen causing human disease. *Journal of Clinical Microbiology* 45:1989-1992.
- Daranas, N., Badosa, E., Francés, J., Montesinos, E., and Bonaterra, A. 2018. Enhancing water stress tolerance improves fitness in biological control strains of *Lactobacillus plantarum* in plant environments. *PLOS ONE* 13:e0190931.
- Daranas, N., Roselló, G., Cabrefiga, J., Donati, I., Francés, J., Badosa, E., Spinelli, F., Montesinos, E., and Bonaterra, A. 2019. Biological control of bacterial plant diseases with *Lactobacillus plantarum* strains selected for their broad-spectrum activity. *Annals of Applied Biology* 174:92-105.
- De Baere, T., Verhelst, R., Labit, C., Verschraegen, G., Wauters, G., Claeys, G., and Vaneechoutte, M. 2004. Bacteremic infection with *Pantoea ananatis*. *Journal of Clinical Microbiology* 42:4393-4395.
- Dean, R., Van Kan, J.A.L., Pretorius, Z.A., Hammond-Kosack, K.E., Di Pietro, A., Spanu, P.D., Rudd, J.J., Dickman, M., Kahmann, R., Ellis, J., and Foster, G.D. 2012. The Top 10 fungal pathogens in molecular plant pathology. *Molecular Plant Pathology* 13:414-430.

- Dickie, G.A., and Bell, C.R. 1995. A full factorial analysis of nine factors influencing in vitro antagonistic screens for potential biocontrol agents. *Canadian journal of microbiology* 41:284-293.
- Durán, P., Thiergart, T., Garrido-Oter, R., Agler, M., Kemen, E., Schulze-Lefert, P., and Hacquard, S. 2018. Microbial Interkingdom Interactions in Roots Promote *Arabidopsis* Survival. *Cell* 175:973-983.e914.
- Failor, K.C., Schmale III, D.G., Vinatzer, B.A., and Monteil, C.L. 2017. Ice nucleation active bacteria in precipitation are genetically diverse and nucleate ice by employing different mechanisms. *The ISME Journal* 11:2740–2753.
- Felnagle, E.A., Jackson, E.E., Chan, Y.A., Podevels, A.M., Berti, A.D., McMahon, M.D., and Thomas, M.G. 2008. Nonribosomal peptide synthetases involved in the production of medically relevant natural products. *Molecular Pharmaceutics* 5:191-211.
- Förster H., McGhee, G.C., Sundin G.W., and Adaskaveg, J.E. . 2015. Characterization of Streptomycin Resistance in Isolates of *Erwinia amylovora* in California. *Phytopathology* 105:1302-1310.
- Gerami, E., Hasanzadeh, N., Abdollahi, H., Ghasemi, A., and Heydari, A. 2013. Evaluation of some bacterial antagonists for biological control of fire blight disease. *Journal of Plant Pathology* 95(1):127-134
- Giddens, S.R., Feng, Y., and Mahanty, H.K. 2002. Characterization of a novel phenazine antibiotic gene cluster in *Erwinia herbicola* Eh1087. *Molecular Microbiology* 45:769-783.
- Özaktan, H., Bora, T., Sukan, S., Sargin, S. and Vardar Sukan, F. 1999. Studies on determination of antagonistic potential and biopreparation of some bacteria against the Fireblight pathogen. *Acta Horticulturae* 489:663 - 668.
- Ishimaru, C.A., Klos, E.J., and Brubaker, R.R. 1988. Multiple Antibiotic Production by *Erwinia herbicola*. *Phytopathology* 78:746-750.

- Jaffuel, G., Imperiali, N., Shelby, K., Campos-Herrera, R., Geisert, R., Maurhofer, M., Loper, J., Keel, C., Turlings, T.C.J., and Hibbard, B.E. 2019. Protecting maize from rootworm damage with the combined application of arbuscular mycorrhizal fungi, *Pseudomonas* bacteria and entomopathogenic nematodes. *Scientific Reports* 9:3127.
- Johnson, K.B., Stockwell, V.O., Sawyer, T.L., and Sugar, D. 2000. Assessment of Environmental Factors Influencing Growth and Spread of *Pantoea agglomerans* on and Among Blossoms of Pear and Apple. *Phytopathology* 90:1285-1294.
- Karnetová, J., Tax, J., Stajner, K., Vaněk, Z., and Krumphanzl, V. 1983. Production of phenazines by *Streptomyces cinnamonensis*. *Folia Microbiologica* 28:51-53.
- Kim, D., Langmead, B., and Salzberg, S.L. 2015. HISAT: a fast spliced aligner with low memory requirements. *Nature Methods* 12:357.
- Kloepper, J.W., Leong, J., Teintze, M., and Schroth, M.N. 1980. *Pseudomonas* siderophores: A mechanism explaining disease-suppressive soils. *Current Microbiology* 4:317-320.
- Koboldt, D.C., Chen, K., Wylie, T., Larson, D.E., McLellan, M.D., Mardis, E.R., Weinstock, G.M., Wilson, R.K., and Ding, L. 2009. VarScan: variant detection in massively parallel sequencing of individual and pooled samples. *Bioinformatics* 25:2283-2285.
- Koczan, J.M., Lenneman, B.R., McGrath, M.J., and Sundin, G.W. 2011. Cell surface attachment structures contribute to biofilm formation and xylem colonization by *Erwinia amylovora*. *Applied and environmental microbiology* 77:7031-7039.
- La Fuente, L.D., Thomashow, L., Weller, D., Bajsa, N., Quagliotto, L., Chernin, L., and Arias, A. 2004. *Pseudomonas Fluorescens* UP61 Isolated From Birdsfoot Trefoil Rhizosphere Produces Multiple Antibiotics and Exerts a Broad Spectrum of Biocontrol Activity. *European Journal of Plant Pathology* 110:671-681.
- Lee, H.B., Hong, J.P., and Kim, S.B. 2010. First Report of Leaf Blight Caused by *Pantoea agglomerans* on Rice in Korea. *Plant Disease* 94:1372-1372.



- Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., Durbin, R., and Genome Project Data Processing, S. 2009. The Sequence Alignment/Map format and SAMtools. *Bioinformatics* 25:2078-2079.
- Lindow, S.E. 1983. The Role of Bacterial ICE Nucleation in Frost Injury to Plants. *Annual Review of Phytopathology* 21:363-384.
- Lindow, S.E. 1991. Determinants of Epiphytic Fitness in Bacteria. In *Microbial Ecology of Leaves*, J.H. Andrews and S.S. Hirano, eds (New York, NY: Springer New York), pp. 295-314.
- Loper, J.E., Henkels, M.D., Roberts, R.G., Grove, G.G., Willet, M.J., and Smith, T.J. 1991. Evaluation of streptomycin, oxytetracycline and copper resistance of *Erwinia amylovora* isolated from pear orchards in Washington State. *Plant Disease* 75:287-290.
- Matyjaszczyk, E. 2015. Products containing microorganisms as a tool in integrated pest management and the rules of their market placement in the European Union. *Pest Management Science* 71:1201-1206.
- McManus, P.S., and Jones, A.L. 1994. Epidemiology and Genetic Analysis of Streptomycin-Resistant *Erwinia amylovora* from Michigan and Evaluation of Oxytetracycline for Control. *Phytopathology* 84:627-633.
- McMurdie, P.J., and Holmes, S. 2013. phyloseq: An R Package for Reproducible Interactive Analysis and Graphics of Microbiome Census Data. *PLOS ONE* 8(4):e61217.
- Medema, M.H., Blin, K., Cimermancic, P., de Jager, V., Zakrzewski, P., Fischbach, M.A., Weber, T., Takano, E., and Breitling, R. 2011. antiSMASH: rapid identification, annotation and analysis of secondary metabolite biosynthesis gene clusters in bacterial and fungal genome sequences. *Nucleic Acids Research* 39:W339-W346.
- Mora, I., Cabrefiga, J., and Montesinos, E. 2015. Cyclic Lipopeptide Biosynthetic Genes and Products, and Inhibitory Activity of Plant-Associated *Bacillus* against Phytopathogenic Bacteria. *PLOS ONE* 10:e0127738.

- Neeno-Eckwall, E.C., Kinkel, L.L., and Schottel, J.L. 2001. Competition and antibiosis in the biological control of potato scab. *Canadian journal of microbiology* 47:332-340.
- Norelli, J.L., Jones, A.L., and Aldwinckle, H.S. 2003. Fire Blight Management in the Twenty-first Century: Using New Technologies that Enhance Host Resistance in Apple. *Plant Disease* 87:756-765.
- Nuclo, R.L., Johnson, K.B., Stockwell, V.O., and Sugar, D. 1998. Secondary Colonization of Pear Blossoms by Two Bacterial Antagonists of the Fire Blight Pathogen. *Plant Disease* 82:661-668.
- Pieterse, C.M.J., Zamioudis, C., Berendsen, R.L., Weller, D.M., Van Wees, S.C.M., and Bakker, P.A.H.M. 2014. Induced Systemic Resistance by Beneficial Microbes. *Annual Review of Phytopathology* 52:347-375.
- Polymenakou, P.N. 2012. Atmosphere: A Source of Pathogenic or Beneficial Microbes? . *Atmosphere* 3:87-102.
- Pujol, M., Badosa, E., Manceau, C., and Montesinos, E. 2006. Assessment of the Environmental Fate of the Biological Control Agent of Fire Blight, *Pseudomonas fluorescens*; EPS62e, on Apple by Culture and Real-Time PCR Methods. *Applied and Environmental Microbiology* 72:2421-2427.
- Pusey, L.P. 1999. Laboratory and field trials with selected microorganisms as biocontrol agents for fire blight. *Acta Horticulturae* 489:655-661.
- Quast, C., Pruesse, E., Yilmaz, P., Gerken, J., Schweer, T., Yarza, P., Peplies, J., and Glöckner, F.O. 2013. The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic Acids Research* 41:D590-D596.
- Reganold, J.P., and Wachter, J.M. 2016. Organic agriculture in the twenty-first century. *Nature Plants* 2:15221.
- Rezzonico, F., Smits, T.H.M., Montesinos, E., Frey, J.E., and Duffy, B. 2009. Genotypic comparison of *Pantoea agglomerans* plant and clinical strains. *BMC Microbiology* 9:204.

- Roselló, G., Bonaterra, A., Francés, J., Montesinos, L., Badosa, E., and Montesinos, E. 2013. Biological control of fire blight of apple and pear with antagonistic *Lactobacillus plantarum*. *European Journal of Plant Pathology* 137:621-633.
- RStudio. 2015. RStudio: Integrated Development for R. RStudio, Inc., Boston, MA.
- Russo, N.L., Burr, T.J., Breth, D.I., and Aldwinckle, H.S. 2008. Isolation of Streptomycin-Resistant Isolates of *Erwinia amylovora* in New York. *Plant Disease* 92:714-718.
- Seemann, T. 2014. Prokka: rapid prokaryotic genome annotation. *Bioinformatics* 30:2068-2069.
- Sharifazizi, M., Harighi, B., and Sadeghi, A. 2017. Evaluation of biological control of *Erwinia amylovora*, causal agent of fire blight disease of pear by antagonistic bacteria. *Biological Control* 104:28-34.
- Smits, T.H., Rezzonico, F., Blom, J., Goesmann, A., Abelli, A., Kron Morelli, R., Vanneste, J.L., and Duffy, B. 2015. Draft Genome Sequence of the Commercial Biocontrol Strain *Pantoea agglomerans* P10c. *Genome announcements* 3.
- Stockwell, V.O., Johnson, K.B., and Loper, J.E. 1998. Establishment of Bacterial Antagonists of *Erwinia amylovora* on Pear and Apple Blossoms as Influenced by Inoculum Preparation. *Phytopathology* 88:506-513.
- Sundin, G.W., and Wang, N. 2018. Antibiotic Resistance in Plant-Pathogenic Bacteria. *Annual Review of Phytopathology* 56:161-180.
- Sundin, G.W., Werner, N.A., Yoder, K.S., and Aldwinckle, H.S. 2009. Field Evaluation of Biological Control of Fire Blight in the Eastern United States. *Plant Disease* 93:386-394.
- Tancos, K.A., Villani, S., Kuehne, S., Borejsza-Wysocka, E., Breth, D., Carol, J., Aldwinckle, H.S., and Cox, K.D. 2015. Prevalence of Streptomycin-Resistant *Erwinia amylovora* in New York Apple Orchards. *Plant Disease* 100:802-809.

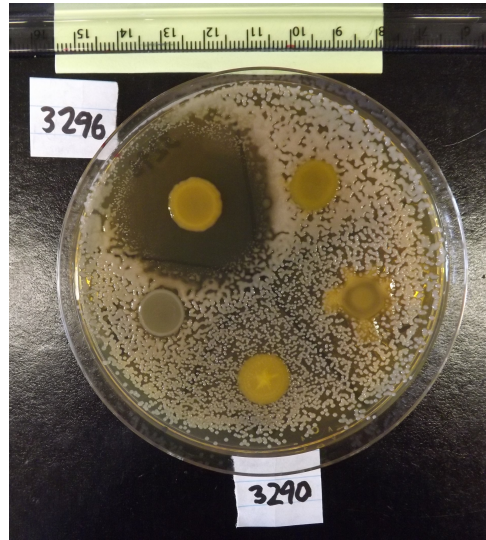
- Thomashow, L.S., and Weller, D.M. 1988. Role of a phenazine antibiotic from *Pseudomonas fluorescens* in biological control of *Gaeumannomyces graminis* var. *tritici*. *Journal of Bacteriology* 170:3499-3508.
- Thomson, S.V. 1985. The Role of the Stigma in Fire Blight Infections. *Phytopathology* 76:476-482.
- Tian, L., Huang, C., Heath, L.S., and Vinatzer, B.A. 2019. LINbase: A Web service for genome-based identification of microbes as members of crowdsourced taxa. *bioRxiv*:752212.
- Temple, T.N., Stockwell, V.O., Loper, J.E. and Johnson, K.B. 2004. Bioavailability of Iron to *Pseudomonas fluorescens* Strain A506 on Flowers of Pear and Apple. *Phytopathology* 94:1286-1294.
- Van Wees, S.C.M., Pieterse, C.M.J., Trijssenaar, A., Van 't Westende, Y.A.M., Hartog, F., and Van Loon, L.C. 1997. Differential Induction of Systemic Resistance in Arabidopsis by Biocontrol Bacteria. *Molecular Plant-Microbe Interactions* 10:716-724.
- Weißhaupt, S., Köhl, L., Kunz, S., Hinze, M., Ernst, M., Schmid, A., and Voegelé, R.T. 2016. Alternative inoculum sources for fire blight: the potential role of fruit mummies and non-host plants. *Plant Pathology* 65:470-483.
- Wick, R.R., Schultz, M.B., Zobel, J., and Holt, K.E. 2015. Bandage: interactive visualization of de novo genome assemblies. *Bioinformatics* 31:3350-3352.
- Wick, R.R., Judd, L.M., Gorrie, C.L., and Holt, K.E. 2017. Unicycler: Resolving bacterial genome assemblies from short and long sequencing reads. *PLOS Computational Biology* 13:e1005595.
- Wilson, M., and Lindow, S.E. 1994. Inoculum Density-Dependent Mortality and Colonization of the Phyllosphere by *Pseudomonas syringae*. *Applied and environmental microbiology* 60:2232-2237.
- Wilson, M., Epton, H.A.S., and Sigeo, D.C. 1992. Biological Control of Fire Blight of Hawthorn (*Crataegus monogyna*) with Fluorescent *Pseudomonas* spp. under Protected Conditions. *Journal of Phytopathology* 136:16-26.

- Wilson M., Epton, H.A.S. and Sigeo, D.C. 1992. Interactions Between *Erwinia herbicola* and *E. amylovora* on the Stigma of Hawthorn Blossoms. *Phytopathology* 82:914-918.
- Wilson, M. and Lindow, S.E. 1992. Interactions between the biological control agent *Pseudomonas fluorescens* A506 and *Erwinia amylovora* in pear blossoms. *Phytopathology* 83:117-123.
- Wright, S.A.I., Zumoff, C.H., Schneider, L., and Beer, S.V. 2001. *Pantoea agglomerans* Strain EH318 Produces Two Antibiotics That Inhibit *Erwinia amylovora* *in vitro*. *Applied and Environmental Microbiology* 67:284.
- Wu, L., Liu, R., Niu, Y., Lin, H., Ye, W., Guo, L., and Hu, X. 2016. Whole genome sequence of *Pantoea ananatis* R100, an antagonistic bacterium isolated from rice seed. *Journal of Biotechnology* 225:1-2.

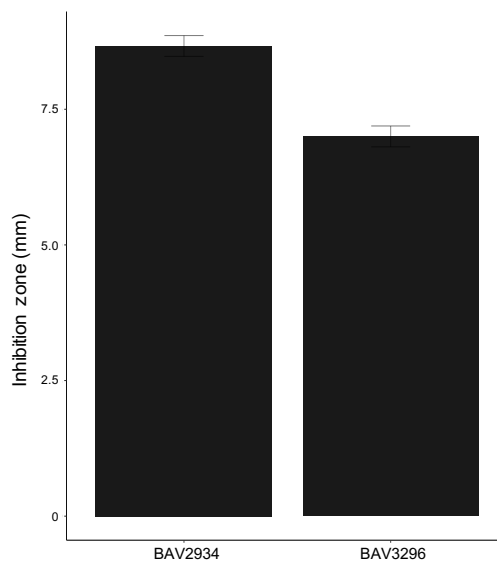
## 11. Figures:

**Figure 1.** Dual culture screening assay showing the inhibitory activity of rain-isolated bacteria against *E. amylovora* represented by A) the halo inhibition where the pathogen did not grow, and B) average size of the inhibition zone induced by BAV2934 and BAV3296.

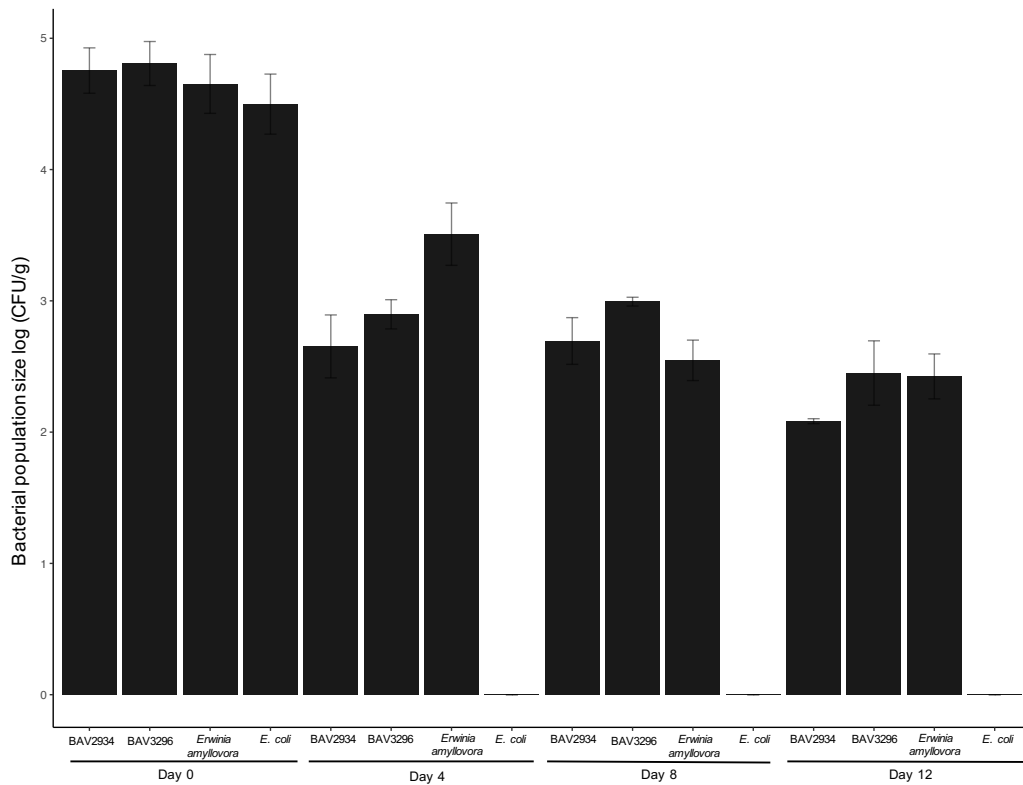
A



B

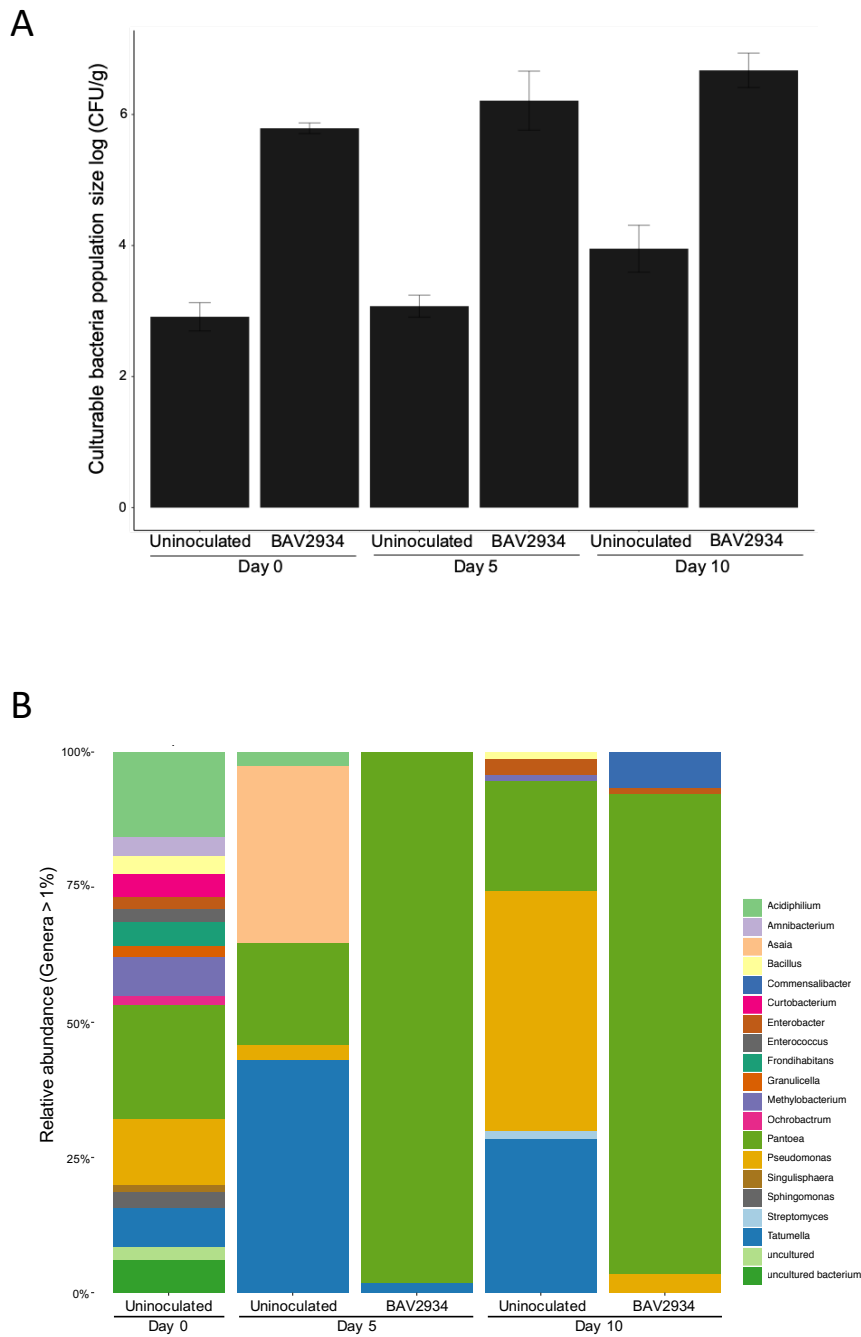


**Figure 2.** Survival, measured as population size, of rain-isolated bacteria BAV2934 and BAV3296 compared with *E. amylovora* on detached apple branches under environmental conditions for up to 12 days. An *Escherichia coli* isolate was included as control.



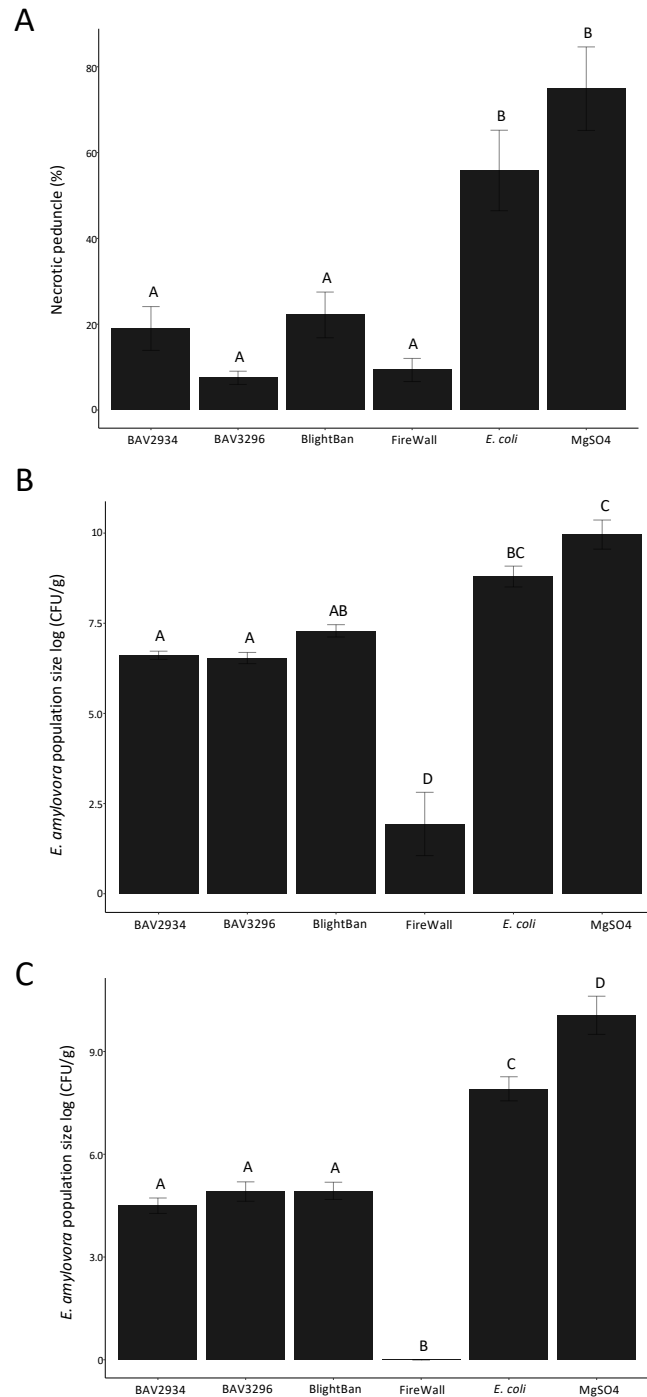
**Figure 3.** Survival of rain-isolated *Pantoea agglomerans* BAV2934 on apple blossoms in an orchard.

A) Culturable bacteria, and B) Relative abundance of bacterial taxa at the genus level (16S rRNA gene) present in blossoms up to 10 days after inoculation with BAV2934.

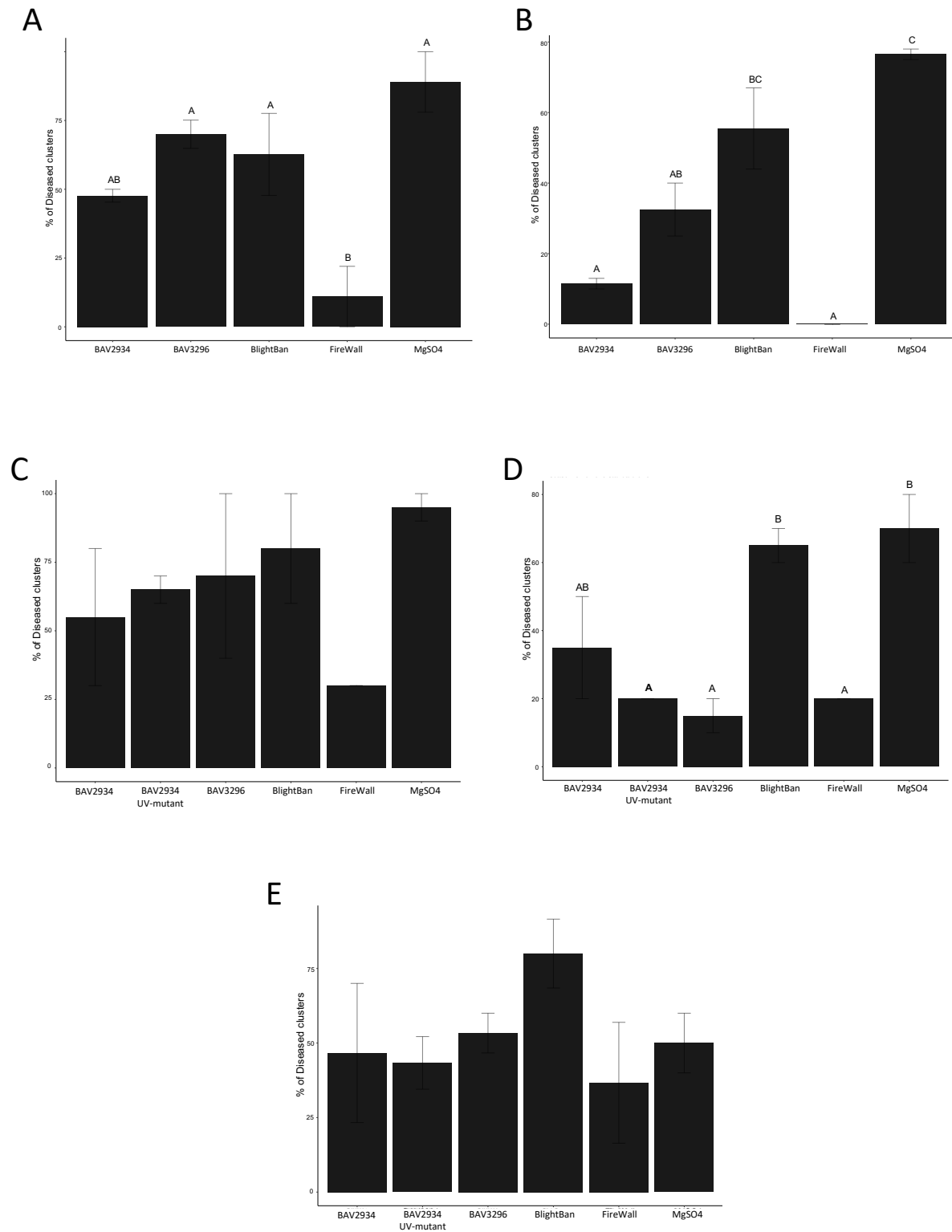




**Figure 4.** Antagonistic effect of rain-isolated bacteria BAV2934 and BAV3296 against *E. amylovora* on detached blossoms under laboratory conditions. A) Disease severity measured by the necrotic tissue caused by *E. amylovora* invasion on the pedicel, B) Pathogen population size on the receptacle, and C) Pathogen population size on the pedicel 7 days after pre-treatment with either rain-isolated bacteria or commercial products.



**Figure 5.** Fire blight control on blossoms in an apple orchard in 2018. A) Disease incidence of Fire blight on ‘Golden Delicious’ and B) Disease incidence of Fire blight on ‘York’.



## 12. Tables:

**Table 1.** Identity of 9 rain-isolated bacteria that showed antagonistic activity against *E. amylovora* in our initial inhibition screen

<b>Isolate</b>	<b>Source of Isolation</b>	<b>Date of Isolation</b>	<b>Location of rain collection</b>	<b>Isolation medium</b>	<b>Affiliation 16S rRNA</b>
BAV 2493	Rain	3/12/13	Private home, Blacksburg, VA, USA	Eosin Methylene Blue	<i>Pseudomonas sp.</i>
BAV 2502	Rain	3/13/13	Private home, Blacksburg, VA, USA	CA	<i>Bacillus sp.</i>
BAV 2572	Rain	4/22/13	Kentland Farm, Blacksburg, VA, USA	Eosin Methylene Blue	<i>Pseudomonas sp.</i>
BAV 2934	Rain	5/15/13	Kentland Farm, Blacksburg, VA, USA	R2A pH5	<i>Pantoea sp.</i>
BAV 3049	Rain	6/26/13	Kentland Farm, Blacksburg, VA, USA	Eosin Methylene Blue	<i>Pantoea sp.</i>
BAV 3226	Rain	9/26/13	Latham Hall, Blacksburg, VA, USA	R2A	<i>Pseudomonas sp.</i>
BAV 3280	Rain	9/28/13	Latham Hall, Blacksburg, VA, USA	Trypticase Soy Agar 10%	<i>Pseudomonas sp.</i>
BAV 3296	Rain	9/29/13	Latham Hall, Blacksburg, VA, USA	CA	<i>Pantoea ananatis</i>
BAV 4579	Rain	3/10/13	Private home, Blacksburg, VA, USA	Trypticase Soy Agar	<i>Pseudomonas sp.</i>

**Table 2.** Inhibition assay of rain-isolated bacteria against bacterial, fungal and oomycetes plant pathogens (part 1).

Rain Isolated bacteria		<i>Xanthomonas</i>	<i>Xanthomonas</i>	<i>Pto</i>	<i>Pto</i>	<i>Pma</i>	<i>A. citrulli</i>	<i>R.</i>	<i>A.</i>
		<i>sp.</i>	<i>sp.</i>	<i>DC3000</i>	<i>K40</i>	<i>ES4326</i>	<i>AAC00-1</i>	<i>andropogonis</i>	<i>tumefaciens</i>
BAV2493	<i>Pseudomonas sp.</i>	-	++	-	-	-	+	-	++
BAV2502	<i>Bacillus sp.</i>	-	-	-	-	++	+	+	++
BAV2572	<i>Pseudomonas sp.</i>	+	+	+	+	+	-	-	+
BAV2934	<i>P. agglomerans</i>	+	++	+	+	+	-	+	+
BAV3049	<i>Pantoea sp.</i>	-	-	-	-	++	+	+	+++
BAV3226	<i>Pseudomonas sp.</i>	+++	+++	++	++	++	+++	+++	+++
BAV3280	<i>Pseudomonas sp.</i>	+++	+	++	++	++	++	++	+
BAV3296	<i>Pantoea ananatis.</i>	+++	+++	++	+	++	++	+++	+++
BAV4579	<i>Pseudomonas sp.</i>	+	+	-	-	-	+	-	++

**Table 2.** Inhibition assay of rain-isolated bacteria against bacterial, fungal and oomycetes plant pathogens (part 2)

Rain Isolated bacteria	<i>Phytophthora</i>	<i>Botrytis</i>	<i>Botrytis</i>	<i>Botrytis</i>	<i>Colletotrichum</i>	<i>Colletotrichum</i>
	<i>capsici</i>	<i>cinerea 132</i>	<i>cinerea 156</i>	<i>cinerea 110/P6</i>	<i>acetatum C15</i>	<i>sp. C2</i>
BAV2493	<i>Pseudomonas sp.</i>	-	-	-	-	-
BAV2502	<i>Bacillus sp.</i>	-	-	-	-	-
BAV2572	<i>Pseudomonas sp.</i>	+	-	-	-	-
BAV2934	<i>P. agglomerans</i>	-	-	-	-	-
BAV3049	<i>Pantoea sp.</i>	-	-	-	-	-
BAV3226	<i>Pseudomonas sp.</i>	+++	+++	+++	++	++
BAV3280	<i>Pseudomonas sp.</i>	+++	+++	+++	++	++
BAV3296	<i>Pantoea ananatis.</i>	-	-	-	-	-
BAV4579	<i>Pseudomonas sp.</i>	++	-	-	-	-

+ = halo inhibition from 0.1 – 5 mm

++ = halo inhibition from 5.5 - 10 mm

+++ = halo inhibition > 10 mm

Table 3. BAV2934 UV-generated non-synonyms mutations

<b>Gene</b>	<b>Identity</b>	<b>Position</b>	<b>Ref</b>	<b>Alt</b>
NOOGOKNH_00074	23S ribosomal RNA	281	T	C
NOOGOKNH_00074	23S ribosomal RNA	1730	T	C
NOOGOKNH_00950	23S ribosomal RNA	281	T	C
NOOGOKNH_00950	23S ribosomal RNA	1169	G	A
NOOGOKNH_00950	23S ribosomal RNA	1730	T	C
NOOGOKNH_01463	Hypothetical protein	290	C	T
NOOGOKNH_02234	D-aminopeptidase	377	T	A
NOOGOKNH_03014	23S ribosomal RNA	281	T	C
NOOGOKNH_03014	23S ribosomal RNA	1169	G	A
NOOGOKNH_03014	23S ribosomal RNA	1730	T	C
NOOGOKNH_03542	23S ribosomal RNA	281	T	C
NOOGOKNH_03542	23S ribosomal RNA	1169	G	A
NOOGOKNH_03542	23S ribosomal RNA	1730	T	C
NOOGOKNH_03576	23S ribosomal RNA	281	T	C
NOOGOKNH_03576	23S ribosomal RNA	1730	T	C
NOOGOKNH_03827	Methionine aminopeptidase	202	C	T
NOOGOKNH_04505	Dimodular nonribosomal peptide synthase	623	G	A

\*Parameters for pooled samples (--min-coverage8 --min-var-freq 0.15 -p-value 0.05)

**Table 4.** Biosynthetic gene clusters predicted in BAV2934 by antiSMASH.

<b>Contig</b>	<b>Predicted product</b>	<b>BGC coordinates</b>	<b>BGC size</b>	<b>Most similar known cluster</b>	<b>Similarity</b>
1	Arylpolyene, Homoserine lactone	2,541,579 - 2,601,389	59,810	APE Ec	78%
1	Thiopeptide	2,667,346 - 2,693,602	26,256	O-antigen	14%
1	Homoserine lactone	3,464,471 - 3,485,109	20,638	-	
1	NRPS	3,575,053 - 3,627,116	52,063	Amonabactin	57%
2	Siderophore	158,889 - 171,243	12,354	Desferrioxamine	100%
2	Terpene	345,771 - 369,332	23,561	Carotenoid	100%
3	Bacteriocin	104,879 - 116,270	11,391	-	
4	Phenazine	78,916 - 99,392	20,476	Pyocyanine	57%

### 13. Supplementary Tables:

**Table S1.** List of rain-isolated bacteria tested in the initial inhibition assay.

**Table S2.** Antagonistic activity of rain-isolated bacteria against plant pathogens.

**Table S3.** Bacterial survival assay on apple branches under environmental conditions.

Experiment #1		
Date	Temperature (Min/Max °C)	Precipitation
18-Nov-16	12/19	-
19-Nov-16	-4/14	Light snow
20-Nov-16	-5/-4	Light snow
21-Nov-16	-6/1	-
22-Nov-16	-5/6	-
23-Nov-16	3/10	-
24-Nov-16	6/12	-
25-Nov-16	3/12	-
26-Nov-16	-2/3	-
27-Nov-16	-3/7	-
28-Nov-16	3/7	-
29-Nov-16	6/14	Heavy Rain
30-Nov-16	9/14	Rain

Experiment #2		
Date	Temperature (Min/Max °C)	Precipitation
13-Feb-17	-5/4	-
14-Feb-17	-2/5	-
15-Feb-17	-4/2	-
16-Feb-17	-7/-1	-
17-Feb-17	-3/12	-
18-Feb-17	-6/13	-
19-Feb-17	4/12	-
20-Feb-17	4/16	-
21-Feb-17	3/10	-
22-Feb-17	6/11	-
23-Feb-17	9/16	-
24-Feb-17	11/18	-
25-Feb-17	2/13	-



Experiment #3		
Date	Temperature (Min/Max °C)	Precipitation
24-Mar-17	1/17	-
25-Mar-17	11/18	-
26-Mar-17	9/14	Light rain
27-Mar-17	10/19	Light rain
28-Mar-17	10/16	Light rain
29-Mar-17	7/18	-
30-Mar-17	4/10	-
31-Mar-17	6/13	Light rain
1-Apr-17	2/7	-
2-Apr-17	0/13	-
3-Apr-17	6/12	Light rain
4-Apr-17	10/17	-
5-Apr-17	8/17	-

**Table S4.** Weather conditions during field treatments at Kentland Farm, VA.

Field Experiment 2018				
Date	Temperature (Min/Max °C)	Humidity (Min/Max %)	Precipitation	Comments
26-Apr-18	4/16	31/88	-	Application 'GD'
27-Apr-18	6/14	50/100	Light rain	-
28-Apr-18	4/17	26/47	-	-
29-Apr-18	0/9	27/64	-	-
30-Apr-18	2/16	6/53	-	Application 'York'
1-May-18	10/23	3/15	-	-
2-May-18	14/25	13/26	-	-
3-May-18	17/27	26/44	-	-
4-May-18	15/25	30/64	-	-
5-May-18	8/15	95/100	-	-
6-May-18	6/16	88/100	Rain	-
7-May-18	9/16	60/81	Light rain	-
8-May-18	11/17	57/95	-	-
9-May-18	11/21	48/66	-	-
10-May-18	14/21	40/68	-	-
11-May-18	14/25	36/62	-	Data collection 'GD'
12-May-18	18/27	29/66	-	-
13-May-18	18/27	29/60	-	-
14-May-18	18/28	41/64	-	Data collection 'York'

Field Experiment 2019				
Date	Temperature (Min/Max °C)	Humidity (Min/Max %)	Precipitation	Comments
15-Apr-19	1/4	84/99	-	Application 'Empire'
16-Apr-19	-1/18	45/100	-	-
17-Apr-19	13/22	46/58	-	-
18-Apr-19	13/22	59/72	Light rain	-
19-Apr-19	12/16	95/100	Light rain	-
20-Apr-19	4/14	76/100	Light rain	-
21-Apr-19	3/11	75/100	-	-
22-Apr-19	4/19	47/99	-	-
23-Apr-19	11/25	41/65	-	-
24-Apr-19	13/21	64/94	-	-
25-Apr-19	14/23	58/90	-	-
26-Apr-19	6/15	91/97	Light rain	-
27-Apr-19	4/15	43/73	-	-
28-Apr-19	8/19	65/84	-	-

29-Apr-19	4/14	78/99	Light snow	Application 'GD', 'Rome'
30-Apr-19	11/24	61/96	-	Data collection 'Empire'
1-May-19	16/25	68/90	-	-
2-May-19	16/23	74/89	-	-
3-May-19	15/23	70/96	Light rain	-
4-May-19	14/21	75/100	Light rain	-
5-May-19	12/18	88/100	-	-
6-May-19	10/20	56/100	-	-
7-May-19	13/21	72/96	-	-
8-May-19	15/23	59/98	-	-
9-May-19	15/20	87/100	Light rain	-
10-May-19	15/23	68/99	-	Data collection 'GD', 'Rome'

Population size of BAV2934 at the Kentland Farm 2018				
Date	Temperature (Min/Max °C)	Humidity (Min/Max %)	Precipitation	Comments
25-Apr-18	4/13	92/100	-	-
26-Apr-18	4/16	31/88	-	-
27-Apr-18	6/14	50/100	Light rain	-
28-Apr-18	4/17	26/47	-	-
29-Apr-18	0/9	27/64	-	-
30-Apr-18	2/16	6/53	-	-
1-May-18	10/23	3/15	-	-
2-May-18	14/25	13/24	-	-
3-May-18	17/27	26/44	-	-
4-May-18	15/25	30/64	-	-
5-May-18	8/15	95/100	Light rain	-

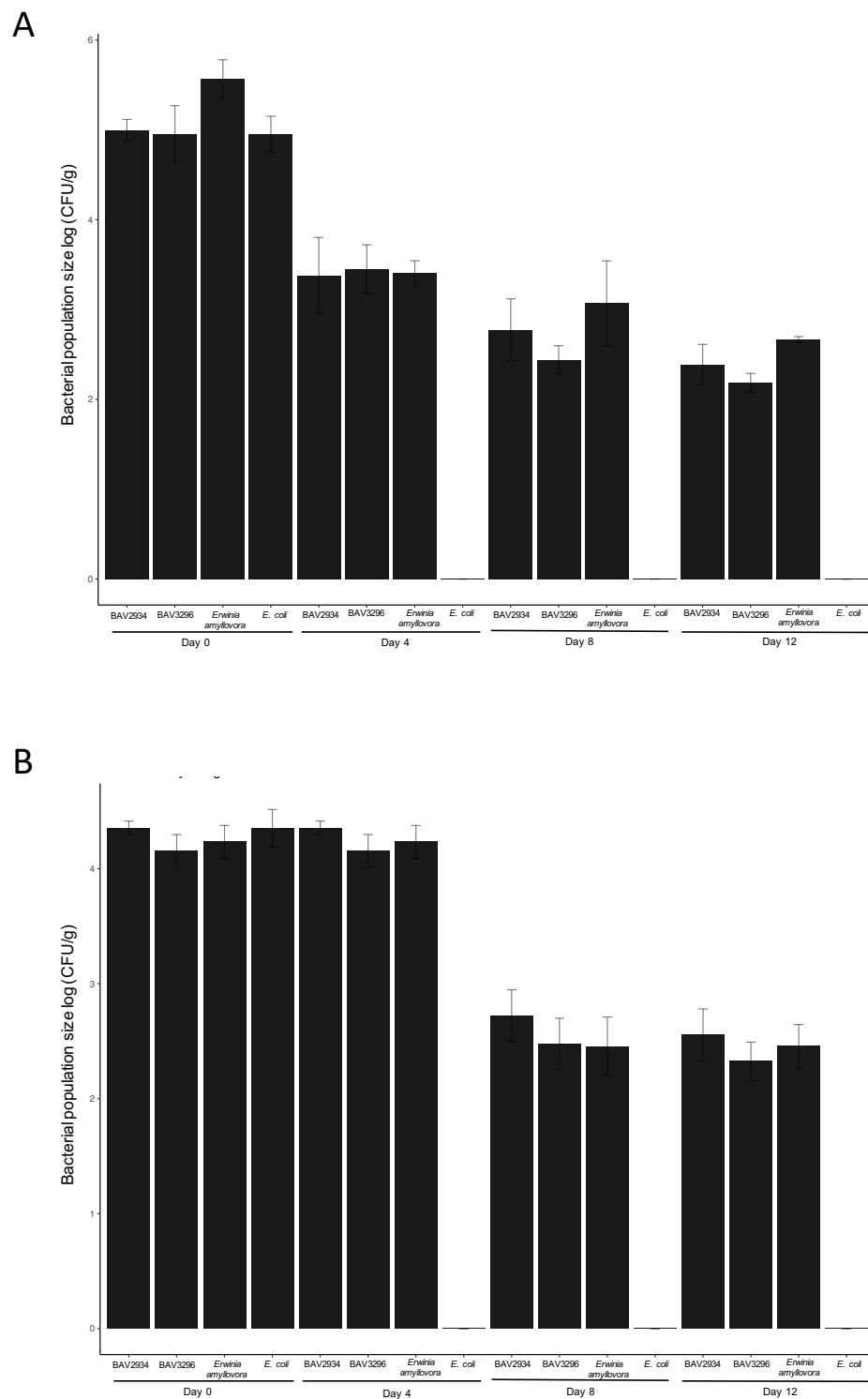
**Table S5.** Infection (% infected clusters) and control (% reduction compared to MgSO<sub>4</sub>) in the field.

2018	BAV2934		BAV3296		BlightBan		FireWall		MgSO <sub>4</sub>	
	% inf	control	% inf	control	% inf	control	% inf	control	% inf	control
G. Del	47	46.6	68	22.7	67	23.9	13	85.2	88	--
York	11	85.5	33	56.6	53	30.3	--	1	76	--
Mean-all	29	66	50.5	39.6	60	27.1	6.5	43.1	82	--

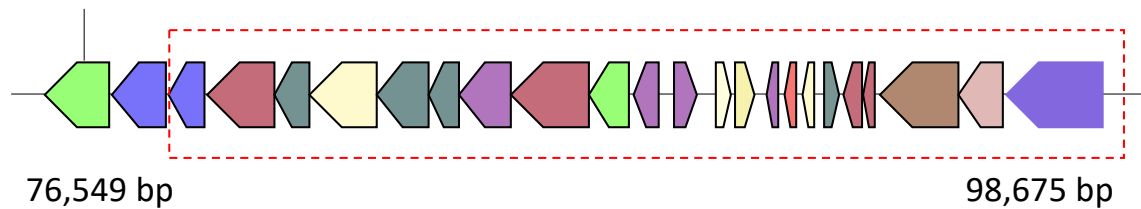
2019	BAV2934		BAV3296		BlightBan		FireWall		MgSO <sub>4</sub>		2934 mutant	
	% inf	control	% inf	control	% inf	control	% inf	control	% inf	control	% inf	control
G. Del	55	31.3	70	12.5	80	0	30	62.5	80	--	65	18.8
Empire	47	6	53	0	80	0	37	26	50	--	43	1
Rome	35	50	15	70	65	7.1	20	98.6	70	--	20	71.4
Mean-all	45.7	31.5	46	31	75	<0	29	56.5	66.7	--	42.7	40
Mean GD&R	45	40	42.5	43.3	72.5	3.3	25	66.7	75	--	42.5	43.3

## 14. Supplementary Figures:

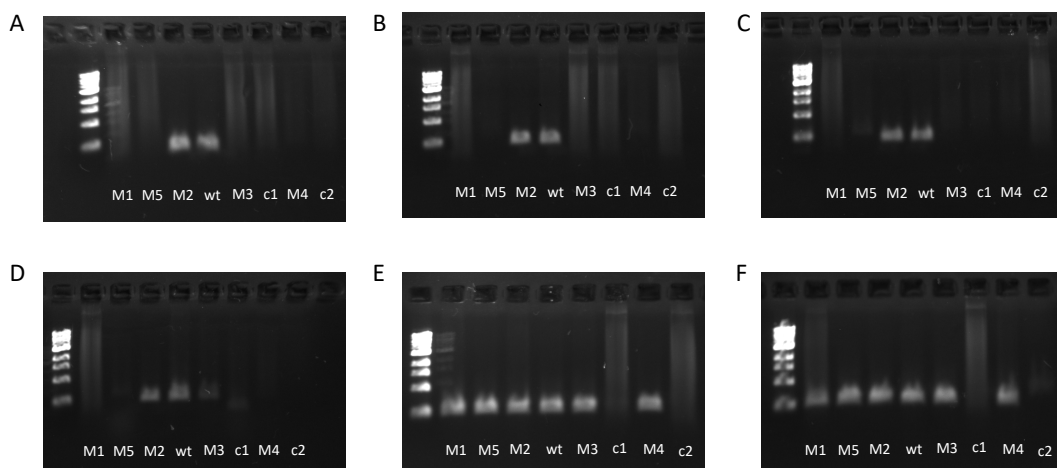
**Supplementary Figure 1.** Bacterial survival rate of rain-isolated BAV2934, BAV3296 and *E. amylovora* BAV5616 on apple branches under environmental conditions in A) November 2016 and B) March 2017 under environmental conditions.



**Supplementary Figure 2.** Identified BGC predicted to encode a phenazine compound. The red dashed box indicates the gene cluster identified by antiSMASH. The black vertical line indicates the location of the base pair mutation. JGI IMG base pair coordinates are indicated.



**Supplementary Figure 3.** PCR detection of contig 4 in BAV2934. A) Contig 4 gene NOOGOKNH\_04505, B) Contig 4 gene NOOGOKNH\_04435, C) Contig 4 gene NOOGOKNH\_04535, D) Contig 4 gene NOOGOKNH\_04609, E) Chromosomal gene NOOGOKNH\_2217, and F) Chromosomal gene NOOGOKNH\_02219.



**Chapter four:**

**Comprehensive characterization of an Aspen (*Populus tremuloides*) leaf litter sample  
that maintained ice nucleation activity for 48 years**

Yalda Vasebi<sup>1,2,\*</sup>, Marco E. Mechan Llontop<sup>1,\*</sup>, Regina Hanlon<sup>1</sup>, David G. Schmale III<sup>1</sup>,  
Russell Schnell<sup>3</sup>, Boris A. Vinatzer<sup>1</sup>

<sup>1</sup>School of Plant and Environmental Sciences, Virginia Tech, Blacksburg, VA 24061, USA

<sup>2</sup>Plant Protection Department, Faculty of Agriculture, University of Tabriz, Tabriz, 51368, Iran

<sup>3</sup>National Oceanic and Atmospheric Administration, Global Monitoring Division, Boulder, CO  
80303, USA

\*These two authors contributed equally

*Correspondence to:* Boris A. Vinatzer ([vinatzer@vt.edu](mailto:vinatzer@vt.edu))

\*\* *Published by the journal Biogeosciences*

Doi: <https://doi.org/10.5194/bg-16-1675-2019>

**Abstract:**

Decaying vegetation was determined to be a potentially important source of atmospheric ice nucleation particles (INPs) in the early 1970s. The bacterium *Pseudomonas syringae* was the first microorganism with ice nucleation activity (INA) isolated from decaying leaf litter in 1974. However, the ice nucleation characteristics of *P. syringae* are not compatible with the characteristics of leaf litter-derived INPs since the latter were found to be sub-micron in size while INA of *P. syringae* depends on much larger intact bacterial cells. Here we determined the cumulative ice nucleation spectrum and microbial community composition of the historic leaf litter sample 70-S-14 collected in 1970 that conserved INA for 48 years. The majority of the leaf litter-derived INPs were confirmed to be sub-micron in size and to be sensitive to boiling. Culture-independent microbial community analysis only identified *Pseudomonas* as potential INA. Culture-dependent analysis identified one *P. syringae* isolate, two isolates of the bacterial species *Pantoea ananatis*, and one fungal isolate of *Mortierella alpina* as having INA among 1170 bacterial colonies and 277 fungal isolates, respectively. Both, *Pa. ananatis* and *M. alpina*, are organisms that produce heat-sensitive sub-micron INPs. They are thus both likely sources of the INPs present in sample 70-S-14 and may represent important terrestrial sources of atmospheric INPs, a conclusion that is in line with other recent results obtained in regard to INPs from soil, precipitation, and the atmosphere.



## 1. Introduction:

Ice-nucleating particles (INPs) are necessary to initiate freezing of cloud droplets in mixed-phase clouds in order for precipitation to form. Therefore, the concentration of atmospheric INPs affects frequency and intensity of precipitation (Mülmenstädt et al., 2015). Identifying and characterizing INPs and their sources is thus the subject of intense research with important implications for modelling climate change (Morris et al., 2014; Coluzza et al., 2017).

In their search for sources of atmospheric INPs in the 1970s, Russell Schnell and Gabor Vali at the University of Wyoming found that decaying leaf litter contained unexpectedly high concentrations of INPs active at temperatures as warm as  $-1.3^{\circ}\text{C}$  (Schnell and Vali, 1972). These very active INPs were called leaf-derived nuclei (LDN). Because LDN were absent in green leaves and increased in concentration during leaf decay, it was hypothesized that they were of microbial origin. In fact, *Pseudomonas syringae*, the most active ice nucleating microbe known to date, was later isolated from leaf litter (Maki et al., 1974). *P. syringae* is a genetically diverse species and includes plant pathogenic as well as non-pathogenic strains (Morris et al., 2013). Parallel to the discovery of *P. syringae* with INA in leaf litter, *P. syringae* with INA was also isolated from dried powdered corn leaves and identified as the culprit of frost damage in crop plants (Arny et al., 1976) and *P. syringae* with INA have been isolated from clouds, precipitations, and surface water (Morris et al., 2013). Because of its ubiquitous presence in compartments of the water cycle and other circumstantial evidence, *P. syringae* might contribute to the formation of precipitation in clouds (Morris et al., 2013). Other bacteria with INA have been identified in additional *Pseudomonas* species (Failor et al., 2017) and in additional genera within the Gammaproteobacteria: *Erwinia* (Phelps et al., 1986), *Pantoea* (Failor et al., 2017), a genus that includes several species classified in the past as members of the genus *Erwinia*, and *Xanthomonas* (Kim et al., 1987). A strain belonging to the genus *Lysinibacillus* was recently identified and characterized as the first ice nucleating Gram-

positive bacterium (Failor et al., 2017). Some fungal species have been found to have INA as well: *Fusarium acuminatum* and *F. avenaceum* (Pouleur et al., 1992), *F. tricinctum* and *F. oxysporum* (Richard et al., 1996), *F. sporotrichioides* (Huffman et al., 2013), *Mortierella alpina* (Fröhlich-Nowoisky et al., 2015), *Isaria farinosa*, and *Acremonium implicatum* (Huffman et al., 2013). An additional fungus with INA is the lichen fungus *Rhizoplaca chrysoleuca* (Kieft, 1988). Finally, pollen was added to this long list of biological INA (Diehl et al., 2001).

Microorganisms with INA produce different types of INA molecules. The Gammaproteobacteria produce a membrane-anchored INA protein that is inserted into the bacterial cell wall and is thus heat-labile and cannot be separated from bacterial cells by filtration through membrane filters with a 0.22  $\mu\text{m}$  pore size (Morris et al., 2004). However, in the case of *Erwinia herbicola* (now known as *Pantoea ananatis*), the INA protein can be released from cells as part of extra-cellular vesicles (EV) (Phelps et al., 1986). Since EVs are submicron in size (tens to hundreds of nm), these vesicles can pass through a 0.22  $\mu\text{m}$  filter.

All biological INPs outside of the Gammaproteobacteria pass through 0.22  $\mu\text{m}$  filters. The INPs produced by the INA isolate of the genus *Lysinibacillus* are moreover very heat-stable (remaining partially active even after an hour of boiling), proteinase-resistant, and they do not pass through a 100kDa protein filter with a pore size of approximately 50 nm. This suggests that they consist in relatively large non-proteinaceous secreted single molecules or stable aggregates of multiple smaller molecules (Failor et al., 2017). Fungal INPs are also mostly retained by 100kDa filters suggesting that they also consist of large secreted single molecules or of large aggregates of smaller molecules (Fröhlich-Nowoisky et al., 2015; O'Sullivan et al., 2015). Many fungal INPs are heat stable up to 60°C but do not resist boiling and are variably susceptible to proteinases and thus believed to be mostly proteinaceous (Fröhlich-Nowoisky et al., 2015; O'Sullivan et al., 2015). Pollen INPs are mostly retained by

100kDa filters, are heat and proteinase resistant (Pummer et al., 2012), and have been proposed to consist in polysaccharides of various chain lengths (Dreischmeier et al., 2017).

Now that all these biological INPs have been identified, and at least partially characterized, we decided to go back to one of the original leaf litter samples, the Aspen (*Populus tremuloides*) leaf litter sample 70-S-14, which had been collected by Russell Schnell in 1970 and which characterization in regard to its INA was published in 1976 (Schnell and Vali, 1976). Intriguingly, sample 70-S-14 maintained a high concentration of INA until today, 48 years after collection, while being stored at room temperature. INPs in this sample had been found to be less than 0.1  $\mu\text{m}$  in size (Schnell and Vali, 1976). Therefore, *Pseudomonas* species can be excluded as the main source of INPs in this sample. While the INPs in the sample could be of plant origin (Pummer et al., 2015), here we tested the specific hypothesis that the INPs in this leaf litter sample were produced by strains of *Lysinibacillus* or by other bacteria or fungi. To test this hypothesis, we performed a comprehensive characterization of sample 70-S-14 in regard to its content in INPs and the composition of its microbial community.

## **2. Materials and Methods:**

### *2.1 Description of leaf litter sample 70-S-14 and how it has been stored:*

Leaf litter sample 70-S-14 was collected in a grove of *Populus tremuloides* (Aspen) about 1 mile west of Penhold Airbase, Alberta, now known as Red Deer Regional Airport (52.1762° N, 113.8870° W) in the summer of 1970. Even though the dominate litter in the grove was from decaying Aspen leaves that fell the previous fall, there were decaying leaves from shrubs and grasses that also were collected in the sample. To collect the litter, handfuls of litter were grasped in a one-meter square to fill a garbage bag. The litter was somewhat moist as it was collected in the shade of the trees. The grove is still there and litter samples from the same area have been collected many times over the intervening decades, most recently in June 2018. The

litter still produces prodigious numbers of ice nuclei with freezing activity beginning at  $-4.5^{\circ}\text{C}$  to  $-5^{\circ}\text{C}$  in concentrations of  $10^8$  per gram of leaf litter active at  $-10^{\circ}\text{C}$ .

Some months after collection, when the litter sample was dry, the twigs and other large debris were removed. Some of the cleaned sample was ground to a fine powder in a fluid energy mill and used for INA tests in the Colorado State University Cloud Chamber where it proved to produce excellent INPs (Schnell and Vali, 1976, Fig. 3).

Sample 70-S-14 was stored at room temperature inside 4 different plastic bags and a paper bag from 1970 to 2018 at room temperature open to the air in Schnell's various offices and a portion mailed to Virginia Tech in 2017. At Virginia Tech, it was again stored at room temperature in a paper bag until processing.

#### *2.2 Characterization of the cumulative ice nucleation spectrum of leaf litter sample 70-S-14:*

One gram of leaf litter was added to 100 ml double-distilled water (DDW) in a 250 ml flask under sterile conditions and stirred for 5 minutes. The leaf litter suspension was passed through a  $2.5\ \mu\text{m}$  pore size Whatman cellulose filter paper (GE Healthcare, USA) to remove large leaf fragments. This primary suspension was then subject to 10 different treatments (see Figure 1) and INA was tested after each treatment to determine the main characteristics of the INPs present in the leaf litter sample.

In short, 10 ml of the primary suspension were passed through a  $0.22\ \mu\text{m}$  filter (Supor® 200 PES membrane Disc Filter, PALL, USA) to select for sub-micron INPs in the filtrate and for intact bacterial and fungal cells in the  $0.22\ \mu\text{m}$  filter retentate resuspended from the filter in the same volume (10 ml). 5 ml of the  $0.22\ \mu\text{m}$  filtrate were then passed through a 100kDa protein filter with an approximate pore size of 50nm (Macrosep Advance Centrifugal Device with 100K MWCO, PALL, USA) for 10 min at 5,000 rpm to separate sub 50 nm-sized INPs in the 100kDa filter filtrate from INPs larger than 50 nm in the 100kDa filter retentate

resuspended in the same volume (5 ml). Portions of the original suspension and of each filtrate and retentate were also boiled for 15 min to determine heat sensitivity of each fraction.

Dilutions from  $10^{-1}$  to  $10^{-6}$  were made after each of the 10 treatments to obtain cumulative ice nucleation spectra. 30 drops of 20  $\mu$ l of each dilution were tested at -8 °C, -10 °C, and -12 °C maintaining each temperature for 10 min on parafilm boats floating in a cryobath (Lauda Alpha RA24, LAUDA-Brinkmann, Delran, NJ, USA). The entire assay, including the ten treatments, was repeated three times. The concentration of ice nuclei was calculated using the approach developed by (Vali, 1971) and described previously (Failor et al., 2017).

### *2.3 Culture-independent microbial community analysis:*

To determine the overall composition of the fungal and bacterial communities in the leaf litter sample, a culture-independent approach was used first. The primary leaf litter suspension described above was vacuum-filtered through a 0.22  $\mu$ m pore-size filter membrane (Supor® 200 PES membrane Disc Filter, PALL, USA). DNA was extracted from the filter with the DNeasy PowerWater kit (Qiagen, USA) according to the manufacturer's protocol. DNA concentration and quality were evaluated by UV spectrophotometry (NanoDrop 1000, Thermo, USA) and visualized on a 1% agarose gel.

For bacterial community analysis, the V4 hypervariable region of the 16S rRNA gene was amplified and sequenced using barcoded versions of the primers 799F (anti-chloroplast, 5'AACMGGATTAGATACCCKG3') and 1115R ("universal," 5'AGGGTTGCGCTCGTTG3'). For fungal community analysis, the ITS 2 region was amplified and sequenced using primers ITS9\_F (GAACGCAGCRAAIIGYGA) and ITS4\_R (TCCTCCGCTTATTGATATGC). All steps from PCR to paired-end ( $2 \times 300$  bp) amplicon sequencing on the Illumina MiSeq platform were performed at Molecular Research LP (MR DNA™, Shallowater, TX, USA).

Paired-end sequences were joined together into a single-sequence read. Quality trimming was performed and barcodes and primer sequences were depleted. Then, sequences shorter than 200bp, sequences with ambiguous base calls, and sequences containing homopolymers longer than 6bp were removed. High-quality sequences were processed using the Quantitative Insights into Microbial Ecology (QIIME) bioinformatic pipeline (Caporaso et al., 2010). Operational Taxonomic Units (OTUs) were assigned using an open-reference approach with a threshold of 97% sequence similarity. OTU picking and taxonomy assignment were performed using UCLUST and the SILVA database (Quast et al., 2013). Fungal ITS paired-end sequences were processed as described above but UNITE was used as the database instead (Abarenkov et al., 2010).

All OTUs assigned to mitochondria or chloroplasts and OTUs with fewer than five reads were excluded from further analysis. QIIME-generated output files were imported to R for data visualization using the Phyloseq 1.19.1 (McMurdie and Holmes, 2013) and ggplot2 2.2.1 (Wickham, 2009) packages.

#### *2.4. Culture-dependent bacterial community analysis:*

In parallel to the culture-independent approach, bacteria were cultured to determine the composition of the bacterial community present in the leaf litter sample using a culture-dependent approach. The undiluted primary leaf litter suspension described above and  $10^{-1}$  and  $10^{-2}$  dilutions were plated on R2A (Reasoner and Geldreich, 1985). Agar plates were incubated at 28 °C for 2-7 days. After incubation, the bacterial population was estimated by counting the bacterial colonies. This assay was performed three times. 100 colonies were randomly selected for identification using PCR followed by Sanger sequencing of the V4 hypervariable region of the 16S rRNA gene as described previously (Failor et al., 2017).

### 2.5. INA testing of individual bacterial colonies:

The above described suspensions and  $10^{-1}$  and  $10^{-2}$  dilutions were plated on R2A and LEM (Lysinibacillus Enrichment Medium). LEM is based on a medium originally developed for *Lysinibacillus sphaericus* (Russell et al., 1989) to enrich for *Lysinibacillus* strains similar to the *Lysinibacillus* strain with INA that we recently isolated (Failor et al., 2017). LEM contains per liter:  $\text{Na}_2\text{HPO}_4$ , 5.57g;  $\text{KH}_2\text{PO}_4$ , 2.4g;  $(\text{NH}_4)_2\text{SO}_4$ , 2.0g;  $\text{MgSO}_4 \cdot 7\text{H}_2\text{O}$ , 50mg;  $\text{MnCl}_2 \cdot 4\text{H}_2\text{O}$ , 4.0mg;  $\text{FeSO}_4 \cdot 7\text{H}_2\text{O}$ , 810 $\mu\text{g}$ ;  $\text{CaCl}_2 \cdot 2\text{H}_2\text{O}$ , 1.5mg;  $\text{H}_2\text{SO}_4$ , 0.3 $\mu\text{l}$ ; sodium acetate, 321.8mg; thiamine 200mg; biotin, 20 $\mu\text{g}$ . After incubation for up to 7 days, random bacterial colonies were resuspended in 140  $\mu\text{l}$  of DDW. Five drops of 20  $\mu\text{l}$  from each colony suspension were tested for INA at  $-8^\circ\text{C}$ ,  $-10^\circ\text{C}$ , and  $-12^\circ\text{C}$ . Colonies for which at least one drop froze at any of the used temperatures were streaked on to new plates. Bacterial suspensions were made from the new plates for these colonies and five drops were tested for INA a second time. If still positive, a third test was performed similar to the one described to determine the cumulative ice nucleation spectrum of leaf litter 70-S-14: 30 drops of 20  $\mu\text{l}$  each were tested for each dilution of a  $10^{-1}$  to  $10^{-6}$  dilution series starting with a bacterial suspension of approximately  $1 \times 10^8$  cfu/ml. The same was done to determine the cumulative ice nucleation spectrum of *Pa. ananatis* BAV 3057. In this case, a bacterial suspension at a concentration of  $3.2 \times 10^8$  cfu/ml was used as starting suspension for the  $10^{-1}$  to  $10^{-6}$  dilution series and the bacterial starting suspension was subject to the same differential filtration and boiling combinations as the leaf litter.

### 2.6. INA testing of individual fungal colonies:

Ten-fold serial dilutions were made from the primary suspension described above. One hundred microliter of each dilution were plated on potato dextrose agar (PDA) medium supplemented with either streptomycin (20mg/L) or lactic Acid (0.1%) to restrict bacterial growth. Plates were incubated for 7 days at room temperature. Single fungal colonies were

transferred to new PDA plates supplemented with streptomycin (20mg/L) and kept at room temperature. For all fungal isolates, mycelium was scraped off the agar and resuspended in 1ml of DDW and 3 drops of 20  $\mu$ l from each fungal suspension were tested for INA as described above. Fungal isolates for which at least one drop froze at any of the used temperatures were transferred to new PDA plates. To confirm INA in fungal isolates, new PDA plates were used to prepare fungal suspensions and thirty drops were tested for a second time. For the cumulative ice nucleation spectrum of *M. alpina* strain LL118, 1g of mycelium was scraped off the PDA plates and suspended in 100ml of DDW and processed as described in the protocol used for the cumulative ice nucleation spectrum of leaf litter 70-S-14.

### **3. Results:**

#### *3.1. Aspen leaf litter 70-S-14 contains mostly INPs of submicron size that are heat-sensitive:*

Differential filtration using a 0.22  $\mu$ m pore size filter and a 100kDa filter (approximately 50 nm pore size) in combination with boiling each filtrate and each retentate allowed us to determine the concentration of total INPs in the leaf litter sample and their approximate size and heat sensitivity (Figure 2). While there was no detectable INA at -6°C and above, INA increased strongly when the temperature was lowered to -8°C, at which temperature we detected  $10^7$  INPs per gram of leaf litter. The concentration of active INPs increased further to almost  $10^8$  per gram at -10°C and -12°C. Based on the earlier tests of ice nucleation on less aged 70-S-14 by Schnell and Vali (1976), it appears that sample of 70-S-14 lost between 1°C -1.5°C of threshold ice nucleation activity and an order of magnitude in total INP concentration active at -10°C to -12°C over the 48 years of storage.

Almost all INPs passed through the 0.22  $\mu$ m filter confirming the results obtained by Schnell in 1976 that leaf litter mostly contains sub-micron INPs (Schnell and Vali, 1976). The retentate of the 0.22  $\mu$ m filter further confirmed this result since it had a very low concentration



of INPs, approximately only 1/100 to 1/1000 of the 0.22  $\mu\text{m}$  filtrate. Boiling the original unfiltered suspension and the 0.22  $\mu\text{m}$  filtrate also reduced the INP concentration to 1/100 to 1/1000 revealing that the majority of INPs present in the leaf litter is heat sensitive. Further passing the 0.22  $\mu\text{m}$  filtrate through the 100kDa filter reduced the concentration of INPs active at  $-8^{\circ}\text{C}$  approximately 100-fold while the concentration of INPs active at  $-12^{\circ}\text{C}$  was reduced approximately 10-fold. This suggests that the majority of sub-micron INPs is larger than 50 nm and that a fraction of INPs is even smaller than 50 nm. Intriguingly, resuspensions of the INPs from the 100kDa filter revealed a concentration of INPs in the 100kDa filter retentate that was even lower than that in the 100kDa filtrate. This could be due to the majority of INPs strongly binding to the filter. When the filtrate and the retentate of the 100kDa filter were boiled, INPs active at  $-8^{\circ}\text{C}$  were almost completely abolished further confirming the heat sensitivity of the majority of the INPs present in the 70-S-14 leaf litter sample. In summary, the majority of INPs in the Aspen leaf litter sample 70-S-14 consisted of heat-sensitive sub-micron sized particles with some being smaller than 50 nm and some being larger than 50 nm.

### *3.2. Microbial population analysis of the Aspen leaf litter 70-S-14 sample reveals few known microbial taxa with INA:*

To identify the possible biological origin of the INPs present in the Aspen leaf litter sample, the overall composition of the bacterial and fungal communities was determined. This was done using a culture-independent analysis by extracting DNA followed by amplification and sequencing of the bacterial V4 hypervariable region of the 16S rRNA gene and the fungal ITS region (Figure 3).

The main bacterial phyla found in the leaf litter were Proteobacteria (49%), Actinobacteria (34%), and Bacteroidetes (16%). Within the Proteobacteria, the genera *Pseudomonas* (13%) and *Sphingomonas* (12%) were the most common genera. Within the Actinobacteria, the most common taxon was an undescribed genus in the Microbacteriaceae

family (22%). Within the Bacteroidetes, *Flavobacterium* was the most common genus (14%). Therefore, of all bacterial taxa known to include strains with INA, only the genus *Pseudomonas* was identified in this culture-independent approach. The fungal leaf litter community contained 42% Basidiomycota, 38% unidentified phyla, and 20% Ascomycota. The most common genus was identified as *Cystofilobasidium* (14%), followed by *Venturia*, *Ceratobasidium*, and unidentified genera in the Helotiales and Tremellomycetes. None of these genera is known to include INA strains.

For bacteria, we also performed a culture-dependent analysis by plating the primary suspension on R2A medium. Based on colony counts, the leaf litter contained a total of  $3.13 \times 10^5$  colony forming units (CFU) per gram of leaf litter. One hundred random colonies were then selected for partially sequencing the 16S rRNA gene. This analysis revealed the presence of bacterial genera that are known to include strains with INA. In fact, not only were 22% of colonies identified as *Pseudomonas* but 1% each were identified as *Pantoea* and *Erwinia*. However, the genus with the highest number of colonies (34%), the Gram-positive bacterium *Clavibacter*, is not known to include any strains with INA.

### 3.3. INA testing of bacterial and fungal colonies reveals presence of INA strains belonging to the bacterial species *P. syringae* and *Pa. ananatis* strains and the fungal species *M. alpina*:

To determine if any culturable bacteria or fungi with INA were still present in the Aspen leaf litter 48 years after collection, a total of 1170 bacterial colonies either grown on R2A (881 colonies including the 100 colonies described above) or LEM (289 colonies) and 277 fungal isolates grown on PDA were tested for INA.

Only three bacterial colonies were found to have stable INA in all tests starting at -8°C or above. The strains were identified by sequencing the hypervariable V4 region of the 16S rRNA gene. Two strains were identified as members of the genus *Pantoea*: one strain had 99% DNA identity over 935 nt with *P. ananatis* strain Ta030 (NCBI accession number MH973238)

and the other strain had 99% DNA identity over 922 nt with *P. ananatis* strain 12WE (NCBI accession number MH010898.1). The third strain was identified as member of the genus *Pseudomonas* since it had 99% DNA identity over 794 nt with *P. syringae* pv. *syringae* strain CFBP4215 (NCBI accession number LT962480).

Among the fungal isolates, only one was found to have stable INA at -8°C. This isolate, LL118, was identified as a member of the species *M. alpina* based on ITS sequencing since it had 99% identity to *M. alpina* isolate F08ID36 with NCBI accession number KJ469836.1 (Fröhlich-Nowoisky et al., 2015).

Unfortunately, the *P. syringae* and the two *Pa. ananatis* strains isolated from the leaf litter were lost and could not be further characterized. However, we characterized the cumulative ice nucleation spectrum of *M. alpina* LL118 and of *Pa. ananatis* strain BAV 3057, which was previously isolated from rain (Failor et al., 2017) and that had over 99% identity in its 16S rRNA sequence to one of the two *Pa. ananatis* strains from leaf litter. Because of the genetic similarity of *Pa. ananatis* BAV 3057 to the two *Pa. ananatis* strains isolated from the leaf litter, these strains can be expected to have very similar cumulative ice nucleation spectra. The number of INPs produced by *M. alpina* LL118 per gram of fungal mycelium active at -7°C (the temperature at which drops of the fungal suspension started freezing) was approximately  $1 \times 10^6$ . The number of INPs produced by *Pa. ananatis* BAV 3057 at -5°C (the temperature at which drops of the bacterial suspension started freezing) was approximately  $1 \times 10^5$  per CFU. Figure 4 shows how the INPs produced by both, *M. alpina* LL118 and *Pa. ananatis* BAV 3057, were mostly heat-sensitive. *M. alpina* INPs were all sub-micron in size and the majority could be resuspended from the 100kDa filter and were thus larger than 50 nm. Only approximately 1/10 of *M. alpina* INPs passed through the 100kDa filter and were thus smaller than 50 nm. *Pa. ananatis* INPs active at -5°C were mostly larger than 0.22  $\mu\text{m}$  but 1/10th of the *Pa. ananatis* INPs active at -9°C and below were sub-micron in size. Similarly to

*M. alpina*, more of the sub-micron INPs could be resuspended from the 100kDa filter (and were thus larger than 50nm) compared to the INPs that passed through the filter (and were thus smaller than 50nm).

#### **4. Discussion and Conclusions:**

Aspen leaf litter sample 70-S-14 (Schnell and Vali, 1976) collected in 1970 has shown remarkably stable INA over 48 years. When its characterization was first published in 1976, *P. syringae* was the only known organism with INA and microbial community analysis was not possible. Therefore, with the discovery of many bacterial, fungal, and plant INA organisms since then and today's ease of determining the composition of microbial communities, this historic sample represented a great opportunity for re-evaluation of its INA and identification of the INA organisms possibly still present and alive in this sample 48 years later.

The cumulative INPs spectra of 70-S-14 obtained after a combination of filtration and boiling (Figure 2) clearly confirmed the result obtained in 1976 by Russel Schnell, *i.e.*, the INPs in this sample are submicron in size. Moreover, the INPs were found to be mostly heat sensitive. Since INPs produced by *Lysinibacillus* are heat resistant (Failor et al., 2017), this result excludes our initial hypothesis that recently discovered INA strains of the genus *Lysinibacillus* might have contributed to the INA of 70-S-14. Also heat-resistant INPs produced by pollen can be excluded (Pummer et al., 2012).

The result is more in line with secreted INPs produced by strains of the species *Erwinia herbicola* (Phelps et al., 1986), of which most strains were later assigned to the species *Pantoea ananatis* (Walterson and Stavrinides, 2015), and with INPs secreted by fungi in the genus *Fusarium* (O'Sullivan et al., 2015) and in the species *M. alpina* (Fröhlich-Nowoisky et al., 2015). In fact, *Erwinia/Pantoea* cells secrete some of the heat-sensitive INA protein as part of sub-micron EVs (Phelps et al., 1986) and *F. avenaceum* and *M. alpina* both secrete sub-micron

INPs that are heat sensitive and probably proteinaceous in nature (Fröhlich-Nowoisky et al., 2015;O'Sullivan et al., 2015). However, culture-independent analysis of 70-S-14 did not identify any of these organisms although a high number of bacterial 16S rRNA fragments (131,796 sequencing reads) and a high number fungal ITS fragments (153,546 sequencing reads) were sequenced. This suggests that these known INA organisms are very minor constituents of the overall microbial communities present in this leaf litter sample.

Interestingly though, sequencing a 16S rRNA fragment of a relatively small number (one hundred) of randomly selected bacterial colonies cultured from 70-S-14 did reveal the presence of *Pantoea/Erwinia* (two colonies). This may be due to enrichment of this genus by the employed culture media. Even more importantly, two *Pa. ananatis* isolates and one *M. alpina* isolate with INA were identified among the 1170 bacterial colonies and the 277 fungal isolates, respectively. The cumulative ice nucleation spectra of *M. alpina* and *Pa. ananatis* fit the overall ice nucleation spectrum of 70-S-14 (with the majority of INPs being heat sensitive and in large part sub-micron ins size) and are thus likely sources of the INPs present in 70-S-14.

Only one *Pseudomonas* strain with INA was isolated from 70-S-14. However, the typical ice nucleation spectrum of *Pseudomonas* species does not fit the leaf litter ice nucleation spectrum since most *Pseudomonas* species do not shed the INA protein as part of EVs but INA of *Pseudomonas* species is generally associated with intact bacterial cells that do not pass through 0.22  $\mu\text{m}$  pores (Failor et al., 2017;Maki et al., 1974;Morris et al., 2004).

Although *Mortierella* was the only fungal genus with known INA identified in 70-S-14, it is possible that other fungal strains are present (and culturable) at very low frequency in this leaf litter sample since the number of tested fungal isolates was relatively low (277 isolates). Therefore, we cannot exclude that testing additional fungal isolates could reveal the presence of additional fungal genera with INA in 70-S-14. However, it is also possible that

some INPs still present in the leaf litter sample 48 years after collection were originally produced by bacteria or fungi that are not culturable anymore and which DNA has been degraded. In fact, it is puzzling that the number of INPs produced per gram of pure culture of the isolated *M. alpina* strain is only tenfold larger than the number of INPs produced per gram of leaf litter. Considering that the total fungal and bacterial mass in the leaf litter probably constitutes only a minor fraction of the total leaf mass in the leaf litter sample suggests that organisms that produce a very high number of INPs must have been present in the leaf litter sample at some point. It is also possible that *M. alpina*, *Pa. ananatis*, and other INA organisms produced a much higher number of INPs when they originally grew in the decaying leaves compared to the number of INPs that they produced when grown on nutrient-rich agar plates in our lab before INA testing. In fact, low nutrient availability has been shown to increase production of INPs by various INA bacteria (Failor et al., 2017).

One important question in regard to the present study is if the presence of viable *Pa. ananatis* and *M. alpina* isolates in a leaf litter sample that has maintained INA for 48 years strengthens the evidence for a role of these organisms as source of atmospheric INPs. Several lines of evidence support a positive answer to this question. In fact, 106 strains among 593 strains found to have stable INA among 33,134 strains isolated from precipitation and tested for INA in our previous work (Failor et al., 2017) were identified as *Pantoea* or *Erwinia* as well. Also, Du and colleagues recently found submicron INPs in precipitation (Du et al., 2017). Although *Pantoea/Erwinia* and *M. alpina* and other known organisms with INA could not be identified in the analyzed precipitation samples using a culture-independent approach (Du et al., 2017), these organisms could still be present since we could not identify them in our culture-independent leaf litter analysis either but still found them by culturing. Finally, Conen and Yakutin recently identified a large fraction of heat-sensitive submicron INPs in soils from various continents (Conen and Yakutin, 2018). Since these INPs were only inactivated through

boiling but not by incubation at 60°C, they are more likely of fungal origin than a product of *Erwinia/Pantoea*. In conclusion, we think that combining the results from these recent studies with our new finding that 70-S-14 still contains viable *Pa. ananatis* and *M. alpina* with INA, supports a role of these organisms as important sources of atmospheric INPs. Importantly, finding that heat-sensitive sub-micron INPs are still active after 48 years in leaf litter suggests that leaf litter might represent an important reservoir of atmospheric INPs. The relative importance of leaf litter compared to live plants and soil as a contributor to the atmospheric pool of INPs is thus an important question that warrants further investigation.

What we could not do in the present study and what we could not do in our previous study of bacterial sources of INPs in precipitation (Failor et al., 2017) was to directly determine the presence of different genes coding for different biological INPs in metagenomic sequences. In fact, while direct culture-independent metagenomic sequencing of environmental samples is possible today (Behzad et al., 2015), the limitation is that the only gene known so far to encode a biological molecule with INA is the INA gene of the Gammaproteobacteria, including *Pseudomonas* species, *Xanthomonas* species, and *Pantoea/Erwinia* species (Edwards et al., 1994). Identifying the genetic basis of biological INPs produced by additional bacteria and by fungi would instead allow determination of the presence of all these various INA genes in environmental samples, such as soil, plants, leaf litter, precipitation, and even clouds. Comparison of presence and abundance of various INA genes between samples could in turn help infer the migration of microbes with INA among environments and their relative contribution to atmospheric INPs.

## **5. Acknowledgements:**

This research was supported by the National Science Foundation under grant IOS-1754721. Funding to Boris A. Vinatzer and David G. Schmale III was also provided in part by the

Virginia Agricultural Experiment Station and the Hatch Program of the National Institute of Food and Agriculture, US Department of Agriculture.

## 6. References:

- Abarenkov, K., Henrik Nilsson, R., Larsson, K.-H., Alexander, I. J., Eberhardt, U., Erland, S., Høiland, K., Kjølner, R., Larsson, E., Pennanen, T., Sen, R., Taylor, A. F. S., Tedersoo, L., Ursing, B. M., Vrålstad, T., Liimatainen, K., Peintner, U., and Kõljalg, U.: The UNITE database for molecular identification of fungi – recent updates and future perspectives, *New Phytologist*, 186, 281-285, doi:10.1111/j.1469-8137.2009.03160.x, 2010.
- Army, D. C., Lindow, S. E., and Upper, C. D.: Frost sensitivity of *Zea mays* increased by application of *Pseudomonas syringae*, *Nature*, 262, 282, 10.1038/262282a0, 1976.
- Behzad, H., Gojobori, T., and Mineta, K.: Challenges and Opportunities of Airborne Metagenomics, *Genome Biology and Evolution*, 7, 1216-1226, 10.1093/gbe/evv064, 2015.
- Caporaso, J. G., Kuczynski, J., Stombaugh, J., Bittinger, K., Bushman, F. D., Costello, E. K., Fierer, N., Peña, A. G., Goodrich, J. K., Gordon, J. I., Huttley, G. A., Kelley, S. T., Knights, D., Koenig, J. E., Ley, R. E., Lozupone, C. A., McDonald, D., Muegge, B. D., Pirrung, M., Reeder, J., Sevinsky, J. R., Turnbaugh, P. J., Walters, W. A., Widmann, J., Yatsunencko, T., Zaneveld, J., and Knight, R.: QIIME allows analysis of high-throughput community sequencing data, *Nature methods*, 7, 335-336, 10.1038/nmeth.f.303, 2010.
- Coluzza, I., Creamean, J., Rossi, M., Wex, H., Alpert, P., Bianco, V., Boose, Y., Dellago, C., Felgitsch, L., Fröhlich-Nowoisky, J., Herrmann, H., Jungblut, S., Kanji, Z., Menzl, G., Moffett, B., Moritz, C., Mutzel, A., Pöschl, U., Schauperl, M., Scheel, J., Stopelli, E.,



- Stratmann, F., Grothe, H., and Schmale, D.: Perspectives on the Future of Ice Nucleation Research: Research Needs and Unanswered Questions Identified from Two International Workshops, *Atmosphere*, 8, 138, 2017.
- Conen, F., and Yakutin, M. V.: Potential relevance of *Mortierella alpina* as a source of ice nucleating particles in soil, *Biogeosciences Discussions*, 1-5, 10.5194/bg-2018-79, 2018.
- Diehl, K., Quick, C., Matthias-Maser, S., Mitra, S. K., and Jaenicke, R.: The ice nucleating ability of pollen: Part I: Laboratory studies in deposition and condensation freezing modes, *Atmospheric Research*, 58, 75-87, [https://doi.org/10.1016/S0169-8095\(01\)00091-6](https://doi.org/10.1016/S0169-8095(01)00091-6), 2001.
- Dreichmeier, K., Budke, C., Wiehemeier, L., Kottke, T., and Koop, T.: Boreal pollen contain ice-nucleating as well as ice-binding 'antifreeze' polysaccharides, *Sci Rep*, 7, 41890, 10.1038/srep41890, 2017.
- Du, R., Du, P., Lu, Z., Ren, W., Liang, Z., Qin, S., Li, Z., Wang, Y., and Fu, P.: Evidence for a missing source of efficient ice nuclei, *Sci Rep*, 7, 39673, 10.1038/srep39673, 2017.
- Edwards, A. R., Van den Bussche, R. A., Wichman, H. A., and Orser, C. S.: Unusual pattern of bacterial ice nucleation gene evolution, *Molecular Biology and Evolution*, 11, 911-920, 10.1093/oxfordjournals.molbev.a040172, 1994.
- Failor, K. C., Schmale, D. G., 3rd, Vinatzer, B. A., and Monteil, C. L.: Ice nucleation active bacteria in precipitation are genetically diverse and nucleate ice by employing different mechanisms, *ISME J*, 11, 2740-2753, 10.1038/ismej.2017.124, 2017.
- Fröhlich-Nowoisky, J., Hill, T. C. J., Pummer, B. G., Yordanova, P., Franc, G. D., and Pöschl, U.: Ice nucleation activity in the widespread soil fungus *Mortierella alpina*, *Biogeosciences*, 12, 1057-1071, 10.5194/bg-12-1057-2015, 2015.

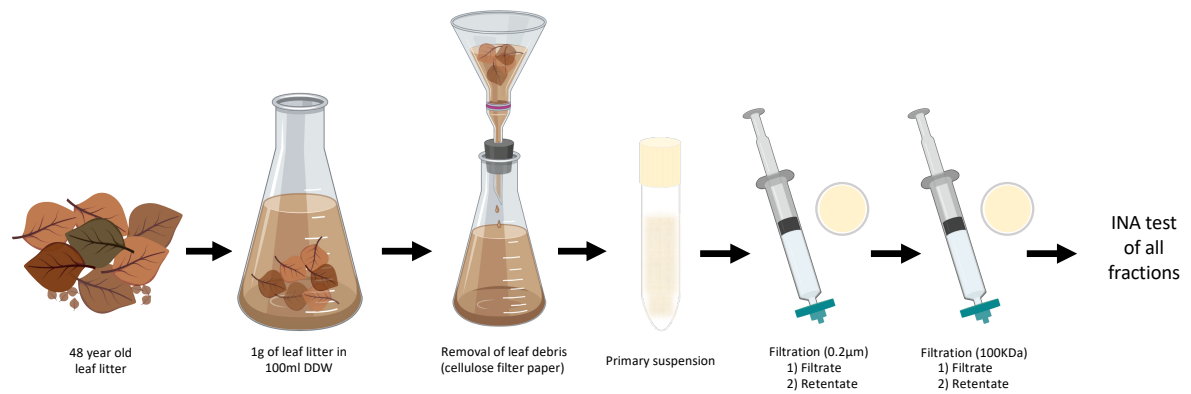
- Huffman, J. A., Prenni, A. J., DeMott, P. J., Pöhlker, C., Mason, R. H., Robinson, N. H., Fröhlich-Nowoisky, J., Tobo, Y., Després, V. R., Garcia, E., Gochis, D. J., Harris, E., Müller-Germann, I., Ruzene, C., Schmer, B., Sinha, B., Day, D. A., Andreae, M. O., Jimenez, J. L., Gallagher, M., Kreidenweis, S. M., Bertram, A. K., and Pöschl, U.: High concentrations of biological aerosol particles and ice nuclei during and after rain, *Atmospheric Chemistry and Physics*, 13, 6151-6164, 10.5194/acp-13-6151-2013, 2013.
- Kieft, T. L.: Ice Nucleation Activity in Lichens, *Applied and Environmental Microbiology*, 54, 1678-1681, 1988.
- Kim, H. K., Orser, C., Lindow, S. E., and Sands, D. C.: *Xanthomonas campestris* pv. *translucens* strains active in ice nucleation, *Plant Disease*, 994-997, 1987.
- Maki, L. R., Galyan, E. L., Chang-Chien, M.-M., and Caldwell, D. R.: Ice Nucleation Induced by *Pseudomonas syringae*, *Applied Microbiology*, 28, 456-459, 1974.
- McMurdie, P. J., and Holmes, S.: phyloseq: An R Package for Reproducible Interactive Analysis and Graphics of Microbiome Census Data, *PLOS ONE*, 8, e61217, 10.1371/journal.pone.0061217, 2013.
- Morris, C. E., Georgakopoulos, D. G., and Sands, D. C.: Ice nucleation active bacteria and their potential role in precipitation, *J. Phys. IV France*, 121, 87-103, 2004.
- Morris, C. E., Monteil, C. L., and Berge, O.: The Life History of *Pseudomonas syringae*: Linking Agriculture to Earth System Processes, *Annual Review of Phytopathology*, 51, 85-104, 10.1146/annurev-phyto-082712-102402, 2013.
- Morris, C. E., Conen, F., Alex Huffman, J., Phillips, V., Pöschl, U., and Sands, D. C.: Bioprecipitation: a feedback cycle linking Earth history, ecosystem dynamics and land use through biological ice nucleators in the atmosphere, *Global Change Biology*, 20, 341-351, doi:10.1111/gcb.12447, 2014.

- Mülmenstädt, J., Sourdeval, O., Delanoë, J., and Quaas, J.: Frequency of occurrence of rain from liquid-, mixed-, and ice-phase clouds derived from A-Train satellite retrievals, *Geophysical Research Letters*, 42, 6502-6509, doi:10.1002/2015GL064604, 2015.
- O'Sullivan, D., Murray, B. J., Ross, J. F., Whale, T. F., Price, H. C., Atkinson, J. D., Umo, N. S., and Webb, M. E.: The relevance of nanoscale biological fragments for ice nucleation in clouds, *Sci Rep*, 5, 8082, 10.1038/srep08082, 2015.
- Phelps, P., Giddings, T. H., Prochoda, M., and Fall, R.: Release of cell-free ice nuclei by *Erwinia herbicola*, *Journal of Bacteriology*, 167, 496-502, 1986.
- Pouleur, S., Richard, C., Martin, J.-G., and Antoun, H.: Ice Nucleation Activity in *Fusarium acuminatum* and *Fusarium avenaceum*, *Applied and Environmental Microbiology*, 58, 2960-2964, 1992.
- Pummer, B. G., Bauer, H., Bernardi, J., Bleicher, S., and Grothe, H.: Suspendable macromolecules are responsible for ice nucleation activity of birch and conifer pollen, *Atmospheric Chemistry and Physics*, 12, 2541-2550, 10.5194/acp-12-2541-2012, 2012.
- Pummer, B. G., Budke, C., Augustin-Bauditz, S., Niedermeier, D., Felgitsch, L., Kampf, C. J., Huber, R. G., Liedl, K. R., Loerting, T., Moschen, T., Schauperl, M., Tollinger, M., Morris, C. E., Wex, H., Grothe, H., Pöschl, U., Koop, T., and Fröhlich-Nowoisky, J.: Ice nucleation by water-soluble macromolecules, *Atmospheric Chemistry and Physics*, 15, 4077-4091, 10.5194/acp-15-4077-2015, 2015.
- Quast, C., Pruesse, E., Yilmaz, P., Gerken, J., Schweer, T., Yarza, P., Peplies, J., and Glöckner, F. O.: The SILVA ribosomal RNA gene database project: improved data processing and web-based tools, *Nucleic Acids Research*, 41, D590-D596, 10.1093/nar/gks1219, 2013.

- Reasoner, D. J., and Geldreich, E. E.: A new medium for the enumeration and subculture of bacteria from potable water, *Applied and Environmental Microbiology*, 49, 1-7, 1985.
- Richard, C., Martin, J. G., and Pouleur, S.: Ice nucleation activity identified in some phytopathogenic *Fusarium* species, *Phytoprotection*, 77, 10.7202/706104ar, 1996.
- Russell, B. L., Jelley, S. A., and Yousten, A. A.: Carbohydrate metabolism in the mosquito pathogen *Bacillus sphaericus* 2362, *Applied and Environmental Microbiology*, 55, 294-297, 1989.
- Schnell, R. C., and Vali, G.: Atmospheric Ice Nuclei from Decomposing Vegetation, *Nature*, 236, 163, 10.1038/236163a0, 1972.
- Schnell, R. C., and Vali, G.: Biogenic Ice Nuclei: Part I. Terrestrial and Marine Sources, *Journal of the Atmospheric Sciences*, 33, 1554-1564, 10.1175/1520-0469(1976)033<1554:binpit>2.0.co;2, 1976.
- Vali, G.: Quantitative Evaluation of Experimental Results on the Heterogeneous Freezing Nucleation of Supercooled Liquids, *Journal of the Atmospheric Sciences*, 28, 402-409, 10.1175/1520-0469(1971)028<0402:Qeoera>2.0.Co;2, 1971.
- Walterson, A. M., and Stavriniades, J.: *Pantoea*: insights into a highly versatile and diverse genus within the Enterobacteriaceae, *FEMS Microbiology Reviews*, 39, 968-984, 10.1093/femsre/fuv027, 2015.
- Wickham, H.: *ggplot2, Use R!*, Springer, New York, 2009.

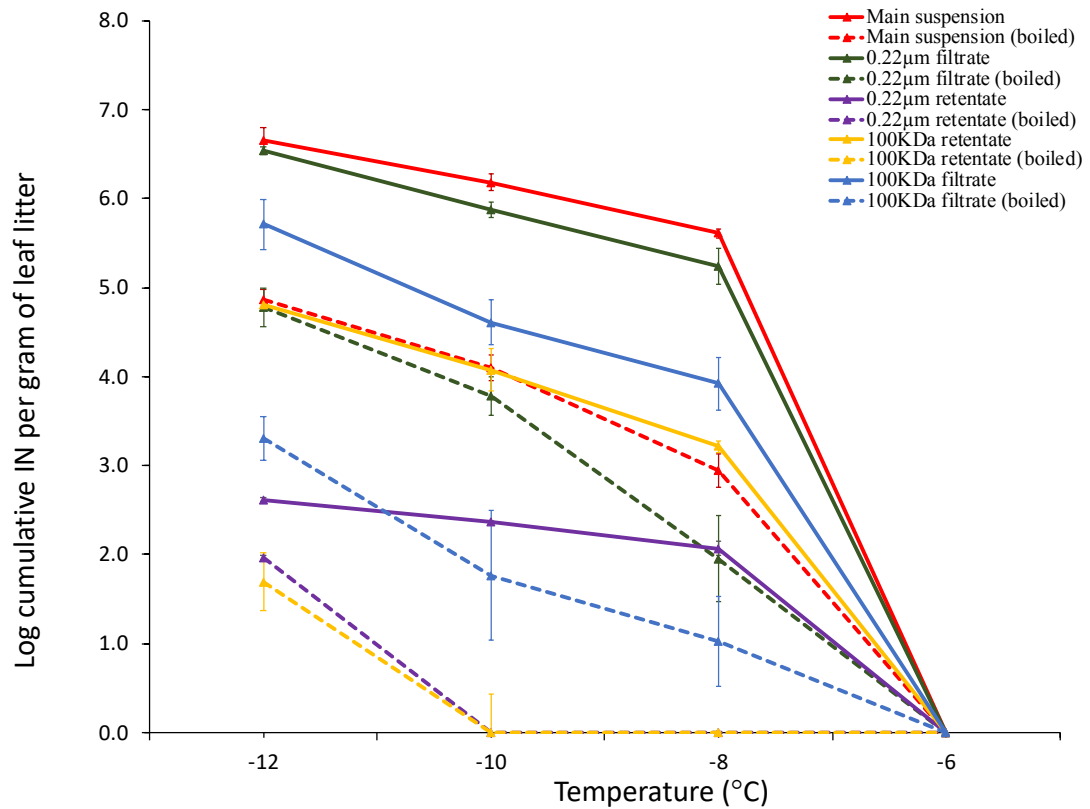
## 7. Figures:

**Figure 1: Characterization of INA of Aspen leaf litter 70-S-14.** For each biological replicate (n=4), 1g of Aspen leaf litter was suspended in 100 ml of sterile double-stilled water (DDW). The primary suspension, the filtrates, and the filter retentates of the 0.22  $\mu\text{m}$  and the 100kDa filters were tested for INA. In parallel, a portion of the primary suspension and of both filtrates and of both filter-retentates were boiled for 15 min and tested for INA as well.



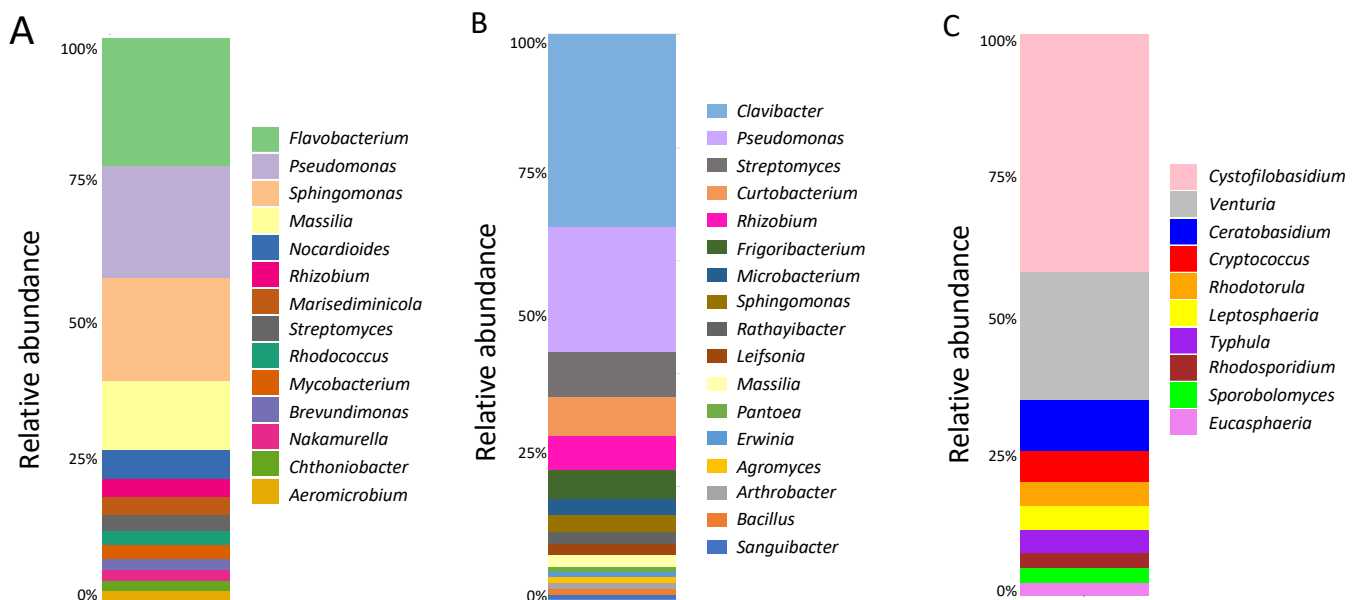
**Figure 2: Cumulative ice nucleation spectra of the Aspen leaf litter sample 70-S-14.**

Results are shown for all fractions shown in Figure 1 based on droplet freezing assays at -6°C, -8°C, -10°C, and -12°C. IN: ice nuclei.

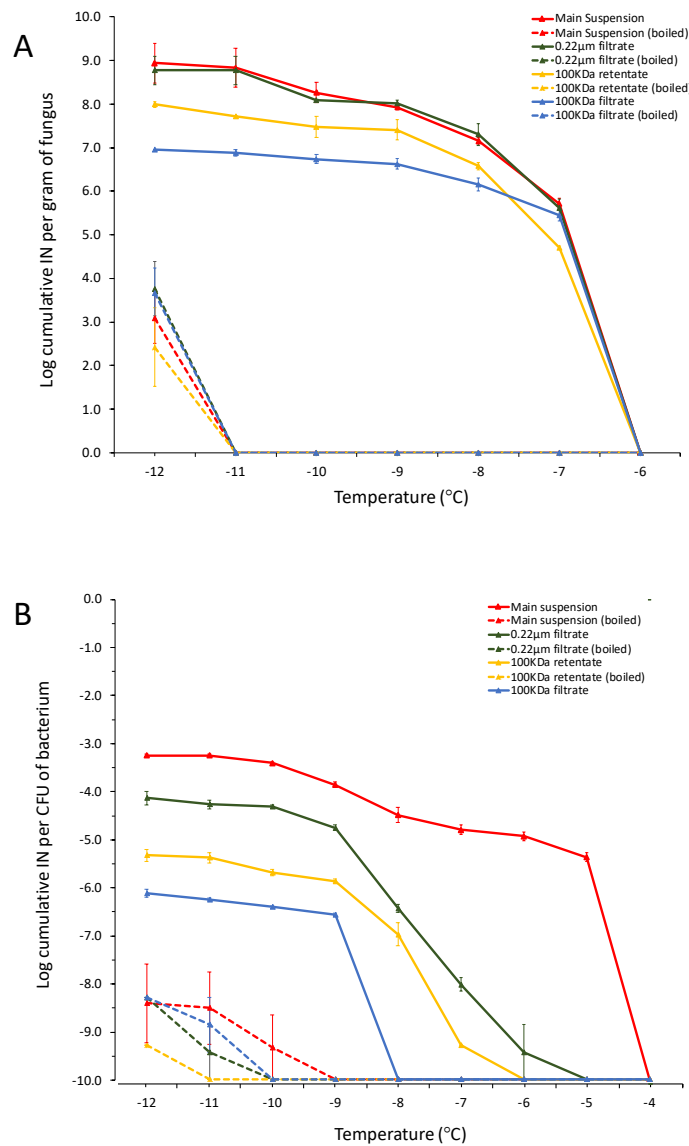


**Figure 3: Analysis of the composition of the microbial communities present in Aspen leaf litter sample 70-S-14.**

(A) Results from the culture-independent analysis of bacterial diversity based on the V4 hypervariable region of the 16S rRNA gene (considering only classified reads); (B) results from sequencing the V4 hypervariable region of the 16S rRNA gene of 100 bacterial isolates cultured from sample 70-S-14; (C) results from the culture-independent analysis of fungal diversity based on the ITS region (considering only classified reads).



**Figure 4: Cumulative ice nucleation spectra of the *M. alpina* strain LL118 (A) isolated from leaf litter sample 70-S-14 and of *Pa. ananatis* BAV 3057 (B) isolated from rain, but over 99% identical in its 16S rRNA sequence to the two *Pa. ananatis* isolates from sample 70-S-14. The analyzed fractions derived from differential filtration and boiling are the same as shown in Figure 1 besides that instead of leaf litter, suspensions of *M. alpina* mycelium and of a *Pa. ananatis* bacterial suspension were used as the starting material. IN: ice nuclei; CFU: colony forming units.**





**Chapter five:**

**Assessing the potential of culture-independent 16S rRNA microbiome analysis in  
disease diagnostics: the example of *Dianthus gratianopolitanus* and *Robbsia  
andropogonis***

Marco Enrique Mechan-Llontop, Long Tian, Vivian Angelica Bernal-Galeano, Ella Reeves,  
Mary Ann Hansen, Elizabeth Bush, and Boris Alexander Vinatzer

School of Plant and Environmental Sciences, Virginia Tech, Blacksburg, VA

Corresponding author: B.A. Vinatzer

E-mail address: [vinatzer@vt.edu](mailto:vinatzer@vt.edu)

Phone number: +1 540 231 2126

ORCID: 0000-0003-4612-225X

\*\* Published by *The European Journal of Plant Pathology*

Doi: <https://doi.org/10.1007/s10658-019-01850-8>

**Abstract:**

The goal of this study was to determine if culture-independent 16S rRNA sequencing of plant-associated microbiomes could facilitate disease diagnosis of cheddar pinks (*Dianthus gratianopolitanus*) with symptoms of leaf spotting at a Virginia nursery. The microbiome composition of cheddar pinks at the same nursery and at a second nursery in the absence of any disease outbreak was determined as well. After the pathogen was identified as *Burkholderia andropogonis* (synonym: *Robbsia andropogonis*) in a parallel culture-dependent study, the microbiome of plants artificially inoculated with *R. andropogonis* was also analyzed. The genus *Robbsia* was found to be ubiquitously present on all *Dianthus gratianopolitanus* nursery plants. However, because of the low resolution of 16S rRNA sequencing, it was not possible to determine the presence or absence of the pathogen at the species level. While relative abundance of *Robbsia* sequences slightly increased during the disease outbreak, symptomatic plants did not have a significantly higher abundance of *Robbsia* than asymptomatic plants. Only microbiomes of artificially inoculated plants were dominated by *Robbsia*. We conclude that culture-independent microbiome analysis using 16S rRNA sequencing was unable to aid disease diagnosis in this specific case. Limitations and potential of the approach in disease diagnosis in general are discussed.

**Keywords**

16S rRNA amplicon sequencing; *Dianthus*; *Robbsia andropogonis*; pathogen identification; disease diagnostics

## 1. Introduction:

Plant disease diagnosis is a key aspect of plant protection and disease control (Riley et al. 2002). Disease diagnosis involves visualization of disease symptoms, followed by culture-dependent and/or culture-independent methods. Culture-dependent methods involve the isolation of a microorganism using selective or non-selective media, often followed by additional tests, such as the analysis of carbon source utilization. Culture-independent methods include nucleic acid-based techniques, such as polymerase chain reaction (PCR), real-time PCR, loop-mediated isothermal amplification, various serological assays, as well as biosensors (Fang and Ramasamy 2015; Khater et al. 2017). All these methods are restricted to the detection of known plant pathogens, which represents a major challenge for the diagnosis of rare or emerging diseases.

Next generation sequencing (NGS) is increasingly used in culture-independent disease diagnostics, in both humans (Thorburn et al. 2015; Flight et al. 2015; Ivy et al. 2018) and animals (Raszek et al. 2016). Because the entire microbial community (*i.e.*, the microbiome) associated with an organism can be identified by NGS, an advantage of this approach in diagnostics is that it is not limited to the detection of common or known pathogens, but it can also be used to identify rare and emerging pathogens (Flight et al. 2015; Schubert et al. 2018). This can be done by sequencing PCR amplicons of either the 16S ribosomal RNA (rRNA) gene of bacteria or the internal transcribed spacer (ITS) region of fungi. Alternatively, the entire DNA (the metagenome) or the entire RNA (the metatranscriptome) present in a sample can be sequenced to identify bacteria, fungi, oomycetes, and viruses (limited to DNA viruses when sequencing the metagenome). Metagenome sequencing has recently been tested for detection of a broad range of pathogens (Chalupowicz et al. 2018). In plant disease diagnosis, RNA analysis has been used in a few cases to detect viruses (Adams et al. 2013; Boonham et al. 2014). While 16S rRNA has been used widely to study plant-associated microbiomes to

investigate many different aspects of plant-microbe interactions (Sergaki et al. 2018), its use in plant disease diagnosis has not been explored.

The genus *Dianthus* includes 15 species, known commonly as carnations or pinks, and is an important commercial ornamental worldwide (Boxriker et al. 2017). Cheddar pink (*Dianthus gratianopolitanus* hybrid Scent First® ‘Passion’) plants displaying severe leaf spotting were submitted to the Virginia Tech Plant Disease Clinic by a commercial nursery in Virginia in 2015. Culture-dependent diagnostic techniques, including proof of pathogenicity through Koch’s postulates, identified *Burkholderia andropogonis* as the causal agent (Reeves et al. 2017). Because *B. andropogonis* has recently been re-classified as *Robbsia andropogonis* (Lopes-Santos et al. 2017), we will use the synonym *R. andropogonis* throughout the remainder of this manuscript. Although this pathogen had never been identified on cheddar pinks in Virginia before, significant disease losses caused by *Burkholderia/Robbsia* species are known to occur in *Dianthus* spp. in other geographic areas (Yagi et al. 2012; Casanova et al. 2004) and *R. andropogonis* is a common pathogen that causes foliar diseases on a wide range of host plants (Lopes-Santos et al. 2017).

The aim of the present study was to take advantage of the disease outbreak on cheddar pinks to explore the use of culture-independent 16S rRNA sequencing of leaf-associated microbiomes for pathogen identification and disease diagnostics. In particular, we tested the hypothesis that the comparison of the microbiome of a small number of symptomatic plants (a typical sample size for samples received by plant disease clinics) with the microbiome of asymptomatic plants could reveal the causal agent of an unknown disease. After this hypothesis was refuted by our data, we extended the study to a more comprehensive analysis of the microbiome composition associated with cheddar pinks at two Virginia nurseries, each sampled twice in two separate years.

## 2. Materials and Methods:

### 2.1. Plant samples and plant inoculations:

All plant samples are listed in Table 1. Two asymptomatic plants and two symptomatic plants (severe leaf spotting) of cheddar pink (*Dianthus gratianopolitanus* hybrid cultivars Scent First® 'Passion' and EverLast™ 'Lavender + Eye') were received from a nursery (nursery 1) in Virginia in fall 2015. Eight additional asymptomatic plants of Scent First® 'Passion' were obtained in summer 2016 from a second nursery (nursery 2), where no disease problems had been observed. Five of these plants were maintained in the greenhouse; these were sprayed with Copper Fungicide Ready to Use® (Bonide) 11 days before inoculation with *R. andropogonis* and Abound® fungicide (active ingredient: azoxystrobin) 9 days post-inoculation to prevent fungal disease. Pathogenicity assays were performed as described previously (Reeves et al. 2017) by spraying asymptomatic plants with suspensions of *R. andropogonis* at a concentration of  $8.8 \times 10^8$  cfu/ml until runoff (besides one plant that was inoculated at  $8.8 \times 10^9$  cfu/ml; see Table 1). Plants were covered with plastic bags to maintain high humidity for 2 days and kept in the greenhouse until the end of the experiment, 24 days post inoculation.

In spring 2018, eight additional plants were obtained from nursery 1 (all Scent First® 'Passion') and four from nursery 2 (two plants of cultivar 'Neon Star' and two plants of cultivar 'Fire Witch'; no Scent First® 'Passion' plants were available from nursery 2). Neither of the two nurseries had any disease problems in 2018 and all plants were healthy. Four of the 2018 plants from nursery 1 were used to repeat the pathogenicity assay with *R. andropogonis*; this time without any pesticide treatments.

### 2.2. DNA extraction:

Ten grams of leaves were randomly collected from each plant. Leaves were neither washed nor surface-sterilized and placed in a zip-lock plastic bag. Sterile distilled water (300ml) was added

and samples were sonicated for 10 minutes using a 1510 BRANSON sonicator (Brandsonic, Mexico). The liquid was vacuum-filtered through 0.22 µm pore-size filter membranes (Supor® 200 PES membrane Disc Filter, PALL, USA) to trap bacteria dislocated from plants during sonication. Filters were stored in sealed plastic bags at -80°C until processing. DNA extraction from filters was performed using the DNeasy PowerWater kit (MoBio Laboratories, Carlsbad, CA) according to the manufacturer's protocol. DNA concentration and quality were assessed by UV spectrophotometry (NanoDrop 1000, Thermo, USA) and visualization on a 1% agarose gel.

### *2.3. 16S rRNA amplicon sequencing:*

The V4 hypervariable region of the 16S rRNA gene was amplified and sequenced using barcoded versions of the primers 799F (anti-chloroplast, 5'AACMGGATTAGATACCCKG3') and 1115R ("universal," 5'AGGGTTGCGCTCGTTG3'). The forward primer was designed to minimize amplification of plant chloroplast 16S rRNA (Shade et al. 2013). For 16S rRNA amplification, a 28 cycle PCR was performed using the HotStarTaq Plus Master Mix Kit (Qiagen, USA) under the following conditions: 94°C for 3 minutes, followed by 28 cycles of 94°C for 30 seconds, 53°C for 40 seconds and 72°C for 1 minute, after which a final elongation step at 72°C for 5 minutes was performed. After amplification, PCR products obtained from the various samples were mixed in equal concentrations and purified using Agencourt Ampure beads (Agencourt Bioscience Corporation, Beverly, MA, USA). All steps from PCR to paired-end (2 × 300 bp) amplicon sequencing on the Illumina MiSeq platform were performed at Molecular Research LP (MR DNA™, Shallowater, TX, USA).

### *2.4. Community composition analysis based on 16S rRNA sequences:*

Raw 16S rRNA paired-end sequences were processed using Molecular Research LP analysis pipeline: 1) reads were joined together after q25 trimming of the ends and reoriented in the 5'-

3' direction, 2) sequences were depleted of barcodes and primer sequences, and 3) sequences shorter than 200bp, sequences with ambiguous base calls, and sequences with homopolymer runs exceeding 6 bp were removed. Further analysis of these processed sequencing data was performed using the open source Quantitative Insights into Microbial Ecology (QIIME) bioinformatic pipeline (Caporaso et al. 2010). Reads that passed quality filtering (phred quality scores above 30) were assigned to operational taxonomic units (OTU) according to the open-reference protocol at 97% sequence identity using UCLUST as the clustering tool and SILVA (Quast et al. 2013) as the database.

All OTUs annotated as mitochondria, chloroplasts, and cyanobacteria, unassigned OTUs, and OTUs with fewer than five reads were removed from the dataset. Remaining reads from each sample were sub-sampled to the smallest sample size (*i.e.*, normalized) to avoid the effect of sample size bias when comparing community composition between samples. The QIIME-generated output file in the Biological Observation Matrix (BIOM) format was used for downstream data analysis and visualization in R version 3.3 using the Phyloseq 1.19.1 and ggplot2 2.2.1 packages (McMurdie and Holmes 2013) (Wickham 2009).

All alpha diversity and beta diversity analyses were performed in QIIME. When comparing alpha diversities between sample groups, statistical significance was determined using a nonparametric t-test with 999 Monte-Carlo permutations and p-values corrected by the Benjamini-Hochberg FDR procedure. For beta diversity analysis, weighted and unweighted UniFrac distances (Lozupone and Knight 2005) between samples were visualized using Principal Coordinates Analysis (PCoA).

### **3. Results**

#### *3.1. Analysis and comparison of the microbiome composition of cheddar pink nursery plants:*

The V4 hypervariable region of the 16S rRNA gene was amplified from microbiome-DNA extracted from leaf tissue of the following nursery plants (Table 1): I) two cheddar pinks plants with symptoms (circular, tan lesions, 1-2 mm in diameter, with purple margins, water soaking, and bacterial streaming) and two cheddar pinks without symptoms, all obtained in fall 2015 from nursery 1 during a disease outbreak (at the time of unknown etiology), II) three asymptomatic cheddar pinks collected from nursery 2 in summer 2016 in the absence of reported disease outbreak, III) eight asymptomatic cheddar pinks collected in spring 2018 (four from nursery 1 and four from nursery 2) also in the absence of any reported disease outbreak.

Amplicon sequencing using the Illumina MiSeq platform resulted in a total of 1,270,133 reads. After performing quality control of the sequences, removing non-bacterial sequences, and normalizing to the smallest sample size (7,255 reads/sample), a total of 108,825 reads remained. These reads were assigned to 4,912 OTUs.

Of the 4,912 bacterial OTUs identified in all samples from both nurseries in all years, only 13 OTUs were present in every plant sample with at least five reads per normalized sample. This very small number of OTUs thus represents the core microbiome of the leaf tissue of cheddar pink plants at the two sampled Virginia nurseries. The 13 OTUs belong to the following genera: *Robbsia* (1 OTU), *Rhodococcus* (4 OTUs), *Pantoea* (1 OTU), *Curtobacterium* (3 OTUs), *Microbacterium* (1 OTU), *Pseudomonas* (1 OTU), *Lysinimonas* (1 OTU), and *Rhizobium* (1 OTU).

At nursery 1, cheddar pinks collected during the disease outbreak in 2015 were characterized by a microbiome structure in which Proteobacteria constituted the only dominant phylum, with the most abundant families being Burkholderiaceae, Enterobacteriaceae, Bacillaceae, and Pseudomonadaceae. The most common genera were *Robbsia* (member of the *Burkholderiaceae* family), *Pantoea*, *Bacillus*, and *Pseudomonas* (Figure 1).

At nursery 2 in 2016, Proteobacteria, Actinobacteria, and Bacteroidetes were found to



dominate the phyllosphere of cheddar pink plants at the phylum level. At the family level, Microbacteriaceae, Mycobacteriaceae, Methylobacteriaceae, Nocardiaceae, Sphingomonadaceae, and Flavobacteriaceae were the dominant taxa. At the genus level, *Mycobacterium*, *Methylobacterium*, *Rhodococcus*, and *Curtobacterium* were the dominant taxa (Figure 1).

In 2018, Proteobacteria, Actinobacteria and Bacteroidetes were the dominant bacterial phyla at both nurseries. At the family level, Enterobacteriaceae, Flavobacteriaceae, Sphingomonadaceae, and Microbacteriaceae were predominant. At the genus level, *Pantoea*, *Sphingomonas*, *Methylobacterium*, *Rhodococcus*, *Chryseobacterium*, *Curtobacterium*, *Robbsia*, and *Pseudomonas* were most abundant.

When comparing the microbial communities between the two symptomatic and the two asymptomatic cheddar pinks collected from nursery 1 in 2015, we could not find a single OTU (or any higher rank bacterial taxon) that would differentiate symptomatic from asymptomatic plants. Therefore, 16S rRNA sequence analysis of cheddar pink-associated microbiomes did not reveal any putative causal agent and our initial hypothesis was rejected. Additionally, none of the 66 OTUs present in all samples collected at nursery 1 during the disease outbreak in 2015 were absent from all samples from nursery 2 collected in 2016 and absent from all samples collected at either nursery in 2018 (Supplementary Table 1). Therefore, no OTU was diagnostic of the disease outbreak at nursery 1 in 2015.

### 3.2. Relative abundance of the confirmed pathogen *R. andropogonis*:

Since *R. andropogonis* was identified as the causal agent of the 2015 disease outbreak at nursery 1 using a culture-dependent approach (Reeves et al. 2017), we specifically looked at the relative abundance of the genus *Robbsia* in all samples. Interestingly, the OTU with the highest average relative abundance (Supplementary Table 1) was found to correspond to *Robbsia* based on the SILVA database (Quast et al. 2013). Moreover, the 16S rRNA consensus

sequence of this *Robbsia* OTU was 100% identical to the 16S rRNA sequence of the *R. andropogonis* strain isolated from one of the symptomatic plants from nursery 1. (This strain was confirmed to be pathogenic on cheddar pink by Reeves et al. [2017].)

The *Robbsia* OTU was present in every sample collected from nursery 1 during the disease outbreak in 2015 (0.99%, 3.39%, 29.71%, and 36.62%). This OTU was also present in every plant collected from nursery 2 in 2016 (0.31%, 1.72% and 4.09%) and in all samples collected from both nurseries in 2018 (0.67%, 0.76%, 0.77%, 0.76%, 1.25%, 1.58%, 7.47%, and 34.54%). At nursery 1 during the 2015 disease outbreak, the *Robbsia* OTU was not even present at a higher level in the two symptomatic plants (0.99% and 29.71%) compared to the two asymptomatic plants (3.39% and 36.62%). However, the average percentage of the *Robbsia* OTU in the samples collected from nursery 1 during the 2015 disease outbreak (17.68%) was higher than the average percentage present in samples from the same nursery three years later in the absence of disease (0.86%) and it was higher than the average percentage present at nursery 2 in 2016 (2.04%) and in 2018 (11.08%).

### 3.3. Analysis and comparison of the diversity of cheddar pink microbiomes:

First, the taxonomic diversity of each individual sample (alpha diversity) was determined. This was done by counting OTUs, calculating commonly used indices of alpha diversity, and constructing rarefaction curves.

OTU counts for each sample are reported in Table 2. Figure 2 shows their statistical analysis. In short, we found an average of 547 observed OTUs per plant at nursery 1 in 2015 compared to 875 OTUs per plant at nursery 2 in 2016. However, this difference was not significant (p-value of 0.09). In 2018, an average of 1082 OTUs were observed per plant at nursery 1 and 854 OTUs were observed per plant at nursery 2. Again, the difference was not significant (p-value of 0.552). Although not significant, these OTU counts suggested that microbiomes associated with nursery plants during the disease outbreak may have been slightly

less diverse compared to the microbiomes associated with nursery plants sampled in the absence of disease. However, this could not be confirmed when calculating two indices of species richness, the Shannon diversity index and the Simpson's diversity index (Hill et al. 2003) (Figure 2). Finally, rarefaction curves did not reveal any lower diversity at nursery 1 during the disease outbreak either (Figure 3A; see also Supplementary Figure 1 for rarefaction curves drawn from raw data of each individual sample). The only observed trend was a generally lower diversity at nursery 1 in both sampling years compared to nursery 2 in both sampling years.

Although we only had two symptomatic and two asymptomatic plants from nursery 1 during the disease outbreak, which did not allow for any meaningful statistical analysis, we still wanted to see if there was any trend in number of OTUs on symptomatic plants compared to asymptomatic plants. Figure 3B shows the respective rarefaction curves. They reveal a lower diversity in the symptomatic plants. When simply counting OTUs in each sample, we found on average 455 OTUs on the symptomatic plants compared to 640 OTUs on the asymptomatic plants. This trend of a lower number of OTUs on symptomatic plants compared to asymptomatic plants was consistent when extending the comparison to all asymptomatic plants at both nurseries in all years (which increased the average number of OTUs per asymptomatic plant to 896).

We then determined the taxonomic diversity of the microbiome (beta diversity) between different plant samples using the UniFrac method, which is based on a phylogenetic composition analysis (Lozupone and Knight 2005). Principal Coordinates Analysis (PCoA) of weighted UniFrac distances was performed; this accounts for the relative abundance of OTUs, but no obvious clustering of samples emerged (Figure 4). However, when testing the strength of the association between the taxonomic composition of samples with the nursery of sample origin (nursery 1 versus nursery 2) and the year of sampling (2015 versus 2016 versus 2018),

year was a significant factor (ANOSIM: R statistics=0.777, p-value=0.023) for nursery 2 samples and for all nursery plants combined (ANOSIM: R statistics=0.501, p-value=0.001) in determining the taxonomic composition of microbiomes. These associations remained significant when using unweighted UniFrac distances, which do not take into account the relative abundance of OTUs. Moreover, when using unweighted UniFrac distances instead of weighted UniFrac distances, sampling year became a significant factor for nursery 1 as well (ANOSIM: R statistics=0.812, p-value=0.02). Interestingly, nursery of origin was not found to be a significant factor.

#### 3.4. The microbiome of cheddar pink plants artificially inoculated with *R. andropogonis*:

When Koch's postulates were performed by artificially inoculating cheddar pinks from nursery 2 with *R. andropogonis*, artificially inoculated plants developed similar disease symptoms to naturally infected nursery plants at 17 days post-inoculation (Reeves et al. 2017). See Supplementary Figure 2. However, the microbiome composition and alpha diversity of the artificially inoculated plants was very different from that of naturally infected plants (Figure 1 and Table 2): while *Robbsia* only represented between 0.99 and 29.71% of the total reads in the symptomatic nursery plants, *Robbsia* represented over 92% of reads in plants artificially inoculated with *R. andropogonis* at 17 days and 24 days post-inoculation. When the inoculation experiment was repeated in 2018 the relative abundance of *Robbsia* was slightly lower, but still much higher than in nursery plants: *Robbsia* represented 77.16% of reads at 17 days post-inoculation and over 59.6% of reads at 24 days post-inoculation.

#### 4. Discussion:

Culture-dependent identification of causal agents of uncommon or emerging plant diseases can take days to weeks, and fulfillment of Koch's postulates may take even longer. Therefore, we reasoned that culture-independent identification of putative pathogens directly from a couple

of symptomatic plants (a typical sample size for plant material received by plant disease clinics for diagnosis) may accelerate disease diagnosis. In particular, we wanted to determine if comparing the microbiome compositions of symptomatic plants with asymptomatic plants by 16S rRNA sequencing could identify the causal agent of a disease outbreak on cheddar pink nursery plants.

Unexpectedly, we found that 16S rRNA sequences with 100% identity to the 16S rRNA sequence of the confirmed pathogen *R. andropogonis* were present on symptomatic as well as asymptomatic plants at nursery 1 collected during the disease outbreak in 2015, on plants at a second nursery without any reported disease problems in 2016, and on all plants three years later at both nurseries (also without any reported disease problems). Moreover, the relative abundance of the *Robbsia* OTU was not even significantly higher in symptomatic compared to asymptomatic plants. This result suggests that *R. andropogonis* is a ubiquitous component of the bacterial community on cheddar pink nursery plants and that the observed disease outbreak was due to environmental conditions and cultural practices that were conducive to disease development. Our initial hypothesis that the pathogen could be identified by its higher relative abundance on symptomatic plants compared to asymptomatic plants could not be confirmed. However, we caution that we randomly collected leaves from each plant. Therefore, it is possible that if we had collected the most symptomatic leaves from the symptomatic plants, the relative abundance of the pathogen may have been significantly higher compared to the asymptomatic plants. Also, results may be very different for other plant pathogens that are not ubiquitous, e.g. for accidentally introduced pathogens, such as the select agent *Ralstonia solanacearum* race 3 biovar 2, which is not present in the USA (Clarke et al. 2015). Finding even a small number of 16S rRNA sequencing reads corresponding to *R. solanacearum* in the microbiome of an imported geranium plant could alert of a possible introduction of *R. solanacearum* race 3 biovar 2, triggering immediate eradication procedures.

A limitation of our 16S rRNA sequencing approach is the limited taxonomic resolution it provides. Different strains belonging to the same species or to closely related species of the same genus cannot be distinguished based on the V4 hypervariable region of the 16S rRNA gene, in particular, when using Illumina sequencing with relatively short read lengths (shorter than 300pb). Therefore, it is possible that the 16S rRNA sequences that were assigned to the *Robbsia* OTU actually originated from several different species of *Robbsia*. For example, we cannot be sure that the *Robbsia* sequences from nursery 1 during the disease outbreak came only from the cultured pathogenic *R. andropogonis* strain (Reeves et al. 2017). Also, *Robbsia* sequences from nursery 2 in 2016 and from both nurseries in 2018 may have originated from closely related but non-pathogenic *Robbsia* strains. Shotgun metagenomic sequencing could be used in the future to avoid this uncertainty. Shotgun metagenomic sequencing has the advantage of higher taxonomic resolution compared to 16S rRNA sequencing because sequencing is not limited to the conserved 16S rRNA gene but extends to entire genome sequences (Ranjan et al. 2016). The power of this approach was recently demonstrated in human disease diagnostics (Huang et al. 2017) whereby shotgun metagenomics was able to reveal that two food-borne disease outbreaks that were suspected to be epidemiologically linked were actually caused by two distinct strains of *Salmonella enterica* subsp. *enterica* serovar Heidelberg. Moreover, since fungal and oomycete DNA sequences are obtained in shotgun metagenomic sequencing as well, this approach has the advantage that many potential pathogens can be detected in a single diagnostic step. On the other hand, shotgun metagenomic sequencing has the disadvantage that plant DNA can comprise the largest fraction of a sample. Therefore, a very high number of sequencing reads must be obtained in order to make sound conclusions. In plant viral disease diagnostics, shotgun metagenomic sequencing, with the modification of a reverse-transcription prior to sequencing for viral genomic RNA has already been used (Massart et al. 2014). Recently, the Oxford Nanopore Technologies MinION

sequencer has been tested in plant pathogen identification (Bronzato Badial et al. 2018; Chalupowicz et al. 2018). The resolution that can be obtained with the MinION is already at species level, and strain-level identification may be possible to achieve after further development of respective software and databases.

Also important to consider is the fact that even though the frequency of the *Robbsia* OTU in the microbiome on the symptomatic plants was not significantly higher compared to all non-symptomatic plants, the absolute population size of *R. andropogonis* may have been higher on these plants. In fact, culture-independent 16S rRNA sequencing does not provide information about the absolute number of any identified bacterial taxon. It only provides information on relative abundance of a taxon compared to all other taxa. Since we were unaware of the identity of the pathogen at the time of the initial sequencing experiments, we were not able to determine the population size of *R. andropogonis* using culturing.

Although the results of our study refuted our initial hypothesis that symptomatic plants contained the causal agent of the observed disease symptoms and the asymptomatic plants did not, interesting results were obtained nonetheless. For example, the tendentially lower taxonomic diversity (alpha diversity) of the microbiomes associated with the plants at nursery 1 compared to nursery 2 is intriguing. This difference in the level of taxonomic diversity (alpha diversity) between the two nurseries may have been caused by different management practices at the two nurseries and/or different environmental conditions due to the different geographic locations of the two nurseries. In fact, fertilizer application was shown to have a significant effect on the ability of microbiomes to suppress disease (Berg and Koskella 2018) and environmental conditions have been found to be an important factor in determining plant-associated microbiomes (Peiffer et al. 2013; Campisano et al. 2017).

In regard to differences in the taxonomic composition between samples (beta diversity), this study suffered from the small number of available symptomatic samples due to the limited

duration of the 2015 disease outbreak, which did not allow for repeated sampling. Therefore, a thorough comparison between symptomatic and asymptomatic samples was not possible. Also, the number of samples taken from each of the available cultivars was too small to determine an effect of genotype on microbiome composition. However, we had enough samples to compare the taxonomic composition of microbiomes from nursery 1 with the taxonomic composition of microbiomes from nursery 2 and between the samples taken in different years. Interestingly, we found year of sampling to have a significant effect on microbiome composition, but nursery of origin did not. This may be due to a strong effect of annual fluctuations in regard to temperature, precipitation, and humidity, but also of season since 2015 samples were taken in fall, 2016 samples were taken in summer, and 2018 samples were taken in spring.

Another important result of our study was the observation that artificially inoculated cheddar pink plants developed disease symptoms similar to the plants affected by the disease outbreak at the nursery but artificially inoculated plants had a very different microbiome composition compared to the naturally infected symptomatic nursery plants. In symptomatic nursery plants, the percentage of *R. andropogonis* only constituted between 0.99 and 29.71% of the total microbial population while in artificially inoculated plants, the percentage of *R. andropogonis* grew to over 90% in our first experiment and to approximately 60-70% in a repeat experiment two years later. Such a dramatic difference in microbiome composition between plants inoculated in a lab compared to plants naturally infected in the field can be expected to have a significant impact on the plant defense response. Therefore, if this result were to be confirmed in additional plant-pathogen systems, it would need to be taken into consideration when trying to extrapolate molecular results obtained in laboratory experiments to plants naturally infected in the field.



In conclusion, based on the results presented here, culture-independent microbiome analysis based on 16S rRNA amplicon sequencing using Illumina technology cannot replace culture-dependent pathogen identification for diagnosis of bacterial diseases on cheddar pink at this time. However, our study has revealed that *R. andropogonis* is ubiquitously present on cheddar pink nursery plants in Virginia and has provided first insights into the microbiome composition of healthy and diseased cheddar pink nursery plants and into the changes in microbiome composition that occur during disease development. Extending our research to additional pathosystems and from 16S rRNA sequencing to shotgun metagenomic sequencing, as recently reported (Chalupowicz et al. 2018), will be necessary to fully test the potential of next generation sequencing in support of routine plant disease diagnostics.

## **5. Funding:**

This work was supported by the Virginia Agricultural Experiment Station and the Hatch Program of the National Institute of Food and Agriculture, US Department of Agriculture, the College of Agriculture and Life Sciences, Virginia Tech, and by the National Science Foundation, grant IOS-1354215. Funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

## **6. Compliance with Ethical Standards:**

### *6.1. Conflict of Interest:*

The authors declare that they have no conflict of interest.

## **7. References:**

Adams, I. P., Miano, D. W., Kinyua, Z. M., Wangai, A., Kimani, E., Phiri, N., et al. (2013).

Use of next-generation sequencing for the identification and characterization of Maize

- chlorotic mottle virus and Sugarcane mosaic virus causing maize lethal necrosis in Kenya. *Plant Pathology*, 62(4), 741-749, doi:10.1111/j.1365-3059.2012.02690.x.
- Berg, M., & Koskella, B. (2018). Nutrient- and Dose-Dependent Microbiome-Mediated Protection against a Plant Pathogen. *Current Biology*, 28(15), 2487-2492.e2483, doi:10.1016/j.cub.2018.05.085.
- Boonham, N., Kreuze, J., Winter, S., van der Vlugt, R., Bergervoet, J., Tomlinson, J., et al. (2014). Methods in virus diagnostics: from ELISA to next generation sequencing. *Virus Research*, 186, 20-31, doi:10.1016/j.virusres.2013.12.007.
- Boxriker, M., Boehm, R., Krezdorn, N., Rotter, B., & Piepho, H.-P. (2017). Comparative transcriptome analysis of vase life and carnation type in *Dianthus caryophyllus* L. *Scientia Horticulturae*, 217(Supplement C), 61-72, doi:<https://doi.org/10.1016/j.scienta.2017.01.015>.
- Bronzato Badial, A., Sherman, D., Stone, A., Gopakumar, A., Wilson, V., Schneider, W., et al. (2018). Nanopore Sequencing as a Surveillance Tool for Plant Pathogens in Plant and Insect Tissues. *Plant Disease*, 102(8), 1648-1652, doi:10.1094/PDIS-04-17-0488-RE.
- Campisano, A., Albanese, D., Yousaf, S., Pancher, M., Donati, C., & Pertot, I. (2017). Temperature drives the assembly of endophytic communities' seasonal succession. *Environ Microbiol*, 19(8), 3353-3364, doi:10.1111/1462-2920.13843.
- Caporaso, J. G., Kuczynski, J., Stombaugh, J., Bittinger, K., Bushman, F. D., Costello, E. K., et al. (2010). QIIME allows analysis of high-throughput community sequencing data. *Nat Methods*, 7(5), 335-336, doi:10.1038/nmeth.f.303.
- Casanova, E., Valdes, A. E., Fernandez, B., Moysset, L., & Trillas, M. I. (2004). Levels and immunolocalization of endogenous cytokinins in thidiazuron-induced shoot organogenesis in carnation. *J Plant Physiol*, 161(1), 95-104.

- Chalupowicz, L., Dombrovsky, A., Gaba, V., Luria, N., Reuven, M., Beerman, A., et al. (2018).  
Diagnosis of plant diseases using the Nanopore sequencing platform. *Plant Pathology*.
- Clarke, C. R., Studholme, D. J., Hayes, B., Runde, B., Weisberg, A., Cai, R., et al. (2015).  
Genome-enabled phylogeographic investigation of the quarantine Pathogen *Ralstonia solanacearum* Race 3 Biovar 2 and screening for sources of resistance against its core effectors. *Phytopathology*, *105*(5), 597-607, doi:10.1094/PHYTO-12-14-0373-R.
- Fang, Y., & Ramasamy, R. P. (2015). Current and prospective methods for plant disease detection. *Biosensors (Basel)*, *5*(3), 537-561, doi:10.3390/bios5030537.
- Flight, W. G., Smith, A., Paisey, C., Marchesi, J. R., Bull, M. J., Norville, P. J., et al. (2015).  
Rapid detection of emerging pathogens and loss of microbial diversity associated with severe lung disease in Cystic Fibrosis. *J Clin Microbiol*, *53*(7), 2022-2029, doi:10.1128/jcm.00432-15.
- Hill, T. C., Walsh, K. A., Harris, J. A., & Moffett, B. F. (2003). Using ecological diversity measures with bacterial communities. *FEMS Microbiol Ecol*, *43*(1), 1-11, doi:10.1111/j.1574-6941.2003.tb01040.x.
- Huang, A. D., Luo, C., Pena-Gonzalez, A., Weigand, M. R., Tarr, C. L., & Konstantinidis, K. T. (2017). Metagenomics of two severe foodborne outbreaks provides diagnostic signatures and signs of coinfection not attainable by traditional methods. *Applied and Environmental Microbiology*, *83*(3), e02577-02516, doi:10.1128/AEM.02577-16.
- Ivy, M. I., Thoendel, M. J., Jeraldo, P. R., Greenwood-Quaintance, K. E., Hanssen, A. D., Abdel, M. P., et al. (2018). Direct Detection and Identification of Prosthetic Joint Infection Pathogens in Synovial Fluid by Metagenomic Shotgun Sequencing. *Journal of Clinical Microbiology*, *56*(9), e00402-00418, doi:10.1128/jcm.00402-18.
- Khater, M., de la Escosura-Muniz, A., & Merkoci, A. (2017). Biosensors for plant pathogen detection. *Biosens Bioelectron*, *93*, 72-86, doi:10.1016/j.bios.2016.09.091.

- Lopes-Santos, L., Castro, D. B. A., Ferreira-Tonin, M., Correa, D. B. A., Weir, B. S., Park, D., et al. (2017). Reassessment of the taxonomic position of *Burkholderia andropogonis* and description of *Robbsia andropogonis* gen. nov., comb. nov. *Antonie Van Leeuwenhoek*, *110*(6), 727-736, doi:10.1007/s10482-017-0842-6.
- Lozupone, C., & Knight, R. (2005). UniFrac: a new phylogenetic method for comparing microbial communities. *Appl Environ Microbiol*, *71*(12), 8228-8235, doi:10.1128/aem.71.12.8228-8235.2005.
- Massart, S., Olmos, A., Jijakli, H., & Candresse, T. (2014). Current impact and future directions of high throughput sequencing in plant virus diagnostics. *Virus Research*, *188*, 90-96, doi:10.1016/j.virusres.2014.03.029.
- McMurdie, P. J., & Holmes, S. (2013). phyloseq: An R Package for Reproducible Interactive Analysis and Graphics of Microbiome Census Data. *PLOS ONE*, *8*(4), e61217, doi:10.1371/journal.pone.0061217.
- Peiffer, J. A., Spor, A., Koren, O., Jin, Z., Tringe, S. G., Dangl, J. L., et al. (2013). Diversity and heritability of the maize rhizosphere microbiome under field conditions. *Proc Natl Acad Sci U S A*, *110*(16), 6548-6553, doi:10.1073/pnas.1302837110.
- Quast, C., Pruesse, E., Yilmaz, P., Gerken, J., Schweer, T., Yarza, P., et al. (2013). The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic Acids Res*, *41*(Database issue), D590-596, doi:10.1093/nar/gks1219.
- Ranjan, R., Rani, A., Metwally, A., McGee, H. S., & Perkins, D. L. (2016). Analysis of the microbiome: Advantages of whole genome shotgun versus 16S amplicon sequencing. *Biochem Biophys Res Commun*, *469*(4), 967-977, doi:10.1016/j.bbrc.2015.12.083.
- Raszek, M. M., Guan le, L., & Plastow, G. S. (2016). Use of genomic tools to improve cattle health in the context of infectious diseases. *Frontiers in Genetics*, *7*, 30, doi:10.3389/fgene.2016.00030.

- Reeves, E., Hansen, M. A., & Bush, E. (2017). First report of bacterial leaf spot of a Hardy Pink (*Dianthus gratianopolitanus* hybrid) caused by *Burkholderia andropogonis* in Virginia. *Plant Disease*, *101*(8), 1540-1540, doi:10.1094/PDIS-02-17-0226-PDN.
- Riley, M., Williamson, M., & Maloy, O. (2002). Plant disease diagnosis.
- Schubert, T., Jeyaprakash, A., & Harmon, C. (2018). Fundamentals and Advances in Plant Problem Diagnostics. In R. J. McGovern, & W. H. Elmer (Eds.), *Handbook of Florists' Crops Diseases* (pp. 13-39). Cham: Springer International Publishing.
- Sergaki, C., Lagunas, B., Lidbury, I., Gifford, M. L., & Schäfer, P. (2018). Challenges and Approaches in Microbiome Research: From Fundamental to Applied. *Frontiers in plant science*, *9*, 1205-1205, doi:10.3389/fpls.2018.01205.
- Shade, A., McManus, P. S., & Handelsman, J. (2013). Unexpected diversity during community succession in the apple flower microbiome. *MBio*, *4*(2), doi:10.1128/mBio.00602-12.
- Thorburn, F., Bennett, S., Modha, S., Murdoch, D., Gunson, R., & Murcia, P. R. (2015). The use of next generation sequencing in the diagnosis and typing of respiratory infections. *J Clin Virol*, *69*, 96-100, doi:10.1016/j.jcv.2015.06.082.
- Wickham, H. (2009). *Ggplot2: Elegant graphics for data analysis*: Springer.
- Yagi, M., Kimura, T., Yamamoto, T., Isobe, S., Tabata, S., & Onozaki, T. (2012). QTL analysis for resistance to bacterial wilt (*Burkholderia caryophylli*) in carnation (*Dianthus caryophyllus*) using an SSR-based genetic linkage map. [journal article]. *Molecular Breeding*, *30*(1), 495-509, doi:10.1007/s11032-011-9639-x.

## 8. Tables:

**Table 1.** Plant samples analyzed in this study:

<b>Plant cultivar (sample name<sup>1</sup>)</b>	<b>Nursery</b>	<b>Collection year</b>	<b>Symptomatic</b>	<b>Days post Inoculation</b>
Everlast Lavender + Eye (n1.15.ns.LE)	n1	2015	ns	NA
Everlast Lavender + Eye (n1.15.s.LE)	n1	2015	s	NA
Scent First ‘Passion’ (n1.15.ns.SF)	n1	2015	ns	NA
Scent First ‘Passion’ (n1.15.s.SF)	n1	2015	s	NA
Scent First ‘Passion’ (n2.16.ns.SF1)	n2	2016	ns	NA
Scent First ‘Passion’ (n2.16.ns.SF2)	n2	2016	ns	NA
Scent First ‘Passion’ (n2.16.ns.SF3)	n2	2016	ns	NA
Scent First ‘Passion’ (n1.18.ns.SF1)	n1	2018	ns	NA
Scent First ‘Passion’ (n1.18.ns.SF2)	n1	2018	ns	NA
Scent First ‘Passion’ (n1.18.ns.SF3)	n1	2018	ns	NA
Scent First ‘Passion’ (n1.18.ns.SF4)	n1	2018	ns	NA
Neon Star (n2.18.ns.NS1)	n2	2018	ns	NA
Neon Star (n2.18.ns.NS2)	n2	2018	ns	NA
Firewitch (n2.18.ns.FW1)	n2	2018	ns	NA

Firewitch (n2.18.ns.FW2)	n2	2018	ns	NA
Scent First 'Passion' (n2.16.ns.SF.11dpi)	n2	2016	ns	11 dpi
Scent First 'Passion' (n2.16.ns.SF.17dpi)	n2	2016	s	17 dpi
Scent First 'Passion' (n2.16.ns.SF1.24dpi)	n2	2016	s	24 dpi
Scent First 'Passion' (n2.16.ns.SF2.24dpi)	n2	2016	s	24 dpi <sup>2</sup>
Scent First 'Passion' (n1.18.ns.SF.11dpi)	n1	2018	s	11 dpi
Scent First 'Passion' (n1.18.ns.SF.17dpi)	n1	2018	s	17 dpi
Scent First 'Passion' (n1.18.ns.SF1.24dpi)	n1	2018	s	24 dpi
Scent First 'Passion' (n1.18.ns.SF2.24dpi)	n1	2018	s	24 dpi <sup>2</sup>

<sup>1</sup> Sample names are based on abbreviations for the nursery of origin (nursery 1 = n1, nursery 2 = n2), the year of sampling, presence of symptoms (ns = no symptoms, s = symptoms), and the name of the plant cultivar (Lavender + Eye = LE, Scent First 'Passion' = SF, Neon Star = NS, Firewitch = FW; if more than one sample was analyzed, the abbreviation of the cultivar name is followed by a number). For inoculated plants, the days post inoculation (dpi) at which the sample was taken is also indicated. Sample names of uninoculated control plants end in neg (negative).

<sup>2</sup> Sample was inoculated with  $8.8 \times 10^9$  cfu/mL while other samples were inoculated with  $8.8 \times 10^8$  cfu/mL

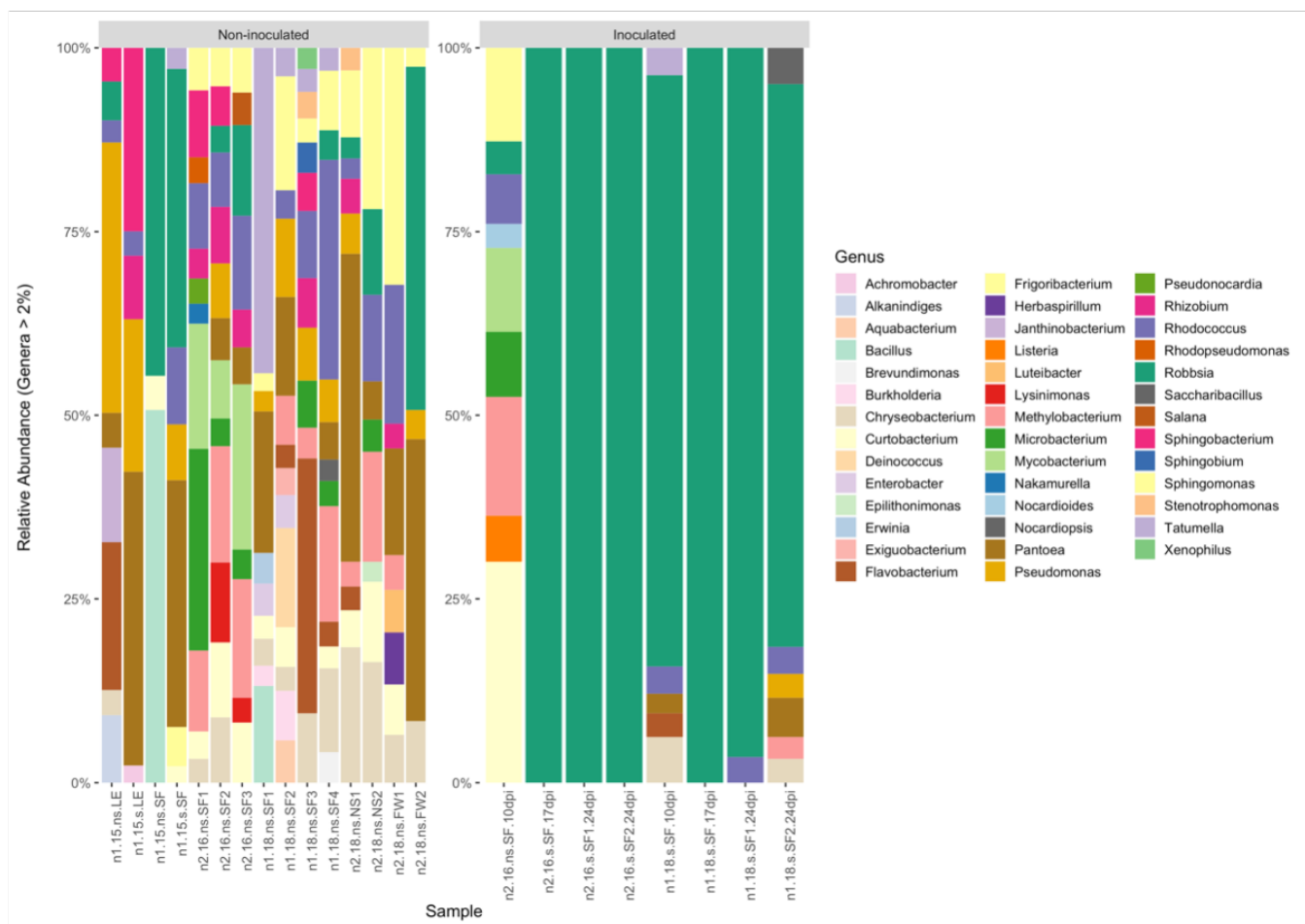
**Table 2.** Diversity indices computed for each individual nursery plant sample:

	OTUs	Shannon	Simpson
n1.15.ns.LE	755	6.71	0.97
n1.15.s.LE	481	5.12	0.91
n1.15.ns.SF	526	3.42	0.72
n1.15.s.SF	429	4.64	0.87
n2.16.ns.SF1	913	7.32	0.98
n2.16.ns.SF2	908	7.26	0.98
n2.16.ns.SF3	805	7.36	0.98
n1.18.ns.SF1	742	5.60	0.92
n1.18.ns.SF2	167	7.94	0.99
n1.18.ns.SF3	1076	7.41	0.97
n1.18.ns.SF4	1245	7.91	0.98
n2.18.ns.FW1	920	7.07	0.98
n2.18.ns.FW2	576	4.79	0.85
n2.18.ns.NS1	882	6.50	0.95
n2.18.ns.NS2	1039	7.00	0.97
<i>Inoculated plants</i>			
n2.16.ns.SF.10dpi	1021	7.36	0.97
n2.16.s.SF.17dpi	84	0.63	0.14
n2.16.s.SF1.24dpi	68	0.55	0.12
n2.16.s.SF2.24dpi	99	0.66	0.14
n1.18.s.SF.10dpi	731	4.06	0.63
n1.18.s.SF.17dpi	492	2.51	0.40
n1.18.s.SF1.24dpi	516	2.50	0.40
n1.18.s.SF2.24dpi	666	4.03	0.64

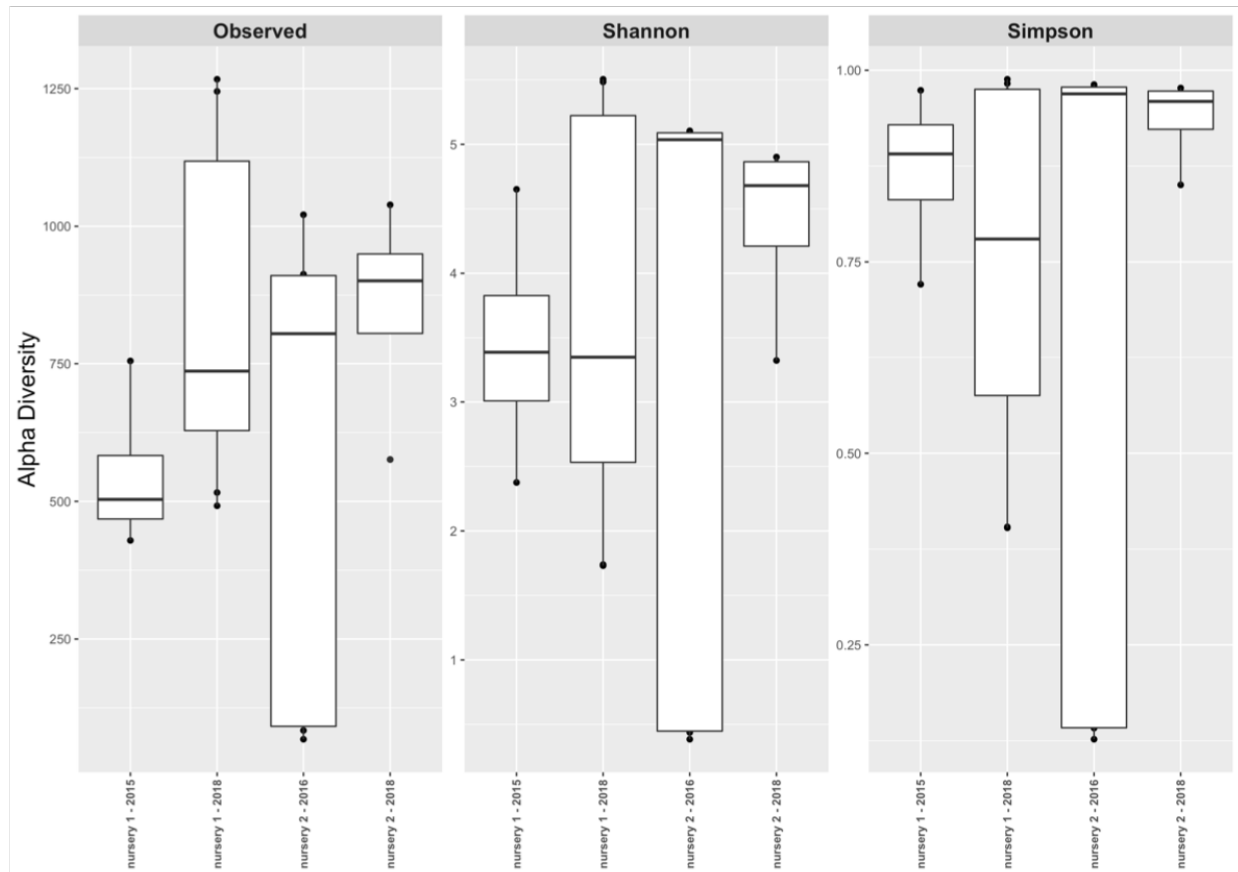


## 9. Figures:

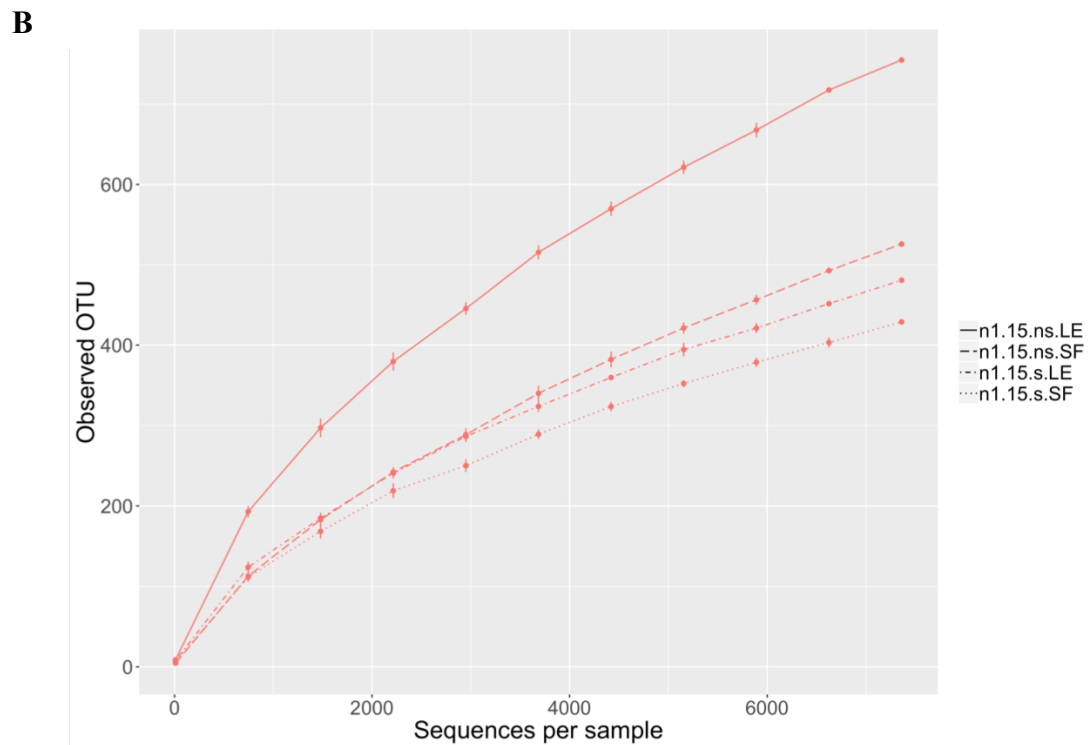
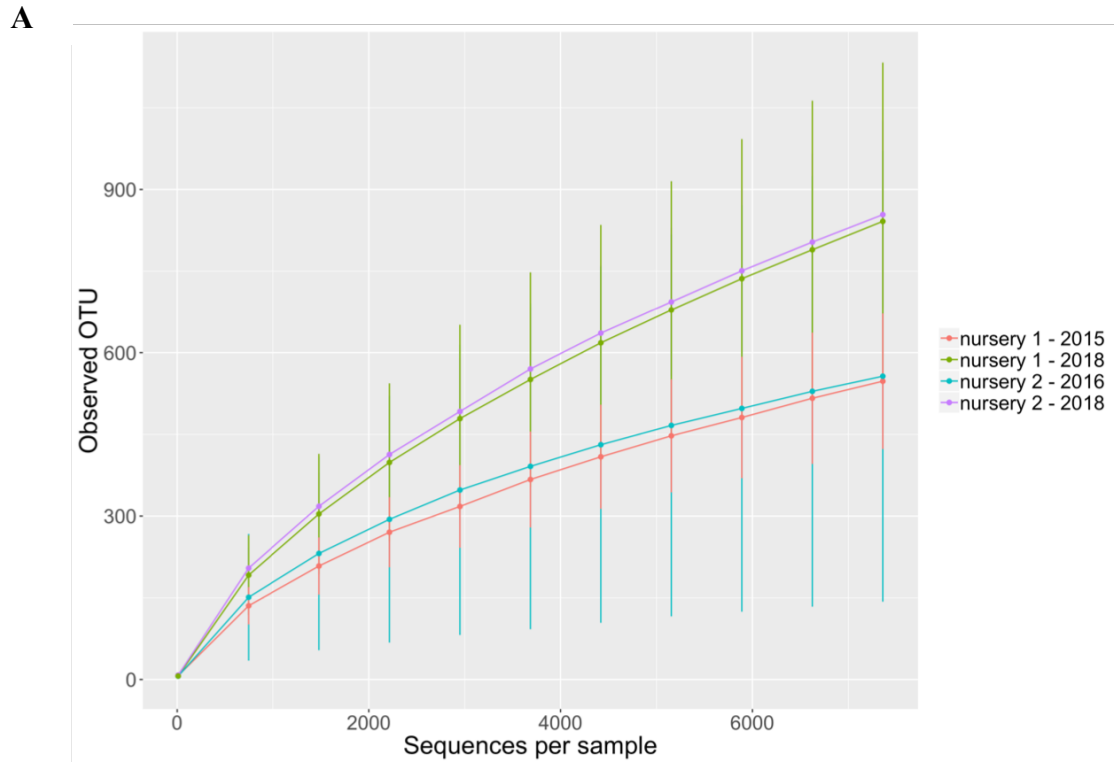
**Fig. 1 Relative abundance of bacterial taxa.** Relative abundance of bacterial taxa at the genus level is shown for symptomatic and asymptomatic cheddar pink plants from nursery 1 during the disease outbreaks in 2015, plants from nursery 2 without a disease outbreak in 2016, and plants from both nurseries in the absence of disease in 2018. Results are also shown for plants after artificial inoculation with *R. andropogonis*. Note that the genus *Robbsia* is present in all samples but below the employed 2% cutoff in some of the samples. Samples are labeled as listed in Table 1.



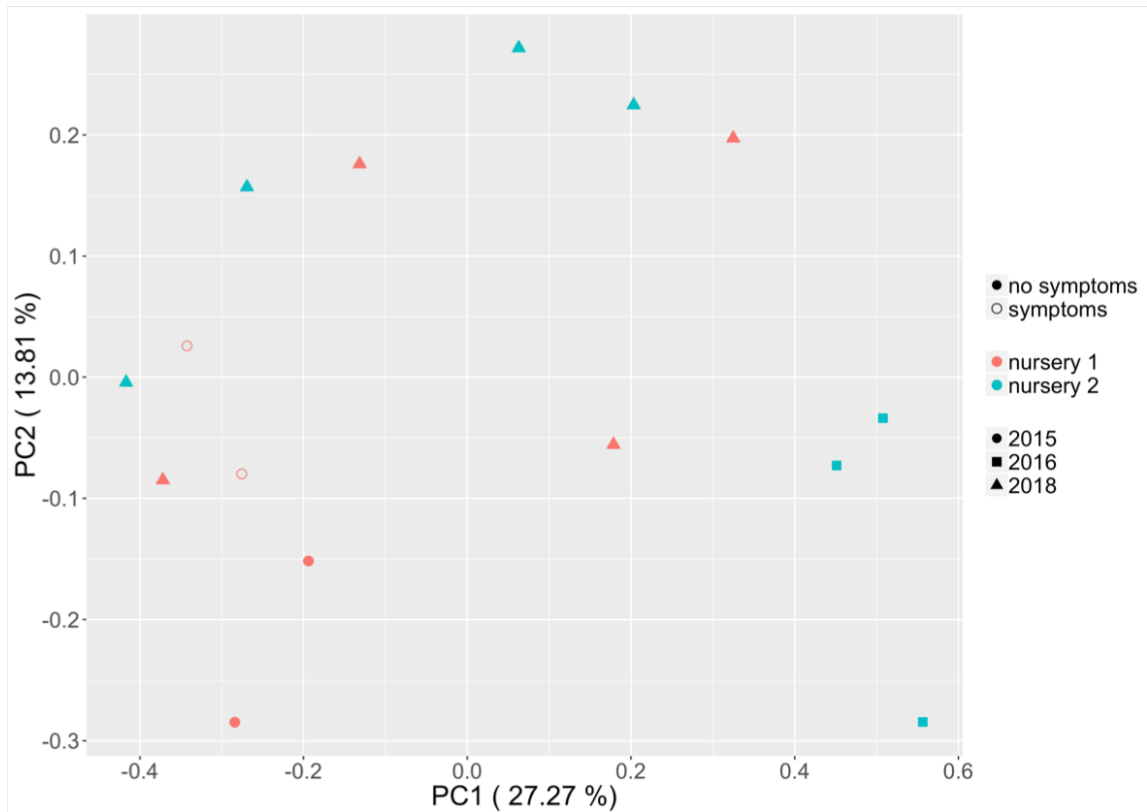
**Fig. 2 Alpha diversity.** Three measures of alpha diversity (observed OTUs, Shannon diversity index, and Simpson diversity index) for nursery 1 during the disease outbreak in 2015, nursery 2 in 2016 (no disease), and both nurseries in 2018 (no disease).



**Fig. 3 Rarefaction curves.** **A** Rarefaction curves for nursery 1 during the disease outbreak in 2015, nursery 2 in 2016 (no disease), and both nurseries in 2018 (no disease). **B** rarefaction curves for each individual plant at nursery 1 during the disease outbreak in 2015.



**Fig. 4 Principal Coordinates Analysis.** Principal Coordinates Analysis based on weighted Unifrac distances is shown for all samples from nursery 1 during the disease outbreak in 2015, nursery 2 in 2016 in the absence of disease, and both nurseries in 2018 in the absence of disease.





**Figure S2. *Dianthus gratianopolitanus* plant with leaf spotting symptoms after artificial inoculation.**

