Unveiling the Dynamics of Plant-Microbe Interactions: From Phyllosphere to Rhizosphere and Beyond

by

Rishi Ram Bhandari

A dissertation submitted to the Graduate Faculty of
Auburn University
in partial fulfillment of the
requirements for the Degree of
Doctor of Philosophy

Auburn, Alabama December 09, 2023

Keywords: *Xanthomonas*, phyllosphere, soil health, microbiome, diversity, metagenomics

Copyright 2023 by Rishi Ram Bhandari

Approved by

Neha Potnis, Chair, Endowed Associate Professor of Entomology and Plant Pathology
Edward J. Sikora, Extension Specialist Professor of Entomology and Plant Pathology
Karen A. Garrett, Professor of Plant Pathology, University of Florida
Leonardo De La Fuente, Professor of Entomology and Plant Pathology
Yucheng Feng, Professor of Crop, Soil and Environmental Sciences

Abstract

Bacterial leaf spot (BLS) is a recurring agricultural issue affecting tomatoes and peppers around the globe. Traditionally, *Xanthomonas perforans* was considered the primary pathogen of tomatoes, while *X. euvesicatoria* was associated with peppers. However, recent studies have indicated a notable shift towards the dominance of *X. perforans* in pepper plants, signifying a potential expansion of its host range. Our research sought to delve into the diversity of the endemic bacterial spot pathogen *Xanthomonas* and uncover the factors driving microbial diversity and pathogen populations. Through a culture-independent approach, we achieved a higher-resolution method for examining pathogen populations and survey of tomato fields indicated that all eight lineages of *X. perforans* found in the samples collected around the globe are also circulating throughout southeastern United States. Co-occurrence of multiple lineages was common among the fields. Furthermore, we employed modeling to analyze *Xanthomonas* populations and disease severity alongside climate variables, emphasizing the critical role of meteorological conditions in shaping disease outcomes. This knowledge is paramount for developing precise predictive models and early warning systems to mitigate disease outbreaks. In addition to studying pathogenic strains, our research delved into the diversity and evolution of nonpathogenic *Xanthomonas* strains, often found alongside their pathogenic counterparts in the phyllosphere. This investigation focused on co-occurrence patterns and phylogenetic relationships to identify genomic traits that underlie their ecological strategies, spanning from commensal to weakly pathogenic to fully pathogenic lifestyles. Our results suggested that the distinction between these lifestyles in *Xanthomonas* is not solely defined by the type III secretion system and effectors. We also identified distinct sets of cell-wall degrading enzymes that differentiate pathogenic from commensal or weakly pathogenic lifestyles. In contrast, pathogens

rely on the type III secretion system and effectors to evade host defense responses, whereas commensal *Xanthomonas* harbor genes that promote stress tolerance rather than avoidance, especially in the absence of the type III secretion system.

The intricate relationships between plants and their associated microbiota, spanning bacteria, fungi, viruses, and protists, have evolved to form the plant microbiota over millions of years. Within this diverse community, only a subset of microbes act as pathogens, impacting specific hosts. These plant-associated microbes can be found in various niches, including the rhizosphere, phyllosphere, or endosphere, and play essential roles in nutrient acquisition, adaptation to stressors, and overall plant growth. Comprehensive comprehension of these complex plant-microbe interactions is vital for the effective management of plant diseases and the stability of ecosystems. For example, the phyllosphere microbiome, comprising microorganisms residing on the aboveground parts of plants, significantly influences plant health, productivity, and resilience to various biotic and abiotic stressors. Unlike the relatively stable rhizosphere, the phyllosphere represents a dynamic environment characterized by rapid environmental fluctuations, including temperature, humidity, UV light, and limited nutrient availability. In a world characterized by global changes such as shifts in climate and land use, these fluctuations significantly impact ecosystems and plant-microbe interactions. To shed light on these influences, we examined how elevated tropospheric ozone (O_3) and *Xanthomonas perforans* infection impact disease outcomes and associated microbiomes in pepper plants. While pathogen infection significantly influenced the microbiome of susceptible cultivars, O_3 stress exacerbated disease severity in resistant cultivars. This alteration in microbial community interactions in both biotic and abiotic stress suggests that microbiomes play a pivotal role in plant-pathogen responses under climate change.

Besides phyllosphere, we also utilized a culture-independent technique to scrutinize the influence of long-term crop management and fertility on soil microbial communities. Our study involved the analysis of nine cropping systems, each employing various fertilization methods and legume cover crops. Our results indicated that long-term balanced nitrogen (N) addition significantly influences fungal communities but has a lesser impact on bacterial communities. Lower soil pH was found to significantly affect bacterial communities, while fungal communities exhibited greater resilience to changes in pH levels. While applying chemical fertilizers has previously been associated with reduced microbial diversity and richness, our research showed relative stability in soil bacterial diversity and richness under standard fertilizer treatment. This stability implies that microbial communities can adapt to prolonged fertilizer use.

Overall, our research provided valuable insights into the diversity, evolution, and ecology of BLS *Xanthomonas* strains and the importance of plant-microbe interactions in plant disease management and adaptation to climate change. These findings contribute to developing sustainable agricultural practices that enhance plant health, productivity, and resilience in the face of evolving pathogens and changing climates.

Acknowledgements

I would like to express my heartfelt gratitude to my major advisor, Dr. Neha Potnis, for providing me with the wonderful opportunity to work on this engaging project and for her invaluable guidance. I am deeply thankful for her continuous encouragement, motivation, patience, and insightful suggestions throughout my research and scientific writing.

I appreciate my committee members, Dr. Karen A. Garrett, Dr. Edward J Sikora, Dr. Leonardo De La Fuente, and Dr. Yucheng Feng, for their guidance, valuable suggestions, and willingness to assist me in my research and thesis writing. Thanks to the external university reader, Dr. Lori Eckhardt, for her constructive feedback.

My sincere gratitude goes to my close friends Anand Tiwari, Gautam Joshi, Nabin Bhandari, and all my friends in Auburn. A special acknowledgment to the members of the Nepalese Students Association for their unwavering support and camaraderie throughout my journey. I also want to thank my current and former lab mates from Potnis Lab Dr. Sivakumar Ramamoorthy, Palash Ghosh, Amanpreet Kaur, Ivory Russel, Bijay Subedi, Kylie Weis, Auston Holland, Dr. Prabha Liyanapathiranage, Shreya Sadhukhan, Destiny Brokaw, David Hill, Dr. Michelle Pena and Dr. Eric Newberry for their suggestions, assistance, and encouragement during my research.

I convey my deep love and profound gratitude by dedicating this dissertation to my late mom, Nandakali Bhandari, who consistently provided unwavering support and encouragement. Mom, your absence is deeply felt today. I am immensely thankful for the steadfast support and countless sacrifices of my father, Dandapani Bhandari, who played a pivotal role in helping me complete my Ph.D.

To my dearest family members, especially my brothers Liladhar and Mohan and my sister Laxmi Bhandari, your unwavering support through thick and thin has been invaluable. I owe a significant part of my success to each of you. Though not explicitly named, every member of my extensive and joyous family has played a role, and I appreciate your love and support.

A special acknowledgment to my wife, Ambika Pokhrel, for her unconditional patience and support throughout my journey. This success is as much yours as it is mine. Your encouragement and dedication have been the driving force behind my achievements, and I cannot thank you enough for being my pillar of strength throughout this challenging journey.

To everyone who has been a part of this journey, thank you from the bottom of my heart. I am infinitely grateful for everything!

"शिक्षायां नैव सन्तुष्टं, ज्ञानं प्राप्नोति नित्यदा।
अन्ते चास्ति न योग्यता, शिक्षणं सर्वदा तु तत्॥"

Table of contents

Abstract ii
Acknowledgements v
Table of contents vii
List of Tables x
List of Figures xii
1. CHAPTER ONE: Introduction and literature review 1
1.1 References for chapter 1 6
2. CHAPTER TWO:* Xanthomonas *infection and ozone stress distinctly influence the microbial community structure and interactions in the pepper phyllosphere 10
Abstract 10
2.1 Introduction 11
2.2 Materials and Methods 15
2.2.1 Experimental site and design 15
2.2.2 Disease severity measurements 17
2.2.3 Sample collection, DNA extraction, sequencing, and quality trimming 17
2.2.4 Taxonomic profiling 18
2.2.5 Culture-dependent method for determining the *Xanthomonas* population 19
2.2.6 Diversity, statistical analysis, and network analysis 20
2.2.7 Functional profiling 21
2.3 Results 22
2.3.1 Influence of O_3 levels on disease severity on resistant and susceptible cultivars 22
2.3.2 Sequencing statistics 22
2.3.3 Microbial diversity and richness are reduced under susceptible response, but are not significantly affected by elevated O_3. 23
2.3.4 The effect of O_3 levels was significant on the eukaryotic community, yet was minimal in shaping bacterial community structure. 24
2.3.5 Influence of pathogen infection and O_3 stress on relative and absolute abundance of microbial taxa 26
2.3.6 Functional composition of phyllosphere communities when exposed to O_3 stress and pathogen infection 29
2.3.7 Microbial network topology is altered under combined pathogen and O_3 stress. 30
2.4 Discussion 32
2.5 Data and code availability 38
2.6 Acknowledgments 38
2.7 References for chapter 2 39
2.8 Supplementary results 54
2.8.1 Taxonomic profiling 54
3. CHAPTER THREE: Long-term fertilization and crop management affects soil microbial communities. 84

Abstract **84**

3.1 Introduction **85**

3.2 Materials and Methods **87**

3.2.1 Study site and sampling 87
3.2.2 Soil DNA extraction and sequencing 88
3.2.3 Sequence data processing 89
3.2.4 Microbial community analyses 90
3.2.5 Biomarker and network analysis 91

3.3 Results **92**

3.3.1 Microbial diversity in response to fertilization 92
3.3.2 Microbial diversity under different fertilizer treatments 93
3.3.3 Influence of soil chemistry parameters on microbial community composition 94
3.3.4 Bacterial community composition among different treatments 95
3.3.5 Effect of fertilizer treatment on soil microbial biomarkers 96

3.4 Discussion **98**

3.5 References for chapter 3 ..... **107**

4. CHAPTER FOUR: Genetic and functional diversity help explain pathogenic, weakly pathogenic, and commensal lifestyles in the genus *Xanthomonas* **132**

Abstract **132**

4.1 Introduction **133**

4.2 Materials and Methods **138**

4.2.1 Bacterial strains collection and genome sequencing 138
4.2.2 Genome-based identification of *Xanthomonas* strains 138
4.2.3 Analysis of the gain and loss dynamics of the T3SS clusters 139
4.2.4 Prediction of Mobile Genetic Elements 141
4.2.5 Comparison of secreted carbohydrate-active enzymes 141
4.2.6 Pathogenicity assays and co-inoculation experiments 142
4.2.7 Identification of lifestyle-associated genes 143

4.3 Results **145**

4.3.1 General features of *Xanthomonas* strains from this study, potential new species, and core genome phylogeny 145
4.3.2 Distribution of T3SS clusters across the phylogeny 147
4.3.3 T3SS was gained and lost multiple times in the genus *Xanthomonas* 148
4.3.4 T3E repertoire ranges from zero to forty-one in crop-associated and environmental strains 149
4.3.5 Genetic exchange between commensal, weakly pathogenic, and pathogenic *Xanthomonas* strains. 152
4.3.6 Lifestyle had a significant effect in determining the repertoires of cell wall-degrading enzymes in *Xanthomonas* 155
4.3.7 Association analysis to identify genes and features that define the lifestyle of *Xanthomonas* species 155

4.4 Discussion **160**

4.5 Acknowledgments **167**

4.6 Data availability **167**

4.7 References for chapter 4 ..... **168**

5. CHAPTER FIVE: Bacterial spot of tomato: strain diversity, dynamics, and environmental drivers in the southeastern United States **223**

Abstract **223**

5.1 Introduction **224**

5.2 Materials and Methods **228**

5.2.1 Reconstruction of the *X. perforans* core genome 228

5.2.2 Sample collection and disease severity estimation 229

5.2.3 DNA extraction and sequencing 229

5.2.4 Taxonomic profiling and diversity estimation 230

5.2.5 Weather data 231

5.2.6 Statistical analysis 232

5.3 Results **234**

5.3.1 Diversity of bacterial leaf spot pathogen *Xp* around the world 234

5.3.2 Abundance of *Xp* is positively associated with BLS disease severity 235

5.3.3 *Xanthomonas* abundance reveals co-occurrence of various *Xp* lineages 235

5.3.4 *Xp* population diversity is positively associated with disease severity 237

5.3.5 Variation in *Xp* lineages during growing season is influenced by various climatic factors 238

5.3.6 Photosynthetically active radiation and wind speed significantly influence BLS disease severity 238

5.3.7 Frequent shifts in wind direction and extreme changes in surface pressure and relative humidity influences *Xp* abundance 239

5.4 Discussion **240**

5.5 Acknowledgments **245**

5.6 References for chapter 5 ..... **246**

List of Tables

Table 5-1A: Results of the ordinal regression model showing parameter estimates and associated statistics for disease severity. 258
Table 5-2: Results of the beta regression model showing parameter estimates and associated statistics for relative abundance of *Xanthomonas perforans* (Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1). 258
Table 5-3: Results of the beta regression model showing parameter estimates and associated statistics for absolute abundance of Xanthomonas perforans (Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1). 259

List of Supplementary Tables

Table S2-1: Multiple pairwise comparisons (Dunn test) of disease severity index across different cultivars and treatment conditions. Significance levels for each treatment combination are indicated by *p < 0.05; **p < 0.01; ***p < 0.001; ****p < 0.0001. 62
Table S2-2: Sample details and sequencing statistics of metagenome reads. 63
Table S2-3A: Pairwise comparison for bacterial Chao1 diversity index for resistant pepper. Significance levels for each treatment combination are indicated by *p < 0.05; **p < 0.01; ***p < 0.001. 64
Table S 2-4A: Pairwise comparison of Eukaryotic Chao1 diversity index for resistant pepper. Significance levels for each treatment combination are indicated by *p < 0.05; **p < 0.01; ***p < 0.001. 68
Table S 2-5A: Permutational multivariant analysis of variance (PERMANOVA) of bacterial community composition based on Bray–Curtis dissimilarities of their relative abundance. Significance levels for each treatment combination are indicated by *P < 0.05; **P < 0.01; ***P < 0.001. 72
Table S 2-6A: Permutational multivariant analysis of variance (PERMANOVA) of eukaryotes community composition based on Bray–Curtis dissimilarities of relative abundance. Significance levels for each treatment combination are indicated by *p < 0.05; **p < 0.01; ***p < 0.001..... 74
Table S 2-7A: Permutational multivariant analysis of variance (PERMANOVA) microbiota density across samples during mid-season and ambient environment based on Bray–Curtis dissimilarities in Control samples. Significance levels for each treatment combination are indicated by *p < 0.05; **p < 0.01; ***p < 0.001. 76
Table S 2-8: Relative abundance of top 18 bacterial genera across the samples based on Kraken2 and Bracken..... 78
Table S 2-9: Normalized RPKS (Reads Per Kilobase of Sequence) value of eukaryotes genera across samples obtained from the EukDetect pipeline. 79
Table S 2-10A: Permutational multivariant analysis of variance (PERMANOVA) of microbial pathways based on Bray–Curtis dissimilarities. Significance levels for each treatment combination are indicated by *p < 0.05; **p < 0.01; ***p < 0.001. 80
Table S 2-11A: Comparison of jaccard index (similarity between the sets of most central nodes) of local network centrality measures between groups. Significance levels for each treatment combination are indicated by *p < 0.05; **p < 0.01; ***p < 0.001.=..... 81

Table S 3-1A: Influence of various factors on bacterial Chao1 richness 128
Table S 3-2A: PERMANOVA Bacterial beta diversity 130
Table S 4-1: Details of Xanthomonas species sequenced in this work. 193
Table S 4-2: Xanthomonas strains sequenced for this work and representative/type strain from NCBI used for comparative genomic analysis 194
Table S 4-3: Taxonomic classification, lifestyle, and variation in flagellin encoding gene in *Xanthomonas* genomes used for the association analysis. 201
Table S 4-4: List of T3Es representing all effector families and putative effectors along with their protein sequence used in the study. 213
Table S 5-1:Farm details and time of sampling 260
Table S 5-2: Climatic parameters used in the study and their meaning 262

List of Figures

Figure 2-1: Elevated O_3 exacerbates bacterial spot disease severity on the resistant cultivar but has no effect on the susceptible cultivar. Box and whisker plots showing the disease severity index (represented as % value) under elevated O_3 and ambient environmental conditions across susceptible and resistant cultivars. Significance levels for each of the treatment combination are indicated by *p < 0.05; **p < 0.01; ***p < 0.001; ****p < 0.0001..... 48
Figure 2-2: Elevated O_3 has little impact on microbial diversity and richness. However, pathogen infection on susceptible cultivar reduces microbial community richness and diversity. (A) Bacterial Chao1 richness and (B) bacterial Shannon diversity index across different environments. (C)Eukaryotic community diversity and (D) richness across different treatments. Inoculated and control samples are indicated with yellow and green bars on the top, while ambient and elevated O_3 treatments are denoted by light blue and red color bars at the bottom, respectively. 49
Figure 2-3: Elevated O_3 changes microbial community structure on susceptible cultivars challenged with pathogen infection, but not on resistant cultivars. (A) Nonmetric Multidimensional Scaling (NMDS) ordination comparing the bacterial community diversity across two cultivars, environmental conditions, and time of sampling. (B) NMDS ordination comparing the eukaryotic community diversity across two cultivars, environmental conditions, and time of sampling. 50
Figure 2-4: The effects of elevated O_3 on disease outcomes are not fully explained by changes in microbiota density and abundance. (A) Box and whisker plot showing microbiota density estimated by microbial DNA quantification (concentration of extracted DNA per mg of leaf samples) for various treatment in two cultivars. (B) Relative (Left) and absolute (right) species abundance of top 15 bacterial taxa across samples. Absolute abundance is obtained by scaling the relative abundance measurements by the microbiota density measurements. (C) Bar plots showing the relative abundance of the top 15 eukaryotic genera across the samples. Inoculated and control samples are indicated with yellow and green bars on the top, while ambient and elevated O_3 treatments are denoted by light blue and red color bars at the bottom, respectively. The time of sampling is indicated by Base (initial samples), Mid (mid-season), and End (end of the season). 51
Figure 2-5: Microbial community functions were affected by host susceptibility to pathogens, while elevated O_3 had little impact. Nonmetric Multidimensional Scaling (NMDS) ordination displaying diversity in (A) metabolic pathways across different treatment conditions in susceptible and resistant cultivars, (B) genes mapped to metabolic pathways across various treatment conditions in susceptible and resistant cultivars. 52
Figure 2-6: Pathogen infection is associated with microbial communities showing positive and stable interactions, but these interactions are random and less predictable with a shift in hub taxa in response to concurrent O_3 stress and pathogen infection. Comparison of bacterial association network across different environments. (A) Bacterial association network for the combined data set of ambient (top) and elevated O_3 (bottom) in both cultivars under control conditions. (B) Bacterial association network for the combined data set from control (top) and inoculated (bottom) samples from both the cultivars under ambient environment. (C) Bacterial association network for the combined data set from control and ambient environment (top) and inoculated and elevated O_3 (bottom) samples from both cultivars. Hub taxa are

highlighted by bold text. Node color represents the cluster determined by greedy modularity optimization. Red edges correspond to negative correlations, while green edges correspond to positive correlations. 53
Figure 3-1: Long term fertility treatment has impact on microbial diversity and richness. (A) Bacterial Chao1 richness (B) bacterial Shannon diversity (C) fungal Chao1 richness and (D) fungal Shannon diversity across different treatments. Nonmetric dimensional scaling (NMDS) analysis of (E) bacterial and (F) fungal community. 117
Figure 3-2: Canonical analysis of principal coordinates (CAP) bi-plot ordination (based upon a Bray-Curtis distance) showing canonical axes (CAP1, CAP2) that best discriminate treatments groups of (A) bacterial and (B) fungal communities. The correlation with canonical axes are only shown when the Pearson's correlation coefficient is >0.6. The length of each vector line is proportional to the strength of the correlation. 118
Figure 3-3: Bacterial and fungal communities differ across fertilizer treatment. (A) Relative abundance of the top 15 bacterial genera and (B) top 15 fungal genera across treatment groups. 119
Figure 3-4: Long term fertilization results in enrichment of various bacterial and fungal taxa across treatments. Differential abundance analysis and identification of microbial markers as predictive signatures through linear discriminant analysis (LDA) in (A) bacterial genera (B) fungal genera. based on effect size measurements (LEfSe) analysis. A taxon is considered as significantly different according to a LDA score of ≥ 3. 120
Figure 4-1: Comparative genome analysis demonstrated the presence of eight novel species in the genus *Xanthomonas*. Maximum-likelihood phylogeny based on the 1005 orthogroups of 134 strains representing the entire *Xanthomonas* genus. The phylogenetic tree was inferred using OrthoFinder v2.5.2 and drawn with R package ggtree. Here, *Xanthomonas* strains from this study are highlighted in red, while the representative/type strains are in black. The tip points are colored orange to show novel species identified from this study, while cyan represents Xanthomonas species with known taxonomy. Blue-colored blocks indicate environmental strains and host-associated strains are indicated by green-colored blocks surrounding the phylogenetic tree. 177
Figure 4-2: Multiple events of gain and loss of T3SS were evident in the genus *Xanthomonas* along with the presence of a novel T3SS cluster. Maximum-likelihood phylogeny based on the core proteome and T3SS cluster and regulators gain-loss prediction inferred for the 134 strains representing the entire *Xanthomonas* genus. The presence/absence of the different types of T3SS and regulators are represented by an ordered vector of size 7, such that a dark grey, a light grey, and a white i^{th} element in the vector indicates a full presence, a partial presence, and an absence of the i^{th} element, respectively. Inferred full and partial T3SS clusters are in color or white background, respectively. Acquisition and loss events are represented by plus and minus signs, respectively. For example, a colored +2 indicates a full acquisition of the 2^{nd} T3SS in this figure, i.e., hrp2. 178
Figure 4-3: Phylogenetic network between pathogens, weak pathogens, and commensals suggests the possibility of several recombination events during their evolutionary history. Neighbor-net tree constructed using SplitsTree software based on concatenated 12 housekeeping gene sequences generated for the 134 *Xanthomonas* strains, indicating diversity and recombination events. 179

Figure 4-4: *Xanthomonas* lifestyle can be explained by altered CAZymes landscape. Principal Component Analysis (PCA) plot showing the contribution of the different bacterial lifestyle in the distribution of gene repertoire for carbohydrate active enzymes (CAZyme). 180
Figure 4-5: Genomic architecture of *Xanthomonas* contains signatures for both phylogenetic placement and their associated lifestyle. A complex heatmap shows the select candidate genes associated with commensal, weakly pathogenic, and pathogenic lifestyles. The candidates obtained from different methods in figure S11 were further narrowed by identifying those present/enriched in commensals and weak pathogens compared to pathogens and vice versa. The matrix shows the presence/absence of genes across these lifestyles along with their functional categories and annotations on the Y-axis. 181
Figure 5-1: Comparative genome analysis demonstrated the presence of eight distinct lineages in *X. perforans*. Midpoint-rooted maximum-likelihood phylogeny of 467 *X. perforans* strains isolated form tomato, pepper, eggplant, and watercress from around the globe based on concatenated alignment of 16,823 core genome SNPs. The clades color coded according to the SCs identified in the first level of the HierBAPS hierarchy. Geographical location and host of isolation of each strain was isolated from, are indicated by different colored blocks within each respective ring surrounding the phylogenetic tree. 252
Figure 5-2: *X. perforans* is the dominant pathogen for BLS disease severity in tomato Box plot showing (A) average disease severity of samples (B) Relative and absolute abundance of *X. perforans* collected from different states during 2020, 2021 and 2022. The box is color coded based on their time of sampling. 253
Figure 5-3: Multiple *X. perforans* lineages co-occur with spatio and temporal variations, involving new lineages, strain turnover, and dominance shifts. Stacked donut chart to show the diversity of *Xp* lineages in different farms from Alabama, North Carolina, South Carolina, and Georgia. The inner two donuts circle represents the year 2020 where inner donut within the year 2020 is for mid-season and outer is for end-season. The outer two donuts are for year 2021 and the outermost two are for the year 2022..... 254
Figure 5-4: **Co-occurrence of more lineages of *X. perforans* in the field results in higher BLS disease severity.** A correlation plot shoeing the interaction between Shannon diversity of *Xp* lineages and disease severity across all samples..... 255
Figure 5-5: Various climatic factors drive *X. perforans* population. Principal coordinate analysis (PCoA) based on the overall structure of Xp lineages in all samples. Each data point represents an individual sample. PCoA was calculated using Bray-Curtis distances. Arrows with the weather parameters indicate the direction and magnitude of variables. 256

List of supplementary figures

Figure S2-1: Study site and treatment design of Atmospheric Deposition Laboratory (AtDep) site at Auburn University. (A) Satellite image of AtDep site with individual chambers label where light blue circles are the chambers with the ambient environment and red circles are the chambers with elevated O_3. Chambers 1-6 marked with yellow color are inoculated with *X. perforans,* and chambers 7-12, with green color are control samples. (B) Individual open-top chamber at the AtDep site (4 x 5m). (C) Daily average O_3 concentration in treatment chambers. Sampling points (mid and end of the season) are marked by a red arrow. (D) Inoculation method showing how plants were inoculated with *Xanthomonas* and control plants with $MgSO_4$ buffer. (E) The plants were then kept inside the OTCs where each chamber had 12 plants (6 each from susceptible and resistant cultivar). Inoculated and control plants were kept in different chambers as per the treatment plan in A. 57
Figure S2-2: Elevated O_3 increases heterogeneity in *Xanthomonas* population counts on the resistant cultivar. Line graph showing the relative and absolute abundance (ng of DNA per mg of leaf tissue) (A) *Xanthomonas* relative and absolute abundance (ng of DNA per mg of leaf tissue) across different treatments in susceptible cultivars. (B) *Xanthomonas* relative and absolute abundance (ng of DNA per mg of leaf tissue) across different treatments in resistant cultivar. end of the season. (C) *Xanthomonas* population growth based on culture-dependent technique on two different cultivars under ambient environment and elevated O_3. 58
Figure S2-3: Box plot showing the relative abundance of top 6 bacterial genera. (A-F) Relative abundance of different bacterial genera on susceptible cultivar across different treatments; (G-L) Relative abundance of different bacterial genera on resistant cultivar across different treatments. Control samples are indicated by a green bar, while the yellow bar represents the inoculated samples. The blue box represents the samples taken during the mid-season, while the orange box is end of the season samples. Different environmental conditions are represented by the ambient environment and elevated O_3 under both inoculation and control treatments. 59
Figure S2-4: Linear Discriminant Analysis Effect Size (LEfSe) of KEGG pathways between susceptible and resistant cultivar. Results were ranked by their Linear Discriminant Analysis (LDA) score. An FDR-adjusted p-value ≤ 0.05, as well as an LDA score ≥ 2.5, were used as thresholds to identify significant features. (A) pathways enriched upon inoculation of the resistant cultivar under elevated O_3 (brown), resistant cultivar under ambient environment (violet), susceptible cultivar under elevated O_3 (cyan), and susceptible cultivar under ambient environment (black). (B) pathways enriched in both cultivars under elevated O_3 (red) and ambient environment (blue), (C) pathways enriched in both the cultivars under combined stress of pathogen infection and elevated O_3 (orange) and ambient environment (lime green). 60
Figure S2-5: Stacked bar plots showing the relative abundance of dominant phyla across the samples. Control samples are indicated by a green bar, while the yellow bar represents the inoculated samples. The time of sampling is indicated by Base (initial samples), Mid (mid-season) and End (end of the season). 61
Figure S 3-1: Study site and treatment design of Cullars Rotation site at Auburn University. (A) Satellite image of Cullars rotation site with individual plots and blocks. (B) Treatment design of each plot in Cullars rotation and treatment plots used in this study. (C) Treatment blocks showing distribution of treatment plan. 121
Figure S 3-2: Pearson correlation analysis to identify the relationship between soil properties in different fertility treatments and bacterial communities. 122

Figure S 3-3: Pearson correlation analysis to identify the relationship between soil properties in different fertility treatments and fungal communities. 123
Figure S 3-4: Distribution of soil (A) bacterial and (B) fungal microbial communities at the phylum level..... 124
Figure S 3-5: Total number of differentially abundant bacterial taxa identified using DESeq2, Corncob, and Maaslin2. 125
Figure S 3-6: Microbial biomarkers as predictive signatures through linear discriminant analysis (LDA) across (A) No P vs Std fertilization, (B) No N/+legume vs No N/No Legume (C) No K vs Std fertilization. A taxon is considered as significantly different according to a LDA score of ≥ 3. 126
Figure S 3-7: Microbial biomarkers as predictive signatures through linear discriminant analysis (LDA) across (D) No N/No Legume vs No winter legumes + N and (E) No Lime vs No soil amendments treatment. A taxon is considered as significantly different according to a LDA score of ≥ 3. 127
Figure S 4-1: Relationship between genome size, GC content, and number of CDS for the genomes sequenced for this study..... 182
Figure S 4-2: Average nucleotide identity (ANI)-based heatmap showing the status of the 134 Xanthomonas strains. The intensity of the color indicates the level of identity of all-versus-all genomes as depicted by the scale..... 183
Figure S 4-3: Average nucleotide identity (ANI)-based heatmap showing the status of the representative *Xanthomonas euroxanthea* strains and strains from this study. The intensity of the color indicates the level of identity of all-versus-all genomes as depicted by the scale..... 184
Figure S 4-4: Heatmap showing the status of T3SS genes in *Xanthomonas* strains used in this study. Here, query genes names from the different species are indicated as gene*_Xeu* (genes from *Xanthomonas campestris* pv. *vesicatoria* 85-10), gene*_Xcc* (genes from *Xanthomonas campestris* pv. *campestris* ATCC33913), gene*_Xtra* (genes from *Xanthomonas translucens* pv. *translucens* DSM18974), gene*_Xalb* (genes from *Xanthomonas albilineans* CFBP 2523), and gene*_X60* (genes from *Xanthomonas* sp. 60). Red color represents the presence while blue color represents the absence of the gene. 185
Figure S 4-5: Schematic representation of the genetic organization of T3SS clusters found in *Xanthomonas campestris* pv. *campestris* ATCC33913 (Xcc), *Xanthomonas campestris* pv. *vesicatoria* 85-10 (*Xeu*), *Xanthomonas translucens* pv. *translucens* DSM18974 (*Xtr*), *Xanthomonas* sp. 60, and *Stenotrophomonas chelatiphaga* DSM 21508. Genes of the same color (except light yellow, which has no COG assignment) are from the same orthologous group. The gene names are according to the Sct T3S injectisome nomenclature and annotation available in the IMG/M system (https://img.jgi.doe.gov/cgi-bin/m/main.cgi). 186
Figure S 4-6: Heatmap showing the status of T3E in *Xanthomonas* strains used in this study. Here the query represents all effectors families, putative effectors (AvrBs1, AvrBs2, AvrBs3, Xop, HpaA, HrpW, and AvrXccA), and their diversity. Red represents the presence, while blue represents the absence of the gene. 187
Figure S 4-7: (A) Pathogenicity and (B) *in-planta* population and (C) distribution of T3E in *X. arboricola* strains (CFBP 6825, CFBP 6826, CFBP 6828, CFBP 6681, CFBP 7681, and CFBP 7680) on 4-5-week-old tomato cv. FL 47R 10 DAI (days after inoculation). The presence of common T3Es across different *Xanthomonas* strains is shown in red, while blue represents the absence of T3Es. 188

Figure S 4-8: Analysis of the presence of type three secreted effector avrBs1 with a transposon insertion among 50 randomly selected transconjugants. Mating experiments were conducted in both compatible (*Capsicum annuum* cv. Early California Wonder) and incompatible (*Capsicum annuum* cv. Early California Wonder 10R) backgrounds with *Xanthomonas euvesicatoria* strain 85-10 and *Xanthomonas* sp. strain T55. Each lane represents the pooled DNA of five to six individual transconjugants. The positive control (*Xanthomonas* sp. T55) shows an approximately 1.0 kb band indicative of avrBS1 with a transposon insertion, transconjugants isolated from the incompatible host in lanes six and eight have acquired the disrupted avirulence gene. 189
Figure S 4-9: Schematic representation and comparison of the genetic organization of the mobile genetic element island identified by MGEfinder in *Xanthomonas campestris* pv. *raphanin* 756C, *Xanthomonas campestris* F24, and *Xanthomonas campestris* F22, highlighting the XopAD effector flanked on both ends by IS5 transposases. 190
Figure S 4-10: Maximum-likelihood phylogeny based on the orthogroups of 337 *Xanthomonas* strains. Commensal *Xanthomonas* strains are highlighted in blue branches, weakly pathogenic strains in green branches, and pathogenic strains in orange. Branches were collapsed to visualize better strain diversity and phylogenetic placement with different lifestyles 191
Figure S 4-11: A complex heatmap showing results from association analysis correlating pathogenic, weakly pathogenic, and commensal phenotypes to the orthologs identified using Orthofinder. The functional categories are indicated for each candidate gene identified based on the intersection of three methods PhyloGLM, Scoary, and hyperglm (gene presence/absence), and those identified by PhyloGLM based on gene copy number. 192
Figure S 5-1: Map of the study area from Alabama, North Carolina, South Carolina, and Georgia from southeast United States. The dots in the figures represents the sampled farm and the size of the dot represents the number of samples collected during 2020, 2021, and 2022. 257

1. CHAPTER ONE: Introduction and literature review

For over 400 million years, plants and their associated microorganisms, including bacteria, fungi, viruses, and protists, have co-evolved to form a collective entity known as a plant microbiota (Trivedi et al. 2020). The interactions between the plants and their associated microbiota range from commensal relationships with no observable effects to complex and mutually beneficial symbiotic associations (Javaux 2006). Among these members of microbiota, only a limited subset can be categorized as pathogens, causing disease in specific hosts (Badet and Croll 2020). These microbial communities found on plant surfaces (rhizosphere and phyllosphere) or within plants (endosphere) has been studied for their role in nutrient acquisition, adaptation to biotic and abiotic stress, and plant growth and productivity (Lindow et al. 1982; Arun K. et al. 2020; Miransari 2011; Teixeira et al. 2021). Along with these plant-associated microbes, studies have also highlighted a vital role of soil in the health of plants, animals and humans, the concept proposed as One Health, recognizing the interconnectedness of human, animal, and plant health within shared ecosystems. Various research demonstrates a systematic link between rhizosphere and phyllosphere microbiomes, primarily driven by the rhizosphere microbiome (Bai et al. 2015; de Vries and Wallenstein 2017). As plant associated microbiomes plays critical roles in mediation plant health and productivity, our understanding of the factors that mold the composition, diversity, evolution, and functional capacities of both the phyllosphere and rhizosphere microbial communities remains limited.

The phyllosphere microbiome, consisting of microorganisms residing on a plant's aboveground parts, plays vital, yet often underestimated, roles in plant health, productivity, ecosystem functioning, and resilience to biotic and abiotic stress (Laforest-Lapointe et al. 2017; Remus-Emsermann and Schlechter 2018; Vorholt 2012; Li et al. 2022). Unlike the relatively

stable rhizosphere, the phyllosphere represents a dynamic and challenging habitat with rapid temperature, humidity, UV light fluctuations, and limited nutrient availability (Koskella 2020). Global changes, including climate shifts and land use, significantly impact ecosystems and plant-microbe interactions. Understanding their effects on phyllosphere microbiomes is essential for bolstering ecosystem resilience and sustainable plant productivity (Vitousek 1994; Zhu et al. 2022). While phyllosphere microbiomes are known to enhance plant resilience against pathogens and abiotic stress (Ehau-Taumaunu and Hockett 2022; Vannier et al. 2019), a comprehensive understanding of how microbial communities respond to and aid in plant adaptation to concurrent biotic and abiotic stressors remains elusive. Therefore, the first research objective of my PhD dissertation is to investigate the individual and combined effects of biotic stress (*Xanthomonas perforans* infection) and abiotic stress (elevated tropospheric ozone) on disease outcomes and their influence on the microbiome structure, function, and interaction networks in pepper.

The soil microbiome, characterized by its vast diversity and spatial heterogeneity, also holds a pivotal role in maintaining plant performance and ecosystem stability (Peiffer et al. 2013; Bakker et al. 2018; de Vries and Wallenstein 2017; Schimel and Schaeffer 2012). These active microbial communities within the rhizosphere are key drivers of soil biochemical processes, influencing carbon and nitrogen nutrient cycling while providing essential resilience to various environmental stresses (Haney et al. 2008; Tardy et al. 2015; Plaza et al. 2013). Soil, as a non-renewable resource, is highly susceptible to climate and agricultural practices that alter its physical and chemical characteristics, impacting microbial communities over time (Tripathi et al. 2015; Ikoyi et al. 2020; Alvarez-Martinez et al. 2020; Lehmann et al. 2020). Fertilization, although enhancing crop yield, can disrupt soil attributes and adversely affect microbial

communities. Preserving soil biodiversity and fertility is essential for ecological balance and sustainable food production, given that soil nurtures microbial communities (Li et al. 2017; Francioli et al. 2016; Kracmarova et al. 2022). Yet, a critical knowledge gap exists in understanding the long-term effects of agricultural management on these communities and their relationship with soil productivity and sustainability. Therefore, the second research objective of my PhD dissertation is to comprehensively understand the impact of long-term cropping systems and fertility management, spanning over 110 years, on soil microbial community structures, interactions, and their interplay with soil characteristics and crop yields.

Plant-microbe interactions within the phyllosphere and rhizosphere are shaped by environmental conditions, stress factors, host genetics, and other microorganisms, offering insights into the plant's role as a meta-organism (Brader et al. 2017). Understanding the roles of plant-associated microbes, their survival strategies, and community assembly factors deepens our knowledge of plant-microbe partnerships. While pathogenic microbes' impact on overall plant health is well-studied, recent molecular and genomic advances reveal that plants rely on resident microbes to suppress pathogens (Berendsen et al. 2012; Hacquard et al. 2017; Banerjee et al. 2018). Many of these microbial taxa provide ecological benefits to plants, enhancing growth and fitness (Arif et al. 2020; Haney et al. 2015). These commensal bacteria can also harbor fitness-related genes such as antibiotic resistance, with the potential to transfer to the pathogenic bacteria (Salyers et al. 2004; Dionisio et al. 2002). Some of these commensals belong to the genera that are otherwise known for pathogenic organisms. One such example is of nonpathogenic xanthomonads that do not cause apparent disease symptoms in their host plants despite their close association but have also been simultaneously isolated along with pathogenic relatives (Vauterin et al. 1990; Merda et al. 2017). The coexistence of nonpathogenic and

pathogenic xanthomonads prompts questions about their contribution to the genus's evolution, pathogenic potential, and plant immunity. Thus, the third research objective of this PhD dissertation aims to unveil the evolution of pathogenic and nonpathogenic *Xanthomonas* strains and identify genes associated with their adaptation to diverse plant and environmental lifestyles. As global plant disease outbreaks threaten food security, endemic bacterial diseases persist with moderate to severe impacts and a lack of clear control strategies complicate the management efforts to mitigate endemic pathogens (Strange and Scott 2005). Resistance genes are thought to be potential driver for disease dynamics/host-pathogen's arms race. But continued outbreaks with diverse pathogen population may be due to host selection pressure, changes in environmental factors, interspecific hybridization, and pathogen lineage mutations (Newberry et al. 2019; Gladieux et al. 2016; Martin et al. 2016). While more research is focused on emerging diseases, understanding the ecological factors driving endemic diseases has gained importance, necessitating further investigation under non-outbreak conditions (Caldwell et al. 2020; Ghelardini et al. 2016; Makiola et al. 2022). Despite extensive documentation of bacterial plant pathogen-environment interactions, the specific influence of environmental factors on endemic bacterial pathogen diversity and disease severity regulation remains unclear. Hence, the fourth research objective of this PhD dissertation centers on exploring the diversity of the endemic bacterial spot pathogen *Xanthomonas* and the driving factors behind microbial diversity and pathogen populations in the southeastern United States.

So, the four objectives of my PhD research are presented as four chapters:

Chapter 2: *Xanthomonas* infection and ozone stress distinctly influence the microbial community structure and interactions in the pepper phyllosphere.

Note – This study was previously published with the following citation:

Bhandari, R., Sanz-Saez, A., Leisner, C. P., and Potnis, N. 2023. *Xanthomonas* infection and ozone stress distinctly influence the microbial community structure and interactions in the pepper phyllosphere. ISME Commun. 3:1–13.

Chapter 3: Long-term fertilization and crop management affects soil bacterial communities.

Chapter 4: Genetic and functional diversity help explain pathogenic, weakly pathogenic, and commensal lifestyles in the genus *Xanthomonas*.

Note – This study is available as preprint (https://doi.org/10.1101/2023.05.31.543148) with the following citation and is currently under review for Genome Biology and Evolution:

Pena, M*, Bhandari, R.*, Bowers, R., Weis, K., Newberry, E., Wagner, N., et al. 2023. Genetic and functional diversity to explain commensal and pathogenic lifestyles in the genus *Xanthomonas*. *bioRxiv.* 2023.05.31.543148.

Chapter 5: Bacterial spot of tomato: strain diversity, dynamics, and environmental drivers in the southeastern United States

.

1.1 References for chapter 1

Alvarez-Martinez, C. E., Sgro, G. G., Araujo, G. G., Paiva, M. R. N., Matsuyama, B. Y., Guzzo, C. R., et al. 2020. Secrete or perish: The role of secretion systems in Xanthomonas biology. Comput. Struct. Biotechnol. J. 19:279–302

Arif, I., Batool, M., and Schenk, P. M. 2020. Plant Microbiome Engineering: Expected Benefits for Improved Crop Growth and Resilience. Trends Biotechnol. 38:1385–1396

Arun K., D., Sabarinathan, K. G., Gomathy, M., Kannan, R., and Balachandar, D. 2020. Mitigation of drought stress in rice crop with plant growth-promoting abiotic stress-tolerant rice phyllosphere bacteria. J. Basic Microbiol. 60:768–786

Badet, T., and Croll, D. 2020. The rise and fall of genes: origins and functions of plant pathogen pangenomes. Curr. Opin. Plant Biol. 56:65–73

Bai, Y., Müller, D. B., Srinivas, G., Garrido-Oter, R., Potthoff, E., Rott, M., et al. 2015. Functional overlap of the Arabidopsis leaf and root microbiota. Nature. 528:364–369

Bakker, P. A. H. M., Pieterse, C. M. J., De Jonge, R., and Berendsen, R. L. 2018. The Soil-Borne Legacy. Cell. 172:1178–1180

Banerjee, S., Schlaeppi, K., and van der Heijden, M. G. A. 2018. Keystone taxa as drivers of microbiome structure and functioning. Nat. Rev. Microbiol. 16:567–576

Berendsen, R. L., Pieterse, C. M. J., and Bakker, P. A. H. M. 2012. The rhizosphere microbiome and plant health. Trends Plant Sci. 17:478–486

Brader, G., Compant, S., Vescio, K., Mitter, B., Trognitz, F., Ma, L.-J., et al. 2017. Ecology and Genomic Insights into Plant-Pathogenic and Plant-Nonpathogenic Endophytes. Annu. Rev. Phytopathol. 55:61–83.

Caldwell, J. M., Aeby, G., Heron, S. F., and Donahue, M. J. 2020. Case-control design identifies ecological drivers of endemic coral diseases. Sci. Rep. 10:2831

Dionisio, F., Matic, I., Radman, M., Rodrigues, O. R., and Taddei, F. 2002. Plasmids Spread Very Fast in Heterogeneous Bacterial Communities. Genetics. 162:1525–1532

Ehau-Taumaunu, H., and Hockett, K. 2022. Passaging phyllosphere microbial communities develop suppression towards bacterial speck disease in tomato. Phytobiomes J.

Francioli, D., Schulz, E., Lentendu, G., Wubet, T., Buscot, F., and Reitz, T. 2016. Mineral vs. Organic Amendments: Microbial Community Structure, Activity and Abundance of Agriculturally Relevant Microbes Are Driven by Long-Term Fertilization Strategies. Front. Microbiol. 7

Ghelardini, L., Pepori, A. L., Luchi, N., Capretti, P., and Santini, A. 2016. Drivers of emerging fungal diseases of forest trees. For. Ecol. Manag. 381:235–246

Gladieux, P., Feurtey, A., Hood, M. E., Snirc, A., Clavel, J., Dutech, C., et al. 2016. The Population Biology of Fungal Invasions. In *Invasion Genetics*, John Wiley & Sons, Ltd, p. 81–100.

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

Haney, C. H., Samuel, B. S., Bush, J., and Ausubel, F. M. 2015. Associations with rhizosphere bacteria can confer an adaptive advantage to plants. Nat. Plants. 1:1–9

Haney, R. L., Brinton, W. H., and Evans, E. 2008. Estimating Soil Carbon, Nitrogen, and Phosphorus Mineralization from Short-Term Carbon Dioxide Respiration. Commun. Soil Sci. Plant Anal. 39:2706–2720

Ikoyi, I., Egeter, B., Chaves, C., Ahmed, M., Fowler, A., and Schmalenberger, A. 2020. Responses of soil microbiota and nematodes to application of organic and inorganic fertilizers in grassland columns. Biol. Fertil. Soils. 56:647–662

Javaux, E. J. 2006. Extreme life on Earth—past, present and possibly beyond. Res. Microbiol. 157:37–48

Koskella, B. 2020. The phyllosphere. Curr. Biol. 30:R1143–R1146

Kracmarova, M., Uhlik, O., Strejcek, M., Szakova, J., Cerny, J., Balik, J., et al. 2022. Soil microbial communities following 20 years of fertilization and crop rotation practices in the Czech Republic. Environ. Microbiome. 17:13

Laforest-Lapointe, I., Paquette, A., Messier, C., and Kembel, S. W. 2017. Leaf bacterial diversity mediates plant diversity and ecosystem function relationships. Nature. 546:145–147

Lehmann, J., Bossio, D. A., Kögel-Knabner, I., and Rillig, M. C. 2020. The concept and future prospects of soil health. Nat. Rev. Earth Environ. 1:544–553

Li, F., Chen, L., Zhang, J., Yin, J., and Huang, S. 2017. Bacterial Community Structure after Long-term Organic and Inorganic Fertilization Reveals Important Associations between Soil Nutrients and Specific Taxa Involved in Nutrient Transformations. Front. Microbiol. 8

Li, P.-D., Zhu, Z.-R., Zhang, Y., Xu, J., Wang, H., Wang, Z., et al. 2022. The phyllosphere microbiome shifts toward combating melanose pathogen. Microbiome. 10:56

Lindow, S. E., Arny, D. C., and Upper, C. D. 1982. Bacterial Ice Nucleation: A Factor in Frost Injury to Plants 1. Plant Physiol. 70:1084–1089

Makiola, A., Holdaway, R. J., Wood, J. R., Orwin, K. H., Glare, T. R., and Dickie, I. A. 2022. Environmental and plant community drivers of plant pathogen composition and richness. New Phytol. 233:496–504

Martin, M. D., Vieira, F. G., Ho, S. Y. W., Wales, N., Schubert, M., Seguin-Orlando, A., et al. 2016. Genomic Characterization of a South American Phytophthora Hybrid Mandates Reassessment of the Geographic Origins of Phytophthora infestans. Mol. Biol. Evol. 33:478–491

Merda, D., Briand, M., Bosis, E., Rousseau, C., Portier, P., Barret, M., et al. 2017. Ancestral acquisitions, gene flow and multiple evolutionary trajectories of the type three secretion system and effectors in Xanthomonas plant pathogens. Mol. Ecol. 26:5939–5952

Miransari, M. 2011. Arbuscular mycorrhizal fungi and nitrogen uptake. Arch. Microbiol. 193:77–81

Newberry, E. A., Bhandari, R., Minsavage, G. V., Timilsina, S., Jibrin, M. O., Kemble, J., et al. 2019. Independent Evolution with the Gene Flux Originating from Multiple Xanthomonas Species Explains Genomic Heterogeneity in Xanthomonas perforans. Appl. Environ. Microbiol. 85:e00885-19

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. 110:6548–6553

Plaza, C., Courtier-Murias, D., Fernández, J. M., Polo, A., and Simpson, A. J. 2013. Physical, chemical, and biochemical mechanisms of soil organic matter stabilization under conservation tillage systems: A central role for microbes and microbial by-products in C sequestration. Soil Biol. Biochem. 57:124–134

Remus-Emsermann, M. N. P., and Schlechter, R. O. 2018. Phyllosphere microbiology: at the interface between microbial individuals and the plant host. New Phytol. 218:1327–1333.

Salyers, A. A., Gupta, A., and Wang, Y. 2004. Human intestinal bacteria as reservoirs for antibiotic resistance genes. Trends Microbiol. 12:412–416

Schimel, J., and Schaeffer, S. 2012. Microbial control over carbon cycling in soil. Front. Microbiol. 3

Strange, R. N., and Scott, P. R. 2005. Plant Disease: A Threat to Global Food Security. Annu. Rev. Phytopathol. 43:83–116

Tardy, V., Spor, A., Mathieu, O., Lévèque, J., Terrat, S., Plassart, P., et al. 2015. Shifts in microbial diversity through land use intensity as drivers of carbon mineralization in soil. Soil Biol. Biochem. 90:204–213

Teixeira, P. J. P. L., Colaianni, N. R., Law, T. F., Conway, J. M., Gilbert, S., Li, H., et al. 2021. Specific modulation of the root immune system by a community of commensal bacteria. Proc. Natl. Acad. Sci. 118:e2100678118

Tripathi, B. M., Kim, M., Tateno, R., Kim, W., Wang, J., Lai-Hoe, A., et al. 2015. Soil pH and biome are both key determinants of soil archaeal community structure. Soil Biol. Biochem. 88:1–8

Trivedi, P., Leach, J. E., Tringe, S. G., Sa, T., and Singh, B. K. 2020. Plant–microbiome interactions: from community assembly to plant health. Nat. Rev. Microbiol. 18

Vannier, N., Agler, M., and Hacquard, S. 2019. Microbiota-mediated disease resistance in plants. PLOS Pathog. 15:e1007740

Vauterin, L., Groningen, R., Kersters, K., Gillis, M., and Mew, T. W. 1990. Towards an Improved Taxonomy of *Xanthomonas*.

Vitousek, P. M. 1994. Beyond Global Warming: Ecology and Global Change. Ecology. 75:1861–1876

Vorholt, J. A. 2012. Microbial life in the phyllosphere. Nat. Rev. Microbiol. 10

de Vries, F. T., and Wallenstein, M. D. 2017. Below-ground connections underlying above-ground food production: a framework for optimising ecological connections in the rhizosphere. J. Ecol. 105:913–920

Zhu, Y., Xiong, C., Wei, Z., Chen, Q., Ma, B., Zhou, S., et al. 2022. Impacts of global change on the phyllosphere microbiome. New Phytol. 234:1977–1986

2. CHAPTER TWO: *Xanthomonas* infection and ozone stress distinctly influence the microbial community structure and interactions in the pepper phyllosphere

Note – This study was previously published with the following citation:

Bhandari, R., Sanz-Saez, A., Leisner, C. P., and Potnis, N. 2023. *Xanthomonas* infection and ozone stress distinctly influence the microbial community structure and interactions in the pepper phyllosphere. ISME Commun. 3:1–13.

Abstract

While the physiological and transcriptional response of the host to biotic and abiotic stresses have been intensely studied, little is known about the resilience of associated microbiomes and their contribution towards tolerance or response to these stresses. We evaluated the impact of elevated tropospheric ozone (O_3), individually and in combination with *Xanthomonas perforans* infection, under open-top chamber field conditions on overall disease outcome on resistant and susceptible pepper cultivars, and their associated microbiome structure, function, and interaction network across the growing season. Pathogen infection resulted in a distinct microbial community structure and functions on the susceptible cultivar, while concurrent O_3 stress did not further alter the community structure, and function. However, O_3 stress exacerbated the disease severity on resistant cultivar. This altered diseased severity was accompanied by enhanced heterogeneity in associated *Xanthomonas* population counts, although no significant shift in overall microbiota density, microbial community structure, and function

was evident. Microbial co-occurrence networks under simultaneous O_3 stress and pathogen challenge indicated a shift in the most influential taxa and a less connected network, which may reflect the altered stability of interactions among community members. Increased disease severity on resistant cultivar may be explained by such altered microbial co-occurrence network, indicating the altered microbiome-associated prophylactic shield against pathogens under elevated O_3. Our findings demonstrate that microbial communities respond distinctly to individual and simultaneous stressors, in this case, O_3 stress and pathogen infection, and can play a significant role in predicting how plant-pathogen interactions would change in the face of climate change.

2.1 Introduction

The phyllosphere (aboveground parts) of plants is a unique, nutrient-poor habitat and is inhabited by various prokaryotic and eukaryotic microorganisms (Lindow and Brandl 2003) that colonize either the leaf surface (epiphytes) or inside the leaf tissue (endophytes) (Leveau 2019; Remus-Emsermann and Schlechter 2018). Leaf microbial community assembly and succession are influenced by deterministic and stochastic processes. Although dispersal from neighboring plants and other demographic factors such as neighbor identity and age are contributing factors toward phyllosphere microbiome diversity (Meyer et al. 2022), plant host factors such as host genotype, developmental stage (Wagner et al. 2016), and host resistance (Karasov and Lundberg 2022) shape the microbiome assembly. This host filtering of the microbiome is observed due to different resource availability on the leaf surface (van der Wal and Leveau 2011), differing physical properties (Hunter et al. 2015), and host defense signaling (Mendes et al. 2018; Lebeis et al. 2015).

Members of the phyllosphere microbiome are known to play a role in nutrient acquisition (Fürnkranz et al. 2008), plant growth and productivity (Abadi et al. 2020) and tolerance to various biotic and abiotic stresses (Li et al. 2022; Lindow et al. 1982; Rico et al. 2014; Chen et al. 2020; Yoshida et al. 2017). Pathogen invasion is one of the most influential biotic stresses affecting the plant microbial assembly in the phyllosphere (Gao et al. 2021). Pathogens can modify the habitat by secretion of virulence factors, biosurfactants, or hormones, thereby increasing resource availability for other resident colonizers including opportunists to flourish (Abdullah et al. 2017; Hoek et al. 2016). Pathogens can also influence resident microflora through niche or resource competition (Gao et al. 2021; Abdullah et al. 2017; Chaudhry et al. 2021; Lindow and Brandl 2003). Plant defense signaling activated in response to pathogen attack has also been indicated as a source of alteration of the phyllosphere community (Xin et al. 2016; Chen et al. 2020). Regardless of the source of change to the phyllosphere community, dominant members are thought to restore stability to this disturbed community (Vincent et al. 2022). Furthermore, increasing evidence has suggested that plants can recruit microbes in the phyllosphere that offer protection against pathogen (Ehau-Taumaunu and Hockett 2022; Vogel et al. 2021; Ritpitakphong et al. 2016), indicating disease-suppressive microbiome assembly in the phyllosphere in response to pathogen similar to what has been observed in the rhizosphere . Phyllosphere microbial community structure and composition is also shaped by host plant's response to abiotic stresses, such as drought (Bechtold et al. 2021; Debray et al. 2022), increase in surface temperature or warming (Aydogan et al. 2018; Manching et al. 2014; Faticov et al. 2021), elevated CO_2 (Ren et al. 2015), and ultraviolet radiation (Kadivar and Stapleton 2003).

Abiotic stressors can alter host susceptibility to pathogens by interfering with defense hormone signaling (Velásquez et al. 2018) and thus influence disease incidence. Exposure of

plants to simultaneous biotic and abiotic stressors can result in positive or negative impacts on plant responses depending on the timing, nature, and severity of each stress, as different defense signaling pathways may interact or inhibit each other (Suzuki et al. 2014; Leisner et al. 2022). Furthermore, recent work has demonstrated that climate change may lead to increased incidence of disease outbreaks due to the spread of pathogens outside their geographical range (Dudney et al. 2021). Taken together, there are many internal and external factors that can shape the phyllosphere microbiome, and work is needed to fully understand the role that phyllosphere microbiome plays in the plant's response to simultaneous biotic and abiotic stressors.

One such abiotic stressor that plants experience is elevated levels of tropospheric ozone (O_3). Global warming caused by greenhouse gases has resulted in the increase of tropospheric O_3 due to the rise in precursors such as nitrogen oxide (NOx), CO, methane, and other volatile organic compounds (Lefohn et al. 2018; Vingarzan 2004). A study across the US predicted that the 5^{th} - 95^{th} percentile for daily 8-hour maximum O_3 will increase from 31-79 parts per billion (ppb) in 2012 to 30-87 ppb in 2050 (Pfister et al. 2014). This increase in O_3 level is significant as O_3 concentrations above 40 ppb are highly phytotoxic (Ainsworth 2017). Elevated O_3 can negatively impact plants and many levels, including visible injury and reduction in photosynthesis which in turn affects plant growth, nutritional value, crop yield, and alterations to carbon allocation (Ashmore 2005; Ainsworth 2017; Burkart et al. 2013). As we learn more about how climate change associated abiotic and biotic stressors influence plant response at the molecular, cellular or transcriptomic level, important questions to address are how associated microbiome would respond to or contribute to plant's response in the presence of individual or simultaneous stressors and whether critical ecological functions of phyllosphere microbial communities would be altered in presence of stressors.

To address these questions, we explicitly focused on the response of the phyllosphere microbiome of two pepper cultivars differing in resistance towards a foliar pathogen, *Xanthomonas perforans*, in presence of ambient and elevated O_3 levels. We used an experimental setup in the field involving open-top chambers (OTCs) that allowed us to manipulate O_3 levels and dissect the influence of genotype x environment (G x E) interactions on the overall outcome of plant disease as well as on microbiome structure and function. The two pepper cultivars used in this study differed in their resistance against *X. perforans*, an emerging pepper pathogen in the southeastern US: one being susceptible cultivar Early Cal Wonder and the other being commercial cultivar PS 09979325, largely deployed in the southeastern US and known to have polygenic quantitative resistance against all eleven races of the bacterial spot pathogen (Kemble et al. 2022). This specific host-pathosystem allowed us to not only study the response of the resistant variety under combined stressors, thereby, assessing its durability under altered climatic conditions, but also to test the response of the emerging pepper pathogenic species, *X. perforans,* (Newberry et al. 2019) on the susceptible and commercial resistant varieties under an altered environment. We hypothesized that phyllosphere microbial communities will show alterations in both taxonomic and functional profiles and altered seasonal dynamics in response to altered O_3 levels, regardless of the cultivars. Interestingly, the influence of elevated O_3 on plant susceptibility depends on the lifestyle of the pathogen. Such differential effects could stem from physiological differences, pathogen biology or differences in defense signaling pathways (Pellegrini et al. 2013; Tao et al. 2023; Temple and Bisessar 1979). We hypothesized that presence of elevated O_3 will increase overall susceptibility of pepper to bacterial spot xanthomonads, even on the resistant cultivar. We also hypothesized that establishment of disease would disrupt seasonal dynamics of the phyllosphere microbiome, and

this effect will be stronger in the environments that support high disease pressure. Our experimental design allowed us to address the influence of elevated O_3 on the overall disease outcome on cultivars differing in their resistance towards pathogen as well as facilitated assessment of taxonomic and functional profiles of the phyllosphere microbiome under simultaneous stressors. Lastly, as studies have indicated the importance of functions rather than species in community structure and assembly (Burke et al. 2011), we compared functional profiles of microbiomes to see whether ecological functions of the community are rather conserved regardless of biotic or abiotic stressors.

2.2 Materials and Methods

2.2.1 Experimental site and design

The experiment was conducted at the Atmospheric Deposition (AtDep) site at Auburn University (Fig. S2-1A) in the 2021 growing season (May-July), where we harnessed OTCs (Fig. S2-1B) that allowed us to test the effect of O_3 stress on plant-pathogen-microbiome interactions and address the complexity of plant defense-development trade-off. We used 12 chambers for fumigation, where six chambers had an ambient environment, and six had elevated O_3 (Fig. S2-1A). Each elevated O_3 chamber contains four O_3 generators (HVAC-1100 Ozone generator, Ozone Technologies, Hull, IA, USA), equipped with two ultraviolet bulbs (Model GPH380T5VH/HO/4P, Ozone Technologies, Hull, IA, USA) to generate the O_3. Generators and bulbs are located within the elevated O_3 chamber fan boxes. To reach the desired set-point of O_3 (~100 ppb), O_3 generators were controlled by 0-10V control wires, which are controlled via an analog output module. To fumigate the plants, the ozonated air was blown from the fan box into the plastic lining of the open-top chamber (Fig. S2-1B). The plastic panel on the lower portion of the chamber is double walled with holes on the inside panel, allowing O_3 to be released over the

plants inside the chamber. Each chamber is connected via plastic tubing to a central gas manifold to which each chamber is opened sequentially by 3-way solenoid valves. A microcontroller cycles through the 12 solenoid valves every 24 minutes (sampling each of the 12 chambers for 2 minutes) to monitor O_3 from each chamber (Model 205 Dual Beam Ozone Monitor, 2B Technologies, Boulder, CO, USA) during the fumigation window (10 am to 6 pm). During this experiment, the average [O_3] in the control chambers was around 30.6 ppb, while the fumigated chambers had an average [O_3] of about 90.3 ppb (Fig. S2-1C). O_3 levels in the elevated chambers were significantly higher during the growing season when compared to the ambient chambers (Kruskal-Wallis, $p = 0.04$) while the O_3 levels between the elevated chambers was similar ($p = 0.62$).

Inoculation was performed on 5–6 weeks old seedlings of both cultivars. Plants were inoculated with a *X. perforans* suspension adjusted to 10^6 CFU/ml in $MgSO_4$ buffer amended with 0.0045% (vol/vol) Silwet L-77 (PhytoTechnology Laboratories, Shawnee Mission, KS, USA). The control plants were dip-inoculated in $MgSO_4$ buffer amended with 0.0045% (vol/vol) Silwet L-77 (Fig S2-1D). The dip-inoculated plants were transplanted into sterile 10-inch plastic pots (The HC Companies, OH) with soil-less potting medium (Premier Tech Horticulture, PA). The pots were then transferred to the above-mentioned OTCs and maintained inside the OTCs throughout the growing season until harvest. In each of the chambers, we had six plants, each of Early Cal Wonder (referred to hereafter as the susceptible cultivar) and PS 09979325 (referred hereafter as the resistant cultivar) (Fig S2-1E). Among the 12 chambers, plants in 6 chambers (three ambient and three elevated O_3) were inoculated with the pathogen *X. perforans* while 6 chambers (three ambient and three elevated O_3) had control plants inoculated with $MgSO_4$ buffer (Fig. S2-1A).

2.2.2 Disease severity measurements

The overall disease development was evaluated by estimating the percentage of disease symptoms caused by bacterial spot after transforming the Horsfall-Barratt ratings (Horsfall and Barratt 1945) to the midpoint of the rating range during both the mid and end of the season (Chiang et al. 2020).

2.2.3 Sample collection, DNA extraction, sequencing, and quality trimming

Pepper leaf samples were collected from both inoculated and control samples of each cultivar separately after inoculation with *Xanthomonas* or $MgSO_4$ buffer and before keeping the plants in the chambers (base samples), followed by two other time points during the growing season (mid and end of the season). For each timepoint, leaves from 6 plants of each cultivar grown inside one chamber were pooled, so we have one sample per cultivar. During sampling, leaves were collected randomly to avoid bias towards diseased leaves and with at least one leaf per plant for each cultivar. 40 grams of leaf samples were sonicated for 15 minutes in phosphate-buffered saline solution (50 mM) amended with 0.02% Tween 20 and the dislodged cells were pelleted down and processed for DNA extraction. Briefly, total DNA was extracted using *Wizard*® Genomic *DNA* Purification *Kit* (Promega, Madison, WI) as per manufacturer instructions with the addition of a phenol:chloroform:isoamyl alcohol (25:24:1) followed by ethanol precipitation. The DNA was quantified using a Qubit 3.0 fluorometer (Thermo Fischer, Waltham, MA) and the DNA samples were submitted to the Duke Center for Genomic and Computational Biology sequencing core (Duke University, Durham, NC) for library preparation, and paired end reads (2 x 150 bp) were sequenced on NovaSeq 6000 S1 flow cell system. The raw reads were then trimmed for quality using BBDuk (http://jgi.doe.gov/data-and-tools/bb-tools/) followed by host contamination

removal with KneadData (https://bitbucket.org/biobakery/kneaddata/) using pepper cv. 59 (GCA_021292125.1) genome as a reference.

2.2.4 Taxonomic profiling

Quality controlled and host decontaminated reads were taxonomically assigned using Kraken2 (v2.1.2) (Wood et al. 2019) against a standard Kraken2 database containing RefSeq libraries (O'Leary et al. 2016) of archaeal, bacterial, human, and viral sequences (as of March 01 2022). Kraken2 is a kmer based short read classification system that assigns a taxonomic identification to each sequencing read by using the lowest common ancestor (LCA) of matching genomes in the database and has been used for high accuracy classification of metagenomic reads (Su et al. 2022; Ye et al. 2019). Kraken2 report files were used as inputs to run Bayesian re-estimation of abundance with the Kraken (Bracken) (v2.6.2) (Lu et al. 2017) to re-estimate abundance at each taxonomic rank for all the samples. Bracken uses the taxonomy labels assigned by Kraken to estimate the abundance of each species. The database for Bracken was subsequently built with the Kraken2 database using the default 35 k-mer length and 100 bp read lengths based on the average read length in our sample with the lowest read length to re-estimate the relative abundance of microbial communities at the species level. The outputs from Bracken were combined using the combine_bracken_outputs.py function for downstream analysis. The kraken-biom tool (https://github.com/smdabdoub/kraken-biom) was used to convert the output from Bracken into BIOM format tables for diversity analyses in R (Team 2022).

In addition to relative abundance for each taxon, we calculated an estimate of absolute abundance based on relative abundance of different bacterial taxa and the total DNA recovered from each sample. Microbiota density described as total DNA (ng) per mg of fresh sample was calculated for each sample, which was then used to calculate the absolute abundance of different

microbial taxa as defined by ng of DNA per mg of sample multiplied with the relative abundance (Contijoch et al. 2019).

The taxonomic composition and diversity of eukaryotes in the samples were accessed using the EukDetect (v1.3) (Lind and Pollard 2021). EukDetect aligns the metagenomic reads to universal marker genes from conserved gene families curated from fungi, protists, non-vertebrate metazoan, and non-streptophyte archaeplastida genomes and transcriptomes followed by low-quality and duplicate reads filtering. The final eukaryotes abundance is calculated by filtering taxa with fewer than four reads and aligning to less than two marker genes. The resulting absolute abundance (Reads Per Kilobase of Sequence) was used to compare the diversity across the samples. The RPKS value was normalized by multiplying with a scaling factor calculated by dividing the median library size by the sample library size, which was then used to compare across the samples.

2.2.5 Culture-dependent method for determining the *Xanthomonas* population

To determine the effect of cultivar and environment on the abundance of *X. perforans*, a culture-dependent method was used for tracking the *in-planta* population of *Xanthomonas*. Plants (6 from each cultivar/chamber) were dip-inoculated as described earlier and kept inside the chambers with ambient and elevated O_3. Leaf samples were taken at day 0, 7 and 14 after inoculation to determine the *in planta* bacterial population. At each sampling time, approximately 4 cm^2 of leaf area was taken using a sterile cork borer and was macerated using a sterile Dremel® in microcentrifuge tubes with 1 ml of sterile 0.01 M $MgSO_4$ buffer. The homogenized suspension was then diluted by ten-fold followed by plating on Nutrient Agar plate using a spiral plater (Neu-tecGroup Inc., NY). Plates were then incubated at 28°C for 3 days and bacterial population was determined as colony forming units per centimeter squared of leaf area.

2.2.6 Diversity, statistical analysis, and network analysis

All statistical and diversity analyses were performed using R (v4.1.3) (Team 2022) and Rstudio (Team 2020) with the *Phyloseq* (v1.38.0) (McMurdie and Holmes 2013), *vegan* (v2.5-7) (Dixon 2003), and *ggplot2* (v3.3.5) (Wickham, Hadley 2016) packages. Before data analysis, the library size was normalized using scaling with ranked subsampling with 'SRS'-function in the 'SRS' R-package (v0.2.2) (Beule and Karlovsky 2020). Alpha diversity measures Chao1 and Shannon index were used to identify community richness and diversity, respectively. The Wilcoxon rank sum test tested significant differences in alpha diversity indices for nonparametric data and the T-test for normally distributed data. The appropriateness of these methods was verified by checking for the normal distribution of residuals based on the Shapiro-Wilk test for normality.

The differences in overall microbial profiles among the cultivars and different environmental conditions (β-diversity) were estimated using the Bray–Curtis distance. To understand the factors contributing to microbial community structure, we performed permutation multivariate analysis of variance (PERMANOVA) (Anderson and Walsh 2013) as implemented in the adonis2 (analysis of variance using distance matrices, ADONIS) with the argument 'by' set to 'margins' and analysis of similarities (ANOSIM) with 1000 permutations ($p = 0.05$) using Bray-Curtis dissimilarity in the *vegan* R package (v 2.5-7). In addition, multivariate homogeneity of group dispersion test (BETADISPER) (Anderson et al. 2006) was performed to determine the homogenous dispersion between the factors in relation to their microbial taxa. Non-metric multidimensional scaling (NMDS) among the sample groups was calculated using Bray-Curtis dissimilarity and visualized using the *ggplot2* package in R.

For our network analysis, the taxonomic data was subsetted to at least 0.5% relative abundance in over 20% of the samples (prevalence) to ensure that all samples had sufficient

sequencing depth to recover most of the diversity. Correlation network analysis was performed using the SPRING (Yoon et al. 2019) approach implemented in the R package *NetCoMi* (v1.1.0) (Peschel et al. 2021). Community structures across the treatment were estimated using the "cluster_fast_greedy" algorithm (Clauset et al. 2004), and hub taxa were determined using the threshold of 0.95. A Jaccard index was used to test for similarities (Jacc = 0, lowest similarity and Jacc = 1, highest similarity) in selected local network centrality measures (degree, betweenness centrality, closeness centrality, and eigenvector centrality) to determine the hub or keystone taxa. A quantitative network assessment was performed with a permutation approach (1000 bootstraps) with an adaptive Benjamini-Hochberg correction for multiple testing.

2.2.7 Functional profiling

Functional profiling of the microbial communities was conducted on concatenated paired-end sequences with HUMAnN3 (v3.0) (Beghini et al. 2021) to quantify gene abundance (UniRef90 gene-families) (Suzek et al. 2015) and MetaCyc pathways (Caspi et al. 2020). ChocoPhlAn nucleotide database v30 was used for functional pathway abundance and coverage estimation. The gene families and pathway abundance tables were sum-normalized to copies per million reads (CPM) to facilitate comparisons between samples with different sequencing depths. The output from HUMAnN3 was then imported into QIIME2 (v2021.11) (Hall and Beiko 2018) to generate nonmetric multidimensional scaling (NMDS) ordinations using Bray-Curtis dissimilarly matrix. To understand the factors driving functional profiles, we performed permutation multivariate analysis of variance (PERMANOVA) (Anderson and Walsh 2013) as implemented in the adonis2 (analysis of variance using distance matrices, ADONIS) and analysis of similarities (ANOSIM) with 1000 permutations ($p = 0.05$) with different factors (cultivars, environment, inoculation status, and time of sampling), as described above. Differentially

abundant pathways across the treatment were identified using the LEfSe (Linear discriminant analysis Effect Size) (v1.1.2) (Segata et al. 2011). Pathways with a corrected *p*-value of 0.05 or less and Linear Discriminant Analysis (LDA) score of log > 2.5 were classified as significantly increased within one of the two groups.

2.3 Results

2.3.1 Influence of O_3 levels on disease severity on resistant and susceptible cultivars

Overall higher disease severity index was recorded on the susceptible cultivar compared to the resistant cultivar. Under ambient conditions, the susceptible cultivar supported an average of 53.01% disease severity index during mid-season, that decreased to 15.11% by the end of the growing season. The resistant cultivar supported minimal disease with disease severity index of 0.37% during mid-season and 0.29% by the end of the season. Elevated O_3 did not impact disease severity on the susceptible cultivar. However, significantly higher disease severity index was observed on the resistant cultivar under elevated O_3 conditions, both at mid-season (12.61%) ($p < 0.001$) and end of the season (2.01%) ($p = 0.01$) compared to the ambient environment (mid-season = 0.37%, end of the season = 0.29%) (Fig. 2-1, Table S2-1).

2.3.2 Sequencing statistics

The samples collected in the beginning of the experiment (base samples) and twice during the growing season (mid-season and end of the season) were subjected to shotgun metagenome sequencing, which produced 2.83 to 17.16 Gbps of raw reads per sample. Adapter trimming and removal of low-quality reads resulted in the loss of 4.3 to 11.3% of the total reads among different samples. Of the quality trimmed reads, 5.78 to 39.09% of the reads were identified as host reads and removed from further analysis. The samples at the early seedling

stage yielded very few reads upon filtering because of higher host contamination (23-39%), indicating minimal microbial colonization in the greenhouse-grown seedlings before transplanting. Around 50.61% to 84.56% of the original total reads were retained for downstream analysis (Table S2-2).

2.3.3 Microbial diversity and richness are reduced under susceptible response but are not significantly affected by elevated O_3.

We next investigated the effect of inoculation and elevated O_3 and their interaction on overall microbial diversity and richness of the phyllosphere communities. Overall bacterial richness and diversity values in both the mid and end of the season samples were higher in control plants when compared with base samples. This could be attributed to low microbial colonization levels on greenhouse-grown base samples that increased in diversity and richness upon exposure to natural field conditions. Eukaryotic diversity in the base samples was not calculated as these samples had reads counts below the threshold (fewer than 4 reads that align to fewer than 2 marker genes) to be considered present in the sample. The O_3 stress alone did not influence bacterial (Table S2-3) and eukaryotic richness and diversity (Table S2-4) in both cultivars. However, pathogen infection led to significantly lower bacterial richness ($p < 0.001$) (Fig. 2A) and diversity (Kruskal-Wallis, $p = 0.01$) (Fig. 2-2B) as well as lower eukaryotic richness (Kruskal-Wallis, $p = 0.01$) (Fig. 2-2C) and diversity (Kruskal-Wallis, $p = 0.02$) (Fig. 2-2D) on the susceptible cultivar under ambient conditions throughout the growing season compared to that on control plants. Under combined stress of pathogen and elevated O_3, there was a significant effect on both richness ($p = 0.01$) and diversity ($p = 0.04$) only during end of the season on the susceptible cultivar. Inoculation and elevated O_3 did not influence bacterial richness (*pinoc* = 0.81, *penv* = 0.07) (Fig. 2-2A) and diversity (*pinoc* = 0.27, *penv* = 0.62) (Fig.

2-2B), or eukaryotic richness (Kruskal-Wallis, *pinoc* = 0.08, penv = 0.31) (Fig. 2-2C) and diversity (Kruskal-Wallis, *pinoc* = 0.23, *penv* = 0.82) (Fig. 2-2D) on the resistant cultivar. Time of sampling had significant influence on bacterial richness and diversity ($p < 0.01$) in both the cultivars.

2.3.4 The effect of O_3 levels was significant on the eukaryotic community yet was minimal in shaping bacterial community structure.

To visualize the differences in bacterial and eukaryotic community structure between samples from two pepper cultivar and two environmental conditions, the taxonomic abundance profiles were used to compute the Bray-Curtis distance matrix and plotted into two dimensions using non-metric multidimensional scaling (NMDS). To understand the relative influence of each factor and their interaction on the overall phyllosphere microbial community structure, we performed a PERMANOVA on Bray-Curtis dissimilarities using cultivar, time of sampling, environment, and inoculation as independent variables. Overall, the effect of cultivar, time of sampling, and inoculation were highly significant in shaping bacterial communities ($p < 0.001$) in addition to the interactions of cultivar, time, and inoculation *(*$p = 0.03$) (Table S2-5A), with separation of inoculated susceptible plants from control susceptible, inoculated and control resistant plants (Fig. 2-3A). We further assessed individual factors' influence and interactions across two sampling points. The effect of the cultivar was significant but diminished over the growing season (mid-season: $R^2 = 0.23$, $p < 0.001$; end of the season: $R^2 = 0.06$, $p = 0.03$). In contrast, effect of inoculation increased over the course of growing season (mid-season: $R^2 = 0.20$, $p < 0.001$; end of the season: $R^2 = 0.55$, $p < 0.001$) (Table S2-5B, Table S2-5C). The effect of interaction among cultivar and inoculation on bacterial communities remained statistically significant over time, although the effect decreased in size by the end of the growing season

(mid-season: $R^2 = 0.15$, $p < 0.01$; end of the season: $R^2 = 0.05$, $p = 0.04$). The effect of elevated O_3 was minimal, with it being not statistically significant by the end of the growing season (mid-season: $R^2 = 0.05$ $p = 0.04$; end of the season: $R^2 = 0.02$, $p = 0.15$) (Table S2-5B, Table S2-5C). The interaction between the environment and other variables was not statistically significant throughout the growing season. An increase in O_3 levels did not alter the bacterial community structure on the susceptible cultivar. However, it influenced bacterial communities on the resistant cultivar ($R^2 = 0.14$, $p = 0.02$) (Table S2-5D) in the absence of *Xanthomonas*. There was no difference in the microbial communities between the chambers with elevated O_3 ($p = 0.69$) or ambient environment ($p = 0.85$) suggesting there is no effect of chamber in overall bacterial diversity (Table S2-5E, S2-5F).

Like bacterial communities, eukaryotic communities diversity was also significantly influenced by the environment, cultivar, time of sampling, and inoculation ($p < 0.01$) (Table S2-6A, Table S2-6B). Cultivar had a significant effect on eukaryotic diversity with more influence during the end of the season (mid-season: $R^2 = 0.12$ (Table S2-6C), $p = 0.007$; end of the season: $R^2 = 0.37$, $p = 0.001$ (Table S2-6D, S2-6E)). An increase in O_3 levels significantly affected the eukaryotic communities during the mid-season, while it was not significant during the end of the season (mid-season: $R^2 = 0.22$, $p = 0.001$ (Table S2-6C); end of the season: $R^2 = 0.06$, $p = 0.19$ (Table S2-6D, Table S2-6E). The effect of inoculation on eukaryotic communities was higher during the mid-season, and it decreased during the end of the season (Figure 2-3B) (mid-season: $R^2 = 0.15$, $p = 0.003$ (Table S2-6C); end of the season: $R^2 = 0.11$ $p = 0.03$ (Table S2-6D, S2-6E)). The influence of time of sampling on clustering was evident in shaping both bacterial and eukaryotic communities (Fig. 2-3A, 2-3B).

These findings indicate that microbial communities on resistant and susceptible cultivars were similar in absence of any stress, either pathogen or elevated O_3, and influence of seasonal succession was evident on both bacterial and eukaryotic communities. Pathogen infection led to a shift in the bacterial community composition on the susceptible cultivar as the growing season progressed. However, despite presence of the *Xanthomonas* population on resistant cultivar, microbial community structure was like that observed on uninoculated plants. Despite increases in disease severity on the resistant cultivar under elevated O_3, bacterial and eukaryotic communities were similar in their composition to that under ambient environment.

2.3.5 Influence of pathogen infection and O_3 stress on relative and absolute abundance of microbial taxa

The presence of *Xanthomonas* on control plants of susceptible and resistant cultivars suggested low levels of natural inoculum in the field. However, the relative abundance of *Xanthomonas* on control plants did not increase significantly over time (< 5% by the end of the season). Both relative and absolute abundance of *Xanthomonas* increased from mid-season to end of the season on inoculated susceptible and resistant cultivars (Fig. S2-2A, S2-2B). Significant variation in the relative (~33-87%) as well as absolute (~13-37%) abundance of *Xanthomonas* on resistant inoculated plants under elevated O_3 conditions was worth noting. However, presence of elevated O_3 did not result in a significant difference in relative (Kruskal-Wallis: $p_{ECW} = 0.12$, $p_{X10R} = 0.78$) or absolute (Kruskal-Wallis: $p_{ECW} = 0.15$, $p_{X10R} = 0.54$) abundance of *Xanthomonas* in either cultivar (Fig. S2-2A, S2-2B). This observation was surprising given that disease severity levels under elevated O_3 conditions on resistant inoculated plants were significantly higher than that under ambient environment.

To further confirm the influence of elevated O_3 and cultivars on *Xanthomonas* population, we analyzed the *in-planta* population of *X. perforans* determined using a culture-dependent method for day 7 and day 14 post-inoculation. While this short-course experiment may not reflect the outcome for the entire growing season, it allowed us to evaluate the effect of elevated O_3 on the *Xanthomonas* population. Similar to the above observations, there was no significant effect of environment (i.e., ambient vs. elevated O_3) on *X. perforans* population in these cultivars ($p_{ECW} = 0.31$, $p_{X10R} = 0.34$) (Fig. S2-2C).

As the increase in the disease severity on the resistant cultivar under elevated O_3 was not the result of changes in *Xanthomonas* population, we hypothesized that this increase was associated with a significant reduction in overall microbial density associated with the resistant cultivar under elevated O_3 compared to ambient environment, referring to an altered prophylactic shield from microbiota under elevated O_3. Microbiota density estimates were obtained based on microbial DNA content per mg of sample, similar to those calculated in gut microbiome studies (Reyes et al. 2013). There was a significant effect of inoculation on microbiota density ($p < 0.001$), while neither cultivar ($p = 0.15$) nor elevated O_3 ($p = 0.19$) had a significant effect on microbiota density (Fig. 2-4A). There was significantly lower microbiota density in mid-season samples on inoculated resistant cultivar under elevated O_3 compared to susceptible cultivar ($p = 0.01$), but not for end of the season samples ($p = 0.13$) (Table S2-7A, S2-7B, S2-7C, S2-7D). We further estimated absolute abundance of each bacterial genus by multiplying its relative abundance (Figure 2-4B) by the total DNA per mg of sample. Overall absolute abundance of microbiota was lower on inoculated resistant cultivar compared to inoculated susceptible cultivar, under both environments, although this difference was not statistically significant, accounting for large variation across samples (Figure 2-4C, Table S2-7E, S2-7F). Total absolute

abundance of microbiota associated with inoculated resistant cultivar under ambient environment was not significantly different compared to that under elevated O_3 environment.

Next, we investigated the temporal dynamics in community assembly and succession in the phyllosphere, and patterns were compared between inoculated and control plants. Detailed taxonomic description of bacteria and eukaryotes across different treatments is given in supplementary information. The taxonomic diversity analysis showed that several bacterial (Fig. 4B, Table S8) and eukaryotic genera (Fig. 2-4C, Table S2-9) monopolizing the phyllosphere environment. These microbial genera are differentially affected by the presence of a pathogen, environmental stress, and their interaction.

Next, we identified genera that showed changes in relative abundance in response to cultivars, or elevated O_3. Bacterial genera *Pseudomonas*, *Pantoea*, *Methylobacterium*, *Sphingomonas*, *Methylobrum*, etc., were negatively affected, while *Microbacterium* was positively influenced in the presence of *Xanthomonas* on susceptible cultivar (Fig. S2-3A to F). In contrast to the susceptible cultivar, the relative abundance of *Pseudomonas* and *Sphingomonas* increased in the presence of the *Xanthomonas* on the resistance cultivar. The bacterial genus *Methylobacterium* was negatively influenced by elevated O_3, while the genera *Pseudomonas* and *Sphingomonas* were positively impacted in resistant cultivar (Fig. S2-3G to L). Regarding eukaryotes, the genus *Bullera* was positively affected by elevated O_3, while the genus *Epicoccum* and *Protomyces* had temporal variation regardless of treatment (Fig. 2-4C).

2.3.6 Functional composition of phyllosphere communities when exposed to O_3 stress and pathogen infection

As the microbial composition was significantly affected by cultivar, inoculation, and time, we sought to investigate whether observed taxonomic differences reflected niche-specific microbial functions. Overall community functions based on the relative abundance of metabolic pathways (Fig. 2-5A) as well as associated gene families that were mapped onto the pathways (Fig. 2-5B) were not affected by these individual factors ($p > 0.05$) (Table S2-10A, S2-10B). However, the interaction between inoculation, cultivar, and sampling time had a significant effect on microbial functions and gene families ($p < 0.01$) (Table S2-10A, S2-10B), as indicated by dissimilarities in the functional composition of both gene families and pathways associated with communities recovered from the inoculated susceptible cultivar compared to the inoculated resistant cultivar. We observed significant effect of cultivar during the end of the season ($p = 0.01$) (Table S2-10C). Elevated O_3 did not alter the functional assemblage of phyllosphere microbiome either on resistant or susceptible cultivars and regardless of the inoculation status. We observed similar functional profiles both in terms of genes as well as pathways across timepoints during the growing season on the respective cultivars despite differences in the species composition in mid vs. end of the season samples. This is likely due to substantial functional redundancy in the metabolic pathways associated with microbial communities over the growing season despite seasonal succession of taxa in the phyllosphere.

To find differentially abundant pathways that explain differences among treatments in response to pathogen infection, elevated O_3, and their interaction, we performed Linear discriminant analysis Effect Size (LEfSe). Upon pathogen infection and elevated O_3, metabolic pathways related to heme scavenging (source of bioavailable iron) were enriched in microbial

communities recovered from the resistant cultivar whereas pathways associated with carbohydrate metabolism (pentose phosphate pathway, gallate degradation, glyoxylate cycle), protection (lipid IV_A biosynthesis), growth and maintenance (phosphatidyl glycerol biosynthesis, CDP-diacylglycerol biosynthesis, GDP-mannose biosynthesis), and metabolism of unsaturated fatty acid (gondoate biosynthesis) were enriched in microbial communities recovered from the resistant cultivar under ambient conditions (Fig. S2-4A). Metabolic pathways that were enriched in microbial communities associated with both the cultivars upon O_3 stress included pathways involved in primary energy production and the degradation of unsaturated fatty acids (beta-oxidation, pentose phosphate), various defense-related pathways against oxygen stress and DNA repair (ubiquinol 7, pyrimidine (deoxy)nucleotides) and pathways related to oxygen-independent respiration (oxygen-independent heme b biosynthesis) (Fig. S2-4B). In the presence of both the pathogen and elevated O_3, pathway related to purine nucleotide production and degradation was enriched (Fig. S2-4C).

2.3.7 Microbial network topology is altered under combined pathogen and O_3 stress.

To assess whether pathogen infection and O_3 stress alone or in combination affected overall microbial association in the phyllosphere, bacterial co-occurrence networks and their topological features across treatments were compared. We assessed local network centrality measures using degree, betweenness, closeness and eigenvector centrality used to determine hub taxa for bacterial co-occurrence networks under elevated O_3 (Fig. 2-6A), inoculation (Fig. 2-6B), and combined stress of elevated O_3 and pathogen (Fig. 2-6C), and compared to ambient, control condition or control condition and ambient environment, respectively. We observed that all treatment comparisons mentioned above showed significant differences for all the local network centrality measures (Table S2-11A). A hub taxon is a highly connected taxon and is known to

have strong impact in the network. There was a significant difference in hub taxa among treatment groups when comparing control with inoculated samples or control and ambient environment with pathogen and O_3 stress (Table S2-11A, S2-11B). However, there was no change in hub taxa on plants exposed to elevated O_3 compared to ambient environment. Comparing the overall similarities of the two networks between the ambient vs. individual stress or combined stress of elevated O_3 and pathogen based on adjusted Rand index (ARI) indicated values close to 0 (ARI = 0.02, $p = 0.07$) for ambient vs. elevated O_3 stress; control vs. inoculated (ARI = 0.03, $p = 0.02$) and control and ambient environment vs combined stress (ARI = 0.10, $p < 0.001$) (Table S2-11C). These observations indicate that the partitions of species into communities show a low degree of similarity in these comparisons. These results, with differences in topology between these networks and dissimilarity in local network centrality measures, indicate that combination of pathogen infection and O_3 stress results shifts in the bacterial community interactions in the phyllosphere.

Next, we assessed global microbial network properties such as number of edges as a measure of complexity, modularity, average path length and clustering coefficient, that compare network topologies across treatments (Dini-Andreote et al. 2014; Hernandez et al. 2021). The current version of *NetCoMi* can only perform 1000 permutations due to the high run time of a single network construction. Since the minimum possible p-value for 1000 permutations is 1/1000, the power is quite low, and this results in large p values after adjusting for multiple testing. Increasing number of permutations may allow evaluating global network properties with sufficient statistical power. Thus, in this study, we focused on absolute differences for each parameter under comparison, rather than associated p-values. Microbial networks under ambient environment showed higher positive edge percentage, higher clustering coefficient, and lower

average path length compared to elevated O_3 (Table S2-11D). This suggests more positive interactions in ambient environments and that O_3 stress may foster less complex and negative associations among community members. On the contrary, the presence of pathogen infection led to a more positive edge percentage, lower path length, higher modularity, and higher clustering coefficient, suggesting that all nodes were highly interlinked within the networks to form a more complex and stable network under pathogen infection (Table S2-11D). However, in the presence of pathogen infection and O_3 stress, more positive interactions were found under ambient environment and control conditions with lower path length and higher clustering coefficient, suggesting that combined stress possibly creates less complex and less stable associations among community members (Table S2-11D).

2.4 Discussion

Changing climate and modern agricultural practices have pre-disposed agro-ecosystems to an increased threat of pests, thus, leaving us with the unpredictability as to how plants will adapt to the simultaneous biotic and abiotic stressors. Many studies have proposed the role of plant-associated microbiomes in contributing towards plant resilience in the changing climate and extending plant immunity against pathogens (Trivedi et al. 2022; Vannier et al. 2019; Debray et al. 2022; Ehau-Taumaunu and Hockett 2022). However, we have yet to fully understand how microbial communities, both respond to as well as contribute towards plant adaptation, in presence of simultaneous biotic and abiotic stressors. In this study, we tested individual and simultaneous effects of elevated O_3 and pathogen stress on phyllosphere bacterial and eukaryotic community structure, function, and stability, and on overall plant disease outcomes on susceptible and resistant pepper cultivars. The resistant pepper cultivar used in this study possesses a compliment of resistance genes that provide an intermediate level of resistance

against all currently known pepper races of bacterial spot *Xanthomonas* (Stall et al. 2009). Our rationale of including this cultivar in this study design was to understand the durability of this resistant cultivar that is currently widely deployed in the southeastern US in response to emerging pathogen species and under elevated O_3, representing future climate.

While the apparent influence of elevated O_3 was not observed on disease severity levels on the susceptible cultivar, the resistant cultivar displayed higher disease severity under elevated O_3 throughout the growing season as compared to ambient environment (Fig. 2-1). This change in disease severity may also be indicative of resistance erosion under elevated O_3 conditions. Unfortunately, the choice of cultivars used in this study not being near-isogenic prevents us from evaluating the influence of resistance loci on microbiome as was done in previous studies (Wagner et al. 2020). The increased disease severity observed on the resistant cultivar under elevated O_3, however, was not associated with the increase in *Xanthomonas* population as estimated by absolute abundance data when compared to the ambient environment. Such a culture-independent DNA sequencing method may not accurately indicate living pathogen cell count and may warrant confirmation of these findings with a culture-dependent pathogen population estimate or with methods such as Quantitative PCR (qPCR) (Bonk et al. 2018; Jian et al. 2020), digital droplet PCR (Morella et al. 2018; Hindson et al. 2011). Although not for the entire growing season, we monitored the dynamics of the *Xanthomonas* population during a short-term two-week experiment and the results supported the previous findings that *Xanthomonas* population was unaffected despite higher disease severity under elevated O_3 on the resistant cultivar. Interestingly, high variability in the *Xanthomonas* population counts on the resistant cultivar under elevated O_3 was worth noting. This may indicate a plastic response of the pathogen during adaptation to the resistant cultivar under altered environment.

A large body of work has indicated that climatic fluctuations can have a profound effect on the outcome of plant-pathogen interactions (Cheng et al. 2013; Huot et al. 2017; Zhou et al. 2019), which may result from alteration of the host environment via modification of host defense pathways, increased pathogen infection efficiency under altered environments, or alteration in the microbiome-provided extended immunity. These three plausible explanations are outlined below that could synergistically drive plant-pathogen-microbiome interactions and help to explain the observation from this study of potential resistance erosion under elevated O_3 conditions.

Studies on plant's response to a combination of abiotic and biotic stress have shown a unique and more complex response than that of individual stresses (Rizhsky et al. 2002, 2004; Zandalinas et al. 2021; Leisner et al. 2022). The effect of combined stress is governed by various factors such as time, degree of stress, plant genotype, and other climatic or environmental factors, thus, not necessarily additive in nature (Omae and Tsuda 2022). Plants respond to biotic and abiotic stresses via complex yet overlapping defense signaling pathways (Zhu 2016; Klessig et al. 2000), with induction of the abscisic acid (ABA) pathway observed upon abiotic stress, which antagonizes the salicylic acid (SA) pathway involved in pathogen defense (Jiang et al. 2010; Yasuda et al. 2008). Simultaneous stresses of pathogen infection and elevated O_3 may result in altered host immune response on the resistant cultivar. Oxidative damage of the plant cuticle caused by elevated [O_3] can increase exposure to pathogens, thus, impacting disease severity (Berner et al. 2015). Complementing this current study with host transcriptomics will explain if such host defense alteration may be what explains the increased susceptibility on resistant cultivar in presence of elevated O_3. Secondarily, increased pathogen virulence via increased effector output (Huot et al. 2017) under altered environment may explain increased

disease severity in absence of significant increase in pathogen population. The increased variation in pathogen population could be due to either host plastic response or plasticity in pathogen population.

Third and the most important explanation for the observations from this study is the alteration in microbiome-mediated protection on the resistant cultivar in response to elevated O_3 and pathogen infection. Microbial communities recruited by the resistant cultivar in the phyllosphere could have a protective role against the pathogen as it has been demonstrated in previous studies (van Dam and Heil 2011; Li et al. 2022) and this protective role may have been altered under elevated O_3, which may have led to increased disease severity. The bacterial and eukaryotic community composition, structure and function on the susceptible cultivar did not differ in the absence of pathogen infection or elevated O_3. However, bacterial community structure on the resistant cultivar were influenced by presence of elevated O_3, but in absence of the pathogen. Whether such differential influence on microbial community structure is due to specific resistance loci remains to be determined since the cultivars that we investigated were not near-isogenic lines for the resistance loci. On the susceptible cultivar, the presence of pathogen infection caused a sizeable shift in the bacterial community structure and function, even though concurrent O_3 stress did not further alter the microbiome structure and function. Although no significant shift in microbiome structure and function was observed on the resistant cultivar upon infection, overall microbiota density associated with infected resistant cultivar was lower compared to infected susceptible cultivar. Furthermore, concurrent O_3 stress resulted in lower total microbiota density during mid-season sampling on the infected resistant cultivar. Whether such reduction reflects impaired prophylactic potential of microbiome associated with resistant cultivar under the combined impact of elevated O_3 and pathogen infection remains to be

investigated. Further experiments to assess the microbiome-mediated protection against pathogen can be designed using synthetic communities associated with the resistant cultivar, similar to the previous studies (Berg and Koskella 2018). These experiments may provide opportunities to dissect the influence of altered environment on absolute abundance of individual members of the community and their interactions and associated functional traits. Interestingly, our data did not reveal any influence of simultaneous stressors on functions of microbial communities associated with the resistant cultivar. This was surprising given the previous studies indicate enrichment of specific metabolic pathways under abiotic or biotic stressors (Muñoz-Elías and McKinney 2006; Nogales et al. 2011; Schiff 1980).

Microbial function in the ecosystem is determined not just because of the number and composition of taxa but also the various positive, negative, direct, or indirect associations among the community members (Wagg et al. 2019). In response to the pathogen challenge, we observed network parameters indicative of a densely connected network. These findings of enhanced positive and complex association among the microbial communities upon pathogen infection have been observed in both the phyllosphere and endosphere (Luo et al. 2019; Hu et al. 2020; Tan et al. 2022). Such densely connected network indicates cooperative association such as facilitation, mutualism or commensalism, and cross-feeding (Faust and Raes 2012; Hernandez et al. 2021). Such connected networks, referred to as small-world networks (Ortiz-Álvarez et al. 2021), are hypothesized to harbor resistance toward disturbances. In contrast, microbial co-occurrence networks across O_3 stress and simultaneous pathogen and O_3 stress showed a similar trend of a relatively unstable random network compared to the control environment. This finding agrees with the notion that varying degrees of environmental stress disturb the stability of microbial communities (Hernandez et al. 2021). The observation from the similarity of the most

central node suggests that microbial communities are considerably different across different treatments. The presence of a pathogen and simultaneous pathogen and O_3 stress considerably affected hub taxa. However, simultaneous stressors, but not individual stresses, had considerable influence on the most influential taxa, suggesting that plants respond to simultaneous stresses by changing the most influential microbial member in the random network. It would be interesting to dissect further the influence of individual cultivar and, thus influence of host defense responses on microbial community networks, as we observed a strong cultivar effect on community composition. However, the present study is limited in sample size, which does not allow sufficient power to compare the network structure across individual cultivars. As we observed that elevated O_3 impacted eukaryotic communities more strongly than bacterial communities and pathogen infection impacted bacterial communities, influence on cross-kingdom interactions cannot be ruled out in this case. Nevertheless, the present study has limitations in determining how specific and concurrent stressors affect cross-kingdom interactions due to the absence of appropriate methods to evaluate relative abundance of eukaryotic communities using shotgun metagenome data. It is possible that elevated O_3 will have an impact on cross-kingdom interactions, as has been shown with other abiotic stressors (Debray et al. 2022).

Overall, our study demonstrated that microbial communities respond to a change by not only altering community composition but also interactions among members and overall community function. This this work provides a base for our understanding of the complex response of microbial communities and their interactions with the host genotype in response to a changing climate. As plants have evolved in association with their phyllosphere microbiome members, the community members identified in this study have shown to be particularly

susceptible to a shift in response to abiotic stress or combined stress. Findings from this study are crucial to evaluate for future work on harnessing the microbiome for stress-tolerant plants.

2.5 Data and code availability

Sequence data generated from this work have been deposited in the SRA (Sequencing Read Achieve) database under the BioProject accessions PRJNA889178. All other data and code used in this study are available in the following GitHub repository (https://github.com/Potnislab/AtDep_2021_metagenome).

2.6 Acknowledgments

This work was supported by the Alabama Agricultural Experiment Station and the Hatch program (Project # 10108601) of the National Institute of Food and Agriculture, U.S. Department of Agriculture. We thank Auston Holland for his assistance in maintaining the plants in the greenhouse and throughout the experiment. We thank the members of Potnis, Leisner, and Sanz-Saez labs for their help with the initial planting. We thank Seth Johnston for setting up and maintaining the OTCs and fumigation. We thank Dr. David Young and the staff at the Alabama Supercomputer Authority for providing the computational resources necessary to conduct this work.

2.7 References for chapter 2

Abadi, V. A. J. M., Sepehri, M., Rahmani, H. A., Zarei, M., Ronaghi, A., Taghavi, S. M., et al. 2020. Role of Dominant Phyllosphere Bacteria with Plant Growth–Promoting Characteristics on Growth and Nutrition of Maize (*Zea mays* L.). J. Soil Sci. Plant Nutr. 20:2348–2363

Abdullah, A. S., Moffat, C. S., Lopez-Ruiz, F. J., Gibberd, M. R., Hamblin, J., and Zerihun, A. 2017. Host–multi-pathogen warfare: pathogen interactions in co-infected plants. Front. Plant Sci. 8:1806.

Ainsworth, E. A. 2017. Understanding and improving global crop response to ozone pollution. Plant J. 90:886–897

Anderson, M. J., Ellingsen, K. E., and McArdle, B. H. 2006. Multivariate dispersion as a measure of beta diversity. Ecol. Lett. 9:683–693.

Anderson, M. J., and Walsh, D. C. I. 2013. PERMANOVA, ANOSIM, and the Mantel test in the face of heterogeneous dispersions: What null hypothesis are you testing? Ecol. Monogr. 83:557–574.

Ashmore, M. R. 2005. Assessing the future global impacts of ozone on vegetation. Plant Cell Environ. 28:949–964.

Aydogan, E. L., Moser, G., Müller, C., Kämpfer, P., and Glaeser, S. P. 2018. Long-Term Warming Shifts the Composition of Bacterial Communities in the Phyllosphere of Galium album in a Permanent Grassland Field-Experiment. Front. Microbiol. 9:144.

Bechtold, E. K., Ryan, S., Moughan, S. E., Ranjan, R., and Nüsslein, K. 2021. Phyllosphere Community Assembly and Response to Drought Stress on Common Tropical and Temperate Forage Grasses. Appl. Environ. Microbiol. 87:e00895-21.

Beghini, F., McIver, L. J., Blanco-Míguez, A., Dubois, L., Asnicar, F., Maharjan, S., et al. 2021. Integrating taxonomic, functional, and strain-level profiling of diverse microbial communities with bioBakery 3 eds. Peter Turnbaugh, Eduardo Franco, and C Titus Brown. eLife. 10:e65088.

Berendsen, R. L., Pieterse, C. M. J., and Bakker, P. A. H. M. 2012. The rhizosphere microbiome and plant health. Trends Plant Sci. 17:478–486.

Berg, M., and Koskella, B. 2018. Nutrient- and Dose-Dependent Microbiome-Mediated Protection against a Plant Pathogen. Curr. Biol. 28:2487-2492.e3.

Berner, J. M., Maliba, B., and Inbaraj, P. 2015. Impact of Elevated Levels of CO2 and O3 on the Yield and Photosynthetic Capabilities of Brassica napus. Procedia Environ. Sci. 29:255

Beule, L., and Karlovsky, P. 2020. Improved normalization of species count data in ecology by scaling with ranked subsampling (SRS): application to microbial communities. PeerJ. 8:e9593.

Bonk, F., Popp, D., Harms, H., and Centler, F. 2018. PCR-based quantification of taxa-specific abundances in microbial communities: Quantifying and avoiding common pitfalls. J. Microbiol. Methods. 153:139–147.

Burkart, S., Bender, J., Tarkotta, B., Faust, S., Castagna, A., Ranieri, A., et al. 2013. Effects of Ozone on Leaf Senescence, Photochemical Efficiency and Grain Yield in Two Winter Wheat Cultivars. J. Agron. Crop Sci. 199:275–285.

Burke, C., Steinberg, P., Rusch, D., Kjelleberg, S., and Thomas, T. 2011. Bacterial community assembly based on functional genes rather than species. Proc. Natl. Acad. Sci. 108:14288–14293

Caspi, R., Billington, R., Keseler, I. M., Kothari, A., Krummenacker, M., Midford, P. E., et al. 2020. The MetaCyc database of metabolic pathways and enzymes - a 2019 update. Nucleic Acids Res. 48:D445–D453

Chaudhry, V., Runge, P., Sengupta, P., Doehlemann, G., Parker, J. E., and Kemen, E. 2021. Shaping the leaf microbiota: plant-microbe-microbe interactions. J. Exp. Bot. 72:36–56

Chen, T., Nomura, K., Wang, X., Sohrabi, R., Xu, J., Yao, L., et al. 2020. A plant genetic network for preventing dysbiosis in the phyllosphere. Nature. 580:653–657

Cheng, C., Gao, X., Feng, B., Sheen, J., Shan, L., and He, P. 2013. Plant immune response to pathogens differs with changing temperatures. Nat. Commun. 4:2530

Chiang, K. S., Liu, H. I., Chen, Y. L., El Jarroudi, M., and Bock, C. H. 2020. Quantitative Ordinal Scale Estimates of Plant Disease Severity: Comparing Treatments Using a Proportional Odds Model. Phytopathology®. 110:734–743

Clauset, A., Newman, M. E. J., and Moore, C. 2004. Finding community structure in very large networks. Phys. Rev. E. 70:066111

Contijoch, E. J., Britton, G. J., Yang, C., Mogno, I., Li, Z., Ng, R., et al. 2019. Gut microbiota density influences host physiology and is shaped by host and microbial factors eds. Wendy S Garrett and Harry Sokol. eLife. 8:e40553

van Dam, N. M., and Heil, M. 2011. Multitrophic interactions below and above ground: en route to the next level. J. Ecol. 99:77–88

Debray, R., Socolar, Y., Kaulbach, G., Guzman, A., Hernandez, C. A., Curley, R., et al. 2022. Water stress and disruption of mycorrhizas induce parallel shifts in phyllosphere microbiome composition. New Phytol. 234:2018–2031

Dini-Andreote, F., de Cássia Pereira e Silva, M., Triadó-Margarit, X., Casamayor, E. O., van Elsas, J. D., and Salles, J. F. 2014. Dynamics of bacterial community succession in a salt marsh chronosequence: evidences for temporal niche partitioning. ISME J. 8:1989–2001

Dixon, P. 2003. VEGAN, a package of R functions for community ecology. J. Veg. Sci. 14:927–930

Dudney, J., Willing, C. E., Das, A. J., Latimer, A. M., Nesmith, J. C. B., and Battles, J. J. 2021. Nonlinear shifts in infectious rust disease due to climate change. Nat. Commun. 12:5102

Ehau-Taumaunu, H., and Hockett, K. 2022. Passaging phyllosphere microbial communities develop suppression towards bacterial speck disease in tomato. Phytobiomes J.

Faticov, M., Abdelfattah, A., Roslin, T., Vacher, C., Hambäck, P., Blanchet, F. G., et al. 2021. Climate warming dominates over plant genotype in shaping the seasonal trajectory of foliar fungal communities on oak. New Phytol. 231:1770–1783

Faust, K., and Raes, J. 2012. Microbial interactions: from networks to models. Nat. Rev. Microbiol. 10:538–550

Fürnkranz, M., Wanek, W., Richter, A., Abell, G., Rasche, F., and Sessitsch, A. 2008. Nitrogen fixation by phyllosphere bacteria associated with higher plants and their colonizing epiphytes of a tropical lowland rainforest of Costa Rica. ISME J. 2:561–570

Gao, M., Xiong, C., Gao, C., Tsui, C. K. M., Wang, M.-M., Zhou, X., et al. 2021. Disease-induced changes in plant microbiome assembly and functional adaptation. Microbiome. 9:187

Hall, M., and Beiko, R. G. 2018. 16S rRNA gene analysis with QIIME2. In *Microbiome analysis*, Springer, p. 113–129.

Hernandez, D. J., David, A. S., Menges, E. S., Searcy, C. A., and Afkhami, M. E. 2021. Environmental stress destabilizes microbial networks. ISME J. 15:1722–1734

Hindson, B. J., Ness, K. D., Masquelier, D. A., Belgrader, P., Heredia, N. J., Makarewicz, A. J., et al. 2011. High-Throughput Droplet Digital PCR System for Absolute Quantitation of DNA Copy Number. Anal. Chem. 83:8604–8610

Hoek, T. A., Axelrod, K., Biancalani, T., Yurtsev, E. A., Liu, J., and Gore, J. 2016. Resource Availability Modulates the Cooperative and Competitive Nature of a Microbial Cross-Feeding Mutualism. PLOS Biol. 14:e1002540

Horsfall, J. G., and Barratt, R. W. 1945. An improved grading system for measuring plant diseases. Phytopathology. 35:655

Hu, Q., Tan, L., Gu, S., Xiao, Y., Xiong, X., Zeng, W., et al. 2020. Network analysis infers the wilt pathogen invasion associated with non-detrimental bacteria. Npj Biofilms Microbiomes. 6:1–8

Hunter, P. J., Pink, D. A. C., and Bending, G. D. 2015. Cultivar-level genotype differences influence diversity and composition of lettuce (Lactuca sp.) phyllosphere fungal communities. Fungal Ecol. 17:183–186

Huot, B., Castroverde, C. D. M., Velásquez, A. C., Hubbard, E., Pulman, J. A., Yao, J., et al. 2017. Dual impact of elevated temperature on plant defence and bacterial virulence in Arabidopsis. Nat. Commun. 8:1808

Jian, C., Luukkonen, P., Yki-Järvinen, H., Salonen, A., and Korpela, K. 2020. Quantitative PCR provides a simple and accessible method for quantitative microbiota profiling. PLOS ONE. 15:e0227285

Jiang, C.-J., Shimono, M., Sugano, S., Kojima, M., Yazawa, K., Yoshida, R., et al. 2010. Abscisic Acid Interacts Antagonistically with Salicylic Acid Signaling Pathway in Rice–*Magnaporthe grisea* Interaction. Mol. Plant-Microbe Interactions®. 23:791–798

Kadivar, H., and Stapleton, A. E. 2003. Ultraviolet radiation alters maize phyllosphere bacterial diversity. Microb. Ecol. 45:353–361.

Karasov, T. L., and Lundberg, D. S. 2022. The changing influence of host genetics on the leaf fungal microbiome throughout plant development. PLOS Biol. 20:e3001748

Kemble, J. M., Bertucci, M. B., Jennings, K. M., Meadows, I. M., Rodrigues, C., Walgenbach, J. F., et al. 2022. *Southeastern US Vegetable Crop Handbook, 2022*. Sevew Group.

Klessig, D. F., Durner, J., Noad, R., Navarre, D. A., Wendehenne, D., Kumar, D., et al. 2000. Nitric oxide and salicylic acid signaling in plant defense. Proc. Natl. Acad. Sci. 97:8849–8855

Lebeis, S. L., Paredes, S. H., Lundberg, D. S., Breakfield, N., Gehring, J., McDonald, M., et al. 2015. Salicylic acid modulates colonization of the root microbiome by specific bacterial taxa. Science. 349:860–864

Lefohn, A. S., Malley, C. S., Smith, L., Wells, B., Hazucha, M., Simon, H., et al. 2018. Tropospheric ozone assessment report: Global ozone metrics for climate change, human health, and crop/ecosystem research eds. Detlev Helmig and Alastair Lewis. Elem. Sci. Anthr. 6:27

Leisner, C. P., Potnis, N., and Sanz-Saez, A. 2022. Crosstalk and trade-offs: Plant responses to climate change-associated abiotic and biotic stresses. Plant Cell Environ. n/a:1–18

Leveau, J. H. 2019. A brief from the leaf: latest research to inform our understanding of the phyllosphere microbiome. Curr. Opin. Microbiol. 49:41–49.

Li, P.-D., Zhu, Z.-R., Zhang, Y., Xu, J., Wang, H., Wang, Z., et al. 2022. The phyllosphere microbiome shifts toward combating melanose pathogen. Microbiome. 10:56

Lind, A. L., and Pollard, K. S. 2021. Accurate and sensitive detection of microbial eukaryotes from whole metagenome shotgun sequencing. Microbiome. 9:58 0168-021-01015-y [Accessed March 7, 2021]

Lindow, S. E., Arny, D. C., and Upper, C. D. 1982. Bacterial Ice Nucleation: A Factor in Frost Injury to Plants 1. Plant Physiol. 70:1084–1089

Lindow, S. E., and Brandl, M. T. 2003. Microbiology of the Phyllosphere. Appl. Environ. Microbiol. 69:1875–1883

Lu, J., Breitwieser, F. P., Thielen, P., and Salzberg, S. L. 2017. Bracken: estimating species abundance in metagenomics data. PeerJ Comput. Sci. 3:e104

Luo, L., Zhang, Z., Wang, P., Han, Y., Jin, D., Su, P., et al. 2019. Variations in phyllosphere microbial community along with the development of angular leaf-spot of cucumber. AMB Express. 9:76

Manching, H. C., Balint-Kurti, P. J., and Stapleton, A. E. 2014. Southern leaf blight disease severity is correlated with decreased maize leaf epiphytic bacterial species richness and the phyllosphere bacterial diversity decline is enhanced by nitrogen fertilization. Front. Plant Sci. 5:403

McMurdie, P. J., and Holmes, S. 2013. phyloseq: An R Package for Reproducible Interactive Analysis and Graphics of Microbiome Census Data ed. Michael Watson. PLoS ONE. 8:e61217

Mendes, L. W., Raaijmakers, J. M., de Hollander, M., Mendes, R., and Tsai, S. M. 2018. Influence of resistance breeding in common bean on rhizosphere microbiome composition and function. ISME J. 12:212–224

Mendes, R., Kruijt, M., de Bruijn, I., Dekkers, E., van der Voort, M., Schneider, J. H. M., et al. 2011. Deciphering the Rhizosphere Microbiome for Disease-Suppressive Bacteria. Science. 332:1097–1100

Meyer, K. M., Porch, R., Muscettola, I. E., Vasconcelos, A. L. S., Sherman, J. K., Metcalf, C. J. E., et al. 2022. Plant neighborhood shapes diversity and reduces interspecific variation of the phyllosphere microbiome. ISME J. 16:1376–1387.

Morella, N. M., Gomez, A. L., Wang, G., Leung, M. S., and Koskella, B. 2018. The impact of bacteriophages on phyllosphere bacterial abundance and composition. Mol. Ecol. 27:2025–2038

Muñoz-Elías, E. J., and McKinney, J. D. 2006. Carbon metabolism of intracellular bacteria. Cell. Microbiol. 8:10–22.

Newberry, E. A., Bhandari, R., Minsavage, G. V., Timilsina, S., Jibrin, M. O., Kemble, J., et al. 2019. Independent evolution with the gene flux originating from multiple Xanthomonas

species explains genomic heterogeneity in Xanthomonas perforans. Appl. Environ. Microbiol. 85:885–904

Nogales, J., Canales, Á., Jiménez-Barbero, J., Serra, B., Pingarrón, J. M., García, J. L., et al. 2011. Unravelling the gallic acid degradation pathway in bacteria: the gal cluster from Pseudomonas putida. Mol. Microbiol. 79:359–374

O'Leary, N. A., Wright, M. W., Brister, J. R., Ciufo, S., Haddad, D., McVeigh, R., et al. 2016. Reference sequence (RefSeq) database at NCBI: current status, taxonomic expansion, and functional annotation. Nucleic Acids Res. 44:D733-745.

Omae, N., and Tsuda, K. 2022. Plant-Microbiota Interactions in Abiotic Stress Environments. Mol. Plant-Microbe Interactions®. 35:511–526

Ortiz-Álvarez, R., Ortega-Arranz, H., Ontiveros, V. J., de Celis, M., Ravarani, C., Acedo, A., et al. 2021. Network Properties of Local Fungal Communities Reveal the Anthropogenic Disturbance Consequences of Farming Practices in Vineyard Soils. mSystems. 6:e00344-21

Pellegrini, E., Trivellini, A., Campanella, A., Francini, A., Lorenzini, G., Nali, C., et al. 2013. Signaling molecules and cell death in Melissa officinalis plants exposed to ozone. Plant Cell Rep. 32:1965–1980

Peschel, S., Müller, C. L., von Mutius, E., Boulesteix, A.-L., and Depner, M. 2021. NetCoMi: network construction and comparison for microbiome data in R. Brief. Bioinform. 22:bbaa290

Pfister, G. G., Walters, S., Lamarque, J.-F., Fast, J., Barth, M. C., Wong, J., et al. 2014. Projections of future summertime ozone over the U.S. J. Geophys. Res. Atmospheres. 119:5559–5582

Remus-Emsermann, M. N. P., and Schlechter, R. O. 2018. Phyllosphere microbiology: at the interface between microbial individuals and the plant host. New Phytol. 218:1327–1333.

Ren, G., Zhang, H., Lin, X., Zhu, J., and Jia, Z. 2015. Response of leaf endophytic bacterial community to elevated CO2 at different growth stages of rice plant. Front. Microbiol. 6:855

Reyes, A., Wu, M., McNulty, N. P., Rohwer, F. L., and Gordon, J. I. 2013. Gnotobiotic mouse model of phage–bacterial host dynamics in the human gut. Proc. Natl. Acad. Sci. 110:20236–20241

Rico, L., Ogaya, R., Terradas, J., and Peñuelas, J. 2014. Community structures of N2-fixing bacteria associated with the phyllosphere of a Holm oak forest and their response to drought. Plant Biol. 16:586–593

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

Rizhsky, L., Liang, H., and Mittler, R. 2002. The Combined Effect of Drought Stress and Heat Shock on Gene Expression in Tobacco. Plant Physiol. 130:1143–1151

Rizhsky, L., Liang, H., Shuman, J., Shulaev, V., Davletova, S., and Mittler, R. 2004. When defense pathways collide. The response of Arabidopsis to a combination of drought and heat stress. Plant Physiol. 134:1683–1696.

Schiff, J. A. 1980. Pathways of assimilatory sulphate reduction in plants and microorganisms. Sulphur Biol. 72:49–64.

Segata, N., Izard, J., Waldron, L., Gevers, D., Miropolsky, L., Garrett, W. S., et al. 2011. Metagenomic biomarker discovery and explanation. Genome Biol. 12:R60.

Stall, R. E., Jones, J. B., and Minsavage, G. V. 2009. Durability of Resistance in Tomato and Pepper to Xanthomonads Causing Bacterial Spot. Annu. Rev. Phytopathol. 47:265–284

Su, P., Wicaksono, W. A., Li, C., Michl, K., Berg, G., Wang, D., et al. 2022. Recovery of metagenome-assembled genomes from the phyllosphere of 110 rice genotypes. Sci. Data. 9:254

Suzek, B. E., Wang, Y., Huang, H., McGarvey, P. B., Wu, C. H., and the UniProt Consortium. 2015. UniRef clusters: a comprehensive and scalable alternative for improving sequence similarity searches. Bioinformatics. 31:926–932

Suzuki, N., Rivero, R. M., Shulaev, V., Blumwald, E., and Mittler, R. 2014. Abiotic and biotic stress combinations. New Phytol. 203:32–43

Tan, L., Xiao, Y., Zeng, W., Gu, S., Zhai, Z., Wu, S., et al. 2022. Network analysis reveals the root endophytic fungi associated with Fusarium root rot invasion. Appl. Soil Ecol. 178:104567

Tao, S., Yin, H., Fang, Y., Zhang, Y., Zhang, N., and Qu, L. 2023. Elevated O3 concentrations alter the compartment-specific microbial communities inhabiting rust-infected poplars. Environ. Microbiol.

Team, R. C. 2022. R: A language and environment for statistical computing.

Team, Rs. 2020. RStudio: Integrated Development for R. http://www.rstudio.com/

Temple, P. J., and Bisessar, S. 1979. Response of White bean to Bacterial Blight, Ozone, and Antioxidant Protection in the Field. Phytopathology. 69:101–103

Trivedi, P., Batista, B. D., Bazany, K. E., and Singh, B. K. 2022. Plant–microbiome interactions under a changing world: responses, consequences and perspectives. New Phytol. 234:1951–1959

Vannier, N., Agler, M., and Hacquard, S. 2019. Microbiota-mediated disease resistance in plants. PLOS Pathog. 15:e1007740

Velásquez, A. C., Castroverde, C. D. M., and He, S. Y. 2018. Plant-Pathogen Warfare under Changing Climate Conditions. Curr. Biol. CB. 28:R619–R634

Vincent, S. A., Ebertz, A., Spanu, P. D., and Devlin, P. F. 2022. Salicylic Acid-Mediated Disturbance Increases Bacterial Diversity in the Phyllosphere but Is Overcome by a Dominant Core Community. Front. Microbiol. 13:809940

Vingarzan, R. 2004. A review of surface ozone background levels and trends. Atmos. Environ. 38:3431–3442

Vogel, C. M., Potthoff, D. B., Schäfer, M., Barandun, N., and Vorholt, J. A. 2021. Protective role of the Arabidopsis leaf microbiota against a bacterial pathogen. Nat. Microbiol. 6:1537–1548

Wagg, C., Schlaeppi, K., Banerjee, S., Kuramae, E. E., and van der Heijden, M. G. A. 2019. Fungal-bacterial diversity and microbiome complexity predict ecosystem functioning. Nat. Commun. 10:4841

Wagner, M. R., Busby, P. E., and Balint-Kurti, P. 2020. Analysis of leaf microbiome composition of near-isogenic maize lines differing in broad-spectrum disease resistance. New Phytol. 225:2152–2165

Wagner, M. R., Lundberg, D. S., del Rio, T. G., Tringe, S. G., Dangl, J. L., and Mitchell-Olds, T. 2016. Host genotype and age shape the leaf and root microbiomes of a wild perennial plant. Nat. Commun. 7:12151

van der Wal, A., and Leveau, J. H. J. 2011. Modelling sugar diffusion across plant leaf cuticles: the effect of free water on substrate availability to phyllosphere bacteria. Environ. Microbiol. 13:792–797

Wickham, Hadley. 2016. *ggplot2: Elegant Graphics for Data Analysis*. New York, NY: Springer.

Wood, D. E., Lu, J., and Langmead, B. 2019. Improved metagenomic analysis with Kraken 2. Genome Biol. 20:257

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

Yasuda, M., Ishikawa, A., Jikumaru, Y., Seki, M., Umezawa, T., Asami, T., et al. 2008. Antagonistic Interaction between Systemic Acquired Resistance and the Abscisic Acid–Mediated Abiotic Stress Response in Arabidopsis. Plant Cell. 20:1678–1692

Ye, S. H., Siddle, K. J., Park, D. J., and Sabeti, P. C. 2019. Benchmarking Metagenomics Tools for Taxonomic Classification. Cell. 178:779–794

Yoon, G., Gaynanova, I., and Müller, C. L. 2019. Microbial networks in SPRING-Semi-parametric rank-based correlation and partial correlation estimation for quantitative microbiome data. Front. Genet. 10:516.

Yoshida, S., Hiradate, S., Koitabashi, M., Kamo, T., and Tsushima, S. 2017. Phyllosphere Methylobacterium bacteria contain UVA-absorbing compounds. J. Photochem. Photobiol. B. 167:168–175

Zandalinas, S. I., Sengupta, S., Fritschi, F. B., Azad, R. K., Nechushtai, R., and Mittler, R. 2021. The impact of multifactorial stress combination on plant growth and survival. New Phytol. 230:1034–1048

Zhou, Y., Van Leeuwen, S. K., Pieterse, C. M. J., Bakker, P. A. H. M., and Van Wees, S. C. M. 2019. Effect of atmospheric CO2 on plant defense against leaf and root pathogens of Arabidopsis. Eur. J. Plant Pathol. 154:31–42

Zhu, J.-K. 2016. Abiotic stress signaling and responses in plants. Cell. 167:313–324 ber 22, 2022]

.

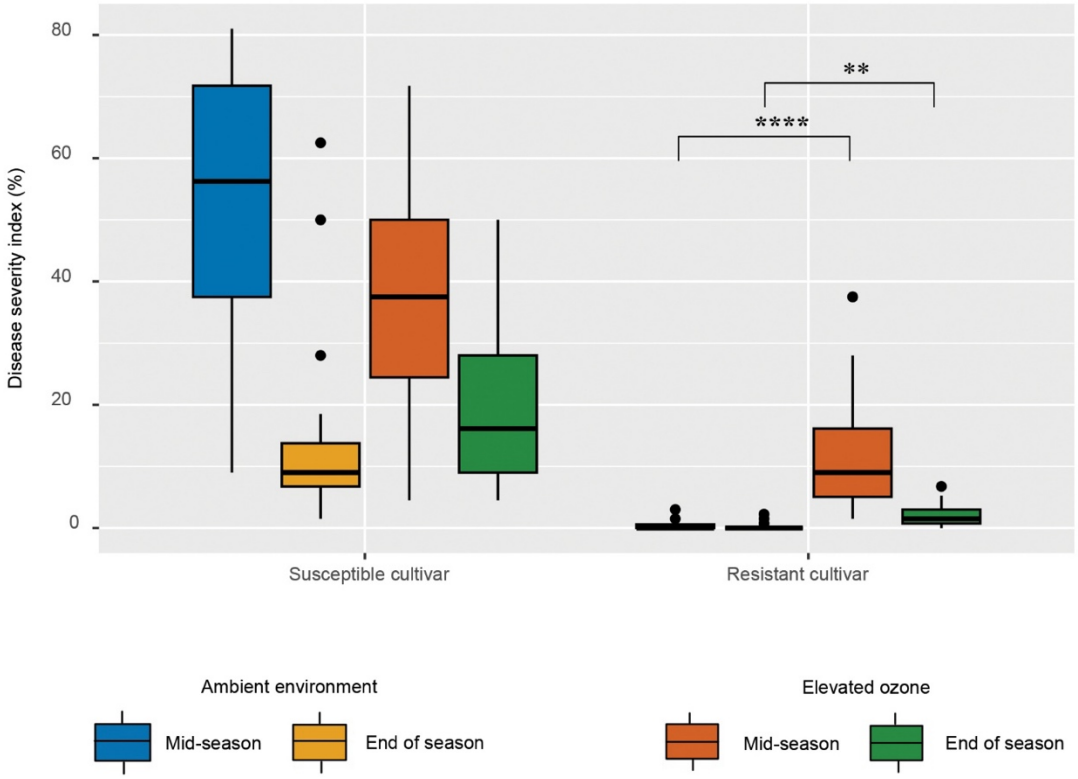

Figure 2-1: Elevated O_3 exacerbates bacterial spot disease severity on the resistant cultivar but has no effect on the susceptible cultivar. Box and whisker plots showing the disease severity index (represented as % value) under elevated O_3 and ambient environmental conditions across susceptible and resistant cultivars. Significance levels for each of the treatment combination are indicated by *p < 0.05; **p < 0.01; ***p < 0.001; ****p < 0.0001.

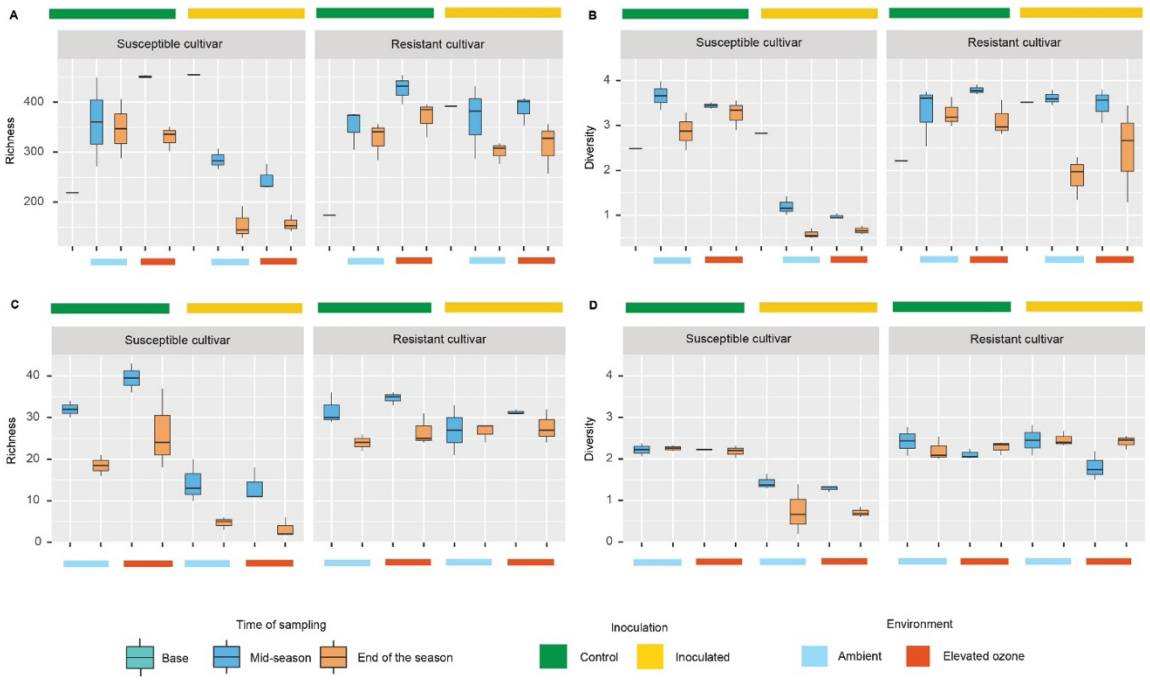

Figure 2-2: Elevated O_3 has little impact on microbial diversity and richness. However, pathogen infection on susceptible cultivar reduces microbial community richness and diversity. (A) Bacterial Chao1 richness and (B) bacterial Shannon diversity index across different environments. (C)Eukaryotic community diversity and (D) richness across different treatments. Inoculated and control samples are indicated with yellow and green bars on the top, while ambient and elevated O_3 treatments are denoted by light blue and red color bars at the bottom, respectively.

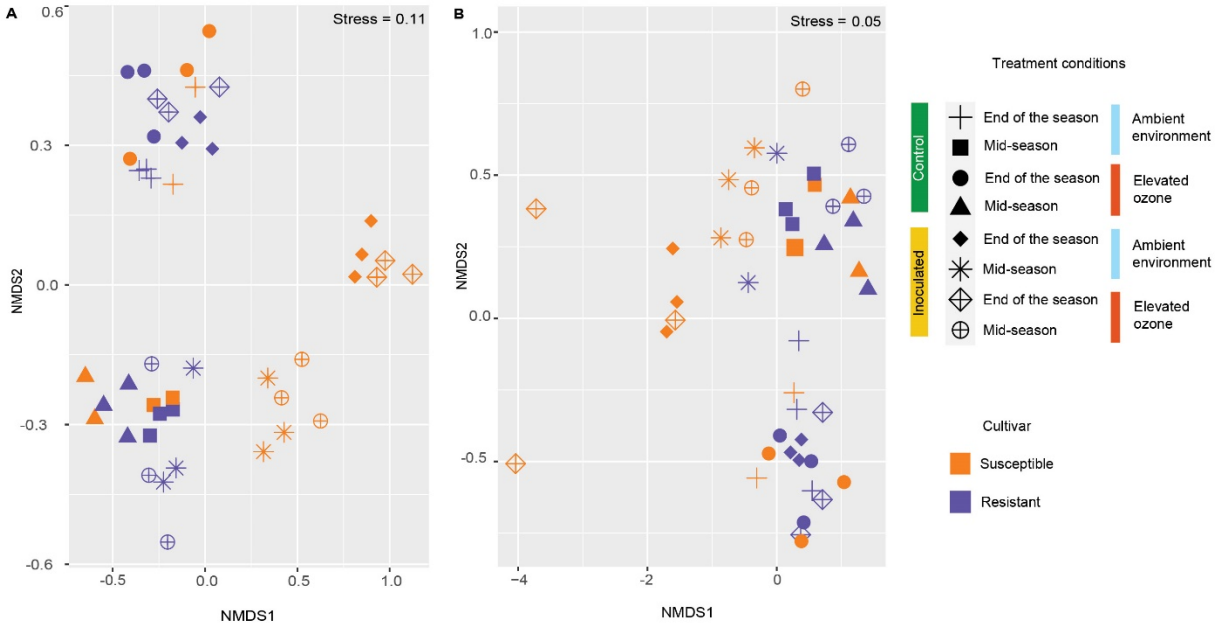

Figure 2-3: Elevated O_3 changes microbial community structure on susceptible cultivars challenged with pathogen infection, but not on resistant cultivars. (A) Nonmetric Multidimensional Scaling (NMDS) ordination comparing the bacterial community diversity across two cultivars, environmental conditions, and time of sampling. (B) NMDS ordination comparing the eukaryotic community diversity across two cultivars, environmental conditions, and time of sampling.

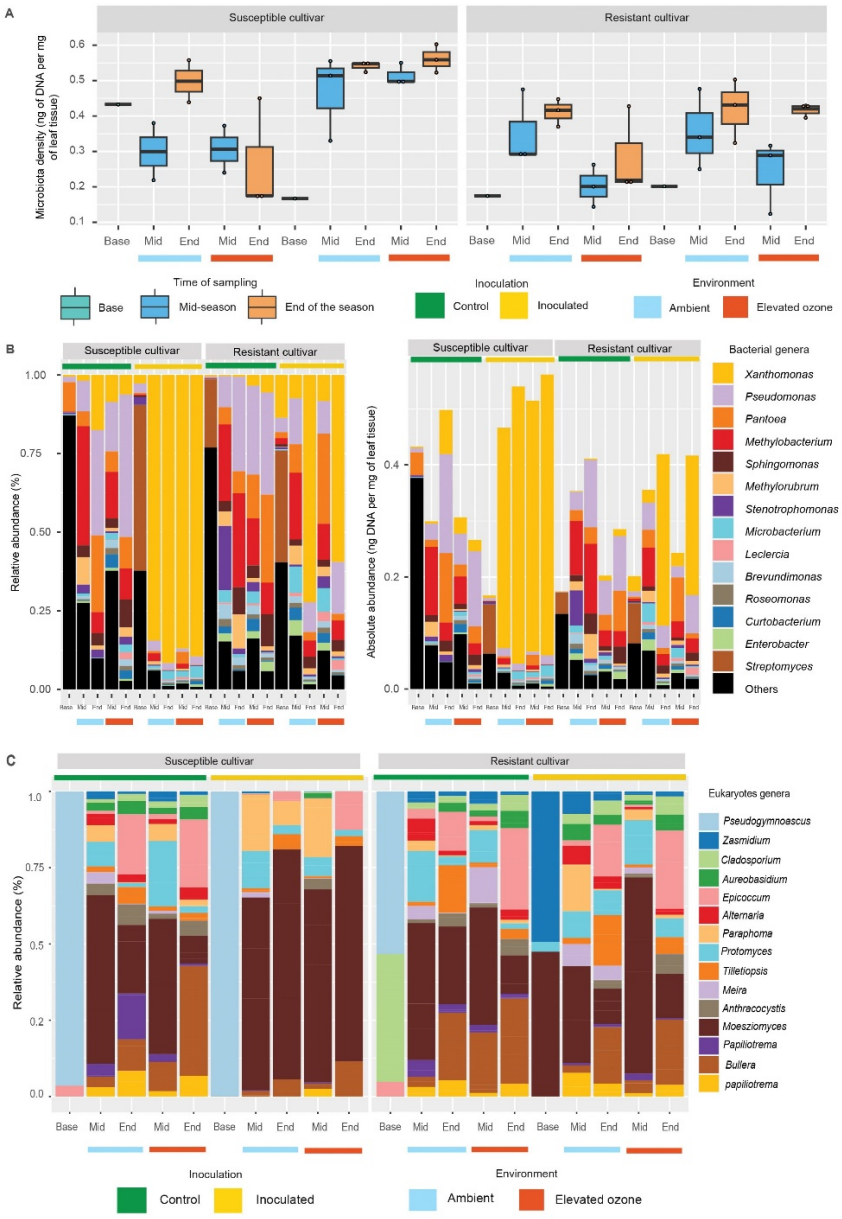

Figure 2-4: The effects of elevated O_3 on disease outcomes are not fully explained by changes in microbiota density and abundance. (A) Box and whisker plot showing microbiota density estimated by microbial DNA quantification (concentration of extracted DNA per mg of leaf samples) for various treatment in two cultivars. (B) Relative (Left) and absolute (right) species abundance of top 15 bacterial taxa across samples. Absolute abundance is obtained by scaling the relative abundance measurements by the microbiota density measurements. (C) Bar plots showing the relative abundance of the top 15 eukaryotic genera across the samples. Inoculated and control samples are indicated with yellow and green bars on the top, while ambient and elevated O_3 treatments are denoted by light blue and red color bars at the bottom, respectively. The time of sampling is indicated by Base (initial samples), Mid (mid-season), and End (end of the season).

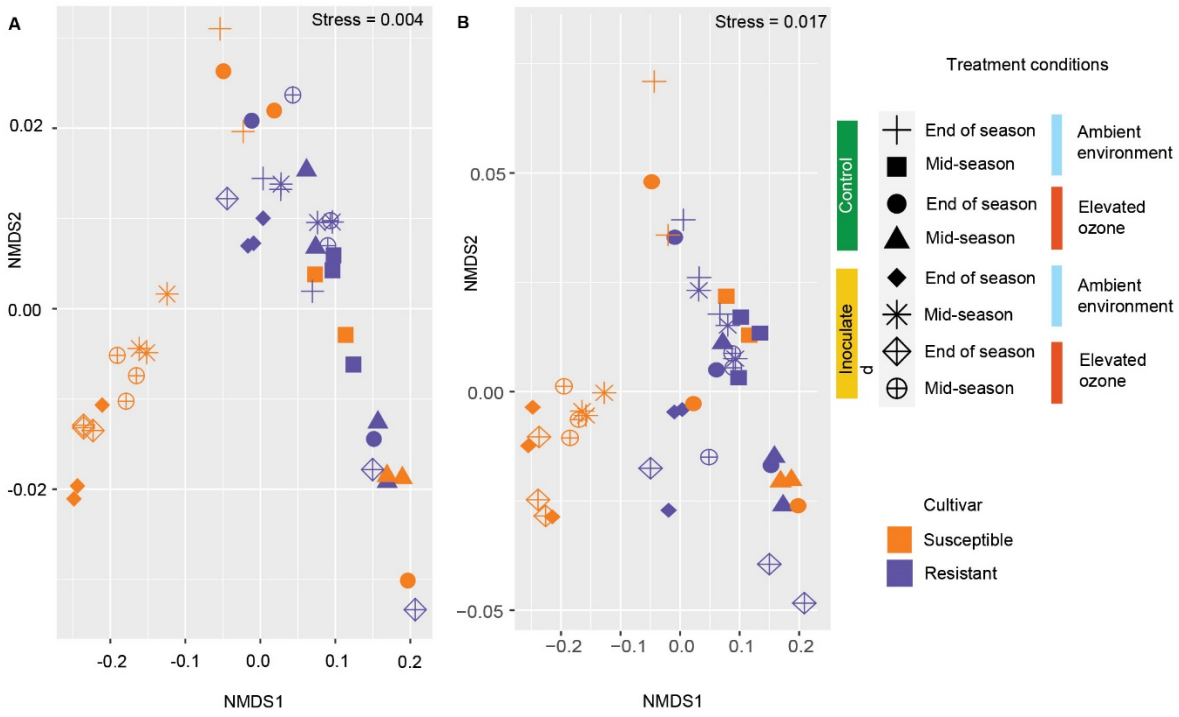

Figure 2-5: Microbial community functions were affected by host susceptibility to pathogens, while elevated O_3 had little impact. Nonmetric Multidimensional Scaling (NMDS) ordination displaying diversity in (A) metabolic pathways across different treatment conditions in susceptible and resistant cultivars, (B) genes mapped to metabolic pathways across various treatment conditions in susceptible and resistant cultivars.

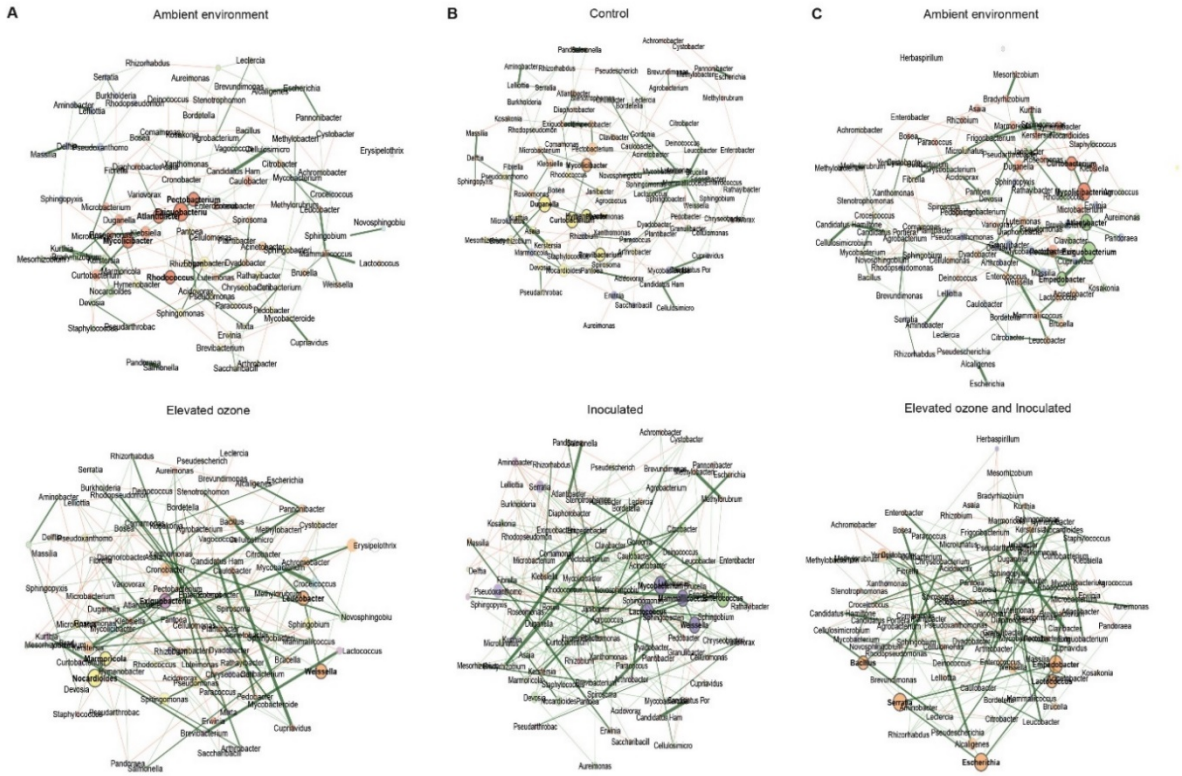

Figure 2-6: Pathogen infection is associated with microbial communities showing positive and stable interactions, but these interactions are random and less predictable with a shift in hub taxa in response to concurrent O_3 stress and pathogen infection. Comparison of bacterial association network across different environments. (A) Bacterial association network for the combined data set of ambient (top) and elevated O_3 (bottom) in both cultivars under control conditions. (B) Bacterial association network for the combined data set from control (top) and inoculated (bottom) samples from both the cultivars under ambient environment. (C) Bacterial association network for the combined data set from control and ambient environment (top) and inoculated and elevated O_3 (bottom) samples from both cultivars. Hub taxa are highlighted by bold text. Node color represents the cluster determined by greedy modularity optimization. Red edges correspond to negative correlations, while green edges correspond to positive correlations.

2.8 Supplementary results

2.8.1 Taxonomic profiling

A temporal pattern was observed in that, *Actinobacteria* were the most abundant (~57%) in the base samples, and their abundance decreased over time both in inoculated (mid-season: 8 %; end of the season: 1%) and control plants (mid-season: 9 %; end of the season: 3%) shifting to Proteobacteria which was dominant both at the mid and end of the season samples (Figure S5). Phylum Bacteroidetes, although accounting for around 3% in the base samples, decreased in abundance to less than 1% at later time points and were negligible on the inoculated susceptible leaves. Phylum Firmicutes contributed to around 2% of the sequences during the mid-season. They were in higher abundance under elevated O_3 conditions on the leaves of the control susceptible cultivar and in both inoculated and control plants in the resistant cultivar. This phylum was affected by inoculation on susceptible plants and was negligible across all the treatment conditions during the end of the season.

We next examined the deeper taxonomic ranks, specifically, genus, for further comparison of patterns affected by the presence of the pathogen and elevated O_3 or a combination of both stresses. Susceptible control plants were dominated mainly by *Pseudomonas* (~5-60%), *Pantoea* (~3-28%), and *Sphingomonas* (3-15%), while *Xanthomonas* (~84-98%) and *Pseudomonas* (~1-4%) were the dominant genera in the inoculated plants. In resistant cultivar, *Pseudomonas* (6-40%), *Pantoea* (~2-44%), and *Methylobacterium* (~9-35%) were the dominant genera in control plants, while *Xanthomonas* (~2-87%), *Pseudomonas* (~2-27%), and *Methylobacterium* (~1-26%) were dominant on inoculated plants (Fig. 4B, Table S8). Other genera contributing to the phyllosphere microbiome included *Pseudomonas*, *Pantoea*, *Methylobacterium*, *Sphingomonas*,

Methylorubrum, *Stenotrophomonas,* and *Microbacterium*. The abundance of *Pseudomonas* increased over the growing season, except for resistant control plants, where abundance levels were higher during the mid-season under elevated O_3 and stayed at similar levels throughout the growing season. *Pseudomonas* abundance was affected by inoculation, where a decrease in the abundance of *Pseudomonas* was proportional to the increase in *Xanthomonas* abundance on either cultivar. The mean relative abundance of *Pantoea* increased over time in control plants of either cultivar. The inoculation shifted the temporal pattern with a decreased mean relative abundance of *Pantoea* from mid to end of the season sampling. This decrease was statistically significant under elevated O_3 on inoculated samples of either cultivar (resistant cultivar: p = 0.031; susceptible cultivar: p = 0.007). Inoculation also had a slight effect on the mean relative abundance of *Methylobacterium* with a significant decrease with inoculation under ambient environment (p = 0.005), under elevated O_3 in the resistant cultivar (Kruskal-Wallis, p = 0.049), and ambient environment (p = 0.022) in the susceptible cultivar. An increase in O_3 concentration has a negative effect on the *Methylobacterium* genus under control conditions and ambient environment (resistant cultivar: p = 0.009, susceptible cultivar: p= 0.025). However, the environment did not affect the *Methylobacterium* genus in inoculated samples in both cultivars (Fig. S3).

Analysis of the eukaryotes diversity across our samples showed that the genera *Moesziomyces*, *Golubevia*, *Paraphoma, Protomyces,* and *Cercospora* were dominant across the samples (Fig. 4C, Table S9). Base samples were dominated by the genus *Pseudogymnoascus* (~64%), which was not observed during later time points. Genus *Bullera* had a higher relative abundance under elevated O_3 both in control (~ 32%) and inoculated plants (~16%) than that of ambient environment (~ 15% in control and 12% in inoculated) during the end of the season in both the

cultivars. Inoculation resulted in lower *Zasmidium*, *Cladosporium*, *Aureobasidium*, *Alternaria,* and *Anthracocystis* genera in susceptible cultivars compared to resistant ones under both environmental conditions. Regardless of treatments, the relative abundance of the genus *Epicoccum* increased at the end of the season (~ 31%) compared to mid-season (~ 4%). In comparison, the genus *Protomyces* decreased during the end of the season (~ 3%) compared to mid-season (~ 11%). Inoculation did not affect the relative abundance of *Moesziomyces* in susceptible cultivars during both mid (~ 60%) and end of the season (~70%). In comparison, the presence of *Xanthomonas* had a negative effect on this genus during the end of the season (~ 12%) in resistant cultivars.

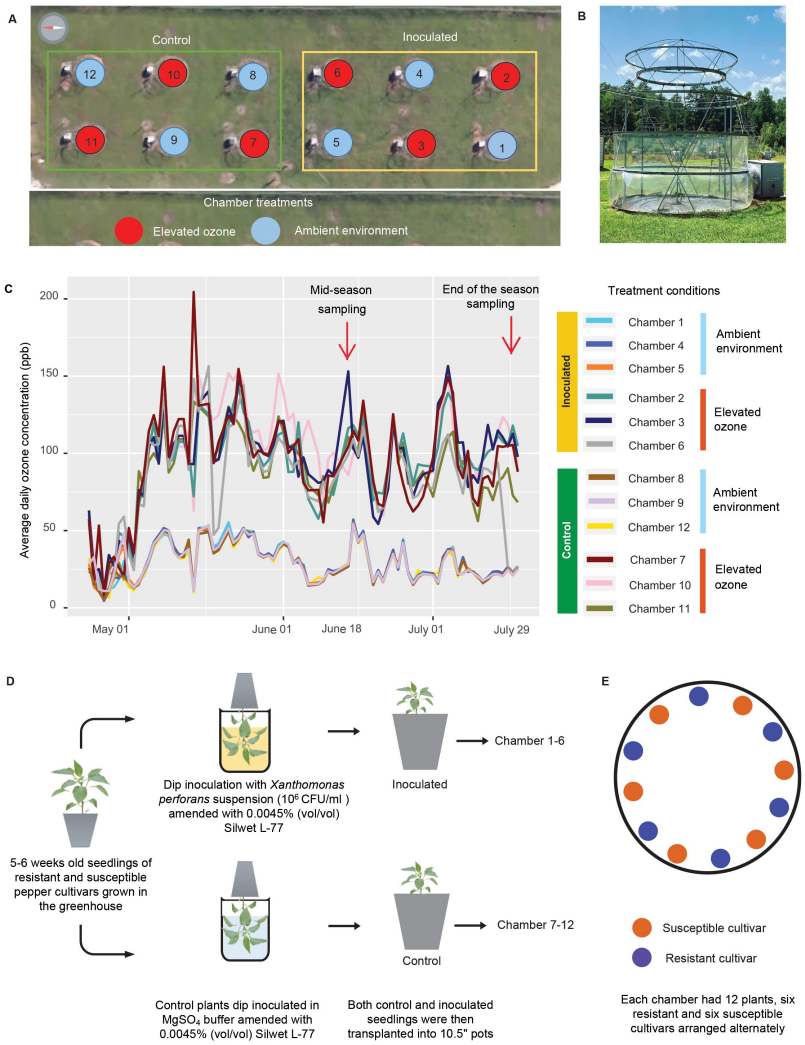

Figure S2-1: Study site and treatment design of Atmospheric Deposition Laboratory (AtDep) site at Auburn University. (A) Satellite image of AtDep site with individual chambers label where light blue circles are the chambers with the ambient environment and red circles are the chambers with elevated O_3. Chambers 1-6 marked with yellow color are inoculated with *X. perforans,* and chambers 7-12, with green color are control samples. (B) Individual open-top chamber at the AtDep site (4 x 5m). (C) Daily average O_3 concentration in treatment chambers. Sampling points (mid and end of the season) are marked by a red arrow. (D) Inoculation method showing how plants were inoculated with *Xanthomonas* and control plants with $MgSO_4$ buffer. (E) The plants were then kept inside the OTCs where each chamber had 12 plants (6 each from susceptible and resistant cultivar). Inoculated and control plants were kept in different chambers as per the treatment plan in A.

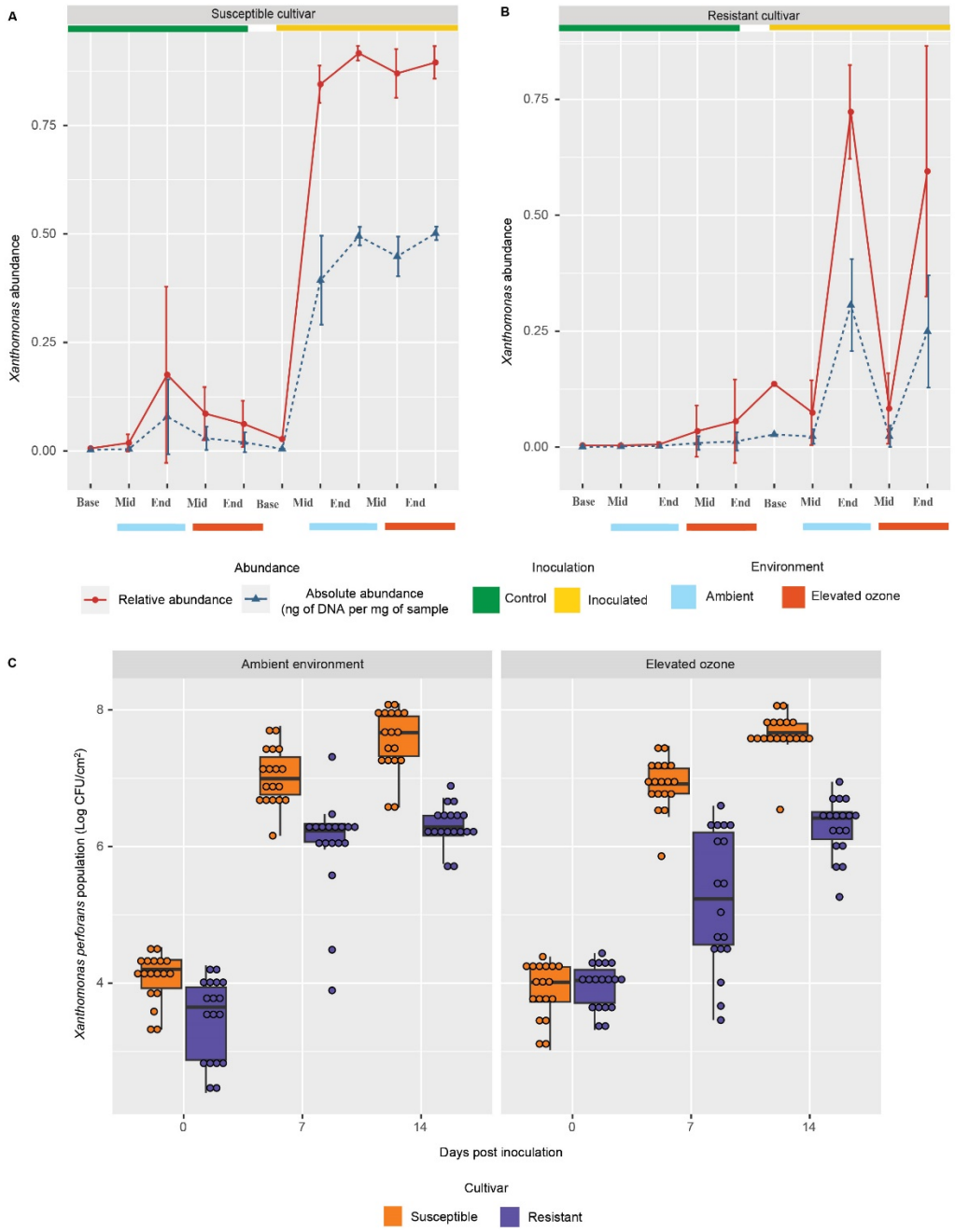

Figure S2-2: Elevated O_3 increases heterogeneity in *Xanthomonas* population counts on the resistant cultivar. Line graph showing the relative and absolute abundance (ng of DNA per mg of leaf tissue) (A) *Xanthomonas* relative and absolute abundance (ng of DNA per mg of leaf tissue) across different treatments in susceptible cultivars. (B) *Xanthomonas* relative and absolute abundance (ng of DNA per mg of leaf tissue) across different treatments in resistant cultivar. end of the season. (C) *Xanthomonas* population growth based on culture-dependent technique on two different cultivars under ambient environment and elevated O_3.

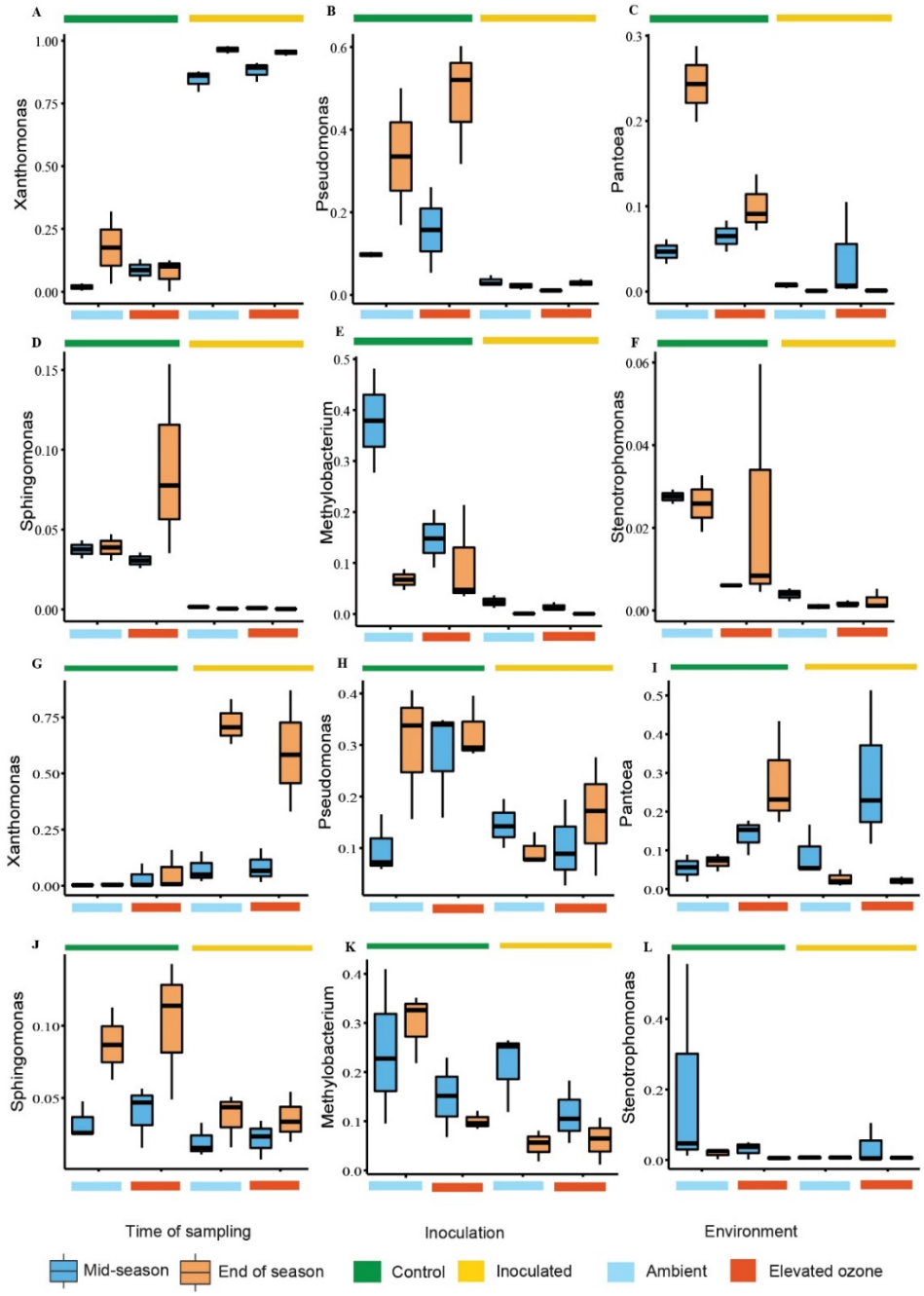

Figure S2-3: Box plot showing the relative abundance of top 6 bacterial genera. (A-F) Relative abundance of different bacterial genera on susceptible cultivar across different treatments; (G-L) Relative abundance of different bacterial genera on resistant cultivar across different treatments. Control samples are indicated by a green bar, while the yellow bar represents the inoculated samples. The blue box represents the samples taken during the mid-season, while the orange box is end of the season samples. Different environmental conditions are represented by the ambient environment and elevated O_3 under both inoculation and control treatments.

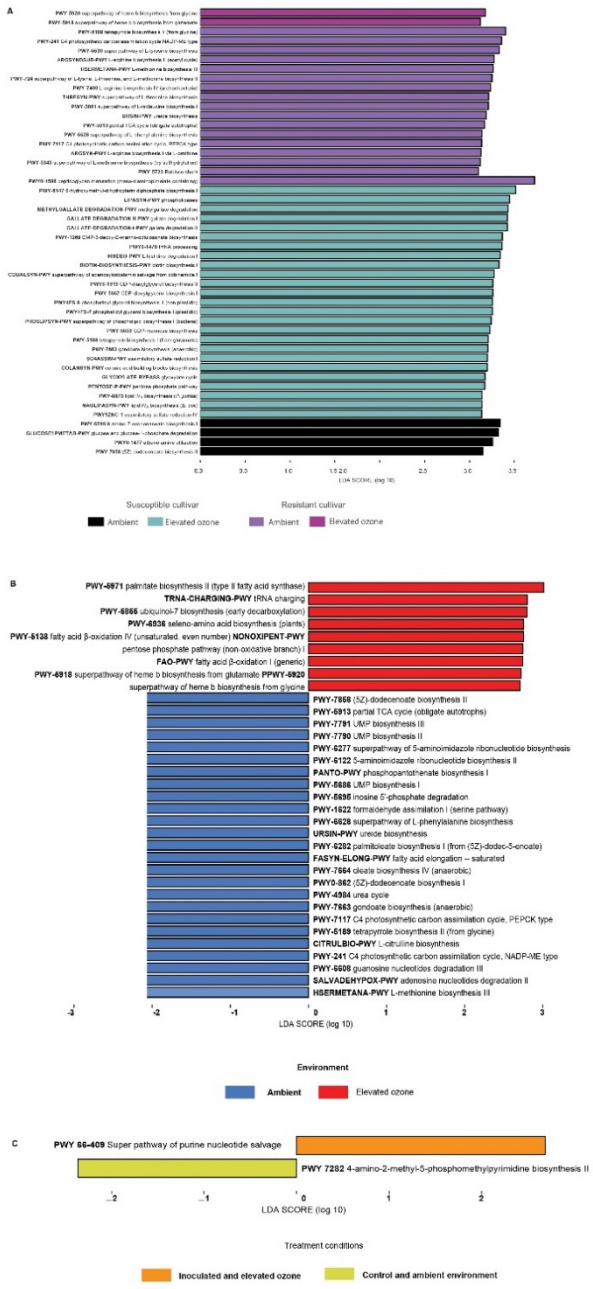

Figure S2-4: Linear Discriminant Analysis Effect Size (LEfSe) of KEGG pathways between susceptible and resistant cultivar. Results were ranked by their Linear Discriminant Analysis (LDA) score. An FDR-adjusted *p*-value ≤ 0.05, as well as an LDA score ≥2.5, were used as thresholds to identify significant features. (A) pathways enriched upon inoculation of the resistant cultivar under elevated O_3 (brown), resistant cultivar under ambient environment (violet), susceptible cultivar under elevated O_3 (cyan), and susceptible cultivar under ambient environment (black). (B) pathways enriched in both cultivars under elevated O_3 (red) and ambient environment (blue), (C) pathways enriched in both the cultivars under combined stress of pathogen infection and elevated O_3 (orange) and ambient environment (lime green).

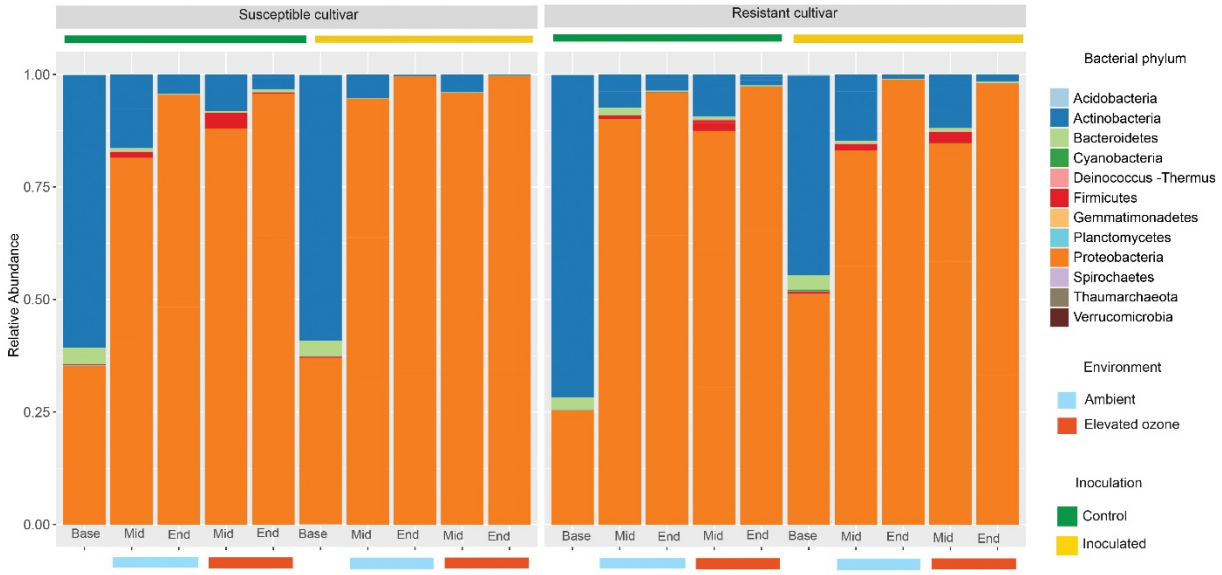

Figure S2-5: Stacked bar plots showing the relative abundance of dominant phyla across the samples. Control samples are indicated by a green bar, while the yellow bar represents the inoculated samples. The time of sampling is indicated by Base (initial samples), Mid (mid-season) and End (end of the season).

Table S2-1: Multiple pairwise comparisons (Dunn test) of disease severity index across different cultivars and treatment conditions. Significance levels for each treatment combination are indicated by *$p < 0.05$; **$p < 0.01$; ***$p < 0.001$; ****$p < 0.0001$.

Cultivar	Comparison	Standard test statistics	*p* value	Adjusted *p*-value	significance
Susceptible	Ambient environment and end of the season vs. Ambient environment and mid-season	4.9074802	9.23E-07	5.54E-06	****
Susceptible	Ambient environment and end of the season vs. Elevated ozone and end of the season	0.8719077	3.83E-01	3.83E-01	
Susceptible	Ambient environment and end of the season vs. Elevated ozone and mid-season	3.4676327	5.25E-04	1.05E-03	**
Susceptible	Ambient environment and mid-season vs. Elevated ozone and end of the season	-4.0355726	5.45E-05	1.63E-04	***
Susceptible	Ambient environment and mid-season vs. Elevated ozone and mid-season	-1.4398475	1.50E-01	1.80E-01	
Susceptible	Elevated ozone and end of the season vs. Elevated ozone and mid-season	2.5957251	9.44E-03	1.42E-02	*
Resistant	Ambient environment and end of the season vs. Ambient environment and mid-season	0.19131	8.48E-01	8.48E-01	
Resistant	Ambient environment and end of the season vs. Elevated ozone and end of the season	2.7448823	6.05E-03	9.08E-03	**
Resistant	Ambient environment and end of the season vs. Elevated ozone and mid-season	6.0969658	1.08E-09	6.49E-09	****
Resistant	Ambient environment and mid-season vs. Elevated ozone and end of the season	2.5535723	1.07E-02	1.28E-02	*
Resistant	Ambient environment and mid-season vs. Elevated ozone and mid-season	5.9056558	3.51E-09	1.05E-08	****
Resistant	Elevated ozone and end of the season vs. Elevated ozone and mid-season	3.3520835	8.02E-04	1.60E-03	**

Table S2-2: Sample details and sequencing statistics of metagenome reads.

Sample id	Cultivar	Inoculation status	Time of sampling	Environment	DNA conc. (ng/µl)	Total raw reads	Total raw bases	Reads after quality trimming	Reads after host trimming	% of final reads for further analysis	Gbp of reads	% lost in quality trimming	% lost in host trimming
1EM	Susceptible	Inoculated	Mid-season	Ambient	44.4	94632028	14012568232	82555398	76943654	81.31	14.01	12.76	6.80
1EE	Susceptible	Inoculated	End of the season	Ambient	43.8	71879368	10853784568	63532100	59509512	82.79	10.85	11.61	6.33
1XM	Resistant	Inoculated	Mid-season	Ambient	27.2	86444924	13053183524	75569647	70279794	81.30	13.05	12.58	7.00
1XE	Resistant	Inoculated	End of the season	Ambient	34.5	83058106	12541774006	73320424	68466214	82.43	12.54	11.72	6.62
2EM	Susceptible	Inoculated	Mid-season	Elevated ozone	44	83619784	12626587384	73105423	68279082	81.65	12.63	12.57	6.60
2EE	Susceptible	Inoculated	End of the season	Elevated ozone	48.2	76613288	11568606488	68214364	64256044	83.87	11.57	10.96	5.80
2XM	Resistant	Inoculated	Mid-season	Elevated ozone	23.1	67954684	10261157284	58945911	54317988	79.93	10.26	13.26	7.85
2XE	Resistant	Inoculated	End of the season	Elevated ozone	31.6	85486336	12908436736	74113643	61375640	71.80	12.91	13.30	17.19
3EM	Susceptible	Inoculated	Mid-season	Elevated ozone	39.5	78103064	11793562664	68629726	64244500	82.26	11.79	12.13	6.39
3EE	Susceptible	Inoculated	End of the season	Elevated ozone	44.7	82463524	12451992124	72822036	68362944	82.90	12.45	11.69	6.12
3XM	Resistant	Inoculated	Mid-season	Elevated ozone	9.88	72073444	10883090044	63021647	58620704	81.33	10.88	12.56	6.98
3XE	Resistant	Inoculated	End of the season	Elevated ozone	34.8	107486130	16230405630	96211494	84545768	78.66	16.23	10.49	12.13
4EM	Susceptible	Inoculated	Mid-season	Ambient	26.4	92419428	13955333628	81432726	76535500	82.81	13.96	11.89	6.01
4EE	Susceptible	Inoculated	End of the season	Ambient	43.9	60405554	9121238654	53906544	50787970	84.08	9.12	10.76	5.79
4XM	Resistant	Inoculated	Mid-season	Ambient	38.1	85511946	12912303846	74096844	68386118	79.97	12.91	13.35	7.71
4XE	Resistant	Inoculated	End of the season	Ambient	40.2	95167264	14370256864	82585708	72069602	75.73	14.37	13.22	12.73
5EM	Susceptible	Inoculated	Mid-season	Ambient	41.1	86126728	13005135928	75442752	70497018	81.85	13.01	12.40	6.56
5EE	Susceptible	Inoculated	End of the season	Ambient	41.9	69052192	10426880992	59171173	54881300	79.48	10.43	14.31	7.25
5XM	Resistant	Inoculated	Mid-season	Ambient	20	87321158	13185494858	76345143	71081610	81.40	13.19	12.57	6.89
5XE	Resistant	Inoculated	End of the season	Ambient	25.9	92131658	13911880358	80518352	73696904	79.99	13.91	12.61	8.47
6EM	Susceptible	Inoculated	Mid-season	Elevated ozone	39.8	83038324	12538786924	72028537	66836114	80.49	12.54	13.26	7.21
6EE	Susceptible	Inoculated	End of the season	Elevated ozone	41.8	85948910	12978285410	74964898	69762754	81.17	12.98	12.78	6.94
6XM	Resistant	Inoculated	Mid-season	Elevated ozone	25.3	72334912	10922571712	62846737	58011850	80.20	10.92	13.12	7.69
6XE	Resistant	Inoculated	End of the season	Elevated ozone	33.6	43692494	6597566594	38063184	29469316	67.45	6.60	12.88	22.58
7EE	Susceptible	Control	End of the season	Elevated ozone	36	80783162	12198257462	70281064	64307588	79.61	12.20	13.00	8.50
7XM	Resistant	Control	Mid-season	Elevated ozone	11.5	82013696	12384068096	73704235	69351830	84.56	12.38	10.13	5.91
7XE	Resistant	Control	End of the season	Elevated ozone	34.2	92715106	13999981006	82266875	74915476	80.80	14.00	11.27	8.94
8EE	Susceptible	Control	End of the season	Ambient	35.1	93149022	14065502322	82364507	76748744	82.39	14.07	11.58	6.82
8XM	Resistant	Control	Mid-season	Ambient	23.4	86912066	13123721966	75428445	69475752	79.94	13.12	13.21	7.89
8XE	Resistant	Control	End of the season	Ambient	33.3	87639550	13233572050	76683529	71598300	81.70	13.23	12.50	6.63
9EM	Susceptible	Control	Mid-season	Ambient	30.4	76739578	11587676278	66009506	60962488	79.44	11.59	13.98	7.65
9EE	Susceptible	Control	End of the season	Ambient	44.6	74029888	11178513088	64536900	59974132	81.01	11.18	12.82	7.07
9XM	Resistant	Control	Mid-season	Ambient	23.4	107540380	16238597380	92551600	84397566	78.48	16.24	13.94	8.81
9XE	Resistant	Control	End of the season	Ambient	35.8	83594546	12622776446	73142122	66505530	79.56	12.62	12.50	9.07
10EM	Susceptible	Control	Mid-season	Elevated ozone	19.2	62672036	9463477436	52745864	45571498	72.71	9.46	15.84	13.60
10EE	Susceptible	Control	End of the season	Elevated ozone	14	79082574	11941468674	70079799	65237834	82.49	11.94	11.38	6.91
10XM	Resistant	Control	Mid-season	Elevated ozone	21	74972812	11320894612	64473571	58494178	78.02	11.32	14.00	9.27
10XE	Resistant	Control	End of the season	Elevated ozone	17.5	98574260	14884713260	85562399	78476258	79.61	14.88	13.20	8.28
11EM	Susceptible	Control	Mid-season	Elevated ozone	29.8	120049118	18127416818	105338540	96783338	80.62	18.13	12.25	8.12
11EE	Susceptible	Control	End of the season	Elevated ozone	13.84	76087608	11489228808	67008309	56202158	73.87	11.49	11.93	16.13
11XM	Resistant	Control	Mid-season	Elevated ozone	16.06	108823052	16432280852	95722349	87635384	80.53	16.43	12.04	8.45
11XE	Resistant	Control	End of the season	Elevated ozone	16.77	76221844	11509498444	66056543	55582900	72.92	11.51	13.34	15.86
12EM	Susceptible	Control	Mid-season	Ambient	17.5	113646040	17160552040	99066065	91624154	80.62	17.16	12.83	7.51
12XM	Resistant	Control	Mid-season	Ambient	38	60168800	9085488800	52245447	48209842	80.12	9.09	13.17	7.72
12XE	Resistant	Control	End of the season	Ambient	29.6	105040296	15861084696	91675213	85177988	81.09	15.86	12.72	7.09
OEC	Susceptible	Control	Base	Green house	34.6	41926072	6330836872	34883887	26811796	63.95	6.33	16.80	23.14
OEI	Susceptible	Inoculated	Base	Green house	13.35	92472636	13963368036	76986190	54694930	59.15	13.96	16.75	28.95
OXC	Resistant	Control	Base	Green house	13.94	18879282	2850771582	15442554	11542140	61.14	2.85	18.20	25.26
OXI	Resistant	Inoculated	Base	Green house	16.11	61371554	9267104654	51009338	31065826	50.62	9.27	16.88	39.10

Table S2-3A: Pairwise comparison for bacterial Chao1 diversity index for resistant pepper. Significance levels for each treatment combination are indicated by *$p < 0.05$; **$p < 0.01$; ***$p < 0.001$.

Comparisons	Mean difference	lwr	upr	Adjusted *p*-value
Control sample, ambient environment, and end of the season vs. Control sample, ambient environment, and base time point	153	-26.281695	332.2816953	0.1298768
Control sample, ambient environment, and mid-season vs. Control sample, ambient environment, and base time point	177.333333	-1.948362	356.6150286	0.0537938
Control sample, elevated ozone, and end of the season vs. Control sample, ambient environment, and base time point	196.333333	17.051638	375.6150286	0.0260945
Control sample, elevated ozone, and mid-season vs. Control sample, ambient environment, and base time point	253.333333	74.051638	432.6150286	0.0028437
Inoculated sample, ambient environment, and base time point vs. Control sample, ambient environment, and base time point	218	-1.574337	437.5743368	0.0524734
Inoculated sample, ambient environment, and end of the season vs. Control sample, ambient environment, and base time point	127	-52.281695	306.2816953	0.3003286
Inoculated sample, ambient environment, and mid-season vs. Control sample, ambient environment, and base time point	193	13.718305	372.2816953	0.0296671
Inoculated sample, elevated ozone, and end of the season vs. Control sample, ambient environment, and base time point	139.666667	-39.615029	318.9483619	0.203141
Inoculated sample, elevated ozone, and mid-season vs. Control sample, ambient environment, and base time point	213.333333	34.051638	392.6150286	0.0134904
Control sample, ambient environment, and mid-season vs. Control sample, ambient environment, and end of the season	24.333333	-102.437969	151.1046358	0.9991099
Control sample, elevated ozone, and end of the season vs. Control sample, ambient environment, and end of the season	43.333333	-83.437969	170.1046358	0.9527894
Control sample, elevated ozone, and mid-season vs. Control sample, ambient environment, and end of the season	100.333333	-26.437969	227.1046358	0.1889544
Inoculated sample, ambient environment, and base time point vs. Control sample, ambient environment, and end of the season	65	-114.281695	244.2816953	0.9343351
Inoculated sample, ambient environment, and end of the season vs. Control sample, ambient environment, and end of the season	-26	-152.771302	100.7713025	0.9985192
Inoculated sample, ambient environment, and mid-season vs. Control sample, ambient environment, and end of the season	40	-86.771302	166.7713025	0.9706473
Inoculated sample, elevated ozone, and end of the season vs. Control sample, ambient environment, and end of the season	-13.333333	-140.104636	113.4379691	0.9999938
Inoculated sample, elevated ozone, and mid-season vs. Control sample, ambient environment, and end of the season	60.333333	-66.437969	187.1046358	0.7636049
Control sample, elevated ozone, and end of the season vs. Control sample, ambient environment, and mid-season	19	-107.771302	145.7713025	0.9998771
Control sample, elevated ozone, and mid-season vs. Control sample, ambient environment, and mid-season	76	-50.771302	202.7713025	0.5029492
Inoculated sample, ambient environment, and base time point vs. Control sample, ambient environment, and mid-season	40.666667	-138.615029	219.9483619	0.9968571
Inoculated sample, ambient environment, and end of the season vs. Control sample, ambient environment, and mid-season	-50.333333	-177.104636	76.4379691	0.8945303
Inoculated sample, ambient environment, and mid-season vs. Control sample, ambient environment, and mid-season	15.666667	-111.104636	142.4379691	0.9999754
Inoculated sample, elevated ozone, and end of the season vs. Control sample, ambient environment, and mid-season	-37.666667	-164.437969	89.1046358	0.9798423
Inoculated sample, elevated ozone, and mid-season vs. Control sample, ambient environment, and mid-season	36	-90.771302	162.7713025	0.9849613
Control sample, elevated ozone, and mid-season vs. Control sample, elevated ozone, and end of the season	57	-69.771302	183.7713025	0.8126418
Inoculated sample, ambient environment, and base time point vs. Control sample, elevated ozone, and end of the season	21.666667	-157.615029	200.9483619	0.9999796
Inoculated sample, ambient environment, and end of the season vs. Control sample, elevated ozone, and end of the season	-69.333333	-196.104636	57.4379691	0.615523
Inoculated sample, ambient environment, and mid-season vs. Control sample, elevated ozone, and end of the season	-3.333333	-130.104636	123.4379691	1
Inoculated sample, elevated ozone, and end of the season vs. Control sample, elevated ozone, and end of the season	-56.666667	-183.437969	70.1046358	0.8172832
Inoculated sample, elevated ozone, and mid-season vs. Control sample, elevated ozone, and end of the season	17	-109.771302	143.7713025	0.9999511
Inoculated sample, ambient environment, and base time point vs. Control sample, elevated ozone, and mid-season	-35.333333	-214.615029	143.9483619	0.9989083
Inoculated sample, ambient environment, and end of the season vs. Control sample, elevated ozone, and mid-season	-126.333333	-253.104636	0.4379691	0.0511777
Inoculated sample, ambient environment, and mid-season vs. Control sample, elevated ozone, and mid-season	-60.333333	-187.104636	66.4379691	0.7636049
Inoculated sample, elevated ozone, and end of the season vs. Control sample, elevated ozone, and mid-season	-113.666667	-240.437969	13.1046358	0.098838
Inoculated sample, elevated ozone, and mid-season vs. Control sample, elevated ozone, and mid-season	-40	-166.771302	86.7713025	0.9706473
Inoculated sample, ambient environment, and end of the season vs. Inoculated sample, ambient environment, and base time point	-91	-270.281695	88.2816953	0.6995067
Inoculated sample, ambient environment, and mid-season vs. Inoculated sample, ambient environment, and base time point	-25	-204.281695	154.2816953	0.9999323
Inoculated sample, elevated ozone, and end of the season vs. Inoculated sample, ambient environment, and base time point	-78.333333	-257.615029	100.9483619	0.8345678
Inoculated sample, elevated ozone, and mid-season vs. Inoculated sample, ambient environment, and base time point	-4.666667	-183.948362	174.6150286	1
Inoculated sample, ambient environment, and mid-season vs. Inoculated sample, ambient environment, and end of the season	66	-60.771302	192.7713025	0.6720025
Inoculated sample, elevated ozone, and end of the season vs. Inoculated sample, ambient environment, and end of the season	12.666667	-114.104636	139.4379691	0.999996
Inoculated sample, elevated ozone, and mid-season vs. Inoculated sample, ambient environment, and end of the season	86.333333	-40.437969	213.1046358	0.3456482
Inoculated sample, elevated ozone, and end of the season vs. Inoculated sample, ambient environment, and mid-season	-53.333333	-180.104636	73.4379691	0.8606658
Inoculated sample, elevated ozone, and mid-season vs. Inoculated sample, ambient environment, and mid-season	20.333333	-106.437969	147.1046358	0.9997862
Inoculated sample, elevated ozone, and mid-season vs. Inoculated sample, elevated ozone, and end of the season	73.666667	-53.104636	200.4379691	0.5418761

Table S2-3B*:* Pairwise comparison for bacterial Shanon diversity index(Dunn's multiple comparison test) for resistant pepper. Significance levels for each treatment combination are indicated by *$p < 0.05$; **$p < 0.01$; ***$p < 0.001$.

Comparisons	Z	Unadjusted *p*	*p* adjusted	Significance
Control sample, ambient environment, and base time point vs. Control sample, ambient environment, and end of the season	-1.13227703	0.257517982	0.57941546	
Control sample, ambient environment, and base time point vs. Control sample, ambient environment, and mid-season	-1.32098987	0.186504741	0.52454458	
Control sample, ambient environment, and end of the season vs. Control sample, ambient environment, and mid-season	-0.26688026	0.789561359	0.88825653	
Control sample, ambient environment, and base time point vs. Control sample, elevated ozone, and end of the season	-0.83033649	0.406348562	0.6772476	
Control sample, ambient environment, and end of the season vs. Control sample, elevated ozone, and end of the season	0.42700841	0.669373202	0.86062269	
Control sample, ambient environment, and mid-season vs. Control sample, elevated ozone, and end of the season	0.69388867	0.487752032	0.70802714	
Control sample, ambient environment, and base time point vs. Control sample, elevated ozone, and mid-season	-2.18906893	0.02859183	0.25732647	
Control sample, ambient environment, and end of the season vs. Control sample, elevated ozone, and mid-season	-1.49452944	0.135037294	0.46743679	
Control sample, ambient environment, and mid-season vs. Control sample, elevated ozone, and mid-season	-1.22764918	0.219578692	0.54894673	
Control sample, elevated ozone, and end of the season vs. Control sample, elevated ozone, and mid-season	-1.92153785	0.054663936	0.35141102	
Control sample, ambient environment, and base time point vs. Inoculated sample, ambient environment, and base time point	-1.01695036	0.309177045	0.63240759	
Control sample, ambient environment, and end of the season vs. Inoculated sample, ambient environment, and base time point	-0.1132277	0.909850033	0.95216864	
Control sample, ambient environment, and mid-season vs. Inoculated sample, ambient environment, and base time point	0.07548514	0.939828724	0.93982872	
Control sample, elevated ozone, and end of the season vs. Inoculated sample, ambient environment, and base time point	-0.41516825	0.678018743	0.82461739	
Control sample, elevated ozone, and mid-season vs. Inoculated sample, ambient environment, and base time point	0.9435642	0.345392396	0.62170631	
Control sample, ambient environment, and base time point vs. Inoculated sample, ambient environment, and end of the season	0.07548514	0.939828724	0.96118847	
Control sample, ambient environment, and end of the season vs. Inoculated sample, ambient environment, and end of the season	1.70803364	0.087630101	0.4381505	
Control sample, ambient environment, and mid-season vs. Inoculated sample, ambient environment, and end of the season	1.9749139	0.04827792	0.3620844	
Control sample, elevated ozone, and end of the season vs. Inoculated sample, ambient environment, and end of the season	1.28102523	0.200184804	0.52990095	
Control sample, elevated ozone, and mid-season vs. Inoculated sample, ambient environment, and end of the season	3.20256308	0.001362105	0.06129471	
Inoculated sample, ambient environment, and base time point vs. Inoculated sample, ambient environment, and end of the season	1.32098987	0.186504741	0.55951422	
Control sample, ambient environment, and base time point vs. Inoculated sample, ambient environment, and mid-season	-1.66067298	0.096779142	0.43550614	
Control sample, ambient environment, and end of the season vs. Inoculated sample, ambient environment, and mid-season	-0.74726472	0.454903785	0.73109537	
Control sample, ambient environment, and mid-season vs. Inoculated sample, ambient environment, and mid-season	-0.48038446	0.630954041	0.83508623	
Control sample, elevated ozone, and end of the season vs. Inoculated sample, ambient environment, and mid-season	-1.17427313	0.240285643	0.56909758	
Control sample, elevated ozone, and mid-season vs. Inoculated sample, ambient environment, and mid-season	0.74726472	0.454903785	0.70588518	
Inoculated sample, ambient environment, and base time point vs. Inoculated sample, ambient environment, and mid-season	-0.41516825	0.678018743	0.84752343	
Inoculated sample, ambient environment, and end of the season vs. Inoculated sample, ambient environment, and mid-season	-2.45529836	0.01407677	0.21115155	
Control sample, ambient environment, and base time point vs. Inoculated sample, elevated ozone, and end of the season	-0.33968311	0.734095182	0.86932324	
Control sample, ambient environment, and end of the season vs. Inoculated sample, elevated ozone, and end of the season	1.12089708	0.262331675	0.5621393	
Control sample, ambient environment, and mid-season vs. Inoculated sample, elevated ozone, and end of the season	1.38777733	0.165204859	0.53101562	
Control sample, elevated ozone, and end of the season vs. Inoculated sample, elevated ozone, and end of the season	0.69388867	0.487752032	0.73162805	
Control sample, elevated ozone, and mid-season vs. Inoculated sample, elevated ozone, and end of the season	2.61542651	0.00891161	0.20051122	
Inoculated sample, ambient environment, and base time point vs. Inoculated sample, elevated ozone, and end of the season	0.90582163	0.365030272	0.63178316	
Inoculated sample, ambient environment, and end of the season vs. Inoculated sample, elevated ozone, and end of the season	-0.58713656	0.557111993	0.78343874	
Inoculated sample, ambient environment, and mid-season vs. Inoculated sample, elevated ozone, and end of the season	1.86816179	0.061739522	0.34728481	
control sample, ambient environment, and base time point vs. Inoculated sample, elevated ozone, and mid-season	-1.50970271	0.131119299	0.49169737	
Control sample, ambient environment, and end of the season vs. Inoculated sample, elevated ozone, and mid-season	-0.53376051	0.593507237	0.80932805	
Control sample, ambient environment, and mid-season vs. Inoculated sample, elevated ozone, and mid-season	-0.26688026	0.789561359	0.91103234	
Control sample, elevated ozone, and end of the season vs. Inoculated sample, elevated ozone, and mid-season	-0.96076892	0.336668368	0.65869898	
Control sample, elevated ozone, and mid-season vs. Inoculated sample, elevated ozone, and mid-season	0.96076892	0.336668368	0.63125319	
Inoculated sample, ambient environment, and base time point vs. Inoculated sample, elevated ozone, and mid-season	-0.26419797	0.791627372	0.86885931	
Inoculated sample, ambient environment, and end of the season vs. Inoculated sample, elevated ozone, and mid-season	-2.24179415	0.024974679	0.28096514	
Inoculated sample, ambient environment, and mid-season vs. Inoculated sample, elevated ozone, and mid-season	0.21350421	0.83093371	0.89028612	
Inoculated sample, elevated ozone, and end of the season vs. Inoculated sample, elevated ozone, and mid-season	-1.65465759	0.097993975	0.40088444	

Table S 2-3C: Pairwise comparison for bacterial Chao1 diversity index for susceptible pepper. Significance levels for each treatment combination are indicated by *$p < 0.05$; **$p < 0.01$; ***$p < 0.001$.

Comparisons	diff	lwr	upr	*p* adjusted	Significance
Control sample, ambient environment, and end of the season vs. Control sample, ambient environment, and base time point	128	-88.71689	344.716891	0.4882834	
Control sample, ambient environment, and mid-season vs. Control sample, ambient environment, and base time point	141.5	-75.21689	358.216891	0.3688167	
Control sample, elevated ozone, and end of the season vs. Control sample, ambient environment, and base time point	110.666667	-93.65598	314.989311	0.5917825	
Control sample, elevated ozone, and mid-season vs. Control sample, ambient environment, and base time point	232	15.28311	448.716891	0.0319968	
Inoculated sample, ambient environment, and base time point vs. Control sample, ambient environment, and base time point	236	-14.24311	486.243111	0.0713791	
Inoculated sample, ambient environment, and end of the season vs. Control sample, ambient environment, and base time point	-63.666667	-267.98931	140.655978	0.964853	
Inoculated sample, ambient environment, and mid-season vs. Control sample, ambient environment, and base time point	66.333333	-137.98931	270.655978	0.9554074	
Inoculated sample, elevated ozone, and end of the season vs. Control sample, ambient environment, and base time point	-62.333333	-266.65598	141.989311	0.9690089	
Inoculated sample, elevated ozone, and mid-season vs. Control sample, ambient environment, and base time point	27.333333	-176.98931	231.655978	0.9999279	
Control sample, ambient environment, and mid-season vs. Control sample, ambient environment, and end of the season	13.5	-163.4486	190.448601	0.9999994	
Control sample, elevated ozone, and end of the season vs. Control sample, ambient environment, and end of the season	-17.333333	-178.86457	144.1979	0.9999887	
Control sample, elevated ozone, and mid-season vs. Control sample, ambient environment, and end of the season	104	-72.9486	280.948601	0.4942194	
Inoculated sample, ambient environment, and base time point vs. Control sample, ambient environment, and end of the season	108	-108.71689	324.716891	0.6846657	
Inoculated sample, ambient environment, and end of the season vs. Control sample, ambient environment, and end of the season	-191.666667	-353.1979	-30.135433	0.015289	*
Inoculated sample, ambient environment, and mid-season vs. Control sample, ambient environment, and end of the season	-61.666667	-223.1979	99.864567	0.894788	
Inoculated sample, elevated ozone, and end of the season vs. Control sample, ambient environment, and end of the season	-190.333333	-351.86457	-28.8021	0.0161128	*
Inoculated sample, elevated ozone, and mid-season vs. Control sample, ambient environment, and end of the season	-100.666667	-262.1979	60.864567	0.4236153	
Control sample, elevated ozone, and end of the season vs. Control sample, ambient environment, and mid-season	-30.833333	-192.36457	130.6979	0.998772	
Control sample, elevated ozone, and mid-season vs. Control sample, ambient environment, and mid-season	90.5	-86.4486	267.448601	0.6567579	
Inoculated sample, ambient environment, and base time point vs. Control sample, ambient environment, and mid-season	94.5	-122.21689	311.216891	0.8085275	
Inoculated sample, ambient environment, and end of the season vs. Control sample, ambient environment, and mid-season	-205.166667	-366.6979	-43.635433	0.0090049	**
Inoculated sample, ambient environment, and mid-season vs. Control sample, ambient environment, and mid-season	-75.166667	-236.6979	86.364567	0.7526888	
Inoculated sample, elevated ozone, and end of the season vs. Control sample, ambient environment, and mid-season	-203.833333	-365.36457	-42.3021	0.0094863	**
Inoculated sample, elevated ozone, and mid-season vs. Control sample, ambient environment, and mid-season	-114.166667	-275.6979	47.364567	0.2815983	
Control sample, elevated ozone, and mid-season vs. Control sample, elevated ozone, and end of the season	121.333333	-40.1979	282.864567	0.2220085	
Inoculated sample, ambient environment, and base time point vs. Control sample, elevated ozone, and end of the season	125.333333	-78.98931	329.655978	0.4425976	
Inoculated sample, ambient environment, and end of the season vs. Control sample, elevated ozone, and end of the season	-174.333333	-318.81126	-29.855406	0.0134579	*
Inoculated sample, ambient environment, and mid-season vs. Control sample, elevated ozone, and end of the season	-44.333333	-188.81126	100.144594	0.9679123	
Inoculated sample, elevated ozone, and end of the season vs. Control sample, elevated ozone, and end of the season	-173	-317.47793	-28.522073	0.0142701	*
Inoculated sample, elevated ozone, and mid-season vs. Control sample, elevated ozone, and end of the season	-83.333333	-227.81126	61.144594	0.5169461	
Inoculated sample, ambient environment, and base time point vs. Control sample, elevated ozone, and mid-season	4	-212.71689	220.716891	1	
Inoculated sample, ambient environment, and end of the season vs. Control sample, elevated ozone, and mid-season	-295.666667	-457.1979	-134.135433	0.000322	*
Inoculated sample, ambient environment, and mid-season vs. Control sample, elevated ozone, and mid-season	-165.666667	-327.1979	-4.135433	0.042541	*
Inoculated sample, elevated ozone, and end of the season vs. Control sample, elevated ozone, and mid-season	-294.333333	-455.86457	-132.8021	0.0003371	***
Inoculated sample, elevated ozone, and mid-season vs. Control sample, elevated ozone, and mid-season	-204.666667	-366.1979	-43.135433	0.0091824	***
Inoculated sample, ambient environment, and end of the season vs. Inoculated sample, ambient environment, and base time point	-299.666667	-503.98931	-95.344022	0.0026616	**
Inoculated sample, ambient environment, and mid-season vs. Inoculated sample, ambient environment, and base time point	-169.666667	-373.98931	34.655978	0.1413235	
Inoculated sample, elevated ozone, and end of the season vs. Inoculated sample, ambient environment, and base time point	-298.333333	-502.65598	-94.010689	0.0027695	**
Inoculated sample, elevated ozone, and mid-season vs. Inoculated sample, ambient environment, and base time point	-208.666667	-412.98931	-4.344022	0.0437245	*
Inoculated sample, ambient environment, and mid-season vs. Inoculated sample, ambient environment, and end of the season	130	-14.47793	274.477927	0.0931871	*
Inoculated sample, elevated ozone, and end of the season vs. Inoculated sample, ambient environment, and end of the season	1.333333	-143.14459	145.811261	1	
Inoculated sample, elevated ozone, and mid-season vs. Inoculated sample, ambient environment, and end of the season	91	-53.47793	235.477927	0.4109739	
Inoculated sample, elevated ozone, and end of the season vs. Inoculated sample, ambient environment, and mid-season	-128.666667	-273.14459	15.811261	0.0985778	
Inoculated sample, elevated ozone, and mid-season vs. Inoculated sample, ambient environment, and mid-season	-39	-183.47793	105.477927	0.9855907	
Inoculated sample, elevated ozone, and mid-season vs. Inoculated sample, elevated ozone, and end of the season	89.666667	-54.81126	234.144594	0.4285686	

:

Table S 2-3D: Pairwise comparison for bacterial Shanon diversity index (Kruskal-Wallis multiple comparisons & p-values adjusted with the Benjamini-Hochberg method) for control susceptible pepper.

Comparisons	Z	*p* unadjusted	*p* adjusted	Significance
Control sample, elevated ozone, and end of the season vs. Inoculated sample, elevated ozone, and end of the season	2.4618298	0.01382302	0.04146907	*
Control sample, elevated ozone, and end of the season vs. Control sample, elevated ozone, and mid-season	-0.2752409	0.78313113	0.78313113	
Control sample, elevated ozone, and mid-season vs. Inoculated sample, elevated ozone, and end of the season	2.4771685	0.01324294	0.07945763	
Control sample, elevated ozone, and end of the season vs. Inoculated sample, elevated ozone, and mid-season	1.3540064	0.17573434	0.2636015	
Control sample, elevated ozone, and mid-season vs. Inoculated sample, elevated ozone, and mid-season	1.4863011	0.1371995	0.274399	
Inoculated sample, elevated ozone, and end of the season vs. Inoculated sample, elevated ozone, and mid-season	-1.1078234	0.26793808	0.3215257	

Table S 2-3E: Pairwise comparison for bacterial Shanon diversity index for inoculated susceptible pepper. Significance levels of treatment combination are indicated by *p < 0.05; **p < 0.01;***p < 0.001.

Comparisons	diff	lwr	upr	*p* adjusted	Significance
Inoculated sample, ambient environment, and end of the season vs. Control sample, ambient environment, and end of the season	-2.2891591	-3.3399312	-1.238387	0.0011631	**
Inoculated sample, ambient environment, and mid-season vs. Control sample, ambient environment, and end of the season	-1.6760916	-2.7268637	-0.6253195	0.0059681	**
Inoculated sample, ambient environment, and end of the season vs. Control sample, ambient environment, and mid-season	-3.080265	-4.1310371	-2.0294929	0.000223	***
Inoculated sample, ambient environment, and mid-season vs. Control sample, ambient environment, and mid-season	-2.4671976	-3.5179697	-1.4164255	0.000772	***
Control sample, ambient environment, and mid-season vs. Control sample, ambient environment, and end of the season	0.7911059	-0.3599572	1.9421691	0.1817845	
Inoculated sample, ambient environment, and mid-season vs. Inoculated sample, ambient environment, and end of the season	0.6130674	-0.3267717	1.5529066	0.2101921	

Table S 2-4A: Pairwise comparison of Eukaryotic Chao1 diversity index for resistant pepper. Significance levels for each treatment combination are indicated by *p < 0.05; **p < 0.01; ***p < 0.001.

Comparisons	diff	lwr	upr	*p* adjusted	Significance
Control, ambient environment, and mid-season vs. Control, ambient environment, and end of the season	7.3333333	-2.5464459	17.213113	0.231451	
Control, elevated ozone, and end of the season vs. Control, ambient environment, and end of the season	2	-7.8797792	11.879779	0.995412	
Control, elevated ozone, and mid-season vs. Control, ambient environment, and end of the season	10	0.1202208	19.879779	0.0462689	*
Inoculated, ambient environment, and end of the season vs. Control, ambient environment, and end of the season	2.6666667	-7.2131125	12.546446	0.9761333	
Inoculated, ambient environment, and mid-season vs. Control, ambient environment, and end of the season	3.6666667	-7.3792623	14.712596	0.93174	
Inoculated, elevated ozone, and end of the season vs. Control, ambient environment, and end of the season	2.3333333	-7.5464459	12.213113	0.988641	
Inoculated, elevated ozone, and mid-season vs. Control, ambient environment, and end of the season	6.3333333	-3.5464459	16.213113	0.383007	
Control, elevated ozone, and end of the season vs. Control, ambient environment, and mid-season	-5.3333333	-15.2131125	4.546446	0.578869	
Control, elevated ozone, and mid-season vs. Control, ambient environment, and mid-season	2.6666667	-7.2131125	12.546446	0.9761333	
Inoculated, ambient environment, and end of the season vs. Control, ambient environment, and mid-season	-4.6666667	-14.5464459	5.213113	0.716115	
Inoculated, ambient environment, and mid-season vs. Control, ambient environment, and mid-season	-3.6666667	-14.7125956	7.379262	0.93174	
Inoculated, elevated ozone, and end of the season vs. Control, ambient environment, and mid-season	-5	-14.8797792	4.879779	0.6482302	
Inoculated, elevated ozone, and mid-season vs. Control, ambient environment, and mid-season	-1	-10.8797792	8.879779	0.9999477	
Control, elevated ozone, and mid-season vs. Control, elevated ozone, and end of the season	8	-1.8797792	17.879779	0.1592666	
Inoculated, ambient environment, and end of the season vs. Control, elevated ozone, and end of the season	0.6666667	-9.2131125	10.546446	0.9999967	
Inoculated, ambient environment, and mid-season vs. Control, elevated ozone, and end of the season	1.6666667	-9.3792623	12.712596	0.9992691	
Inoculated, elevated ozone, and end of the season vs. Control, elevated ozone, and end of the season	0.3333333	-9.5464459	10.213113	1	
Inoculated, elevated ozone, and mid-season vs. Control, elevated ozone, and end of the season	4.3333333	-5.5464459	14.213113	0.7800355	
Inoculated, ambient environment, and end of the season vs. Control, elevated ozone, and mid-season	-7.3333333	-17.2131125	2.546446	0.231451	
Inoculated, ambient environment, and mid-season vs. Control, elevated ozone, and mid-season	-6.3333333	-17.3792623	4.712596	0.5107304	
Inoculated, elevated ozone, and end of the season vs. Control, elevated ozone, and mid-season	-7.6666667	-17.5464459	2.213113	0.1926158	
Inoculated, elevated ozone, and mid-season vs. Control, elevated ozone, and mid-season	-3.6666667	-13.5464459	6.213113	0.8867681	
Inoculated, ambient environment, and mid-season vs. Inoculated, ambient environment, and end of the season	1	-10.0459289	12.045929	0.9999754	
Inoculated, elevated ozone, and end of the season vs. Inoculated, ambient environment, and end of the season	-0.3333333	-10.2131125	9.546446	1	
Inoculated, elevated ozone, and mid-season vs. Inoculated, ambient environment, and end of the season	3.6666667	-6.2131125	13.546446	0.8867681	
Inoculated, elevated ozone, and end of the season vs. Inoculated, ambient environment, and mid-season	-1.3333333	-12.3792623	9.712596	0.9998302	
Inoculated, elevated ozone, and mid-season vs. Inoculated, ambient environment, and mid-season	2.6666667	-8.3792623	13.712596	0.9871275	
Inoculated, elevated ozone, and mid-season vs. Inoculated, elevated ozone, and end of the season	4	-5.8797792	13.879779	0.8375864	

Table S 2-4B Pairwise comparison of Eukaryotic Shannon diversity index for resistant pepper. Significance levels for each treatment combination are indicated by *p < 0.05; **p < 0.01; ***p < 0.001.

Comparisons	diff	lwr	upr	*p* adjusted	Significance
Control, ambient environment, and mid-season vs. Control, ambient environment, and end of the season	0.20217503	-0.5332479	0.93759798	0.9736648	
Control, elevated ozone, and end of the season vs. Control, ambient environment, and end of the season	0.060349817	-0.6750731	0.79577277	0.9999874	
Control, elevated ozone, and mid-season vs. Control, ambient environment, and end of the season	-0.098997082	-0.83442	0.63642587	0.9996517	
Inoculated, ambient environment, and end of the season vs. Control, ambient environment, and end of the season	0.26322271	-0.4722002	0.99864566	0.9034352	
Inoculated, ambient environment, and mid-season vs. Control, ambient environment, and end of the season	0.262638007	-0.5595899	1.08486587	0.9432662	
Inoculated, elevated ozone, and end of the season vs. Control, ambient environment, and end of the season	0.168106746	-0.5673162	0.9035297	0.9905876	
Inoculated, elevated ozone, and mid-season vs. Control, ambient environment, and end of the season	-0.394374595	-1.1297975	0.34104836	0.5862019	
Control, elevated ozone, and end of the season vs. Control, ambient environment, and mid-season	-0.141825214	-0.8772482	0.59359774	0.9965823	
Control, elevated ozone, and mid-season vs. Control, ambient environment, and mid-season	-0.301172112	-1.0365951	0.43425084	0.830105	
Inoculated, ambient environment, and end of the season vs. Control, ambient environment, and mid-season	0.06104768	-0.6743753	0.79647063	0.9999864	
Inoculated, ambient environment, and mid-season vs. Control, ambient environment, and mid-season	0.060462976	-0.7617649	0.88269083	0.9999941	
Inoculated, elevated ozone, and end of the season vs. Control, ambient environment, and mid-season	-0.034068284	-0.7694912	0.70135467	0.9999998	
Inoculated, elevated ozone, and mid-season vs. Control, ambient environment, and mid-season	-0.596549625	-1.3319726	0.13887333	0.1579671	
Control, elevated ozone, and mid-season vs. Control, elevated ozone, and end of the season	-0.159346899	-0.8947699	0.57607606	0.9931169	
Inoculated, ambient environment, and end of the season vs. Control, elevated ozone, and end of the season	0.202872893	-0.5325501	0.93829585	0.9731763	
Inoculated, ambient environment, and mid-season vs. Control, elevated ozone, and end of the season	0.20228819	-0.6199397	1.02451605	0.9856766	
Inoculated, elevated ozone, and end of the season vs. Control, elevated ozone, and end of the season	0.107756929	-0.627666	0.84317988	0.9993951	
Inoculated, elevated ozone, and mid-season vs. Control, elevated ozone, and end of the season	-0.454724412	-1.1901474	0.28069854	0.4239381	
Inoculated, ambient environment, and end of the season vs. Control, elevated ozone, and mid-season	0.362219792	-0.3732032	1.09764275	0.6758122	
Inoculated, ambient environment, and mid-season vs. Control, elevated ozone, and mid-season	0.361635089	-0.4605928	1.18386295	0.7778266	
Inoculated, elevated ozone, and end of the season vs. Control, elevated ozone, and mid-season	0.267103828	-0.4683191	1.00252678	0.8969574	
Inoculated, elevated ozone, and mid-season vs. Control, elevated ozone, and mid-season	-0.295377513	-1.0308005	0.44004544	0.8426747	
Inoculated, ambient environment, and mid-season vs. Inoculated, ambient environment, and end of the season	-0.000584704	-0.8228126	0.82164315	1	
Inoculated, elevated ozone, and end of the season vs. Inoculated, ambient environment, and end of the season	-0.095115964	-0.8305389	0.64030699	0.9997321	
Inoculated, elevated ozone, and mid-season vs. Inoculated, ambient environment, and end of the season	-0.657597305	-1.3930203	0.07782565	0.0966907	
Inoculated, elevated ozone, and end of the season vs. Inoculated, ambient environment, and mid-season	-0.094531261	-0.9167591	0.7276966	0.9998772	
Inoculated, elevated ozone, and mid-season vs. Inoculated, ambient environment, and mid-season	-0.657012601	-1.4792405	0.16521526	0.1692467	
Inoculated, elevated ozone, and mid-season vs. Inoculated, elevated ozone, and end of the season	-0.562481341	-1.2979043	0.17294161	0.2048296	

Table S 2-4C: Pairwise comparison of Eukaryotic Chao1 diversity index for susceptible pepper. Significance levels for each treatment combination are indicated by *p < 0.05; **p < 0.01; ***p < 0.001; ****p < 0.0001.

Comparisons	diff	lwr	upr	*p* adjusted	Significance
Control, ambient environment, and mid-season vs. Control, ambient environment, and end of the season	13	-4.108058	30.1080581	0.2015098	
Control, elevated ozone, and end of the season vs. Control, ambient environment, and end of the season	8.333333	-7.284116	23.9507822	0.5700742	
Control, elevated ozone, and mid-season vs. Control, ambient environment, and end of the season	19.5	2.391942	36.6080581	0.021199	*
Inoculated, ambient environment, and end of the season vs. Control, ambient environment, and end of the season	-13.333333	-28.950782	2.2841156	0.1193093	
Inoculated, ambient environment, and mid-season vs. Control, ambient environment, and end of the season	-4	-19.617449	11.6174489	0.9790401	
Inoculated, elevated ozone, and end of the season vs. Control, ambient environment, and end of the season	-14.666667	-30.284116	0.9507822	0.072205	
Inoculated, elevated ozone, and mid-season vs. Control, ambient environment, and end of the season	-5	-20.617449	10.6174489	0.9347312	
Control, elevated ozone, and end of the season vs. Control, ambient environment, and mid-season	-4.666667	-20.284116	10.9507822	0.9532884	
Control, elevated ozone, and mid-season vs. Control, ambient environment, and mid-season	6.5	-10.608058	23.6080581	0.8610813	
Inoculated, ambient environment, and end of the season vs. Control, ambient environment, and mid-season	-26.333333	-41.950782	-10.7158844	0.0008164	***
Inoculated, ambient environment, and mid-season vs. Control, ambient environment, and mid-season	-17	-32.617449	-1.3825511	0.0290765	*
Inoculated, elevated ozone, and end of the season vs. Control, ambient environment, and mid-season	-27.666667	-43.284116	-12.0492178	0.0005076	***
Inoculated, elevated ozone, and mid-season vs. Control, ambient environment, and mid-season	-18	-33.617449	-2.3825511	0.0195972	*
Control, elevated ozone, and mid-season vs. Control, elevated ozone, and end of the season	11.166667	-4.450782	26.7841156	0.2550805	
Inoculated, ambient environment, and end of the season vs. Control, elevated ozone, and end of the season	-21.666667	-35.635338	-7.6979957	0.0017689	**
Inoculated, ambient environment, and mid-season vs. Control, elevated ozone, and end of the season	-12.333333	-26.302004	1.6353376	0.1007084	
Inoculated, elevated ozone, and end of the season vs. Control, elevated ozone, and end of the season	-23	-36.968671	-9.031329	0.0010214	**
Inoculated, elevated ozone, and mid-season vs. Control, elevated ozone, and end of the season	-13.333333	-27.302004	0.6353376	0.0658351	
Inoculated, ambient environment, and end of the season vs. Control, elevated ozone, and mid-season	-32.833333	-48.450782	-17.2158844	0.0000893	****
Inoculated, ambient environment, and mid-season vs. Control, elevated ozone, and mid-season	-23.5	-39.117449	-7.8825511	0.0023201	**
Inoculated, elevated ozone, and end of the season vs. Control, elevated ozone, and mid-season	-34.166667	-49.784116	-18.5492178	0.0000585	****
Inoculated, elevated ozone, and mid-season vs. Control, elevated ozone, and mid-season	-24.5	-40.117449	-8.8825511	0.0015964	**
Inoculated, ambient environment, and mid-season vs. Inoculated, ambient environment, and end of the season	9.333333	-4.635338	23.3020043	0.3220228	
Inoculated, elevated ozone, and end of the season vs. Inoculated, ambient environment, and end of the season	-1.333333	-15.302004	12.6353376	0.9999565	
Inoculated, elevated ozone, and mid-season vs. Inoculated, ambient environment, and end of the season	8.333333	-5.635338	22.3020043	0.445405	
Inoculated, elevated ozone, and end of the season vs. Inoculated, ambient environment, and mid-season	-10.666667	-24.635338	3.3020043	0.1974943	
Inoculated, elevated ozone, and mid-season vs. Inoculated, ambient environment, and mid-season	-1	-14.968671	12.968671	0.9999939	
Inoculated, elevated ozone, and mid-season vs. Inoculated, elevated ozone, and end of the season	9.666667	-4.302004	23.6353376	0.2865034	

Table S 2-4D: Pairwise comparison of Eukaryotic Shanon diversity index for susceptible pepper. Significance levels for each treatment combination are indicated by *p < 0.05; **p < 0.01; ***p< 0.001.

Comparisons	diff	lwr	upr	*p* adjusted	Significance
Control, ambient environment, and mid-season vs. Control, ambient environment, and end of the season	-0.044112784	-0.9977797	0.90955417	0.9999997	
Control, elevated ozone, and end of the season vs. Control, ambient environment, and end of the season	-0.070985068	-0.9415599	0.79958977	0.9999851	
Control, elevated ozone, and mid-season vs. Control, ambient environment, and end of the season	-0.043791967	-0.9974589	0.90987498	0.9999997	
Inoculated, ambient environment, and end of the season vs. Control, ambient environment, and end of the season	-1.506356533	-2.3769314	-0.6357817	0.0006376	***
Inoculated, ambient environment, and mid-season vs. Control, ambient environment, and end of the season	-0.837815566	-1.7083904	0.03275927	0.0627951	
Inoculated, elevated ozone, and end of the season vs. Control, ambient environment, and end of the season	-1.543782445	-2.4143573	-0.67320761	0.0005027	***
Inoculated, elevated ozone, and mid-season vs. Control, ambient environment, and end of the season	-0.971433688	-1.8420085	-0.10085885	0.0245729	*
Control, elevated ozone, and end of the season vs. Control, ambient environment, and mid-season	-0.026872285	-0.8974471	0.84370255	1	
Control, elevated ozone, and mid-season vs. Control, ambient environment, and mid-season	0.000320816	-0.9533461	0.95398777	1	
Inoculated, ambient environment, and end of the season vs. Control, ambient environment, and mid-season	-1.462243749	-2.3328186	-0.59166891	0.0008469	***
Inoculated, ambient environment, and mid-season vs. Control, ambient environment, and mid-season	-0.793702783	-1.6642776	0.07687205	0.0850128	
Inoculated, elevated ozone, and end of the season vs. Control, ambient environment, and mid-season	-1.499669661	-2.3702445	-0.62909483	0.0006655	***
Inoculated, elevated ozone, and mid-season vs. Control, ambient environment, and mid-season	-0.927320904	-1.7978957	-0.05674607	0.0335637	*
Control, elevated ozone, and mid-season vs. Control, elevated ozone, and end of the season	0.027193101	-0.8433817	0.89776794	1	
Inoculated, ambient environment, and end of the season vs. Control, elevated ozone, and end of the season	-1.435371465	-2.2140373	-0.65670566	0.0003432	***
Inoculated, ambient environment, and mid-season vs. Control, elevated ozone, and end of the season	-0.766830498	-1.5454963	0.01183531	0.0548337	
Inoculated, elevated ozone, and end of the season vs. Control, elevated ozone, and end of the season	-1.472797377	-2.2514632	-0.69413157	0.0002653	***
Inoculated, elevated ozone, and mid-season vs. Control, elevated ozone, and end of the season	-0.900448619	-1.6791144	-0.12178282	0.0191385	*
Inoculated, ambient environment, and end of the season vs. Control, elevated ozone, and mid-season	-1.462564566	-2.3331394	-0.59198973	0.0008451	***
Inoculated, ambient environment, and mid-season vs. Control, elevated ozone, and mid-season	-0.794023599	-1.6645984	0.07655124	0.0848274	
Inoculated, elevated ozone, and end of the season vs. Control, elevated ozone, and mid-season	-1.499990478	-2.3705653	-0.62941564	0.0006641	***
Inoculated, elevated ozone, and mid-season vs. Control, elevated ozone, and mid-season	-0.92764172	-1.7982166	-0.05706689	0.0334878	*
Inoculated, ambient environment, and mid-season vs. Inoculated, ambient environment, and end of the season	0.668540967	-0.1101248	1.44720677	0.1160321	
Inoculated, elevated ozone, and end of the season vs. Inoculated, ambient environment, and end of the season	-0.037425912	-0.8160917	0.74123989	0.9999996	
Inoculated, elevated ozone, and mid-season vs. Inoculated, ambient environment, and end of the season	0.534922846	-0.243743	1.31358865	0.2937759	
Inoculated, elevated ozone, and end of the season vs. Inoculated, ambient environment, and mid-season	-0.705966879	-1.4846327	0.07269892	0.0875989	
Inoculated, elevated ozone, and mid-season vs. Inoculated, ambient environment, and mid-season	-0.133618121	-0.9122839	0.64504768	0.9979966	
Inoculated, elevated ozone, and mid-season vs. Inoculated, elevated ozone, and end of the season	0.572348758	-0.206317	1.35101456	0.2299224	

Table S 2-5A*:* Permutational multivariant analysis of variance (PERMANOVA) of bacterial community composition based on Bray–Curtis dissimilarities of their relative abundance. Significance levels for each treatment combination are indicated by *P < 0.05; **P < 0.01; ***P < 0.001.

Variable	Df	Sum of sq	Mean Sqs	F.Model	R2	Pr(>F)	Significance
Cultivar	1	1.5289	1.5289	13.6497	0.12077	0.000999	***
Time of sampling	1	1.2936	1.2936	11.5487	0.10218	0.000999	***
Inoculation status	1	3.2369	3.2369	28.8983	0.25568	0.000999	***
Environment	1	0.3373	0.3373	3.0112	0.02664	0.023976	*
Cultivar:Time of sampling	1	0.2654	0.2654	2.3694	0.02096	0.052947	***
Cultivar:Inoculation status	1	0.9932	0.9932	8.8667	0.07845	0.000999	***
Time of sampling:Inoculation status	1	0.5124	0.5124	4.5742	0.04047	0.008991	**
Cultivar:Environment	1	0.1136	0.1136	1.0142	0.00897	0.330669	
Time of sampling:Environment	1	0.1164	0.1164	1.0389	0.00919	0.336663	
Inoculation status:Environment	1	0.1682	0.1682	1.5015	0.01328	0.172827	
Cultivar:Time of sampling:Inoculation status	1	0.3345	0.3345	2.986	0.02642	0.032967	*
Cultivar:Time of sampling:Environment	1	0.1147	0.1147	1.0242	0.00906	0.333666	
Cultivar:Inoculation status:Environment	1	0.0735	0.0735	0.6565	0.00581	0.632368	
Time of sampling:Inoculation status:Environment	1	0.1436	0.1436	1.2823	0.01135	0.241758	
Cultivar:Time of sampling:Inoculation status:Environment	1	0.1796	0.1796	1.6034	0.01419	0.16983	
Residuals	29	3.2483	0.112		0.25658		
Total	48	14.31			1		

Table S 2-5B: Permutational multivariant analysis of variance (PERMANOVA) of bacterial community composition based on Bray–Curtis dissimilarities of relative abundance during the mid- season. Significance levels for each treatment combination are indicated by *p < 0.05; **p < 0.01; ***p < 0.001.

Variable	Df	Sum of sq	Mean Sqs	F.Model	R2	Pr(>F)	Significance
Cultivar	1	1.1685	1.16851	11.9388	0.23261	0.000999	***
Environment	1	0.2725	0.27248	2.784	0.05424	0.048951	*
Inoculation status	1	1.0519	1.05187	10.7471	0.20939	0.000999	***
Cultivar:Environment	1	0.0731	0.07313	0.7472	0.01456	0.515485	
Cultivar:Inoculation status	1	0.7896	0.78957	8.0671	0.15717	0.001998	**
Environment:Inoculation status	1	0.1676	0.16764	1.7127	0.03337	0.143856	
Cultivar:Inoculation status:Environment	1	0.1301	0.13012	1.3294	0.0259	0.208791	
Residuals	14	1.3703	0.09788		0.27276		
Total	21	5.0236			1		

Table S 2-5C: Permutational multivariant analysis of variance (PERMANOVA) of bacterial community composition based on Bray–Curtis dissimilarities of relative abundance during end of the season. Significance levels for each treatment combination are indicated by *p < 0.05; **p < 0.01; ***p < 0.001.

Variable	Df	Sum of sq	Mean Sqs	F.Model	R2	Pr(>F)	Significance
Cultivar	1	0.3615	0.3615	4.038	0.06432	0.032967	*
Environment	1	0.1627	0.16272	1.818	0.02895	0.152847	
Inoculation status	1	3.1114	3.1114	34.755	0.55363	0.000999	***
Cultivar:Environment	1	0.1342	0.13425	1.5	0.02389	0.190809	
Cultivar:Inoculation status	1	0.2961	0.29615	3.308	0.05269	0.042957	*
Environment:Inoculation status	1	0.0881	0.08808	0.984	0.01567	0.336663	
Cultivar:Inoculation status:Environment	1	0.1231	0.12309	1.375	0.0219	0.223776	
Residuals	15	1.3429	0.08952		0.23894		
Total	22	5.62			1		

Table S 2-5D: Permutational multivariant analysis of variance (PERMANOVA) of bacterial community composition based on Bray–Curtis dissimilarities of relative abundance in control samples of the resistant cultivar. Significance levels for each treatment combination are indicated by *p < 0.05; **p < 0.01; ***p < 0.001.

Variable	Df	Sum of sq	Mean Sqs	F.Model	R2	Pr(>F)	Significance
Environment	1	0.31574	0.31574	2.3301	0.1417	0.027	*
Time of sampling	1	0.64875	0.64875	4.7876	0.29115	0.001	***
Environment:Time of sampling	1	0.17973	0.17973	1.3264	0.08066	0.228	
Residuals	8	1.08405	0.13551		0.4865		
Total	11	2.22828			1		

Table S 2-5E: Permutational multivariant analysis of variance (PERMANOVA) of bacterial community composition based on Bray–Curtis dissimilarities of relative abundance in samples from the elevated ozone. Significance levels for each treatment combination are indicated by *p < 0.05; **p < 0.01; ***p < 0.001.

Variable	Df	Sum of sq	Mean Sqs	F.Model	R2	Pr(>F)	Significance
Inoculation status	1	2.0005	2.00048	10.659	0.32704	0.001998	**
Chamber	4	0.5767	0.14417	0.7682	0.09428	0.691309	
Time of Sampling	1	0.7791	0.77909	4.1512	0.12737	0.015984	*
Residuals	11	2.0645	0.0912				
Total	22	6.117	0.18768				

Table S 2-5F: Permutational multivariant analysis of variance (PERMANOVA) of bacterial community composition based on Bray–Curtis dissimilarities of relative abundance in samples from the ambient environment. Significance levels for each treatment combination are indicated by *p < 0.05; **p < 0.01; ***p < 0.001.

Variable	Df	Sum of sq	Mean Sqs	F.Model	R2	Pr(>F)	Significance
Inoculation status	1	2.06467	2.06467	10.9547	0.36955	0.000999	***
Chamber	4	0.4095	0.10239	0.5432	0.0733	0.856144	
Time of Sampling	1	0.7042	0.70423	3.7365	0.12605	0.026973	*
Residuals	10	1.8847	0.07308				
Total	21	5.5869	0.18847				

Table S 2-6A: Permutational multivariant analysis of variance (PERMANOVA) of eukaryotes community composition based on Bray–Curtis dissimilarities of relative abundance. Significance levels for each treatment combination are indicated by *p < 0.05; **p < 0.01; ***p < 0.001.

Variable	Df	Sum of sq	R2	F	Pr(>F)	Significance
Time	1	1.8147	0.16834	9.3572	0.001	***
Environment	1	0.8229	0.07634	4.243	0.002	**
Time:Environment	1	0.3847	0.03569	1.9837	0.06	
Residual	40	7.7575	0.71963			
Total	43	10.7799	1			

Table S 2-6B: Analysis of similarity (ANOSIM) of eukaryotes community composition based on Bray–Curtis dissimilarities of relative abundance influenced by different factors. Significance levels for each treatment combination are indicated by *p < 0.05; **p < 0.01; ***p < 0.001.

Variable	R-Value	P-Value	Significance
Susceptible vs Resistant	0.2092	0.001	**
Inoculation vs Control	0.1189	0.002	**

Table S 2-6C: Permutational multivariant analysis of variance (PERMANOVA) of eukaryote community composition based on Bray–Curtis dissimilarities of relative abundance during mid- season. Significance levels for each treatment combination are indicated by *p < 0.05; **p < 0.01; ***p < 0.001.

Variable	Df	Sum of sq	R2	F	Pr(>F)	Significance
Cultivar	1	0.4475	0.1208	5.68	0.007	**
Environment	1	0.8238	0.22238	10.4565	0.001	***
Inoculation	1	0.5713	0.15421	7.2508	0.003	**
Cultivar:Environment	1	0.1401	0.03781	1.7779	0.134	
Cultivar:Inoculation	1	0.3258	0.08795	4.1352	0.012	*
Environment:inoculation	1	0.2071	0.05591	2.629	0.062	
Cultivar:Environment:Inoculation	1	0.1647	0.04447	2.0909	0.098	
Residual	13	1.0242	0.27648			
Total	20	3.7046	1			

Table S 2-6D: Permutational multivariant analysis of variance (PERMANOVA) of eukaryote community composition based on Bray–Curtis dissimilarities of relative abundance during the end season. Significance levels for each treatment combination are indicated by *p < 0.05; **p < 0.01; ***p < 0.001.

Variable	Df	Sum of sq	R2	F	Pr(>F)	Significance
Environment	1	0.3513	0.06679	1.5029	0.19	
Residual	21	4.9092	0.93321			
Total	22	5.2606	1			

Table S 2-6E: Analysis of similarity (ANOSIM) of eukaryotes community composition based on Bray–Curtis dissimilarities of relative abundance during end season. Significance levels for each treatment combination are indicated by *p < 0.05; **p < 0.01; ***p < 0.001.

Variable	R-Value	P-Value	Significance
Susceptible vs. Resistant	0.3747	0.001	**
Inoculation vs. Control	0.1157	0.035	*

Table S 2-7A: Permutational multivariant analysis of variance (PERMANOVA) microbiota density across samples during mid-season and ambient environment based on Bray–Curtis dissimilarities in Control samples. Significance levels for each treatment combination are indicated by *p < 0.05; **p < 0.01; ***p < 0.001.

Variable	Df	Sum of sq	R2	F	Pr(>F)	Significance
Treatments	3	0.03898	0.01299	1.012	0.442	
Residuals	7	0.08985	0.01284			

Table S 2-7B: Pairwise comparison of microbiota density across samples during mid-season under elevated ozone. Significance levels for each treatment combination are indicated by *p < 0.05;**p < 0.01; ***p < 0.001.

Comparisons	diff	lwr	upr	*p* adjusted	Significance
Control resistant cultivar vs. Control susceptible cultivar	-0.10391667	-0.33077446	0.12294113	0.4770532	
Inoculated susceptible vs. Control susceptible plant	0.20833333	-0.01852446	0.43519113	0.0710621	
Inoculated resistant cultivar vs. Control susceptible cultivar	-0.06341667	-0.29027446	0.16344113	0.7929435	
Inoculated susceptible cultivar vs. Control resistant cultivar	0.31225	0.10934222	0.51515778	0.0059349	**
Inoculated resistant cultivar vs. Control resistant cultivar	0.0405	-0.16240778	0.24340778	0.908505	
Inoculated resistant cultivar vs. Inoculated susceptible cultivar	-0.27175	-0.47465778	-0.06884222	0.012501	*

Table S 2-7C: Permutational multivariant analysis of variance (PERMANOVA) microbiota density across samples during mid-season and ambient environment based on Bray–Curtis dissimilarities in Control samples. Significance levels for each treatment combination are indicated by *p < 0.05; **p < 0.01; ***p < 0.001.

Variable	Df	Sum of sq	R2	F	Pr(>F)	Significance
Treatments	3	0.03403	0.011345	2.976	0.106	
Residuals	7	0.02668	0.003812			

Table S 2-7D: Pairwise comparison of microbiota density across samples during the end of the season and elevated ozone. Significance levels for each treatment combination are indicated by *p< 0.05; **p < 0.01; ***p < 0.001.

Comparisons	diff	lwr	upr	*p* adjusted	Significance
Control resistant cultivar vs. Control susceptible cultivar	0.01929167	-0.250525082	0.2891084	0.9954232	
Inoculated susceptible cultivar vs. Control susceptible cultivar	0.29525	0.025433252	0.5650667	0.0327813	*
Inoculated resistant cultivar vs. Control susceptible cultivar	0.15066667	-0.119150082	0.4204834	0.3446105	
Inoculated susceptible cultivar vs. Control resistant cultivar	0.27595833	0.006141585	0.5457751	0.0451318	*
Inoculated resistant cultivar vs. Control resistant cultivar	0.131375	-0.138441748	0.4011917	0.4499963	
Inoculated resistant cultivar vs. Inoculated susceptible cultivar	-0.14458333	-0.414400082	0.1252334	0.3758588	

Table S 2-7E: Permutational multivariant analysis of variance (PERMANOVA) on the total absolute abundance of microbial communities based on Bray–Curtis dissimilarities in Control samples. Significance levels for each treatment combination are indicated by *p < 0.05; **p < 0.01; ***p < 0.001.

Variable	Df	Sum of sq	R2	F	Pr(>F)	Significance
Treatments	7	0.1481	0.02116	1.881	0.154	
Residuals	13	0.1462	0.01125			

Table S 2-7F: Pairwise comparison of the total absolute abundance of microbial communities in inoculated samples (only the significant treatment combinations are shown in the table). Significance levels for each treatment combination are indicated by *p < 0.05; **p < 0.01; ***p < 0.001.

Comparisons	diff	lwr	upr	*p* adjusted	Significance
Ambient environment, mid-season, and resistant cultivar vs. Ambient environment, end of the season, and susceptible cultivar	-0.24672431	-0.43374918	-0.05969943	0.0059222	**
Elevated ozone, mid-season, and resistant cultivar vs. Ambient environment, end of the season, and susceptible cultivar	-0.31988045	-0.50690532	-0.13285558	0.0004468	***
Elevated ozone, mid-season, and resistant cultivar vs. Ambient environment, end of the season, and resistant cultivar	-0.19830876	-0.38533363	-0.01128389	0.0336884	*
Elevated ozone, mid-season, and resistant cultivar vs. Ambient environment, mid-season, and susceptible cultivar	-0.22308397	-0.41010884	-0.0360591	0.0139018	*
Elevated ozone, end of the season, and susceptible cultivar vs. Ambient environment, mid-season, and resistant cultivar	0.26977115	0.08274628	0.45679602	0.0025868	**
Elevated ozone, mid-season, and susceptible cultivar vs. Ambient environment, mid-season, and resistant cultivar	0.21757628	0.03055141	0.40460115	0.0169484	*
Elevated ozone, mid-season, and resistant cultivar vs. Elevated ozone, end of the season, and susceptible cultivar	-0.34292729	-0.52995217	-0.15590242	0.0002051	***
Elevated ozone, mid-season, and resistant cultivar vs. Elevated ozone, mid-season, and susceptible cultivar	-0.29073242	-0.4777573	-0.10370755	0.0012295	**

Table S 2-8: Relative abundance of top 18 bacterial genera across the samples based on Kraken2 and Bracken

Sample id	Cultivar	Inoculation status	Time of sampling	Environment	Relative abundance of top 10 bacterial genera									
					Xanthomonas	**Pseudomonas**	**Pantoea**	**Methylobacterium**	**Sphingomonas**	**Methylorubrum**	**Stenotrophomonas**	**Microbacterium**	**Leclercia**	**Brevundimonas**
1XM	Resistant	Inoculated	Mid-season	Ambient	0.021	0.142	0.166	0.252	0.011	0.056	0.003	0.079	0.005	0.008
1XE	Resistant	Inoculated	End of the season	Ambient	0.632	0.131	0.051	0.057	0.044	0.019	0.010	0.002	0.000	0.014
2XM	Resistant	Inoculated	Mid-season	Elevated ozone	0.166	0.027	0.513	0.056	0.008	0.009	0.005	0.036	0.011	0.009
2XE	Resistant	Inoculated	End of the season	Elevated ozone	0.583	0.172	0.021	0.065	0.034	0.028	0.006	0.006	0.015	0.011
3XM	Resistant	Inoculated	Mid-season	Elevated ozone	0.017	0.194	0.117	0.105	0.023	0.027	0.105	0.086	0.055	0.031
3XE	Resistant	Inoculated	End of the season	Elevated ozone	0.871	0.047	0.011	0.012	0.020	0.003	0.009	0.001	0.001	0.002
4XM	Resistant	Inoculated	Mid-season	Ambient	0.049	0.100	0.053	0.119	0.033	0.043	0.010	0.134	0.024	0.032
4XE	Resistant	Inoculated	End of the season	Ambient	0.831	0.076	0.009	0.018	0.016	0.007	0.003	0.000	0.000	0.004
5XM	Resistant	Inoculated	Mid-season	Ambient	0.153	0.195	0.054	0.264	0.015	0.080	0.007	0.036	0.005	0.022
5XE	Resistant	Inoculated	End of the season	Ambient	0.705	0.078	0.020	0.081	0.051	0.032	0.007	0.001	0.000	0.006
6XM	Resistant	Inoculated	Mid-season	Elevated ozone	0.067	0.089	0.229	0.183	0.034	0.056	0.004	0.102	0.026	0.019
6XE	Resistant	Inoculated	End of the season	Elevated ozone	0.331	0.276	0.032	0.107	0.054	0.018	0.002	0.011	0.080	0.010
7XM	Resistant	Control	Mid-season	Elevated ozone	0.002	0.159	0.177	0.152	0.047	0.024	0.037	0.052	0.004	0.061
7XE	Resistant	Control	End of the season	Elevated ozone	0.001	0.396	0.174	0.085	0.114	0.020	0.002	0.005	0.006	0.015
8XM	Resistant	Control	Mid-season	Ambient	0.003	0.165	0.089	0.227	0.025	0.073	0.013	0.073	0.003	0.053
8XE	Resistant	Control	End of the season	Ambient	0.012	0.338	0.090	0.326	0.087	0.074	0.003	0.009	0.001	0.018
9XM	Resistant	Control	Mid-season	Ambient	0.002	0.073	0.019	0.096	0.048	0.013	0.555	0.015	0.008	0.026
9XE	Resistant	Control	End of the season	Ambient	0.002	0.157	0.074	0.351	0.113	0.123	0.025	0.011	0.001	0.024
10XM	Resistant	Control	Mid-season	Elevated ozone	0.098	0.348	0.153	0.068	0.057	0.008	0.050	0.027	0.010	0.016
10XE	Resistant	Control	End of the season	Elevated ozone	0.159	0.295	0.231	0.096	0.143	0.007	0.009	0.004	0.000	0.012
11XM	Resistant	Control	Mid-season	Elevated ozone	0.004	0.339	0.088	0.230	0.016	0.008	0.002	0.035	0.006	0.009
11XE	Resistant	Control	End of the season	Elevated ozone	0.007	0.284	0.434	0.121	0.049	0.016	0.006	0.001	0.000	0.004
12XM	Resistant	Control	Mid-season	Ambient	0.006	0.060	0.056	0.409	0.026	0.051	0.047	0.028	0.005	0.062
12XE	Resistant	Control	End of the season	Ambient	0.004	0.406	0.046	0.218	0.063	0.127	0.027	0.008	0.001	0.030
OXC	Resistant	Control	Base	Green house	0.004	0.005	0.002	0.001	0.001	0.000	0.002	0.000	0.000	0.000
OXI	Resistant	Inoculated	Base	Green house	0.136	0.046	0.019	0.019	0.009	0.001	0.005	0.001	0.000	0.003
1EM	Susceptible	Inoculated	Mid-season	Ambient	0.878	0.024	0.004	0.012	0.002	0.005	0.002	0.001	0.000	0.007
1EE	Susceptible	Inoculated	End of the season	Ambient	0.978	0.013	0.000	0.001	0.000	0.000	0.000	0.018	0.079	0.010
2EM	Susceptible	Inoculated	Mid-season	Elevated ozone	0.893	0.011	0.003	0.010	0.001	0.005	0.001	0.001	0.021	0.005
2EE	Susceptible	Inoculated	End of the season	Elevated ozone	0.940	0.038	0.002	0.000	0.000	0.000	0.005	0.020	0.005	0.006
3EM	Susceptible	Inoculated	Mid-season	Elevated ozone	0.836	0.009	0.105	0.009	0.001	0.004	0.001	0.053	0.000	0.004
3EE	Susceptible	Inoculated	End of the season	Elevated ozone	0.955	0.027	0.001	0.000	0.001	0.000	0.001	0.067	0.003	0.004
4EM	Susceptible	Inoculated	Mid-season	Ambient	0.861	0.048	0.008	0.023	0.002	0.010	0.004	0.004	0.000	0.002
4EE	Susceptible	Inoculated	End of the season	Ambient	0.966	0.024	0.001	0.001	0.001	0.000	0.001	0.023	0.001	0.003
5EM	Susceptible	Inoculated	Mid-season	Ambient	0.796	0.027	0.010	0.037	0.001	0.017	0.005	0.013	0.002	0.002
5EE	Susceptible	Inoculated	End of the season	Ambient	0.949	0.029	0.001	0.001	0.000	0.000	0.002	0.044	0.004	0.002
6EM	Susceptible	Inoculated	Mid-season	Elevated ozone	0.911	0.012	0.007	0.023	0.001	0.010	0.002	0.000	0.001	0.001
6EE	Susceptible	Inoculated	End of the season	Elevated ozone	0.965	0.022	0.000	0.000	0.000	0.000	0.001	0.018	0.000	0.002
7EE	Susceptible	Control	End of the season	Elevated ozone	0.101	0.602	0.137	0.035	0.035	0.022	0.008	0.002	0.000	0.001
8EE	Susceptible	Control	End of the season	Ambient	0.319	0.170	0.199	0.088	0.047	0.020	0.033	0.001	0.000	0.000
9EM	Susceptible	Control	Mid-season	Ambient	0.004	0.091	0.032	0.481	0.043	0.093	0.029	0.002	0.001	0.000
9EE	Susceptible	Control	End of the season	Ambient	0.032	0.501	0.288	0.047	0.031	0.007	0.019	0.001	0.000	0.000
10EM	Susceptible	Control	Mid-season	Elevated ozone	0.043	0.054	0.047	0.091	0.036	0.010	0.006	0.008	0.003	0.033
10EE	Susceptible	Control	End of the season	Elevated ozone	0.125	0.520	0.091	0.047	0.078	0.005	0.005	0.069	0.005	0.046
11EM	Susceptible	Control	Mid-season	Ambient	0.129	0.261	0.083	0.205	0.026	0.010	0.006	0.035	0.001	0.016
11EE	Susceptible	Control	End of the season	Ambient	0.001	0.317	0.072	0.213	0.154	0.063	0.060	0.003	0.063	0.016
12EM	Susceptible	Control	Mid-season	Ambient	0.033	0.104	0.061	0.277	0.032	0.082	0.026	0.018	0.000	0.016
OEC	Susceptible	Control	Base	Green house	0.006	0.017	0.094	0.001	0.001	0.000	0.006	0.003	0.000	0.000
OEI	Susceptible	Inoculated	Base	Green house	0.027	0.030	0.005	0.003	0.005	0.000	0.025	0.000	0.000	0.000

Table S 2-9: Normalized RPKS (Reads Per Kilobase of Sequence) value of eukaryotes genera across samples obtained from the EukDetect pipeline.

Sample id	Cultivar	Inoculation status	Time of sampling	Environment	Top 10 eukaryotic genera									
					Pseudogymnoascus	Zasmidium	Cladosporium	Aureobasidium	Epicoccum	Protomyces	Meira	Moesziomyces	Bullera	Papiliotrema
1EM	Susceptible	Inoculated	Mid-season	Ambient	0.00	0.03	0.01	0.00	0.00	0.36	0.12	2.70	0.04	0.00
1EE	Susceptible	Inoculated	End of the season	Ambient	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.67	0.00	0.00
1XM	Resistant	Inoculated	Mid-season	Ambient	0.00	0.72	0.19	0.59	0.06	0.60	0.40	1.54	0.15	0.36
1XE	Resistant	Inoculated	End of the season	Ambient	0.00	0.17	0.34	0.55	1.78	0.60	0.00	1.60	1.81	0.55
2EM	Susceptible	Inoculated	Mid-season	Elevated ozone	0.00	0.04	0.06	0.00	0.03	0.74	0.03	8.19	0.23	0.00
2EE	Susceptible	Inoculated	End of the season	Elevated ozone	0.00	0.00	0.00	0.00	0.08	0.02	0.00	0.69	0.07	0.00
2XM	Resistant	Inoculated	Mid-season	Elevated ozone	0.00	0.49	0.83	0.86	0.35	5.58	1.37	36.07	2.17	0.78
2XE	Resistant	Inoculated	End of the season	Elevated ozone	0.00	0.43	1.20	0.65	5.61	0.95	0.05	5.30	2.74	0.94
3EM	Susceptible	Inoculated	Mid-season	Elevated ozone	0.00	0.00	0.03	0.34	0.00	0.52	0.00	2.73	0.09	0.00
3EE	Susceptible	Inoculated	End of the season	Elevated ozone	0.00	0.00	0.00	0.00	0.02	0.00	0.00	0.00	0.00	0.00
3XM	Resistant	Inoculated	Mid-season	Elevated ozone	0.00	0.34	0.56	0.57	0.27	4.84	0.37	11.04	1.29	0.46
3XE	Resistant	Inoculated	End of the season	Elevated ozone	0.00	0.18	0.65	0.45	2.57	0.50	0.02	0.48	4.09	0.37
4EM	Susceptible	Inoculated	Mid-season	Ambient	0.00	0.00	0.00	0.00	0.01	0.39	0.01	1.48	0.04	0.00
4EE	Susceptible	Inoculated	End of the season	Ambient	0.00	0.00	0.00	0.00	0.04	0.04	0.00	0.76	0.04	0.00
4XE	Resistant	Inoculated	End of the season	Ambient	0.00	0.68	0.83	0.60	2.31	0.97	0.10	1.11	2.09	0.48
5EM	Susceptible	Inoculated	Mid-season	Ambient	0.00	0.03	0.00	0.00	0.02	0.35	0.03	1.49	0.08	0.00
5EE	Susceptible	Inoculated	End of the season	Ambient	0.00	0.00	0.00	0.00	0.03	0.03	0.00	0.45	0.10	0.00
5XM	Resistant	Inoculated	Mid-season	Ambient	0.00	0.09	0.18	0.00	0.14	0.36	0.41	1.96	0.12	0.50
5XE	Resistant	Inoculated	End of the season	Ambient	0.00	0.25	0.53	0.00	2.05	1.37	1.64	1.47	2.79	0.53
6EM	Susceptible	Inoculated	Mid-season	Elevated ozone	0.00	0.00	0.00	0.00	0.01	0.06	0.00	2.49	0.04	0.53
6EE	Susceptible	Inoculated	End of the season	Elevated ozone	0.00	0.00	0.00	0.00	0.02	0.00	0.00	0.00	0.05	0.00
6XM	Resistant	Inoculated	Mid-season	Elevated ozone	0.00	0.39	0.65	0.00	0.30	5.24	0.46	22.11	1.00	0.00
6XE	Resistant	Inoculated	End of the season	Elevated ozone	0.00	0.28	1.34	1.67	5.49	1.88	0.00	1.95	4.57	0.74
7EE	Susceptible	Control	End of the season	Elevated ozone	0.00	0.15	0.79	0.53	4.27	0.28	0.00	0.86	4.22	0.39
7XM	Resistant	Control	Mid-season	Elevated ozone	0.00	0.37	0.69	0.46	0.37	2.65	0.27	10.66	1.38	0.66
7XE	Resistant	Control	End of the season	Elevated ozone	0.00	0.12	1.00	0.66	5.34	0.32	0.00	1.22	3.63	0.44
8EE	Susceptible	Control	End of the season	Ambient	0.00	0.10	0.21	0.74	2.16	0.13	0.00	2.89	1.38	0.82
8XM	Resistant	Control	Mid-season	Ambient	0.00	0.41	0.17	0.00	0.55	2.40	0.59	3.42	0.37	0.33
8XE	Resistant	Control	End of the season	Ambient	0.00	0.22	0.28	0.00	1.77	0.24	0.05	3.14	5.29	0.74
9EM	Susceptible	Control	Mid-season	Ambient	0.00	0.36	0.20	0.00	0.11	1.07	0.15	4.55	0.57	0.42
9EE	Susceptible	Control	End of the season	Ambient	0.00	0.04	0.18	0.00	1.19	0.09	0.03	0.92	0.38	0.60
9XM	Resistant	Control	Mid-season	Ambient	0.00	0.50	0.42	0.00	0.34	1.66	0.55	2.77	0.24	0.42
9XE	Resistant	Control	End of the season	Ambient	0.00	0.29	0.47	0.79	2.46	0.79	0.05	2.18	1.71	0.77
10EM	Susceptible	Control	Mid-season	Elevated ozone	0.00	0.96	0.83	0.71	0.56	6.34	0.32	19.23	1.17	0.78
10EE	Susceptible	Control	End of the season	Elevated ozone	0.00	0.01	0.15	0.53	0.30	0.01	0.00	1.17	1.55	0.79
10XM	Resistant	Control	Mid-season	Elevated ozone	0.00	1.17	0.87	1.03	0.68	1.53	12.65	18.08	1.66	0.38
10XE	Resistant	Control	End of the season	Elevated ozone	0.00	0.03	0.31	0.62	0.78	0.08	0.00	1.35	2.29	0.40
11EM	Susceptible	Control	Mid-season	Elevated ozone	0.00	1.77	0.88	0.93	0.83	11.20	0.46	17.09	6.71	0.64
11EE	Susceptible	Control	End of the season	Elevated ozone	0.00	0.57	1.28	1.22	8.07	0.95	0.02	3.14	14.65	2.67
11XM	Resistant	Control	Mid-season	Elevated ozone	0.00	3.17	0.83	0.77	0.70	7.95	0.55	15.08	19.48	0.34
11XE	Resistant	Control	End of the season	Elevated ozone	0.00	0.31	0.69	0.89	4.13	0.29	0.00	2.23	4.82	0.77
12EM	Susceptible	Control	Mid-season	Ambient	0.00	0.28	0.11	0.68	0.15	0.98	0.80	9.53	0.31	0.36
12XM	Resistant	Control	Mid-season	Ambient	0.00	0.30	0.12	0.00	0.20	1.51	0.33	8.85	0.47	0.32
12XE	Resistant	Control	End of the season	Ambient	0.00	0.11	0.15	0.37	0.84	0.13	0.00	4.78	1.79	0.61
OEC	Susceptible	Control	Base	Green house	5.13	0.00	0.00	0.00	0.19	0.00	0.00	0.00	0.00	0.00
OEI	Susceptible	Inoculated	Base	Green house	0.08	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
OXC	Resistant	Control	Base	Green house	3.40	0.00	2.67	0.00	0.31	0.00	0.00	0.00	0.00	0.00
OXI	Resistant	Inoculated	Base	Green house	0.00	1.01	0.00	0.00	0.00	0.07	0.00	0.98	0.00	0.00

Table S 2-10A: Permutational multivariant analysis of variance (PERMANOVA) of microbial pathways based on Bray–Curtis dissimilarities. Significance levels for each treatment combination are indicated by *p < 0.05; **p < 0.01; ***p < 0.001.

Variable	Df	Sum of sq	Mean Sqs.	F.Model	R2	Pr(>F)	Significance
Environment	1	0.0215	0.021499	1.3167	0.02512	0.2584	
Cultivar	1	0.00262	0.002619	0.1604	0.00306	0.7028	
Inoculation	1	0.0019	0.001898	0.1163	0.00222	0.7612	
Time of sampling	1	0.01013	0.01013	0.6204	0.01184	0.4332	
Environment:Cultivar	1	0.00129	0.001292	0.0791	0.00151	0.814	
Environment:Inoculation status	1	0.00161	0.001607	0.0984	0.00188	0.7756	
Cultivar:Inoculation status	1	0.02972	0.029721	1.8202	0.03473	0.1764	
Environment:Time of sampling	1	0.0315	0.031496	1.9289	0.03681	0.1756	
Cultivar:Time of sampling	1	0.00811	0.008114	0.4969	0.00948	0.4892	
Innoculation status:Time of sampling	1	0.00137	0.001366	0.0837	0.0016	0.8046	
Environment:Cultivar:Inoculation status	1	0.00115	0.001146	0.0702	0.00134	0.8326	
Environment:Cultivar:Time of sampling	1	0.00398	0.00398	0.2438	0.00465	0.6352	
Environment:Inoculation status:Time of sampling	1	0.01283	0.012834	0.786	0.015	0.3814	
Cultivar:Inoculation status:Time of sampling	1	0.20899	0.208991	12.7991	0.24422	0.0014	**
Environment:Cultivar:Inoculation status:Time of sampling	1	0.04551	0.045513	2.7873	0.05319	0.104	
Residuals	29	0.47353	0.016329		0.55336		
Total	44	0.85574			1		

Table S 2-10B: Permutational multivariant analysis of variance (PERMANOVA) of microbial genes based on Bray–Curtis dissimilarities. Significance levels for each treatment combination are indicated by *p < 0.05; **p < 0.01; ***p < 0.001.

Variable	Df	Sum Of Sqs	R2	F	Pr(>F)	Significance
Environment	1	0.02586	0.02767	1.1657	0.2854	
Cultivar	1	0.00379	0.00406	0.171	0.7336	
Inoculation	1	0.00401	0.00429	0.1806	0.727	
Time of sampling	1	0.01453	0.01555	0.6552	0.4262	
Environment:Cultivar	1	0.00154	0.00164	0.0696	0.9034	
Environment:Inoculation status	1	0.00295	0.00315	0.1338	0.7818	
Cultivar:Inoculation status	1	0.02723	0.02915	1.241	0.2664	
Environment:Time of sampling	1	0.03592	0.03844	1.6454	0.1956	
Cultivar:Time of sampling	1	0.00812	0.00869	0.3661	0.5712	
Innoculation status:Time of sampling	1	0.00165	0.00176	0.0739	0.895	
Environment:Cultivar:Inoculation status	1	0.00138	0.00147	0.0588	0.9234	
Environment:Cultivar:Time of sampling	1	0.00449	0.0048	0.1976	0.7074	
Environment:Inoculation status:Time of sampling	1	0.01168	0.0125	0.5156	0.4874	
Cultivar:Inoculation status:Time of sampling	1	0.21206	0.22697	11.834	6.00E-04	***
Environment:Cultivar:Inoculation status:Time of sampling	1	0.05069	0.05426	2.8391	0.0954	
Residuals	29	0.51778	0.5542			
Total	44	0.93429	1			

Table S 2-10C: Permutational multivariant analysis of variance (PERMANOVA) of microbial pathways based on Bray–Curtis dissimilarities during mid and end of the season. Significance levels for each treatment combination are indicated by *p < 0.05; **p < 0.01; ***p < 0.001.

Variable	Mid-season			End of the season		
	R2	Pr(>F)	Significance	R2	Pr(>F)	Significance
Environment	0.00088	0.9288		0.00034	0.9744	
Cultivar	0.14094	0.1278		0.29088	0.0194	*
Inoculation	0.04592	0.371		0.02146	0.5004	
Environment:Cultivar	0.01013	0.6776		0.00231	0.8574	
Environment:Inoculation status	0.03693	0.431		0.03231	0.3902	
Cultivar:Inoculation status	0.00561	0.7722		0.00888	0.6762	
Environment:Cultivar:Inoculation status	-0.0003	0.9994		0.00038	0.9798	

Table S 2-11A: Comparison of jaccard index (similarity between the sets of most central nodes) of local network centrality measures between groups. Significance levels for each treatment combination are indicated by *p < 0.05; **p < 0.01; ***p < 0.001.=

Pairwise comparison between networks												
	Ambient vs. Elevated ozone				Inoculated vs. Control				Control and ambient environment vs. Combined stress			
Centrality measures	Jaccard index	p(<=Jaccard)	Significance	p(>=Jaccard)	Jaccard index	p(<=Jaccard)	Significance	p(>=Jaccard)	Jaccard index	p(<=Jaccard)	Significance	p(>=Jaccard)
Degree	0.069	0.000915	***	0.999879	0.056	0.000081	***	0.999991	0.133	0.01223	*	0.996703
Betweenness centrality	0.136	0.002762	**	0.999208	0.163	0.010359	*	0.996387	0.111	0.00059	***	0.999865
Closeness centrality	0.19	0.031715	*	0.986817	0.163	0.010359	*	0.996387	0.163	0.010359	*	0.996387
eigenvector centrality	0.25	0.171435		0.903429	0.25	0.171435		0.903429	0.02	0	***	1
Hub taxa	0.111	0.143068		0.973988	0	0.017342	*	1	0	0.017342	*	1

Table S 2-11B: Microbial hub taxa across treatment groups

Comparison between ambient and elevated ozone		**Comparison between control and inoculated**		**Comparison between control and ambient environment vs. combined stress**	
Ambient	**Ozone**	**Control**	**Inoculated**	**Ambient**	**Combined**
Atlantibacter	Exiguobacterium	Brevibacterium	Enterococcus	Acinetobacter	Alcaligenes
Exiguobacterium	Leucobacter	Curtobacterium	Lactococcus	Altererythrobacter	Empedobacter
Mycolicibacterium	Marmoricola	Duganella	Mammaliicoccus	Asaia	Escherichia
Pectobacterium	Nocardioides	Hymenobacter	Mycobacterium	Brucella	Humibacter
Rhodococcus	Weissella	Mycolicibacterium	Weissella	Kocuria	Lactococcus
				Sorangium	Serratia
				Sphingobacterium	Weissella
				Urbifossiella	Yersinia

Table S 2-11C: Adjusted Rand index (similarity between clustering) across treatment groups

Comparisons	**ARI**	**p-value**
Ambient vs. Elevated ozone	0.027	0.07
Control vs. Inoculated	0.034	0.021
Control and ambient environment vs. Combined stress	0.103	0

Table S 2-11D: Topological features of the taxonomic networks across various treatments

	Ambient vs. Elevated ozone				Control vs. Inoculated				Ambient vs. Combined stress			
Topological features	**Ambient**	**Elevated ozone**	**abs.diff.**	***p*-value**	**Control**	**Inoculated**	**abs.diff.**	***p*-value**	**Control and ambient environment**	**Combined stress**	**abs.diff.**	***p*-value**
Number of components	3	1	2	0.181818	1	1	0	1	1	3	2	0.18981
Clustering coefficient	0.125	0.082	0.043	0.345654	0.12	0.169	0.048	0.410589	0.182	0.121	0.061	0.253746
Moduarity	0.561	0.562	0.002	0.976024	0.544	0.601	0.057	0.168831	0.56	0.611	0.051	0.164835
Positive edge percentage	68.156	65.497	2.659	0.681319	59.563	82.258	22.695	0.102897	83.152	73.62	9.533	0.207792
Edge density	0.036	0.035	0.002	0.904096	0.037	0.038	0.001	0.911089	0.037	0.035	0.002	0.481518
Natural connectivity	0.012	0.012	0	0.723277	0.012	0.013	0	0.535465	0.013	0.013	0	0.713287
Vertex connectivity	1	1	0	1	1	1	0	1	1	1	0	1
Edge connectivity	1	1	0	1	1	1	0	1	1	1	0	1
Average dissimilarity	0.988	0.989	0.001	0.539461	0.989	0.988	0.001	0.669331	0.988	0.989	0.001	0.333666
Average path length	2.702	2.956	0.254	0.703297	2.821	2.765	0.056	0.874126	2.859	3.064	0.205	0.355644

3. CHAPTER THREE: Long-term fertilization and crop management affects soil microbial communities.

Abstract

Soil microbial communities plays a crucial role in plant health and agricultural productivity. Yet, there remains limited knowledge about how soil microbiomes respond to long-term fertilization in various cropping systems. In this study, we investigated the impact of over a century of cover cropping and fertility management on soil microbial community structure, abundance, and activity in 8 treatments and one cropping systems (cotton-corn-wheat-soybean rotation). These treatment includes N, P, and K fertilization, lime, and legume cover crop variations. Our results, obtained through Illumina MiSeq sequencing (16S and ITS rRNA), reveal that addition of legume cover crops in the absence of any chemical fertilizer showed the highest microbial diversity and richness. Although short term N fertilization is associated with lower microbial diversity and richness, but our study indicates that long-term N addition leads to higher bacterial diversity and richness. However, fungal diversity and richness decrease in response to N, suggesting a negative impact on fungal communities. Interestingly, pH significantly influences bacterial communities, but fungal communities appear more resilient to pH changes. Furthermore, soil biomarker analysis uncovers distinct microbial biomarkers associated with different treatments, highlighting the influence of specific cropping and fertility management practices on microbial communities. Despite previous findings that long-term chemical fertilizer use can reduce soil bacterial richness and diversity, potentially disrupting microbial community ecology, our study suggests that bacterial diversity and richness remain

relatively stable in the standard treatment over time. This stability may indicate the adaptation of microbial communities to long-term fertilizer application.

3.1 Introduction

Soil microbes play an integral role in plant growth and development and soil fertility management through nutrient cycling, inhibiting soil-borne plant diseases, determining soil structure, and supporting plant productivity by regulating many ecological processes (Kumar and Verma 2019; Coban et al. 2022). With the recent advancements in genomic and microbial metagenomic technology enabling comprehensive analysis of both culturable and unculturable microorganisms, this can provide insight into soil microbial community structure, diversity, and activity which are considered key indicators of the overall soil health and productivity potential (Fierer 2017; Mbuthia et al. 2015). Soil microbial communities are sensitive to changes as they are constantly exposed to natural fluctuations in environmental conditions (Zhou et al. 2020). Hence, alterations in soil physicochemical properties due to land use and various management practices can influence the composition and structure of microbial communities (Hu et al. 2021; Wang et al. 2019).

The soil microbiome is one of the Earth's most intricate and ever-changing microbiomes (Fierer 2017). These microbial communities are connected as a complex ecological web through positive (e.g., mutualism), negative (e.g., competition), or neutral (e.g., commensalism) interactions. Correlation-based network inference methods have been developed to understand these complex associations to predict potential microbial relationships (Faust and Raes 2012; Röttjers and Faust 2018; Proulx et al. 2005). Bacteria and fungi can have both direct and indirect effects on each other, and they often coexist, forming intricate interaction networks (Boer et al. 2005; Frey-Klett et al. 2011). Co-occurrence patterns of community members are known to be

influenced by various soil properties (Kumar et al. 2018; Liang et al. 2021; Chen et al. 2021). As the soil's physical and chemical properties are influenced by long-term fertilization and inter-kingdom interaction, the identification of putative keystone bacterial and fungal taxa across different fertilizer treatments can lead to understanding of the influence of various fertilizer treatments in microbial coexistence over long durations.

Crop rotation and cover cropping improve soil health and alleviate the repercussions of soil degradation on a temporal scale (Saleem et al. 2020; Dabney et al. 2001; Mendes et al. 2015). The effect of crop rotational sequence and cover crops on soil microbial communities are known to be variable due to the influence of the different plant residues (Venter et al. 2016; Nevins et al. 2018). Several studies have explored the positive effects of cover cropping on the soil microbial biomass and microbial diversity by altering the dynamics of microbial communities, suppressing pathogens and stimulating beneficial microorganisms (Kim et al. 2020a; Iriarte et al. 2007; Manici et al. 2018; Qi et al. 2020). However, this effect by cover crops has been shown to depend on the different species of cover crop used(Calderón et al. 2016; Finney et al. 2017), and their effect was conditioned by various crop and residue management practices, time, and method of termination (Romdhane et al. 2019; Liang et al. 2014; Nevins et al. 2018; Schmidt et al. 2018).Traditionally, legumes have been used as cover crops as they increase the availability of organic carbon and fix atmospheric nitrogen in the soil (Hubbard et al. 2013; Villamil et al. 2006). However, their effect on long-term fertility management and response to microbial diversity is yet to be understood.

To provide a comprehensive understanding of the impact of long-term cropping systems and fertility management on soil health indicators, changes in soil microbial community structure and their relationship with soil characteristics and crop yield were analyzed in a >110 years

cropping and fertilization treatment experiment. We harnessed Illumina MiSeq sequencing, which was used to analyze the prokaryotic 16S rDNA and fungal 18S rDNA from soil across different cropping systems and fertilization regimes, along with complete fertilization as a control. Soil physicochemical properties, bacterial and fungal community composition and diversity, bioindicators, and co-occurrence were determined. In this study, we hypothesized that (i) different fertilization regimes drive the composition and abundance of the soil microbiome, (ii) soil nutrient content is the main determinant for bacterial and fungal structure, whereas winter legumes are more helpful for the development of microbial community. Our findings provide a basis for further understanding the impact of different fertilization and cropping patterns on microbial diversity and soil properties in the long-term. This could provide important information on why reducing the excessive use of fertilizers and increasing the sustainability of currently used fertilizers is needed.

3.2 Materials and Methods

3.2.1 Study site and sampling

The experiment was conducted in Cullars Rotation (32° 35' 15.35'4" N lat" 85° 28' 56.28'2" W lon") (Fig. S3-1A) at Auburn University, Auburn, Alabama. The Cullars Rotation, also known as the Alvis field, was started in 1911 by J. A. Cullars and John P. Alvis. Mr. Cullars and Mr. Alvis allowed Professor George F. Atkinson of the Agricultural and Mechanical College of Alabama, now Auburn University, to conduct cotton research projects. The experiment has been running for 110 years and is the oldest continuous soil fertility experiment in the southern United States. The experimental site comprises three blocks: east, middle, and west. Each block represented a different stage of a cotton-corn-wheat-soybean rotation. These three blocks were

divided into 14 treatments (Fig. S3-1B). Each experimental plot measured 6.1 x 30.2 m and was separated by a 1 m alley, while a 6 m alley separated the blocks. A selection of eight treatments was utilized for this study from each block. These included the following: plot A (No N fertilizer applied), plot B (no N fertilizer applied or legume cover crop planted), plot C (no soil amendments and no legume cover crop), plot 1 (no legume cover crop with N applied), plot 2 (no P fertilizer applied), plot 3 (complete fertilization), plot 6 (no K fertilizer applied), and plot 8 (no agricultural limestone used).

Soil samples were obtained from each plot on three different dates: April 27, 2020, July 7, 2020, and October 6, 2020. A 2.5 cm diameter soil probe was used for each sample to collect soil from 0-10 cm depth. Around 15 soil cores were taken from each plot and combined in a bucket to create a composite sample. The probes and corresponding buckets were surface-sterilized using 70% ethanol to prevent potential cross-contamination between samples before sampling each plot. To ensure proper handling, the collected soil samples were stored in a cooler until further processing. Immediately after sampling, each sample was sieved to a particle size of 4 mm for subsequent analysis. A portion of the sieved sample was stored at a temperature of 4°C for microbial analysis, while a second portion was stored at -80°C for DNA analysis. Additionally, a third portion of the sample was air-dried at room temperature for chemical and physical analysis. Furthermore, a subset of the air-dried soil was finely ground using a coffee grinder to prepare it for total carbon analysis through dry combustion.

3.2.2 Soil DNA extraction and sequencing

Soil DNA was isolated using the DNeasy PowerSoil kit (Qiagen, Germany). 0.5 grams of moist soil samples were added to the PowerBead tube and subjected to bead beating for 10

minutes (with 30 second break after 5 minutes) using the Mini-Beadbeater-96 (BioSpec, Bartlesville, OK, USA). The manufacturer's instructions were followed for the remaining DNA extraction steps, except for using 50 µl of nuclease-free water for elution in the final step. The concentration and purity of the nucleic acids were evaluated using the Qubit dsDNA high sensitivity assay kits on a Qubit 3.0 Fluorometer (Life Technologies, Carlsbad, CA) and a Nanodrop spectrophotometer 2000 (Thermo Scientific, Wilmington, DE, USA), respectively. The extracted DNA was stored at -80 °C until further processing.

The V3-V4 hypervariable region of the 16S rRNA gene was amplified from the extracted DNA using the 341F (CCTACGGGNGGCWGCAG) and 806R (GGACTACNNGGGTATCTAAT) primer. Moreover, the ITS 2 region of the fungal ITS gene was amplified using the primers ITS3-2024F (GCATCGATGAAGAACGCAGC) and ITS4-2409R (TCCTCCGCTTATTGATATGC). The library construction and sequencing were conducted on the Illumina NovaSeq platform at Novogene Corporation (Beijing, China) and paired end reads of 250 bp were generated.

3.2.3 Sequence data processing

Raw reads obtained from MiSeq Illumina were subjected to processing using the high-resolution Divisive Amplicon Denoising Algorithm 2 (DADA2) package (v1.20.0) (Callahan et al. 2016) to generate amplicon sequence variants (ASVs) with minor modifications (filterAndTrim: maxEE =2, truncQ = 5) based on the results of an analysis of mock communities. To estimate the error rates in amplicon sequencing data, learnErros function was used for forward and reverse reads, which represent the probability of a sequencing error occurring at each position in the reads. Chimeric sequences were detected and removed using the consensus method with the remote Bimera Denovo function in DADA2. The taxonomy

assignment was performed with the assignTaxonomy function using the SILVA database (silva_nr99_v138.1) (Quast et al. 2013) for 16S rRNA gene sequences and the UNITE database (sh_general_release_10.05.2021) (Nilsson et al. 2019) for ITS region sequences.

The ASV table, phylogenetic tree, and taxonomic placement were joined in a phyloseq object in R (Team 2022) within R studio (Team 2020) using packages PhyloSeq (v1.30) (McMurdie and Holmes 2013). The PhyloSeq object was filtered to account for ASVs detected in negative and extraction kit controls. ASVs pointing to nonbacterial taxa and unassigned phylum were removed for further bacterial community analysis. Similarly, ASVs pointing to noneukaryotic taxa and assigned phyla were removed for fungal community analyses.

3.2.4 Microbial community analyses

PhyloSeq objects obtained above for bacterial and eukaryotic data were filtered for samples with less than 1000 counts and samples with ASVs with overall abundance higher than 0.005% of the total sampling depth. The Phyloseq object was then rarefied to the smallest sample size for even depth using the "rarefy" command in *vegan* (v2.6-4) (Dixon 2003) R package. The datasets were assessed for normality using the Shapiro-Wilk normality test. Chao1 (Chao 1984) and Shannon diversity (Magurran 2013; Ludwig and Reynolds 1988) index was used to assess the effects of treatments and sampling time on microbial richness and diversity, respectively. Significant differences in alpha diversity indexes between the treatment groups ($p < 0.05$) were determined either by an analysis of variance (ANOVA) or the Kruskal-Wallis rank-sum test, depending on the distribution of the data.

Microbial beta-diversity on the normalized reads counts was assessed by computing weighted UniFrac distance matrices. The ordinations for β-diversity between treatments were

estimated using Non-Metric Multidimensional Scale (NMDS) and principal coordinate analysis (PCoA) based on Bray-Curtis dissimilarity matrices as implemented in the *plot_ordinatio*n() function in PhyloSeq package. The effect of different fertilizer treatments on both the bacterial and fungal microbial community dissimilarity was tested with permutational multivariate analysis of variance (PERMANOVA) followed by pairwise comparisons using adonis2() function in vegan (v2.6-4) (Dixon 2003) R package. To explore the relationships between microbial taxa and soil properties, we employed a constrained analysis of principal coordinates (CAP) (Anderson and Willis 2003). This analysis allowed us to assess the influence of soil properties on microbial taxa composition. The significance of the relationships was evaluated using a permutation test. A pairwise multi-level comparison was performed using *pairwise.adonis()* function of pairwiseAdonis (v0.4.1) (Martinez Arbizu 2020) package to further identify differences among treatments. Multivariate dispersion among the treatments (deviation from centroids) was performed with the function *betadisper()* and *permutest()* from the vegan package.

To Identify significant differentially abundant ASVs at genus level between different fertilizer treatments was conducted using three different tools DESeq2 (1.40.2) (Love et al. 2014), Corncob (v0.3.1) (Martin et al. 2020), and MaAsLin2 (v1.14.1) (Mallick et al. 2021) package in R. Venn diagrams were drawn to illustrate the number of shared ASVs across the different treatment types using the vennDiagram package (v1.7.3).

3.2.5 Biomarker and network analysis

To identify the potential biomarker, the linear discriminant analysis effect size (LEfSe) (Segata et al. 2011) was applied (Wilcoxon $p < 0.05$, logarithmic LDA (linear discriminant

analysis) score > 3) to identify the biomarker across different fertilizer treatments at the genus level. We performed a network analysis using a subset of the taxonomic data to ensure sufficient sequencing depth for capturing most diversity. Specifically, we retained taxa with a relative abundance of at least 0.5% in over 20% of the samples (prevalence criterion). This approach ensured that all samples had adequate sequencing depth to capture the most taxonomic diversity in the analysis. The cross-domain microbial co-occurrence of bacteria and fungi was constructed using SpiecEasi (v1.1.2) (Kurtz et al. 2015) and was compared between the fertilizer treatments using NetCoMi (v1.1.0) (Peschel et al. 2021). To examine community structures among the treatments, we employed the "cluster_fast_greedy" algorithm (Clauset et al. 2004). The association matrices across the treatments were compared using the netAnalyse function within NetCoMi. Hub taxa, which play crucial roles in the networks, were identified using a threshold of 0.95 on eigenvector centrality, and a comparison of the hub taxa across the treatment groups was based on the Jaccard similarity index.

3.3 Results

3.3.1 Microbial diversity in response to fertilization

From soil samples collected from east, middle, and west plots of eight different fertilizer treatments across three times during the season, a total of 72 samples were collected for 16S rRNA and ITS gene high-throughput sequencing. 26,325 16S rRNA and 20,077 ITS rRNA ASVs from 7 taxonomic ranks were obtained at >97% sequence similarity. Samples with less than 1000 counts and ASVs with an overall abundance higher than 0.005% of the total sampling depth were pruned. 2,303 unique 16S rRNA ASVs for bacteria and 2,790 I18S rRNA ASVs for fungi were obtained for downstream analysis.

We observed a significant difference in bacterial Chao1 richness between the treatments ($p < 0.001$), while sampling time and interaction between treatments and sampling time had no significant influences ($p > 0.05$) (Fig. 3-1A, Table S3-1A, Table S3-1B). Bacterial Shannon diversity was significantly affected by treatment (Kruskal-Wallis: $p < 0.01$) and the interaction of treatment and sampling time was also found to be significant (Kruskal-Wallis $p = 0.002$) (Fig. 3-1B, Table S3-1C, S3-1D). However, sampling time had no significant influence on bacterial diversity (Kruskal-Walis: $p > 0.05$). Regarding the fungal community richness, there was a significant difference between treatment ($p = 0.02$) and time of sampling ($p < 0.01$). In contrast, the interaction of sampling time and treatment did not have a significant effect on overall richness ($p > 0.05$) (Fig. 3-1C, Table S3-1E, Table S3-1F, Table S3-1G). However, treatment did not influence fungal Shannon diversity ($p > 0.05$). Yet, time of sampling (Kruskal-Walis: $p < 0.001$) and interaction between time of sampling and treatment (Kruskal-Walis: $p = 0.02$) had significant effect on the diversity (Fig. 3-1D, Table S3-1H, Table S3-1I).

3.3.2 Microbial diversity under different fertilizer treatments

Our results based on NMDS ordination of bacterial communities indicated that samples clustered according to the treatment but not by time or by location (Fig. 3-1E). PERMANOVA analysis of the bacterial diversity suggested that the variation in the bacterial community was mainly explained by fertilizer treatment ($R^2 = 0.48$, $p < 0.001$) followed by time of sampling ($R^2 = 0.04$, $p < 0.01$) (Table S3-2A). Bacterial communities associated with treatments No soil amendments, No Lime, No P, and No winter legumes + N clustered distinctly from other treatments based on NMDS ordination (Fig. 3-1E). The non-significant beta dispersion tests across treatments indicate that the effects of treatments on soil bacterial community composition are likely due to actual differences in community composition rather than group dispersions

(Table S3-2B). Further pairwise comparison of bacterial communities across treatments indicated significant differences among the following pairs: No N/+legume vs Std fertilization (p = 0.012), No N/+legume vs No K (p = 0.028), and No P vs. No lime (p = 0.029) (Table S3-2C). Regarding fungal communities, we only had a significant effect ($p < 0.001$) at sampling time. In contrast, treatments and the interaction between treatment and time had no significant differences in community diversity (Fig. 3-1F, Table S3-2D). Among the time points, overall fungal communities were significantly different in early sampling time when compared to mid (p = 0.001) and end (p = 0.001). At the same time, there were no differences in fungal communities when comparing between mid and end timepoint (Table S3-2E).

3.3.3 Influence of soil chemistry parameters on microbial community composition

Pearson correlation analysis was used to identify the relationship between soil properties in different fertility treatments following the removal of the highly correlated variables (>0.80) (Fig. S3-2, Fig. S3-3). Canonical analysis of principal coordinates (CAP) plot that incorporates both the microbiome composition (Bray–Curtis dissimilarity) and soil health parameters was constructed to determine the interaction between these variables.

The CAP ordination explains approximately 34% (21.3% in CAP1 and 12.6% CAP2) and 3% (1.8% in CAP1 and 1.6% in CAP2) of the total variance observed in bacterial (Fig. 3-2A) and fungal (Fig. 3-2B) communities respectively. The constrained axes capture the bacterial microbiome variation is associated with continuous variables showed to have the strongest influence on bacterial community composition along the primary y-axis (CAP2, 12.6% of variation explained). At the same time, P and ACE primarily drive the community composition in “No Lime” and “No winter legume /+N” treatment (CAP1, 21.3% variance) (Figure 2A).

However, soil chemical properties weakly influenced fungal community diversity (CAP1, 1.8% of variation; CAP2, 1.6% of variation) (Fig. 3-2B).

All the soil chemistry parameters significantly affected the bacterial community ($p < 0.05$). The pH was the strongest predictor for bacterial communities ($R^2 = 0.70$, $p = 0.001$), followed by ACE ($R^2 = 0.52$, $p = 0.001$) (Table S3-3). However, none of the soil health parameters significantly affected fungal diversity, although TC was a weak predictor for fungal diversity ($R^2 = 0.532$, $p = 0.056$) (Table S3-3).

3.3.4 Bacterial community composition among different treatments

Given the PERMANOVA significance test results, we identified taxa that have driven changes in soil bacterial communities between treatments. Proteobacteria is the dominant bacterial phylum across all treatments, followed by Acidobacteria, Actinobacteria, and Chloroflexi (Fig. S3-4A). Members of the bacterial candidate phylum WPS-2 (or Eremiobacterota) were abundant only in "No Lime" and "No soil amendment" treatments. In terms of fungal community, phylum Ascomycota was the most dominant phylum across all treatment, while Basidiomycota, Glomeromycota, and Monoblepharomycota were seen in low abundance during the end sampling (Fig. S4B). In terms of genus, bacterial genus *Sphingomonas*, *Rhodoplanes*, *Microvirga*, *Candidatus* koribacter, and *Bryobacter* are dominant across all treatments (Fig. 3-3A). In terms of fungal genera, *Fusarium*, *Cladosporium*, *Talaromyces*, *Penicillium*, and *Humicola* are dominant across all treatments (Fig. 3-3B).

Further analysis using three methods, DESeq2, MaAsLin2, and Corncob, for assessing differential enrichment of taxa among various fertilizer treatments led to the identification of several ASVs enriched across the treatment groups at the family level (Fig. S3-5). A single ASV

belonging to Xanthobacteraceae was enriched compared to the treatments “No N/+legumes” and “No N/ No legume”. Comparison of the fertilizer treatment “No P” with “Std Fertilization”, ASVs belonging to Beijerinckiaceae, Xanthobacteraceae, Rhizobiaceae, and Nitrososphaeraceae were differently abundant. Intriguingly, neither “No Lime” nor “No soil amendments” treatments have low pH environments (pH range 4.5-5.0). Yet, the community compositions are distinct in the two treatments, as indicated by the beta-diversity analyses above. The differential abundance analysis between “No Lime” and “No soil amendments” resulted in 132 differentially abundant ASVs mostly belonging to Ktedonobacteraceae, Xanthobacteraceae, Bryobacteraceae, Acetobacteraceae, Koribacteraceae (*Candidatus* Koribacter), Ktedonobacteraceae, Bryobacteraceae (*Bryobacter*), Solibacteraceae (*Candidatus* Solibacter), and Sphingomonadaceae (*Sphingomonas*). ASVs belonging to Ktedonobacteraceae, Xanthobacteraceae (*Rhodoplanes*), and Nitrososphaeraceae (*Candidatus Nitrocosmicus*) were differentially abundant when compared between “No K” and “Standard fertilization”. The effect of nitrogen suggests ASVs belonging to Bryobacteraceae (*Bryobacter*), Xanthobacteraceae, Acetobacteraceae, Ilumatobacteraceae (*Ilumatobacter*), Ktedonobacteraceae (1921-3), Streptosporangiaceae, Koribacteraceae (*Candidatus* Koribacter), Nitrososphaeraceae, and Nitrosotaleaceae (*Candidatus* Nitrosotalea) were differentially abundant when compared to treatments between “No winter legume /+N” and “No N/No legume treatments”.

3.3.5 Effect of fertilizer treatment on soil microbial biomarkers

A Linear Discriminant Effect Size (LEfSe) analysis was performed to identify and select distinct microbial taxa significantly associated with each fertilizer treatment. The LEfSe analysis with "all-against-all" computation for all the fertilizer treatments identified 71 bacterial (Fig. 3-4A) and 42 fungal biomarkers (Fig. 3-4B). In “Std Fertilization” treatment, bacterial genera

Lamia, Pseudolabrys, *Candidatus* Nitrocosmicus and fungal genera *Stagonosporopsis, Humicola, Gibellulopsis, Striaticonidium* and *Nectriopsis* were identified as biomarkers. In "No winter legume+N" treatment, bacterial genera *Devosiaceae* and *Candidatus* Nitrososphaera, and fungal genera *Penicillium* were identified as biomarkers. Bacterial genera *Acidothermus*, *Actinomadura,* and *Rhodoplanes* and fungi belonging to genera *Aspergillus*, *Arthocladium*, *Chaetomium, Arcopilus, Macrophomina, stagonospora, Arcopilus*, and *Collariella* were identified as biomarkers under "No soil amendment" treatment. The fertilizer treatment "No N/+legume" had the highest number of both bacterial and fungal biomarkers across all the treatment where bacterial genera *Candidatus* Nitrosotalea, *Bryobacter, Roseiarcus*, and *Occallatibacter* were identified as biomarkers. Fungal genera *Neophaeococcomyces, Polyschema, Ochroconis, Stainwardia, Phaeoacremonium, Didymocyrtis*, and *Dictyosporium* were identified as biomarkers. In the absence of nitrogen ("No N/No legume"), bacterial genera *Pedomicrobium, Methyloceanibacter, Actinocorallia, Nonomuraea, Streptosporangium, Sphaerisporangium,* and *Dongia* were the biomarkers. Similarly, fungal genera *Paraphoma* and *Paraphaeosphaeria* were identified as biomarkers. The addition of legume as a source of nitrogen ("No N/+legume") identifies bacterial family Xanthobacteraceae and Reyranellaceae and fungal genera *Fusarium, Codinaea,* and *Purpureocillium* as bioindicators. In "No K" treatment, the bacterial family Koribacteraceae, Solibacteraceae, fungal genera *Emericellopsis*, and *Paramyrothecium* were identified as biomarkers. Comparison of microbial markers across the standard fertilizer treatment leads to the identification of various biomarkers in "No P" soil (22 markers), "No K" (26 markers), and "No soil amendments" (74 markers) (Fig. S3-6A, S3-6C, S3-6D). Similarly, comparing the treatments to see the effect of legume ("No N/+legume" vs. "No N/No legume") leads to identifying 13 biomarkers (Fig. S3-6B). In comparison, the

effect of N (“No N/No legume” vs. “No winter legume +N”) suggests 38 biomarkers (Fig. S3-6E). The low pH soil treatments (“No Lime” vs. “No soil amendments”) led to the identification of 46 biomarkers (Fig. S3-6F).

3.4 Discussion

Agricultural viability in the southeastern United States is threatened by land degradation. The soils in this region, initially developed under high-rainfall forests, have deteriorated over time due to intensive land use, becoming acidic, infertile, and degraded (Sikora and Moore 2014). The repeated application of fertilizers has the potential to induce lasting alterations in soil functionality and quality. This can encompass changes in the soil's physical, chemical, and biological characteristics, as well as modifications in nutrient availability. These changes, which can persist for over a century under consistent treatment, can profoundly impact soil fertility.

Despite its significance, the comprehension of how long-term soil management practices influence shifts in microbial communities remains incompletely recognized. The present study delved into this aspect by investigating the impacts of different long-term fertilization approaches on crop yields and the soil's physicochemical and microbiological attributes. The timing of soil sample collection did not demonstrate a noteworthy impact on soil bacterial diversity and richness. Conversely, distinct fertilizer treatments exhibited a significant influence on these factors. The absence of potassium (K) displayed no discernible effect on the diversity and richness of bacterial microbial communities. However, the "No lime" treatment exhibited the lowest levels of diversity and richness across all time points. Interestingly, the "No soil amendments" treatment, despite having lower calcium levels than the "No lime" treatment,

showed higher bacterial diversity and richness. This suggests that the quantity of calcium might not directly dictate bacterial diversity and richness. The reduced pH associated with the "No lime" treatment could be responsible for the decreased bacterial diversity and richness. Similar findings of higher bacterial diversity in neutral pH and lower in acidic soil samples suggest the negative effect of lower pH (Wu et al. 2017; Griffiths et al. 2011; Ren et al. 2018). Despite having a lower pH, higher bacterial diversity in no soil amendments was noteworthy. The heightened diversity observed in the "No soil amendments" treatment could be attributed to winter legumes. Crop rotation with legumes improves soil's physical, chemical, and biological properties (Aschi et al. 2017).

Notably, there were significant disparities in bacterial diversity and richness between the "No winter legumes +N" and "No N/No legumes" treatments. The introduction of nitrogen led to increased bacterial diversity and richness. Prolonged nitrogen addition, however, has been associated with reduced bacterial richness due to induced soil acidification and decreased soil microbial biomass (Dai et al. 2018; Wang et al. 2023; Yang et al. 2020). Nonetheless, the impact of nitrogen deposition has been recognized as seasonal (Huang et al. 2021) and dependent on fertilizer gradient (Wan et al. 2021). Thus, the heightened bacterial diversity observed in the nitrogen-added treatment might be linked to factors like soil moisture, fertilizer quantity, and fluctuations in soil pH during sampling.

Regarding fungal richness, nitrogen fertilization demonstrated a significant decrease in richness. Although not statistically significant, fungal diversity declined with nitrogen treatment during the early and mid-season. This echoes similar observations of varied responses in fungal diversity and richness across different nitrogen treatments (Zhou et al. 2016; Lin et al. 2023;

Liao et al. 2021), underscoring the influence of soil type, fertilizer quantity, and crop involvement in shaping microbial diversity.

In addressing the global concern of soil acidification and enhancing crop productivity, lime application emerges as a promising agricultural approach. Bacterial diversity and richness were notably lower in the "No Lime" treatment (pH = 4.51) than in other fertilizer treatments. Standard fertilization (pH = 6.73) resulted in a significant elevation in bacterial diversity through pH increment. Our findings align with the previously posited hypothesis that soil acidification can constrain the growth of a wide array of bacterial species (Xie et al. 2023). Interestingly, there appeared to be no variance in fungal community diversity and richness across treatments with varying pH, suggesting fungal communities might be resilient to pH changes (Rousk et al. 2010). Similar observations of fungal communities exhibiting resilience against diverse fertilizer treatments have been documented (Wen et al. 2020).

Cationic antagonism is known to exist between calcium (Ca), magnesium (Mg), and potassium (K), whereby elevated levels of one or more of these nutrients can lead to reduced uptake of another nutrient (Rhodes et al. 2018; Garcia et al. 1999; Toumi et al. 2016). In scenarios lacking potassium fertilization, a notable increase in bacterial diversity and richness was observed compared to standard fertilization practices. Notably, the soil's calcium (740 mg/kg) and magnesium (86.8 mg/kg) content proved significantly higher in the "No K" treatment when contrasted with complete fertilization, underscoring the antagonistic interactions of these cations with potassium. The heightened bacterial diversity within this treatment might be attributed to the greater availability of calcium and magnesium in the soil, while the fungal communities remained unaffected. Similar findings of magnesium's positive impact on bacterial diversity have been documented in tea cultivation (Yang et al. 2023). Moreover, positive

influences on bacterial diversity and composition have been observed in soils amended with calcium, as indicated by previous studies (Sridevi et al. 2012; Silveira et al. 2021).

Our study revealed Actinobacteria, Proteobacteria, Acidobacteria, and Chloroflexi as the prevailing entities among the bacterial species. In contrast, the presence of the Firmicutes phylum was noted in limited quantities across various samples. Acidobacteria, recognized as oligotrophic bacteria, commonly thrive in environments with low nutrient availability and high soil acidity. They also can degrade complex and resistant carbon sources (Fierer et al. 2003; Kalam et al. 2020). Our study also indicates that Acidobacteria exhibited the highest abundance in soils subjected to the "No amendment" treatment, as opposed to the "No lime" treatment, despite both treatments exhibiting lower pH levels. While a positive correlation between Acidobacteria and pH has been established (Kalam et al. 2020), it is noteworthy that a significantly negative correlation with nitrogen fertilizer treatments has also been reported (Ren et al. 2020). Specifically, the "No Lime" and "No Soil amendments" treatments solely demonstrated the presence of the candidate bacterial phylum WPS-2 (now identified as *Candidatus* Eremiobacterota). This abundant candidate phylum in soils remains uncultivated and encompasses bacteria displaying diverse metabolic capabilities (Sheremet et al. 2020). Existing research has illustrated that soils characterized by coarse textures, acidic conditions ($pH < 6$), and reduced fertility exhibit notably elevated proportions of Amplicon Sequence Variants (ASVs) affiliated with *Candidatus* Eremiobacterota (Ji et al. 2021; Kim et al. 2019; Ward et al. 2019). Regarding fungal taxa, the Ascomycota species emerged as the dominant fungal entity across all examined samples. In contrast, the presence of the Basidiomycota phylum was observed in specific fertilizer treatments during the concluding stages of the growing season. These two

fungal phyla, Ascomycota and Basidiomycota, command prominence in agricultural soils (Ding et al. 2017; Egidi et al. 2019).

Soil microbial biomarkers help to understand how microorganisms respond to changes in environmental conditions, land management practices, and disturbances. Lefse analysis to identify the microbial biomarkers across various treatments reveals the prevalence of the *Candidatus* Koribacter and *Candidatus* Solibacter genera (both belonging to the Acidobacteria phylum) in soil treatments "No K." Despite their widespread presence, the physiological and ecological roles of the Acidobacteria phylum remain enigmatic due to the challenges associated with culturing these microorganisms (Kielak et al. 2016). Genomic investigations have indicated these bacteria's involvement in nitrogen (N_2) cycling through processes such as nitrate, nitrite, and potentially nitric oxide reduction (Ward et al. 2009). Our findings also shed fresh insights into the role of Acidobacteria in nutrient and carbon cycling within agriculture-affected soils characterized by low potassium and high phosphorus levels. Likewise, fungal genera *Emericellopsis* and *Paramyrothecium* enrichment were observed in treatments lacking potassium. The *Emericellopsis* fungi have long been recognized as producers of antimicrobial peptides from the peptaibol group, with the potential to combat resistant pathogens (Baranova et al. 2017). In the absence of soil amendments, specific microbial taxa, including *Acidothermus*, *Actinomadura, Rhodoplanes*, and several families such as Acidobacteriae, Thermonosporaceae, and Ktedonobacteraceae, were identified as biomarkers. *Acidothermus* bacteria are renowned for thriving in acidic conditions and breaking down plant tissues (Ogola et al. 2021; Kim et al. 2016). *Rhodoplanes* spp. are associated with nitrate reduction and denitrification processes that improve nitrogen availability and soil enhancement (Buckley et al. 2007). In the "No N/+legume" treatment, biomarkers included N-fixing soil genera like *Bradyrhizobium,*

Hyphomicrobium, Reyranella, and ASVs from the family Xanthobacteraceae. These methanotrophic bacteria are recognized for their ability to nodulate and fix nitrogen in association with legumes (Sprent et al. 2017; Bender et al. 2022), implying that in the absence of inorganic nitrogen, these bacterial genera become enriched to facilitate nitrogen fixation from legume crops. Under standard fertilization, bacterial ASVs from the genera *Lamia, Pseudolabrys, Candidatus* Nitrocosmicus, and fungal genera *Humicola, Nitriopsos, Gibellulopsis*, and *Stagonosporopsis* were identified as microbial biomarkers. Members of the *Pseudolabrys* genus are known for their capacity to metabolize organic acids (Miao et al. 2019; Kämpfer et al. 2006) and are associated with enrichment in healthy soils and potassium-deficient treatments (Wang et al. 2017; Eo and Park 2016). The *Candidatus* Nitrosocosmicus genus plays a significant role in soil ammonia oxidation, particularly in nitrogen fertilizer input conditions (Hink et al. 2017; Prosser and Nicol 2012; Hink et al. 2018). The *Pseudolabrys* genus has been reported as a beneficial biocontrol microorganism in the plant rhizosphere, capable of forming symbiotic relationships with plant roots to facilitate nitrogen fixation (Kämpfer et al. 2006). The *Humicola* fungal genus is a potential biological control agent against plant diseases (Yang et al. 2014; Ko et al. 2011). However, the identification of pathogenic biomarkers, such as *Stagonosporopsis* (Stewart et al. 2015) and *Striaticonidium* (Lombard et al. 2016), suggests that antagonistic activities related to phytopathogens are influenced by specific ecological conditions or variations in the functional relationships between species or hosts. The NMDS analysis yielded distinct bacterial community compositions influenced by fertilization. Notably, the fertilizer treatments without lime, potassium, phosphorous, and soil amendments formed different clusters separate from the other samples. Remarkably, even though the treatments “No soil amendments” and “No Lime” exhibited lower pH, they still exhibited significant differences

in bacterial communities. Furthermore, the absence of inorganic nitrogen and the utilization of cover crops led to unique microbial community patterns compared to standard fertilization, highlighting the potential role of inorganic nitrogen in shaping microbial communities. During the early season, microbial diversity and richness were lower in the legume treatment, gradually increasing as the season progressed. This trend might stem from the rate of cover crop decomposition and subsequent biomass nutrient release, which is closely linked to microbial activity (Murungu et al. 2011). The lower diversity early in the sampling period could indicate limited nutrient availability, with diversity rising as degradation progresses throughout the season.

Long-term use of chemical fertilizers has been shown to diminish soil bacterial richness and diversity, disrupting the ecological balance of microbial communities (Sun et al. 2015; Ramirez et al. 2010). In contrast, our findings suggest that bacterial diversity and richness remained relatively stable in the standard treatment throughout the sampling period, surpassing other nutrient-deficient treatments. The positive soil health data and higher yields associated with this treatment further underscore its stability. Although the treatments involving other fertilizers lacking winter legume demonstrated higher bacterial diversity and richness than standard fertilization, these differences were not statistically significant. The impact of cover crops on soil microbial diversity is known to hinge on factors such as climate, cover crop species, termination methods, and tillage practices (Kim et al. 2020b). A similar decrease in microbial diversity due to cover crops during summer was observed in long-term apple experiments (Yang et al. 2019). This decline could be attributed to intra- and inter-kingdom competition from shifts in soil chemical properties, mirroring findings associated with green manure application (Hu et al. 2018; Vida et al. 2020).

Agricultural practices can alter soil mineral composition, influencing microbial community structure and functional diversity (Dai et al. 2020; Carson et al. 2007). Earlier reports have highlighted shifts in soil bacterial and fungal community structures due to various fertilizer treatments in short-term fertility trials, driven by modifications in soil nutrients for microbial growth and colonization (Zeng et al. 2016; Tao et al. 2015). The impacts of inorganic fertilizers can vary based on application duration, with establishing equilibrium within the soil ecosystem requiring considerable time (Moscatelli et al. 2008). Consequently, numerous studies emphasize the importance of long-term field experiments to comprehensively assess the effects of diverse farming systems on soil quality and productivity (Francioli et al. 2016; Hartmann et al. 2015; Merbach and Schulz 2013). While soil pH is commonly recognized as a key factor influencing bacterial diversity, soil fungal diversity tends to be less affected by pH variations (Rousk et al. 2010). Additionally, using autoclaved citrate extractable (ACE) protein has been identified as a potential indicator of soil health (Das et al. 2023; Hurisso et al. 2018).

For sustainable agriculture, it is an utmost need to study how long-term fertilization impacts soil nutrients, microbiomes, and crop productivity in different cropping systems and understand their functional mechanisms. Our study involving century-long cropping and fertility management studies intricate relationships between long-term fertilization, soil microbial communities, and their effects on agricultural systems. We found that adding legume cover crops, even without chemical fertilizers, significantly increased microbial diversity and richness. While short-term nitrogen (N) fertilization was associated with reduced microbial diversity, our results indicate that balanced, long-term N addition led to higher bacterial diversity and richness. However, it's noteworthy that fungal diversity and richness declined with N application, suggesting a detrimental impact on fungal communities. Despite earlier concerns that long-term

chemical fertilizer use might disrupt soil bacterial richness and diversity, our findings indicate that bacterial diversity and richness remained relatively stable in the standard treatment over time. This stability implies that microbial communities have adapted to prolonged fertilizer application. Our research underscores the importance of understanding how soil microbiomes respond to long-term fertilization in diverse cropping systems, providing valuable insights for sustainable agriculture and ecosystem management.

3.5 References for chapter 3

Anderson, M. J., and Willis, T. J. 2003. Canonical Analysis of Principal Coordinates: A Useful Method of Constrained Ordination for Ecology. Ecology. 84:511–525

Aschi, A., Aubert, M., Riah-Anglet, W., Nélieu, S., Dubois, C., Akpa-Vinceslas, M., et al. 2017. Introduction of Faba bean in crop rotation: Impacts on soil chemical and biological characteristics. Appl. Soil Ecol. 120:219–228

Baranova, A. A., Georgieva, M. L., Bilanenko, E. N., Andreev, Ya. A., Rogozhin, E. A., and Sadykova, V. S. 2017. Antimicrobial potential of alkalophilic micromycetes Emericellopsis alkalina. Appl. Biochem. Microbiol. 53:703–710

Bender, F. R., Alves, L. C., da Silva, J. F. M., Ribeiro, R. A., Pauli, G., Nogueira, M. A., et al. 2022. Microbiome of Nodules and Roots of Soybean and Common Bean: Searching for Differences Associated with Contrasting Performances in Symbiotic Nitrogen Fixation. Int. J. Mol. Sci. 23:12035

Boer, W. de, Folman, L. B., Summerbell, R. C., and Boddy, L. 2005. Living in a fungal world: impact of fungi on soil bacterial niche development⋆. FEMS Microbiol. Rev. 29:795–811

Buckley, D. H., Huangyutitham, V., Hsu, S.-F., and Nelson, T. A. 2007. Stable Isotope Probing with 15N2 Reveals Novel Noncultivated Diazotrophs in Soil. Appl. Environ. Microbiol. 73:3196–3204

Calderón, F. J., Nielsen, D., Acosta-martínez, V., Vigil, M. F., and Lyon, D. 2016. Cover Crop and Irrigation Effects on Soil Microbial Communities and Enzymes in Semiarid Agroecosystems of the Central Great Plains of North America. Pedosphere. 26:192–205

Callahan, B. J., McMurdie, P. J., Rosen, M. J., Han, A. W., Johnson, A. J. A., and Holmes, S. P. 2016. DADA2: High-resolution sample inference from Illumina amplicon data. Nat. Methods. 13:581–583

Carson, J. K., Rooney, D., Gleeson, D. B., and Clipson, N. 2007. Altering the mineral composition of soil causes a shift in microbial community structure. FEMS Microbiol. Ecol. 61:414–423

Chao, A. 1984. Nonparametric Estimation of the Number of Classes in a Population. Scand. J. Stat. 11:265–270

Chen, B., Jiao, S., Luo, S., Ma, B., Qi, W., Cao, C., et al. 2021. High soil pH enhances the network interactions among bacterial and archaeal microbiota in alpine grasslands of the Tibetan Plateau. Environ. Microbiol. 23:464–477

Clauset, A., Newman, M. E. J., and Moore, C. 2004. Finding community structure in very large networks. Phys. Rev. E. 70:066111

Coban, O., De Deyn, G. B., and van der Ploeg, M. 2022. Soil microbiota as game-changers in restoration of degraded lands. Science. 375:abe0725

Dabney, S. M., Delgado, J. A., and Reeves, D. W. 2001. Using Winter Cover Crops to Improve Soil and Water Quality. Commun. Soil Sci. Plant Anal. 32:1221–1250

Dai, Z., Liu, G., Chen, H., Chen, C., Wang, J., Ai, S., et al. 2020. Long-term nutrient inputs shift soil microbial functional profiles of phosphorus cycling in diverse agroecosystems. ISME J. 14:757–770

Dai, Z., Su, W., Chen, H., Barberán, A., Zhao, H., Yu, M., et al. 2018. Long-term nitrogen fertilization decreases bacterial diversity and favors the growth of Actinobacteria and Proteobacteria in agro-ecosystems across the globe. Glob. Change Biol. 24:3452–3461

Das, S., Liptzin, D., and Maharjan, B. 2023. Long-term manure application improves soil health and stabilizes carbon in continuous maize production system. Geoderma. 430:116338

Ding, J., Jiang, X., Guan, D., Zhao, B., Ma, M., Zhou, B., et al. 2017. Influence of inorganic fertilizer and organic manure application on fungal communities in a long-term field experiment of Chinese Mollisols. Appl. Soil Ecol. 111:114–122

Dixon, P. 2003. VEGAN, a package of R functions for community ecology. J. Veg. Sci. 14:927–930

Egidi, E., Delgado-Baquerizo, M., Plett, J. M., Wang, J., Eldridge, D. J., Bardgett, R. D., et al. 2019. A few Ascomycota taxa dominate soil fungal communities worldwide. Nat. Commun. 10:2369

Eo, J., and Park, K.-C. 2016. Long-term effects of imbalanced fertilization on the composition and diversity of soil bacterial community. Agric. Ecosyst. Environ. 231:176–182

Faust, K., and Raes, J. 2012. Microbial interactions: from networks to models. Nat. Rev. Microbiol. 10:538–550

Fierer, N. 2017. Embracing the unknown: disentangling the complexities of the soil microbiome. Nat. Rev. Microbiol. 15:579–590

Fierer, N., Schimel, J. P., and Holden, P. A. 2003. Variations in microbial community composition through two soil depth profiles. Soil Biol. Biochem. 35:167–176

Finney, D. M., Buyer, J. S., and Kaye, J. P. 2017. Living cover crops have immediate impacts on soil microbial community structure and function. J. Soil Water Conserv. 72:361–373

Francioli, D., Schulz, E., Lentendu, G., Wubet, T., Buscot, F., and Reitz, T. 2016. Mineral vs. Organic Amendments: Microbial Community Structure, Activity and Abundance of Agriculturally Relevant Microbes are Driven by Long-Term Fertilization Strategies. Front. Microbiol. 7

Frey-Klett, P., Burlinson, P., Deveau, A., Barret, M., Tarkka, M., and Sarniguet, A. 2011. Bacterial-Fungal Interactions: Hyphens between Agricultural, Clinical, Environmental, and Food Microbiologists. Microbiol. Mol. Biol. Rev. MMBR. 75:583–609

Garcia, M., Daverede, C., Gallego, P., and Toumi, M. 1999. Effect of various potassium-calcium ratios on cation nutrition of grape grown hydroponically. J. Plant Nutr. 22:417–425

Griffiths, R. I., Thomson, B. C., James, P., Bell, T., Bailey, M., and Whiteley, A. S. 2011. The bacterial biogeography of British soils. Environ. Microbiol. 13:1642–1654

Hartmann, M., Frey, B., Mayer, J., Mäder, P., and Widmer, F. 2015. Distinct soil microbial diversity under long-term organic and conventional farming. ISME J. 9:1177–1194

Hink, L., Gubry-Rangin, C., Nicol, G. W., and Prosser, J. I. 2018. The consequences of niche and physiological differentiation of archaeal and bacterial ammonia oxidisers for nitrous oxide emissions. ISME J. 12:1084–1093

Hink, L., Nicol, G. W., and Prosser, J. I. 2017. Archaea produce lower yields of N2O than bacteria during aerobic ammonia oxidation in soil. Environ. Microbiol. 19:4829–4837

Hu, X., Liu, J., Zhu, P., Wei, D., Jin, J., Liu, X., et al. 2018. Long-term manure addition reduces diversity and changes community structure of diazotrophs in a neutral black soil of northeast China. J. Soils Sediments. 18:2053–2062

Hu, Y., Wang, Z., Zhang, Z., Song, N., Zhou, H., Li, Y., et al. 2021. Alteration of desert soil microbial community structure in response to agricultural reclamation and abandonment. CATENA. 207:105678

Huang, T., Liu, W., Long, X.-E., Jia, Y., Wang, X., and Chen, Y. 2021. Different Responses of Soil Bacterial Communities to Nitrogen Addition in Moss Crust. Front. Microbiol. 12

Hubbard, R. K., Strickland, T. C., and Phatak, S. 2013. Effects of cover crop systems on soil physical properties and carbon/nitrogen relationships in the coastal plain of southeastern USA. Soil Tillage Res. 126:276–283

Hugerth, L. W., Muller, E. E. L., Hu, Y. O. O., Lebrun, L. A. M., Roume, H., Lundin, D., et al. 2014. Systematic Design of 18S rRNA Gene Primers for Determining Eukaryotic Diversity in Microbial Consortia. PLoS ONE. 9:e95567

Hurisso, T. T., Moebius-Clune, D. J., Culman, S. W., Moebius-Clune, B. N., Thies, J. E., and van Es, H. M. 2018. Soil Protein as a Rapid Soil Health Indicator of Potentially Available Organic Nitrogen. Agric. Environ. Lett. 3:180006

Iriarte, F. B., Balogh, B., Momol, M. T., Smith, L. M., Wilson, M., and Jones, J. B. 2007. Factors Affecting Survival of Bacteriophage on Tomato Leaf Surfaces. Appl. Environ. Microbiol. 73:1704–1711

Ji, M., Williams, T. J., Montgomery, K., Wong, H. L., Zaugg, J., Berengut, J. F., et al. 2021. Candidatus Eremiobacterota, a metabolically and phylogenetically diverse terrestrial phylum with acid-tolerant adaptations. ISME J. 15:2692–2707

Kalam, S., Basu, A., Ahmad, I., Sayyed, R. Z., El-Enshasy, H. A., Dailin, D. J., et al. 2020. Recent Understanding of Soil Acidobacteria and Their Ecological Significance: A Critical Review. Front. Microbiol. 11

Kämpfer, P., Young, C.-C., Arun, A. B., Shen, F.-T., Jäckel, U., Rosselló-Mora, R., et al. 2006. Pseudolabrys taiwanensis gen. nov., sp. nov., an alphaproteobacterium isolated from soil. Int. J. Syst. Evol. Microbiol. 56:2469–2472

Kielak, A. M., Barreto, C. C., Kowalchuk, G. A., van Veen, J. A., and Kuramae, E. E. 2016. The Ecology of Acidobacteria: Moving beyond Genes and Genomes. Front. Microbiol. 7

Kim, M., Lim, H.-S., Hyun, C.-U., Cho, A., Noh, H.-J., Hong, S. G., et al. 2019. Local-scale variation of soil bacterial communities in ice-free regions of maritime Antarctica. Soil Biol. Biochem. 133:165–173

Kim, N., Zabaloy, M. C., Guan, K., and Villamil, M. B. 2020a. Do cover crops benefit soil microbiome? A meta-analysis of current research. Soil Biol. Biochem. 142:107701

Kim, N., Zabaloy, M. C., Guan, K., and Villamil, M. B. 2020b. Do cover crops benefit soil microbiome? A meta-analysis of current research. Soil Biol. Biochem. 142:107701

Kim, S.-K., Chung, D., Himmel, M. E., Bomble, Y. J., and Westpheling, J. 2016. Heterologous expression of family 10 xylanases from Acidothermus cellulolyticus enhances the exoproteome of Caldicellulosiruptor bescii and growth on xylan substrates. Biotechnol. Biofuels. 9:176

Klindworth, A., Pruesse, E., Schweer, T., Peplies, J., Quast, C., Horn, M., et al. 2013. Evaluation of general 16S ribosomal RNA gene PCR primers for classical and next-generation sequencing-based diversity studies. Nucleic Acids Res. 41:e1

Ko, W.-H., Yang, C.-H., Lin, M.-J., Chen, C.-Y., and Tsou, Y.-J. 2011. Humicola phialophoroides sp. nov. from soil with potential for biological control of plant diseases. Bot. Stud. 52.

Kumar, A., and Verma, J. P. 2019. The Role of Microbes to Improve Crop Productivity and Soil Health. In *Ecological Wisdom Inspired Restoration Engineering*, EcoWISE, eds. Varenyam Achal and Abhijit Mukherjee. Singapore: Springer, p. 249–265

Kumar, U., Kumar Nayak, A., Shahid, M., Gupta, V. V. S. R., Panneerselvam, P., Mohanty, S., et al. 2018. Continuous application of inorganic and organic fertilizers over 47 years in paddy soil alters the bacterial community structure and its influence on rice production. Agric. Ecosyst. Environ. 262:65–75

Kurtz, Z. D., Müller, C. L., Miraldi, E. R., Littman, D. R., Blaser, M. J., and Bonneau, R. A. 2015. Sparse and Compositionally Robust Inference of Microbial Ecological Networks. PLOS Comput. Biol. 11:e1004226

Liang, S., Grossman, J., and Shi, W. 2014. Soil microbial responses to winter legume cover crop management during organic transition. Eur. J. Soil Biol. 65:15–22

Liang, Y., Pan, F., Ma, J., Yang, Z., and Yan, P. 2021. Long-term forest restoration influences succession patterns of soil bacterial communities. Environ. Sci. Pollut. Res. 28:20598–20607

Liao, L., Wang, X., Wang, J., Liu, G., and Zhang, C. 2021. Nitrogen fertilization increases fungal diversity and abundance of saprotrophs while reducing nitrogen fixation potential in a semiarid grassland. Plant Soil. 465:515–532

Lin, X., Song, B., Adil, M. F., Lal, M. K., Jia, Q., Wang, Q., et al. 2023. Response of the rhizospheric soil microbial community of sugar beet to nitrogen application: A case of black soil in Northeast China. Appl. Soil Ecol. 191:105050

Lombard, L., Houbraken, J., Decock, C., Samson, R. A., Meijer, M., Réblová, M., et al. 2016. Generic hyper-diversity in Stachybotriaceae. Persoonia - Mol. Phylogeny Evol. Fungi. 36:156–246.

Love, M. I., Huber, W., and Anders, S. 2014. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. 15:550

Ludwig, J. A., and Reynolds, J. F. 1988. *Statistical ecology: a primer in methods and computing*. John Wiley & Sons.

Magurran, A. E. 2013. Measuring biological diversity. Hoboken.

Mallick, H., Rahnavard, A., McIver, L. J., Ma, S., Zhang, Y., Nguyen, L. H., et al. 2021. Multivariable association discovery in population-scale meta-omics studies. PLOS Comput. Biol. 17:e1009442

Manici, L. M., Caputo, F., Nicoletti, F., Leteo, F., and Campanelli, G. 2018. The impact of legume and cereal cover crops on rhizosphere microbial communities of subsequent vegetable crops for contrasting crop decline. Biol. Control. 120:17–25

Martin, B. D., Witten, D., and Willis, A. D. 2020. Modeling microbial abundances and dysbiosis with beta-binomial regression. Ann. Appl. Stat. 14:94–115

Martinez Arbizu, P. 2020. pairwiseAdonis: Pairwise multilevel comparison using adonis. R Package Version 04. 1.

Mbuthia, L. W., Acosta-Martínez, V., DeBruyn, J., Schaeffer, S., Tyler, D., Odoi, E., et al. 2015. Long term tillage, cover crop, and fertilization effects on microbial community structure, activity: Implications for soil quality. Soil Biol. Biochem. 89:24–34

McMurdie, P. J., and Holmes, S. 2013. phyloseq: An R Package for Reproducible Interactive Analysis and Graphics of Microbiome Census Data ed. Michael Watson. PLoS ONE. 8:e61217

Mendes, L. W., Tsai, S. M., Navarrete, A. A., de Hollander, M., van Veen, J. A., and Kuramae, E. E. 2015. Soil-Borne Microbiome: Linking Diversity to Function. Microb. Ecol. 70:255–265.

Merbach, I., and Schulz, E. 2013. Long-term fertilization effects on crop yields, soil fertility and sustainability in the Static Fertilization Experiment Bad Lauchstädt under climatic conditions 2001–2010. Arch. Agron. Soil Sci. 59:1041–1057

Miao, Y., Johnson, N. W., Gedalanga, P. B., Adamson, D., Newell, C., and Mahendra, S. 2019. Response and recovery of microbial communities subjected to oxidative and biological treatments of 1,4-dioxane and co-contaminants. Water Res. 149:74–85

Moscatelli, M. C., Lagomarsino, A., De Angelis, P., and Grego, S. 2008. Short- and medium-term contrasting effects of nitrogen fertilization on C and N cycling in a poplar plantation soil. For. Ecol. Manag. 255:447–454

Murungu, F. S., Chiduza, C., Muchaonyerwa, P., and Mnkeni, P. N. S. 2011. Decomposition, nitrogen and phosphorus mineralization from winter-grown cover crop residues and suitability for a smallholder farming system in South Africa. Nutr. Cycl. Agroecosystems. 89:115–123

Nevins, C. J., Nakatsu, C., and Armstrong, S. 2018. Characterization of microbial community response to cover crop residue decomposition. Soil Biol. Biochem. 127:39–49

Nilsson, R. H., Larsson, K.-H., Taylor, A. F. S., Bengtsson-Palme, J., Jeppesen, T. S., Schigel, D., et al. 2019. The UNITE database for molecular identification of fungi: handling dark taxa and parallel taxonomic classifications. Nucleic Acids Res. 47:D259–D264

Ogola, H. J. O., Selvarajan, R., and Tekere, M. 2021. Local Geomorphological Gradients and Land Use Patterns Play Key Role on the Soil Bacterial Community Diversity and Dynamics in the Highly Endemic Indigenous Afrotemperate Coastal Scarp Forest Biome. Front. Microbiol. 12

Peschel, S., Müller, C. L., Von Mutius, E., Boulesteix, A.-L., and Depner, M. 2021. NetCoMi: network construction and comparison for microbiome data in R. Brief. Bioinform. 22:bbaa290

Prosser, J. I., and Nicol, G. W. 2012. Archaeal and bacterial ammonia-oxidisers in soil: the quest for niche specialisation and differentiation. Trends Microbiol. 20:523–531

Proulx, S. R., Promislow, D. E. L., and Phillips, P. C. 2005. Network thinking in ecology and evolution. Trends Ecol. Evol. 20:345–353

Qi, G., Chen, S., Ke, L., Ma, G., and Zhao, X. 2020. Cover crops restore declining soil properties and suppress bacterial wilt by regulating rhizosphere bacterial communities and improving soil nutrient contents. Microbiol. Res. 238:126505

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:D590–D596

Ramirez, K. S., Craine, J. M., and Fierer, N. 2010. Nitrogen fertilization inhibits soil microbial respiration regardless of the form of nitrogen applied. Soil Biol. Biochem. 42:2336–2338

Ren, B., Hu, Y., Chen, B., Zhang, Y., Thiele, J., Shi, R., et al. 2018. Soil pH and plant diversity shape soil bacterial community structure in the active layer across the latitudinal gradients in continuous permafrost region of Northeastern China. Sci. Rep. 8:5619

Ren, N., Wang, Y., Ye, Y., Zhao, Y., Huang, Y., Fu, W., et al. 2020. Effects of Continuous Nitrogen Fertilizer Application on the Diversity and Composition of Rhizosphere Soil Bacteria. Front. Microbiol. 11

Rhodes, R., Miles, N., and Hughes, J. C. 2018. Interactions between potassium, calcium and magnesium in sugarcane grown on two contrasting soils in South Africa. Field Crops Res. 223:1–11

Romdhane, S., Spor, A., Busset, H., Falchetto, L., Martin, J., Bizouard, F., et al. 2019. Cover Crop Management Practices Rather Than Composition of Cover Crop Mixtures Affect Bacterial Communities in No-Till Agroecosystems. Front. Microbiol. 10

Röttjers, L., and Faust, K. 2018. From hairballs to hypotheses–biological insights from microbial networks. FEMS Microbiol. Rev. 42:761–780

Rousk, J., Bååth, E., Brookes, P. C., Lauber, C. L., Lozupone, C., Caporaso, J. G., et al. 2010. Soil bacterial and fungal communities across a pH gradient in an arable soil. ISME J. 4:1340–1351

Saleem, M., Pervaiz, Z. H., Contreras, J., Lindenberger, J. H., Hupp, B. M., Chen, D., et al. 2020. Cover crop diversity improves multiple soil properties via altering root architectural traits. Rhizosphere. 16:100248

Schmidt, R., Gravuer, K., Bossange, A. V., Mitchell, J., and Scow, K. 2018. Long-term use of cover crops and no-till shift soil microbial community life strategies in agricultural soil. PLOS ONE. 13:e0192953

Segata, N., Izard, J., Waldron, L., Gevers, D., Miropolsky, L., Garrett, W. S., et al. 2011. Metagenomic biomarker discovery and explanation. Genome Biol. 12:R60

Sheremet, A., Jones, G. M., Jarett, J., Bowers, R. M., Bedard, I., Culham, C., et al. 2020. Ecological and genomic analyses of candidate phylum WPS-2 bacteria in an unvegetated soil. Environ. Microbiol. 22:3143–3157

Sikora, F. J., and Moore, K. P. 2014. Soil test methods from the southeastern United States. South. Coop. Ser. Bull. 419:54–58.

Silveira, R., de Mello, T. de R. B., Silva, M. R. S. S., Krüger, R. H., and Bustamante, M. M. da C. 2021. Long-term liming promotes drastic changes in the composition of the microbial community in a tropical savanna soil. Biol. Fertil. Soils. 57:31–46

Sprent, J. I., Ardley, J., and James, E. K. 2017. Biogeography of nodulated legumes and their nitrogen-fixing symbionts. New Phytol. 215:40–56

Sridevi, G., Minocha, R., Turlapati, S. A., Goldfarb, K. C., Brodie, E. L., Tisa, L. S., et al. 2012. Soil bacterial communities of a calcium-supplemented and a reference watershed at the Hubbard Brook Experimental Forest (HBEF), New Hampshire, USA. FEMS Microbiol. Ecol. 79:728–740

Stewart, J. E., Turner, A. N., and Brewer, M. T. 2015. Evolutionary history and variation in host range of three Stagonosporopsis species causing gummy stem blight of cucurbits. Fungal Biol. 119:370–382

Sun, R., Zhang, X.-X., Guo, X., Wang, D., and Chu, H. 2015. Bacterial diversity in soils subjected to long-term chemical fertilization can be more stably maintained with the addition of livestock manure than wheat straw. Soil Biol. Biochem. 88:9–18

Tao, R., Liang, Y., Wakelin, S. A., and Chu, G. 2015. Supplementing chemical fertilizer with an organic component increases soil biological function and quality. Appl. Soil Ecol. 96:42–51

Team, R. C. 2022. R: A language and environment for statistical computing.

Team, Rs. 2020. RStudio: Integrated Development for R. http://www.rstudio.com/

Toumi, M., Nedjimi, B., Halitim, A., and Garcia, M. 2016. Effects of K-Mg ratio on growth and cation nutrition of Vitis vinifera L. cv. “Dattier de Beiruth” grafted on SO4 rootstock. J. Plant Nutr. 39:904–911

Venter, Z. S., Jacobs, K., and Hawkins, H.-J. 2016. The impact of crop rotation on soil microbial diversity: A meta-analysis. Pedobiologia. 59:215–223

Vida, C., de Vicente, A., and Cazorla, F. M. 2020. The role of organic amendments to soil for crop protection: Induction of suppression of soilborne pathogens. Ann. Appl. Biol. 176:1–15

Villamil, M. B., Bollero, G. A., Darmody, R. G., Simmons, F. W., and Bullock, D. G. 2006. No-Till Corn/Soybean Systems Including Winter Cover Crops. Soil Sci. Soc. Am. J. 70:1936–1944

Wan, Y., Li, W., Wang, J., and Shi, X. 2021. Bacterial Diversity and Community in Response to Long-Term Nitrogen Fertilization Gradient in Citrus Orchard Soils. Diversity. 13:282

Wang, R., Zhang, H., Sun, L., Qi, G., Chen, S., and Zhao, X. 2017. Microbial community composition is related to soil biological and chemical properties and bacterial wilt outbreak. Sci. Rep. 7:343

Wang, X., Feng, J., Ao, G., Qin, W., Han, M., Shen, Y., et al. 2023. Globally nitrogen addition alters soil microbial community structure, but has minor effects on soil microbial diversity and richness. Soil Biol. Biochem. 179:108982

Wang, Z., Liu, Y., Zhao, L., Zhang, W., and Liu, L. 2019. Change of soil microbial community under long-term fertilization in a reclaimed sandy agricultural ecosystem. PeerJ. 7:e6497

Ward, L. M., Cardona, T., and Holland-Moritz, H. 2019. Evolutionary Implications of Anoxygenic Phototrophy in the Bacterial Phylum Candidatus Eremiobacterota (WPS-2). Front. Microbiol. 10

Ward, N. L., Challacombe, J. F., Janssen, P. H., Henrissat, B., Coutinho, P. M., Wu, M., et al. 2009. Three Genomes from the Phylum Acidobacteria Provide Insight into the Lifestyles of These Microorganisms in Soils. Appl. Environ. Microbiol. 75:2046–2056

Wen, Y.-C., Li, H.-Y., Lin, Z.-A., Zhao, B.-Q., Sun, Z.-B., Yuan, L., et al. 2020. Long-term fertilization alters soil properties and fungal community composition in fluvo-aquic soil of the North China Plain. Sci. Rep. 10:7198

Wu, Y., Zeng, J., Zhu, Q., Zhang, Z., and Lin, X. 2017. pH is the primary determinant of the bacterial community structure in agricultural soils impacted by polycyclic aromatic hydrocarbon pollution. Sci. Rep. 7:40093

Xie, L., Li, W., Pang, X., Liu, Q., and Yin, C. 2023. Soil properties and root traits are important factors driving rhizosphere soil bacterial and fungal community variations in alpine Rhododendron nitidulum shrub ecosystems along an altitudinal gradient. Sci. Total Environ. 864:161048

Yang, C.-H., Lin, M.-J., Su, H.-J., and Ko, W.-H. 2014. Multiple resistance-activating substances produced by Humicola phialophoroides isolated from soil for control of Phytophthora blight of pepper. Bot. Stud. 55:40

Yang, J., Zhang, T., Zhang, R., Huang, Q., and Li, H. 2019. Long-term cover cropping seasonally affects soil microbial carbon metabolism in an apple orchard. Bioengineered. 10:207–217

Yang, W., Ji, Z., Wu, A., He, D., Rensing, C., Chen, Y., et al. 2023. Inconsistent responses of soil bacterial and fungal community's diversity and network to magnesium fertilization in tea (Camellia sinensis) plantation soils. Appl. Soil Ecol. 191:105055

Yang, Y., Cheng, H., Gao, H., and An, S. 2020. Response and driving factors of soil microbial diversity related to global nitrogen addition. Land Degrad. Dev. 31:190–204

Zeng, J., Liu, X., Song, L., Lin, X., Zhang, H., Shen, C., et al. 2016. Nitrogen fertilization directly affects soil bacterial diversity and indirectly affects bacterial community composition. Soil Biol. Biochem. 92:41–49

Zhou, J., Jiang, X., Zhou, B., Zhao, B., Ma, M., Guan, D., et al. 2016. Thirty four years of nitrogen fertilization decreases fungal diversity and alters fungal community composition in black soil in northeast China. Soil Biol. Biochem. 95:135–143

Zhou, Z., Wang, C., and Luo, Y. 2020. Meta-analysis of the impacts of global change factors on soil microbial diversity and functionality. Nat. Commun. 11:3072

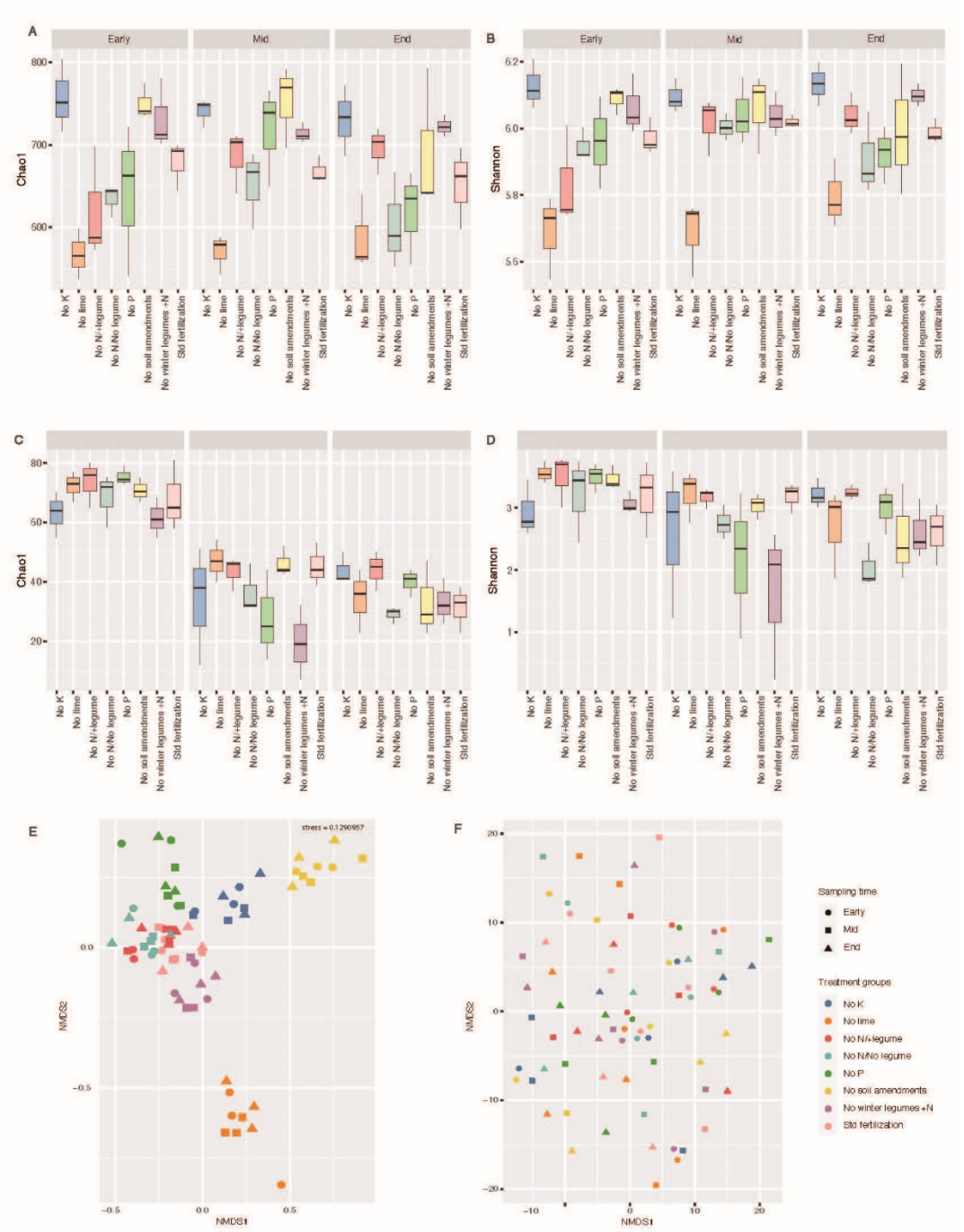

Figure 3-1: Long term fertility treatment has impact on microbial diversity and richness. (A) Bacterial Chao1 richness (B) bacterial Shannon diversity (C) fungal Chao1 richness and (D) fungal Shannon diversity across different treatments. Nonmetric dimensional scaling (NMDS) analysis of (E) bacterial and (F) fungal community.

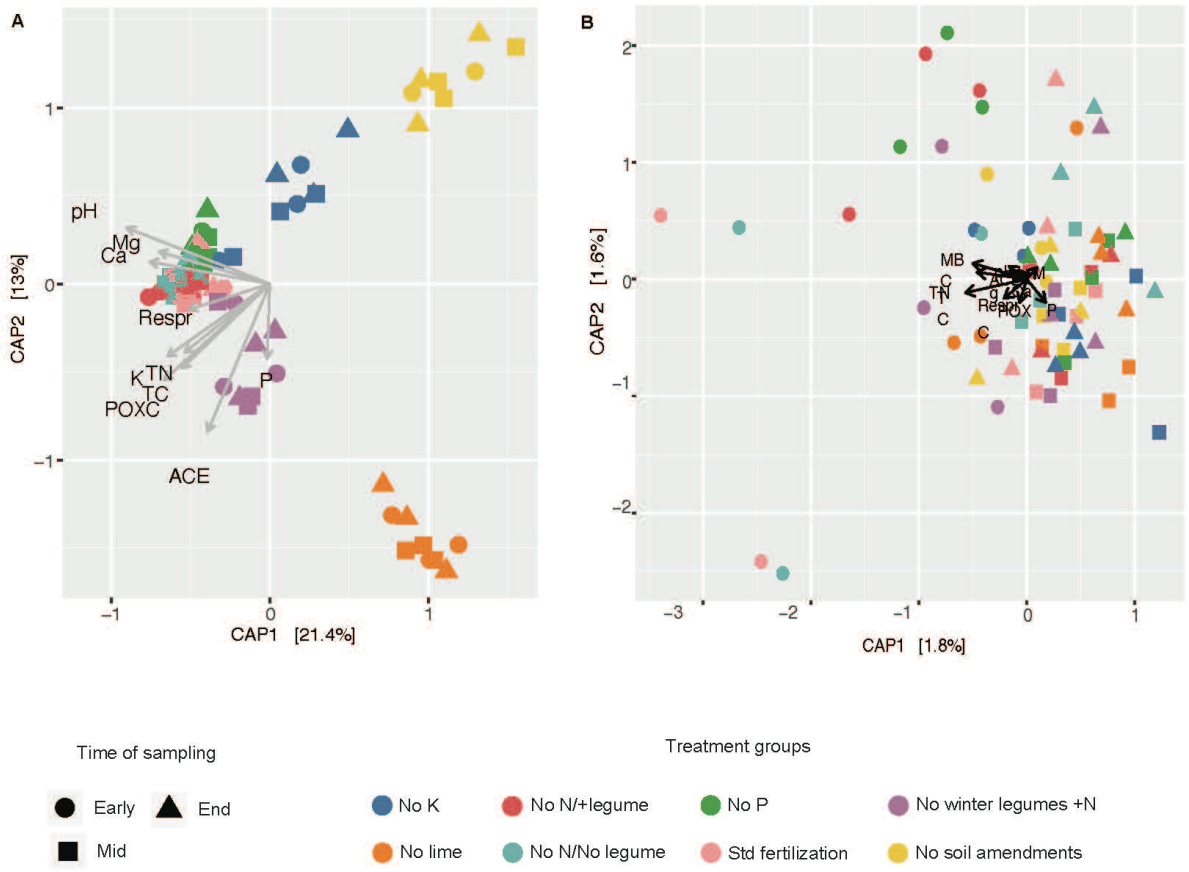

Figure 3-2: Canonical analysis of principal coordinates (CAP) bi-plot ordination (based upon a Bray-Curtis distance) showing canonical axes (CAP1, CAP2) that best discriminate treatments groups of (A) bacterial and (B) fungal communities. The correlation with canonical axes are only shown when the Pearson's correlation coefficient is >0.6. The length of each vector line is proportional to the strength of the correlation.

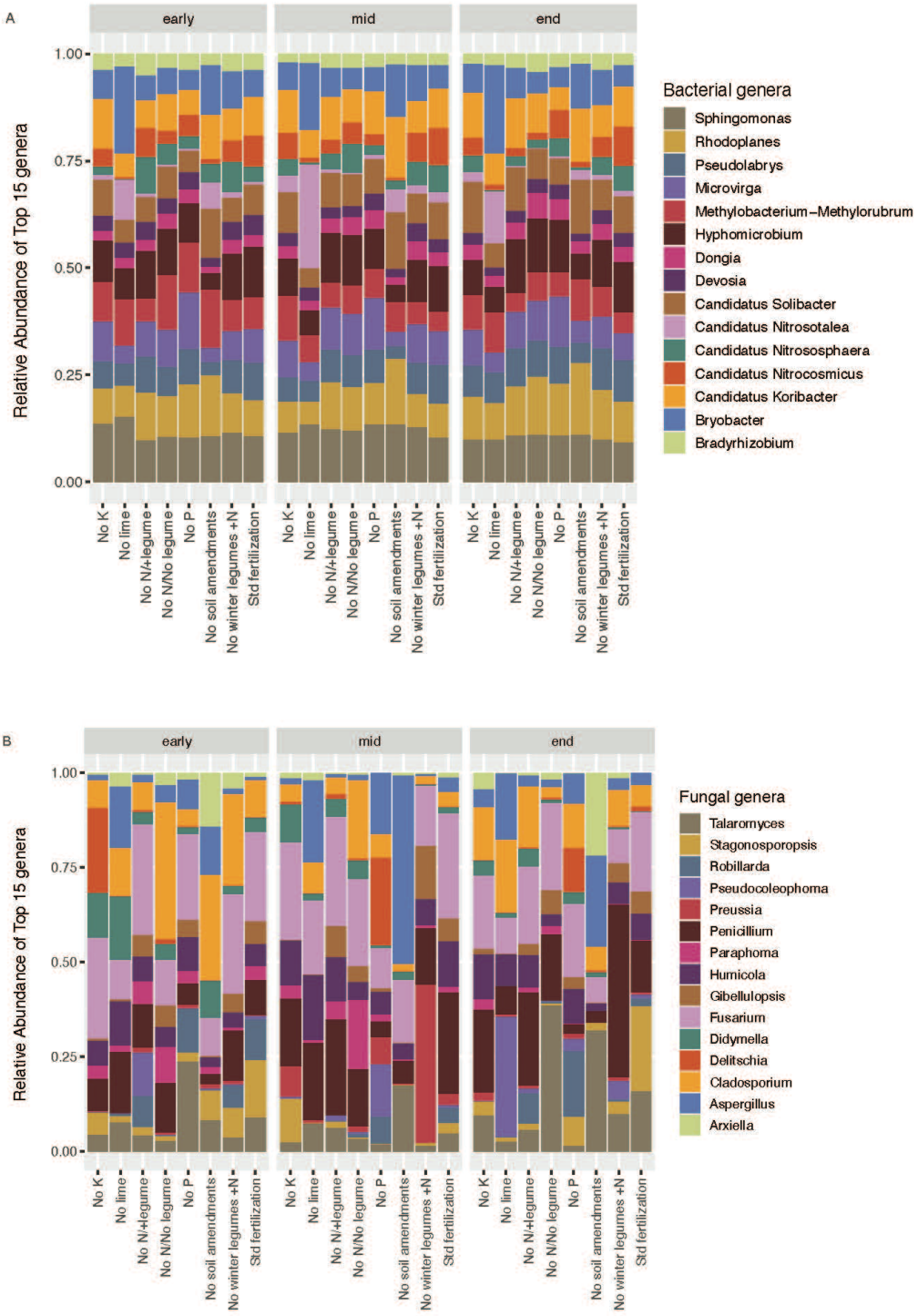

Figure 3-3: Bacterial and fungal communities differ across fertilizer treatment. (A) Relative abundance of the top 15 bacterial genera and (B) top 15 fungal genera across treatment groups.

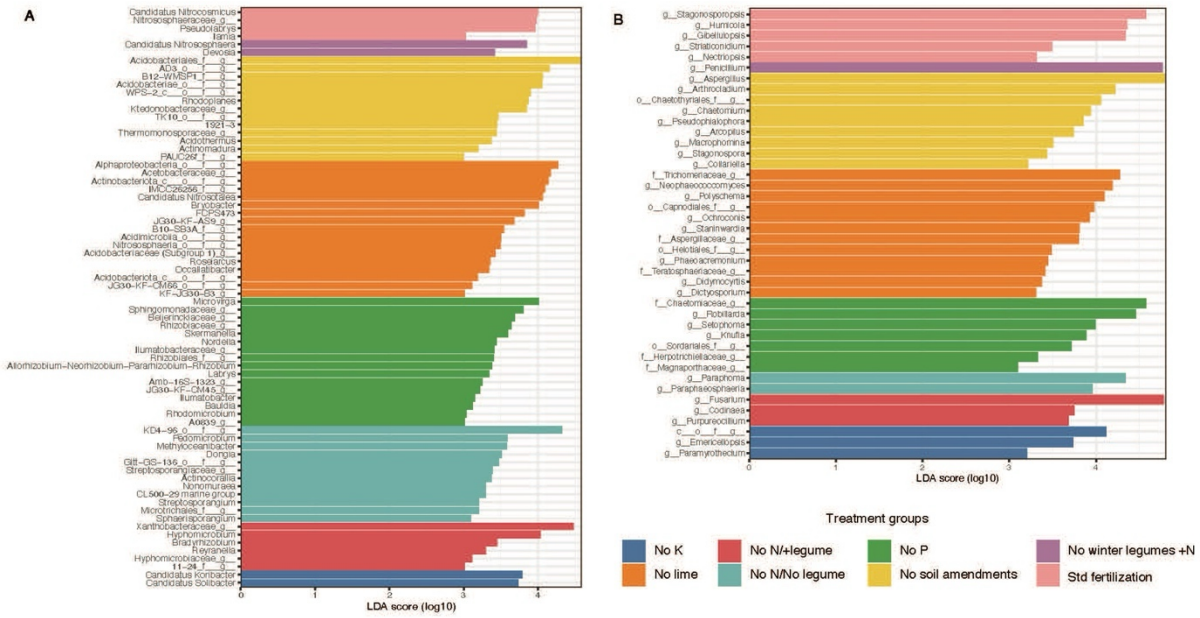

Figure 3-4: Long term fertilization results in enrichment of various bacterial and fungal taxa across treatments. Differential abundance analysis and identification of microbial markers as predictive signatures through linear discriminant analysis (LDA) in (A) bacterial genera (B) fungal genera. based on effect size measurements (LEfSe) analysis. A taxon is considered as significantly different according to a LDA score of ≥ 3.

A

B

Treatment Key	
A*	No N/ +legume
B*	No N/ -legume
C*	No amendments
1*	N/ -legume
2*	No P
3*	Complete
4	4/3 K
5	Rock Phosphate
6*	No K
7	2/3 K
8*	No lime
9	No sulfur
10	Complete + micronutrients
11	1/3 K

*Treatments used in this study.

C

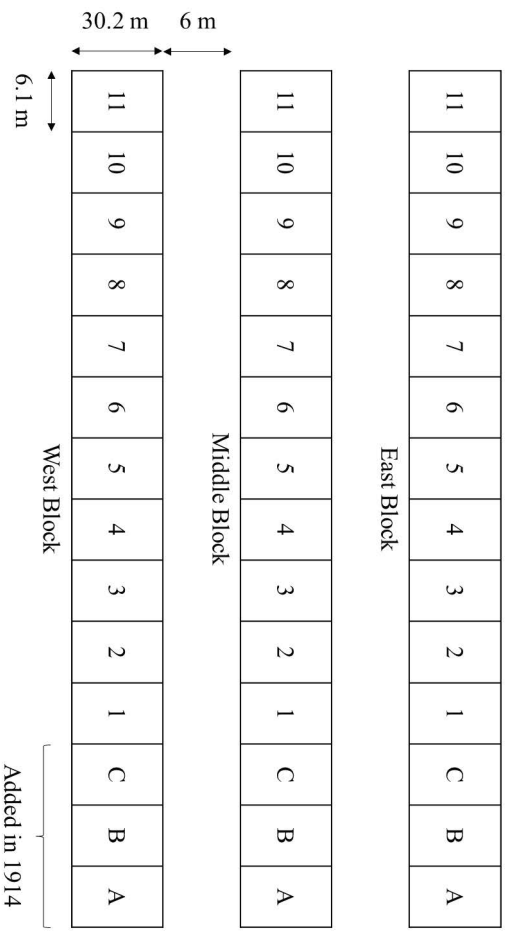

Figure S 3-1: Study site and treatment design of Cullars Rotation site at Auburn University. (A) Satellite image of Cullars rotation site with individual plots and blocks. (B) Treatment design of each plot in Cullars rotation and treatment plots used in this study. (C) Treatment blocks showing distribution of treatment plan.

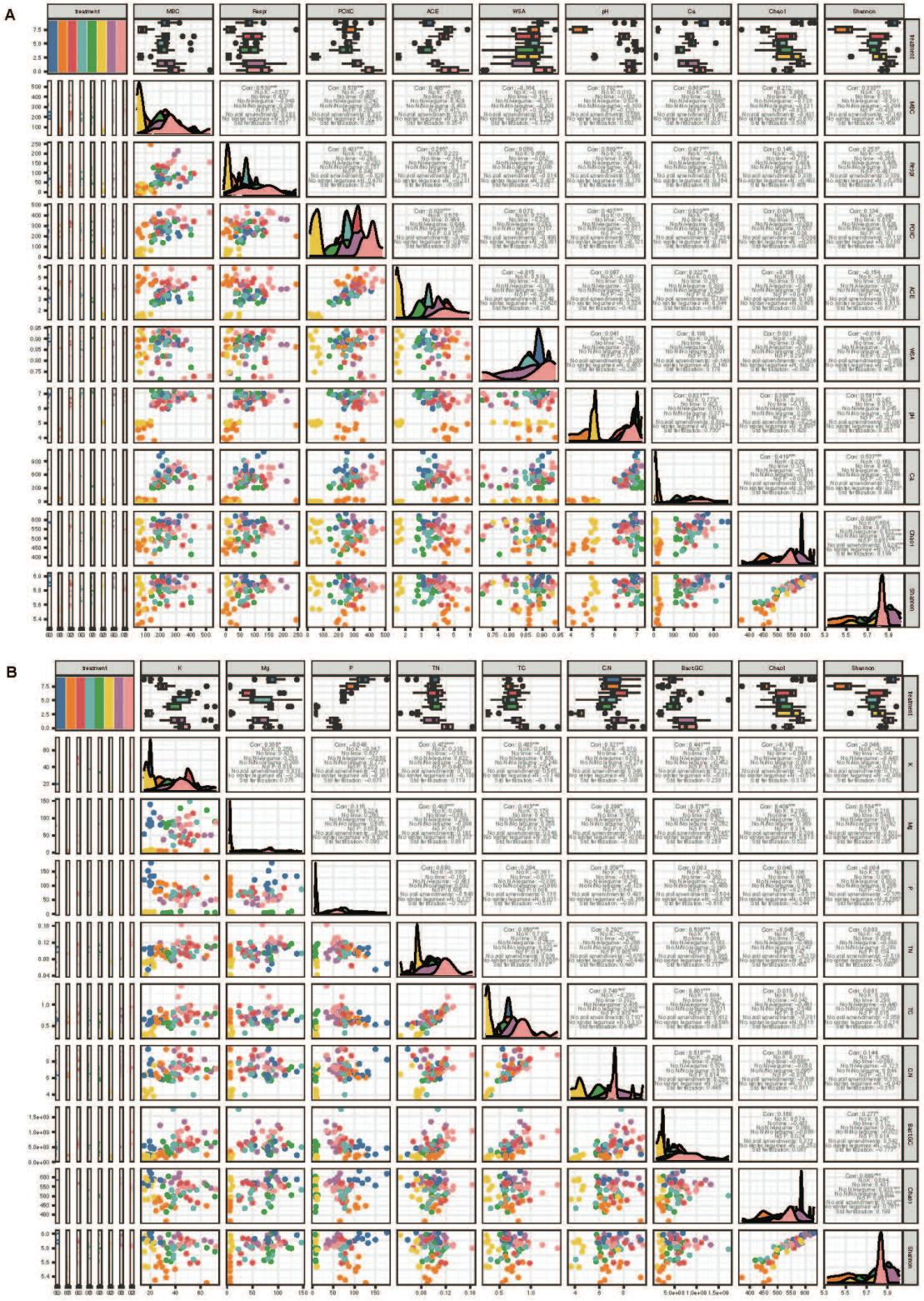

Figure S 3-2: Pearson correlation analysis to identify the relationship between soil properties in different fertility treatments and bacterial communities.

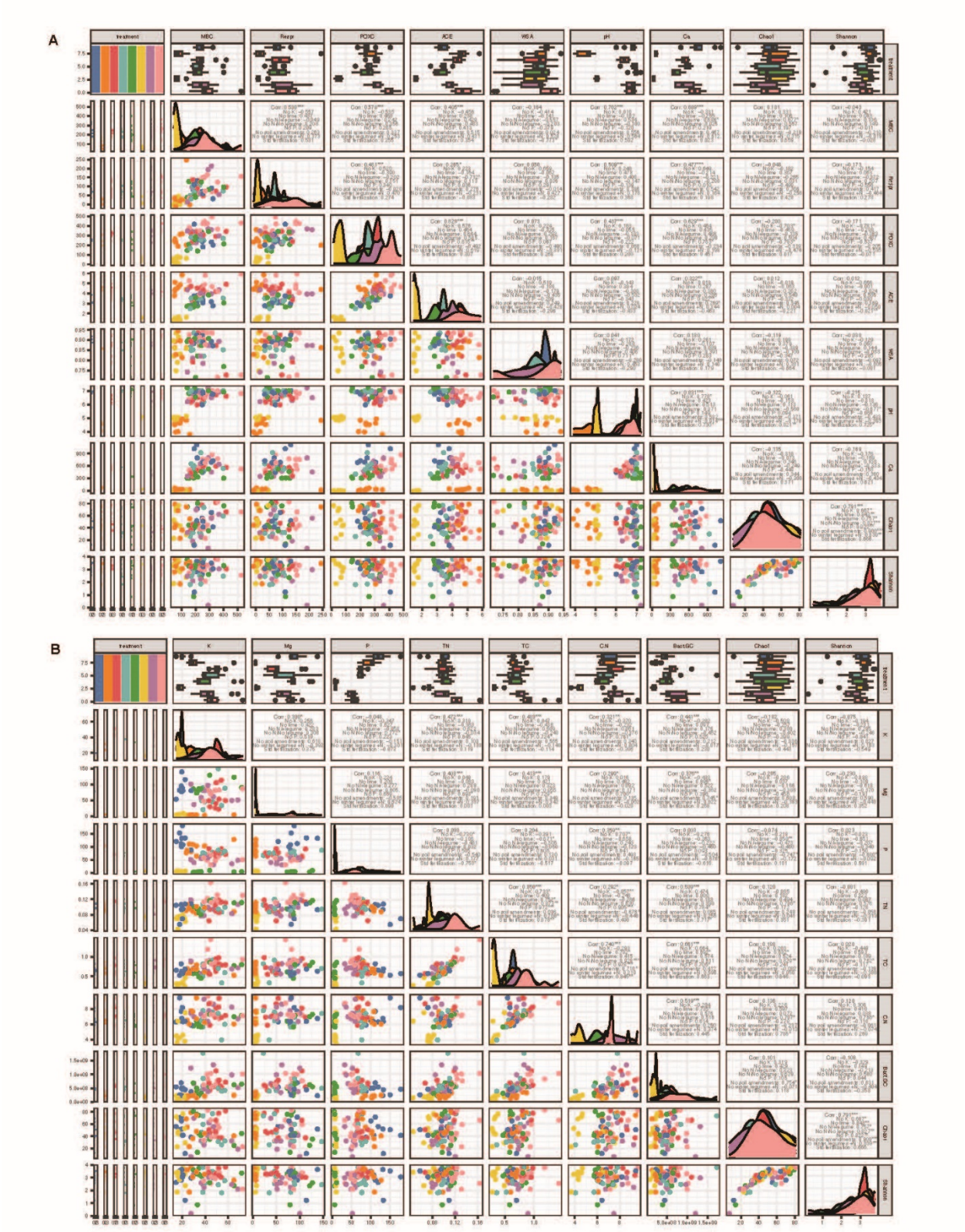

Figure S 3-3: Pearson correlation analysis to identify the relationship between soil properties in different fertility treatments and fungal communities.

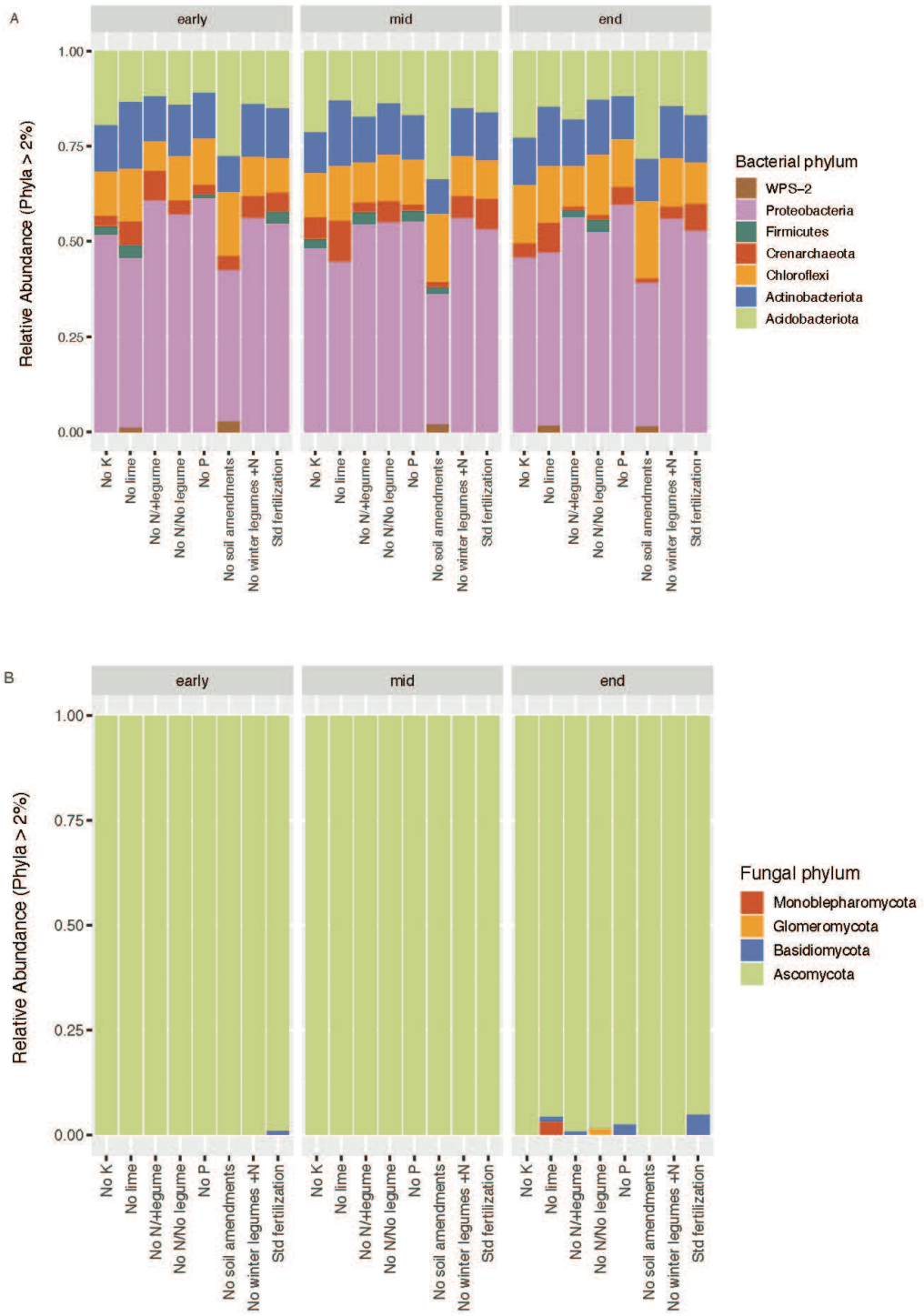

Figure S 3-4: Distribution of soil (A) bacterial and (B) fungal microbial communities at the phylum level.

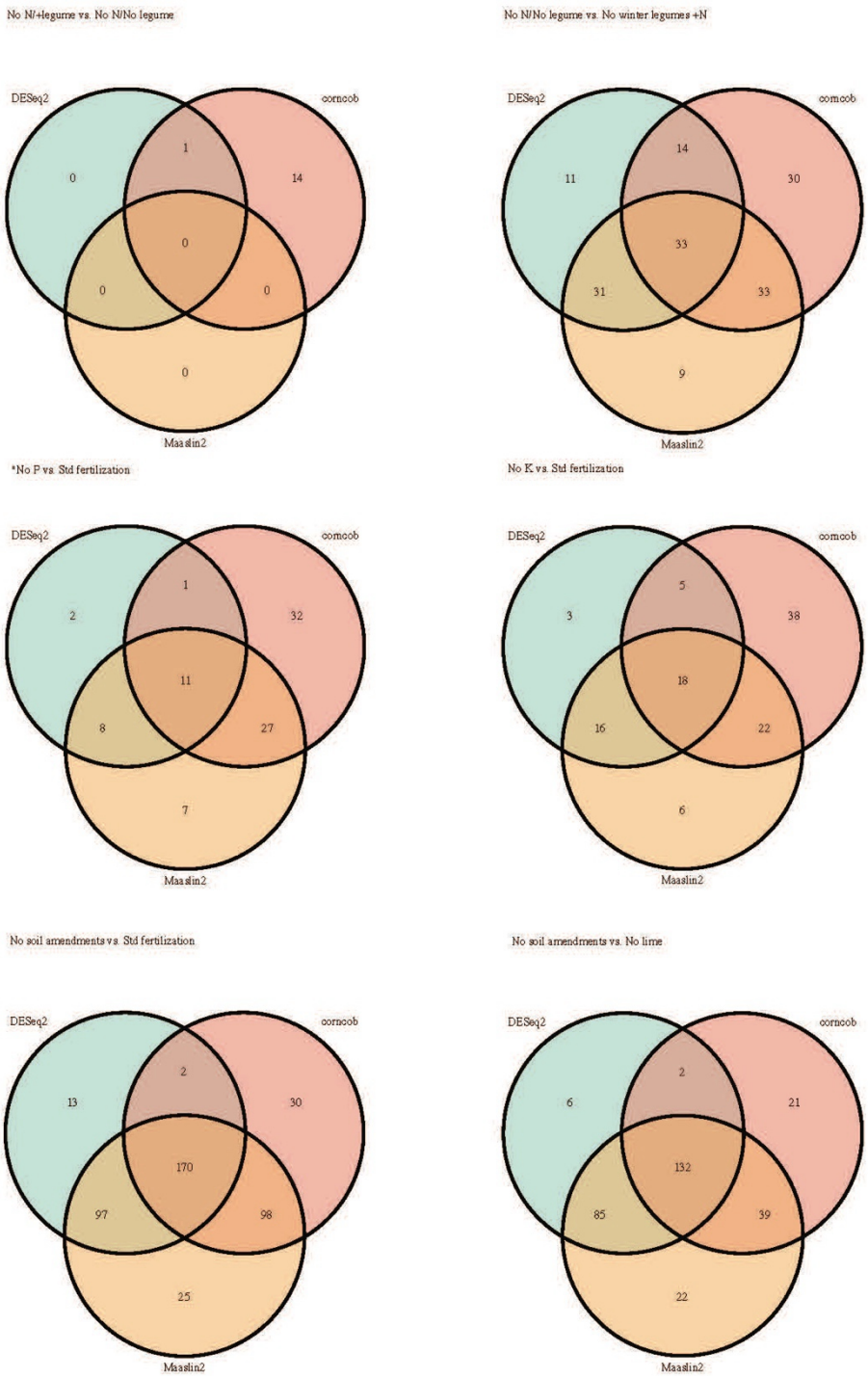

Figure S 3-5: Total number of differentially abundant bacterial taxa identified using DESeq2, Corncob, and Maaslin2.

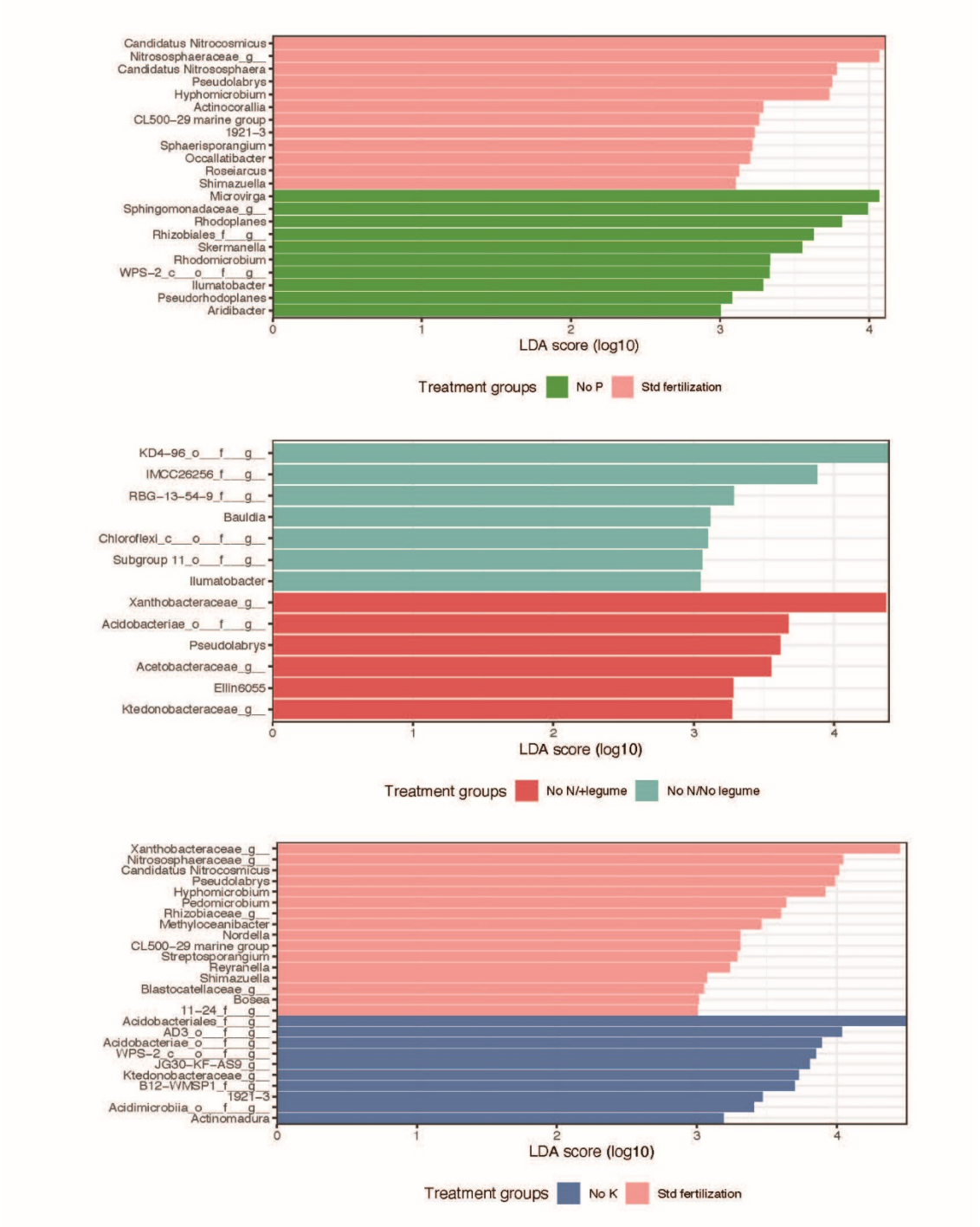

Figure S 3-6: Microbial biomarkers as predictive signatures through linear discriminant analysis (LDA) across (A) No P vs Std fertilization, (B) No N/+legume vs No N/No Legume (C) No K vs Std fertilization. A taxon is considered as significantly different according to a LDA score of ≥ 3.

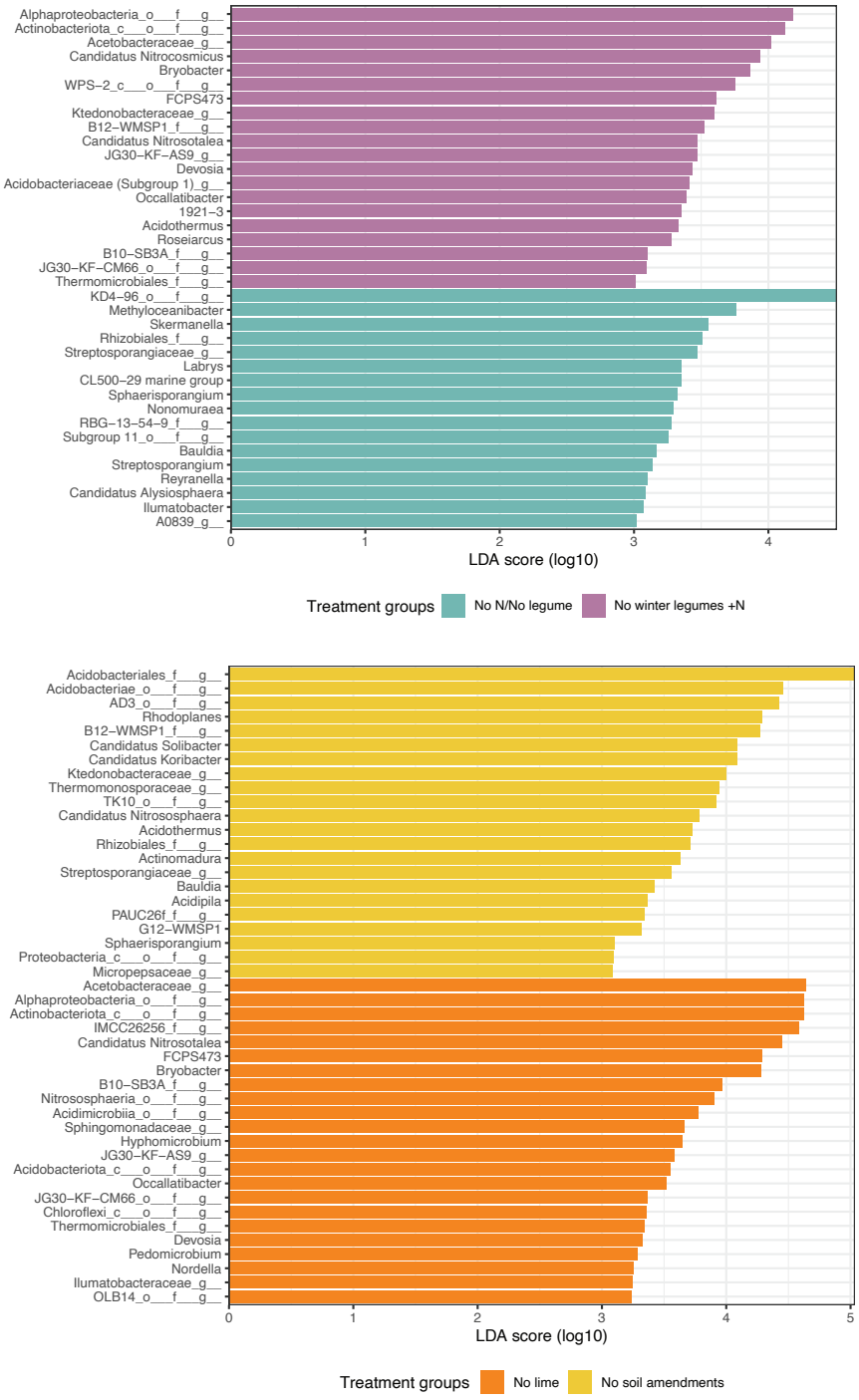

Figure S 3-7: Microbial biomarkers as predictive signatures through linear discriminant analysis (LDA) across (D) No N/No Legume vs No winter legumes + N and (E) No Lime vs No soil amendments treatment. A taxon is considered as significantly different according to a LDA score of ≥ 3.

Table S 3-1A: Influence of various factors on bacterial Chao1 richness

	Df	Sum sq	Mean Sq	F value	Pr(>F)
treatment	7	130582	18655	13.333	2.09E-09
time	2	2102	1051	0.751	0.477
treatment:time	14	18094	1292	0.924	0.541
Residuals	48	67157	1399		
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1					

Table S 3-1B: Bacterial Chao1 richness between treatments

Treatments	diff	lwr	upr	Padj	
No lime-No K	-142.48228	-198.34769	-86.616864	0	***
No N/+legume-No K	-56.911267	-112.77668	-1.045852	0.0430404	*
No N/No legume-No K	-89.758033	-145.62345	-33.892617	0.0001513	***
No P-No K	-79.746849	-135.61227	-23.881434	0.0009785	**
No soil amendments-No K	-44.917416	-100.78283	10.947999	0.2010368	
No winter legumes +N-No K	-9.421554	-65.28697	46.443861	0.9993972	
Std fertilization-No K	-53.9924	-109.85782	1.873015	0.0649577	
No N/+legume-No lime	85.571012	29.705596	141.436427	0.0003337	**
No N/No legume-No lime	52.724246	-3.141169	108.589662	0.0771555	
No P-No lime	62.73543	6.870014	118.600845	0.0178624	**
No soil amendments-No lime	97.564863	41.699448	153.430278	0.0000336	***
No winter legumes +N-No lime	133.060725	77.19531	188.92614	0	***
Std fertilization-No lime	88.489879	32.624464	144.355294	0.0001925	**
No N/No legume-No N/+legume	-32.846765	-88.712181	23.01865	0.5820193	
No P-No N/+legume	-22.835582	-78.700997	33.029833	0.8962496	
No soil amendments-No N/+legume	11.993851	-43.871564	67.859266	0.9971667	
No winter legumes +N-No N/+legume	47.489713	-8.375702	103.355128	0.1496854	
Std fertilization-No N/+legume	2.918867	-52.946548	58.784282	0.9999998	
No P-No N/No legume	10.011183	-45.854232	65.876599	0.9991051	
No soil amendments-No N/No legume	44.840616	-11.024799	100.706032	0.2027471	
No winter legumes +N-No N/No legume	80.336478	24.471063	136.201894	0.0008788	**
Std fertilization-No N/No legume	35.765633	-20.099783	91.631048	0.474426	
No soil amendments-No P	34.829433	-21.035982	90.694848	0.5085303	
No winter legumes +N-No P	70.325295	14.45988	126.19071	0.0051518	**
Std fertilization-No P	25.754449	-30.110966	81.619864	0.8238365	
No winter legumes +N-No soil amendments	35.495862	-20.369553	91.361277	0.4841909	
Std fertilization-No soil amendments	-9.074984	-64.940399	46.790431	0.9995285	
Std fertilization-No winter legumes +N	-44.570846	-100.43626	11.294569	0.2088379	

Table S 3-1C: Bacterial Shannon diversity index based on Kruskall walis rank sum test

	Kruskal chi square	df	p value	Remarks
Treatment	40.23	7	**1.14E-06**	***
Time	0.80613	2	0.6352	
Treatment:time	46.432	23	0.002642	**

Table S 3-1D: Influence of various treatments in bacterial communities based on pairwise wilcox test for treatment.

	No K	No lime	No N/+legume	No N/No leg	No P	No soil amer	No winter legumes +N
No lime	0.00038	-	-	-	-	-	-
No N/+legume	0.03501	0.00154	-	-	-	-	-
No N/No legume	0.00058	0.00066	0.51534	-	-	-	-
No P	0.00931	0.00066	0.73629	0.75749	-	-	-
No soil amendments	0.02838	0.00471	0.75749	0.62291	0.79617	-	-
No winter legumes +N	0.51534	0.00038	0.11348	0.00101	0.01675	0.08286	-
Std fertilization	0.00066	0.00038	0.75749	0.05507	0.09726	0.75749	0.00346

Table S 3-1E: Influence of various factors in fungal chao1 richness

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
treatment	7	1429	204	2.561	0.0252
time	2	16835	8417	105.635	<2e-16
treatment:time	14	1914	137	1.716	0.0835
Residuals	48	3825	80		

Table S 3-1F: Pairwise comparison of Fungal Chao1 richness between treatments

Treatments comparison	diff	lwr	upr	P adj	Remarks
No lime-No K	4.3333333	-8.998922	17.66558829	0.9674189	
No N/+legume-No K	6.7777778	-6.554477	20.11003274	0.7416481	
No N/No legume-No K	-2.1388889	-15.471144	11.19336607	0.9995656	
No P-No K	0.8333333	-12.498922	14.16558829	0.9999993	
No soil amendments-No K	3.1666667	-10.165588	16.49892163	0.9947048	
No winter legumes +N-No K	-8.9722222	-22.304477	4.36003274	0.4099675	
Std fertilization-No K	1.3333333	-11.998922	14.66558829	0.9999819	
No N/+legume-No lime	2.4444444	-10.887811	15.77669941	0.9989624	
No N/No legume-No lime	-6.4722222	-19.804477	6.86003274	0.7831183	
No P-No lime	-3.5	-16.832255	9.83225496	0.9903268	
No soil amendments-No lime	-1.1666667	-14.498922	12.16558829	0.9999928	
No winter legumes +N-No lime	-13.3055556	-26.637811	0.02669941	0.0508002	
Std fertilization-No lime	-3	-16.332255	10.33225496	0.9962034	
No N/No legume-No N/+legume	-8.9166667	-22.248922	4.41558829	0.417966	
No P-No N/+legume	-5.9444444	-19.276699	7.38781052	0.8469738	
No soil amendments-No N/+legume	-3.6111111	-16.943366	9.72114385	0.9883701	
No winter legumes +N-No N/+legume	-15.75	-29.082255	-2.41774504	0.0105977	*
Std fertilization-No N/+legume	-5.4444444	-18.776699	7.88781052	0.8967141	
No P-No N/No legume	2.9722222	-10.360033	16.30447718	0.9964162	
No soil amendments-No N/No legume	5.3055556	-8.026699	18.63781052	0.9085309	
No winter legumes +N-No N/No legume	-6.8333333	-20.165588	6.49892163	0.7338119	
Std fertilization-No N/No legume	3.4722222	-9.860033	16.80447718	0.9907736	
No soil amendments-No P	2.3333333	-10.998922	15.66558829	0.9992324	
No winter legumes +N-No P	-9.8055556	-23.137811	3.52669941	0.2991374	
Std fertilization-No P	0.5	-12.832255	13.83225496	1	
No winter legumes +N-No soil amendments	-12.1388889	-25.471144	1.19336607	0.0985215	
Std fertilization-No soil amendments	-1.8333333	-15.165588	11.49892163	0.9998439	
Std fertilization-No winter legumes +N	10.3055556	-3.026699	23.63781052	0.2422346	

Table S 3-1G: Fungal Chao1 richness between sampling time

Time comparison	diff	lwr	upr	P adj	Remarks
end-early	-33.083333	-39.315497	-26.851169	0	***
mid-early	-31.75	-37.982164	-25.517836	0	***
mid-end	1.333333	-4.898831	7.565497	0.8632367	

Table S 3-1H: Fungal Shannon diversity index based on Kruskall walis rank sum test

Factors	Kruskal chi square	df	p value	Remarks
Treatment	11.118	7	**1.34E-01**	
Time	15.485	2	0.0004341	***
Treatment:time	37.524	23	0.02863	*

Table S 3-1I: Pairwise comparison of fungal communities and time of sampling

	early	end
end	0.00062	-
mid	0.00191	0.631

Table S 3-2A: PERMANOVA Bacterial beta diversity

Factors	Df	SumOfSqs	R2	F	Pr(>F)	Remarks
treatment	7	5.3873	0.4843	8.5487	0.000999	***
time	2	0.4527	0.04069	2.5141	0.000999	***
treatment:time	14	0.9626	0.08653	0.7637	0.991009	
Residual	48	4.3213	0.38847			
Total	71	11.1239	1			

Table S 3-2B: Permutation test for homogeneity of multivariate dispersions to determine if the variance differ by treatment group.

	No K	No lime	No N/+legume	No N/No legume	No P	No soil amendments	No winter legumes +	Std fertilization
No K		0.521479	0.113886	0.324675	0.582418	0.131868	0.062937	0.1389
No lime	0.539273		0.452547	0.719281	0.98002	0.405594	0.383616	0.5195
No N/+legume	0.112973	0.435823		0.762238	0.421578	0.777223	0.935065	0.9131
No N/No legume	0.317624	0.70011	0.761374		0.678322	0.638362	0.703297	0.8262
No P	0.571065	0.977793	0.432661	0.686924		0.380619	0.341658	0.5025
No soil amendments	0.140549	0.388989	0.782238	0.632537	0.385405		0.829171	0.7383
No winter legumes +N	0.06716	0.362944	0.934771	0.699387	0.364098	0.806086		0.8462
Std fertilization	0.148754	0.499965	0.915765	0.829727	0.493982	0.728605	0.842355	

Table S 3-2C: Pairwise adonis for pairwise comparison against various treatment and time

No N/+legume vs standard fertilization

Factors	Df	SumOfSqs	R2	F	Pr(>F)	Remarks
treatment	1	2027405	0.13017	2.1241	0.012	*
time	2	1017417	0.06532	0.533	0.189	
treatment:time	2	1076698	0.06913	0.564	0.946	
Residual	12	11453734	0.73538			
Total	17	15575254	1			

Factors	Df	SumOfSqs	R2	F	Pr(>F)	Remarks
treatment	1	1942881	0.11652	1.8681	0.028	*
time	2	1208597	0.07248	0.581	0.653	
treatment:time	2	1042355	0.06251	0.5011	0.984	
Residual	12	12480140	0.74848			
Total	17	16673973	1			

No P vs No lime

Factors	Df	SumOfSqs	R2	F	Pr(>F)	Remarks
treatment	1	1918752	0.1093	1.7407	0.029	*
time	2	1268892	0.07228	0.5756	0.728	
treatment:time	2	1139900	0.06493	0.5171	0.993	
Residual	12	13227726	0.75349			
Total	17	17555269	1			

Table S 3-2D: PERMANOVA of fungal diversity

Factors	Df	SumOfSqs	R2	F	Pr(>F)	Remarks
treatment	7	3.429	0.09679	0.9798	1	
time	2	1.141	0.03221	1.1412	0.001998	**
treatment:time	14	6.859	0.19359	0.9798	0.964036	
Residual	48	24	0.6774			
Total	71	35.429	1			

Table S 3-2E: Pairwise permanova between timepoints for fungal communities

Early vs mid-season

Factors	Df	SumOfSqs	R2	F	Pr(>F)	Remarks
time	1	25711087	0.02241	1.0543	0.001	***
Residual	46	1121792244	0.97759			
Total	47	1147503330	1			

Early vs end season

Factors	Df	SumOfSqs	R2	F	Pr(>F)	Remarks
time	1	21262479	0.02266	1.0664	0.001	***
Residual	46	917156278	0.97734			
Total	47	938418756	1			

Mid vs end-season

Factors	Df	SumOfSqs	R2	F	Pr(>F)
time	1	29605501	0.02128	1	1
Residual	46	1361853033	0.97872		
Total	47	1391458533	1		

4. CHAPTER FOUR: Genetic and functional diversity help explain pathogenic, weakly pathogenic, and commensal lifestyles in the genus *Xanthomonas*

Note – This study is available as a preprint (https://doi.org/10.1101/2023.05.31.543148) with the following citation and is currently under review for Genome Biology and Evolution.

Pena, M. M.*, **Bhandari, R**.*, Bowers, R. M., Weis, K., Newberry, E., Wagner, N., et al. 2023. Genetic and functional diversity help explain pathogenic, weakly pathogenic, and commensal lifestyles in the genus *Xanthomonas*. *bioRxiv*:2023.05.31.543148

***Michelle M. Pena and Rishi Bhandari contributed equally to this manuscript.**

This work performed in collaboration with other authors. My part of the collaboration was to process the samples and run various analysis. I wrote the final manuscript and formatted the figures and tables.

Abstract

The primary focus of research on the genus *Xanthomonas* has been its pathogenic interactions with plants, with distinct pathovars demonstrating a high level of specificity towards particular hosts and tissues. However, this genus also comprises nonpathogenic strains from diverse hosts and environments, including rainwater, which are frequently isolated with the pathogenic strains. Based on their capacity to cause no or minor symptoms on the host of isolation under favorable conditions, nonpathogenic xanthomonads can be further characterized as commensal and weakly pathogenic. The diversity and evolution of these diverse lifestyles within this genus are poorly characterized. This study aimed to understand the diversity and

evolution of nonpathogenic xanthomonads compared to their pathogenic counterparts based on their co-occurrence or phylogenetic relationship and to identify genomic traits that form the basis of a life-history framework that groups xanthomonads by ecological strategies. We sequenced genomes of 83 strains spanning the genus phylogeny and identified eight novel species, indicating unexplored diversity. While some nonpathogenic species have experienced a recent loss of type III secretion system, specifically, *hrp2* cluster, we observed an apparent lack of association of *hrp2* cluster with lifestyles of diverse species. We gathered evidence for gene flow among co-occurring pathogenic and nonpathogenic strains, suggesting the potential of nonpathogenic strains to act as a reservoir of adaptive traits for pathogenic strains. We further identified traits enriched in nonpathogens that suggest their strategies of tolerance rather than avoidance of stressors they may experience during their association with a broad range of host plants.

4.1 Introduction

The genus *Xanthomonas*, traditionally considered to group plant pathogenic bacteria, encompasses bacterial strains that although they maintain close association with plants, do not cause apparent disease symptoms in their host of isolation (Bansal et al., 2021; Essakhi et al., 2015; Garita-Cambronero et al., 2017; Martins et al., 2020; Merda et al., 2016, 2017; Vauterin et al., 1996). Nonpathogenic xanthomonads have a varied lifestyle with the ability to colonize the plant hosts and survive in various environments outside the plants, such as rain and aerosols (Mechan Llontop et al., 2021; Vauterin et al., 1996). Although referred to as nonpathogenic in the context of their phenotype based on artificial inoculation on the host of isolation, it cannot be ruled out that these strains may cause disease in other hosts. Some of these nonpathogenic *Xanthomonas* strains have been isolated together with pathogenic relatives from a diversity of host plants, at

times, from the same lesion in infected plants or asymptomatic hosts or seed-lots or transplants (Gitaitis, 1987; Vauterin et al., 1996). Some of these nonpathogenic xanthomonads are opportunistic pathogens under favorable conditions (Vauterin et al., 1996) and have amylolytic and/or pectolytic activity, which allows them to cause soft rot on their host (Gitaitis, 1987; Zarei et al., 2022). Vauterin et al. (1996) systematically characterized seventy diverse nonpathogenic xanthomonads based on fatty acid methyl ester (FAME) and sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) protein patterns. This pioneering study indicated potential new species and the need to address the diversity and relatedness of nonpathogenic strains to the pathogenic strains of xanthomonads from an ecological viewpoint and the practical aspects of disease diagnostics and management strategies. In the last two decades, several studies have addressed this question of diversity, focusing on individual species, using more advanced methods of multi-locus sequence typing and genome sequencing (Bansal et al., 2020, 2021; Cesbron et al., 2015; Essakhi et al., 2015; Gonzalez et al., 2002; T. Li et al., 2020; Triplett et al., 2015). Whole genome-based phylogeny placed the crop-associated nonpathogenic xanthomonads in the species *arboricola*, *cannabis*, belonging to Group 2 (Cesbron et al., 2015; Jacobs et al., 2015; Merda et al., 2016, 2017) and some in the newly described species, such as *X. sontii* (Bansal et al., 2021), belonging to the early branching clade, Group 1. Apart from *X. arboricola* and *X. campestris*, which house both pathogenic and nonpathogenic strains simultaneously isolated from symptomatic hosts (Lee et al., 2020; Martins et al., 2020), other nonpathogenic strains are only distantly related to the co-colonizing pathogenic strains (Vauterin et al., 1996).

These nonpathogenic strains are diverse in their phylogenetic placements and vary in their makeup of type III secretion systems (T3SS) and associated effectors. The T3SS, encoded by the *hrp2* cluster and type III effectors and/or their repertoires, are important determinants of

pathogenicity in xanthomonads, and individual effectors or their repertoires have been hypothesized to contribute towards host specificity (Hajri et al., 2009; Jacques et al., 2016; White et al., 2009). Pathogenic xanthomonads belonging to *X. maliensis*, *X. cannabis*, *X. pseudoalbilineans,* and *X. sacchari* are an exception in that they lack a *hrp2* T3SS, but some possess regulators of T3SS (Triplett et al. 2015; Jacobs et al. 2015, Studholme et al. 2011, Pieretti et al. 2015). Knowing the importance of T3SS and associated effectors in the pathogenicity of xanthomonads, it is unsurprising that most nonpathogenic xanthomonads lack T3SS. Such nonpathogenic xanthomonads lacking T3SS, specifically *X. arboricola*, have been previously referred to as commensal xanthomonads. However, Merda et al. (2017) further showed that some commensal strains of *X. arboricola* possess T3SS but contain only 3-4 effectors. Some commensal strains lack *hrp2* cluster but have up to four effectors (Cesbron et al., 2015).

Interestingly, some *X. arboricola* strains have been reported as weak pathogens based on pathogenicity tests showing faint or mild water-soaking symptoms on the host of isolation (Roach et al., 2018; Sawada et al., 2011). Further genome analysis of weakly pathogenic *X. arboricola* strains revealed a limited set of effectors compared to pathogenic strains, in some cases, up to 16 effectors (Roach et al., 2019), suggesting the possibility of these strains be potentially pathogenic on other hosts. This heterogenic distribution of T3SSs and variable T3E repertoires make nonpathogenic strains a suitable model to study the evolutionary history of the *hrp2* T3SS family, effectors, and associated regulators in xanthomonads (Cesbron et al., 2015; Merda et al., 2017). Merda et al. (2017) inferred ancestral acquisitions of the *hrp2* cluster in *Xanthomonas* and indicated loss or subsequent gain in certain clades in 82 genomes spanning the genus *Xanthomonas*. As we uncover additional diversity in the genus *Xanthomonas*, evolutionary gain and loss of different types of T3SSs, including atypical T3SSs (Pesce et al., 2017; Pieretti et al.,

2015) in xanthomonads along with effectors and regulators can be a valuable approach to understand the role of T3SSs in allowing intimate association of xanthomonads with plants and to further determining their lifestyle.

The presence of nonpathogenic xanthomonads in association with pathogenic xanthomonads from the same infected tissues raises the important question of whether the presence of nonpathogenic xanthomonads influences the population dynamics of pathogenic xanthomonads and vice versa, either through the exchange of genetic material or through sharing of public goods, such as cell-wall degrading enzymes or their products of degradation (Sadhukhan et al., 2023). Profiling of the mobile genetic elements (MGEs) shared between pathogenic and nonpathogenic strains can provide critical information to understand how genes and their encoded functions can be exchanged via horizontal gene transfer. These MGEs can influence selection pressure-driven changes in the population dynamics of pathogens and nonpathogens. Such events linked to pathoadaptation have been proposed (Cesbron et al., 2015; Meline et al., 2019).

Xanthomonas strains were recently found in the endosphere of *Arabidopsis* as a part of the At-LSPHERE collection. Evaluation of these strains' pathogenicity on wild-type *Arabidopsis* and immunocompromised plants confirmed their opportunistic or conditional pathogenic nature based on aggressive symptoms on immunocompromised plants lacking plant NADPH oxidase (RBOHD) (Pfeilmeier et al., 2021). This study further highlighted the role of microbial community members and microbiota-induced plant immunity in reducing the prevalence of opportunistic strains. On the other hand, a closely related endophytic *Xanthomonas* strain, WCS2014-23, was identified as a member of the consortium recruited to the rhizosphere of *Arabidopsis thaliana* upon foliar infection with the biotrophic pathogen and played a role in

induced systemic resistance against the biotrophic pathogen and enhancing plant growth (Berendsen et al., 2018). These findings raise the question of whether nonpathogenic xanthomonads play an important role as resident and functional members of the phyllosphere microbiome or are transient nonfunctional community members. A related question is if the so far uncharacterized xanthomonads that have been isolated from rainwater are *bona fide* phyllosphere microbiome members that are only transiently present in the atmosphere or if they present a separate population of non-plant-associated xanthomonads (Failor et al., 2017).

In this study, we set out to address the above-mentioned knowledge gaps on the overall diversity of nonpathogenic xanthomonads regarding the genetic diversity of strains, associated virulence factors, and mobile genetic elements. We also focused on understanding the evolution of pathogenic and nonpathogenic strains and deciphering genes associated with the adaptation of these diverse strains to different lifestyles in association with plants and the environment. We sequenced a collection of 83 presumptive nonpathogenic *Xanthomonas* strains from diverse hosts, environments, and geographical locations. The phylogenetic placement of these strains spanned the entire genus, including the identification of potentially new species. The heterogeneous distribution of T3SS and associated effectors across pathogenic, weakly pathogenic, and commensal strains suggested a lack of apparent association of these important pathogenicity factors with the lifestyle of *Xanthomonas* strains. Thus, we used an integrated approach of comparative genomics and association analysis to identify the genomic attributes associated with these lifestyles of strains spanning the genus phylogeny.

4.2 Materials and Methods

4.2.1 Bacterial strains collection and genome sequencing

Nonpathogenic *Xanthomonas* strains collected from different plant hosts and environmental samples (Table S4-1) were used for genomic DNA extraction using the CTAB-NaCl method (William et al., 2012). Degradation and contamination of the genomic DNA were monitored on 0.5% agarose gels. DNA concentration was measured using a Qubit® DNA Assay Kit on a Qubit® 2.0 Fluorometer (Life Technologies, CA, USA) and submitted to the Joint Genome Institute (JGI) for library preparation and sequencing. Paired-end reads were generated by multiplexing 12 libraries in a single lane on the Illumina NovaSeq (PE150) platform. Raw reads, annotation data, and final assembly are in the JGI data portal (http://genome.jgi.doe.gov). The information for 83 newly sequenced genomes from this study is in Table S4-1.

4.2.2 Genome-based identification of *Xanthomonas* strains

Comparative genomic analysis was performed among 134 *Xanthomonas* strains, including 83 strains from this study and 51 representative *Xanthomonas* strains from NCBI (Table S4-2). Average nucleotide identity (ANI) was estimated using all-versus-all strategies using FastANI (v1.1) (Jain et al., 2018) and pyani (v0.2.12) (Pritchard et al., 2015). We also used the ANI values from the web server LINbase (Tian et al., 2020) and Microbial Species Identifier (MiSI) (Varghese et al., 2015) as additional tools in species circumscription. The MiSI method addresses inconsistencies based on ANI alone and includes alignment fractions and genome-wide ANI values. Additionally, twenty strains representing the novel species diversity were subject to the Type (Strain) Genome Server (https://tygs.dsmz.de/) to calculate digital DNA-DNA hybridization (dDDH) values based on the Genome BLAST Distance Phylogeny (Meier-Kolthoff et al., 2013).

A combination of ANI, dDDH, and MiSI was used to designate the "new species" status to a given strain only when values were below the accepted threshold (≤95% for ANI and ≤70% for dDDH) (Kim et al., 2014).

The whole proteome of the 134 *Xanthomonas* strains was compared by OrthoFinder (v2.5.2) (Emms & Kelly, 2019) to identify orthogroups using the original algorithm (Emms & Kelly, 2015). The identified orthogroups were used to infer unrooted gene trees using the BLAST-based hierarchical clustering algorithm DendroBLAST (Kelly & Maini, 2013). Using this set of unrooted gene trees, the STAG algorithm identified the closest pair of genes from those species to infer an unrooted species tree (Emms & Kelly, 2018). The unrooted species tree inferred from STAG was then rooted using the STRIDE algorithm (Emms & Kelly, 2017) by identifying well-supported gene duplication events. The resulting cladogram was visualized with R package *ggtree* (G. Yu et al., 2017).

To identify and visualize possible conflicting signals that would suggest recombination events and evolutionary relationships within the *Xanthomonas* sequence data, Multi-locus sequence analysis (MLSA) was carried out for 12 housekeeping genes fragments (*gyr*B, *gap*A, *lac*F, *glt*A, *fsu*A, *lep*A, *atp*D, *rpo*D, *gln*A, *ef*P, *dna*K, and *fyu*A) using autoMLSA2 (v0.7.1) (https://github.com/davised/automlsa2). The resultant splits.nex file was used in SplitsTree4 (v4.17.0) (Huson et al., 2008) to develop a phylogenetic network. The possibility of recombination events was identified by the branches that form parallelograms (Joseph & Forsythe, 2012).

4.2.3 Analysis of the gain and loss dynamics of the T3SS clusters

Protein sequences of T3Es representing all effector families, putative effectors, and their diversity were also identified in genome sequences using tBLASTn searches. The T3SS-coding genes from five different *Xanthomonas* species (*X. campestris* pv. *vesicatoria* 85-10, *X. campestris*

pv. *campestris* ATCC33913, *X. translucens* pv. *translucens* DSM18974, *X. albilineans* CFBP 2523, and *Xanthomonas* sp. 60), were used as query to perform BLASTn searches on *Xanthomonas* strains genomes using autoMLSA2 (v0.7.1) with the cut-offs set to 40% identity and 30% coverage. Heatmaps for the blast searches were generated using the *R* package *Pheatmap* (v1.0.12).

Branch-specific gain and loss probabilities of *Xanthomonas* T3SS genes during the evolution were inferred with the species tree and presence/absence in the 134 genomes using GLOOME (Cohen et al., 2010). GLOOME analyzes presence and absence profiles (phyletic patterns) and accurately infers branch-specific and site-specific gain and loss events. We first inferred the gain and loss dynamics of all genes encoding components of the T3SS. Specifically, we searched the genes of four T3SS clusters: (1) *hrp2* cluster derived from *X. campestris* pv. *campestris* ATCC33913 (*Xcc* cluster, 22 genes); (2) *X. translucens* pv. *translucens* DSM18974 (*Xtra* cluster, 18 genes); (3) *X. albilineans* CFBP 2523 (*Xalb* cluster, 11 genes); (4) *Xanthomonas.* sp. 60 (strain 60 cluster from this study, 11 genes). We additionally searched for the presence and absence of genes encoding transcription factors involved in T3SS-related pathogenicity: HrpG, HrpX, HpaR, and HpaS (four genes). These 66 genes were searched in the 134 genomes using BLASTp. A hit was considered if the identity percentage was at least 50%, the E-value was lower than 10^{-10}, and the coverage was at least 30%. When a hit was detected to two or more genes from different T3SS clusters, the one with the highest bit-score was retained. As the phylogenetic tree for the analysis, we used the species tree generated by OrthoFinder. The final GLOOME analysis was performed with default parameter values. Graphical visualization of the tree was done using FigTree (v1.4.4) (http://tree.bio.ed.ac.uk/software/figtree/).

4.2.4 Prediction of Mobile Genetic Elements

To study the genomic differences driven by mobile genetic elements (MGE) in the genus *Xanthomonas*, we used Mobile Genetic Element Finder (MGEfinder) (v1.0.6) (Durrant et al., 2020). MGEfinder assembles the short reads and aligns them to a reference genome to find insertions (Durrant et al., 2020). *Xanthomonas* reads from this study and representative short reads were downloaded from NCBI's SRA database and trimmed using Trim Galore (v0.6.6) (https://github.com/FelixKrueger/TrimGalore). Reference genomes were indexed, and the cleaned reads were aligned with BWA-MEM (v0.7.17) (H. Li & Durbin, 2009). Target strains were assigned to pathogenic reference genomes according to their phylogenetic placement generating eight clusters (Table S4-2). Each cluster contains a representative/type strain and the strains from this study. The predicted mobile genetic islands were annotated using a consensus from BLASTx results in NCBI, JGI, and UniProt. Islands containing carrier genes of potential interest were further analyzed using JGI BLASTp and gene neighborhood viewer to locate transposases or phage-related genes associated with or flanking the island. Additionally, EasyFig (v2.2.2) (Sullivan et al., 2011) was used to visualize the insertion location of mobile genetic elements within genomes.

4.2.5 Comparison of secreted carbohydrate-active enzymes

We screened the genomes of commensal, weakly pathogenic, and pathogenic *Xanthomonas* strains for the presence of various genes involved in breakdowns (CEs, PLs, GHs) and assembly (GTs) of carbohydrates, lignin degradation (AAs), and the carbohydrate-binding module (CBM) (Kaoutari et al., 2013; Lairson et al., 2008). Carbohydrate active enzymes (CAZymes) were assigned to Prokka (v1.14.5) (Seemann, 2014) protein output files (.faa files)

using run_dbcan command (https://github.com/linnabrown/run_dbcan) against the HMMER, DIAMOND, and eCAMI databases with default settings. Final CAZyme domain annotations were the best hits based on the outputs of at least two databases to investigate the genomic potential of various species for carbohydrate utilization. To assess the impacts of different lifestyles on the secreted CAZyme count while considering phylogenetic signals, pairwise phylogenetic distances were created between the genomes using the function tree.distance() from the package biopython phylo (Cock et al., 2009), which was then used to build a principal component analysis (PCA). The CAZyme count for each genome from the run_dbcan step was then converted into a distance matrix with the function vegdist (method="jaccard") from the *Vegan* (v2.6-4) R package (Dixon, 2003) (Oksanen et al. 2009). Permutational multivariate analysis of variance (PERMANOVA) was performed to determine the effect of phylogeny and microbial lifestyle on the distribution of CAZymes with the function adonis2 from the *Vegan* R package. Pairwise comparisons between the lifestyles were carried out using the function *pairwise.perm.manova* (from R package *RVaidememoire*) to understand the difference between lifestyles in terms of their genome content as described in Miyauchi et al. (2020).

4.2.6 Pathogenicity assays and co-inoculation experiments

A subset of *Xanthomonas* strains isolated from tomato were further tested for their pathogenicity by dip-inoculating tomato plants (susceptible cultivar FL 47R). 4-5 weeks old tomato plants were dip-inoculated with an inoculum of overnight grown *Xanthomonas* strains adjusted to 10^6 CFU/ml in $MgSO_4$ buffer amended with 0.0025% (vol/vol) Silwet L-77 (PhytoTechnology Laboratories, Shawnee Mission, KS, USA) and maintained under greenhouse conditions. The symptoms on leaves and *in-planta* bacterial population were recorded after ten days. The *in-planta* bacterial population was estimated by sampling $2cm^2$ leaf tissue using a cork-

borer, homogenizing the tissue in 1ml of $MgSO_4$ buffer, and plating on Nutrient Agar using a spiral plater (Neu-tecGroup Inc., NY). Plates were incubated at 28°C for three days, and the population for each *Xanthomonas* strain was determined as colony-forming units per centimeter squared of leaf area.

A coinfection experiment was conducted to gain experimental evidence for the exchange of genetic material among commensal and pathogenic *Xanthomonas* under host selection pressure. Pathogenic *Xanthomonas*, *X. euvesicatoria* 85-10 carrying the intact *avrBs1* gene and the commensal *Xanthomonas* strain T55 were co-inoculated onto pepper cv. Early Cal Wonder (ECW, susceptible cultivar) and ECW-10R (carrying the *Bs1* resistance gene, resistant cultivar) using the dip-inoculation method described above at the inoculum concentration of 10^7 CFU/ml. The plants were maintained under greenhouse conditions for two weeks, and the leaf tissue containing lesions was sampled from resistant and susceptible cultivars. Inactivation of the *avrBs1* gene by transposons was tested by PCR using primers flanking transposons on transconjugants obtained from these samples.

4.2.7 Identification of lifestyle-associated genes

We retrieved 1,834 *Xanthomonas* genomes from the NCBI GenBank RefSeq database to ensure a high-quality and minimally biased set of genomes. These genomes were then de-replicated using dRep (v3.2.2) (Olm et al., 2017) to de-replicate the complete dataset using a 95% minimum genome completeness cut-off and 5% maximum contamination. The ANI threshold to form primary clusters (-pa) was set at 0.95 (species level) and 0.99 (strain level) for the secondary cluster. During secondary comparisons, a minimum level of overlap between genomes was set to 80% coverage. The de-replicated genomes were manually curated to include the 83 new genomes generated in this work in addition to representative genomes from diverse *Xanthomonas* species

belonging to groups 1 and 2. A total of 337 genomes were selected for downstream analysis (Table S4-3). OrthoFinder was used as a clustering approach to compare the whole proteome of the 337 *Xanthomonas* strains with the default settings as previously described. To determine significantly enriched or depleted protein clusters in different *Xanthomonas* lifestyles, we used the hypergeometric test, PhyloGLM, and Scoary as described in Levy et al. (2018). Among the three methods, the hypergeometric test looks for the overall enrichment of genes without considering the dataset's phylogenetic structure. PhyloGLM is a phylogenetic-aware method that eliminates enrichments related to shared ancestry (Ives & Garland, 2010), while Scoary combines the phylogeny-aware test, Fisher's exact test, and empirical label-switching permutation analysis (Brynildsrud et al., 2016). All these approaches were used on gene presence/absence and gene copy-number data and used for PhyloGLM test. A gene was considered significant a) if it had a q-value < 0.05 for Fisher's exact test and an empirical *p*-value < 0.05 for Scoary; b) if it had a corrected *p*-value with FDR with q < 0.01 for hypergeometric test; and c) a *p*-value < 0.01 along with an estimate of < -1.5 or > 1.5 in copy number analysis for PhyloGLM. We used eggNOG-mapper (v2.1.7) (Cantalapiedra et al., 2021) to address COG categories to each significantly enriched and depleted protein cluster from the combination of two or more methods (hypergeometric, PhyloGLM, and Scoary) in the different *Xanthomonas* lifestyles. In addition, each ortholog id was queried across the IMG database to obtain annotation based on COG, KO, TIGRFAM, and Pfam. Heatmaps were generated using the R package ggplot2.

4.3 Results

4.3.1 General features of *Xanthomonas* strains from this study, potential new species, and core genome phylogeny

Presumptive nonpathogenic *Xanthomonas* strains sequenced in this study were isolated from various symptomatic and asymptomatic crops, including tomato (*Solanum lycopersicum*) (14 strains), pepper (*Capsicum annuum*) (8 strains), and common bean (*Phaseolus vulgaris*) (27 strains), along with other plant species such as radish (*Raphanus sativus*), walnut (*Juglans*), orange (*Citrus sinensis)*, and sunflower (*Helianthus*). In addition to the strains isolated from plant hosts, 18 strains were recovered from rainwater (Table S4-1). The genome sizes among the sequenced strains varied from 3.6 Mb for strain 60 to 5.3 Mb for strain F5. The percent GC content ranged from 64.60% for strain 3075 to 69.31% for strain F10. Furthermore, the number of coding sequences (CDS) varied from 3,223 in strain 60 to 4,535 in strain F5. There was no apparent correlation between genome size, CDS, %GC and status as pathogenic and nonpathogenic (Figure S1). While a median genome size of 4.87 Mb and a median number of genes of 4208 are found in the genus *Xanthomonas*, one strain (60) showed a reduced genome size of 3.6 Mb and a reduced number of 3223 coding genes (Table S4-1).

To determine the taxonomic placement of the 83 newly sequenced strains, ANI, dDDH, and MiSI were used. ANI values of the strains varied from 79% to 100% compared to the representative *Xanthomonas* strains. Nonpathogenic strains sequenced in this study belonged to both *Xanthomonas* groups. For group 1, nine strains belonged to *X. euroxanthea*. For group 2, 34 strains belonged to *X. arboricola*, 16 to *X. cannabis*, 9 to *X. campestris*, and two to *X. euvesicatoria* (Table S4-4 and Figure S4-2, S4-3). The remaining thirteen strains showed ANI values between 85-94% when compared with known species of *Xanthomonas*. These strains were assigned to eight

cluster-type cliques or singletons according to the MiSI method, indicating the presence of at least eight potentially novel species. Given the findings of potentially novel species adding diversity to the existing *Xanthomonas* genus phylogeny, we established a robust phylogenetic tree based on the single-copy genes of these newly sequenced strains along with type or representative strains of the *Xanthomonas* genus (Figure 4-1). The OrthoFinder analysis assigned most genes (544,723; 99% of the total) to 11,456 orthogroups. There were 1005 orthogroups in all species, and 819 of these consisted entirely of single-copy genes. The phylogenetic reconstruction showed a considerable diversity of strains isolated from both plant hosts and the environment, and these strains were broadly distributed throughout the genus (Figure 4-1).

X. sontii, *X. sacchari*, *X. albilineans*, *X. hyacinthi*, *X. translucens*, *X. theicola,* and the two recently described species *X. bonasiae* and *X. youngii* were the only known species belonging to *Xanthomonas* group 1 (Bansal et al., 2021; Mafakheri et al., 2022; Rodriguez-R et al., 2012). The collection sequenced here adds four new *Xanthomonas* species to group 1, species I (strain F5), species II (strain F1), species IV (strain F10), and species VIII (strains 3307, 3498, and F4), all isolated from citrus plants and rainwater. Another new *Xanthomonas* species, species III (strain 60), from this collection clustered with early branching species at the base of the phylogenetic tree, along with *X. retroflexus*.

Our collection also added three new species to *Xanthomonas* group 2 (Figure 4-1). Most strains sequenced here belong to clade A, specifically to *X. arboricola* (S4-5) and *X. euroxanthea*. Strains isolated from bean seeds (CFBP 8151 and CFBP 8152) were closely related to *X. arboricola,* and strains isolated from rainwater (3075 and 3058) belong to potentially novel species group, species VI and species V, respectively, within clade A (Figure 4-1). Two of our sequenced strains belonged to clade B and were identified as *X. euvesicatoria* (CFBP 7921 and CFBP 7922).

Strains in clade C isolated from bean seed, tomato, and nightshade plants belong to *X. cannabis* species (16 strains). This clade also harbors one novel *Xanthomonas* species, species VII (strains 3793 and 4461), isolated from rainwater. Crop-associated *Xanthomonas campestris* strains (9 strains) isolated from radish, bean, and tomato plants belong to clade D, including pathogenic *X. campestris* strains (Figure 4-1).

4.3.2 Distribution of T3SS clusters across the phylogeny

We screened three types of T3SS clusters known within the genus *Xanthomonas* across the set of genomes: (1) the *hrp2* cluster present in group 2 xanthomonads (Tampakaki et al., 2010), (2) the SPI-1 type present in *X. albilineans* (Pieretti et al., 2009), and (3) the noncanonical T3SS cluster present in *X. translucens* (Wichmann et al., 2013). In addition, we also included the T3SS cluster from *Stenotrophomonas* sp. to represent the *sct*-type cluster from a closely related genus. Among the 83 newly sequenced *Xanthomonas* genomes, 24% contained a functional T3SS cluster (Figure S4-4). Strain *Xanthomonas* sp. 60 showed the presence of a unique *sct*-type T3SS cluster with a gene organization comparable to the one found in *Stenotrophomonas chelatiphaga* DSM 21508 (Figure S4-5). Apart from partial T3SS clusters in strains of *X. fragariae*, we also identified single gene encoding protein V of sct-type cluster in *X. cannabis* CFB P8595, *X. cannabis* 8600, *X. cannabis* CFBP 8600, *X. arboricola* F12, *X. arboricola* 84A, and *Xanthomonas* sp. 4461 and a single *hrpF* in *X. arboricola* CFBP 7681 (Figure 4-2). T3SS clusters were missing in *X. pisi*, *X. floridensis*, *X. melonis*, *X. maliensis*, *X. sontii*, *X. sacchari*, and *X. retroflexus. X. phaseoli* pv. *phaseoli* CFBP 412 showed the presence of two types of T3SS, the *hrp2* cluster and the SPI-1 type cluster present in *X. albilineans*, like previous findings on *X. phaseoli* pv. *phaseoli* CFBP 6164 (Alavi et al., 2008) (Figure 4-2).

4.3.3 T3SS was gained and lost multiple times in the genus *Xanthomonas*

To better understand the evolution of T3SS clusters in the genus *Xanthomonas*, we next studied the presence and absence patterns of T3SS clusters among the analyzed genomes. The phyletic pattern generated above was input to GLOOME, which maps gain and loss events onto the phylogeny. According to the scenario estimated with GLOOME, there were several independent acquisition and loss events of T3SS clusters during the evolution of xanthomonads. Independent acquisition of the *sct*-type T3SS cluster was inferred to have occurred in *Xanthomonas* sp. 60 (Figure 4-2). The ancestor of *X. translucens, X. hyacinthi, Xanthomonas* sp. *F5,* and *X. theicola* acquired the *Xtr*-type T3SS cluster. This cluster was found to be conserved in the *X. translucens* clade. However, it has been subsequently lost in *Xanthomonas* sp. F5 and partially lost in *X. hyacinthi* and *X. theicola,* as indicated by a partial *Xtr*-type cluster (Figure S4-5 and Figure 4-2). The SPI-1 T3SS cluster was independently acquired in *X. albilineans* and *X. phaseoli* CFBP 412.

Next, in addition to the probabilities of gene-gain/loss estimated by GLOOME analysis, we noted the genomic context and sequence identities to identify *hrp2* cluster gain/loss events associated with group 2 xanthomonads. Based on the sequence identities of *hrp2* cluster genes, clades A, C, and D were observed to possess the *Xcc*-type *hrp2* cluster, while clade B, except for *X. nasturtii*, possessed the *Xeu*-type *hrp2* cluster. It is possible that replacement of the *Xcc*-type *hrp2* cluster by the *Xeu*-type *hrp2* cluster occurred within clade B strains, except in *X. nasturtii*, through rearrangements.

According to GLOOME analysis, regulators of T3SS, HrpX, and HrpG, were acquired by a common ancestor of *Xanthomonas* before the split into group 1 and group 2. A single loss event of these regulators occurred in group 2, where these genes were lost on the branch leading to *X.*

pisi DSM18956. In group 1 these genes were lost in several independent events: (1) on the branch leading to the common ancestor of *X. albilineans*, *X. sontii*, *X. sacchari*, and *Xanthomonas* sp. strains (F10, F1, F4, 3307, and 3498); (2) on the branch leading to *Xanthomonas* sp. F5; (3) on the branch leading to the common ancestor of *X. retroflexus* and *Xanthomonas* sp. 60. An alternative less parsimonious explanation is that these regulators were only acquired by group 2 *Xanthomonas* and were independently acquired by the cluster of *X. hyacinthi*, *X. translucens*, *X. theicola, and Xanthomonas* sp. F5, followed by loss of these genes together with loss of *Xtra*-type T3SS genes in *Xanthomonas* sp. F5. Similar to HrpX and HrpG, the regulators HpaR and HpaS were also acquired by the common ancestor of *Xanthomonas* before the split of group 1 and group 2. These regulators were lost independently on numerous occasions: (1) on the branch leading to *X. translucens* in group 1; (2) on the branch leading to *X. populi* in group 2, and (3) on the branch leading to *X. oryzae* in group 2.

4.3.4 T3E repertoire ranges from zero to forty-one in crop-associated and environmental strains

Previous studies involving genome screening for T3E of nonpathogenic *X. arboricola* strains indicated low effector gene loads, with a reduced core effector set ranged from zero to four effectors, namely, XopR, HpaA, XopF1, and AvrBs2 (Merda et al., 2017). Given the diversity of commensal xanthomonads spanning the entire phylogeny of the genus *Xanthomonas* and the fact that this study also included environmental strains, we hypothesized that low effector loads might be widespread among nonpathogenic xanthomonads that are either plant-associated or environmental strains, owing to their global broad host range. However, we caution that host range tests have not been conducted for each strain sequenced in this study. Thus, we refer to a global broad host range based on previous studies that either recovered nonpathogenic isolates from

diverse plant hosts or tested their host range on diverse plants. Our analysis indicated that effector repertoires vary greatly from zero to 41 (Figure S4-6).

X. arboricola strains sequenced in this study, although belonging to a monophyletic group, showed a considerable variation in the presence/absence of T3SS and the size of effector repertoires, ranging from one to twelve known effectors. Strains lacking T3SS but containing effectors, XopAW and XopAX, included crop-associated and environmental strains. Some crop-associated strains lacking a T3SS possessed an additional T3E, AvrBs2. Two *X. arboricola* strains (F21 and CFBP 6681) isolated from tomato lacked T3SS but possessed two effectors, AvrBs1 and XopH, in addition to XopAW and XopAX. AvrBs1 and XopH have been identified as plasmid-borne effectors in *X. euvesicatoria*, a tomato pathogen. Interestingly, a set of strains isolated from rainwater and from diverse crops such as walnut, pepper, and bean seed shared the same effector repertoire (XopR, XopF1, XopF2, AvrBs2, and XopAW), in addition to a functional T3SS. Strains of *X. arboricola* (CFBP 6825, CFBP 6826, and CFBP 6828) isolated from pepper possessed unusually large effector repertoires comprising 10 T3Es (XopZ2, XopR, XopP, XopF1, XopF2, XopAW, XopAR, XopAL1, XopAD, AvrBs2) comparable to those found in *X. arboricola* strains pathogenic on walnut (Figure S4-6).

We hypothesized that *X. arboricola* strains with intact T3SSs and larger effector repertoires (i.e.,> seven effectors) would be weakly pathogenic on their host of isolation. Pathogenicity assays using the dip-inoculation method mimicking a natural infection and *in-planta* population growth were performed using six *X. arboricola* strains (CFBP 6825, CFBP 6826, CFBP 6828, CFBP 6681, CFBP 7681, and CFBP 7680) possessing different repertoires of T3Es and presence/absence of a T3SS in tomato cv. FL47R. Strains CFBP 6825, CFBP 6826, and CFBP 6828 triggered slight disease symptoms on tomato leaves at 10 days after inoculation (DAI) (Figure S7A). Strains CFBP

6681 and CFBP 7681 were unable to colonize the host tissue, while strains CFBP 6825, CFBP 6826, and CFBP 6828 maintained a population of ~ 10^5 CFU/cm^2 at 10 DAI (Figure S7B). No visible symptoms were observed for strain CFBP 7680, despite maintaining a population of ~ 10^4 CFU/cm^2 at 10 DAI (Figure S4-7B). A comparison of the T3Es among these strains indicated that strains that caused disease symptoms had a repertoire of T3E of similar size (11 effectors). In contrast, strains that were unable to cause disease lacked T3SS and had few effectors (<7 effectors) (Figure S4-7C).

Along with the variable presence of T3SSs, crop-associated nonpathogenic strains also varied in effector repertoire size, which ranged from one to 10 effectors in *X. cannabis* and reached 25 effectors in *X. campestris* (Figure S4-6). These strains with a higher number of effectors (>7) may suggest their possible pathogenic status, although their host range needs further exploration. Rain-derived *X. euroxanthea* strains lacked T3SSs and possessed a single effector, XopR. Crop-associated *X. euroxanthea* strains isolated from tomato and bean contained a T3SS and effectors, XopF1, XopF2, XopZ2, XopAK, and XopR. Rain-derived novel *Xanthomonas* sp. strains (3058, 3075, 3793, and 4461), although lacking canonical T3SSs, possessed orthologs of the HrpG/X master regulators. Strains 3058 and 3075 lacked any known effectors, while strains 3793 and 4461 contained homologs of AvrXccA1 and AvrXccA2. XopAW and AvrXccA1 were observed in the crop-associated strains 3793 and 4461 lacking T3SSs but containing sequences homologous of HrpG/X.

Based on the work described so far, we defined the lifestyle of strains based on the presence of the T3SS gene cluster and putative effectors. Strains that possessed an intact T3SS and 13 or more T3Es were considered as potentially pathogenic strains. *Xanthomonas* strains with a number of T3SS effectors ranging from 7 to 12 were considered potentially weak pathogens, while strains

lacking an intact T3SS and having <7 or no T3Es were considered commensals (Table S4-3). These three types of lifestyles defined here will be used for the following downstream analyses to infer genetic exchange between commensal and pathogenic *Xanthomonas* and to identify the features that define these lifestyles.

4.3.5 Genetic exchange between commensal, weakly pathogenic, and pathogenic *Xanthomonas* strains

To determine whether environmental (i.e., rainwater-associated) or crop-associated commensals or weakly pathogenic strains exchange genetic material with crop-associated pathogenic strains, we examined phylogenetic networks inferred from concatenated sequences of twelve housekeeping genes using SplitsTree (Figure 4-3). This analysis revealed reticulated events between commensal and pathogenic *Xanthomonas* strains, suggesting recombination. Two evident reticulation events were identified at the intersections of the network, one between pathogenic and commensal or weakly pathogenic strains belonging to *X. arboricola*, and another one between species encompassing clades B, C, and D of group 2 and species belonging to group 1, indicating the flow of genetic information between them (highlighted gray in Figure 4-3). Three intrinsic events can be observed by closely analyzing the parallelograms between Group 1 and Group 2. The first event is localized in the central part of the entire network. This reticulated event links the two main branches involving all species belonging to clades B, C, and D of group 2 and species belonging to group 1 (highlighted purple in Figure 4-3). The second event involves clade D (*X. campestris*), some species belonging to clade C and the entire group 1 (highlighted orange in Figure 4-3). Finally, the third event is exclusively shared between species from clade B and some species belonging to clade C (highlighted green in Figure 4-3). According to the neighbor-net tree, at least two strains isolated from rainwater appear to result from the above-mentioned putative

recombination events. Within the clade A parallelogram, *X. arboricola* strains (3790, 2768, 3140, 3272, 3046, and 3376) isolated from rainwater showed genetic recombination with type strain *X. arboricola* pv. *juglandis* CFBP 2528 and other crop-associated commensals/weakly pathogenic *X. arboricola* strains. A similar observation was found for *Xanthomonas* sp. strains (3793, 4461, 3307, and 3498) as these environmental strains exchanged genomic content with the ancestors of crop-associated pathogenic and commensal species belonging to clade C and of species belonging to Group 1, respectively.

Many exchange events were observed between the ancestors of crop-associated commensal, weakly pathogenic, and pathogenic strains isolated from the same crop hosts. Within the *X. arboricola* clade, two commensal strains (CFBP 7629 and CFBP 7634) and one weakly pathogenic strain (CFBP 7652) isolated from walnut and the pv. *juglandis* pathotype strain (CFBP 2528) showed a reticulated network. *X. campestris* ATCC33913, pathogenic on crucifers also appears to result from the genomic exchange with the ancestor of the commensal strains *X. campestris* (CFBP 13567 and CFBP 13568), both isolated from radish plants.

The sympatric association of commensals with pathogens, the presence of shared T3Es in commensals, and the observation of recombination signals between commensals and pathogens suggest that commensals and pathogens may act as repositories of fitness traits for each other. To test this hypothesis, we co-inoculated *X. cannabis* T55, a nonpathogenic strain with the T3E gene *avrBs1* with a transposon insertion, IS476, and a plant pathogenic strain, *X. euvesicatoria* 85-10, on pepper plants. *X. euvesicatoria* 85-10 harbors a functional *avrBs1* gene. AvrBs1 induces a hypersensitive response (HR) when infiltrated in the resistant pepper genotype Early California Wonder 10R containing the *Bs1* resistance gene. When screening for putative *X. euvesicatoria* transconjugants, we found that none of the strains isolated from the susceptible pepper genotype

(Early California Wonder) had acquired the disrupted avirulence gene from strain T55. In contrast, of nine pools each containing 50 transconjugants, at least two pools isolated from the resistant pepper genotype tested positive for an IS476 transposon insertion (Figure S4-8).

By screening for mobile genetic elements and associated genes using a computational approach, we further tested the hypothesis that commensal xanthomonads may act as reservoirs carrying fitness and virulence factors that can potentially be transferred to other strains through mobile genetic elements. MGEfinder identified at least one predicted mobile genetic island for each phylogenetic cluster and over 300 unique MGEs. Predicted mobile islands containing genes of possible interest and all islands identified as having a mobility gene, such as transposases or integrases. While many predicted mobile genetic elements that we found contain housekeeping genes, some contained genes that may play a role in increasing fitness or virulence. The predicted mobile islands contained genes associated with antimicrobial resistance and genes for bacterial and fungal competition, such as multi-drug efflux pumps, type IV secretion system genes, chitinase, and virulence factors. Most mobile genetic elements also contained integrase and phage-related genes. Type III effector XopAD, flanked on both ends by IS5 transposases, was observed in *X. campestris* strains F24 and F22. This island was also found in *X. campestris* pv. *raphani* 756C (the insertion location is shown in Figure S4-9). The type III effector XopAA, flanked by a putative transposase, was found in *X. euvesicatoria* strains CFBP 7922 and CFBP 7921. An endopolygalacturonase, known for degrading pectin in plant cell walls (Federici et al., 2001), was identified in *Xanthomonas* sp. F1. A peptidoglycan O-acetylase, which may alter bacterial cell walls to avoid lysis by an innate immune response (Sychantha et al., 2018), was detected in *X. arboricola* 3272.

4.3.6 Lifestyle had a significant effect in determining the repertoires of cell wall-degrading enzymes in *Xanthomonas*

We hypothesized that nonpathogenic *Xanthomonas* strains possess a distinct repertoire of cell-wall degrading enzymes compared to that from pathogens. Such repertoire might confer them the ability to utilize a wide range of carbohydrate substrates and colonize diverse host plants. Within the genus *Xanthomonas*, different species generally had similar types of CAZymes, but with large variations in the absolute numbers of genes within each category in the CAZy profiles (Figure 4-4). We assessed the effect of bacterial lifestyle on repertoires of genes coding for CAZymes. Distance-based redundancy analyses (dbRDA) of Jaccard distances and permutational multivariate analysis of variance (PERMANOVA) calculated on the genomic compositions of CAZyme family revealed a significant contribution of lifestyle to the distribution of gene repertoires for CAZyme ($R^2 = 0.06$, $p < 0.05$). Furthermore, a pairwise comparison of genomes of different lifestyles revealed CAZyme gene repertoire composition across commensals, weak pathogens, and pathogens to be significantly different ($p < 0.05$), suggesting that lifestyle plays an important role in determining the distribution of CAZymes. Principal component analysis on a matrix containing CAZymes using Jaccard distances showed increased separation of pathogenic from weakly pathogenic strains and commensals.

4.3.7 Association analysis to identify genes and features that define the lifestyle of *Xanthomonas* species

Next, we were interested in identifying genes either involved in the adaptation of *Xanthomonas* species to a commensal or a pathogenic lifestyle. Association analysis was performed on the 26,812 orthogroups from a set of 337 *Xanthomonas* genomes containing

representative strains from different clades that comprised both pathogens, commensals, and weak pathogens (Table S4-4, Figure S4-10). We hypothesized that commensals or weak pathogens and pathogens possess genes or gene families that define their adaptation to the respective lifestyles and have evolved these genes in a phylogeny-independent manner. Presence/absence as well as copy number matrix of orthologs, were used as input to conduct association analysis using three methods, Scoary, PhyloGLM, and hyperglm. Candidate genes present/absent or enriched/depleted in commensal, weakly pathogenic, and pathogenic strains were identified (Figure S4-11). Overall, each lifestyle category had unique genomic attributes. Interestingly, weak pathogens, as defined here by the presence of a T3SS and 7-12 effectors, contained overlapping genes with both pathogens and commensals (Figure S4-11).

Among the genes identified as enriched in commensals compared to pathogens and weak pathogens, we found those that belonged to the following functional categories: intracellular trafficking and secretion (type VI secretion system proteins and putative effectors), carbohydrate metabolism and transport, phages and transposons, post-translational modification, protein turnover and chaperones, replication and repair and defense mechanisms (specifically those encoding multi-drug efflux proteins) (Figure 4-5). Enrichment of genes involved in carbohydrate metabolism in commensals may indicate their ability to utilize broader energy sources than pathogens. A distinct set of DNA-binding transcriptional regulators in commensals may allow them to easily switch between different energy sources depending on their availability on a diverse group of hosts. Other genes that may also impart stress tolerance related to a much broader range of conditions associated with plants or environments outside plants include multi-drug efflux proteins, chaperones, outer-membrane transporters, and type I restriction-modification system genes. During epiphytic colonization of a broad range of hosts, enrichment of genes involved in

type VI secretion systems and effectors may allow the commensal xanthomonads to mediate interactions with the resident epiphytic community or pathogenic species (Drebes Dörr & Blokesch, 2018).

We next identified the genomic attributes shared by commensals and weak pathogens but absent from pathogens that could explain strategies evolved by nonpathogens during their association with a broad range of hosts (Figure 4-5). Glyoxalase bleomycin resistance protein dioxygenase was identified to be enriched in both the commensal and weakly pathogenic xanthomonads. This protein might be involved in providing tolerance to toxins produced by other phyllosphere colonizers, as seen recently with the clone expressing a putative glyoxylase/ bleomycin resistance dioxygenase showing neutralizing activity against toxoflavin, produced by *Burkholderia gladioli* (Choi et al., 2018). SGNH hydrolase-like domain containing acetyltransferase protein, AlgX, is involved in alginate biosynthesis in *Pseudomonas aeruginosa* and in virulence during cystic fibrosis. Although alginate is known to be involved in epiphytic fitness in *Pseudomonas syringae* (Yu et al., 1999), its contribution towards stress tolerance in xanthomonads remains to be investigated. Another gene belonging to PD-(D/E)XK nuclease superfamily, also characterized in contact-dependent DNA-degrading effectors (Sirias et al., 2020), was identified in both commensals and weak pathogens. Such nuclease effectors may impart a competitive advantage to the commensals and weak pathogens in a mixed species community as they establish their niche. At least three TonB-dependent receptors were identified to be associated with commensals and weak pathogens. Overall, TonB-dependent receptors are over-represented in xanthomonads and are thought to impart the adaptive ability to xanthomonads to utilize a wide variety of carbohydrates (Blanvillain et al., 2007). Another gene associated with commensal and weak pathogens belonged to the NRAMP (natural resistance-associated

macrophage proteins) family, conserved from bacteria to eukaryotes. These proteins function as proton-dependent manganese transporters in bacteria and have evolved to adapt to oxidative environments (Cellier et al., 2001; Kehres et al., 2000). Enrichment of these proteins in commensals and weak pathogens may explain why they can be simultaneously isolated along with pathogens from infected tissue and withstand host defenses, including reactive oxygen species.

We also screened the genomes for *fliC*, flagellin encoding gene, and specifically, the canonical flg22 epitope sequence to see the extent of variation within xanthomonads. It is not known whether commensal and weakly pathogenic xanthomonads can evade MAMP-triggered immunity induced by flagellin. Polymorphism in the 43rd amino acid residue in the *fliC* sequence, a part of the flg22 epitope, was previously observed in pathogenic xanthomonads (Malvino et al., 2022). More specifically, a change from the canonical residue Val^{43} to Asp^{43} allowed pathogenic xanthomonads to escape detection by FLS2 in *Arabidopsis* and tomato (Malvino et al., 2022; Sun et al., 2006). We observed that all commensals and weak pathogens carried the canonical Val^{43} in the flg22 epitope, except for *X. maliensis* LMG27592. In addition, some pathogenic xanthomonads also had the immunogenic version of flagellin (Table S4-3). Some commensals and weak pathogens possessed multiple copies (up to 3) of *fliC* in their genome. Based on their similarity to the canonical flg22 epitope residues, they can be expected to possess immunogenic properties. However, variations in other parts of the gene may have a role in modulating immunity in various hosts.

We also identified genes that were enriched in pathogens compared to commensals. As expected, most such genes were genes involved in T3SSs and effectors (Figure S4-11). Other features enriched in pathogens included glycosyl hydrolases, phage proteins, and chemotaxis proteins. Certain phage-related proteins and transposases were also identified as exclusively

present in pathogens but absent in commensals, indicating their role in mobilizing pathogen fitness and virulence genes. This finding suggests that it might be possible to devise phage-based control strategies using phages selective towards pathogens. Pathogens also exclusively contained certain transcriptional regulators, genes involved in c-di-GMP signaling pathway and chemotaxis, lipid metabolism, and carbohydrate and amino acid metabolism compared to commensals.

Genes that were enriched in pathogens compared to nonpathogens (commensals and weak pathogens) included a TonB-dependent receptor involved in Fe transport, a chemotaxis protein, specific cell-wall degrading enzymes, and an anti-sigma factor and AraC transcriptional regulator (Figure 4-5). Among the genes depleted in commensals and weak pathogens was an adenylosuccinate synthetase (S-AMPS) involved in *de novo* purine biosynthesis in the cytoplasm. This pathway is connected to the TCA cycle and, thus, to the central metabolism. This enzyme links GTP hydrolysis to inosine monophosphate (IMP) condensation with L-aspartate to produce adenylosuccinate (S-AMP). It can be hypothesized that pathogens may experience limitations of these compounds in their niche, for example, apoplast or xylem, and, thus, have evolved the ability to synthesize S-AMP. This enzyme was found in the periplasm-enriched fraction of *X. citri* but not in *X. fuscans* subsp. *aurantifolii* type B (Zandonadi et al., 2020). Another gene involved in lipid metabolism, annotated as phosphatidylserine/phosphatidylglycerophosphate/cardiolipin synthase or related enzyme was associated exclusively with pathogenic *Xanthomonas*. *X. campestris* was characterized for its unique lipid metabolism pathways with a bifunctional enzyme for PE synthesis that functions in a serine-dependentor an ethanolamine-dependent pathway, depending upon the availability of substrates *in-planta* (Aktas & Narberhaus, 2015).

4.4 Discussion

Since the identification of nonpathogenic *Xanthomonas* strains isolated simultaneously with pathogenic strains and in close association with plants or plants debris, the interest in exploring the role of these strains in the evolution of the genus, their potential for pathogenicity, and their contribution towards the microbiota-mediated extended plant immunity have increased (Bansal et al., 2020, 2021; Cesbron et al., 2015; Entila et al., 2023; Essakhi et al., 2015; Gonzalez et al., 2002; T. Li et al., 2020; Martins et al., 2020; Pfeilmeier et al., 2021, 2023; Triplett et al., 2015; Vauterin et al., 1996). In this study, we attempted to address these aspects to understand how the diversity spanning the commensal to pathogenic lifestyles across the genus *Xanthomonas* has evolved. We harnessed the unexplored diversity that the collection sequenced in this study brought, particularly the previously poorly explored group 1 xanthomonads species (Studholme et al., 2011). These new genomes were combined with a collection of 1,834 high-quality publicly available *Xanthomonas* genomes to obtain a phylogenetically representative set of 337 genomes spanning the different lifestyles present in the genus for comparative analysis.

We observed variable, but not random, presence/absence of T3SS clusters as indicated in the previous studies (Cesbron et al., 2015; Fang et al., 2015; Merda et al., 2017; Triplett et al., 2015). This study extends these findings to include some atypical T3SS clusters not limited to group 1 xanthomonads. This finding of gain of atypical T3SSs in addition to the *hrp2* cluster in some group 2 strains raises important questions about its functional significance and ecological role. The variable presence of T3SSs and effector repertoire sizes in xanthomonads ranging from 0 to 41 effectors indicate a plethora of diversity in the lifestyles of xanthomonads associated with plants along a continuum commensal endophyte to opportunistic pathogens or weak pathogens to pathogens.

As a genus for which research has been highly focused on pathogenic potential, commensal or nonpathogenic xanthomonads represent a largely ignored component (Vauterin et al., 1996). Unlike pseudomonads, this side of the continuum that spans from commensal to opportunistic lifestyles has not been well studied in xanthomonads. We lack understanding as to what extent nonpathogenic xanthomonads play a role in being evolutionary partners of plants with adaptive value contributing to overall plant health or whether they are just the members that contribute to niche filling, a process influenced by plant traits but are of minor adaptive importance in terms of fitness or growth benefit to the plant host. Hacquard et al. (2017) proposed a system involving multiple layers of barriers for establishing homeostasis with plants. Here, we assess how xanthomonads may have established association with the plants across the continuum of lifestyles from commensals to pathogens using this framework. The first two protective layers are microbiota exhibiting nutritional and niche competition and plant physical barriers. Based on the association analysis conducted in this study, we identified several traits in commensals, such as enrichment in type VI secretion system genes, transcriptional regulators, and carbohydrate metabolism genes, which may indicate their ability to overcome nutritional and niche competition with other microflora members associated with a wide range of plant hosts. Distinct cell-wall degrading enzyme repertoire in commensals, weak pathogens, and pathogens may also indicate their differential ability to overcome epidermal cell barriers. The next layer in maintaining homeostasis with plants is MAMP-triggered immunity (MTI). We found that commensal, weakly pathogenic, and few pathogenic xanthomonads possess canonical flg22 immunogenic epitopes, indicating that plants recognize and mount an innate immune response against them. Whether they can suppress this defense response may depend on the presence of T3SSs and effectors. It is hypothesized that a minimal repertoire of effectors present in some nonpathogenic *X. arboricola*

may help them suppress the MTI (Merda et al., 2017). Some commensal strains from our collection had no known effectors, indicating that MTI may explain their low abundance and lower in planta population. Their simultaneous isolation with pathogenic xanthomonads may also suggest that association with pathogens allows them to take advantage of innate immune response suppression performed by the pathogens. However, the ability of commensal and weakly pathogenic xanthomonads to activate MTI may also suggest their contribution towards extending plant immunity against other pathogenic bacteria or fungi, as seen in *Sphingomonas* (Innerebner et al., 2011). Apart from this epitope conservation, we observed variation in the rest of the flagellin sequence of commensals and weak pathogens. The importance of such sequence diversification in commensals and multiple copies of flagellin in evading MTI needs to be further explored. MTI suppression by type III secreted effectors has been demonstrated in many pathogenic xanthomonads. Analysis in this study showed that weak pathogens could also cross this MTI barrier due to the presence of a larger set of effectors (7-12 effectors) and intact T3SS. Whether minimized ancestral T3E repertoire of commensal xanthomonads is sufficient for them to overcome MTI barrier needs to be explored as commensals lack T3SS and thus secretion and translocation of these effectors might be of question. However, Merda et al. (2017) indicated the possibility of the secretion of effectors, specifically *xopR* and *avrBs2,* mediated by the flagellar apparatus (Journet et al., 2005). These effectors may also have functions independent of T3SS. Simultaneous colonization of commensals and pathogenic xanthomonads may also mean that commensal xanthomonads coordinate the action of these effectors and share these effectors as a public good with the pathogenic members, similar to the phenomenon demonstrated with studies using effectorless *Pseudomonas* strains (Ruiz-Bedoya et al., 2023). Experiments evaluating the effect of coinfection on their collective virulence may further our understanding of the importance

of reduced effector repertoires in commensal xanthomonads. These effectors in commensal strains may also have a role outside the host, similar to that shown for AvrBs1, IS476, and the associated plasmid offering enhanced overwintering potential (O'Garro et al., 1997). Further functional assessment of MTI-inducing and MTI-suppressive abilities of the commensals and weak pathogens from diverse clades across the phylogeny may clarify whether they can actively or passively cross the MTI barrier and how they may contribute towards microbiota-host homeostasis and ultimately towards plant growth-defense tradeoff and plant fitness (Ma et al., 2021). Although the opportunistic nature of nonpathogens has been documented (Gitaitis et al., 1987), the contribution of T3SS and associated minimized effector repertoire or distinct CWDE repertoire towards such conditional pathogenicity has not been experimentally validated. Alternatively, conditional pathogenicity may result from altered host-microbiota homeostasis and a compromised immune response rather than the involvement of T3SS or CWDEs alone. This important question of the contribution of nonpathogenic xanthomonads as a member of the microbiota has been investigated by two independent studies that emphasized the role of T2SS and cell wall degrading enzymes in mediating the shift in microbiota and conditional pathogenicity (Entila et al. 2023; Pfeilmeier et al. 2023). Further, Entila et al. (2023) showed that conditional pathogenicity of a nonpathogenic xanthomonad strain, lacking *hrp2* cluster, is kept in check by suppression of cell-wall degrading enzymes by plant NADPH oxidase respiratory burst oxidase homolog D (RBOHD) through a negative feedback loop between DAMP-triggered immunity-led reactive oxygen species production by Arabidopsis and T2SS/cell-wall degrading enzymes.

The diversity of commensals and weak pathogens across the genus phylogeny present opportunities to study evolution of xanthomonads specifically in the context of important virulence factors such as T3SS, effectors, and CWDEs that explain their plant-associated lifestyles. While

CWDEs have been proposed to be acquired by the ancestor of xanthomonads, they seemed to have diversified along the course of evolution as xanthomonads established associations with diverse hosts. The differential repertoire of CWDEs accompanied by unique TonB-dependent receptors among pathogens and commensals may explain the differential niche colonization. While we did not study patterns of gain/loss of specific CWDEs, we observed that commensal and weak pathogens have distinct sets of CWDEs compared to pathogens, similar to the observation with pathogenic and nonpathogenic *X. arboricola* (Cesbron et al., 2015). Different methyl-accepting chemotaxis proteins in commensals and pathogens may also suggest how the perception of the environment may be lifestyle-dependent. Like Merda et al. (2016, 2017) observations, we found evidence for an ancestral gain of *hrp2* cluster before the split of Group 2 xanthomonads. Group 1 xanthomonads have independently acquired different T3SS clusters. Thus, commensals belonging to Group 1 xanthomonads may possess ancestral traits that allowed them to closely associate with plants, such as diverse regulators, carbohydrate metabolism, and defense/repair-related genes. Subsequently, upon diversification of Group 2 xanthomonads into subsequent species, each displaying a high degree of host specificity, the effector repertoire reshuffling was observed (Hajri et al., 2009). Examining the genomic context of *hrp2* cluster and identities of core *hrp2* genes led us to hypothesize that replacement of *Xcc*-type *hrp2* cluster by *Xeu*-type *hrp2* cluster occurred within clade B strains, except in *X. nasturtii*, through rearrangements. Merda et al. (2016, 2017) also observed gene flow among T3SS and effectors of Group 2 xanthomonads. The subsequent loss of *hrp2* cluster in certain lineages representing commensal xanthomonads was observed (Figure 2). T3SS and associated effector loss may suggest that T3SS may impose a fitness cost. Thus, under this model, commensals belonging to Group 2 may have been derived because of regressive evolution from pathogenic strains. We also assessed patterns of gain/loss of *hrp2*

regulators, *hrpG/X*, analyzed in this study to identify clues to their involvement in the regulation of T3SS, associated limited repertoire in weak pathogens, or their existence in commensals in the absence of T3SS. Pfeilmeier et al. (2023) found that these master regulators do not regulate T2SS in nonpathogenic *Xanthomonas* strains lacking T3SS. This suggests a distinct regulatory network in play in commensal and pathogenic xanthomonads upon colonization of the host. Our analysis also indicated that many commensal and pathogenic strains engage in gene exchange and possess shared mobile genetic elements and carrier genes. Although experiments demonstrating the transfer of T3SSs from pathogenic to nonpathogenic xanthomonads were insufficient to impart pathogenicity phenotype to nonpathogenic strains (Meline et al., 2019), gene flow among certain strains may explain the re-acquisition of T3SS. Such gain/loss events of T3SS and associated effectors may explain the heterogenous distribution of *hrp2* cluster and diversity in effector repertoires across strains. Merda et al. (2016) showed that genetic exchange involving life history traits between pathogenic and nonpathogenic *Xanthomonas* strains occurred likely through horizontal gene transfer and suggested a possibility of nonpathogenic strains acting as reservoirs of traits that allow the emergence of novel pathogenic strains (Meline et al., 2019). Here, we gathered empirical evidence for the transfer of an IS element from the nonpathogenic to the pathogenic strain upon co-inoculation of the same plant. This exchange disrupted the avirulence gene in the pathogen under selection pressure allowing it to overcome host resistance. This finding further highlights the role of commensals as a reservoir of traits that may contribute to the adaptation of pathogens resulting in new outbreaks, as demonstrated by Lee et al. (2020). HGT between pathogen and commensal strains has been demonstrated in some bacteria, converting nonpathogenic strains into pathogenic (Brouwer et al., 2013). Such adaptive traits may have been subject to gene transfer among commensal and pathogenic strains.

However, this study also demonstrates that T3SS and effectors are not the sole determinants of the lifestyle of xanthomonads. The commensals and weak pathogens have evolved strategies for tolerance to stresses with distinct sets of chemotaxis proteins, type VI secretion system, TonB-dependent receptors, chaperone, and transcriptional regulators. We previously examined gain/loss patterns of type VI secretion system clusters across *Xanthomonas* phylogeny and found evidence for non-random acquisition of T6SS and gene flow in core genes and effectors (Liyanapathiranage et al., 2022). Further examination of gain/loss patterns of other non-pathogen-enriched traits may help assess support for the regressive evolution model to explain the origin of commensal strains.

Finally, systematic evaluation of genomic traits associated with lifestyles might guide us as we develop diagnostic strategies to differentiate pathogenic and nonpathogenic strains associated with seed samples or infected field samples. A recent machine-learning approach developed to predict the phenotype of plant-associated xanthomonads indicated specific domains associated with pathogenic and nonpathogenic lifestyles (Molder et al., 2021), many of which were found to overlap with the candidates identified in this study. Our study further confirms that diagnostic methods cannot rely on type III secretion system gene markers alone to identify pathogenic xanthomonads.

This study highlights the diversity of lifestyles across the genus phylogeny along the continuum of commensal, weakly pathogenic, and pathogenic strains. We find that type III secretion system and effectors are not the only factors that define these lifestyles. Several niche adaptative factors were identified to be associated with each lifestyle. Commensals establish themselves on different hosts in the presence of various host defenses and competing microflora, as well as derive complex nutrients from a wide range of hosts while sustaining their populations

on a broad range of hosts. We also observed distinct cell-wall degrading enzyme repertoires that distinguish pathogenic vs. commensal or weakly pathogenic lifestyles. Conversely, pathogens rely on type III secretion system and associated effectors to subvert defense responses. Overall, commensals possess genes that allow them to tolerate stresses rather than avoid them in the absence of T3SS.

4.5 Acknowledgments

The work (proposal: 10.46936/10.25585/60001156) conducted by the US Department of Energy Joint Genome Institute (https://ror.org/04xm1d337), a DOE Office of Science User Facility, is supported by the Office of Science of the US Department of Energy operated under Contract No. DE-AC02-05CH11231.

We thank CIRM-CFBP (Beaucouzé, INRAE, France http://www6.inra.fr/cirm_eng/CFBP-Plant-Associated-Bacteria) for strain preservation and supply. We also thank Abi S. A. Marques and Marisa A. S. V. Ferreira for participating in the strain sampling from bean.

4.6 Data availability

Sequence data generated from this work have been deposited in the DOE JGI and GenBank database under the JGI taxon id and Bioproject accession number as given in Table S4-1.

4.7 References for chapter 4

Aktas, M., & Narberhaus, F. (2015). Unconventional membrane lipid biosynthesis in *Xanthomonas campestris*. *Environmental Microbiology*, *17*(9), 3116-3124.

Alavi, S. M., Sanjari, S., Durand, F., Brin, C., Manceau, C., & Poussier, S. (2008). Assessment of the genetic diversity of *Xanthomonas axonopodis* pv. *phaseoli* and *Xanthomonas fuscans* subsp. *fuscans* as a basis to identify putative pathogenicity genes and a type III secretion system of the SPI-1 family by multiple suppression subtractive hybridizations. *Applied and Environmental Microbiology*, *74*(10), 3295–3301.

Bansal, K., Kaur, A., Midha, S., Kumar, S., Korpole, S., & Patil, P. B. (2021). *Xanthomonas sontii* sp. Nov., a non-pathogenic bacterium isolated from healthy basmati rice (*Oryza sativa*) seeds from India. *Antonie van Leeuwenhoek*, *114*(11), 1935–1947.

Bansal, K., Kumar, S., & Patil, P. B. (2020). Phylogenomic insights into diversity and evolution of nonpathogenic *Xanthomonas* strains associated with Citrus. *MSphere*, *5*(2), e00087-20.

Berendsen, R. L., Vismans, G., Yu, K., Song, Y., de Jonge, R., Burgman, W. P., Burmølle, M., Herschend, J., Bakker, P. A. H. M., & Pieterse, C. M. J. (2018). Disease-induced assemblage of a plant-beneficial bacterial consortium. *The ISME Journal*, *12*(6), 1496-1507.

Blanvillain, S., Meyer, D., Boulanger, A., Lautier, M., Guynet, C., Denancé, N., Vasse, J., Lauber, E., & Arlat, M. (2007). Plant carbohydrate scavenging through TonB-dependent receptors: a feature shared by phytopathogenic and aquatic bacteria. *PLoS ONE*, *2*(2), e224.

Brouwer, M. S., Roberts, A. P., Hussain, H., Williams, R. J., Allan, E., & Mullany, P. (2013). Horizontal gene transfer converts non-toxigenic *Clostridium difficile* strains into toxin producers. *Nature Communications*, *4*(1), 2601.

Brynildsrud, O., Bohlin, J., Scheffer, L., & Eldholm, V. (2016). Rapid scoring of genes in microbial pan-genome-wide association studies with Scoary. *Genome Biology*, *17*(1),1-9.

Cantalapiedra, C. P., Hernández-Plaza, A., Letunic, I., Bork, P., & Huerta-Cepas, J. (2021). eggNOG-mapper v2: functional annotation, orthology assignments, and domain prediction at the metagenomic scale. *Molecular Biology and Evolution*, *38*(12), 5825–5829.

Cellier, M. F., Bergevin, I., Boyer, E., & Richer, E. (2001). Polyphyletic origins of bacterial Nramp transporters. *Trends in Genetics: TIG*, *17*(7), 365-370.

Cesbron, S., Briand, M., Essakhi, S., Gironde, S., Boureau, T., Manceau, C., Fischer-Le Saux, M., & Jacques, M.-A. (2015). Comparative genomics of pathogenic and nonpathogenic strains of *Xanthomonas arboricola* unveil molecular and evolutionary events linked to pathoadaptation. *Frontiers in Plant Science*, *6*, 1126.

Choi, J.-E., Nguyen, C. M., Lee, B., Park, J. H., Oh, J. Y., Choi, J. S., Kim, J.-C., & Song, J. K. (2018). Isolation and characterization of a novel metagenomic enzyme capable of degrading bacterial phytotoxin toxoflavin. *PLOS ONE*, *13*(1), e0183893.

Cock, P. J. A., Antao, T., Chang, J. T., Chapman, B. A., Cox, C. J., Dalke, A., Friedberg, I., Hamelryck, T., Kauff, F., Wilczynski, B., & de Hoon, M. J. L. (2009). Biopython: freely available Python tools for computational molecular biology and bioinformatics. *Bioinformatics (Oxford, England)*, *25*(11), 1422–1423.

Cohen, O., Ashkenazy, H., Belinky, F., Huchon, D., & Pupko, T. (2010). GLOOME: gain loss mapping engine. *Bioinformatics (Oxford, England)*, *26*(22), 2914–2915.

Dixon, P. (2003). VEGAN, a package of R functions for community ecology. *Journal of Vegetation Science*, *14*(6), 927–930.

Drebes Dörr, N. C., & Blokesch, M. (2018). Bacterial type VI secretion system facilitates niche domination. *Proceedings of the National Academy of Sciences of the United States of America*, *115*(36), 8855–8857.

Durrant, M. G., Li, M. M., Siranosian, B. A., Montgomery, S. B., & Bhatt, A. S. (2020). A bioinformatic analysis of integrative mobile genetic elements highlights their role in bacterial adaptation. *Cell Host & Microbe*, *27*(1), 140-153.

Emms, D. M., & Kelly, S. (2015). OrthoFinder: solving fundamental biases in whole genome comparisons dramatically improves orthogroup inference accuracy. *Genome Biology*, *16*(1), 1-14.

Emms, D. M., & Kelly, S. (2017). STRIDE: species tree root inference from gene duplication events. *Molecular Biology and Evolution*, *34*(12), 3267–3278.

Emms, D. M., & Kelly, S. (2018). STAG: species tree inference from all genes. *BioRxiv*, 267914.

Emms, D. M., & Kelly, S. (2019). OrthoFinder: phylogenetic orthology inference for comparative genomics. *Genome Biology*, *20*(1),1-14.

Entila, F., Han, X., Mine, A., Schulze-Lefert, P., & Tsuda, K. (2023). Commensal lifestyle regulated by a negative feedback loop between *Arabidopsis* ROS and the bacterial T2SS. *bioRxiv*, 2023-05.

Essakhi, S., Cesbron, S., Fischer-Le Saux, M., Bonneau, S., Jacques, M. A., & Manceau, C. (2015). Phylogenetic and variable-number tandem-repeat analyses identify nonpathogenic *Xanthomonas arboricola* lineages lacking the canonical type III secretion system. *Applied and Environmental Microbiology*, *81*(16), 5395–5410.

Failor, K. C., Schmale, D. G., Vinatzer, B. A., & Monteil, C. L. (2017). Ice nucleation active bacteria in precipitation are genetically diverse and nucleate ice by employing different mechanisms. *The ISME Journal*, *11*(12), 2740–2753.

Fang, Y., Lin, H., Wu, L., Ren, D., Ye, W., Dong, G., Zhu, L., & Guo, L. (2015). Genome sequence of *Xanthomonas sacchari* R1, a biocontrol bacterium isolated from the rice seed. *Journal of Biotechnology*, *206*, 77–78.

Federici, L., Caprari, C., Mattei, B., Savino, C., Di Matteo, A., De Lorenzo, G., Cervone, F., & Tsernoglou, D. (2001). Structural requirements of endopolygalacturonase for the interaction with PGIP (polygalacturonase-inhibiting protein). *Proceedings of the National Academy of Sciences of the United States of America*, *98*(23), 13425–13430.

Garita-Cambronero, J., Palacio-Bielsa, A., López, M. M., & Cubero, J. (2017). Pan-genomic analysis permits differentiation of virulent and non-virulent strains of *Xanthomonas arboricola* that cohabit *Prunus* spp. and elucidate bacterial virulence factors. *Frontiers in Microbiology*, *8, 573*.

Gitaitis, R. D. (1987). Pectolytic xanthomonads in mixed infections with *Pseudomonas syringae* pv. *syringae, P. syringae* pv. *tomato,* and *Xanthomonas campestris* pv. *vesicatoria* in tomato and pepper transplants. *Phytopathology*, *77*(4), 611-615.

Gitaitis, R. D., Sasser, M., & Stall, R. E. (1987). Opportunistic xanthomonads in mixed infections with *Pseudomonas syringae* in tomato and pepper. In *Plant Pathogenic Bacteria: Proceedings of the Sixth International Conference on Plant Pathogenic Bacteria, Maryland, June 2–7, 1985* (pp. 767–767). Springer Netherlands.

Gonzalez, C., Restrepo, S., Tohme, J., & Verdier, V. (2002). Characterization of pathogenic and nonpathogenic strains of *Xanthomonas axonopodis* pv. *manihotis* by PCR-based DNA fingerprinting techniques. *FEMS Microbiology Letters*, *215*(1), 23–31.

Hacquard, S., Spaepen, S., Garrido-Oter, R., & Schulze-Lefert, P. (2017). Interplay between innate immunity and the plant microbiota. *Annual Review of Phytopathology*, *55*, 565–589.

Hajri, A., Brin, C., Hunault, G., Lardeux, F., Lemaire, C., Manceau, C., Boureau, T., & Poussier, S. (2009). A «Repertoire for Repertoire» hypothesis: repertoires of type three effectors are candidate determinants of host specificity in *Xanthomonas*. *PLOS ONE*, *4*(8), e6632.

Huson, D. H., Kloepper, T., & Bryant, D. (2008). SplitsTree 4.0-computation of phylogenetic trees and networks. *Bioinformatics*, *14*, 68-73.

Innerebner, G., Knief, C., & Vorholt, J. A. (2011). Protection of *Arabidopsis thaliana* against leaf-pathogenic *Pseudomonas syringae* by *Sphingomonas* strains in a controlled model system. *Applied and Environmental Microbiology*, *77*(10), 3202-3210.

Ives, A. R., & Garland, T., Jr. (2010). Phylogenetic logistic regression for binary dependent variables. *Systematic Biology*, *59*(1), 9-26.

Jacobs, J. M., Pesce, C., Lefeuvre, P., & Koebnik, R. (2015). Comparative genomics of a cannabis pathogen reveals insight into the evolution of pathogenicity in *Xanthomonas*. *Frontiers in Plant Science*, *6,* 431.

Jacques, M. A., Arlat, M., Boulanger, A., Boureau, T., Carrère, S., Cesbron, S., Chen, N. W. G., Cociancich, S., Darrasse, A., Denancé, N., Fischer-Le Saux, M., Gagnevin, L., Koebnik, R., Lauber, E., Noël, L. D., Pieretti, I., Portier, P., Pruvost, O., Rieux, A., … Vernière, C. (2016). Using ecology, physiology, and genomics to understand host specificity in *Xanthomonas*. *Annual Review of Phytopathology*, *54*, 163–187.

Jain, C., Rodriguez-R, L. M., Phillippy, A. M., Konstantinidis, K. T., & Aluru, S. (2018). High throughput ANI analysis of 90K prokaryotic genomes reveals clear species boundaries. *Nature Communications*, *9*(1), 5114.

Joseph, S., & Forsythe, S. (2012). Insights into the emergent bacterial pathogen *Cronobacter* spp., generated by multilocus sequence typing and analysis. *Frontiers in Microbiology*, *3,* 397.

Journet, L., Hughes, K. T., & Cornelis, G. R. (2005). Type III secretion: a secretory pathway serving both motility and virulence. *Molecular Membrane Biology*, *22*(1–2), 41–50.

Kaoutari, A. E., Armougom, F., Gordon, J. I., Raoult, D., & Henrissat, B. (2013). The abundance and variety of carbohydrate-active enzymes in the human gut microbiota. *Nature Reviews Microbiology*, *11*(7), 497–504.

Kehres, D. G., Zaharik, M. L., Finlay, B. B., & Maguire, M. E. (2000). The NRAMP proteins of *Salmonella typhimurium* and *Escherichia coli* are selective manganese transporters involved in the response to reactive oxygen. *Molecular Microbiology*, *36*(5), 1085-1100.

Kelly, S., & Maini, P. K. (2013). DendroBLAST: Approximate phylogenetic trees in the absence of multiple sequence alignments. *PLoS ONE*, *8*(3), e58537.

Kim, M., Oh, H. S., Park, S. C., & Chun, J. (2014). Towards a taxonomic coherence between average nucleotide identity and 16S rRNA gene sequence similarity for species demarcation of prokaryotes. *International Journal of Systematic and Evolutionary Microbiology*, *64*(Pt_2), 346–351.

Lairson, L. L., Henrissat, B., Davies, G. J., & Withers, S. G. (2008). Glycosyltransferases: structures, functions, and mechanisms. *Annu. Rev. Biochem*., *77,* 521–555.

Lee, Y.-A., Yang, P.-Y., & Huang, S.-C. (2020). Characterization, phylogeny, and genome analyses of nonpathogenic *Xanthomonas campestris* strains isolated from Brassica Seeds. *Phytopathology*, *110*(5), 981-988.

Levy, A., Gonzalez, I. S., Mittelviefhaus, M., Clingenpeel, S., Paredes, S. H., Miao, J., Wang, K., Devescovi, G., Stillman, K., Monteiro, F., Alvarez, B. R., Lundberg, D. S., Lu, T.-Y., Lebeis, S., Jin, Z., McDonald, M., Klein, A. P., Feltcher, M. E., del Rio, T. G., … Dangl, J. L. (2018). Genomic features of bacterial adaptation to plants. *Nature Genetics*, *50*(1), 138–150.

Li, H., & Durbin, R. (2009). Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics (Oxford, England)*, *25*(14), 1754–1760.

Li, T., Mann, R., Sawbridge, T., Kaur, J., Auer, D., & Spangenberg, G. (2020). Novel *Xanthomonas* species from the perennial ryegrass seed microbiome – assessing the bioprotection activity of non-pathogenic relatives of pathogens. *Frontiers in Microbiology*, *11,* 1991.

Liyanapathiranage, P., Wagner, N., Avram, O., Pupko, T., & Potnis, N. (2022). Phylogenetic distribution and evolution of type VI secretion system in the genus *Xanthomonas*. *Frontiers in Microbiology*, *13*, 840308.

Ma, K. W., Niu, Y., Jia, Y., Ordon, J., Copeland, C., Emonet, A., Geldner, N., Guan, R., Stolze, S. C., Nakagami, H., Garrido-Oter, R., & Schulze-Lefert, P. (2021). Coordination of microbe-host homeostasis by crosstalk with plant innate immunity. *Nature Plants*, *7*(6), 814–825.

Mafakheri, H., Taghavi, S. M., Zarei, S., Rahimi, T., Hasannezhad, M. S., Portier, P., Fischer-Le Saux, M., Dimkić, I., Koebnik, R., Kuzmanović, N., & Osdaghi, E. (2022). Phenotypic and molecular-phylogenetic analyses reveal distinct features of crown gall-associated *Xanthomonas* strains. *Microbiology Spectrum*, *10*(1), e0057721.

Malvino, M. L., Bott, A. J., Green, C. E., Majumdar, T., & Hind, S. R. (2022). Influence of flagellin polymorphisms, gene regulation, and responsive memory on the motility of *Xanthomonas* species that cause bacterial spot disease of solanaceous plants. *Molecular Plant-Microbe Interactions*, *35*(2), 157-169.

Martins, L., Fernandes, C., Blom, J., Dia, N. C., Pothier, J. F., & Tavares, F. (2020). *Xanthomonas euroxanthea* sp. nov., a new xanthomonad species including pathogenic and non-pathogenic strains of walnut. *International Journal of Systematic and Evolutionary Microbiology*, *70*(12), 6024–6031.

Mechan Llontop, M. E., Tian, L., Sharma, P., Heflin, L., Bernal-Galeano, V., Haak, D. C., Clarke, C. R., & Vinatzer, B. A. (2021). Experimental evidence pointing to rain as a reservoir of tomato phyllosphere microbiota. *Phytobiomes Journal*, *5*(4), 382–399.

Meier-Kolthoff, J. P., Auch, A. F., Klenk, H. P., & Göker, M. (2013). Genome sequence-based species delimitation with confidence intervals and improved distance functions. *BMC Bioinformatics*, *14*(1), 1-14.

Meline, V., Delage, W., Brin, C., Li-Marchetti, C., Sochard, D., Arlat, M., Rousseau, C., Darrasse, A., Briand, M., Lebreton, G., Portier, P., Fischer-Le Saux, M., Durand, K., Jacques, M.-A., Belin, E., & Boureau, T. (2019). Role of the acquisition of a type 3 secretion system in the emergence of novel pathogenic strains of *Xanthomonas*. *Molecular Plant Pathology*, *20*(1), 33-50.

Merda, D., Bonneau, S., Guimbaud, J.-F., Durand, K., Brin, C., Boureau, T., Lemaire, C., Jacques, M. A., & Fischer-Le Saux, M. (2016). Recombination-prone bacterial strains form a reservoir from which epidemic clones emerge in agroecosystems. *Environmental Microbiology Reports*, *8*(5), 572–581.

Merda, D., Briand, M., Bosis, E., Rousseau, C., Portier, P., Barret, M., Jacques, M. A., & Fischer-Le Saux, M. (2017). Ancestral acquisitions, gene flow and multiple evolutionary trajectories of the type three secretion system and effectors in Xanthomonas plant pathogens. *Molecular Ecology*, *26*(21), 5939–5952.

Miyauchi, S., Kiss, E., Kuo, A., Drula, E., Kohler, A., Sánchez-García, M., Morin, E., Andreopoulos, B., Barry, K. W., Bonito, G., Buée, M., Carver, A., Chen, C., Cichocki, N., Clum, A., Culley, D., Crous, P. W., Fauchery, L., Girlanda, M., … Martin, F. M. (2020). Large-scale genome sequencing of mycorrhizal fungi provides insights into the early evolution of symbiotic traits. *Nature Communications*, *11*(1), 5125.

O'Garro, L. W., Gibbs, H., & Newton, A. (1997). Mutation in the *avrBs1* avirulence gene of *Xanthomonas campestris* pv. *vesicatoria* influences survival of the bacterium in soil and detached leaf tissue. *Phytopathology*, *87*(9), 960-966.

Olm, M. R., Brown, C. T., Brooks, B., & Banfield, J. F. (2017). dRep: a tool for fast and accurate genomic comparisons that enables improved genome recovery from metagenomes through de-replication. *The ISME Journal*, *11*(12), 2864.

Pesce, C., Jacobs, J. M., Berthelot, E., Perret, M., Vancheva, T., Bragard, C., & Koebnik, R. (2017). Comparative genomics identifies a novel conserved protein, HpaT, in proteobacterial type III secretion systems that do not possess the putative translocon protein HrpF. *Frontiers in Microbiology*, *8,* 1177.

Pfeilmeier, S., Petti, G. C., Bortfeld-Miller, M., Daniel, B., Field, C. M., Sunagawa, S., & Vorholt, J. A. (2021). The plant NADPH oxidase RBOHD is required for microbiota homeostasis in leaves. *Nature Microbiology*, *6*(7), 852-864.

Pfeilmeier, S., Werz, A., Ote, M., Bortfeld-Miller, M., Kirner, P., Keppler, A., Hemmerle, L., Gäbelein, C. G., Pestalozzi, C. M., & Vorholt, J. A. (2023). Dysbiosis of a leaf microbiome is caused by enzyme secretion of opportunistic *Xanthomonas* strains. *bioRxiv*, 2023-05.

Pieretti, I., Pesic, A., Petras, D., Royer, M., Süssmuth, R. D., & Cociancich, S. (2015). What makes *Xanthomonas albilineans* unique amongst xanthomonads? *Frontiers in Plant Science*, *6*, 289.

Pieretti, I., Royer, M., Barbe, V., Carrere, S., Koebnik, R., Cociancich, S., Couloux, A., Darrasse, A., Gouzy, J., Jacques, M. A., Lauber, E., Manceau, C., Mangenot, S., Poussier, S., Segurens, B., Szurek, B., Verdier, V., Arlat, M., & Rott, P. (2009). The complete genome sequence of *Xanthomonas albilineans* provides new insights into the reductive genome evolution of the xylem-limited Xanthomonadaceae. *BMC Genomics*, *10*(1), 1-15.

Pritchard, L., Glover, R. H., Humphris, S., Elphinstone, J. G., & Toth, I. K. (2015). Genomics and taxonomy in diagnostics for food security: soft-rotting enterobacterial plant pathogens. *Analytical Methods*, *8*(1), 12–24.

Roach, R., Mann, R., Gambley, C. G., Chapman, T., Shivas, R. G., & Rodoni, B. (2019). Genomic sequence analysis reveals diversity of Australian *Xanthomonas* species associated with bacterial leaf spot of tomato, capsicum and chilli. *BMC Genomics*, *20*(1), 1-22.

Roach, R., Mann, R., Gambley, C. G., Shivas, R. G., & Rodoni, B. (2018). Identification of *Xanthomonas* species associated with bacterial leaf spot of tomato, capsicum and chilli crops in eastern Australia. *European Journal of Plant Pathology*, *150*(3), 595–608.

Rodriguez-R, L. M., Grajales, A., Arrieta-Ortiz, M. L., Salazar, C., Restrepo, S., & Bernal, A. (2012). Genomes-based phylogeny of the genus *Xanthomonas*. *BMC Microbiology*, *12*(1), 1-14.

Ruiz-Bedoya, T., Wang, P. W., Desveaux, D., & Guttman, D. S. (2023). Cooperative virulence via the collective action of secreted pathogen effectors. *Nature Microbiology*, 1–11.

Sadhukhan, S., Jacques, M. A., & Potnis, N. (2023). Influence of co-occurring weakly pathogenic bacterial species on bacterial spot disease dynamics on tomato. *bioRxiv*, 2023-04.

Sawada, H., Kunugi, Y., Watauchi, K., Kudo, A., & Sato, T. (2011). Bacterial spot, a new disease of grapevine (*Vitis vinifera*) caused by *Xanthomonas arboricola*. *Japanese Journal of Phytopathology*, *77*(1), 7–22.

Seemann, T. (2014). Prokka: rapid prokaryotic genome annotation. *Bioinformatics*, *30*(14), 2068–2069.

Sirias, D., Utter, D. R., & Gibbs, K. A. (2020). A family of contact-dependent nuclease effectors contain an exchangeable, species-identifying domain. *bioRxiv,* 2020-02.

Studholme, D. J., Wasukira, A., Paszkiewicz, K., Aritua, V., Thwaites, R., Smith, J., & Grant, M. (2011). Draft genome sequences of *Xanthomonas sacchari* and two banana-associated xanthomonads reveal insights into the *Xanthomonas* group 1 clade. *Genes*, *2*(4), 1050–1065.

Sullivan, M. J., Petty, N. K., & Beatson, S. A. (2011). Easyfig: a genome comparison visualizer. *Bioinformatics*, *27*(7), 1009–1010.

Sun, W., Dunning, F. M., Pfund, C., Weingarten, R., & Bent, A. F. (2006). Within-species flagellin polymorphism in *Xanthomonas campestris* pv. *campestris* and its impact on elicitation of *Arabidopsis* FLAGELLIN SENSING2–dependent defenses. *The Plant Cell*, *18*(3), 764-779.

Sychantha, D., Brott, A. S., Jones, C. S., & Clarke, A. J. (2018). Mechanistic pathways for peptidoglycan O-acetylation and de-O-acetylation. *Frontiers in Microbiology*, *9*, 2332.

Tampakaki, A. P., Skandalis, N., Gazi, A. D., Bastaki, M. N., Panagiotis F., S., Charova, S. N., Kokkinidis, M., & Panopoulos, N. J. (2010). Playing the "Harp": evolution of our understanding of *hrp/hrc* genes. *Annual Review of Phytopathology*, *48*(1), 347–370.

te Molder, D., Poncheewin, W., Schaap, P. J., & Koehorst, J. J. (2021). Machine learning approaches to predict the plant-associated phenotype of *Xanthomonas* strains. *BMC Genomics*, *22*(1), 1-14.

Tian, L., Huang, C., Mazloom, R., Heath, L. S., & Vinatzer, B. A. (2020). LINbase: A web server for genome-based identification of prokaryotes as members of crowdsourced taxa. *Nucleic Acids Research*, *48*(W1), W529–W537.

Triplett, L. R., Verdier, V., Campillo, T., Van Malderghem, C., Cleenwerck, I., Maes, M., Deblais, L., Corral, R., Koita, O., Cottyn, B., & Leach, J. E. (2015). Characterization of a novel clade of *Xanthomonas* isolated from rice leaves in Mali and proposal of *Xanthomonas maliensis* sp. nov. *Antonie Van Leeuwenhoek*, *107*(4), 869–881.

Varghese, N. J., Mukherjee, S., Ivanova, N., Konstantinidis, K. T., Mavrommatis, K., Kyrpides, N. C., & Pati, A. (2015). Microbial species delineation using whole genome sequences. *Nucleic Acids Research*, *43*(14), 6761–6771.

Vauterin, L., Yang, P., Alvarez, A., Takikawa, Y., Roth, D. A., Vidaver, A. K., Stall, R. E., Kersters, K., & Swings, J. (1996). Identification of non-pathogenic *Xanthomonas* strains associated with plants. *Systematic and Applied Microbiology*, *19*(1), 96–105.

White, F. F., Potnis, N., Jones, J. B., & Koebnik, R. (2009). The type III effectors of *Xanthomonas*. *Molecular Plant Pathology*, *10*(6), 749–766.

Wichmann, F., Vorhölter, F., Hersemann, L., Widmer, F., Blom, J., Niehaus, K., Reinhard, S., Conradin, C., & Kölliker, R. (2013). The noncanonical type III secretion system of *Xanthomonas translucens* pv. *graminis* is essential for forage grass infection. *Molecular Plant Pathology*, *14*(6), 576–588.

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

Yu, G., Smith, D. K., Zhu, H., Guan, Y., & Lam, T. T.-Y. (2017). ggtree: An R package for visualization and annotation of phylogenetic trees with their covariates and other associated data. *Methods in Ecology and Evolution*, *8*(1), 28–36.

Yu, J., Peñaloza-Vázquez, A., Chakrabarty, A. M., & Bender, C. L. (1999). Involvement of the exopolysaccharide alginate in the virulence and epiphytic fitness of *Pseudomonas syringae* pv. *syringae*. *Molecular Microbiology*, *33*(4), 712-720.

Zandonadi, F. S., Ferreira, S. P., Alexandrino, A. V., Carnielli, C. M., Artier, J., Barcelos, M. P., Nicolela, N. C. S., Prieto, E. L., Goto, L. S., Belasque, J., & Novo-Mansur, M. T. M. (2020). Periplasm-enriched fractions from *Xanthomonas citri* subsp. *citri* type A and *X.*

fuscans subsp. *aurantifolii* type B present distinct proteomic profiles under in vitro pathogenicity induction. *PLoS ONE*, *15*(12), e0243867.

Zarei, S., Taghavi, S. M., & Osdaghi, E. (2022). Reassessment of the occurrence and distribution of *Xanthomonas arboricola* on fruit trees in Iran. *Plant Pathology*, *71*(9), 1859–1869.

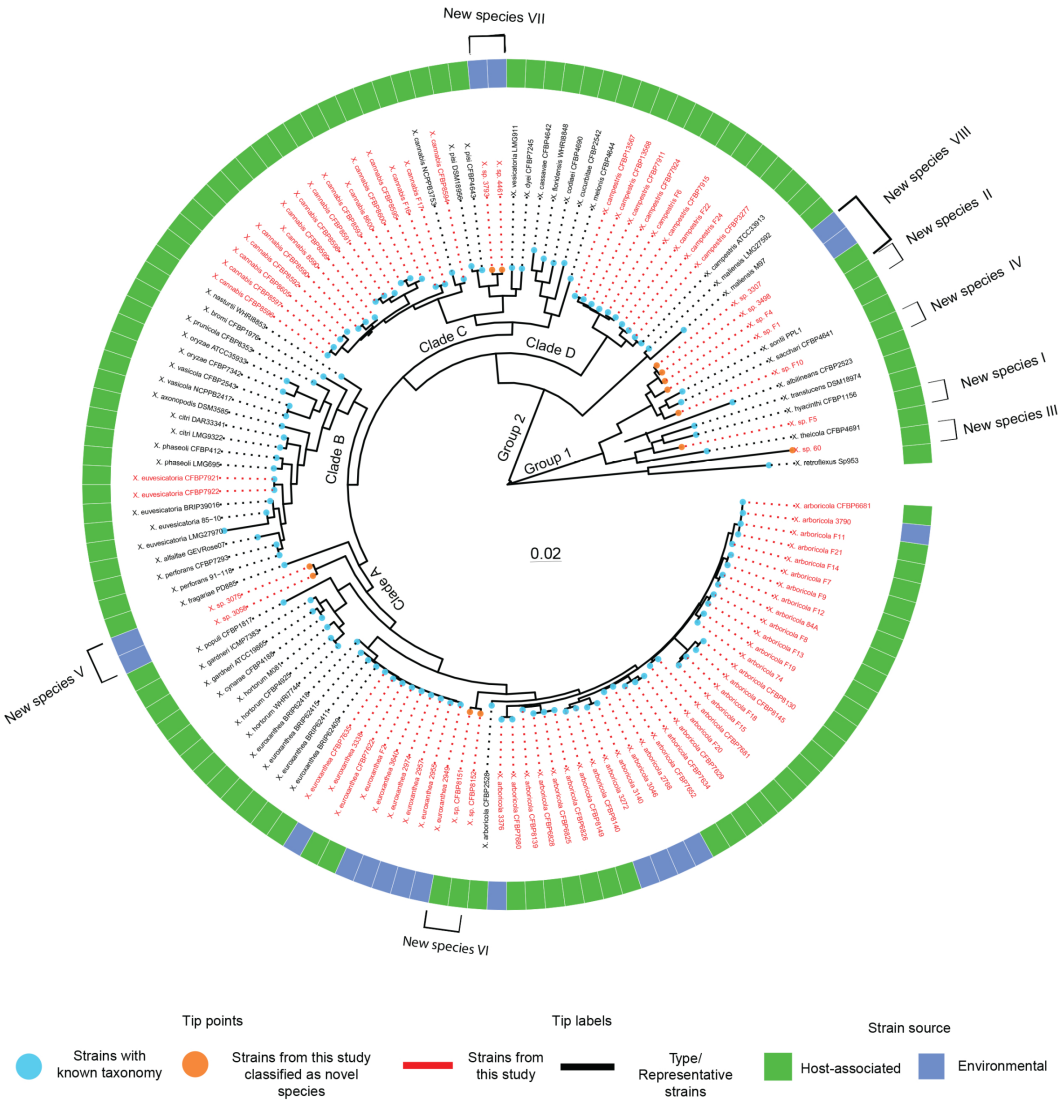

Figure 4-1: Comparative genome analysis demonstrated the presence of eight novel species in the genus *Xanthomonas*. Maximum-likelihood phylogeny based on the 1005 orthogroups of 134 strains representing the entire *Xanthomonas* genus. The phylogenetic tree was inferred using OrthoFinder v2.5.2 and drawn with R package ggtree. Here, *Xanthomonas* strains from this study are highlighted in red, while the representative/type strains are in black. The tip points are colored orange to show novel species identified from this study, while cyan represents *Xanthomonas* species with known taxonomy. Blue-colored blocks indicate environmental strains and host-associated strains are indicated by green-colored blocks surrounding the phylogenetic tree.

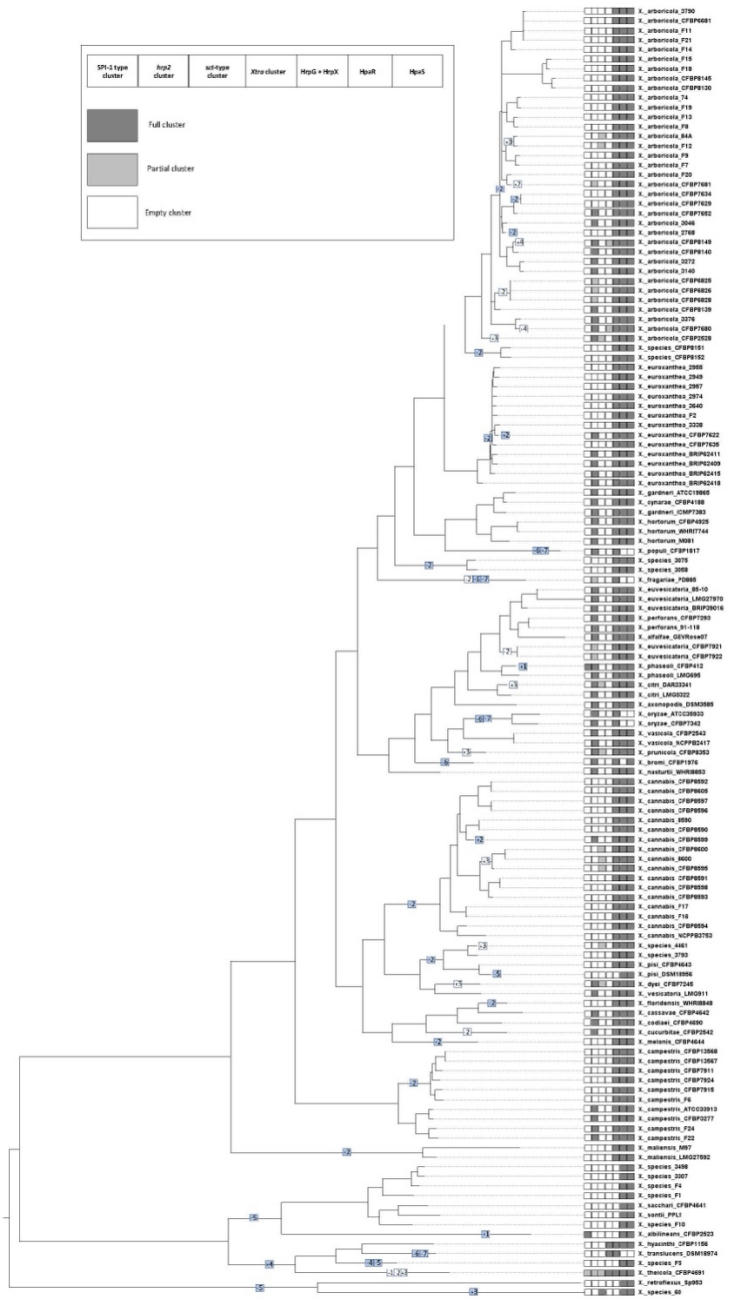

Figure 4-2: Multiple events of gain and loss of T3SS were evident in the genus *Xanthomonas* along with the presence of a novel T3SS cluster. Maximum-likelihood phylogeny based on the core proteome and T3SS cluster and regulators gain-loss prediction inferred for the 134 strains representing the entire *Xanthomonas* genus. The presence/absence of the different types of T3SS and regulators are represented by an ordered vector of size 7, such that a dark grey, a light grey, and a white i^{th} element in the vector indicates a full presence, a partial presence, and an absence of the i^{th} element, respectively. Inferred full and partial T3SS clusters are in color or white background, respectively. Acquisition and loss events are represented by plus and minus signs, respectively. For example, a colored +2 indicates a full acquisition of the 2nd T3SS in this figure, i.e., *hrp2*.

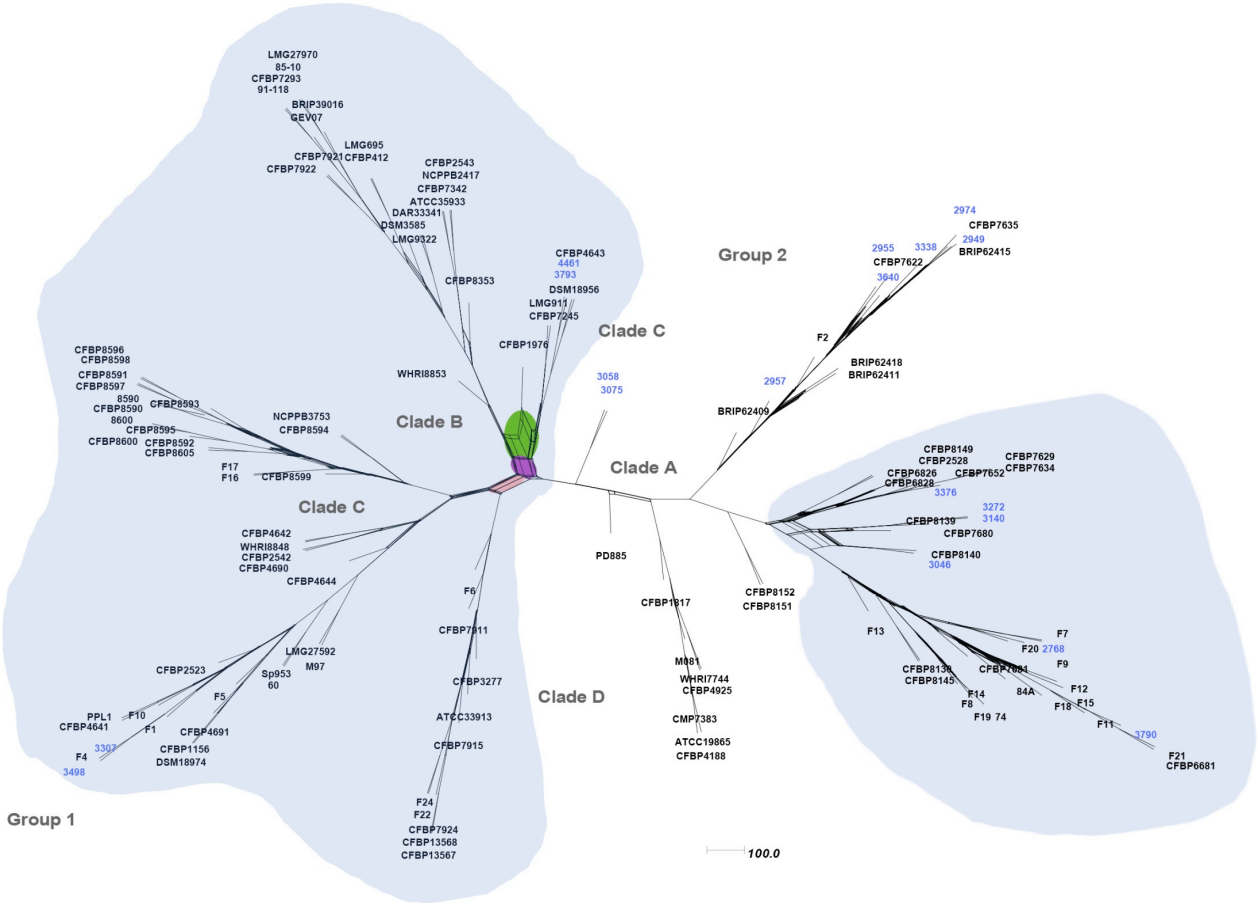

Figure 4-3: Phylogenetic network between pathogens, weak pathogens, and commensals suggests the possibility of several recombination events during their evolutionary history. Neighbor-net tree constructed using SplitsTree software based on concatenated 12 housekeeping gene sequences generated for the 134 *Xanthomonas* strains, indicating diversity and recombination events.

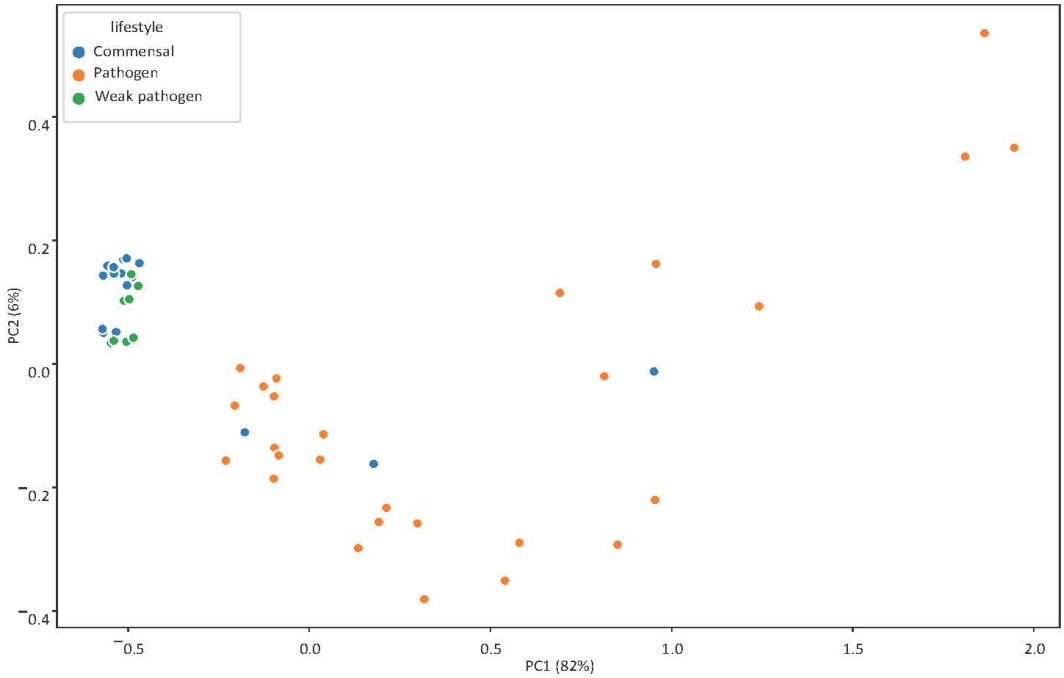

Figure 4-4: *Xanthomonas* lifestyle can be explained by altered CAZymes landscape. Principal Component Analysis (PCA) plot showing the contribution of the different bacterial lifestyle in the distribution of gene repertoire for carbohydrate active enzymes (CAZyme).

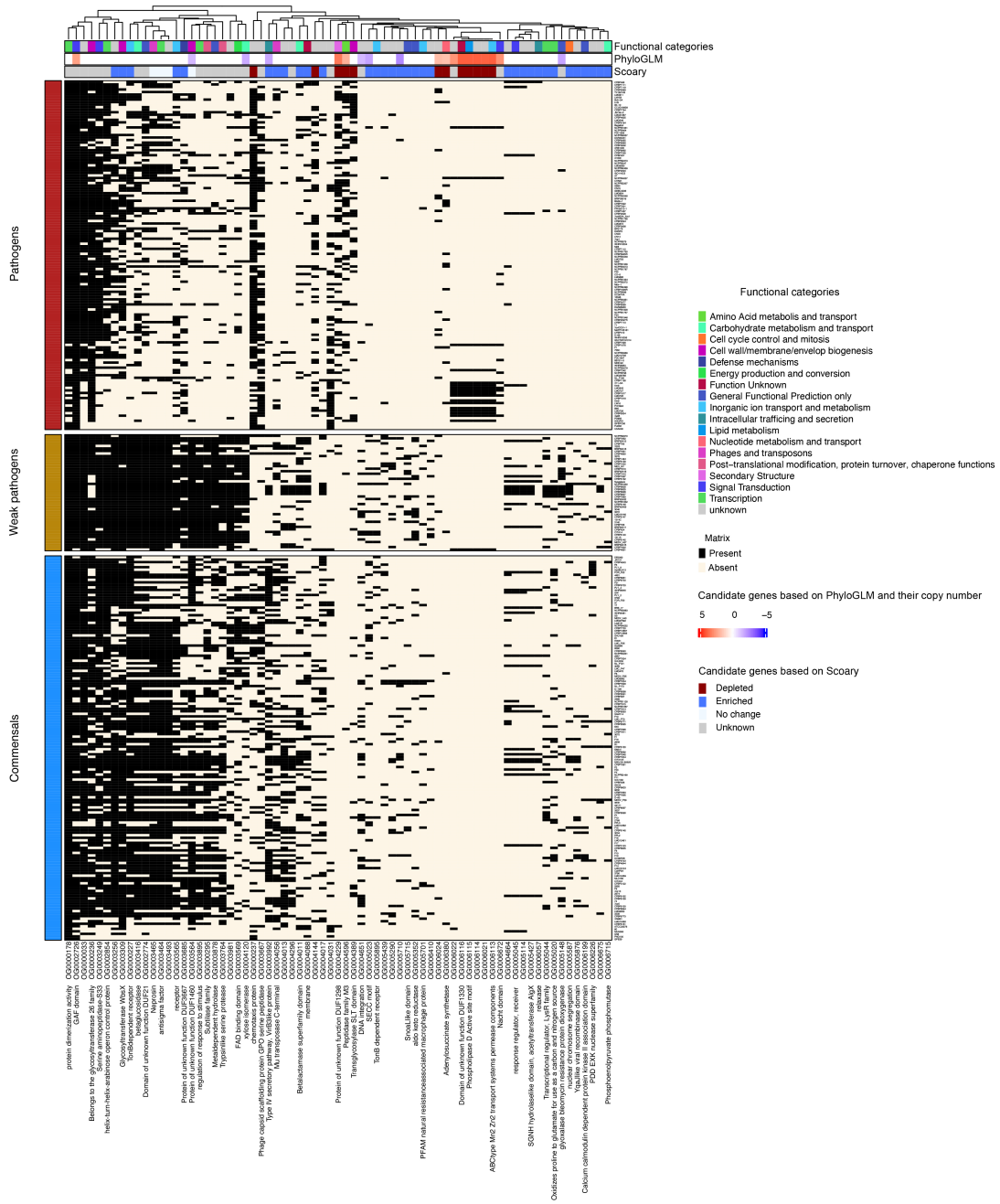

Figure 4-5: Genomic architecture of *Xanthomonas* contains signatures for both phylogenetic placement and their associated lifestyle. A complex heatmap shows the select candidate genes associated with commensal, weakly pathogenic, and pathogenic lifestyles. The candidates obtained from different methods in figure S11 were further narrowed by identifying those present/enriched in commensals and weak pathogens compared to pathogens and vice versa. The matrix shows the presence/absence of genes across these lifestyles along with their functional categories and annotations on the Y-axis.

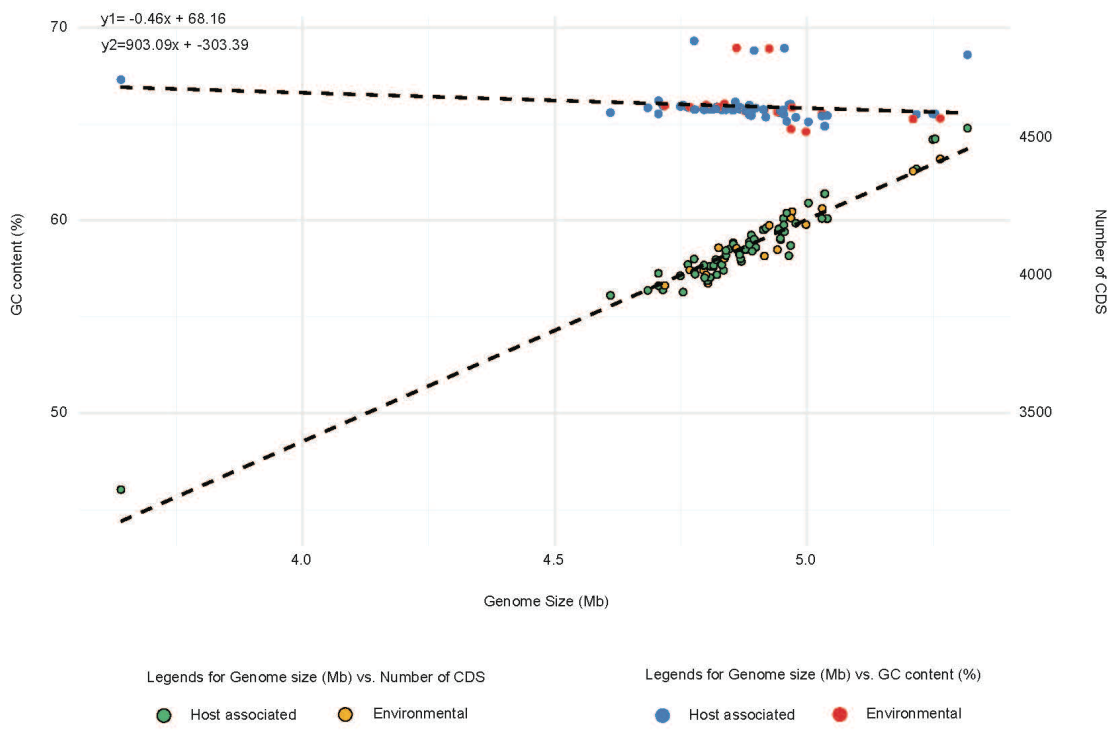

Figure S 4-1: Relationship between genome size, GC content, and number of CDS for the genomes sequenced for this study.

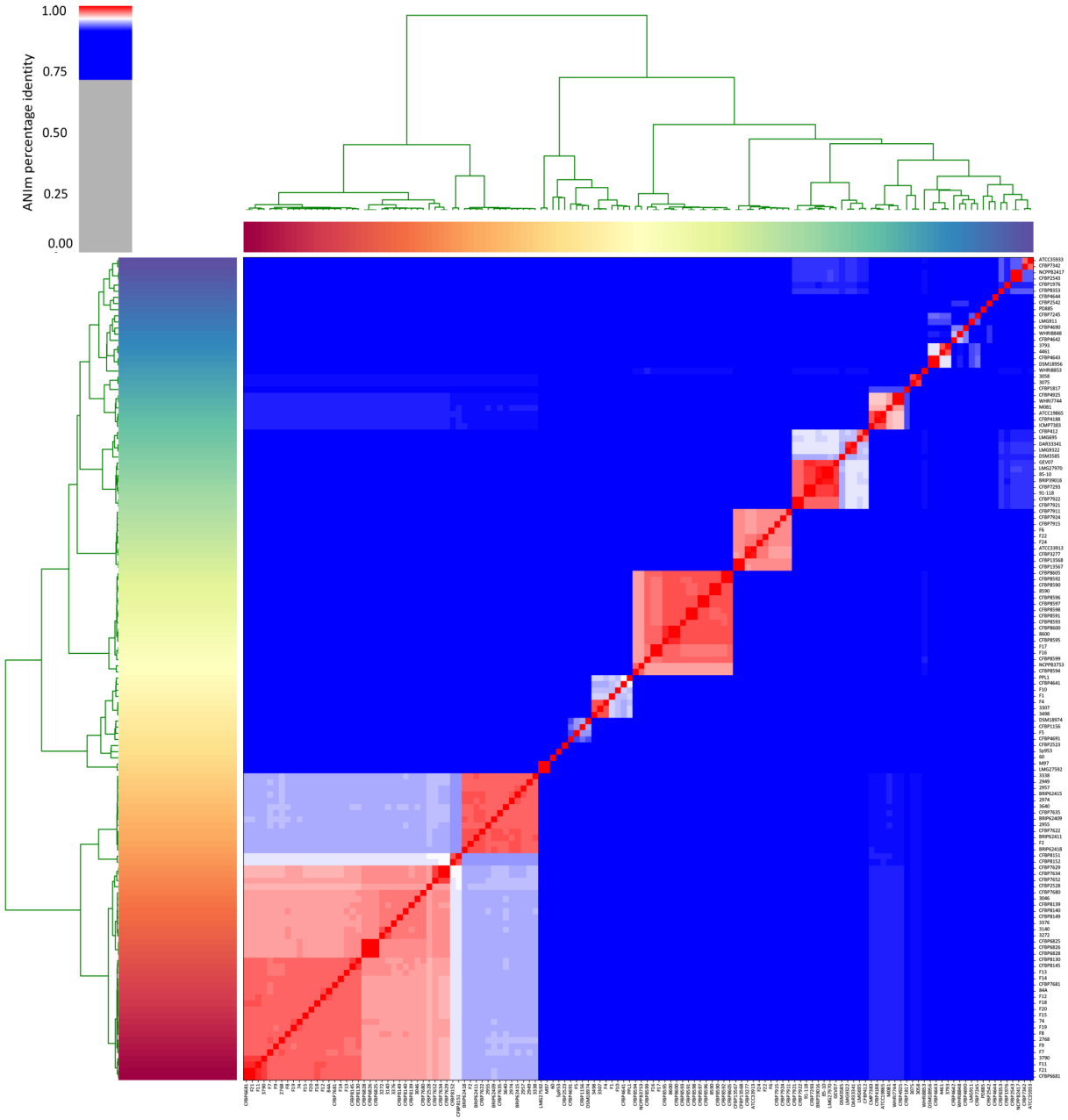

Figure S 4-2: Average nucleotide identity (ANI)-based heatmap showing the status of the 134 *Xanthomonas* strains. The intensity of the color indicates the level of identity of all-versus-all genomes as depicted by the scale.

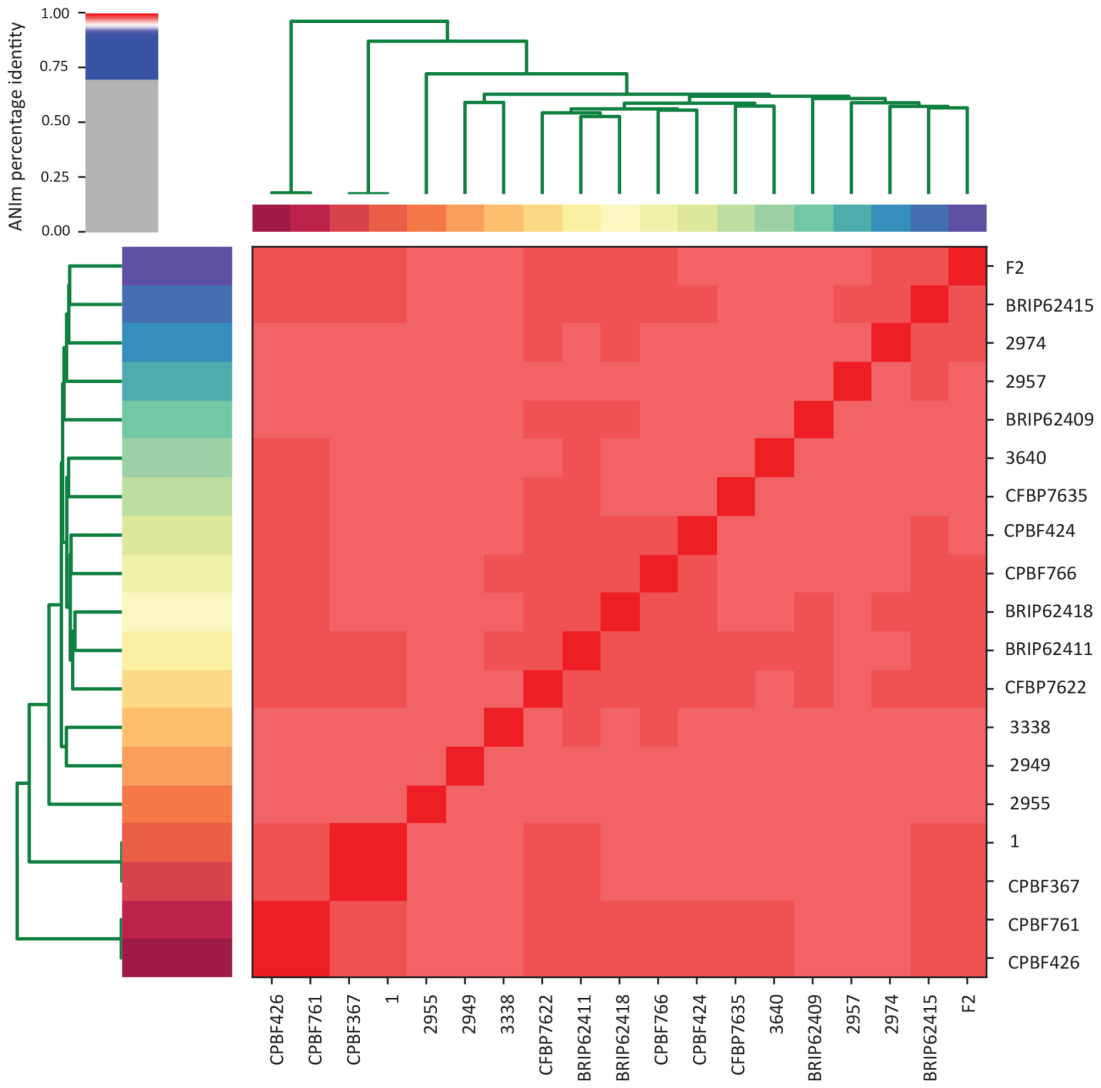

Figure S 4-3: Average nucleotide identity (ANI)-based heatmap showing the status of the representative *Xanthomonas euroxanthea* strains and strains from this study. The intensity of the color indicates the level of identity of all-versus-all genomes as depicted by the scale.

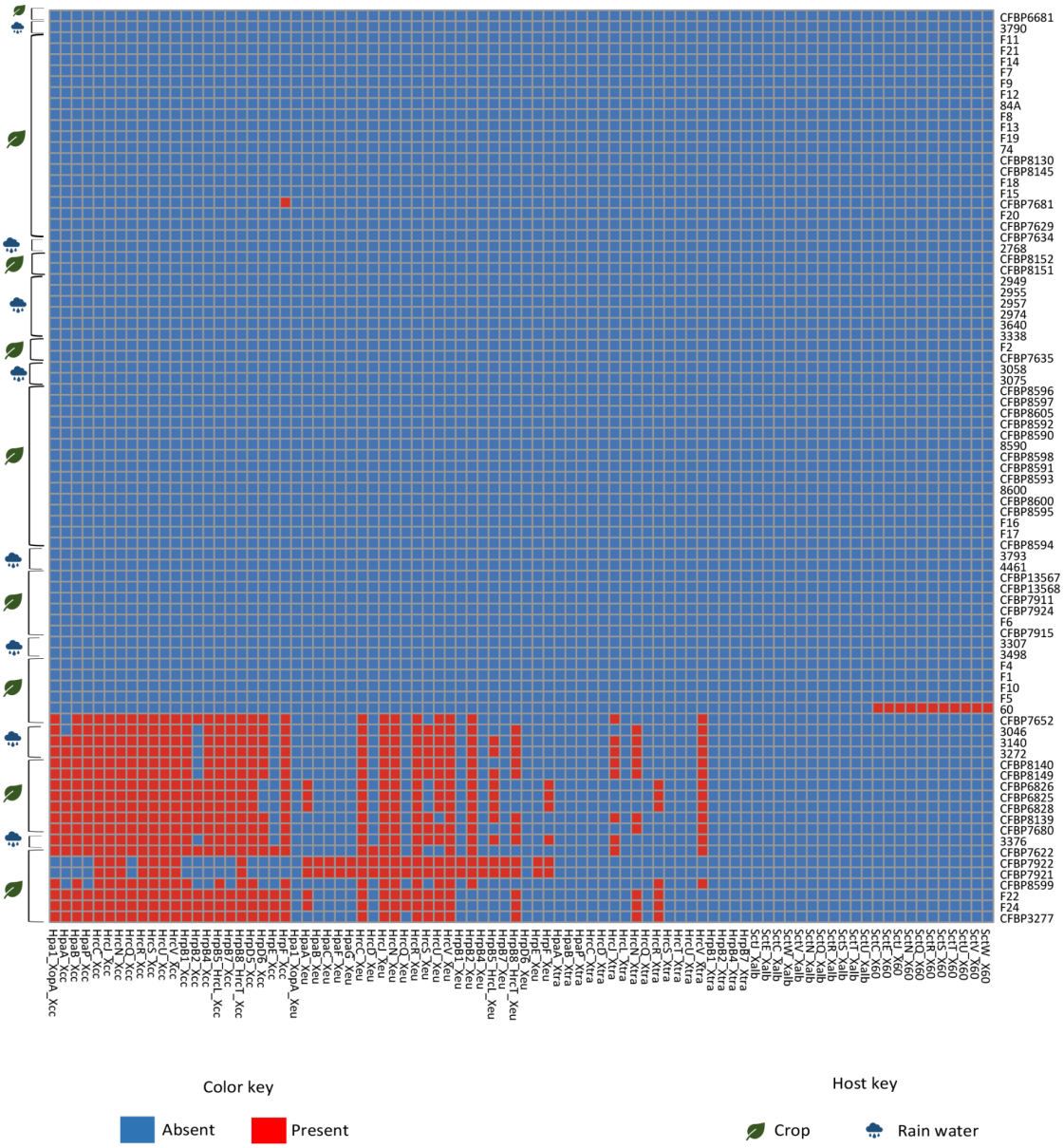

Figure S 4-4: Heatmap showing the status of T3SS genes in *Xanthomonas* strains used in this study. Here, query genes names from the different species are indicated as gene_*Xeu* (genes from *Xanthomonas campestris* pv. *vesicatoria* 85-10), gene_*Xcc* (genes from *Xanthomonas campestris* pv. *campestris* ATCC33913), gene_*Xtra* (genes from *Xanthomonas translucens* pv. *translucens* DSM18974), gene_*Xalb* (genes from *Xanthomonas albilineans* CFBP 2523), and gene_*X60* (genes from *Xanthomonas* sp. 60). Red color represents the presence while blue color represents the absence of the gene.

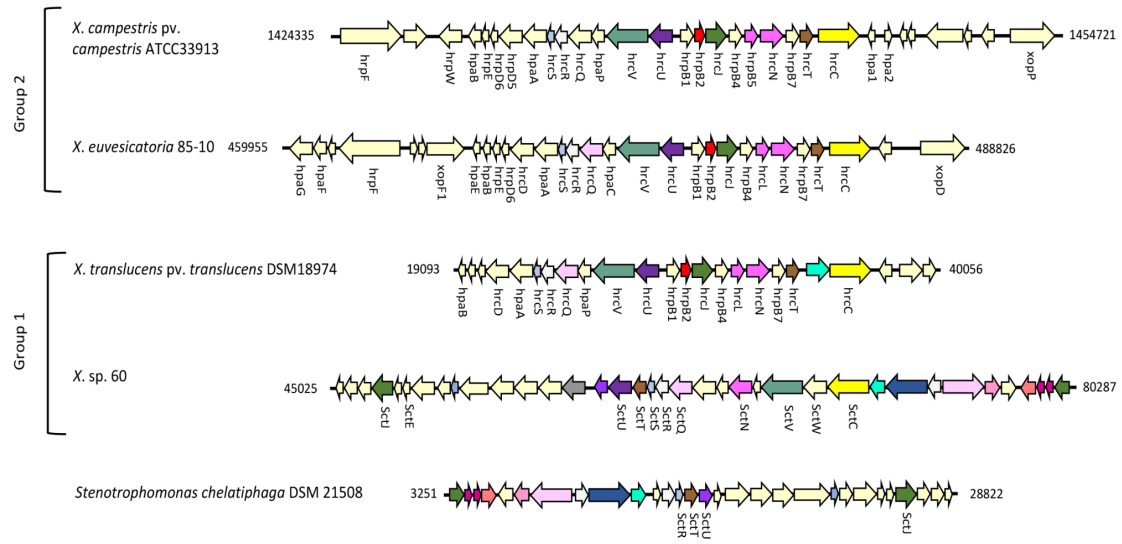

Figure S 4-5: Schematic representation of the genetic organization of T3SS clusters found in *Xanthomonas campestris* pv. *campestris* ATCC33913 (Xcc), *Xanthomonas campestris* pv. *vesicatoria* 85-10 (Xeu), *Xanthomonas translucens* pv. *translucens* DSM18974 (Xtr), *Xanthomonas* sp. 60, and *Stenotrophomonas chelatiphaga* DSM 21508. Genes of the same color (except light yellow, which has no COG assignment) are from the same orthologous group. The gene names are according to the Sct T3S injectisome nomenclature and annotation available in the IMG/M system (https://img.jgi.doe.gov/cgi-bin/m/main.cgi).

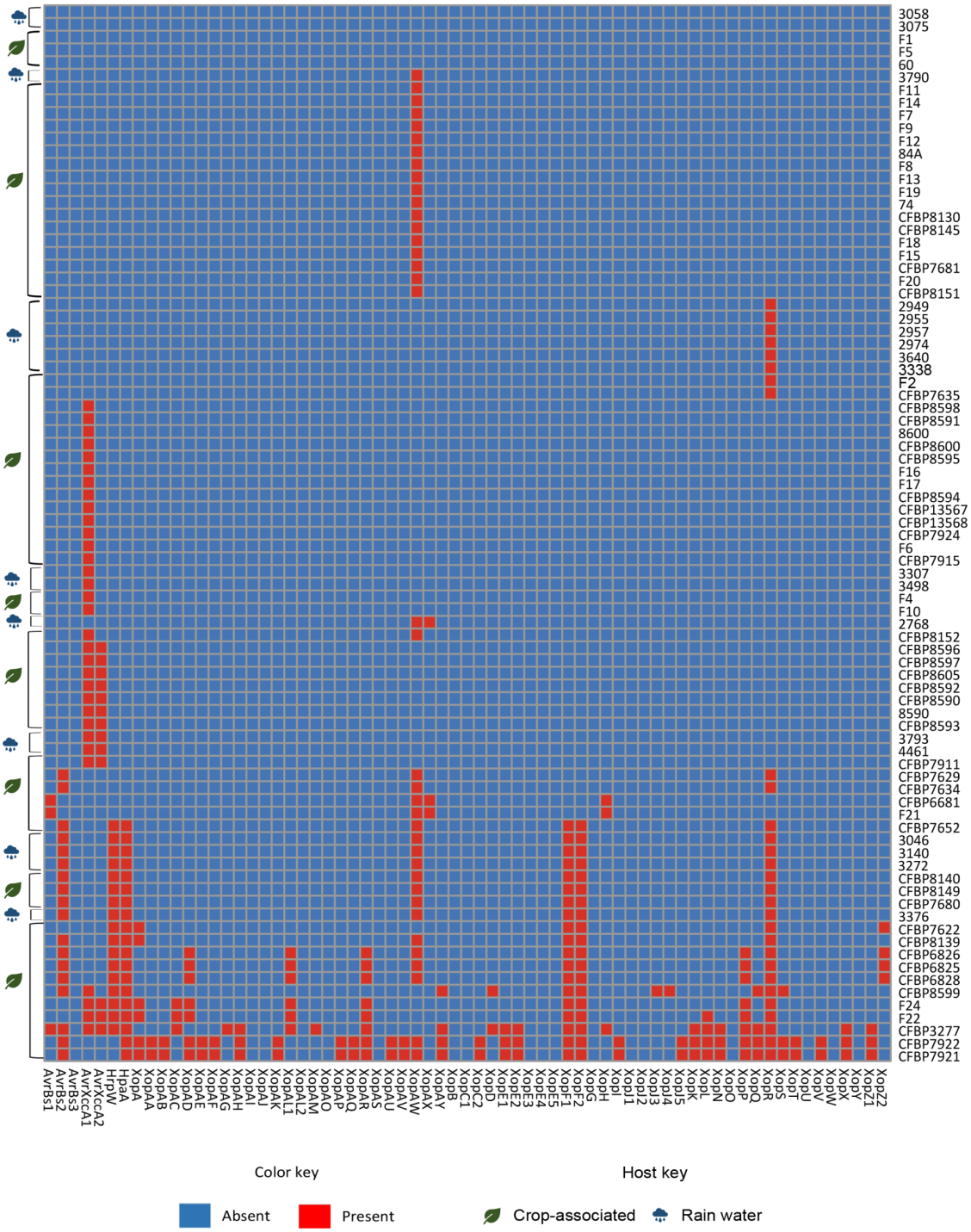

Figure S 4-6: Heatmap showing the status of T3E in *Xanthomonas* strains used in this study. Here the query represents all effectors families, putative effectors (AvrBs1, AvrBs2, AvrBs3, Xop, HpaA, HrpW, and AvrXccA), and their diversity. Red represents the presence, while blue represents the absence of the gene.

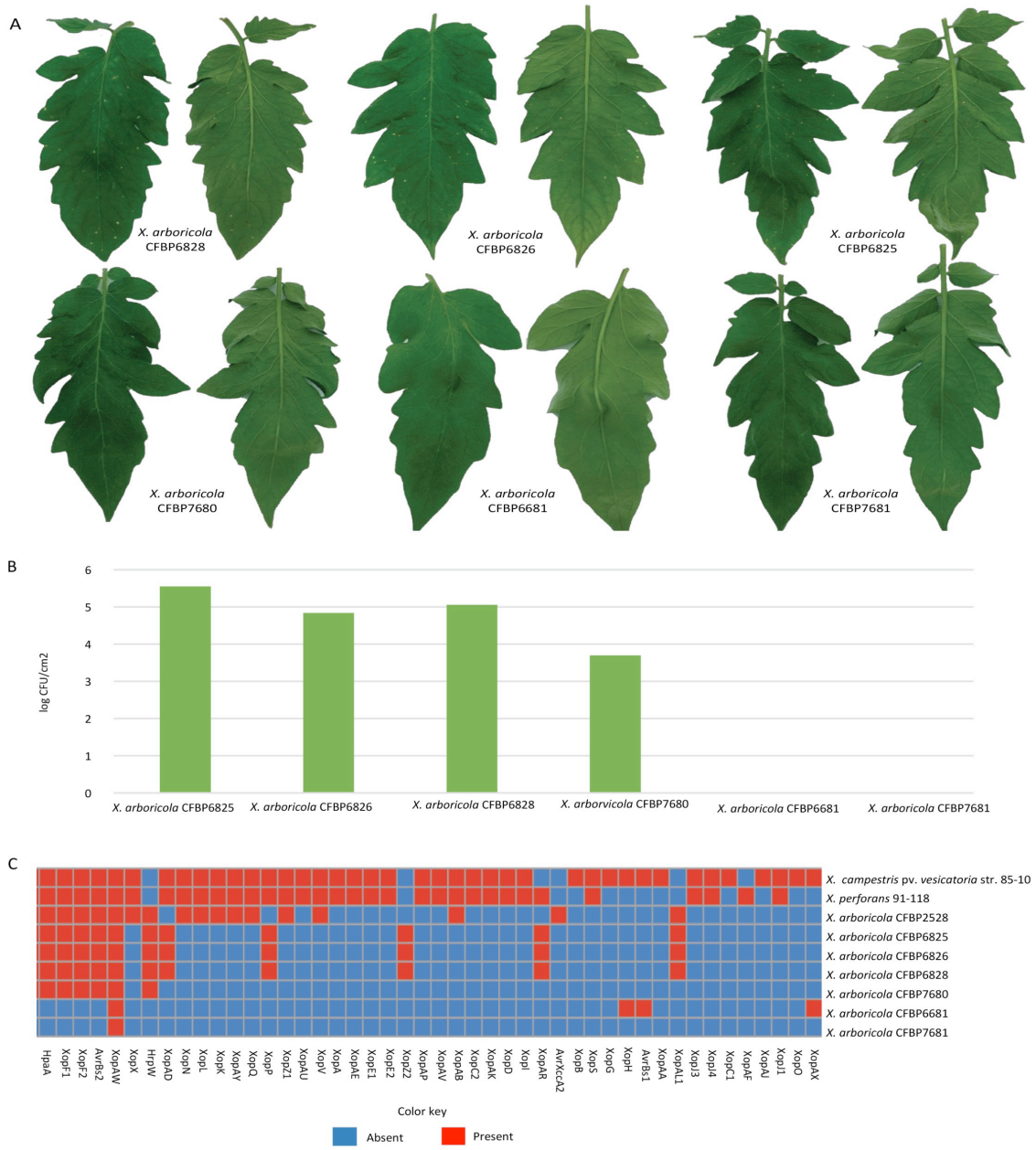

Figure S 4-7: (A) Pathogenicity and (B) *in-planta* population and (C) distribution of T3E in *X. arboricola* strains (CFBP 6825, CFBP 6826, CFBP 6828, CFBP 6681, CFBP 7681, and CFBP 7680) on 4-5-week-old tomato cv. FL 47R 10 DAI (days after inoculation). The presence of common T3Es across different *Xanthomonas* strains is shown in red, while blue represents the absence of T3Es.

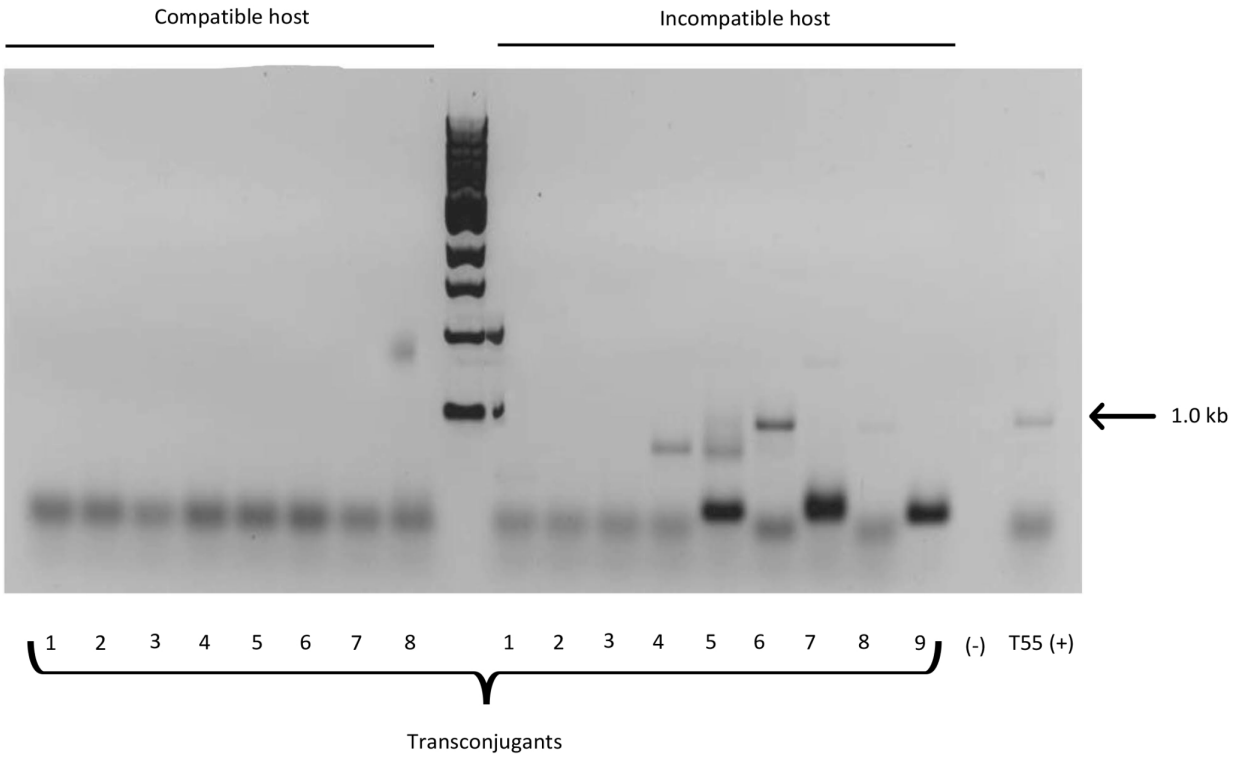

Figure S 4-8: Analysis of the presence of type three secreted effector avrBs1 with a transposon insertion among 50 randomly selected transconjugants. Mating experiments were conducted in both compatible (Capsicum annuum cv. Early California Wonder) and incompatible (Capsicum annuum cv. Early California Wonder 10R) backgrounds with *Xanthomonas euvesicatoria* strain 85-10 and *Xanthomonas* sp. strain T55. Each lane represents the pooled DNA of five to six individual transconjugants. The positive control (*Xanthomonas* sp. T55) shows an approximately 1.0 kb band indicative of avrBS1 with a transposon insertion, transconjugants isolated from the incompatible host in lanes six and eight have acquired the disrupted avirulence gene.

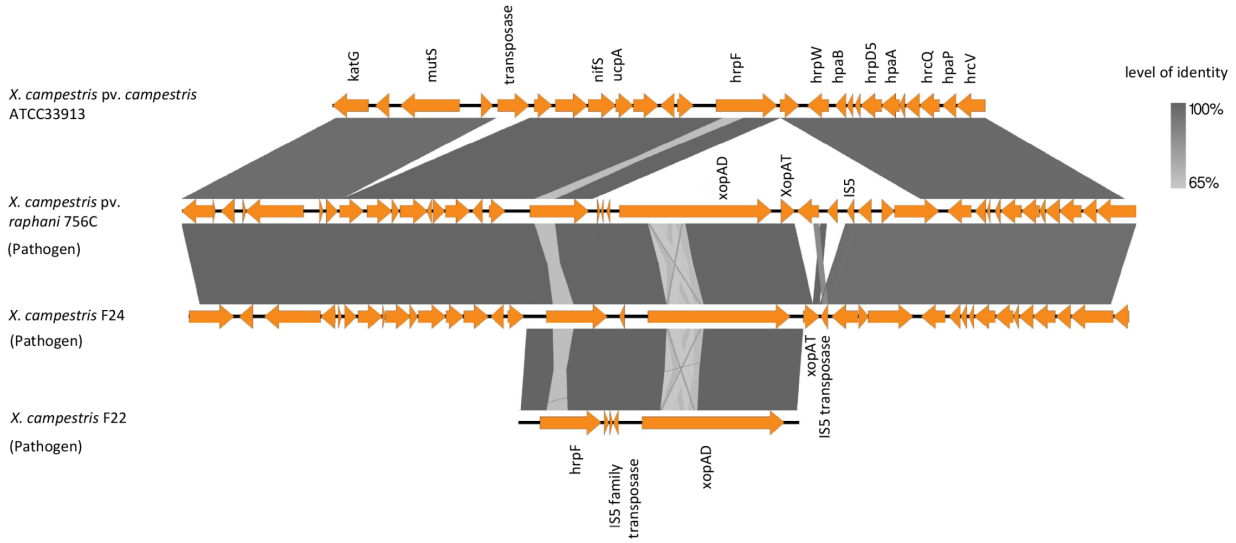

Figure S 4-9: Schematic representation and comparison of the genetic organization of the mobile genetic element island identified by MGEfinder in *Xanthomonas campestris* pv. *raphanin* 756C, *Xanthomonas campestris* F24, and *Xanthomonas campestris* F22, highlighting the XopAD effector flanked on both ends by IS5 transposases.

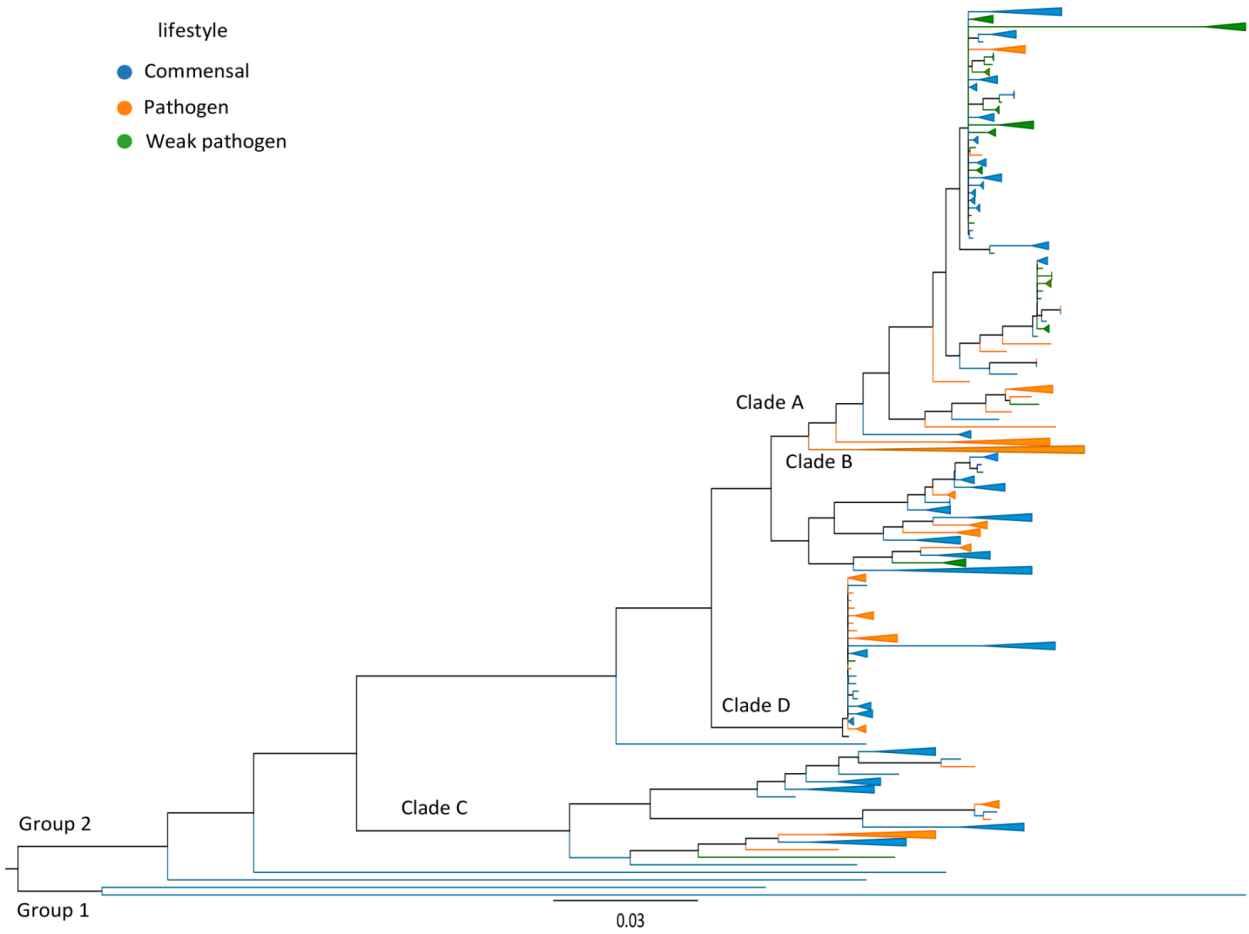

Figure S 4-10: Maximum-likelihood phylogeny based on the orthogroups of 337 *Xanthomonas* strains. Commensal *Xanthomonas* strains are highlighted in blue branches, weakly pathogenic strains in green branches, and pathogenic strains in orange. Branches were collapsed to visualize better strain diversity and phylogenetic placement with different lifestyles

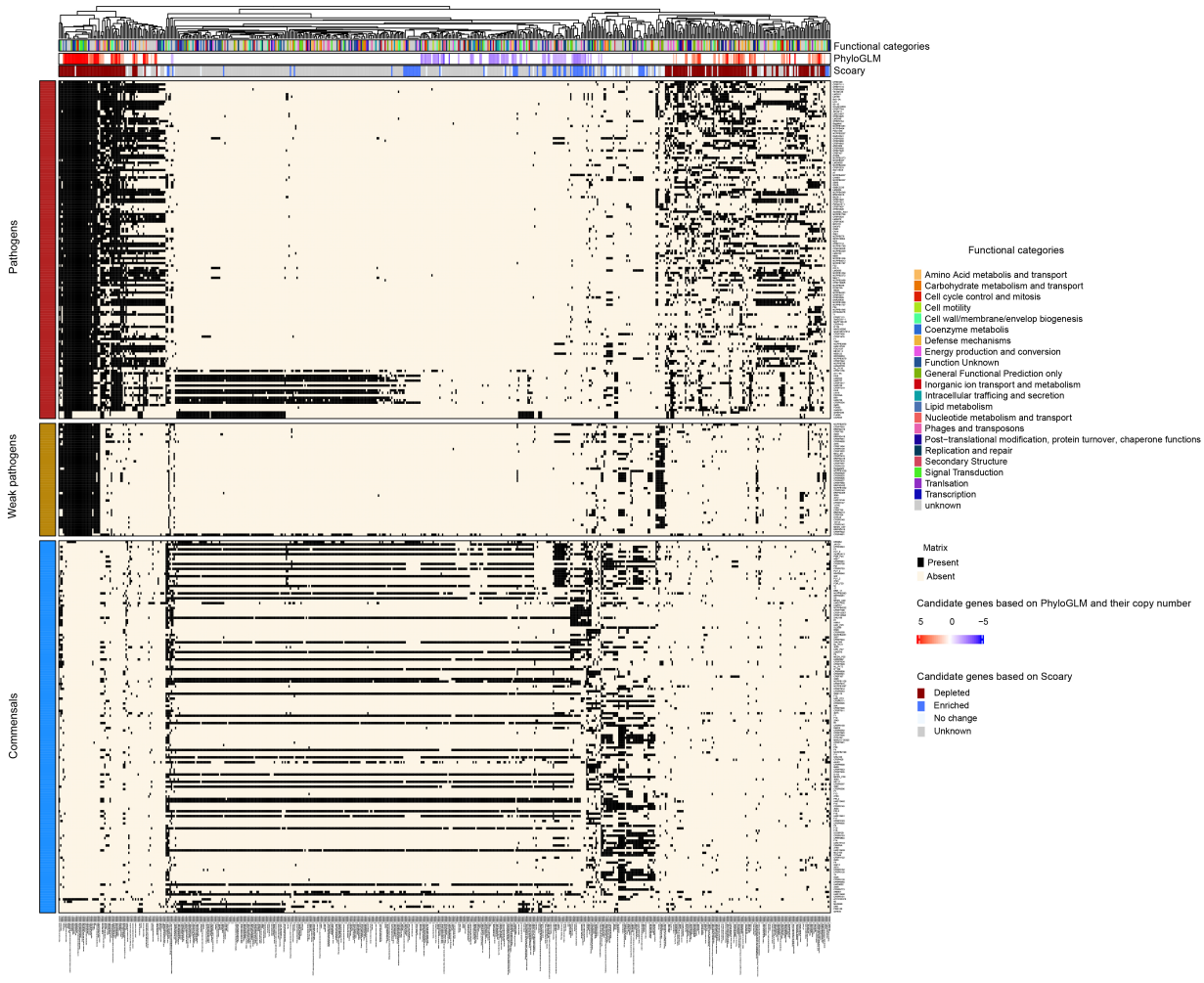

Figure S 4-11: A complex heatmap showing results from association analysis correlating pathogenic, weakly pathogenic, and commensal phenotypes to the orthologs identified using Orthofinder. The functional categories are indicated for each candidate gene identified based on the intersection of three methods PhyloGLM, Scoary, and hyperglm (gene presence/absence), and those identified by PhyloGLM based on gene copy number.

Table S 4-1: Details of *Xanthomonas* species sequenced in this work.

S.N.	Strain id	Taxonomic classification	Host of isolation	Country of isolation	Collected by	Genome size (bp)	GC content (%)	No. of tRNA	No. of rRNA	No. of gene	No. of CDS	No. of scaffolds	Genome coverage (x)	CheckM completeness (%)	CheckM contamination (%)	IMG Taxon ID	NCBI Bioproject Accession	NCBI Biosample Accession	Genbank ID
1	2768	Xanthomonas arboricola 2768	Rainwater	USA	Boris A Vinatzer	4,795,040	65.89	53	3	4129	4018	16	313	98.77	0.6	2828658931	PRJNA583330	SAMN13173792	JACHOE000000000
2	3046	Xanthomonas arboricola 3046	Rainwater	USA	Boris A Vinatzer	4,877,301	65.67	53	3	4207	4093	6	307	97.62	0.8	2828036968	PRJNA547322	SAMN12024095	JAASRI000000000
3	3140	Xanthomonas arboricola 3140	Rainwater	USA	Boris A Vinatzer	4,915,557	65.71	50	2	4183	4071	10	305	97.89	0.91	2828041176	PRJNA547323	SAMN12024092	JAASRH000000000
4	3272	Xanthomonas arboricola 3272	Rainwater	USA	Boris A Vinatzer	4,941,414	65.63	51	3	4210	4094	4	303	97.95	1.16	2828671728	PRJNA583334	SAMN13190211	JAATIV000000000
5	3376	Xanthomonas arboricola 3376	Rainwater	USA	Boris A Vinatzer	5,029,964	65.55	52	3	4355	4243	5	298	97.48	1.03	2828680123	PRJNA583336	SAMN13190212	JAATIW000000000
6	3790	Xanthomonas arboricola 3790	Rainwater	USA	Boris A Vinatzer	4,803,630	65.91	53	3	4084	3972	17	312	99.02	0.25	2828049435	PRJNA547325	SAMN12024080	JACICJ000000000
7	74	Xanthomonas arboricola 74	Cucurbit	USA	Neha Potnis	4,714,362	65.95	53	3	4061	3948	5	318	96.85	0.25	2828016237	PRJNA547318	SAMN12024071	JAASRL000000000
8	84A	Xanthomonas arboricola 84A	Sunflower	USA	Neha Potnis	4,898,161	65.82	54	3	4213	4102	7	306	99.02	0.75	2828020299	PRJNA547330	SAMN12024093	JACICG000000000
9	CFBP6681	Xanthomonas arboricola CFBP 6681	Tomato	Cuba	CFBP, Marie-Agnès Jacques	5,249,229	65.53	55	5	4611	4494	80	291	99.91	0.05	2828121528	PRJNA547335	SAMN12024074	JACICS000000000
10	CFBP6825	Xanthomonas arboricola CFBP 6825	Pepper	USA	CFBP, Marie-Agnès Jacques	4,947,835	65.69	52	3	4249	4130	3	300	99.89	0.24	2828126140	PRJNA547336	SAMN12024076	JAASRC000000000
11	CFBP6826	Xanthomonas arboricola CFBP 6826	Pepper	USA	CFBP, Marie-Agnès Jacques	4,946,363	65.69	52	3	4252	4133	5	300	99.56	0.24	2828130390	PRJNA547337	SAMN12024085	JACICT000000000
12	CFBP6828	Xanthomonas arboricola CFBP 6828	Pepper	USA	CFBP, Marie-Agnès Jacques	4,947,178	65.69	52	3	4254	4135	5	301	99.56	0.24	2828134643	PRJNA546814	SAMN12025056	JACICU000000000
13	CFBP7629	Xanthomonas arboricola CFBP 7629	Walnut	France	CFBP, Marie-Agnès Jacques	4,943,431	65.6	54	4	4288	4164	11	300	99.39	0.65	2828142967	PRJNA546816	SAMN12025187	JACICV000000000
14	CFBP7634	Xanthomonas arboricola CFBP 7634	Walnut	France	CFBP, Marie-Agnès Jacques	4,943,858	65.6	54	4	4293	4169	11	303	99.1	1.11	2828147256	PRJNA546817	SAMN12025210	JACICW000000000
15	CFBP7652	Xanthomonas arboricola CFBP 7652	Walnut	France	CFBP, Marie-Agnès Jacques	5,216,680	65.5	53	4	4511	4387	78	285	99.8	0.05	2828155758	PRJNA546819	SAMN12025058	JACHNJ000000000
16	CFBP7680	Xanthomonas arboricola CFBP 7680	Pepper	France	CFBP, Marie-Agnès Jacques	4,891,176	65.7	53	3	4209	4088	11	307	99.91	0.05	2828692919	PRJNA583339	SAMN13173797	JACJNK000000000
17	CFBP7681	Xanthomonas arboricola CFBP 7681	Pepper	China	CFBP, Marie-Agnès Jacques	4,963,904	66	54	3	4189	4072	19	302	98.84	0.54	2828697129	PRJNA583340	SAMN13172221	JACHOI000000000
18	CFBP8130	Xanthomonas arboricola CFBP 8130	Bean seed	France	CFBP, Marie-Agnès Jacques	4,754,280	65.99	53	3	4052	3940	13	315	98.77	0.5	2828173407	PRJNA547342	SAMN12024090	JACHNM000000000
19	CFBP8139	Xanthomonas arboricola CFBP 8139	Bean seed	N/A	CFBP, Marie-Agnès Jacques	4,953,125	65.74	54	3	4324	4207	8	303	97.81	0.65	2828177460	PRJNA547343	SAMN12024081	JAASRA000000000
20	CFBP8140	Xanthomonas arboricola CFBP 8140	Bean seed	N/A	CFBP, Marie-Agnès Jacques	4,830,954	65.73	53	3	4162	4039	9	310	99.72	0.75	2828270531	PRJNA546808	SAMN12025522	JACIIT000000000
21	CFBP8145	Xanthomonas arboricola CFBP 8145	Bean seed	N/A	CFBP, Marie-Agnès Jacques	4,834,499	65.89	53	3	4131	4019	11	310	99.91	0.05	2828181785	PRJNA546809	SAMN12025540	JAASSC000000000
22	CFBP8149	Xanthomonas arboricola CFBP 8149	Bean seed	N/A	CFBP, Marie-Agnès Jacques	4,878,272	65.69	51	3	4208	4095	7	307	99.72	0.5	2828185917	PRJNA546810	SAMN12025186	JACHNN000000000
23	F11	Xanthomonas arboricola F11	Citrus orange	USA	R. E. Stall and J. B. Jones	4,869,765	65.82	53	4	4163	4051	11	308	99.02	1.01	2828215148	PRJNA546658	SAMN12025545	JACHNQ000000000
24	F12	Xanthomonas arboricola F12	Tomato	USA	R. E. Stall and J. B. Jones	4,967,385	66.03	53	3	4218	4109	19	302	98.81	1.1	2828219312	PRJNA546659	SAMN12025199	JACHNR000000000
25	F13	Xanthomonas arboricola F13	Tomato	USA	R. E. Stall and J. B. Jones	4,808,302	65.76	55	3	4150	4034	4	312	96.74	0.65	2828722391	PRJNA583343	SAMN13173072	JAATIY000000000
26	F14	Xanthomonas arboricola F14	Citrus orange	USA	R. E. Stall and J. B. Jones	4,748,859	65.91	52	3	4108	3999	10	312	99.91	0.05	2828726542	PRJNA583344	SAMN13190100	JACHOM000000000
27	F15	Xanthomonas arboricola F15	Linaria sp.	USA	R. E. Stall and J. B. Jones	4,868,193	65.75	51	3	4172	4061	12	308	98.97	0.25	2828480307	PRJNA546660	SAMN12025150	JACIJV000000000
28	F18	Xanthomonas arboricola F18	Tomato	USA	R. E. Stall and J. B. Jones	4,821,314	65.89	51	3	4113	4003	14	311	98.72	0.05	2828227655	PRJNA546663	SAMN12025087	JACHNT000000000
29	F19	Xanthomonas arboricola F19	Pepper	USA	R. E. Stall and J. B. Jones	4,846,312	65.78	52	3	4208	4096	7	309	99.91	0.05	2828231769	PRJNA546664	SAMN12025543	JACJNJ000000000
30	F20	Xanthomonas arboricola F20	Tomato	USA	R. E. Stall and J. B. Jones	4,866,293	65.88	52	4	4188	4076	16	308	98.69	0.75	2828235978	PRJNA546665	SAMN12025125	JACIIP000000000
31	F21	Xanthomonas arboricola F21	Tomato	USA	R. E. Stall and J. B. Jones	5,253,819	65.53	55	5	4613	4496	75	285	98.77	0.99	2828240167	PRJNA546666	SAMN12025546	JACIIQ000000000
32	F7	Xanthomonas arboricola F7	Tomato	USA	R. E. Stall and J. B. Jones	4,763,808	65.86	53	3	4161	4040	12	313	99.91	0.05	2828714124	PRJNA583349	SAMN13172344	JAATLO000000000
33	F8	Xanthomonas arboricola F8	Tomato	USA	R. E. Stall and J. B. Jones	4,807,951	65.78	54	3	4104	3993	4	312	97.06	0.06	2828718286	PRJNA583350	SAMN13172246	JAATIK000000000
34	F9	Xanthomonas arboricola F9	Tomato	USA	R. E. Stall and J. B. Jones	4,802,478	65.83	53	3	4094	3981	9	312	99.91	0.06	2833698404	PRJNA546669	SAMN12025045	JACCEJ000000000
35	CFBP13567	Xanthomonas campestris CFBP 13567	Radish seed	France	CFBP, Marie-Agnès Jacques	5,040,092	65.44	54	3	4383	4206	9	297	99.64	0.24	2828488598	PRJNA547331	SAMN12024066	JANUPS000000000
36	CFBP13568	Xanthomonas campestris CFBP 13568	Radish seed	France	CFBP, Marie-Agnès Jacques	5,040,178	65.44	54	3	4385	4207	9	297	99.64	0.24	2828274694	PRJNA547333	SAMN12024089	JANUPR000000000
37	CFBP3277	Xanthomonas campestris CFBP 3277	Tomato	Guadeloupe (French west indies)	CFBP, Marie-Agnès Jacques	5,002,720	65.11	54	3	4444	4263	84	300	99.64	0.05	2828117083	PRJNA547334	SAMN12024072	JACICR000000000
38	CFBP7911	Xanthomonas campestris CFBP 7911	Bean seed	N/A	CFBP, Marie-Agnès Jacques	4,885,581	65.47	54	3	4282	4112	7	304	99.64	0.05	2828160270	PRJNA547338	SAMN12024075	JANUPO000000000
39	CFBP7915	Xanthomonas campestris CFBP 7915	Bean seed	China	CFBP, Marie-Agnès Jacques	4,889,592	65.43	56	3	4322	4147	22	307	98.07	2.45	2828701319	PRJNA583341	SAMN13172233	JANTYW000000000
40	CFBP7924	Xanthomonas campestris CFBP 7924	Bean seed	N/A	CFBP, Marie-Agnès Jacques	4,977,629	65.35	53	3	4368	4190	9	301	99.64	0.05	2828266162	PRJNA547341	SAMN12024073	JANUPQ000000000
41	F22	Xanthomonas campestris F22	Tomato	USA	R. E. Stall and J. B. Jones	4,959,759	65.14	54	3	4394	4227	33	299	99.64	0.24	2828730651	PRJNA583346	SAMN13172197	
42	F24	Xanthomonas campestris F24	Pepper	USA	R. E. Stall and J. B. Jones	4,918,247	65.36	54	3	4344	4172	38	305	99.64	0.24	2828735046	PRJNA583347	SAMN13172198	JANUPT000000000
43	F6	Xanthomonas campestris F6	Tomao	USA	R. E. Stall and J. B. Jones	5,029,390	65.42	54	3	4380	4207	23	295	99.68	0.05	2828206614	PRJNA546668	SAMN12025061	JANUPP000000000
44	8590	Xanthomonas cannabis 8590	Bean leaf	Brazil	CFBP, Marie-Agnès Jacques	4,913,203	65.75	55	3	4276	4166	8	302	99.8	0.41	2828062588	PRJNA547328	SAMN12024094	JAASRG000000000
45	8600	Xanthomonas cannabis 8600	Bean leaf	Brazil	CFBP, Marie-Agnès Jacques	4,795,610	65.74	53	3	4145	4036	17	311	99.73	0.05	2828104512	PRJNA547329	SAMN12024084	JACICP000000000
46	CFBP8590	Xanthomonas cannabis CFBP 8590	Bean leaf	Brazil	CFBP, Marie-Agnès Jacques	4,914,203	65.75	55	3	4276	4166	8	305	97.52	0.77	2828108658	PRJNA546813	SAMN12025062	JAASRD000000000
47	CFBP8591	Xanthomonas cannabis CFBP 8591	Bean leaf	Brazil	CFBP, Marie-Agnès Jacques	4,853,759	65.71	53	4	4224	4114	18	309	97.52	0.57	2828066865	PRJNA546648	SAMN12025170	JACICK000000000
48	CFBP8592	Xanthomonas cannabis CFBP 8592	Bean leaf	Brazil	CFBP, Marie-Agnès Jacques	4,838,975	65.74	53	3	4182	4073	22	310	97.48	0.52	2828071090	PRJNA546649	SAMN12025542	JAASRF000000000
49	CFBP8593	Xanthomonas cannabis CFBP 8593	Bean leaf	Brazil	CFBP, Marie-Agnès Jacques	4,839,134	65.75	55	3	4203	4091	18	307	99.87	0.05	2828075273	PRJNA546650	SAMN12025064	JAASRE000000000
50	CFBP8594	Xanthomonas cannabis CFBP 8594	Bean leaf	Brazil	CFBP, Marie-Agnès Jacques	4,684,449	65.84	52	3	4058	3946	8	317	99.75	0.05	2828079477	PRJNA546651	SAMN12025075	JACICL000000000
51	CFBP8595	Xanthomonas cannabis CFBP 8595	Bean leaf	Brazil	CFBP, Marie-Agnès Jacques	4,819,314	65.78	53	3	4164	4058	19	311	97.59	0.67	2828083536	PRJNA546652	SAMN12025144	JACICM000000000
52	CFBP8596	Xanthomonas cannabis CFBP 8596	Bean leaf	Brazil	CFBP, Marie-Agnès Jacques	4,814,602	65.77	52	3	4139	4032	16	308	99.87	0.05	2828087701	PRJNA546653	SAMN12025171	JACBZB000000000
53	CFBP8597	Xanthomonas cannabis CFBP 8597	Bean leaf	Brazil	CFBP, Marie-Agnès Jacques	4,814,889	65.77	52	3	4144	4037	15	308	99.71	0.05	2828091841	PRJNA546654	SAMN12025123	JACBZC000000000
54	CFBP8598	Xanthomonas cannabis CFBP 8598	Bean leaf	Brazil	CFBP, Marie-Agnès Jacques	4,853,479	65.71	53	4	4229	4119	15	306	99.76	0.05	2828095986	PRJNA546655	SAMN12025188	JACICN000000000
55	CFBP8599	Xanthomonas cannabis CFBP 8599	Bean leaf	Brazil	CFBP, Marie-Agnès Jacques	4,953,994	65.51	54	3	4295	4184	18	301	99.78	0.17	2828100216	PRJNA546656	SAMN12025100	JACICO000000000
56	CFBP8600	Xanthomonas cannabis CFBP 8600	Bean leaf	Brazil	CFBP, Marie-Agnès Jacques	4,795,634	65.74	53	3	4147	4038	19	310	99.71	0.05	2828112935	PRJNA546806	SAMN12025110	JACICQ000000000
57	CFBP8605	Xanthomonas cannabis CFBP 8605	Bean leaf	Brazil	CFBP, Marie-Agnès Jacques	4,796,975	65.74	53	3	4103	3992	25	313	97.48	0.52	2828198290	PRJNA546807	SAMN12025111	JAASZY000000000
58	F16	Xanthomonas cannabis F16	Nightshade (*Solanum nigrum*)	USA	R. E. Stall and J. B. Jones	4,776,640	65.75	54	3	4123	4011	18	314	97.48	0.27	2828223531	PRJNA546661	SAMN12025151	JACHNS000000000
59	F17	Xanthomonas cannabis F17	Tomato	USA	R. E. Stall and J. B. Jones	4,778,094	65.75	54	3	4117	4005	19	314	97.53	0.27	2828484480	PRJNA546662	SAMN12025124	JACIJW000000000
60	2949	Xanthomonas euroxanthea 2949	Rainwater	USA	Boris A Vinatzer	4,824,595	65.87	52	3	4211	4101	7	311	96.76	0.4	2828024513	PRJNA547319	SAMN12024087	JAASRK000000000
61	2955	Xanthomonas euroxanthea 2955	Rainwater	USA	Boris A Vinatzer	4,767,670	65.88	53	3	4130	4020	7	314	96.94	0.15	2828028725	PRJNA547320	SAMN12024091	JAASRJ000000000
62	2957	Xanthomonas euroxanthea 2957	Rainwater	USA	Boris A Vinatzer	4,970,240	65.85	55	3	4346	4232	8	302	96.45	0.41	2828663061	PRJNA583331	SAMN13173081	JAATIU000000000
63	2974	Xanthomonas euroxanthea 2974	Rainwater	USA	Boris A Vinatzer	4,799,859	65.98	53	3	4111	4002	20	312	96.45	0.4	2828032856	PRJNA547321	SAMN12024078	JACICH000000000
64	3338	Xanthomonas euroxanthea 3338	Rainwater	USA	Boris A Vinatzer	4,718,025	65.95	51	3	4074	3964	11	318	96.83	0.06	2828045360	PRJNA547324	SAMN12024079	JACICI000000000
65	3640	Xanthomonas euroxanthea 3640	Rainwater	USA	Boris A Vinatzer	4,836,250	66.04	52	3	4168	4062	13	310	97.24	0.77	2828688750	PRJNA583338	SAMN13173795	JACHOI000000000
66	CFBP7622	Xanthomonas euroxanthea CFBP 7622	Bean	USA	CFBP, Marie-Agnès Jacques	4,705,769	66.21	53	3	4068	3963	6	319	97.4	0.13	2828138898	PRJNA546815	SAMN12025169	JAASRB000000000
67	CFBP7635	Xanthomonas euroxanthea CFBP 7635	Walnut (cv. Franquette)	France	CFBP, Marie-Agnès Jacques	4,857,910	66.15	49	3	4207	4100	10	306	99.67	0.05	2828151550	PRJNA546818	SAMN12025097	JACHNI000000000
68	F2	Xanthomonas euroxanthea F2	Peruvian groundcherry (*Physalis peruviana)*	N/A	R. E. Stall and J. B. Jones	4,885,591	65.98	52	3	4232	4122	10	304	99.57	0.23	2828709891	PRJNA583345	SAMN13172987	JACHOL000000000
69	CFBP7921	Xanthomonas euvesicatoria CFBP 7921	Bean seed	N/A	CFBP, Marie-Agnès Jacques	5,035,348	64.89	56	3	4426	4297	21	298	99.78	0.43	2828164553	PRJNA547339	SAMN12024088	JACHNK000000000
70	CFBP7922	Xanthomonas euvesicatoria CFBP 7922	Bean seed	N/A	CFBP, Marie-Agnès Jacques	5,034,570	64.89	56	3	4426	4297	21	298	98.22	0.43	2828168980	PRJNA547340	SAMN12024083	JACHNL000000000
71	3058	Xanthomonas sp. 3058	Rainwater	USA	Boris A Vinatzer	4,968,012	64.74	54	3	4319	4209	60	302	97.09	0.37	2828667408	PRJNA583332	SAMN13173779	JACHOF000000000
72	3075	Xanthomonas sp. 3075	Rainwater	USA	Boris A Vinatzer	4,997,784	64.6	52	3	4294	4185	61	300	97.17	0.82	2831322938	PRJNA583333	SAMN13172773	JACIFI000000000
73	3307	Xanthomonas sp. 3307	Rainwater	USA	Boris A Vinatzer	4,860,261	68.94	51	3	4183	4099	20	308	95.43	3.44	2828675939	PRJNA583335	SAMN13172743	JACHOG000000000
74	3498	Xanthomonas sp. 3498	Rainwater	USA	Boris A Vinatzer	4,925,041	68.91	55	3	4270	4182	13	304	95.57	3.73	2828684479	PRJNA583337	SAMN13190338	JACHOH000000000
75	3793	Xanthomonas sp. 3793	Rainwater	USA	Boris A Vinatzer	5,210,011	65.26	55	3	4513	4379	14	288	97.67	0.6	2828053520	PRJNA547326	SAMN12024082	JANTYU000000000
76	4461	Xanthomonas sp. 4461	Rainwater	USA	Boris A Vinatzer	5,263,960	65.29	55	3	4553	4423	12	285	98.05	1.03	2828058034	PRJNA547327	SAMN12024077	JANTYV000000000
77	CFBP8151	Xanthomonas sp. CFBP 8151	Bean seed	N/A	CFBP, Marie-Agnès Jacques	4,610,370	65.59	52	3	4036	3928	9	325	96.29	1.31	2828190126	PRJNA546811	SAMN12025089	JAASQZ000000000
78	CFBP8152	Xanthomonas sp. CFBP 8152	Bean seed	N/A	CFBP, Marie-Agnès Jacques	4,705,389	65.53	53	3	4126	4008	15	319	96.69	0.56	2828194163	PRJNA546812	SAMN12025148	JACHNO000000000
79	F1	Xanthomonas sp. F1	Citrus orange	USA	R. E. Stall and J. B. Jones	4,955,224	68.93	53	3	4248	4159	24	300	99.96	0.74	2828705642	PRJNA583342	SAMN13190379	JACHOK000000000
80	F10	Xanthomonas sp. F10	Citrus orange	USA	R. E. Stall and J. B. Jones	4,776,209	69.31	51	3	4152	4060	6	314	95.45	3.47	2828210995	PRJNA546657	SAMN12025141	JACHNP000000000
81	F4	Xanthomonas sp. F4	Citrus orange	USA	R. E. Stall and J. B. Jones	4,895,058	68.81	51	4	4219	4132	13	303	99.54	0.34	2828202394	PRJNA546667	SAMN12024050	JAASQX000000000
82	F5	Xanthomonas sp. F5	Pepper	USA	R. E. Stall and J. B. Jones	5,317,945	68.59	55	3	4621	4535	17	279	99.62	0.5	2833702499	PRJNA583348	SAMN13172314	JACIFN000000000
83	60	Xanthomonas sp. 60	Tomato	USA	Neha Potnis	3,640,246	67.3	65	4	3318	3223	13	408	99.74	0.39	2828262843	PRJNA547317	SAMN12024202	JAASQW000000000

Table S 4-2: *Xanthomonas* strains sequenced for this work and representative/type strain from NCBI used for comparative genomic analysis

S.N.	Strain ID	Genome source	Taxonomic classification	Cluster id for MGE finder
1	CFBP6681	This study	Xanthomonas arboricola CFBP 6681	8
2	3790	This study	Xanthomonas arboricola 3790	8
3	F11	This study	Xanthomonas arboricola F11	8
4	F21	This study	Xanthomonas arboricola F21	8
5	F14	This study	Xanthomonas arboricola F14	8
6	F7	This study	Xanthomonas arboricola F7	8
7	F9	This study	Xanthomonas arboricola F9	8
8	F12	This study	Xanthomonas arboricola F12	8
9	84A	This study	Xanthomonas arboricola 84A	8
10	F8	This study	Xanthomonas arboricola F8	8
11	F13	This study	Xanthomonas arboricola F13	8
12	F19	This study	Xanthomonas arboricola F19	8
13	74	This study	Xanthomonas arboricola 74	8
14	CFBP8130	This study	Xanthomonas arboricola CFBP 8130	8
15	CFBP8145	This study	Xanthomonas arboricola CFBP 8145	8
16	F18	This study	Xanthomonas arboricola F18	8
17	F15	This study	Xanthomonas arboricola F15	8
18	CFBP7681	This study	Xanthomonas arboricola CFBP 7681	8

S.N.	Strain ID	Genome source	Taxonomic classification	Cluster id for MGE finder
19	F20	This study	Xanthomonas arboricola F20	8
20	CFBP7629	This study	Xanthomonas arboricola CFBP 7629	8
21	CFBP7634	This study	Xanthomonas arboricola CFBP 7634	8
22	CFBP7652	This study	Xanthomonas arboricola CFBP 7652	8
23	2768	This study	Xanthomonas arboricola 2768	8
24	3046	This study	Xanthomonas arboricola 3046	8
25	3140	This study	Xanthomonas arboricola 3140	8
26	3272	This study	Xanthomonas arboricola 3272	8
27	CFBP8140	This study	Xanthomonas arboricola CFBP 8140	8
28	CFBP8149	This study	Xanthomonas arboricola CFBP 8149	8
29	CFBP6826	This study	Xanthomonas arboricola CFBP 6826	8
30	CFBP6825	This study	Xanthomonas arboricola CFBP 6825	8
31	CFBP6828	This study	Xanthomonas arboricola CFBP 6828	8
32	CFBP8139	This study	Xanthomonas arboricola CFBP 8139	8
33	CFBP7680	This study	Xanthomonas arboricola CFBP 7680	8
34	3376	This study	Xanthomonas arboricola 3376	8
35	CFBP2528	Type/Representative	Xanthomonas arboricola CFBP 2528	8
36	CFBP8152	This study	Xanthomonas sp. CFBP 8152	8
37	CFBP8151	This study	Xanthomonas sp. CFBP 8151	8
38	2949	This study	Xanthomonas euroxanthea 2949	

S.N.	Strain ID	Genome source	Taxonomic classification	Cluster id for MGE finder
39	2955	This study	Xanthomonas euroxanthea 2955	8
40	2957	This study	Xanthomonas euroxanthea 2957	
41	2974	This study	Xanthomonas euroxanthea 2974	8
42	3640	This study	Xanthomonas euroxanthea 3640	8
43	F2	This study	Xanthomonas euroxanthea F2	8
44	CFBP7622	This study	Xanthomonas euroxanthea CFBP 7622	8
45	3338	This study	Xanthomonas euroxanthea 3338	8
46	CFBP7635	This study	Xanthomonas euroxanthea CFBP 7635	8
47	BRIP62409	Type/Representative	Xanthomonas euroxanthea BRIP62409	
48	BRIP62411	Type/Representative	Xanthomonas euroxanthea BRIP62411	
49	BRIP62415	Type/Representative	Xanthomonas euroxanthea BRIP62415	
50	BRIP62418	Type/Representative	Xanthomonas euroxanthea BRIP62418	
51	WHRI7744	Type/Representative	Xanthomonas hortorum WHRI7744	
52	CFBP4925	Type/Representative	Xanthomonas hortorum CFBP 4925	
53	M081	Type/Representative	Xanthomonas hortorum M081	
54	CFBP4188	Type/Representative	Xanthomonas cynarae CFBP 4188	
55	ATCC19865	Type/Representative	Xanthomonas gardneri ATCC19865	
56	ICMP7383	Type/Representative	Xanthomonas gardneri ICMP7383	
57	CFBP1817	Type/Representative	Xanthomonas populi CFBP 1817	
58	3058	This study	Xanthomonas sp. 3058	8

S.N.	Strain ID	Genome source	Taxonomic classification	Cluster id for MGE finder
59	3075	This study	Xanthomonas sp. 3075	8
60	PD885	Type/Representative	Xanthomonas fragariae PD885	
61	91-118	Type/Representative	Xanthomonas perforans 91-118	
62	CFBP7293	Type/Representative	Xanthomonas perforans CFBP 7293	
63	GEV07	Type/Representative	Xanthomonas alfalfae GEVRose07	1
64	LMG27970	Type/Representative	Xanthomonas euvesicatoria LMG27970	
65	85-10	Type/Representative	Xanthomonas euvesicatoria 85-10	
66	BRIP39016	Type/Representative	Xanthomonas euvesicatoria BRIP39016	
67	CFBP7922	This study	Xanthomonas euvesicatoria CFBP 7922	1
68	CFBP7921	This study	Xanthomonas euvesicatoria CFBP 7921	1
69	LMG695	Type/Representative	Xanthomonas phaseoli LMG695	
70	CFBP412	Type/Representative	Xanthomonas phaseoli CFBP 412	
71	LMG9322	Type/Representative	Xanthomonas citri LMG9322	
72	DAR33341	Type/Representative	Xanthomonas citri DAR33341	
73	DSM3585	Type/Representative	Xanthomonas axonopodis DSM3585	
74	NCPPB2417	Type/Representative	Xanthomonas vasicola NCPPB2417	
75	CFBP2543	Type/Representative	Xanthomonas vasicola CFBP 2543	
76	CFBP7342	Type/Representative	Xanthomonas oryzae CFBP 7342	
77	ATCC35933	Type/Representative	Xanthomonas oryzae ATCC35933	
78	CFBP8353	Type/Representative	Xanthomonas prunicola CFBP 8353	

S.N.	Strain ID	Genome source	Taxonomic classification	Cluster id for MGE finder
79	CFBP1976	Type/Representative	Xanthomonas bromi CFBP 1976	
80	WHRI8853	Type/Representative	Xanthomonas nasturtii WHRI8853	
81	CFBP8596	This study	Xanthomonas cannabis CFBP 8596	3
82	CFBP8597	This study	Xanthomonas cannabis CFBP 8597	3
83	CFBP8605	This study	Xanthomonas cannabis CFBP 8605	3
84	CFBP8592	This study	Xanthomonas cannabis CFBP 8592	3
85	CFBP8590	This study	Xanthomonas cannabis CFBP 8590	3
86	8590	This study	Xanthomonas cannabis 8590	3
87	CFBP8599	This study	Xanthomonas cannabis CFBP 8599	3
88	CFBP8598	This study	Xanthomonas cannabis CFBP 8598	3
89	CFBP8591	This study	Xanthomonas cannabis CFBP 8591	3
90	CFBP8593	This study	Xanthomonas cannabis CFBP 8593	3
91	8600	This study	Xanthomonas cannabis 8600	3
92	CFBP8600	This study	Xanthomonas cannabis CFBP 8600	3
93	CFBP8595	This study	Xanthomonas cannabis CFBP 8595	3
94	F16	This study	Xanthomonas cannabis F16	3
95	F17	This study	Xanthomonas cannabis F17	3
96	NCPPB3753	Type/Representative	Xanthomonas cannabis NCPPB3753	3
97	CFBP8594	This study	Xanthomonas cannabis CFBP 8594	3
98	DSM18956	Type/Representative	Xanthomonas pisi DSM18956	

S.N.	Strain ID	Genome source	Taxonomic classification	Cluster id for MGE finder
99	CFBP4643	Type/Representative	Xanthomonas pisi CFBP 4643	2
100	3793	This study	Xanthomonas sp. 3793	2
101	4461	This study	Xanthomonas sp. 4461	2
102	LMG911	Type/Representative	Xanthomonas vesicatoria LMG911	
103	CFBP7245	Type/Representative	Xanthomonas dyei CFBP 7245	
104	CFBP4642	Type/Representative	Xanthomonas cassavae CFBP 4642	
105	WHRI8848	Type/Representative	Xanthomonas floridensis WHRI8848	
106	CFBP4690	Type/Representative	Xanthomonas codiaei CFBP 4690	
107	CFBP2542	Type/Representative	Xanthomonas cucurbitae CFBP 2542	
108	CFBP4644	Type/Representative	Xanthomonas melonis CFBP 4644	
109	CFBP13567	This study	Xanthomonas campestris CFBP 13567	7
110	CFBP13568	This study	Xanthomonas campestris CFBP 13568	7
111	CFBP7911	This study	Xanthomonas campestris CFBP 7911	7
112	CFBP7924	This study	Xanthomonas campestris CFBP 7924	7
113	F6	This study	Xanthomonas campestris F6	7
114	CFBP7915	This study	Xanthomonas campestris CFBP 7915	7
115	F22	This study	Xanthomonas campestris F22	7
116	F24	This study	Xanthomonas campestris F24	7
117	CFBP3277	This study	Xanthomonas campestris CFBP 3277	7
118	ATCC33913	Type/Representative	Xanthomonas campestris ATCC33913	7

S.N.	Strain ID	Genome source	Taxonomic classification	Cluster id for MGE finder
119	LMG27592	Type/Representative	Xanthomonas maliensis LMG27592	
120	M97	Type/Representative	Xanthomonas maliensis M97	
121	3307	This study	Xanthomonas sp. 3307	6
122	3498	This study	Xanthomonas sp. 3498	6
123	F4	This study	Xanthomonas sp. F4	6
124	F1	This study	Xanthomonas sp. F1	6
125	PPL1	Type/Representative	Xanthomonas sontii PPL1	
126	CFBP4641	Type/Representative	Xanthomonas sacchari CFBP 4641	6
127	F10	This study	Xanthomonas sp. F10	6
128	CFBP2523	Type/Representative	Xanthomonas albilineans CFBP 2523	
129	DSM18974	Type/Representative	Xanthomonas translucens DSM18974	5
130	CFBP1156	Type/Representative	Xanthomonas hyacinthi CFBP 1156	
131	F5	This study	Xanthomonas sp. F5	
132	CFBP4691	Type/Representative	Xanthomonas theicola CFBP 4691	5
133	AL60	This study	Xanthomonas sp. 60	4
134	Sp953	Type/Representative	Xanthomonas retroflexus Sp953	4

•

Table S 4-3: Taxonomic classification, lifestyle, and variation in flagellin encoding gene in *Xanthomonas* genomes used for the association analysis.

Strain ID	Taxonomic classification	Lifestyle	flg22-43rd amino acid residue
17	Xanthomonas arboricola 17	Commensal	D
74	Xanthomonas arboricola 74	Commensal	D
2768	Xanthomonas arboricola 2768	Commensal	D
2949	Xanthomonas euroxanthea 2949	Commensal	D
2955	Xanthomonas euroxanthea 2955	Commensal	D
2957	Xanthomonas euroxanthea 2957	Commensal	D
2974	Xanthomonas euroxanthea 2974	Commensal	D
3004	Xanthomonas arboricola 3004	Commensal	D
3058	Xanthomonas sp. 3058	Commensal	D
3075	Xanthomonas sp. 3075	Commensal	D
3307	Xanthomonas sp. 3307	Commensal	D
3338	Xanthomonas euroxanthea 3338	Commensal	D
3498	Xanthomonas sp. 3498	Commensal	D
3640	Xanthomonas euroxanthea 3640	Commensal	D
3790	Xanthomonas arboricola 3790	Commensal	D
8590	Xanthomonas cannabis CFBP 8590	Commensal	D
84A	Xanthomonas arboricola 84A	Commensal	D
A2111	Xanthomonas sp. A2111	Commensal	D
AmX2	Xanthomonas sp.	Commensal	D
ATCC23378	Xanthomonas cucurbitae	Commensal	D
BRE_17	Xanthomonas campestris	Commensal	D
CaNP6A	Xanthomonas melonis	Commensal	D
CFBP3122	Xanthomonas arboricola pv. populi	Commensal	D
CFBP3123	Xanthomonas arboricola pv. populi	Commensal	D
CFBP426	Xanthomonas sp. CPBF 426	Commensal	D
CFBP4643	Xanthomonas pisi CFBP 4643	Commensal	D

CFBP4644	Xanthomonas melonis CFBP 4644	Commensal	D
CFBP6771	Xanthomonas arboricola	Commensal	D
CFBP6773	Xanthomonas arboricola pv. fragariae	Commensal	D
CFBP7604	Xanthomonas arboricola	Commensal	D
CFBP7614	Xanthomonas arboricola	Commensal	D
CFBP7634	Xanthomonas arboricola CFBP 7634	Commensal	D
CFBP7645	Xanthomonas arboricola	Commensal	D
CFBP7653	Xanthomonas arboricola	Commensal	D
CFBP7681	Xanthomonas arboricola	Commensal	D
CFBP7698	Xanthomonas sp. CFBP 7698	Commensal	D
CFBP7700	Xanthomonas campestris	Commensal	D
CFBP8130	Xanthomonas arboricola CFBP 8130	Commensal	D
CFBP8138	Xanthomonas arboricola	Commensal	D
CFBP8150	Xanthomonas arboricola	Commensal	D
CFBP8151	Xanthomonas sp. CFBP 8151	Commensal	D
CFBP8152	Xanthomonas sp. CFBP 8152	Commensal	D
CFBP8153	Xanthomonas arboricola	Commensal	D
CFBP8591	Xanthomonas cannabis CFBP 8591	Commensal	D
CFBP8595	Xanthomonas cannabis CFBP 8595	Commensal	D
CFBP8605	Xanthomonas cannabis CFBP 8605	Commensal	D
CFBP8700	Xanthomonas bonasiae	Commensal	D
CFBP8703	Xanthomonas bonasiae	Commensal	D
CITA124	Xanthomonas arboricola	Commensal	D
CITA44	Xanthomonas arboricola CITA 44	Commensal	D
CPBF367	Xanthomonas arboricola pv. juglandis	Commensal	D
D-109	Xanthomonas sp. D-109	Commensal	D
D-93	Xanthomonas sp. D-93	Commensal	D
DC06P2B	Xanthomonas arboricola	Commensal	D
DE0062	Xanthomonas citri pv. mangiferaeindicae	Commensal	D
DMCX	Xanthomonas melonis	Commensal	D
E1	Xanthomonas campestris strain:E1	Commensal	D

F10	Xanthomonas sp. F10	Commensal	D
F12	Xanthomonas arboricola F12	Commensal	D
F13	Xanthomonas arboricola F13	Commensal	D
F15	Xanthomonas arboricola F15	Commensal	D
F16	Xanthomonas cannabis F16	Commensal	D
F18	Xanthomonas arboricola F18	Commensal	D
F20	Xanthomonas arboricola F20	Commensal	D
F8	Xanthomonas arboricola F8	Commensal	D
FOR_F20	Xanthomonas arboricola	Commensal	D
FOR_F23	Xanthomonas arboricola	Commensal	D
GPE39	Xanthomonas sp. GPE 39	Commensal	D
GW	Xanthomonas sp. GW	Commensal	D
HWA1	Xanthomonas campestris	Commensal	D
JAI131	Xanthomonas sp. JAI131	Commensal	D
Leaf131	Xanthomonas sp. Leaf131	Commensal	D
LMC_P25	Xanthomonas campestris	Commensal	D
LMC_P47	Xanthomonas campestris	Commensal	D
LMC_P73	Xanthomonas campestris	Commensal	D
LMG12459	Xanthomonas sp. LMG 12459	Commensal	D
LMG12460	Xanthomonas sp. LMG 12460	Commensal	D
LMG12461	Xanthomonas sp. LMG 12461	Commensal	D
LMG12462	Xanthomonas sp. LMG 12462	Commensal	D
LMG19144	Xanthomonas arboricola	Commensal	D
LMG27592	Xanthomonas maliensis LMG 27592	Commensal	V
LMG476	Xanthomonas sacchari LMG 476	Commensal	D
LMG8992	Xanthomonas sp. LMG 8992	Commensal	D
LMG9002	Xanthomonas sp. LMG 9002	Commensal	D
MEDV_A40	Xanthomonas campestris	Commensal	D
MEDV_P25	Xanthomonas campestris	Commensal	D
MEDV_P39	Xanthomonas arboricola	Commensal	D
MLO165	Xanthomonas sp. MLO165	Commensal	D

MUS060	Xanthomonas sp. MUS 060	Commensal	D
MWU16-30325	Xanthomonas sp. MWU16-30325	Commensal	D
NCPPB1067	Xanthomonas sp. NCPPB 1067	Commensal	D
NCPPB1128	Xanthomonas sp. NCPPB 1128	Commensal	D
NCPPB2190	Xanthomonas campestris pv. esculenti	Commensal	D
NCPPB2983	Xanthomonas campestris pv. phormiicola	Commensal	D
NCPPB4231	Xanthomonas campestris	Commensal	D
NCPPB4232	Xanthomonas campestris	Commensal	D
NL_P121	Xanthomonas campestris	Commensal	D
NL_P172	Xanthomonas campestris	Commensal	D
PLY_2	Xanthomonas dyei	Commensal	D
PLY_4	Xanthomonas arboricola	Commensal	D
PLY_9	Xanthomonas arboricola	Commensal	D
PNG130	Xanthomonas albilineans PNG130	Commensal	D
PPL2	Xanthomonas sontii	Commensal	D
PPL3	Xanthomonas sontii	Commensal	D
R1	Xanthomonas arboricola pv. pruni	Commensal	D
Sa3BUA13	Xanthomonas surreyensis	Commensal	D
SAM114	Xanthomonas sontii	Commensal	D
SHU166	Xanthomonas sp. SHU 166	Commensal	S
SHU199	Xanthomonas sp. SHU199	Commensal	D
SHU308	Xanthomonas sp. SHU308	Commensal	D
SI	Xanthomonas sp. SI	Commensal	D
SL2098	Xanthomonas arboricola	Commensal	D
SN8	Xanthomonas massiliensis SN8	Commensal	D
SS	Xanthomonas sp. SS	Commensal	D
WHRI8481	Xanthomonas campestris	Commensal	D
WHRI8848	Xanthomonas floridensis WHRI 8848	Commensal	D
CFBP6681	Xanthomonas arboricola CFBP 6681	Commensal	D
CFBP7629	Xanthomonas arboricola CFBP 7629	Commensal	D
CFBP8145	Xanthomonas arboricola CFBP 8145	Commensal	D

F11	Xanthomonas arboricola F11	Commensal	D
F14	Xanthomonas arboricola F14	Commensal	D
F19	Xanthomonas arboricola F19	Commensal	D
F21	Xanthomonas arboricola F21	Commensal	D
F7	Xanthomonas arboricola F7	Commensal	D
F9	Xanthomonas arboricola F9	Commensal	D
CFBP13567	Xanthomonas campestris CFBP 13567	Commensal	D
CFBP13568	Xanthomonas arboricola CFBP 13568	Commensal	D
CFBP7911	Xanthomonas arboricola CFBP 8139	Commensal	D
CFBP7915	Xanthomonas campestris CFBP 7915	Commensal	D
CFBP7924	Xanthomonas campestris CFBP 7924	Commensal	D
F6	Xanthomonas campestris F6	Commensal	D
8600	Xanthomonas cannabis CFBP 8600	Commensal	D
CFBP8590	Xanthomonas cannabis 8590	Commensal	D
CFBP8592	Xanthomonas cannabis CFBP 8592	Commensal	D
CFBP8593	Xanthomonas cannabis CFBP 8593	Commensal	D
CFBP8594	Xanthomonas cannabis CFBP 8594	Commensal	D
CFBP8596	Xanthomonas cannabis CFBP 8596	Commensal	D
CFBP8597	Xanthomonas cannabis CFBP 8597	Commensal	D
CFBP8598	Xanthomonas cannabis CFBP 8598	Commensal	D
CFBP8600	Xanthomonas cannabis 8600	Commensal	D
F17	Xanthomonas cannabis F17	Commensal	D
Xcz13	Xanthomonas cannabis	Commensal	D
Xcz5	Xanthomonas arboricola	Commensal	D
CFBP7635	Xanthomonas euroxanthea CFBP 7635	Commensal	D
F2	Xanthomonas euroxanthea F2	Commensal	D
XNM01	Xanthomonas sp. XNM01	Commensal	D
3793	Xanthomonas sp. I 3793	Commensal	D
4461	Xanthomonas sp. I 4461	Commensal	D
60	Xanthomonas sp. 60	Commensal	D
F1	Xanthomonas sp. F1	Commensal	D

F4	Xanthomonas sp. F4	Commensal	D
F5	Xanthomonas sp. F5	Commensal	D
1	Xanthomonas euroxanthea 1	Pathogenic	V
85-10	Xanthomonas campestris pv. vesicatoria 85-10	Pathogenic	V
A7	Xanthomonas arboricola pv. corylina	Pathogenic	V
B99	Xanthomonas translucens pv. poae	Pathogenic	V
BA29-1	Xanthomonas vesicatoria	Pathogenic	V
BB151-3	Xanthomonas oryzae BB151-3	Pathogenic	V
BP5178	Xanthomonas hortorum	Pathogenic	V
Bv5-3A	Xanthomonas vesicatoria	Pathogenic	V
CCUG18839	Xanthomonas cissicola	Pathogenic	V
CFBP1817	Xanthomonas populi CFBP 1817	Pathogenic	V
CFBP1976	Xanthomonas bromi CFBP 1976	Pathogenic	V
CFBP2528	Xanthomonas arboricola pv. juglandis CFBP 2528	Pathogenic	V
CFBP2533	Xanthomonas hortorum pv. pelargonii	Pathogenic	V
CFBP410	Xanthomonas hortorum pv. taraxaci	Pathogenic	V
CFBP4642	Xanthomonas cassavae CFBP 4642	Pathogenic	V
CFBP4925	Xanthomonas hortorum pv. hederae CFBP 4925	Pathogenic	V
CFBP498	Xanthomonas hortorum pv. vitians	Pathogenic	V
CFBP7112	Xanthomonas citri pv. vignicola CFBP 7112	Pathogenic	V
CFBP7245	Xanthomonas dyei CFBP 7245	Pathogenic	V
CFBP7319	Xanthomonas oryzae pv. oryzae CFBP 7319	Pathogenic	V
CFBP7342	Xanthomonas oryzae pv. oryzicola CFBP 7342	Pathogenic	V
CFBP7408	Xanthomonas arboricola pv. guizotiae	Pathogenic	V
CFBP8304	Xanthomonas translucens	Pathogenic	V
CFBP8355	Xanthomonas prunicola	Pathogenic	V
CN03	Xanthomonas campestris pv. campestris str. CN03	Pathogenic	V
CN14	Xanthomonas campestris pv. campestris CN14	Pathogenic	V
CN18	Xanthomonas campestris pv. campestris CN18	Pathogenic	V
CO-5	Xanthomonas vasicola	Pathogenic	V
CPBF427	Xanthomonas arboricola pv. juglandis	Pathogenic	V

D-99	Xanthomonas sp. D-99	Pathogenic	V
DW3F3	Xanthomonas arboricola pv. juglandis	Pathogenic	V
F1	Xanthomonas sp. F1	Pathogenic	V
GBBC3236	Xanthomonas campestris pv. campestris	Pathogenic	V
GSXT20191014	Xanthomonas campestris pv. campestris	Pathogenic	V
JS749-3	Xanthomonas hortorum pv. gardneri JS749-3	Pathogenic	V
Km8	Xanthomonas translucens pv. Translucens	Pathogenic	V
Km9	Xanthomonas translucens pv. translucens	Pathogenic	V
LH3	Xanthomonas perforans LH3	Pathogenic	V
LM159	Xanthomonas vesicatoria LM159	Pathogenic	V
LMG31887	Xanthomonas hydrangeae LMG 31887	Pathogenic	V
LMG727	Xanthomonas translucens pv. arrhenatheri LMG 727	Pathogenic	V
LMG728	Xanthomonas translucens pv. poae LMG 728	Pathogenic	V
LMG843	Xanthomonas translucens pv. phleipratensis	Pathogenic	V
LMG911	Xanthomonas vesicatoria LMG 911	Pathogenic	V
LW16	Xanthomonas translucens LW16	Pathogenic	V
M081	Xanthomonas hortorum pv. carotae M081	Pathogenic	V
M28	Xanthomonas campestris	Pathogenic	V
Mex-1	Xanthomonas vasicola	Pathogenic	V
NCPPB1061	Xanthomonas campestris pv. plantaginis	Pathogenic	V
NCPPB1334	Xanthomonas campestris pv. ionidii	Pathogenic	V
NCPPB2372	Xanthomonas campestris pv. fici	Pathogenic	V
NCPPB2373	Xanthomonas campestris pv. carissae	Pathogenic	V
NCPPB2498	Xanthomonas campestris pv. convolvuli	Pathogenic	V
NCPPB3888	Xanthomonas campestris pv. parthenii	Pathogenic	V
NCPPB4013	Xanthomonas campestris pv. asclepiadis	Pathogenic	V
NCPPB4037	Xanthomonas campestris	Pathogenic	V
NCPPB4349	Xanthomonas campestris pv. pennamericanum	Pathogenic	V
NCPPB796	Xanthomonas axonopodis pv. vasculorum	Pathogenic	V
NEB122	Xanthomonas campestris pv. badrii NEB122	Pathogenic	V
PD885	Xanthomonas fragariae PD885	Pathogenic	V

PXO99A	Xanthomonas oryzae pv. oryzae PXO99A	Pathogenic	V
SB80	Xanthomonas campestris pv. campestris	Pathogenic	V
SHQP01	Xanthomonas fragariae	Pathogenic	V
X11-5A	Xanthomonas oryzae X11-5A	Pathogenic	V
Xa85	Xanthomonas axonopodis	Pathogenic	V
CFBP3277	Xanthomonas campestris CFBP 3277	Pathogenic	V
Xcc8004_Xcc1	Xanthomonas campestris pv. campestris	Pathogenic	V
Xcp1	Xanthomonas arboricola pv. Pruni	Pathogenic	V
BRIP39016	Xanthomonas euvesicatoria BRIP39016	Pathogenic	V
71	Xanthomonas campestris	Pathogenic	D
18048	Xanthomonas campestris pv. incanae	Pathogenic	D
756C	Xanthomonas campestris pv. raphani 756C	Pathogenic	D
A1809	Xanthomonas sp. A1809	Pathogenic	D
AR81009	Xanthomonas citri malvacearum AR81009	Pathogenic	D
Bagalkot	Xanthomonas citri pv. punicae	Pathogenic	D
CFBP1156	Xanthomonas hyacinthi CFBP 1156	Pathogenic	D
CFBP1606R	Xanthomonas campestris pv. incanae	Pathogenic	D
CFBP2524	Xanthomonas axonopodis pv. begoniae	Pathogenic	D
CFBP2527R	Xanthomonas campestris pv. incanae	Pathogenic	D
CFBP3836	Xanthomonas euvesicatoria pv. alfalfae CFBP 3836	Pathogenic	D
CFBP5825R	Xanthomonas campestris CFBP 5825R	Pathogenic	D
CFBP6164	Xanthomonas phaseoli pv. phaseoli CFBP 6164	Pathogenic	D
CFBP6369	Xanthomonas euvesicatoria pv. allii CFBP 6369	Pathogenic	D
CFBP6988	Xanthomonas citri sv. phaseoli fuscans CFBP 6988R	Pathogenic	D
CFBP6992	Xanthomonas citri sv. phaseoli fuscans CFBP 6992	Pathogenic	D
CFBP7111	Xanthomonas citri pv. vignicola CFBP 7111	Pathogenic	D
CFBP7113	Xanthomonas citri pv. vignicola CFBP 7113	Pathogenic	D
CFBP7119	Xanthomonas citri pv. glycines CFBP 7119	Pathogenic	D
CFBP7407	Xanthomonas arboricola pv. Arracaciae	Pathogenic	D
CFBP7764	Xanthomonas citri	Pathogenic	D
CFBP7921	Xanthomonas euvesicatoria CFBP 7921	Pathogenic	D

CHN01	Xanthomonas phaseoli pv. manihotis	Pathogenic	D
DAR26930	Xanthomonas euvesicatoria	Pathogenic	D
DC06T4A	Xanthomonas campestris	Pathogenic	D
FDC1559	Xanthomonas citri aurantifolii FDC 1559	Pathogenic	D
FDC1637	Xanthomonas euvesicatoria pv. citrumelonis	Pathogenic	D
FIJ080	Xanthomonas albilineans FIJ080	Pathogenic	D
FPH2013-1	Xanthomonas hortorum	Pathogenic	D
GPEPC86	Xanthomonas albilineans GPE PC86	Pathogenic	D
ISO118C5	Xanthomonas fuscans fuscans ISO118C5	Pathogenic	D
LMG12749	Xanthomonas euvesicatoria LMG12749	Pathogenic	D
LMG26789	Xanthomonas axonopodis pv. Commiphoreae	Pathogenic	D
LMG548	Xanthomonas axonopodis pv. Bauhiniae LMG548	Pathogenic	D
LMG695	Xanthomonas phaseoli pv. dieffenbachiae LMG 695	Pathogenic	D
LMG726	Xanthomonas translucens pv. graminis	Pathogenic	D
LMG753	Xanthomonas axonopodis pv. khayae LMG753	Pathogenic	D
LMG872	Xanthomonas citri pv. thirumalacharii	Pathogenic	D
LMG9050	Xanthomonas axonopodis pv. Melhusii LMG9050	Pathogenic	D
LMG954	Xanthomonas campestris pv. vitiswoodrowii	Pathogenic	D
MAFF106181	Xanthomonas campestris pv. raphani	Pathogenic	D
NCPPB1336	Xanthomonas euvesicatoria pv. alangii	Pathogenic	D
NCPPB1757	Xanthomonas campestris pv. blepharidis	Pathogenic	D
NCPPB1758	Xanthomonas campestris pv. coriandri	Pathogenic	D
NCPPB1760	Xanthomonas campestris pv. spermacoces	Pathogenic	D
NCPPB1787	Xanthomonas campestris pv. veroniae	Pathogenic	D
NCPPB1828	Xanthomonas campestris pv. euphorbiae	Pathogenic	D
NCPPB1946	Xanthomonas campestris pv. raphani	Pathogenic	D
NCPPB2057	Xanthomonas campestris pv. heliotropii	Pathogenic	D
NCPPB2337	Xanthomonas dyei pv. eucalypti	Pathogenic	D
NCPPB2439	Xanthomonas campestris pv. zinniae	Pathogenic	D
NCPPB3079	Xanthomonas campestris pv. paulliniae	Pathogenic	D
NCPPB347	Xanthomonas campestris pv. armoraciae	Pathogenic	D

NCPPB4348	Xanthomonas campestris pv. mirabilis	Pathogenic	D
NCPPB4351	Xanthomonas campestris pv. viegasii	Pathogenic	D
NCPPB464	Xanthomonas campestris pv. olitorii	Pathogenic	D
NCPPB579	Xanthomonas campestris pv. lawsoniae	Pathogenic	D
NCPPB586	Xanthomonas campestris pv. uppalii	Pathogenic	D
NL_P126	Xanthomonas arboricola	Pathogenic	D
TX160149	Xanthomonas citri pv. citri TX160149	Pathogenic	D
USA048	Xanthomonas albilineans USA048	Pathogenic	D
WHRI10004	Xanthomonas campestris pv. campestris	Pathogenic	D
WHRI10006	Xanthomonas campestris pv. raphani	Pathogenic	D
WHRI8853	Xanthomonas nasturtii WHRI 8853	Pathogenic	D
F22	Xanthomonas campestris F22	Pathogenic	D
F24	Xanthomonas campestris F24	Pathogenic	D
CFBP8599	Xanthomonas cannabis CFBP 8599	Pathogenic	D
XcitriDAR33341	Xanthomonas sp. DAR33341	Pathogenic	D
XcvDC91-1	Xanthomonas campestris	Pathogenic	D
CFBP7922	Xanthomonas euvesicatoria CFBP 7922	Pathogenic	D
3046	Xanthomonas arboricola 3046	Weaklypathogenic	D
3140	Xanthomonas arboricola 3140	Weaklypathogenic	D
3272	Xanthomonas arboricola 3272	Weaklypathogenic	D
3376	Xanthomonas arboricola 3376	Weaklypathogenic	D
1311A	Xanthomonas arboricola	Weaklypathogenic	D
1314C	Xanthomonas arboricola	Weaklypathogenic	D
BRIP62409	Xanthomonas sp. BRIP62409	Weaklypathogenic	D
BRIP62411	Xanthomonas sp. BRIP62411	Weaklypathogenic	D
BRIP62412	Xanthomonas arboricola	Weaklypathogenic	D
BRIP62415	Xanthomonas sp. BRIP62415	Weaklypathogenic	D
BRIP62416	Xanthomonas arboricola	Weaklypathogenic	D
BRIP62418	Xanthomonas sp. BRIP62418	Weaklypathogenic	D
BRIP62432	Xanthomonas arboricola	Weaklypathogenic	D
CFBP1022	Xanthomonas arboricola	Weaklypathogenic	D

CFBP4690	Xanthomonas codiaei CFBP 4690	Weaklypathogenic	D
CFBP4691	Xanthomonas theicola CFBP 4691	Weaklypathogenic	D
CFBP6827	Xanthomonas arboricola	Weaklypathogenic	D
CFBP7410	Xanthomonas arboricola pv. zantedeschiae	Weaklypathogenic	D
CFBP7610	Xanthomonas arboricola	Weaklypathogenic	D
CFBP7622	Xanthomonas euroxanthea CFBP 7622	Weaklypathogenic	D
CFBP7651	Xanthomonas arboricola CFBP 7651	Weaklypathogenic	D
CFBP7652	Xanthomonas arboricola CFBP 7652	Weaklypathogenic	D
CFBP7697	Xanthomonas arboricola	Weaklypathogenic	D
CFBP8132	Xanthomonas arboricola	Weaklypathogenic	D
CFBP8139	Xanthomonas arboricola CFBP 8139	Weaklypathogenic	D
CFBP8140	Xanthomonas arboricola CFBP 8140	Weaklypathogenic	D
CFBP8142	Xanthomonas arboricola	Weaklypathogenic	D
CFBP8147	Xanthomonas arboricola	Weaklypathogenic	D
CFBP8149	Xanthomonas arboricola CFBP 8149	Weaklypathogenic	D
CITA14	Xanthomonas arboricola	Weaklypathogenic	D
CPBF1494	Xanthomonas arboricola pv. juglandis	Weaklypathogenic	D
CPBF424	Xanthomonas euroxanthea	Weaklypathogenic	D
CPBF765	Xanthomonas arboricola pv. juglandis	Weaklypathogenic	D
CPBF766	Xanthomonas euroxanthea	Weaklypathogenic	D
LMG19146	Xanthomonas arboricola pv. fragariae	Weaklypathogenic	D
MEDV_A37	Xanthomonas arboricola	Weaklypathogenic	D
MEU_M1	Xanthomonas arboricola MEU_M1	Weaklypathogenic	D
NCPPB1630	Xanthomonas arboricola pv. Celebensis NCPPB 1630	Weaklypathogenic	D
NCPPB1832	Xanthomonas arboricola pv. celebensis NCPPB 1832	Weaklypathogenic	D
NCPPB2970	Xanthomonas campestris pv. papavericola	Weaklypathogenic	D
Nyagatare	Xanthomonas cannabis pv. phaseoli Nyagatare	Weaklypathogenic	D
X203	Xanthomonas codiaei	Weaklypathogenic	D
CFBP6825	Xanthomonas arboricola CFBP 6825	Weaklypathogenic	D
CFBP6826	Xanthomonas arboricola CFBP 6826	Weaklypathogenic	D
CFBP6828	Xanthomonas arboricola CFBP 6828	Weaklypathogenic	D

CFBP7680	Xanthomonas arboricola CFBP 7680	Weaklypathogenic	D

.

Table S 4-4: List of T3Es representing all effector families and putative effectors along with their protein sequence used in the study.

Family	Synonyms			Source	T3E id
AvrBs1			AvrA	AvrBs1 [Xanthomonas campestris pv. vesicatoria str. 85-10]	AvrBs1_Xeu
AvrBs2				avrA_family_Xg101	AvrBs1_Xg
				AvrBs2 [Xanthomonas campestris pv. vesicatoria str. 85-10]	AvrBs2_Xeu
AvrBs3			PthA, PthB, PthN, PthXo, TALE, AvrXa7, AvrXa27	avrBs3[Xanthomonas vesicatoria]	AvrBs3_Xve
				pthB	AvrBs3_Xac1
				pthC	AvrBs3_Xac2
				avXa7[Xanthomonas oryzae pv. oryzae KACC10331]	AvrBs3_Xoo1
				pthA3[Xanthomonas axonopodis pv. citri str. 306]	AvrBs3_Xac3
				pthA4[Xanthomonas oryzae pv. oryzae KACC10331]	AvrBs3_Xoo2
				pthA4[Xanthomonas axonopodis pv. citri str. 306]	AvrBs3_Xac4
				pthA2[Xanthomonas oryzae pv. oryzae KACC10331]	AvrBs3_Xoo3
				pthA2 [Xanthomonas axonopodis pv. citri str. 306]	AvrBs3_Xac5
				pthA1 [Xanthomonas axonopodis pv. citri str. 306]	AvrBs3_Xac6

XopB			HopD1, HopPtoD1, AvrPphD1	xopB [Xanthomonas campestris pv. vesicatoria str. 85-10]	XopB_Xeu
XopC	C1			xopC [Xanthomonas campestris pv. vesicatoria str. 85-10]	XopC1_Xeu
	C2		RSp1239	XopC2_XPE3585	XopC2_Xp
XopD				xopD [Xanthomonas campestris pv. vesicatoria str. 85-10]	XopD_Xeu
				xopD_XPE2945	XopD_Xp
XopE	E1		AvrXacE1, HopX(AvrPphE)	avrXacE1[Xanthomonas axonopodis pv. citri str. 306]	XopE1_Xac
				XopE1_XPE1224	XopE1_Xp
				gi|78034280|emb|CAJ21925.1| Xanthomonas outer protein E1 [Xanthomonas campestris pv. vesicatoria str. 85-10] {XopE1}	XopE1_Xeu
	E2		AvrXacE3, AvrXccE1, HopX (AvrPphE)	avrXccE1 [Xanthomonas campestris pv. campestris str. ATCC 33913]	XopE2_Xcc
				gi|78036266|emb|CAJ23957.1| Xanthomonas outer protein E2 [Xanthomonas campestris pv. vesicatoria str. 85-10] {XopE2}	XopE2_Xeu
				avrXacE3 [Xanthomonas axonopodis pv. citri str. 306]	XopE2_Xac
	E3		AvrXacE2, HopX (AvrPphE)	avrXacE2 [Xanthomonas axonopodis pv. citri str. 306]	XopE3_Xac

	E4		Putative HopX (AvrPphE)	SBV51443.1 putative type III effector protein XopE class [Xanthomonas bromi] {XopE4}	XopE4_Xbr
				xopE4_C-aurantifolii_strain	XopE4_Xau
	E5			QDI05960.1 type III effector HopX1 [Xanthomonas translucens pv. cerealis] {XopE5}	XopE5_Xtr
XopF	F1		Hpa4	xopF1 [Xanthomonas campestris pv. vesicatoria str. 85-10]	XopF1_Xeu
				XopF1_XPE2922	XopF1_Xp
	F2			xopF2[Xanthomonas campestris pv. vesicatoria str. 85-10]	XopF2_Xeu
				XopF2_XPE1639	XopF2_Xp
XopG	G1		HopH (HopPtoH), HopAP	xopG	xopG_Xph
XopH				xopH-avrBs1.1	xopH_X
XopI	I1			XopI_Fbox_XAC0754	XopI_Xac
				XopI_XPE3711	XopI_Xp
XopJ	J1		XopJ	xopJ [Xanthomonas campestris pv. vesicatoria str. 85-10]	xopJ1_Xeu
	J2		AvrBst, XopJ, HopZ2	AvrBsT [Xanthomonas campestris pv. vesicatoria]	XopJ2_Xeu
	J3		AvrRxv, XopJ	AvrRxv [Xanthomonas campestris pv. vesicatoria str. 85-10]	XopJ3_Xeu
	J4		AvrXv4, XopJ, PopP1	AvrXv4 [Xanthomonas campestris pv. vesicatoria]	XopJ4_Xeu
	J5		AvrXccB, XopJ	avrXccB[Xanthomonas campestris pv. campestris str. ATCC 33913]	xopJ5_Xcc

XopK				Xoo1669	XopK_Xoo
				XopK_XPE1077	XopK_Xp
				xopK_XAC3085	xopK_Xac
XopL				XCC4186_LRR	XopL_Xcc
				XopL_XPE1073	XopL_Xp
				xopL_XAC3090_LRR	xopL_Xac
XopN			HopAU1	xopN [Xanthomonas campestris pv. vesicatoria str. 85-10]	xopN_Xeu
				XopN_XPE1640	XopN_Xp
XopO			HopK1, (HopPtoK, HoIPtoAB), AvrRps4 N-terminal domain	xopO [Xanthomonas campestris pv. vesicatoria str. 85-10]	xopO_Xeu
XopP			HlK	xopP [Xanthomonas campestris pv. vesicatoria str. 85-10]	xopP_Xeu
				xopP_XPE3586	xopP_Xp1
				XopP_XPE4695	XopP_Xp2
XopQ			HopQ1 (HoIPtoQ), RipB	xopQ [Xanthomonas campestris pv. vesicatoria str. 85-10]	xopQ_Xeu
				XopQ_XPE0810	XopQ_Xp
XopR				Xoo4134_XopR	XopR_Xoo
				xopR	xopR_Xeu

				XopR_XPE3295	XopR_Xp
XopS				xopS	xopS_Xeu
XopT				Xoo2210	XopT_Xoo1
				xopT	xopT_Xoo2
XopU				xopU	xopU_Xoo
XopV				Xoo3803_XopV	XopV_Xoo
				xopV	xopV_Xeu
				XopV_XPE4158	XopV_Xp
XopW				Xoo0037	XopW_Xoo
XopX			HopAE1 (HoIPsyAE)	xopX [Xanthomonas campestris pv. vesicatoria str. 85-10]	xopX_Xeu
				xopX_XPE1488	xopX_Xp
XopY				Xoo1488	XopY_Xoo
XopZ	Z1		HopAS1	xopZ_XAC2009	xopZ1_Xac
				XopZ1_XPE2869	XopZ1_Xp
	Z2			xopZ2	xopZ2_X
XopA	A		Ecf, HopAE1 (HoIPsyAE)	holPsyAE/Xoo3022	XopAA_Xoo
				early chlorosis factor_Xeuves	XopAA_Xeu
	B		weakly related to XopN	xopAB	xopAB_X
	C		AvrAC	avrAC_XCC2565	XopAC_Xcc

	D			gi|188519096|gb|ACD57041.1| skwp protein 4 [Xanthomonas oryzae pv. oryzae PXO99A]	XopAD_Xoo
				xopAD	xopAD_Xac
				xopAD_XPE4269	xopAD_Xp
	E			XopAE_XPE2919_HpaF/G	XopAE_Xp
				xopAE	xopAE_Xac
	F	F1	AvrXv3, HopAF1 (HopPtoJ)	avrXv3_XopAF	XopAF_Xeu
	G		AvrGf1, HopG1 (HopPtoG), HoIPtoW	avrGf1_Xac306	XopAG_Xac1
				avrGf2	XopAG_Xac2
	H		AvrXccC, AvrB, AvrC (AvrPphC)	avrXccC[Xanthomonas campestris pv. campestris str. ATCC 33913]	XopAH_Xcc
	I		HopO1 (HopPtoO, HopPtoS), HopAI1 (HopIPtoAI)	xopAI_XAC3230	xopAI_Xac
	J		AvrRxo1	avrRxo1	XopAJ_X

	K		HopK1 (HopPtoK, HoIPtoAB)	xopAK_XAC3666	xopAK_Xac
				xopAK_XPE4569	xopAK_Xp
	L	L1	Eop3	xopAL1	xopAL1_Xcc
		L2	Eop3	xopAL2	xopAL2_Xcc
	M		HopR1	xopAM_avrRpm1family	xopAM_Xhor
	O		AvrRpm1	xopAO	xopAO_X
	P			XopAP_XPE_1567	XopAP_Xp
				xopAP	xopAP_Xcc
	Q		Rip6, Rip11	XopAQ_XGA2091	XopAQ_Xg
	R			XopAR_XPE_2975	XopAR_Xp
	S		HopAS1	XopAS_XGA0764/0765(HopAS_homolog)	XopAS_Xg
	U			xopAU	xopAU_Xcc
	V			xopAV	xopAV_Xcc
	W			xopAW	xopAW_Xcc
	X			xopAX	xopAX_Xcc
	Y			xopAY	xopAY_Xcc
			Hpa1, Hpa6	xopA [Xanthomonas campestris pv. vesicatoria str. 85-10]	xopA_Xeu
HpaA				HpaA [Xanthomonas campestris pv. vesicatoria]	HpaA_Xeu
HrpW				hrpW_306	hrpW_Xac
				hrpW	hrpW_Xeu

AvrXccA	1		AvrXca	avrXccA1 [Xanthomonas campestris pv. campestris str. ATCC 33913]	avrXccA1_Xcc
	2			avrXccA2 [Xanthomonas campestris pv. campestris str. ATCC 33913]	avrXccA2_Xcc

Table S 4-5: Designation of new Xanthomonas species from this study based on the Microbial Species Identifier (MiSI) method.

S.N.	Taxonomic classification	Species designation	MiSI		FASTANI	
			Cluster ID	Cluster Type	Most similar species	ANI %
1	Xanthomonas sp. F5	New species I	16927	singleton	Xanthomonas translucens	92
2	Xanthomonas sp. F1	New species II	3459	clique	Xanthomonas spp.	96
3	Xanthomonas sp. 60	New species III	16552	singleton	Xanthomonas retroflexus	83
4	Xanthomonas sp. F10	New species IV	16550	singleton	Xanthomonas sacchari	94
5	Xanthomonas sp. 3058	New species V	3626	clique	Xanthomonas arboricola	89
6	Xanthomonas sp. 3075					
7	Xanthomonas sp. CFBP 8152	New species VI	3417	clique	Xanthomonas arboricola	94
8	Xanthomonas sp. CFBP 8151					
9	Xanthomonas sp. 3793	New species VII	3442	clique	Xanthomonas pisi	94

10	Xanthomonas sp. 4461					
11	Xanthomonas sp. 3307	New species VIII	3418	clique	Xanthomonas arboricola	94
12	Xanthomonas sp. 3498					
13	Xanthomonas sp. F4					

5. CHAPTER FIVE: Bacterial spot of tomato: strain diversity, dynamics, and environmental drivers in the southeastern United States

Abstract

Bacterial leaf spot is a significant economic constraint to tomato production worldwide and is caused by the bacterium *Xanthomonas perforans* (*Xp*). Understanding the genetic variations of the pathogen and the factors driving its diversity and disease severity is essential for insights into successful disease management. We conducted a comprehensive three-year study in Southeastern US tomato fields, employing a high-resolution metagenomics approach to track the dynamics of *Xp* in tomato fields. Our findings revealed the presence of up to five co-occurring *X*p lineages within individual fields and that the composition and diversity of pathogen strains exhibit variability across space and time. Correlation analysis between the pathogen diversity and disease severity suggests that the diversity of *Xp* strains positively correlated with disease severity, suggesting that multiple lineages may increase the pathogen's ability to cause disease. A modeling approach to predict the drivers of disease severity found that environmental factors, such as photosynthetically active radiation (PAR) and wind direction, were significant predictors. Additionally, extremes in atmospheric pressure and relative humidity influenced the pathogen abundance. This study provides valuable information for disease diagnostics, epidemiological studies, and predictive modeling, emphasizing the importance of high-resolution metagenomics and considering meteorological parameter distribution in disease management and early warning systems.

5.1 Introduction

In natural ecosystems, continual interactions exist among plants, pathogens, and the environment, which are influenced by co-evolutionary processes (Burdon and Thrall 2009). Diseases are not of frequent occurrence in the natural environment, primarily due to host diversity and connectivity that result in stabilized co-evolutionary interactions among host and the pathogen. In contrast, the interactions between the pathogen and the host are altered under agricultural conditions. The monoculture system, global trade, widespread chemical use, and poor cultural practices create a perfect environment for pathogens to evolve rapidly and cause recurring outbreaks (Zhan et al. 2002; McDonald and Linde 2002; Bartoli et al. 2016; Sundin and Wang 2018; Wichmann et al. 2005). Isolate genome sequencing efforts over the years have provided insights into how specific virulence factors have evolved in response to host selection pressure due to deployment of resistant cultivars or how pathogen acquires resistance genes in response to chemicals (Velásquez et al. 2018). However, how environmental factors or other cultural practices foster pathogen diversification is not yet clear. The continuous input of different pathogen genotypes on seeds or plant materials that are circulated worldwide additionally contributes to the diversity observed under field conditions (Elmer 2001). As pathogen genetic diversity enhances fitness, adaptability, and evolution, enabling evasion of control measures, studying pathogen diversity is crucial for understanding infectivity, evolution, and adaptability in a changing environment.

Among various factors influencing the disease development and pathogen diversity, environmental conditions have a substantial impact. Plant disease epidemiologists often correlate disease outbreaks with the weather patterns, particularly in a changing climate (Shah et al. 2019). One of the most significant approaches for managing plant disease epidemics is genetic

resistance within the host plant. Yet, in the absence of resistant cultivars, farmers often resort to strategically applying crop protection chemicals to prevent diseases from reaching levels that would economically damage crop quality and yield, which are costly (Lechenet et al. 2017). Climate change has a variable effect on pesticide efficacy and exacerbates pest pressure (Ma et al. 2021). So, the relationship between weather and plant disease is commonly employed for predicting and controlling disease epidemics, as the disease epidemics can vary based on climatic fluctuations (Chakraborty et al. 2000; Seherm and Coakley 2003). Predictive models are being developed for many crops, which helps farmers evaluate the disease risk and make management decisions based on these climatic factors (Shah et al. 2019). A plant disease model is a simplified representation of complex interactions within the pathosystem, encompassing pathogens, crops, and the environment, which lead to epidemic development occurring over time and space (van Maanen and Xu 2003; Soubeyrand et al. 2008; Lee et al. 2020). These disease models have been developed since the middle 1900s, showing the relationship between components of disease cycles and their associated weather conditions (Zadoks 1971; Plank 2013). But a significant problem in predicting plant diseases over time and space is understanding the intricate relationship between the pathogens and how they respond to various disease-driving factors in the face of climate change. Climate change can significantly impact plant disease dynamics by altering pathogen evolution, host-pathogen interactions, and the emergence of new pathogen strains, potentially compromising host-plant resistance (Cohen and Leach 2020; Newbery et al. 2016). Additionally, it can lead to shifts in the geographic ranges of pathogens and hosts, potentially expanding the spread of plant diseases into new areas (Dudney et al. 2021; Chaloner et al. 2021). In case of endemic diseases, such climatic fluctuations can lead to alterations in disease severities. While the potential effects of variable climatic factors in agriculture have been

extensively discussed, environmental changes can alter the disease triangle by favoring new climatic preferences and altering pathogen dynamics in plant-pathogen interactions (Velásquez et al. 2018). Despite being often overlooked in epidemiological modeling, pathogen diversity and variations in genotype fitness play a significant role in shaping the course of epidemic diseases.

Bacterial leaf spot (BLS) is a disease affecting tomato and pepper plants, with economic implications for processing and fresh market produce. It can impact various aerial parts of these plants and is found in tropical, subtropical, and temperate regions globally (Jones et al. 2004; Hamza et al. 2010).(Jones et al. 2004; Hamza et al. 2010). BLS is considered an endemic problem in the United States, particularly in the Southeast, being known to be present now for more than a century. This disease results from a genetically diverse group of *Xanthomonas* bacteria, including *X. euvesicatoria* pv. *euvesicatoria (Xeu)*, *X. euvesicatoria* pv. *perforans (Xp)*, *X. hortorum* pv. *gardneri*, and *X. vesicatoria* (Jones et al. 2004). Although *Xeu* was the dominant pathogen of tomato in the mid 1990s, *Xp* is a dominant species in the southeastern US for the past three decades (Timilsina et al. 2015). BLS is a seed-borne pathogen, primarily spreading through contaminated seeds and transplants (Dutta et al. 2014). Research into the population structure of BLS pathogens in the United States has revealed considerable diversity, marked by species shifts, race changes, and an expansion of host ranges (Timilsina et al. 2016; Schwartz et al. 2015). Genomic studies based on the Multilocus Sequence Analysis (MLSA) as well as core genome have indicated that multiple recombination events and horizontal transfer events are responsible for the emergence of new lineages of Xp over the period of past three decades (Timilsina et al. 2019; Newberry et al. 2019). Moreover, the co-occurrence of multiple lineages of *Xp* in the same field was recently observed by our group when we used a high-resolution metagenome sequencing approach (Newberry et al. 2020). The presence of multiple

co-occurring pathogen genotypes in the field at a given sampling time suggest that it is not just one single genotype-host binary interaction that we need to focus on, but rather consider the diversity existing in the field when considering management strategies. Considering this, recognizing the impact of environmental conditions on plant-*Xanthomonas* interactions is essential for predicting disease outbreaks, associated pathogen diversity and developing resilient crop plants capable of withstanding current and future climate changes.

Host resistance is the primary and most effective approach to managing BLS disease, however the emergence of pathogenic races capable of overcoming resistance genes in tomatoes has led to a lack of commercially available resistant tomato cultivar. The efficacy of copper-based bactericides is diminishing due to the development of copper-resistant bacterial pathogens resulting in disease outbreaks around the world. As the development of effective management strategies including resistance breeding, relies upon the accurate identification of pathogens and an understanding of their diversity and pathogenicity, this study describes the large-scale survey of pathogen diversity using high resolution metagenome profiling and then studying strain dynamics during growing season and three years across the fields in southeast United States. We screened the metagenome data for the presence of eight *X. perforans* lineages that we identified based on phylogenetic analysis of isolate genome collected globally. The high-resolution shotgun metagenome analysis was useful in studying patterns of co-occurrence of lineages across fields spatially and temporally. It is currently unknown whether presence of such multiple lineages has influence on the disease severity levels, although co-operation among genotypes may be quite common leading to collective performance of pathogen population with high disease severity values. So, we looked whether the diversity of lineages present in the field is associated with disease severity. Moreover, as isolate genome sequencing results have shown the

dominance of *Xp* in tomato after mid 1990s, we wanted to see if high resolution approach, in fact, detects any low levels of *Xeu*. This study also addresses the knowledge gap regarding how climate components and their interaction with human activities affect endemic bacterial pathogen dynamics and disease severity. This information can inform management practices and targeted breeding programs to mitigate the influence of these pathogens.

5.2 Materials and Methods

5.2.1 Reconstruction of the *X. perforans* core genome

Sequencing reads for 497 *Xp* strains downloaded from NCBI (as of 08/15/2023) along with quality trimmed reads from strains collected from a pepper field in Georgia (sequenced by our group, unpublished), was used to study the global diversity of *Xp* isolate genomes. To ensure high-quality and minimally biased genomes, the genomes were de-replicated using dRep (v3.2.2) (Olm et al. 2017) using a 95% minimum genome completeness and 5% maximum contamination. We manually discarded the genomes with more than 400 contigs to ensure better alignment with the reference genome and avoid bias while creating the SNV profile. Four hundred sixty seven quality-controlled genomes collected from tomato, and pepper (with one eggplant isolate) s around the globe were individually aligned against the completed genome of *Xp* strain LH3. Single-nucleotide polymorphisms (SNPs) common to all genomes were extracted to generate a concatenated set of high-quality core-genome SNPs using Parsnp (Treangen et al. 2014). The phylogenetic tree was visualized and annotated using iTol (v.6.7.5) (Letunic and Bork 2007)

5.2.2 Sample collection and disease severity estimation

To study the diversity of BLS *Xanthomonas* in the southeastern United States, we collected 66 samples from 22 fields. These samples were collected from symptomatic and asymptomatic tomato plants grown in Alabama, Georgia, South Carolina, and North Carolina over 2020, 2021, and 2022 (Fig. S5-1, Table S5-1). Samples were taken during mid-season, end-season, and, when available, in the winter to capture the temporal dynamics of the pathogen population. While sampling, two to three leaflets were randomly taken from individual plants in the field. Multiple plants were sampled across the field to obtain a representative 50-100 grams leaf tissue sample. Assessments of the disease were made by estimating the percent of disease symptoms and defoliation caused by bacterial spots using the Horsfall-Barratt scale, which ranges from 1 to 12, with scale of 1 being no disease and 12 being 100% defoliation (Horsfall and Barratt 1945).

5.2.3 DNA extraction and sequencing

Metagenomic DNA extraction, quantification and sequencing of the samples was done as described earlier (Newberry et al. 2020). Approximately 40 grams of the leaf sample was taken in a Ziploc bag (Ziploc®) with 50 ml 0.1% wash buffer (0.05M PBS, 8.5 g NaCl, and 0.2 ml tween 20 per liter of water), sonicated for 20 minutes, followed by transfer of the buffer to a 50 ml falcon tube with the help of a sterile pipette. The tubes were centrifuged at 4000 rpm, 4°C for 20 minutes (Eppendorf® 5418 R) to collect all the cells. The cells were then washed twice with sterile deionized water, and the metagenomic DNA extraction was done using Wizard® Genomic DNA Purification Kit (Promega) as per the manufacturer's instructions with the addition of a phenol-chloroform step to remove protein contamination. The extracted DNA was quantified using a

Qubit® fluorometer and dsDNA high-sensitivity assay kit (Life Technologies, Carlsbad, CA, USA) and kept at -80^{0}C until further processing. Sequencing was performed on the Illumina HiSeq 4000 platform. The raw reads were processed and trimmed for the adapter using BBDuk (https://jgi.doe.gov/data-and-tools/bbtools/bb-tools-user-guide/bbduk-guide/). DynamicTrim command of SolexaQa was used to trim the reads based on quality values with Qphred less than 20, followed by filtering the reads shorter than 50bp (Cox et al. 2010). Host reads were separated and removed from the metagenomic samples using KneadData tools (https://bitbucket.org/biobakery/kneaddata/wiki/Home) using tomato cv. Heinz 1706 (GCA_022405115.1) as a reference.

5.2.4 Taxonomic profiling and diversity estimation

The relative abundance of *Xanthomonas* sp. in quality-controlled and host-decontaminated reads was performed using Kraken2 (v2.1.2) (Wood et al., 2019) against a standard Kraken2 database. The database contained RefSeq libraries of archaeal, bacterial, human, and viral sequences as of March 1, 2022 (O'Leary et al., 2016). The resulting Kraken2 report files were used as inputs for Bayesian re-estimation of abundance using Kraken (Bracken) (v2.6.2) (Lu et al., 2017). We also calculated the absolute abundance estimates for *Xp* by combining their relative abundance with the total DNA content in each sample. Microbiota density, expressed as total DNA (ng) per mg of fresh sample, was computed for each sample. This information was utilized to determine the absolute abundance of *Xp* by multiplying the ng of DNA per mg of the sample with their relative abundance.

To explore the intraspecific diversity and dynamics of *Xp* in the metagenomic samples, we employed the StrainEst pipeline (Albanese and Donati, 2017). StrainEst is a reference-based

method that utilizes single-nucleotide variant (SNV) profiles from carefully selected reference genomes. This approach enables us to estimate the presence and relative abundance of strains within a given sample. Initially, we conducted SNV profiling of representative strains from *Xp*, *Xeu*, and other closely related pathovars belonging to *Xeu* species complex to determine their presence. A second round of profiling was done to identify the diversity of Xp strains. A representative genome from each *Xp* SC (Fig. 5-1) was selected and aligned against the completed reference genome of *Xp* strain 91-118 using the MUMmer algorithm (Kurtz et al. 2004). This genome alignment was used to construct a SNV matrix, where each row corresponded to a variable position in the reference genome, and columns contained allelic variants present in the reference strains. This matrix was used as a reference in modeling strain level abundance.

To assess the diversity of *Xp* SCs within the samples, we computed the Shannon diversity index (Shannon-Wiener diversity index) using the "diversity()" function from the R package *vegan* (v2.6-4) (Dixon 2003). The index considers both the count of SCs present (richness) and their proportional distribution (evenness). A higher index value indicates greater species diversity in the habitat. When the Shannon diversity index of a sample is equal to 0, it implies that only a single SC is detected in that sample.

5.2.5 Weather data

Daily climate data spanning three years for all 22 fields were compiled from the NASA Power database using the nasapower (v4.0.12) package (Sparks 2018) in R. Data were collected during mid-season, end-season, and winter samples as per the sampling time in Table S1. The gridded weather data used in this analysis are based on satellite observations and models and are offered globally at a horizontal resolution of 0.5° X 0.625° latitude–longitude grid cell. Weather

variables, such as temperature, precipitation, solar radiation, relative humidity, wind, and others (Table S5-2), were consolidated over time and space to generate data points for each sampling time, resulting in an average value. The average for mid-season was calculated as an average from the time of planting to mid-season sampling time, which is usually 60 days after transplanting, and the average for end-season is calculated as the average from the time of sampling to end-of-season sampling time, which is generally 30 days after mid-season sampling.

Variability among the weather parameters were measured using the descriptive statistics (standard deviation, skewness, entropy, and kurtosis) for all the climatic variables to see if the extreme temperature events during the growing season have an impact on disease severity, pathogen diversity, and abundance. Standard deviation measures the deviation from the mean value. Higher the standard deviation indicates greater variability in the weather parameters while lower standard deviation suggests more consistent pattern. Skewness measures the degree and direction of asymmetry. A positive skewness implies that there are more extreme high weather parameters while a negative skewness suggests more extreme low with zero suggesting a symmetric distribution. Kurtosis is a measure of tail extremity reflecting either the presence of outliers in a distribution (Westfall 2014). Higher positive kurtosis suggests that presence of extreme weather values or outliers while negative kurtosis indicates the flatter distribution with less extreme weather values. Similarly, entropy quantifies the level or unpredictability in the weather values. Higher the entropy, more will be the unpredictability in the weather fluctuations.

5.2.6 Statistical analysis

We quantified the relative abundance of various *Xp* lineages across the samples and analyzed their distribution patterns in relation to various climatic factors. We then did a principal

component analysis using a biplot to know the main climatic factors that are related to the variation observed among *Xp* lineages and computed a pairwise average bi-distance (Bray-Curtiss using ggplot2 package in R. In our study, we harnessed the power of the Least Absolute Shrinkage and Selection Operator (LASSO) regression models to identify the most influential factors affecting disease severity. LASSO is the preferred feature selection method in our analysis, as it combines the benefits of ridge regression and subset selection, enhancing model accuracy and interpretability (Tibshirani 1996). This model can also simultaneously conduct variable selection and regularization, improving predictive accuracy and model interpretability, rendering it the top choice among regression models (Fonti and Belitser 2017). LASSO imposes an upper limit on the absolute values of model parameters. The method employs a shrinkage, or regularization, process that penalizes regression variable coefficients, effectively reducing some to zero. Feature selection was performed on the variables that still have non-zero coefficients after the shrinkage process using the "cv.glmnet" function to prevent the overfitting of the model. This study determined the tuning parameter for regularization using 10-fold cross-validation. The path length (min_lambda/max-lambda) was chosen to be 0.001, and 100 default values along the regularization pass were tested to find the best lambda value. The effects of the climatic parameters on *Xp* relative and absolute abundance were analyzed using a beta regression analysis in R using the *betareg* package (Cribari-Neto and Zeileis 2010). The beta regression model is employed for non-normally distributed data with a range between 0 and 1, addressing heteroskedasticity and asymmetry issues (Ferrari and Cribari-Neto 2004), and is specifically chosen here due to the non-normal distribution of *Xp* relative abundance in the sample, which conforms to a beta distribution.

5.3 Results

5.3.1 Diversity of bacterial leaf spot pathogen *Xp* around the world

More than 500 genomes of *Xp* are now available for the strains collected around the world over the last three decades. Our group had identified 6 lineages (also known as sequence clusters) based on core single nucleotide polymorphisms (SNPs) phylogeny using the set of published genomes available then in 2019 (Newberry et al. 2019). However, after this study, although there was exponential increase in sequenced *Xp* genomes, the studies did not include all the sequenced genomes in informing population structure or used other methods to characterize pathogen population structure, thereby, leading to a knowledge gap of relationship of existing strains to the newly sequenced strains and the global population structure. Here, utilizing a maximum likelihood phylogenetic approach, a phylogenetic tree was constructed based on core genome SNPs, and a Bayesian analysis of population structure was performed. This comprehensive analysis identified eight distinct lineages within *Xp* strains referred here as sequence clusters (SC) (Fig. 5-1). SC1 and SC2 represent previously characterized population structures of *Xp* strains primarily collected from tomatoes, primarily in Florida but also from Ohio, Indiana, and Georgia. Most strains in SC3 originate from tomatoes in Florida, and this group also includes six strains isolated from peppers, including three new strains from this study (GAJa1, GAJa2, and GACu1). SC4 primarily consists of strains collected from tomatoes in the United States, specifically Florida, Ohio, and Alabama. It includes a single strain isolated from tomatoes in Australia (BRIP62398), one from eggplants in Turkey (Tu-04), and a small number of tomato strains isolated from Canada. SC5 consists exclusively of tomato strains collected from Alabama and Florida. SC6, on the other hand, is a diverse cluster primarily composed of strains collected from peppers in Alabama, Georgia, and Florida. Additionally, this cluster includes

eight pepper strains obtained during this study. Furthermore, SC6 contains a subset of tomato strains from Turkey and one strain each from Alabama and Australia. In contrast, SC7 is entirely composed of *Xp* strains collected from tomatoes in Australia, while SC8 consists of tomato strains collected from Florida.

5.3.2 Abundance of *Xp* is positively associated with BLS disease severity

Out of the 66 tomato phyllosphere samples gathered over three years, six exhibited no symptoms of BLS. Among 66 fields, not all the field were sampled for both mid-season and end season and for all three years. A total of 11 fields were sampled during both mid and end season. The average disease severity in samples from Alabama was lower (2.7) during mid-season when compared to other states. However, the average disease severity during end season was higher in Alabama (6) as compared to other states (Fig. 5-2A). Beside Alabama and North Carolina, average disease severity during end season was lower in samples from Georgia and South Carolina when compared to mid-season. The relative abundance of *Xp* was positively correlated with disease severity both during the mid and end season (mid-season: $R^2 = 0.5084$, $p < 0.001$, end season: $R^2 = 0.378$, $p < 0.001$). Similar results were observed with absolute abundance of *Xp* during both timepoints (mid-season: $R^2 = 0.3814$, $p < 0.001$, end season: $R^2 = 0.2833$, $p < 0.01$). The pathogen population had more significant influence on disease severity during early season compared to late season based on R^2 value (Fig. 5-2B). SNV profiling of *Xp, Xeu* and *Xeu* sister cluster using StrainEST showed absence of of *Xeu* in the sampled tomato fields. The presence of *Xp* alone suggest *Xp* is the dominant pathogen of BLS in tomatoes in southeastern United States.

5.3.3 *Xanthomonas* abundance reveals co-occurrence of various *Xp* lineages

Subsequently, we conducted a second round of SNV profiling to examine which Xp lineages that are identified in the global collection are present in the individual fields of

southeastern US. From our analysis, we identified the presence of all lineages in our samples. Notably, SC1, SC7, and SC8 were found to be present in lower abundance, appearing in only 8 of the samples. Co-occurrence of different *Xp* lineages in the samples was observed, as was also noted in our previous limited metagenome survey. Interestingly, fields had minimum of 1 to maximum of 5 lineages present at a given sampling time and this diversity of co-occurring lineages was different across different years and samples from neighboring states (Fig. 5-3). Among the *Xp* lineages, SC3 was dominant in Georgia across all three years with fewer abundance of SC4, SC2, and SC1. Similarly, SC4, SC5, and SC6 were the prevalent SCs in South Carolina. In Alabama, in most of the samples, SC3, SC4, SC5, and SC6 were prevalent. Although we had only one field sample with *Xanthomonas* from North Carolina, the sample was dominated with SC4 and SC2. In most of the samples, there was a mixed occurrence of multiple SCs where SC3, SC4 and SC6 was the most prevalent combination across all samples.

Changes in strain composition were not only evident across different fields but also within individual fields throughout the growing season, spanning a period of three years. These dynamics in strain composition were marked by fluctuations in their relative abundance, with a notable prevalence of introduction of new lineages during the end season, introduction of completely new strains, and the dominance of one strain over others. For example, in Farm 3 during the year 2022, only SC4 was present. However, as the season transitioned from mid to end, the abundance of SC4 decreased significantly, coinciding with the introduction of new lineages of, SC3 and SC2. A similar pattern was observed in Farm 22, where the abundance of SC3 decreased from mid-season to end season, accompanied by the introduction of a new lineage, SC5, during the end season in 2020. In Farm 10, a comparable reduction in the abundance of SC4 occurred with the introduction of SC3 and SC5 during the end season.

Moreover, in Farm 7, a new lineage SC3, was introduced, leading to a decrease in the abundance of SC4 during the end season.

Notably, introduction of completely new lineages (SCs) was also observed in some fields over this three-year period in Alabama. For instance, in Farm 1 during the year 2020, only SC6 was prevalent. In 2021, SC6 dominated the mid-season, while SC5 and SC7 were introduced and became prominent during the end season. However, in 2022, a new strain, SC4, was introduced, and it became the dominant strain in the samples. A similar pattern was observed in Farm 4, where SC6 was the predominant strain in 2020. In 2021, SC5 was introduced during the mid-season and later dominated the end-season sample. However, in 2022, a novel strain, SC3, emerged and completely dominated the mid-season sample. This result suggests these dynamics in the co-occurrence of these SCs might be influenced by various climatic, host, and anthropogenic factors.

5.3.4 *Xp* population diversity is positively associated with disease severity

Next, we investigated whether the presence of multiple co-occurring lineages of *Xp* within individual fields influenced disease severity. We estimated shannon diversity index that captures evenness indicative of the extent of diversity of *Xp*. Correlation analysis between shannon diversity index of pathogen population and disease severity revealed a positive association ($R^2 = 0.52$, $p < 0.001$) (Fig. 5-4). This suggests that the presence of multiple SCs in the field leads to more severe disease.

5.3.5 Variation in *Xp* lineages during growing season is influenced by various climatic factors

To investigate the factors driving the composition and diversity of various *Xp* lineages in our samples, we used Principal Coordinate Analysis (PCoA) plots, utilizing the compositional variations in *Xp* lineages (assessed using Bray-Curtis distances). The results revealed distinct sample clustering patterns based on the timing of sample collection, distinguishing between mid, end, and winter seasons (Fig. 5-5). The ordination plot further suggested that UV index, temperature, and precipitation are the factors that drive the *Xp* lineages diversity during end-season while wind speed is the major driver of *Xp* lineages during mid-season and winter (Fig. 5-5).

5.3.6 Photosynthetically active radiation and wind speed significantly influence BLS disease severity

Next, we focused on understanding which climatic factors drive disease severity. We subsequently employed a disease modeling approach. We used disease severity as our response variable and considered climatic factors, sampling time, year, and farm-scale as predictors, treating the latter three as categorical variables. In addition to routine climatic variables, their mean or median values, we also considered upscaling approach, that allows considering the extreme climatic values using indexes that capture the spread (standard deviation), asymmetries, and tail-heaviness (skewness and kurtosis). We employed an ordinal logistic regression model to analyze the relationship between their predictors and ordinal disease severity. We assessed statistical significance using t-values, with values exceeding 2 or falling below -2 indicating significance. This approach helped identify key factors influencing the severity of BLS disease.

Although not significant, the size of the sampling field positively influenced the disease severity (t value = 1.78), suggesting that an increased sampling field size would increase the possibility of increased disease severity. A similar observation was made about sampling year (t value = 1.93), suggesting that in 2020, the BLS disease severity in the sampled field was higher than in the years 2021 and 2022. Among the climatic variables, the predictor average of clear sky surface photosynthetically active radiation (PAR), indicative of the photosynthetic capability of plants (t value = 2.3826) and standard deviation of the average of the wind direction at 10 meters above the surface of the earth (t value = 2.4679) had the significant positive influence on BLS disease severity (Table 5-1A).

5.3.7 Frequent shifts in wind direction and extreme changes in surface pressure and relative humidity influences *Xp* abundance

A beta regression model was designed to predict the relative and absolute abundance of *Xp* (response variable) and the influence of climatic factors, sampling time, year, and farm-scale as predictors, treating the latter three as categorical variables as described earlier. Results from the beta regression demonstrate that the standard deviation of the average of the wind direction at 10 meters above the surface of the earth ($p < 0.01$) and farm scale where commercial farm ($p < 0.01$) has a significant positive effect on *Xp* relative abundance (Table 5-1B). Moreover, the skewness of surface pressure ($p < 0.01$) and Kurtosis of relative humidity ($p < 0.01$) showed a significant negative effect on *Xp* relative abundance (Table 5-1B). A similar observation of negative interaction between Kurtosis of relative humidity at 2 meters ($p < 0.01$) and skewness of surface pressure ($p < 0.001$) was observed on the absolute abundance of *Xp*. In comparison, only the standard deviation of the wind direction at 10 meters above the surface of the earth had significant positive interaction ($p < 0.01$) (Table 5-1C) with *Xp* absolute abundance.

5.4 Discussion

BLS pathogen, *X. perforans*, has shown extraordinary diversity over time in various parts of the world (Schwartz et al. 2015; Timilsina et al. 2016; Newberry et al. 2019; Timilsina et al. 2019; Klein-Gordon et al. 2021). Different strains of the same species differ significantly in their gene content, single nucleotide polymorphism (SNPs), etc. These differences in different strains within the same species for the basis for understanding microbial adaptation and evolution (Zhu et al. 2015; Schloissnig et al. 2012; Myers et al. 1993). While previous studies using isolate genome sequencing have indicated genomic changes or processes that may have contributed towards diversification of the pathogen into several lineages (eight, to be precise, based on the comprehensive analysis using core genome SNP phylogeny built in this study), we lack understanding of the extent of diversity existing in the fields at a given time. Being endemic disease, presence of the disease in the tomato fields in the southeastern US is not surprising. But what drives the differences in disease severity across fields in the neighboring states is a complex question to answer. This is primarily because of multiple factors involved in driving the disease dynamics, ranging from climatic factors, choice of host genotypes, scale of farm operations, source of seeds or transplants, field sanitation, other cultural practices, to the management strategies used by the growers to tackle bacterial spot disease or other fungal or nematode diseases. We attempted to address these knowledge gaps in this study by using strain-resolved metagenomics approach that we optimized on phyllosphere samples in 2019 (Newberry et al. 2020). This high-resolution method allowed us to track strain dynamics of *Xp* in individual fields in the southeastern US over the course of three years. Beyond pathogen population survey, a rich metadata collected alongside samples from the field and climate data allowed us to identify

variable importance of individual factors in contributing towards disease severity and pathogen diversity.

Our survey of fields in Alabama, Georgia, North and South Carolina in the growing seasons of 2020-2022 indicated overall higher disease pressures during mid-season compared to the end season. The disease severity was positively correlated with *Xanthomonas* abundance. Overall higher severity values in mid-season compared to end season could be due to age-related host resistance (Hu and Yang 2019, Whalen 2005; Sharabani et al. 2013), or prevalence of other fungal diseases later during the growing season, or changes in climatic conditions less conducive for further disease development of bacterial spot.

Strain-resolved metagenomics approach indicated prevalent lineages in the fields. Interestingly, all eight lineages are currently circulating in the fields of the southeastern US, although some lineages are always found in low abundance. The spatio-temporal analysis of strain dynamics allowed us to identify persisting lineages across fields. As previously noted in our pilot study on limited samples, we observed co-occurrence of up to five lineages of *Xp* in individual fields. Intraspecific co-occurrences are common in the natural environment due to stabilizing host-pathogen interactions but are assumed to be of rare occurrence under agricultural environment (Susi et al. 2015; Walkowiak et al. 2015; Karasov et al. 2018). Co-occurrence of microbial communities can ~~infections~~ result in either increased or decreased pathogen virulence, depending on the specific interactions among the co-infecting bacterial strains. The disease severity values obtained in this survey were found to be positively correlated with *Xp* diversity, indicating the overall pattern of higher disease severity in presence of multiple lineages. The positive association between the diversity of the SCs and disease severity suggest the possible mutualistic or competitive interactions among the SCs (Fig. 5-3) (West and Buckling 2003;

Buckling and Brockhurst 2008, Kinnula et al. 2017; Inglis et al. 2009; Carvalho et al. 2023). Increased in diversity of the co-occurring strains might have resulted in more diverse gene pool with more accessory genes which might provide benefit to the pathogens to adapt to fluctuating climatic conditions and host genotypes. Increased in fitness of the co-occurring microbial communities by sharing the public good molecules which includes virulence factors, toxins, signaling molecules is widespread in bacterial communities (Dimitriu et al. 2014; West et al. 2007). The possible increased co-operation and sharing of resources among the co-occurring SCs and the inability of the host defense system to defend against multiple infections, might be linked with increased disease severity.

The composition and diversity of pathogen strains was found to exhibit variability across both space and time. Seasonal fluctuations in the abundance of these strain clusters may be attributed to the process of strain assembly and the differential survival of strains within their local environment (Krasnov et al. 2015). With the onset of new growing season, new strain clusters are introduced to the field through seeds or transplants, alongside strains from the previous season, which may originate from either the same seed source or overwintering pathogens. This results in the introduction of new strain clusters or the persistence of similar ones across multiple years in the field. During the growing season, the spread of infection among crops occurs through various means, including aerosols, wind-driven rain, and human activities. Spatial variation of the SCs across different states might be driven by heterogeneity in tomato cultivars and varied environmental conditions. Variation in the environmental conditions and genotype specific response of pathogen to different cultivars is known to promote the differentiation in the pathogen population (Thompson and Burdon 1992).

Environmental factors are known to exert a substantial influence on the development of diseases in field crops. Consequently, assessing the connection between epidemiological variables and disease severity is crucial. This assessment provides an early warning system to predict the disease's onset and understand the influence of different factors that drives BLS disease severity. In this study, we examined the effect of climatic factors and the prevalence of extreme weather events on BLS disease severity and *Xp* relative and absolute abundance. The ordinal regression model indicated that average clear sky surface photosynthetically active radiation (PAR) and standard deviation of wind direction at 10 Meters (WD10M) are significant predictors of higher disease severity. Photosynthetically active radiation is in the spectral range of 400 to 700 nm and is known to play a key role in plant growth and biomass production through photosynthesis (Frolking et al. 1998; McCree 1981). It is known that the change in the global climate affects plant's ability to make defense hormones, making them more prone to disease (Kim et al. 2022; Bhandari et al. 2023). As sun is the main source of energy, the rate of energy from the sun influences the climate variability in the earth including temperature, precipitation, etc. (Bhargawa and Singh 2019). So, a shift in other climatic factors due to change in PAR might exacerbate the disease. Therefore, further investigation with consideration of *Xanthomonas*-host interactions and their adaption/evolution to gradually changed environmental factors is necessary to understand their response to global climate changes better. Wind plays an essential role in disseminating various plant bacterial pathogens (Vidaver and Lambrecht 2004). The direction of the wind can play a crucial role in the dispersal and drift of *Xanthomonas* within and between the fields. Thus, finding of shifts in wind direction influencing disease severity levels is not surprising.

The abundance of *Xp* was negatively influenced by skewness of surface pressure and kurtosis of relative humidity, suggesting that frequent changes in surface pressure and relative humidity from normal levels may be less conducive for pathogen multiplication. Atmospheric pressure, though often overlooked, is a significant factor that can influence other climatic factors such as temperature. The atmospheric pressure might lead to changes in oxygen and carbon dioxide partial pressures, which can lead to physiological and growth impacts for plants (Armarego-Marriott 2021). Atmospheric pressure is also positively associated with temperature suggesting increase in air pressure results in increase of temperature. Studies on common dandelion (*Taraxacum officinale*) have revealed reduced plant growth under lower atmospheric pressure and higher production of defense metabolites under higher pressure (Arce et al. 2021). The negative interaction of the surface pressure and *Xp* abundance might be because of increased defense response with the increase in surface pressure. Similarly, higher kurtosis of relative humidity was significantly associated with a lower abundance of *Xp*. Higher kurtosis of relative humidity values indicates that periods of particularly high or low humidity (extreme) negatively affect the abundance of *Xp*. *Xp* population is positively associated with high humidity and temperatures, commonly found in tropical and subtropical regions (Abrahamian et al. 2021; Obradovic et al. 2008). Humidity also alters the interspecies interaction in tomatoes infected with *Xp,* giving advantages to other phyllosphere weak colonizers (Sadhukhan et al. 2023). These findings emphasize the importance of considering the extreme values of meteorological parameters and their distributional characteristics when assessing their influence on disease severity and pathogen abundance.

In summary, we observed diversity of the BLS pathogen, *Xp*, in various crop hosts and regions worldwide, characterized by two novel lineages. Our high-resolution metagenomics

approach tracked strain dynamics of *Xp* over three years in Southeastern US tomato fields, revealing 1-5 co-occurring lineages of *Xp*, with disease severity positively correlated with diversity among them. Modelling approach to identify the drivers of BLS disease severity predicts that environmental factors, including photosynthetically active radiation (PAR) and wind direction, significantly influenced disease severity. Similarly, extreme values of atmospheric pressure and relative humidity also played roles in pathogen abundance, highlighting the complex interplay of climate variables in driving disease dynamics. Our results suggest that the culture independent higher resolution method can be employed for disease diagnostics and epidemiological studies. Modeling of pathogen population and disease severity with the climatic variables underscores the importance of considering the distributional characteristics of meteorological parameters when evaluating their impact on disease severity and pathogen population. This understanding is instrumental in developing more precise predictive models and early warning systems for disease outbreaks.

5.5 Acknowledgments

"These data were obtained from the NASA Langley Research Center (LaRC) POWER Project funded through the NASA Earth Science/Applied Science Program."

5.6 References for chapter 5

Abrahamian, P., Klein-Gordon, J. M., Jones, J. B., and Vallad, G. E. 2021. Epidemiology, diversity, and management of bacterial spot of tomato caused by Xanthomonas perforans. Appl. Microbiol. Biotechnol. 105:6143–6158

Anderson, R. M., and May, R. M. 1982. Coevolution of hosts and parasites. Parasitology. 85:411–426

Arce, C. C. M., Bont, Z., Machado, R. A. R., Cristaldo, P. F., and Erb, M. 2021. Adaptations and responses of the common dandelion to low atmospheric pressure in high-altitude environments. J. Ecol. 109:3487–3501

Armarego-Marriott, T. 2021. Plants under pressure. Nat. Clim. Change. 11:719–719

Barrett, L. G., and Heil, M. 2012. Unifying concepts and mechanisms in the specificity of plant–enemy interactions. Trends Plant Sci. 17:282–292

Barrett, L. G., Zala, M., Mikaberidze, A., Alassimone, J., Ahmad, M., McDonald, B. A., et al. 2021. Mixed infections alter transmission potential in a fungal plant pathogen. Environ. Microbiol. 23:2315–2330

Bartoli, C., Roux, F., and Lamichhane, J. R. 2016. Molecular mechanisms underlying the emergence of bacterial pathogens: an ecological perspective. Mol. Plant Pathol. 17:303–310

Bhandari, R., Sanz-Saez, A., Leisner, C. P., and Potnis, N. 2023. Xanthomonas infection and ozone stress distinctly influence the microbial community structure and interactions in the pepper phyllosphere. ISME Commun. 3:1–13

Bhargawa, A., and Singh, A. K. 2019. Solar irradiance, climatic indicators and climate change – An empirical analysis. Adv. Space Res. 64:271–277

Buckling, A., and Brockhurst, M. A. 2008. Kin selection and the evolution of virulence. Heredity. 100:484–488

Burdon, J. J., and Thrall, P. H. 2009. Co-evolution of plants and their pathogens in natural habitats. Science. 324:755–756

Carvalho, T., Medina, D., P. Ribeiro, L., Rodriguez, D., Jenkinson, T. S., Becker, C. G., et al. 2023. Coinfection with chytrid genotypes drives divergent infection dynamics reflecting regional distribution patterns. Commun. Biol. 6:1–10

Chakraborty, S., Tiedemann, A. V., and Teng, P. S. 2000. Climate change: potential impact on plant diseases. Environ. Pollut. 108:317–326

Chaloner, T. M., Gurr, S. J., and Bebber, D. P. 2021. Plant pathogen infection risk tracks global crop yields under climate change. Nat. Clim. Change. 11:710–715

Cohen, S. P., and Leach, J. E. 2020. High temperature-induced plant disease susceptibility: more than the sum of its parts. Curr. Opin. Plant Biol. 56:235–241

Cribari-Neto, F., and Zeileis, A. 2010. Beta Regression in R. J. Stat. Softw. 34:1–24 Available at: https://doi.org/10.18637/jss.v034.i02

Dimitriu, T., Lotton, C., Bénard-Capelle, J., Misevic, D., Brown, S. P., Lindner, A. B., et al. 2014. Genetic information transfer promotes cooperation in bacteria. Proc. Natl. Acad. Sci. 111:11103–11108

Dixon, P. 2003. VEGAN, a package of R functions for community ecology. J. Veg. Sci. 14:927–930

Dudney, J., Willing, C. E., Das, A. J., Latimer, A. M., Nesmith, J. C. B., and Battles, J. J. 2021. Nonlinear shifts in infectious rust disease due to climate change. Nat. Commun. 12:5102

Dutta, B., Gitaitis, R., Sanders, H., Booth, C., Smith, S., and Langston, D. B. 2014. Role of Blossom Colonization in Pepper Seed Infestation by *Xanthomonas euvesicatoria*. Phytopathology. 104:232–239

Elmer, W. H. 2001. Seeds as vehicles for pathogen importation. Biol. Invasions. 3:263–271.

Ferrari, S., and Cribari-Neto, F. 2004. Beta Regression for Modelling Rates and Proportions. J. Appl. Stat. 31:799–815

Fonti, V., and Belitser, E. 2017. Feature selection using lasso. VU Amst. Res. Pap. Bus. Anal. 30:1–25

Frolking, S. E., Bubier, J. L., Moore, T. R., Ball, T., Bellisario, L. M., Bhardwaj, A., et al. 1998. Relationship between ecosystem productivity and photosynthetically active radiation for northern peatlands. Glob. Biogeochem. Cycles. 12:115–126

Hamza, A. A., Robène-Soustrade, I., Jouen, E., Gagnevin, L., Lefeuvre, P., Chiroleu, F., et al. 2010. Genetic and Pathological Diversity Among *Xanthomonas* Strains Responsible for Bacterial Spot on Tomato and Pepper in the Southwest Indian Ocean Region. Plant Dis. 94:993–999

He, D., Zhan, J., and Xie, L. 2016. Problems, challenges and future of plant disease management: from an ecological point of view. J. Integr. Agric. 15:705–715

Horsfall, J. G., and Barratt, R. W. 1945. An improved grading system for measuring plant diseases. Phytopathology. 35:655

Hu, L., and Yang, L. 2019. Time to Fight: Molecular Mechanisms of Age-Related Resistance. Phytopathology®. 109:1500–1508

Inglis, R. F., Gardner, A., Cornelis, P., and Buckling, A. 2009. Spite and virulence in the bacterium Pseudomonas aeruginosa. Proc. Natl. Acad. Sci. 106:5703–5707

Jones, J. B., Lacy, G. H., Bouzar, H., Stall, R. E., and Schaad, N. W. 2004. Reclassification of the Xanthomonads Associated with Bacterial Spot Disease of Tomato and Pepper. Syst. Appl. Microbiol. 27:755–762

Karasov, T. L., Horton, M. W., and Bergelson, J. 2014. Genomic variability as a driver of plant-pathogen coevolution? Curr. Opin. Plant Biol. 18:24–30

Kim, J. H., Castroverde, C. D. M., Huang, S., Li, C., Hilleary, R., Seroka, A., et al. 2022. Increasing the resilience of plant immunity to a warming climate. Nature. 607:339–344

Kinnula, H., Mappes, J., and Sundberg, L.-R. 2017. Coinfection outcome in an opportunistic pathogen depends on the inter-strain interactions. BMC Evol. Biol. 17:77

Klein-Gordon, J. M., Xing, Y., Garrett, K. A., Abrahamian, P., Paret, M. L., Minsavage, G. V., et al. 2021. Assessing Changes and Associations in the Xanthomonas perforans Population Across Florida Commercial Tomato Fields Via a Statewide Survey. Phytopathology. 111:1029–1041.

Krasnov, B. R., Shenbrot, G. I., Khokhlova, I. S., Stanko, M., Morand, S., and Mouillot, D. 2015. Assembly rules of ectoparasite communities across scales: combining patterns of abiotic factors, host composition, geographic space, phylogeny and traits. Ecography. 38:184–197

Laine, A.-L., Burdon, J. J., Dodds, P. N., and Thrall, P. H. 2011. Spatial variation in disease resistance: from molecules to metapopulations. J. Ecol. 99:96–112

Lechenet, M., Dessaint, F., Py, G., Makowski, D., and Munier-Jolain, N. 2017. Reducing pesticide use while preserving crop productivity and profitability on arable farms. Nat. Plants. 3:1–6

Lee, S. H., Goëau, H., Bonnet, P., and Joly, A. 2020. New perspectives on plant disease characterization based on deep learning. Comput. Electron. Agric. 170:105220

Letunic, I., and Bork, P. 2007. Interactive Tree Of Life (iTOL): an online tool for phylogenetic tree display and annotation. Bioinforma. Oxf. Engl. 23:127–128.

Ma, C.-S., Zhang, W., Peng, Y., Zhao, F., Chang, X.-Q., Xing, K., et al. 2021. Climate warming promotes pesticide resistance through expanding overwintering range of a global pest. Nat. Commun. 12:5351

van Maanen, A., and Xu, X.-M. 2003. Modelling Plant Disease Epidemics. Eur. J. Plant Pathol. 109:669–682

McCree, K. J. 1981. Photosynthetically Active Radiation. In *Physiological Plant Ecology I: Responses to the Physical Environment*, Encyclopedia of Plant Physiology, eds. O. L. Lange, P. S. Nobel, C. B. Osmond, and H. Ziegler. Berlin, Heidelberg: Springer, p. 41–55.

McDonald, B. A., and Linde, C. 2002. Pathogen population genetics, evolutionary potential, and durable resistance. Annu. Rev. Phytopathol. 40:349–379

McDonald, B. A., and Stukenbrock, E. H. 2016. Rapid emergence of pathogens in agro-ecosystems: global threats to agricultural sustainability and food security. Philos. Trans. R. Soc. Lond. B. Biol. Sci. 371

Myers, L., Terranova, M., Ferentz, A., Wagner, G., Verdine, G., Szabó, G., et al. 1993. Repair of DNA methylphosphotriesters through a metalloactivated cysteine nucleophile. Science. 261:1164–1167

Newberry, E. A., Bhandari, R., Minsavage, G. V., Timilsina, S., Jibrin, M. O., Kemble, J., et al. 2019. Independent Evolution with the Gene Flux Originating from Multiple Xanthomonas Species Explains Genomic Heterogeneity in Xanthomonas perforans. Appl. Environ. Microbiol. 85:e00885-19

Newberry, E., Bhandari, R., Kemble, J., Sikora, E., and Potnis, N. 2020. Genome-resolved metagenomics to study co-occurrence patterns and intraspecific heterogeneity among plant pathogen metapopulations. Environ. Microbiol. 22:2693–2708

Newbery, F., Qi, A., and Fitt, B. D. 2016. Modelling impacts of climate change on arable crop diseases: progress, challenges and applications. Curr. Opin. Plant Biol. 32:101–109

Obradovic, A., Jones, J. B., Balogh, B., and Momol, M. T. 2008. Integrated Management of Tomato Bacterial Spot. In *Integrated Management of Diseases Caused by Fungi, Phytoplasma and Bacteria*, Dordrecht: Springer Netherlands, p. 211–223.

Olm, M. R., Brown, C. T., Brooks, B., and Banfield, J. F. 2017. dRep: a tool for fast and accurate genomic comparisons that enables improved genome recovery from metagenomes through de-replication. ISME J. 11:2864

Plank, J. E. V. D. 2013. *Plant Diseases: Epidemics and Control*. Elsevier.

Roach, R., Mann, R., Gambley, C. G., Shivas, R. G., and Rodoni, B. 2018. Identification of Xanthomonas species associated with bacterial leaf spot of tomato, capsicum and chilli crops in eastern Australia. Eur. J. Plant Pathol. 150:595–608

Sadhukhan, S., Jacques, M.-A., and Potnis, N. 2023. Influence of co-occurring weakly pathogenic bacterial species on bacterial spot disease dynamics on tomato. :2023.04.25.538297

Schloissnig, S., Arumugam, M., Sunagawa, S., Mitreva, M., Tap, J., Zhu, A., et al. 2012. Genomic variation landscape of the human gut microbiome. Nature. 493:45–50

Schwartz, A. R., Potnis, N., Timilsina, S., Wilson, M., Patané, J., Martins, J., et al. 2015. Phylogenomics of Xanthomonas field strains infecting pepper and tomato reveals diversity in effector repertoires and identifies determinants of host specificity. Front. Microbiol. 6:535

Seherm, H., and Coakley, S. M. 2003. Plant pathogens in a changing world. Australas. Plant Pathol. 32:157–165

Shah, D. A., Paul, P. A., De Wolf, E. D., and Madden, L. V. 2019. Predicting plant disease epidemics from functionally represented weather series. Philos. Trans. R. Soc. B Biol. Sci. 374:20180273

Sharabani, G., Shtienberg, D., Borenstein, M., Shulhani, R., Lofthouse, M., Sofer, M., et al. 2013. Effects of plant age on disease development and virulence of Clavibacter michiganensis subsp. michiganensis on tomato. Plant Pathol. 62:1114–1122 Available at: https://onlinelibrary.wiley.com/doi/abs/10.1111/ppa.12013.

Soubeyrand, S., Held, L., Höhle, M., and Sache, I. 2008. Modelling the Spread in Space and Time of An Airborne Plant Disease. J. R. Stat. Soc. Ser. C Appl. Stat. 57:253–272

Sparks, A. H. 2018. nasapower: a NASA POWER global meteorology, surface solar energy and climatology data client for R [Journal of Open Source Software, vol. 3, no. 30, 1035.

Stukenbrock, E. H., and Bataillon, T. 2012. A Population Genomics Perspective on the Emergence and Adaptation of New Plant Pathogens in Agro-Ecosystems ed. Joseph Heitman. PLoS Pathog. 8:e1002893

Subedi, A., Kara, S., Aysan, Y., Minsavage, G. V., Timilsina, S., Roberts, P. D., et al. 2023a. Draft genome sequences of 11 Xanthomonas strains associated with bacterial spot disease in Turkey. Access Microbiol. 5:acmi000586.v3.

Subedi, A., Minsavage, G. V., Jones, J. B., Goss, E. M., and Roberts, P. D. 2023b. Exploring Diversity of Bacterial Spot Associated Xanthomonas Population of Pepper in Southwest Florida. Plant Dis. :PDIS10222484RE.

Sundin, G. W., and Wang, N. 2018. Antibiotic Resistance in Plant-Pathogenic Bacteria. Annu. Rev. Phytopathol. 56:161–180

Susi, H., Barrès, B., Vale, P. F., and Laine, A.-L. 2015. Co-infection alters population dynamics of infectious disease. Nat. Commun. 6:5975

Thompson, J. N., and Burdon, J. J. 1992. Gene-for-gene coevolution between plants and parasites. Nature. 360:121–125

Tibshirani, R. 1996. Regression Shrinkage and Selection Via the Lasso. J. R. Stat. Soc. Ser. B Methodol. 58:267–288

Timilsina, S., Abrahamian, P., Potnis, N., Minsavage, G. V., White, F. F., Staskawicz, B. J., et al. 2016. Analysis of Sequenced Genomes of *Xanthomonas perforans* Identifies Candidate Targets for Resistance Breeding in Tomato. Phytopathology. 106:1097–1104

Timilsina, S., Jibrin, M. O., Potnis, N., Minsavage, G. V., Kebede, M., Schwartz, A., et al. 2015. Multilocus Sequence Analysis of Xanthomonads Causing Bacterial Spot of Tomato and

Pepper Plants Reveals Strains Generated by Recombination among Species and Recent Global Spread of Xanthomonas gardneri ed. H. Goodrich-Blair. Appl. Environ. Microbiol. 81:1520–1529

Timilsina, S., Pereira-Martin, J. A., Minsavage, G. V., Iruegas-Bocardo, F., Abrahamian, P., Potnis, N., et al. 2019. Multiple Recombination Events Drive the Current Genetic Structure of Xanthomonas perforans in Florida. Front. Microbiol. 10:448

Velásquez, A. C., Castroverde, C. D. M., and He, S. Y. 2018. Plant-Pathogen Warfare under Changing Climate Conditions. Curr. Biol. CB. 28:R619–R634

Vidaver, A. K., and Lambrecht, P. A. 2004. Bacteria as plant pathogens. Plant Health Instr.

Walkowiak, S., Bonner, C. T., Wang, L., Blackwell, B., Rowland, O., and Subramaniam, R. 2015. Intraspecies Interaction of Fusarium graminearum Contributes to Reduced Toxin Production and Virulence. Mol. Plant-Microbe Interactions®. 28:1256–1267

West, S. A., and Buckling, A. 2003. Cooperation, virulence and siderophore production in bacterial parasites. Proc. R. Soc. Lond. B Biol. Sci. 270:37–44

West, S. A., Diggle, S. P., Buckling, A., Gardner, A., and Griffin, A. S. 2007. The Social Lives of Microbes. Annu. Rev. Ecol. Evol. Syst. 38:53–77

Westfall, P. H. 2014. Kurtosis as Peakedness, 1905 - 2014. R.I.P. Am. Stat. 68:191–195.

Whalen, M. C. 2005. Host defence in a developmental context. Mol. Plant Pathol. 6:347–360.

Wichmann, G., Ritchie, D., Kousik, C. S., and Bergelson, J. 2005. Reduced genetic variation occurs among genes of the highly clonal plant pathogen Xanthomonas axonopodis pv. vesicatoria, including the effector gene avrBs2. Appl. Environ. Microbiol. 71:2418–32

Zadoks, J. C. 1971. Systems analysis and the dynamics of epidemics. Phytopahology. 61:600–610

Zhan, J., Mundt, C. C., Hoffer, M. E., and McDonald, B. A. 2002. Local adaptation and effect of host genotype on the rate of pathogen evolution: an experimental test in a plant pathosystem. J. Evol. Biol. 15:634–647

Zhu, A., Sunagawa, S., Mende, D. R., and Bork, P. 2015. Inter-individual differences in the gene content of human gut bacterial species. Genome Biol. 16:82

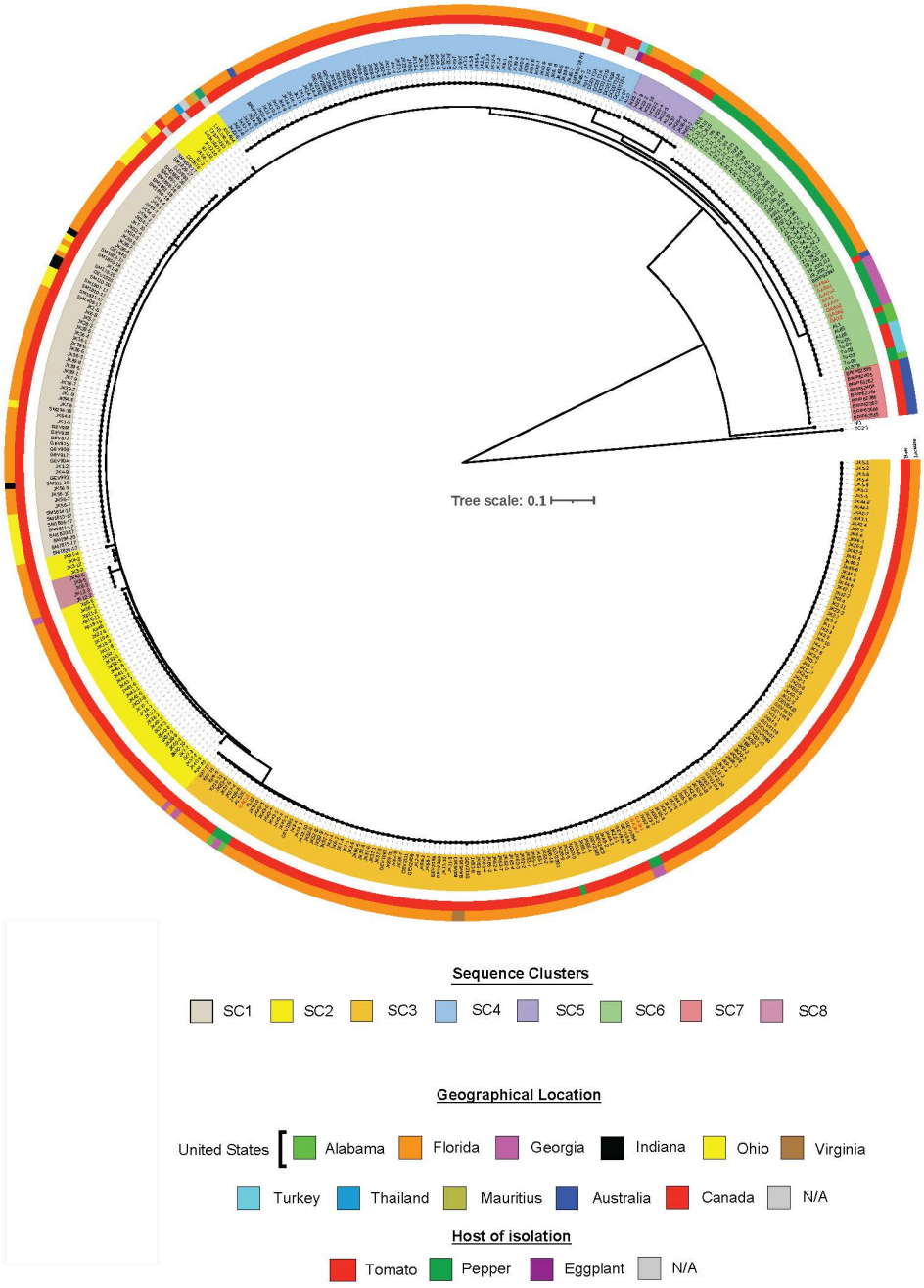

Figure 5-1: Comparative genome analysis demonstrated the presence of eight distinct lineages in *X. perforans*. Midpoint-rooted maximum-likelihood phylogeny of 467 *X. perforans* strains isolated form tomato, pepper, eggplant, and watercress from around the globe based on concatenated alignment of 16,823 core genome SNPs. The clades color coded according to the SCs identified in the first level of the HierBAPS hierarchy. Geographical location and host of isolation of each strain was isolated from, are indicated by different colored blocks within each respective ring surrounding the phylogenetic tree.

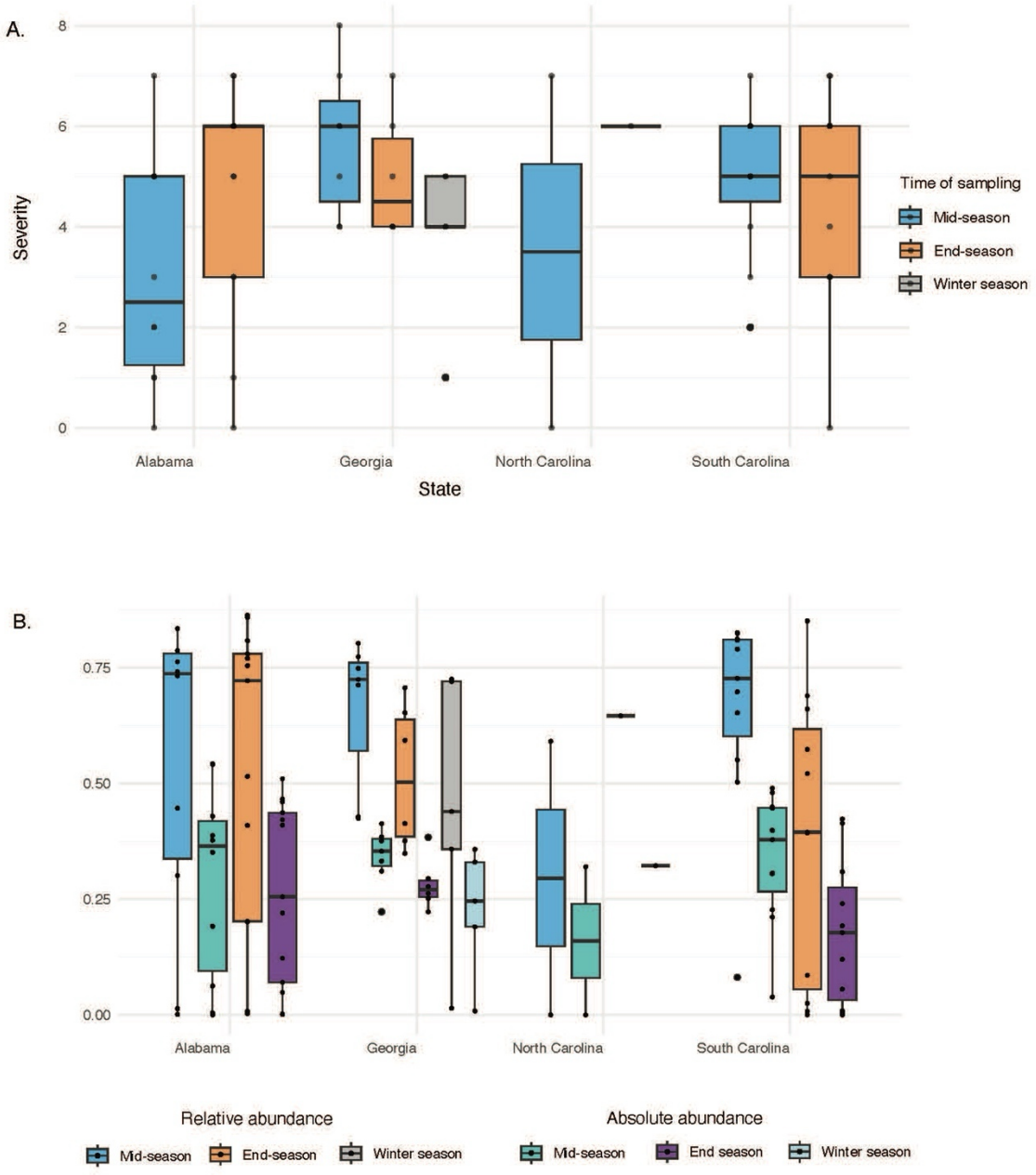

Figure 5-2: *X. perforans* is the dominant pathogen for BLS disease severity in tomato Box plot showing (A) average disease severity of samples (B) Relative and absolute abundance of *X. perforans* collected from different states during 2020, 2021 and 2022. The box is color coded based on their time of sampling.

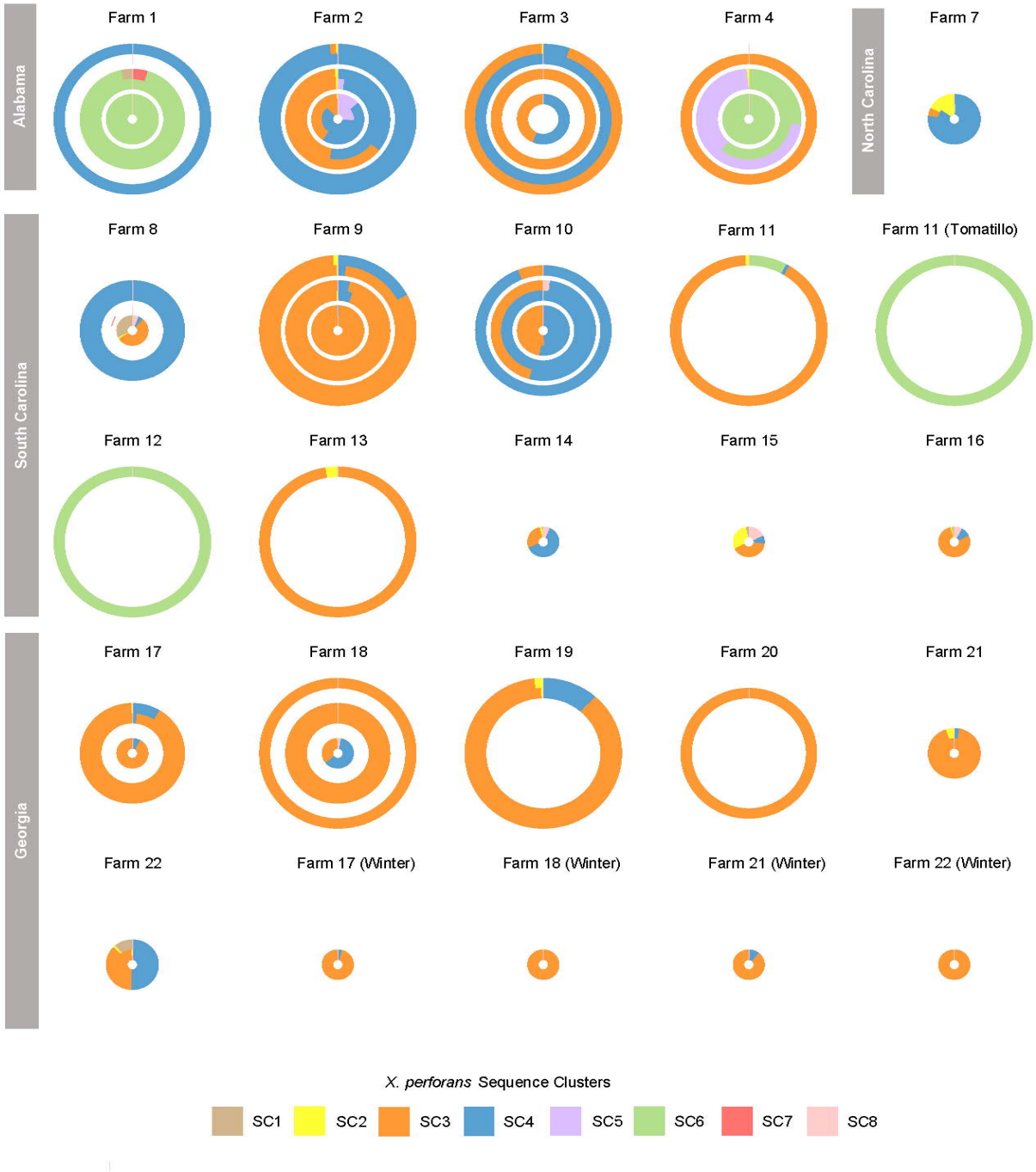

Figure 5-3: Multiple *X. perforans* lineages co-occur with spatio and temporal variations, involving new lineages, strain turnover, and dominance shifts. Stacked donut chart to show the diversity of *Xp* lineages in different farms from Alabama, North Carolina, South Carolina, and Georgia. The inner two donuts circle represents the year 2020 where inner donut within the year 2020 is for mid-season and outer is for end-season. The outer two donuts are for year 2021 and the outermost two are for the year 2022.

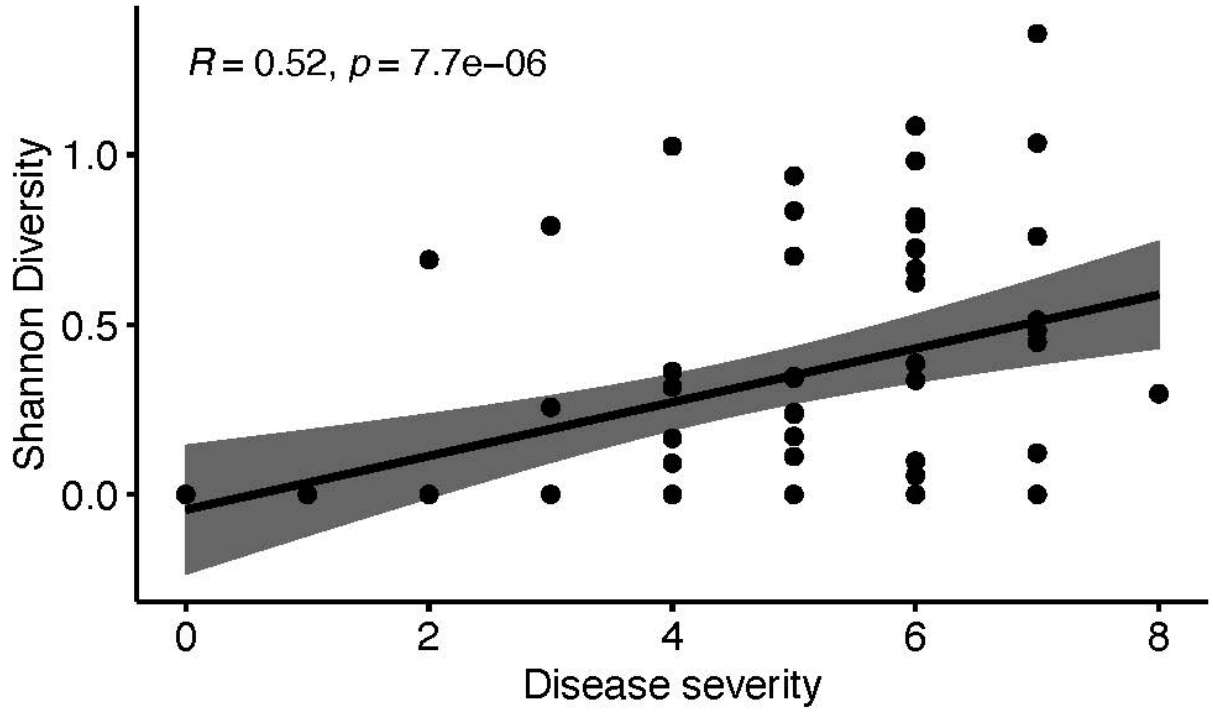

Figure 5-4: **Co-occurrence of more lineages of *X. perforans* in the field results in higher BLS disease severity.** A correlation plot shoeing the interaction between Shannon diversity of *Xp* lineages and disease severity across all samples.

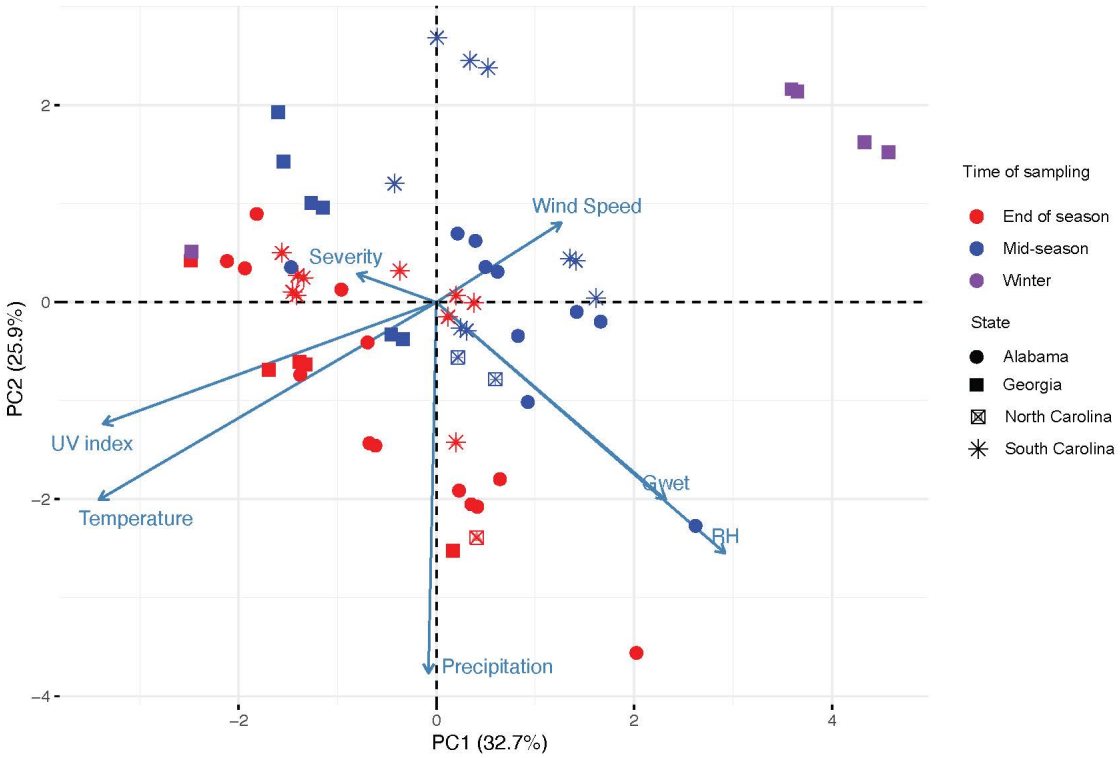

Figure 5-5: Various climatic factors drive *X. perforans* population. Principal coordinate analysis (PCoA) based on the overall structure of *Xp* lineages in all samples. Each data point represents an individual sample. PCoA was calculated using Bray-Curtis distances. Arrows with the weather parameters indicate the direction and magnitude of variables.

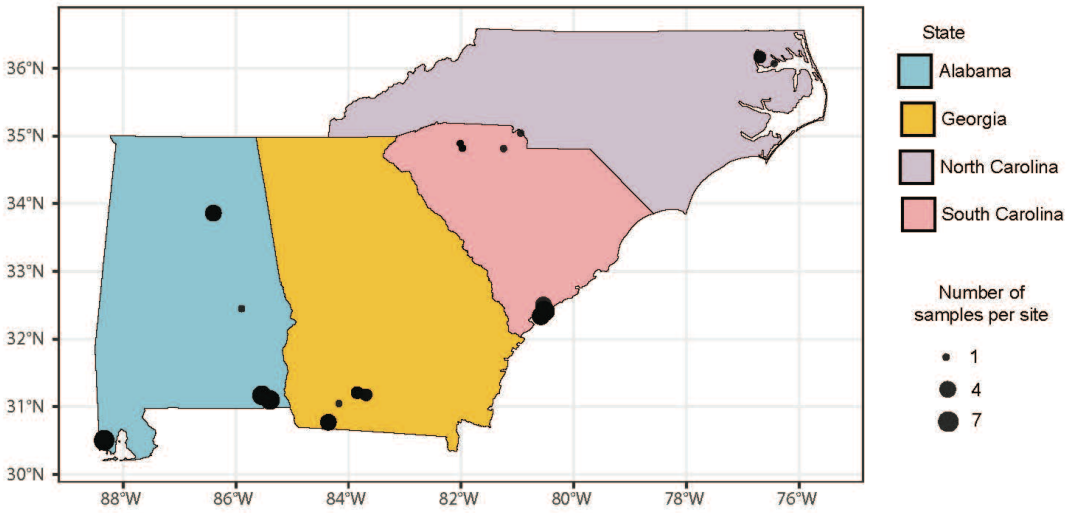

Figure S 5-1: Map of the study area from Alabama, North Carolina, South Carolina, and Georgia from southeast United States. The dots in the figures represents the sampled farm and the size of the dot represents the number of samples collected during 2020, 2021, and 2022.

Table 5-1A: Results of the ordinal regression model showing parameter estimates and associated statistics for disease severity.

Parameters	Value	Std. Error	t value
Av_CLRSKY_SFC_PAR_TOT	0.08753	0.03674	2.3826
Sd_WD10M	0.09293	0.03765	2.4679
Skew_CLRSKY_SFC_PAR_TOT	-0.3143	0.4417	-0.7116
sampled.field.size	0.22126	0.12428	1.7803
Year_2020	1.20354	0.62204	1.9348

Table 5-2: Results of the beta regression model showing parameter estimates and associated statistics for relative abundance of *Xanthomonas perforans* (Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1).

	Estimate	Std. Error	z value	Pr(>|t|)	Remarks
(Intercept)	-4.7595	1.15335	-4.127	3.68E-05	***
Sd_WD10M	0.0391	0.01429	2.736	0.006218	**
Skew_RH2M	0.41746	0.26863	1.554	0.120177	
Skew_PS	-1.37198	0.4113	-3.336	0.000851	***
Kur_RH2M	-0.63686	0.21604	-2.948	0.0032	**
Year_2020	-0.07795	0.22443	-0.347	0.728365	
Year_2022	-0.17922	0.27699	-0.647	0.517621	

Table 5-3: Results of the beta regression model showing parameter estimates and associated statistics for absolute abundance of *Xanthomonas perforans* (Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1).

	Estimate	Std. Error	z value	Pr(>|t|)	Remarks
(Intercept)	-4.7595	1.15335	-4.127	3.68E-05	***
Sd_WD10M	0.0391	0.01429	2.736	0.006218	**
Skew_RH2M	0.41746	0.26863	1.554	0.120177	
Skew_PS	-1.37198	0.4113	-3.336	0.000851	***
Kur_RH2M	-0.63686	0.21604	-2.948	0.0032	**
Year_2020	-0.07795	0.22443	-0.347	0.728365	
Year_2022	-0.17922	0.27699	-0.647	0.517621	

Table S 5-1:Farm details and time of sampling

Farm ID	Year	State	Growing Season Start Date	Mid Sample Date	End Sample Date
Farm 1	2020	AL	4/6/20	5/21/20	7/17/20
Farm 1	2021	AL	4/12/21	5/27/21	6/17/21
Farm 1	2022	AL	4/13/22	5/28/22	6/30/22
Farm 2	2020	AL	4/6/20	5/21/20	6/17/20
Farm 2	2021	AL	4/12/21	5/27/21	6/17/21
Farm 2	2022	AL	4/13/22	5/28/22	6/30/22
Farm 3	2020	AL	5/19/20	7/3/20	7/28/20
Farm 3	2021	AL	5/24/21	7/8/21	7/17/21
Farm 3	2022	AL	5/27/22	7/11/22	8/17/22
Farm 4	2020	AL	4/13/20	5/28/20	7/3/20
Farm 4	2021	AL	4/17/21	6/1/21	7/1/21
Farm 4	2022	AL	4/9/22	5/24/22	6/16/22
Farm 5	2022	AL	5/12/22		8/10/22
Farm 6	2020	NC	5/31/20	7/15/20	7/30/20
Farm 7	2020	NC	5/31/20	7/15/20	8/3/20
Farm 8	2020	SC	4/5/20	5/20/20	6/20/20
Farm 8	2021	SC	4/6/21	5/21/21	6/21/21
Farm 9	2020	SC	4/4/20	5/19/20	6/17/20
Farm 9	2021	SC	4/6/21	5/21/21	6/24/21
Farm 9	2022	SC	4/5/22	5/20/22	6/14/22
Farm 10	2020	SC	4/4/20	5/19/20	6/17/20
Farm 10	2021	SC	5/3/21	6/17/21	7/17/21
Farm 10	2022	SC	4/5/22	5/20/22	6/18/22
Farm 11-1	2022	SC	5/17/22		8/16/22
Farm 11-2	2022	SC	5/17/22		8/16/22
Farm 12	2022	SC	5/17/22		8/16/22

Farm ID	Year	State	Growing Season Start Date	Mid Sample Date	End Sample Date
Farm 13	2022	SC	5/17/22		8/16/22
Farm 14	2020	SC	6/7/20	7/22/20	
Farm 15	2020	SC	6/7/20	7/22/20	
Farm 16	2020	SC	6/7/20	7/22/20	
Farm 17	2020	GA	5/3/20	6/17/20	
Farm 17	2021	GA	5/3/21	6/17/21	7/17/21
Farm 17-W	2020	GA	8/15/21	9/30/20	
Farm 18	2020	GA	5/3/20	6/17/20	
Farm 18-W	2020	GA	8/15/20	9/30/20	
Farm 18	2021	GA	5/3/21	6/17/21	7/17/21
Farm 18	2022	GA	5/4/22		7/6/22
Farm 19	2022	GA	5/1/22	6/15/22	7/7/22
Farm 20	2022	GA	8/15/22		10/6/22
Farm 21	2020	GA	5/12/20	6/26/20	7/26/20
Farm 21-W	2020	GA	8/15/20	9/30/20	
Farm 22	2020	GA	5/12/20	6/26/20	7/26/20
Farm 22-W	2020	GA	8/15/20	9/30/20	

Table S 5-2: Climatic parameters used in the study and their meaning

Parameters	Meaning
T2M	Temperature at 2 Meters (C)
T2MDEW	Dew/Frost Point at 2 Meters (C)
T2MWET	Wet Bulb Temperature at 2 Meters (C)
TS	Earth Skin Temperature (C)
T2M_RANGE	Temperature at 2 Meters Range (C)
QV2M	Specific Humidity at 2 Meters (g/kg)
RH2M	Relative Humidity at 2 Meters (%)
PRECTOTCORR	Precipitation Corrected (mm/day)
CLRSKY_SFC_PAR_TOT	Clear Sky Surface PAR Total (W/m^2)
ALLSKY_SFC_PAR_TOT	All Sky Surface PAR Total (W/m^2)
PS	Surface Pressure (kPa)
WS10M	Wind Speed at 10 Meters (m/s)
WD10M	Wind Direction at 10 Meters (Degrees)
ALLSKY_SFC_UV_INDEX	All Sky Surface UV Index
ALLSKY_SFC_LW_DWN	All Sky Surface Longwave Downward