Understanding the Microbiome of Puget Prairies:
Community composition of bacteria in a hemiparasitic plant system

Victoria G. J. Fox

A thesis
submitted in partial fulfillment of the
requirements for the degree of

Master of Science

University of Washington
2020

Committee:
Jonathan D. Bakker
Sharon L. Doty

Program Authorized to Offer Degree:
Environmental and Forest Sciences

University of Washington

Abstract

Understanding the microbiome of Puget prairies:
community composition of bacteria in a hemiparasitic plant system

Victoria G. J. Fox

Chair of the Supervisory Committee:
Professor Jonathan D. Bakker
School of Environmental and Forest Sciences

Recent advances in the field of metagenomics have allowed for a boom of research in the field of microbial community ecology. Using DNA extraction techniques, Illumina sequencing, and advanced statistical software, scientists are now able to examine the community composition of microbiomes existing throughout the world. My research examines the microbial communities of Puget prairie plants, which have remained largely unexplored until now. I performed a field study to identify the bacterial communities that comprise the stem microbiomes of 16 native prairie plant species. I discovered that the bacterial communities within Puget prairie plants often differ significantly between plant species, but plant species belonging to the same family often have similar bacterial communities. Additionally, I discovered that bacterial communities differed between samples taken from different sampling locations. I also found that bacterial communities are only affected by disturbances applied several years prior to sampling, and in disturbance regimes applied continuously to research plots, for *Cerastium arvense*. I explored the theory that the bacterial community within Puget prairie plants could be influenced by parasitic

root connections established by *Castilleja levisecta*, a hemiparasitic plant that attaches root connections to other prairie plants. Testing all samples that could be assigned to trios regardless of species sample size, I found that plant parasitism significantly affects the bacterial communities of Puget prairie plants overall. Testing individual species with large sample sizes, I found an effect of plant parasitism on the microbiomes of parasitic plant *C. levisecta* for *Eriophyllum lanatum* and *Lomatium utriculatum*, and further study of this system with larger sample sizes could reveal an effect of parasitism for *Balsamorhiza deltoidea* and *Festuca roemeri*. This research provides valuable information about the types of bacteria that exist within the stem tissues of native Puget prairie plants, and insights into the role that parasitic plants may play in the colonization of bacteria across the Puget prairie ecosystem.

Table of Contents

List of Abbreviations

Abbreviation	Explanation
GHP	Glacial Heritage Preserve
ID	Identity
NMDS	Non-Metric Multidimensional Scaling
OTU	Operational Taxonomic Unit
PERMANOVA	Permutational Multivariate Analysis of Variance
SM	Smith Prairie
UFWS	U.S. Fish and Wildlife Service
UMGC	University of Minnesota Genomics Center
UW	University of Washington

List of Figures

Chapter 1

Chapter 2

List of Tables

Chapter 1

Acknowledgements

First and foremost, I am immensely grateful for the steadfast assistance of my advisors, Dr. Jon Bakker and Dr. Sharon Doty. Their combined knowledge of ecological theory and practical hands-on research techniques were instrumental to the success of my thesis. I thank Ellie Reese and Dr. Laura Prugh for providing the Shared Genetics Lab for student use, and offering their unfailing support as I learned to navigate the field of genetics. This study would not have been possible without many lessons from Andrew Sher and Robert Tournay, who patiently guided me through the tricky process of culturing bacteria. The Center for Natural Lands Management and the Pacific Rim Institute not only allowed me to collect my samples, but also continue to protect and steward our priceless prairies, for which we should all be thankful. I thank the members of the Terrestrial Restoration Ecology Lab, especially Jasna Hodzic, Chloe May, Will Braks, Helen Ganahl, and Julianna Hoza for their honesty, creative ideas, and genuine friendship. Of course, my time spent in college would be far less meaningful without the friends and family who saw me through from the beginning to the end of my journey at UW. Special thanks to my parents, Kelly Lam and Michael Fox, and to my friends, Tierra Gogue-Garcia and Jill Aneri Shah, for being there for me when I needed them most. Finally, the following organizations made this work possible with the generous funding they contributed to this research endeavor: Society for Ecological Restoration Northwest Chapter; University of Washington School of Environmental and Forest Sciences Program; Royalty Research Fund; Hall Conservation Genetics Research Fund; Rachel A. Woods Professorship Fund.

Introduction

On March 1st, 2019, the United Nations General Assembly declared 2021-2030 to be the UN Decade on Ecological Restoration (United Nations 2019). This statement from the United Nations, as a respected and powerful international institution, reflects a societal scale recognition of the importance of ecological restoration in the modern era. The concept of ecological restoration has been documented throughout history and has existed for centuries, yet had only just begun to be defined as a practice in the early 1980's (Martin 2017). The Society for Ecological Restoration (SER) is often referred to as the current authority on aspects involving restoration ecology. The SER defines ecological restoration as, "the process of assisting the recovery of an ecosystem that has been degraded, damaged, or destroyed" (Gann et al. 2019). The definition of ecological restoration is contrasted against the definition of restoration ecology, which is described as, "the science that supports the practice of ecological restoration, and from other forms of environmental repair in seeking to assist recovery of native ecosystems and ecosystem integrity" (Gann et al. 2019). While ecological restoration is the process of assisting the recovery of ecosystems, restoration ecology is the science that supports these recovery efforts. The field of restoration ecology has increased in popularity in recent years, and has become a foundational source of information for landscape managers seeking to restore native ecosystems.

Ecosystems that have been threatened and endangered by environmental degradation are of particular focus for ecological restoration. One such threatened system is the Puget prairie ecosystem, which exists in the Pacific Northwest region of the United States. These landscapes occur in Mediterranean climate systems, which experience hot, dry summers and mild, warm winters (Klausmeyer and Shaw, 2009). Puget prairies, also known as South Sound prairies, are

rich in biodiversity, but have declined to less than 10% of their historical range (UFWS 2010). Altered fire regimes, climate change, land use change, invasions of non-native species, and habitat fragmentation, amongst other threats, imperil the survival of Puget prairie ecosystems. Without immediate changes in management, it is likely that Puget prairie ecosystems will continue to decline and these systems may fail to persist into the future (Dunwiddie and Bakker 2011). As a result, natural resource management organizations across the Pacific Northwest have considered prairie ecosystems to be high priority areas for ecological restoration (UFWS 2010).

Puget prairie ecosystems are high in species diversity and host many threatened and endangered plants and animals. These Prairie ecosystems are renowned for their spectacular spring blooms, and are celebrated annually on the second Saturday of May during Prairie Appreciation Day ("Prairie Appreciation Day" 2020). Modern Puget prairie ecosystems are typically comprised of at least 190 native herbaceous plant species, and given the historical decline of these prairies in recent decades, it is expected that many more plant species once occupied these ecosystems (Dunwiddie et al. 2014). *Euphydryas editha taylori* is a butterfly species listed as endangered that frequents Puget prairie ecosystems (UFWS 2010). *Castilleja levisecta* is a perennial plant that inhabits Puget prairies and is listed as a threatened species. *C. levisecta* plants support important pollinators, such as the endangered butterfly *E. editha taylori* (Dunwiddie et al. 2016). Several other butterfly and plant species that occur in Puget prairies are also considered as either candidates for the Endangered Species List or are considered a conservation concern.

Parasitic plants -plants that are able to derive nutrients, energy, and other resources from their host plants- are components of ecosystems found throughout the globe, including Puget prairie ecosystems (Kuijt 1969; Heide-Jørgensen 2008; Westwood et al. 2010). While many

parasitic plants depend entirely on their host plant for resources, others are hemiparasites: plants that are able to both photosynthesize and take up resources from host plants. *Castilleja levisecta* is one such hemiparasitic plant. Relatively little information about community interactions between hemiparasitic plants and their hosts exists in current peer-reviewed literature, demonstrating a gap in scientific knowledge concerning hemiparasites. To fill this gap, the Terrestrial Restoration Ecology Lab at the University of Washington (UW) has studied community interactions between *C. levisecta* and its host plants as a model hemiparasitic plant (Rafay 2018; Dunwiddie et al. 2016; Delvin et al. 2012; Schmidt 2016). Research on *C. levisecta* interactions has been used to assist the recovery of *C. levisecta* populations throughout Western Washington, as this species is listed as threatened under the Endangered Species List (Wentworth 1994; Clark 2015). In previous studies, the Terrestrial Restoration Ecology Lab has identified several potential mechanisms that drive *C. levisect*a growth and reproduction success (Dunwiddie et al. 2016).

One exciting mechanism that may influence hemiparasite performance, and plant performance in general, is the plant microbiome. The microbiome of a plant has been described as an extension of the host genome, as these microbes can have considerable effects on plant protein synthesis, chemical signaling, nutrient acquisition, and other crucial biological processes (Turner et al. 2013; Vandenkoornhuyse et al. 2015; Rho et al. 2018). The microbiome of a plant is comprised of microorganisms that exhibit pathogenic, non-pathogenic, or beneficial traits that influence the growth of the plant in which it lives. Non-pathogenic and beneficial bacteria and fungi that live within plant tissues are referred to as endophytes, and many endophytes are known to have plant growth promoting properties (Glick 2012).

Ecologists and land managers aiming to restore Puget prairie ecosystems have conducted research experiments on Puget prairies for decades, accruing a wealth of information on subjects such as plant and animal species composition, applications of land management techniques, interspecific interactions, the effects of land use change, and future projections for Puget prairie ecosystems, among other studies (Bachelet et al 2001; Stanley et al. 2011; Delvin 2013; Klausmeyer and Shaw 2009; Dunwiddie and Bakker 2011). However, there remains a lack of knowledge of community interactions on smaller scales; the microbial ecology of Puget prairie ecosystems remains an understudied aspect of these systems. Bacterial endophytes and pathogens have been detected in the plant tissues of every plant ever surveyed for the presence of bacteria (Afzal et al. 2019). These bacteria are known to have complex interactions with their hosts and with the other microbes that share inner plant tissues, and can have profound effects on the health of individual plants. Thus, it is of critical importance to understand the microbial community of Puget prairie plants.

Bacterial endophytes are species of bacteria that are able to colonize and exist within plant tissues without causing disease. Common definitions of an endophyte generally include bacteria, fungi, and other microorganisms that inhabit plant tissues. Bacteria that negatively impact plants are considered to be pathogens and do not fall within the definition of endophytes; instead, endophytes either have neutral or positive effects on their host plants. Bacterial endophytes that are able to sustain and supplement plant physiological processes are considered to have plant growth promoting traits. Plant growth promoting traits encompass a diverse array of properties, including nutrient provisioning, nutrient solubilization, disease resistance, modulation of phytohormone levels, and production of cytokinins, among other direct and indirect mechanisms (Glick 2012). Nutrient deficiencies, water limitations, and pathogenic

bacteria induce stress in plants that can be ameliorated by plant growth promoting bacteria (Mei and Flinn 2010). Bacterial endophytes that promote plant growth in their host plants are of particular research interest for their potential application to the fields of agriculture, horticulture, and restoration ecology.

As well as promoting plant growth by producing hormones and acquiring nutrients, bacterial endophytes can also promote plant growth by competing with pathogenic bacteria. Losses in crop yield of every agricultural product can be attributable to plant pathogens, many of which are bacterial plant pathogens. While advances in biological pathogen resistance methods have led to recoveries in crop yields, pesticides and artificial fertilizers are still the most prevalent mechanisms used to protect crop yields. However, bacterial endophytes have also been developed for use as biological control agents to reduce the spread of pathogenic bacteria, since bacterial endophytes occupy similar niches within plant tissues as pathogenic bacteria (Ryan 2008; Compant et al. 2005). Competition for space and substrates contained within plant tissues and production of anti-bacterial compounds are the primary ways in which plant growth promoting endophytic bacteria are able to limit pathogenic bacteria living within plant tissues (Compant et al. 2005).

Bacterial endophytes are known to occur in an extensive number of plant species, and have been recorded living in the space between cells within plant stem, leaf, and root tissues. Many of these stem and leaf inhabiting bacterial endophyte species have been determined to have nutrient provisioning plant growth promoting traits (Hardoim et al. 2008). The microbiome that can be found within the stem and leaf tissue is typically less diverse than that of root tissue, and generally hosts a smaller abundance of bacteria than root tissue (Zhang et al. 2019; Liu, 2017). The majority of bacterial endophytes discovered within plant tissues are derived from the

surrounding environment, since vertical transmission (transmission of bacteria from parent plant to seed) is selective in the type and amount of bacteria that colonize the seed (Walitang et al. 2018). The rhizosphere acts as a main contact zone for root inhabiting endophytes (Yan et al. 2016). Exposed entrances to inner plant tissues, such as stomata or wounds, allow both pathogenic and endophytic bacteria to colonize the intercellular space (Frank et al. 2017).

In this study, I focus specifically on bacteria (including endophytes) that inhabit the plant stem, as these bacteria are thought to readily disperse throughout the plant via xylem and phloem (Frank et al. 2017). When root targeting hemiparasites like *C. levisecta* attach to a host plant, they form haustoria. Haustoria are specialized root connections that facilitate the movement of xylem solute from the host plant to the hemiparasitic plant (Yoshida et al. 2016). Bacteria that disperse throughout a host plant via xylem may be able to use xylem connections between hemiparasitic *C. levisecta* and its host plant to travel between these plants. Furthermore, hemiparasites generally have reduced root systems, and thus are less likely than non-parasitic plants to acquire bacteria from the rhizosphere. As a result, host plants may have significant influence over the number and type of bacteria that colonize the hemiparasite, and thus impact the "extended genome" of hemiparasitic plants.

However, even if no evidence of bacterial transfer was found, the information derived from this study generates important fundamental knowledge about the microbiomes of numerous plant species that have never been studied in this fashion. Characterizing the bacterial taxa that can be found to naturally occur within *C. levisecta* and its hosts stems improves our understanding of this threatened species and its ecological interactions. A greater understanding of the microbiome within a healthy *C. levisecta* population may help us improve the resilience of less healthy remaining populations and contribute to the recovery and delisting of this threatened

species. Just as restoration project managers often use reference ecosystems to determine what flora and fauna should characterize their restoration site, restoration project managers can use reference bacterial taxa information to shape their restoration management techniques (Gann et al. 2019).

The objective of my field study was twofold. First, I intended to characterize the microbial communities that exist in 16 Puget prairie plant species. These bacterial communities have never been examined using non-culture dependent techniques, and thus this research serves as the first Illumina based investigation of bacteria existing in the stems of these 16 Puget prairie plants. Additionally, I intended to discover if bacterial communities arrange themselves in particular patterns across plant species. I theorized that plant samples derived from the same species would likely share similar bacterial Operational Taxonomic Init (OTU) compositions. I also theorized that microbial communities may differ between plant species, as plants are likely to have coevolved with certain species of beneficial endophytes and pathogens and thus associate more often with some bacteria over others. Additionally, I investigated if bacterial communities of Puget prairie plants differ based on the type of disturbance treatment that is applied to research plots. I theorized that the different soil conditions generated by different restoration treatments would create unique challenges and opportunities for bacteria, and thus bacterial communities even within the same plant species may differ based on disturbance treatment.

The second objective of my field study was to investigate if hemiparasitic plants can exchange bacteria with their host plants. Haustorial root connections may provide a pathway for bacteria to travel between hemiparasitic *Castilleja levisecta* and its various host plant species. Therefore, I expected that the microbiomes of host and hemiparasitic plants connected via haustoria would more closely resemble each other than the microbiomes of these same plant

species where parasitism via haustoria does not occur. The application of microbiome genetic analysis to restoration ecology will develop the field of conservation genetics in a novel research direction, providing all three fields with crucial information that can be used to help preserve the health of threatened and endangered plant communities.

Literature Cited

Afzal, I., Z. K. Shinwari, S. Sikandar, S. Shahzad. 2019. Plant beneficial endophytic bacteria: Mechanisms, diversity, host range and genetic determinants. Microbiological Research, 221, 36–49. doi: 10.1016/j.micres.2019.02.001

Bachelet, D., B. R. Johnson, S. D. Bridgham, P. V. Dunn, H. E. Anderson, and B. M. Rogers. 2011. Climate change impacts on western Pacific Northwest prairies and savannas. Northwest Science 85:411-429.

Carthey, A. J. R, D. T. Blumstein, R. V. Gallagher, S. G. Tetu, and M. R. Gillings. 2020. Conserving the holobiont. Functional Ecology, Vol. 34, iss. 4. Pp. 764-776.

Clark, L. A. 2015. Bee-crossed Lovers and a Forbidden Castilleja Romance: Cross-breeding between C. hispida and endangered C. levisecta in prairie restoration sites. Master of Science thesis, University of Washington, Seattle.

Compant, S. B. Duffy, J. Nowak, C. Clement, E. A. Barka. 2005. Use of Plant Growth-Promoting Bacteria for Biocontrol of Plant Diseases: Principles, Mechanisms of Action, and Future Prospects. Applied and Environmental Microbiology, ;71(9):4951-9

Delvin, E. G., J. D. Bakker, P. W. Dunwiddie. 2012. Investigating the role of host plants in recovering golden paintbrush (*Castilleja levisecta*). 97th ESA Annual Convention 2012 Conference Paper.

Dunwiddie, P. W. and J. D. Bakker. 2011. The Future of Restoration and Management of Prairie-Oak Ecosystems in the Pacific Northwest. Northwest Science 85:83–92.

Dunwiddie, P. W., N. L. Haan, M. Linders, J. D. Bakker, C. Fimbel, T. B. Thomas. 2016. Intertwined Fates: Opportunities and Challenges in the Linked Recovery of Two Rare Species. Natural Areas Journal, 36(2):207-215.

Dunwiddie, P. W., R. A. Martin, E. R. Alverson. 2014. Annual Species in Native Prairies of South Puget Sound, Washington. Northwest Science 88(2):94-105. DOI: 10.3955/046.088.0205

Frank, A. C., J. P. Guzman, J. E. Shay. 2017. Transmission of Bacterial Endophytes. Microorganisms, 10;5(4).

Gann, G. D., T. McDonald, B. Walder, J. Aronson, C. R. Nelson, J. Jonson, J. G. Hallett, C. Eisenberg, M. R. Guariguata, J. Liu, F. Hua, C. Echeverria, E. Gonzales, N. Shaw, K. Decleer, K. W. Dixon. 2019. International Principles and standards for the practice of ecological restoration. Second Edition. Restoration Ecology, Vol. 27, Iss. S1.

Glick, B. R. 2012. Plant Growth-Promoting Bacteria: Mechanisms and Applications. Scientifica, vol. 2012, pp. 1–15., doi:10.6064/2012/963401.

Hardoim, P. R., L. S. Van Overbeek, and J. D. Elsas. 2008. Properties of bacterial endophytes and their proposed role in plant growth. *Trends Microbiol* 16: 463– 471.

Heide-Jørgensen, H.S. 2008. Parasitic Flowering Plants. Brill, Leiden, The Netherlands. 421 p.

Klausmeyer K. R. and M. R. Shaw. 2009. Climate change, habitat loss, protected areas, and the climate adaptation potential of species in Mediterranean ecosystems worldwide. PLoS ONE, 4(7), e6392

Kuijt, J. 1969. The Biology of Parasitic Flowering Plants. University of California Press, Berkeley, CA, USA.

Liu, H., L. C. Carvalhais, M. Crawford, E. Singh, P. G. Dennis, C. M. J. Pieterse and P. M. Schenk. 2017. Inner Plant Values: Diversity, Colonization and Benefits from Endophytic Bacteria, *Frontiers in Microbiology,* 10.3389

Martin, D. M. 2017. Ecological restoration should be redefined for the twenty-first century. Restoration Ecology, Volume 21, Iss. 5. https://doi.org/10.1111/rec.12554

Mei, C. and B. Flinn. 2010. The Use of Beneficial Microbial Endophytes for Plant Biomass and Stress Tolerance Improvement. Recent Patents on Biotechnology, vol. 4, no. 1, pp. 81–95., doi:10.2174/187220810790069523.

“Prairie Appreciation Day: Explore the South Sound Prairies”. Prairie Appreciation Day, http://www.prairieappreciationday.org/

Rafay, L. S. 2018. Try it with fire and lime: phytochemical responses to prescribed fire, soil amendments, and simulated herbivory. Master of Science thesis, University of Washington, Seattle.

Rho, H., M. Hsieh, S. L. Kandel, J. Cantillo, S. L. Doty, and S-H. Kim. 2018. Do endophytes promote growth of host plants under stress? A meta-analysis on plant stress mitigation by endophytes. Microbial Ecology 75:407-418.

Turner, T. R., E. K. James, P. S. Poole. 2013. The Plant Microbiome. Genome Biology, 14 (6): 209.

United Nations, Department of Public Information. 2019. *New UN Decade on Ecosystem Restoration offers unparalleled opportunity for job creation, food security and addressing climate change*. SC/11018, 1 March 2019. https://www.unenvironment.org/news-and-stories/press-release/new-un-decade-ecosystem-restoration-offers-unparalleled-opportunity

U.S. Fish and Wildlife Service (UFWS). 2010. Recovery plan for the prairie species of western Oregon and southwestern Washington.

Vandenkoornhuyse, P., A. Quaiser, M. Duhamel, A. Le Van, A. Dufresne. 2015. The importance of the microbiome of the plant holobiont. New Phytologist 206:1196.1206.

Walitang, D. I., C. G. Kim, S. Jeon, Y. Kang, T. Sa. 2018. Conservation and transmission of seed bacterial endophytes across generations following crossbreeding and repeated inbreeding of rice at different geographic locations. Microbiology Open, 8(3). 10.1002/mbo3.662

Wentworth, J.B. 1994. The demography and population dynamics of *Castilleja levisecta*, an endangered perennial. Master of Science thesis, University of Washington, Seattle, Washington.

Westwood, J.H., J.I. Yoder, M.P. Timko, C.W. dePamphilis. 2010. The evolution of parasitism in plants. Trends in Plant Science 15:227-235.

Yan Y., E. E. Kuramae, M. de Hollander, P. G. L. Klinkhamer, J. A. van Veen. 2016. Functional traits dominate the diversity-related selection of bacterial communities in the rhizosphere. ISME J., 11 pp. 56-66.

Yoshida, S., S. Cui, Y. Ichihashi, K. Shirasu. 2016. The Haustorium, a Specialized Invasive Organ in Parasitic Plants. Annual Review of Plant Biology, 67:643-67.

Zhang, Q., J. J. Acuña, N. G. Inostroza. 2019. Endophytic Bacterial Communities Associated with Roots and Leaves of Plants Growing in Chilean Extreme Environments. *Sci Rep* **9,** 4950

Chapter 1: Bacterial Composition of Puget Prairie Plants

Abstract

The Puget prairie ecosystem is a charismatic and ecologically important feature of North America's Pacific Northwest ecosystems, but faces mounting threats from land use change, invasion of non-native plant species, and climate change. Solutions to these threats require enhanced knowledge of these systems, and novel approaches to the particular challenges that impede the recovery of prairie ecosystems. Ecological restoration efforts are beginning to develop and implement practices that enhance plant growth by capitalizing on beneficial relationships formed between plants and endophytes. However, the bacterial endophyte community of many plant systems remains unexplored, as well as the ways in which bacterial endophytes may be traveling within these systems.

I performed a field study to identify what bacteria exist in Puget Sound prairie systems, and to investigate if disturbance treatments affect the bacterial community contained within Puget prairie plants. I processed 335 plant stems of 16 different Puget prairie plant species from research plots in Glacial Heritage Preserve and Smith Prairie in Washington State. I extracted bacterial DNA from these samples and sequenced the 16s rRNA gene to identify bacteria existing within these stems. I used Illumina sequencing, CLC Workbench, and R programming technologies to compare the community profile of bacterial Operational Taxonomic Units (OTUs) between species, and to investigate similarities between the bacterial community profiles of hemiparasitic plants and their hosts.

7,365 different bacterial OTUs were identified across 292 plant samples, and nearly half of these OTUs were not previously identified (de-novo OTUs). I discovered that the bacterial communities within Puget prairie plants often differ significantly between plant species, but

often not between plant species belonging to the same family. I also found that there were significant differences in bacterial OTU composition based on sampling location (Glacial Heritage Preserve and Smith Prairie). Finally, I found that these bacterial communities did not consistently reflect disturbances applied several years prior to sampling, nor to disturbance regimes applied continuously to research plots; only *Cerastium arvense* revealed an effect of initial disturbance or continuous disturbance regime. This work provides the first survey of bacterial diversity within plants in the Puget prairie ecosystem and highlights the importance of spatial distance between sampling locations. With further investigation into the identity of these bacterial OTUs, the knowledge gained through this research may one day benefits land managers who assist the recovery of ecosystems containing hemiparasitic plants, as well as land managers applying microbial diversity and community interaction enhancing techniques within restoration sites.

Introduction

Ecologists and land managers aiming to restore Puget prairie ecosystems have conducted research and experiments on Puget prairies for decades, accruing a wealth of information on subjects such as climate conditions, plant species composition, successful land management techniques, and interspecific interactions, among other subjects (Stanley et al. 2011; Delvin 2013; Klausmeyer and Shaw 2009; UFWS 2010; Dunwiddie and Bakker 2011). However, there remains a lack of knowledge of community interactions on smaller scales; the microbial ecology of Puget prairie ecosystems remains an understudied aspect of these systems. Bacterial endophytes and pathogens have been detected in the plant tissues of every plant ever surveyed for their presence (Afzal et al. 2019; Santoyo et al. 2017; Stone et al. 2000). These bacteria have complex interactions with their hosts, and can have profound effects -both positive and negative- on the health of individual plants (Vandenkoornhuyse et al. 2015; Hardoim et al. 2008; Dheilly 2014). Thus, it is of critical importance to understand the microbial community of Puget prairie plants (Carthey et al. 2020).

Common definitions of an endophyte generally include bacteria, fungi, and other microorganisms that inhabit plant tissues (Wani et al. 2015). Specifically, bacterial endophytes are bacteria that are able to colonize and exist within plant tissues without causing disease (Wani et al. 2015). Bacteria that negatively impact plants are considered to be pathogens and do not fall within the definition of endophytes; instead, endophytes either have neutral or positive effects on their host plants. Bacterial endophytes that are able to sustain and supplement plant physiological processes are considered to have plant growth promoting traits (Berg 2009). Plant growth promoting traits encompass a diverse array of properties, including nutrient provisioning, nutrient solubilization, disease resistance, modulation of phytohormone levels, and production of

cytokinins, among other direct and indirect mechanisms of influence (Glick 2012). Nutrient deficiencies, water limitations, and pathogenic bacteria induce stress in plants that can be ameliorated by several different known plant growth promoting bacteria (Mei and Flinn 2010). Bacterial endophytes that promote plant growth in their host plants are of particular research interest for their potential application to the fields of agriculture, plant nurseries, and restoration ecology.

Bacterial endophytes are known to occur in an extensive number of plant species, and have been recorded living in the space between cells within plant stem, leaf, and root tissues (Afzal et al. 2019). Many of these stem and leaf inhabiting bacterial endophyte species have been determined to have nutrient provisioning plant growth promoting traits (Santoyo et al. 2016). The microbiome that can be found within the stem and leaf tissue is typically less diverse than that of root tissue, and generally hosts a smaller abundance of bacteria than root tissue (Zhang et al. 2019; Liu et al. 2017). The majority of bacterial endophytes discovered within plant tissues are derived from the surrounding environment, since vertical transmission (transmission of bacteria from parent plant to seed) is selective in the species of bacteria that colonize the seed (Walitang et al. 2018). The rhizosphere acts as a main contact zone for root inhabiting endophytes (Yan et al. 2016). Exposed entrances to inner plant tissues, such as stomata or wounds, allow both pathogenic and endophytic bacteria to colonize the intercellular space (Frank et al. 2017).

Soil conditions and chemistry play a critical role in influencing the composition of the plant microbiome (Burns et al. 2015; Yan et al. 2016). As horizontal transmission of bacteria (transmission of bacteria from the surrounding environment to plant tissues) is the most common method of bacterial transfer, the bacterial communities present in the soil largely determine the

species and abundances of bacteria a plant may acquire. Soil pH has been investigated as a factor that influences rhizosphere inhabiting bacterial communities, as certain bacterial biological processes depend on a specific range of pH values; in a study of North American soils, different species of bacteria were found to occupy different ranges in pH values (Lauber et al. 2009). Soil moisture is also a factor that effects the composition of rhizosphere inhabiting bacterial communities; in a study of wheat plants, *Pseudomonas* bacteria were abundant at low and medium soil moisture levels, while *Arthrobacter*, *Bacillus*, and *Cytophaga* were abundant at high soil moisture levels (Peterson et al. 1965). Bacteria are also preferential to different nutrient concentrations in the soil, where bacterial species can be found occupying different nutrient gradients in the soil (Buee et al. 2009). Plants capitalize on this relationship between bacteria and soil nutrient levels by producing root exudates which selectively benefit certain bacteria over others, attracting bacteria which may have plant growth promoting properties (Haichar et al. 2008). In sum, the rhizosphere -which provides plants with many of the bacteria that colonize their inner tissues- is capable of hosting a range of bacterial species which have unique preferences for soil pH, moisture, and nutrient concentrations.

Bacterial communities are not always stable after initial colonization of plant tissues; several communities have been shown to change in abundance and composition from season to season in several studies (Shen and Fulthorpe 2015; Ou et al. 2019). In a study of Mulberry cultivars, bacterial OTU abundance, alpha diversity, and bacterial community complexity were significantly higher for bacterial endophytes collected from branch samples collected in spring than from branch samples collected in autumn (Ou et al. 2019). It is thought that the bacterial community that colonizes plant tissues, particularly above ground plant masses, could be affected by seasonal changes in abiotic environmental conditions such as temperature. Also,

seasonality is theorized to influence bacterial community composition as plant physiology responds to changing seasons, such as changes in the availability of sugars, amino acids, and other crucial nutrients within the plant (Cox and Stushnoff 2001). As the bacterial community changes seasonally, it would follow that disturbances to systems applied in different seasons may influence the bacterial community in different ways.

Bacterial community assemblages are driven by a wide variety of factors, where an individual plant can be thought of as its own "ecosystem" that presents different opportunities and challenges for potential bacterial colonizers. Host plant specificity is one factor that varies between endophyte species; while some endophytes are found to quickly colonize plants where the endophyte has not been known to naturally occur, other endophytes have strong specificity for individual plants, and even to particular organs within a plant (Afzal et al. 2019). In a recent review of the literature, it was found that bacterial communities existing within plants exhibiting different growth patterns interact with tissues differently; in woody plants, stem tissue was rich in bacteria while in graminoids, the roots were the richest tissues (Harrison and Griffin 2020). Within species, plants of different genotypes have been observed to accommodate different bacterial communities (Rodríguez-Blanco et al. 2015). In sum, bacterial interactions with potential host plants are complex, and depend on a variety of biological factors including colonization specificity, host plant identity, genotype, and growth pattern.

While it is recognized that microbes play important roles in many ecosystem functions and have dynamic interactions with plant species, the microbiome of Puget prairie plants remains understudied. A study of bacterial endophytes in plants exposed to PHC's included an analysis of the culturable bacterial endophyte community of *Achillea millefolium*, which revealed the relative proportions of cultured bacteria found in *A. millefolium* plant stems (Lumactud et al.

2016). However, as the *A. millefolium* plants in Lumactud et al. 2016 were collected in Ontario, only surveyed culturable bacteria, and plants were exposed to surface oil deposits, there are likely large differences between *A. millefolium* plants existing in Puget prairie ecosystems that makes this study irreflective of Puget prairie *A. millefolium* bacterial endophytes. So far as I can determine, information on the fungal endophytes of *Festuca roemeri* has been investigated, but not the bacterial microbiome of this species (Bailes et al. 2020). Information on bacteria associated with *Lupine* spp. have been investigated, but commonly with a tight focus on the bacterial endophytes associated with root nodules developed by this leguminous plant (Ferchichi et al. 2019). I was unable to find information on the bacterial microbiome that may comprise any of the other 13 Puget prairie plants that I examined in my study.

To better understand the bacterial communities of the Puget prairie ecosystem, I collected bacterial DNA from 16 different Puget prairie species. First, I wanted to investigate if there were observable differences in the composition of the bacterial OTUs contained within different plant species. Second, I wanted to investigate if large scale differences in sampling location, between Glacial Heritage Preserve and Smith Prairie, generated differences in bacterial OTU compositions between plants of the same species. Third, I wanted to investigate if site disturbance treatments, either applied at the beginning of the restoration treatment study or applied continuously throughout the study, generated differences in bacterial OTU composition between plants of the same species. After reviewing the scientific literature, I posed the following hypothesis about the composition of the bacterial communities of the Puget prairie ecosystem:

H1: Effect of plant species on bacterial community composition. I predicted that bacterial OTU composition will significantly differ between Puget prairie plant species.

H2: Effect of sampling location on bacterial community composition.

H2a: I predicted that bacterial OTU composition will differ between plants of the same species collected at different sampling sites (Glacial Heritage Preserve and Smith Prairie).

H2b: I predicted that bacterial OTU composition will differ within plants of the same species collected from plots that received different initial disturbance treatments.

H2c: I predicted that bacterial OTU composition will differ within plants of the same species that were taken from plots that received different disturbance treatment regimes. I also expected these results to be stronger than the effect of the initial disturbance treatments, as the disturbance regimes were applied to the sites closer in time to plant sampling and subsequent bacterial OTU composition analysis.

Methods

Study Area

I studied two locations in western Washington State (Figure 1.1). The primary study site, from which the majority of the samples were collected, are research plots that had already been established in the Glacial Heritage Preserve (GHP) (Figure 1.2). GHP is owned by Thurston County and the Washington Department of Fish and Wildlife, and managed by the Center for Natural Lands Management. The second study site is at Smith Prairie (SM), on Whidbey Island in Island County. SM is owned and managed by the Pacific Rim Institute for Environmental Stewardship (Figure 1.3). Experimental restoration plots were established at both sites about a decade ago (Figures 1.2, 1.3; Appendix 1.A, Appendix 1.B). Research plots were established for use as restoration experiments in July 2008 and are a part of an ongoing study of Puget prairie restoration. Site preparation and seeding mix differ between plots within the prairie; data on the plot that each plant sample was collected was recorded in the metadata. Plants removed from these plots would not have a detrimental impact on one of the few remaining natural Puget prairies existing in Washington State.

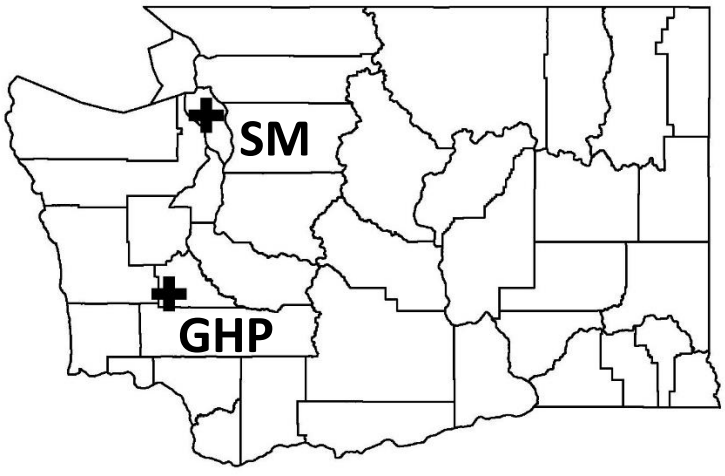


Figure 1.1: Map of Washington State Counties, featuring the locations of Glacial Heritage Preserve (GHP) and Smith Prairie (SM). The Glacial Heritage Preserve is located at 46.8655° N, 123.0537° W. Smith Prairie is located at 48.2043° N, 122.6310° W.

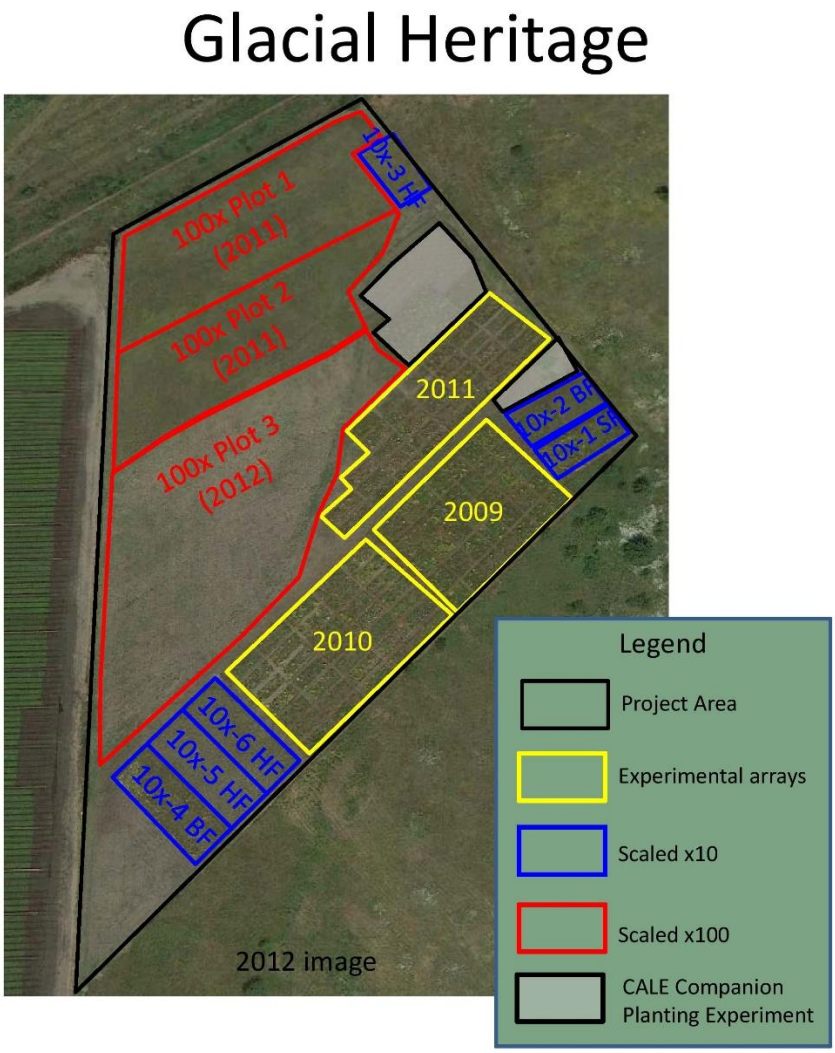


Figure 1.2: Map of Glacial Heritage Preserve research plots. The Collection Site codes used in the metadata refer to this map and the Smith Prairie plot map. A more detailed view of the plots is available in Appendix 1.A.

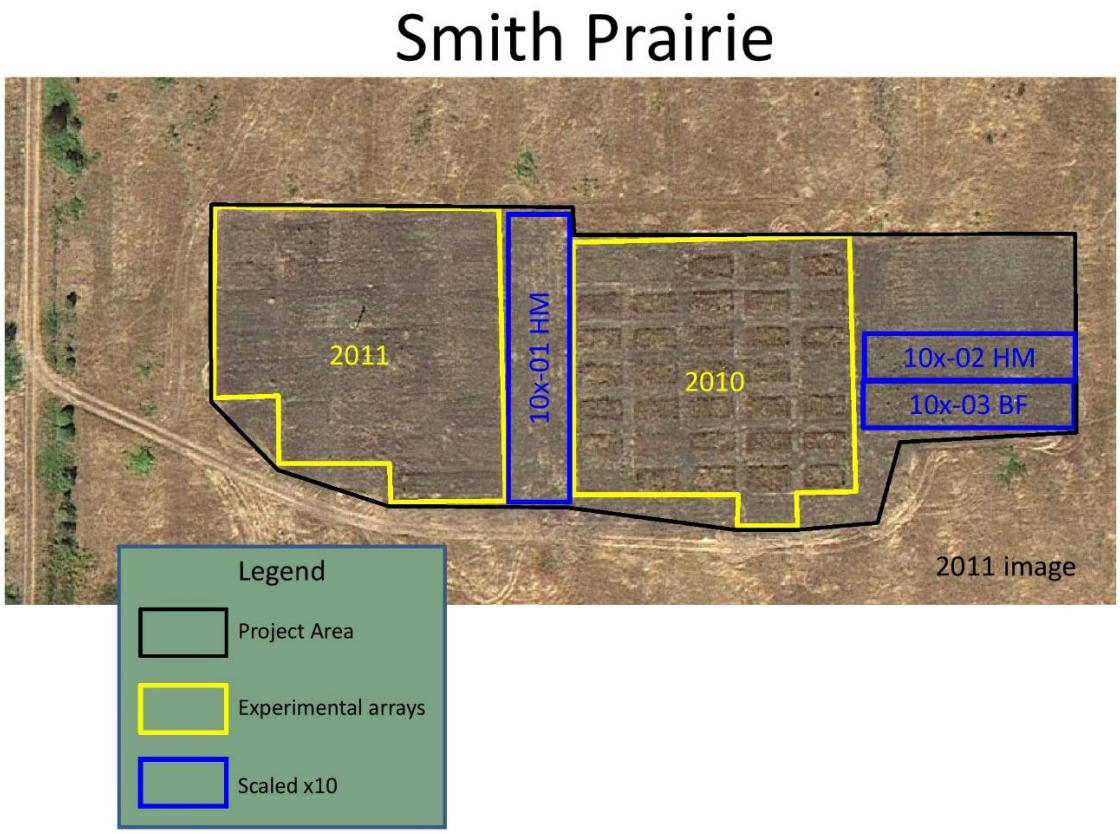


Figure 1.3: Map of Smith Prairie research plots. The Collection Site field of the metadata refer to this map and the Glacial Heritage plot map. A detailed map of the initial site treatment and continuous disturbance regime is available in Appendix 1.B.

Initial disturbance treatments were applied to Glacial Heritage Preserve and Smith Prairie sites in 2009, 2010 and 2011 to examine the prairies response to restoration treatments. Initial disturbance treatments were applied to Glacial Heritage Preserve and Smith Prairie sites in 2009, 2010 and 2011. An array of 35 plots was established at each site in each year, for a total of five arrays (GHP 2009, GHP 2010, GHP 2011, SM 2010, SM 2011). Each plot was 40 m^2 at GHP and 25 m^2 at SM. Three experimental initial disturbance treatments were applied to the plots: solarization, two-year herbicide, and broadcast burning. Beginning in 2014, a fire frequency experiment was overlaid onto the plots at GHP (J.D. Bakker, *unpub. data*). Arrays entered the fire frequency experiment in different years, so plots differ in terms of how long they have been

treated. Five continuing disturbance treatments are being tested: broadcast burned annually in early summer, annually in late summer, triannually in early summer, triannually in late summer, or mowed annually. Detailed maps of the fire frequency treatments applied to each plot within each array are available in Appendix 1.A.

Sample Collection

In May and June 2019, 328 prairie plant stem samples were collected from GHP and 59 samples were collected from SM. Each sample was either a leaf or a stem of a plant, but only stems were used in the set of samples submitted for sequencing. I recorded data on the date the sample was collected, its collection location (site, array, and plot number), and the taxonomic identity of the plant. Plant samples were taken from 16 different prairie plant species (Appendix 1.C).

The sampling process was as follows. Eight trips to the Glacial Heritage Preserve were made throughout the months of May and June. A healthy plant was identified and selected for use in the field (plants with unknown identity were collected and preserved for later identification upon return to Seattle). Each sample was collected by taking a stem cutting of the plant with sterilized scissors, close to where the stem reaches the roots. As much stem material as could fit in one Eppendorf tube was collected. Samples were surface sterilized in the field to remove external bacteria present on the surface of the plant. Surface sterilization was performed by soaking the stem in 70% ethanol for 10 minutes then rinsing the plant in sterile water before placing the stem immediately in a sterile Eppendorf tube. Samples were temporarily preserved for transport in a cooler, and held for long term storage in -20^{o}C in an industrial freezer until they were processed. I attempted to collect at least 25 of samples from each plant species. However, due to the nature of the Puget prairie system, not all plant species occurred in the research plots

in equal numbers. *Erigeron speciosus* and *Symphoricarpos albus* were among the species that were the most difficult to find, and thus I was unable to collect many samples from these species.

Sample Processing

Samples were screened for quality of preservation and relevance for the questions asked. Because there was a budgetary limit to the number of samples that I could sequence, I choose only to sequence samples that were well preserved in sterile conditions and that allowed me to investigate my hypothesis. Plant samples that were stored in cracked Eppendorf tubes, samples that thawed before processing, and samples that were processed under questionably sterile conditions were not selected for sequencing by the UMGC. Additionally, samples from plants that had an abundance of replicates and samples from plant species that did not have enough replicates were not selected for processing or sequencing. The samples that were not selected for processing or sequencing, but that were still preserved in sterile conditions, were prepared for long term storage at -80^{o}C for potential use in future studies. Of the 328 samples that were collected from the Glacial Heritage Preserve, 293 were selected for processing and analysis. Of the 59 samples that were collected from the Smith Prairie, 42 were selected for processing and analysis. A total of 335 plant samples were processed. 13 negative controls ("Blanks") were also submitted for sequencing as controls to check for sterility during processing and sequencing of the plant samples.

Table 1.1: Total number of samples processed from each species. Samples are identified as either derived from Glacial Heritage Preserve or Smith Prairie. The code translation for each species, as well as species taxonomic information, can be found in Appendix 1.C.

Scientific Name	Species Code	Samples Processed GHP	Samples Processed SM	Total
Achillea millefolium	ACMI	18	10	28

Aquilegia formosa	AQFO	15	0	15
Aster curtisii	ASCU	14	0	14
Balsamorhiza deltoidea	BADE	17	0	17
Castilleja levisecta	CALE	40	12	42
Camassia quamash	CAQU	20	0	20
Cerastium arvense	CEAR	30	0	30
Delphinium menziesii	DEME	15	0	15
Eriophyllum lanatum	ERLA	26	9	35
Festuca roemeri	FERO	10	11	21
Lomatum triternatum	LOTR	18	0	18
Lomatium utriculatum	LOUT	20	0	20
Lupinus lepidus	LULE	18	0	18
Potentilla gracilius	POGR	18	0	18
Symphoricarpos albus	SYAL	12	0	12
Blank	BLANK	N/A	N/A	13
TOTAL		**293**	**42**	**335**

Samples were processed throughout September and December 2019. Plant samples were ground into powder by immersing the stems in liquid nitrogen and crushed using sterilized mortars and pestles. Mortars and pestles were only used on one sample per batch, and were washed in hot water and wiped with paper towels soaked in 70% ethanol before being placed in autoclavable plastic bags and sterilized via autoclave after each use. In batches 1 and 2 (removed from analysis due to contamination), mortars and pestles were not autoclaved in plastic bags, and

were instead autoclaved with tin foil sealing the top of the mortars and pestles wrapped in tin foil. I decided to autoclave mortars and pestles in plastic bags after batches 1 and 2 were found to be contaminated, as it was thought that small breaks in the tin foil could have allowed bacteria to contaminate samples from the lab environment. The following procedure was performed as described in the DNeasy PowerSoil Pro Kit Handbook, which accompanies the Qiagen DNeasy PowerSoil Pro Kit (Qiagen 2019). Ground plant samples were placed immediately in a PowerBead Pro Tube containing Solution CD1 and microbeads. Solution CD1 protects nucleic acids from degradation and dissolves humic acids. Samples were vortexed using a Vortex Genie with a horizontal plastic clip microtube holder attachment at maximum speed for 15 minutes, in order to further break down the plant material. Mechanical shaking and chemical agents in Solution CD1 lyse bacterial cells, which can increase DNA yields. The PowerBead Pro Tube was centrifuged at 15,000*g for 1 minute to concentrate the excess plant material and beads to the bottom of the tube.

600 ul of supernatant was transferred to clean 1 ml microcentrifuge tubes, and 200 ul of Solution CD2 was added to the supernatant, then vortexed for 5 seconds. Solution CD1 precipitates non-DNA organic and inorganic material, removing contaminating substances that reduce DNA purity and interfere with downstream DNA applications. The microcentrifuge tubes were centrifuged at 15,000*g for 1 minute to pellet the non-DNA organic and inorganic material. 600 ul of supernatant was transferred to a clean microcentrifuge tube. 600 ul of Solution CD3 was added to the supernatant and vortexed for 5 seconds. Solution CD3 contains a high concentration of salt, which allows for the binding of silica to DNA but not to non-DNA organic and inorganic material. 600 ul of this supernatant and Solution CD3 lysate was added to a silica

membrane containing MB Spin Column and centrifuged at 15,000*g for 1 minute. This allows contaminants to pass through the filter membrane, leaving only DNA bound to the filter.

The flow through was discarded, and an additional 600 ul of the solution -or the rest of the solution if less than 600 ul- was added to the MB Spin Column and centrifuged again at 15,000*g for 1 minute. The flow through was discarded, and the MB Spin Column was placed into a clean 2ml Collection Tube. 500 ul of Solution EA was added to the MB Spin Column and centrifuged at 15,000*g for 1 minute. Solution EA is designed to wash protein and other non-aqueous contaminants from the filter membrane, further purifying the DNA on the filter. The flow through was discarded, and 500 ul of Solution C5 was added to the MB Spin Column then centrifuged at 15,000*g for 1 minute. Solution C5 removes residual salts, humic acids, and other contaminants from the filter membrane. The flow through was discarded and the MB Spin Colum was placed into a clean 2 ml Collection Tube. The MB Spin Column was centrifuged at 16,000*g for 2 minutes, to ensure that all residual solutions were removed from the filter as ethanol contained in Solution C5 can interfere with downstream DNA applications. 100 ul of Solution C6 was added to the MB Spin Column and centrifuged at 15,000*g for 1 minute. Solution C6 contains no salt, which allows the DNA that was bound to the filter to release into solution. The MB Spin Column was discarded, and the 100 ul of DNA extract was stored in an industrial freezer at -20^{o}C until it was needed for further use.

After all samples were processed, the DNA extracts were thawed and DNA concentration was calculated using a NanoDrop Spectrophotometer (ND1000). The NanoDrop Spectrophotometer readouts provide data on the quantity and purity of the nucleic acids present in each sample. A concentration of 1-100 ng/ul was required for sequencing; all samples were quantified, and no samples with lower than 1 ng/ul were present. 30 ul of each extract was

loaded into 96-well plates. Samples were submitted to the University of Minnesota Genomics Center (UMGC) for sequencing.

PCR and Sequencing

Sample extractions were submitted to the UMGC for Dual-Index microbiome amplification and sequencing in September and December 2019. Based on established protocols developed by the Earth Microbiome Project, the sample extractions were sequenced using primers 515F/806R, targeting the hypervariable V4 region of the conserved 16s bacterial ribosomal RNA gene ("16s Illumina Amplicon Protocol", 2018). The 16s gene is frequently used for identification of bacteria in microbiome studies (Patwardhan et al. 2014). mPNA and pPNA blockers were used during PCR to prevent mitochondria and chloroplast from interfering with DNA sequencing (Table 1.2). The UMGC workflow for sample processing is available in Appendix 1.D. The UMGC completed indexing, library preparation and Illumina protocols for sequencing. The Miseq Standard v.3 Chemistry 2x300bp sequencing platform was used to sequence pooled DNA.

Table 1.2: Primer/Blocker Names, Sequences and Purpose. The following primers and blockers were used during PCR by the University of Minnesota Genomics Center.

Primer/Blocker Name	Sequence	Purpose
515F Primer	GTGCCAGCMGCCGCGGTAA	Forward Primer (16s-specific portion)
806R Primer	GGACTACHVGGGTWTCTAAT	Reverse Primer (16s-specific portion
Meta_V4_515F Primer	TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGGTGCCAGCMGCCGCGGTAA	Full Forward Primer sequence
Meta_V4_806R Primer	GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAGGGACTACHVGGGTWTCTAAT	Full Reverse Primer sequence
Forward Indexing Primer	AATGATACGGCGACCACCGAGATCTACAC[i5]TCGTCGGCAGCGTC	Forward Indexing Primer

Reverse Indexing Primer	CAAGCAGAAGACGGCATACGAGAT [i7]GTCTCGTGGGCTCGG	Reverse Indexing Primer
mPNA	GGCAAGTGTTCTTCGGA	Mitochondria Blocker
pPNA	GGCTCAACCCTGGACAG	Chloroplast Blocker
Nextera Adapter Read 1	CTGTCTCTTATACACATCTCCGAGC CCACGAGACNNNNNNNNATCTCGT ATGCCGTCTTCTGCTTG	Sequences used for post-run trimming
Nextera Adapter Read 2	CTGTCTCTTATACACATCTGACGCT GCCGACGANNNNNNNNGTGTAGAT CTCGGTGGTCGCCGTATCATT	Sequences used for post-run trimming

Data Processing

Two fastq files were generated per sample: a pair of forward sequence and reverse sequence reads for each sample. Compressed fastq files (.gz) were retrieved from the UMGC. The raw sequence reads were processed using a genomic pipeline generated in CLC Genomics Workbench 12.0.3, a data analysis package created by Qiagen. The Microbial Genomics Module for CLC Genomics Workbench software is designed to process and analyze "16s rRNA and other commonly used metagenome derived amplicon data." (CLC Microbial Genomics Module User Manual). The Microbial Genomics Module was used to trim, filter, and cluster reads into OTUs. The process for read editing is described below.

First, I uploaded the forward and reverse paired-end Illumina files to Workbench. In the Import wizard, the import type was set to Paired Reads, the minimum distance was set to 200, the maximum distance was set to 550, and quality scores associated with the reads were imported as well. Then, reads with quality scores less than 0.05 were trimmed. The Trim Reads tool was also used to trim ambiguous nucleotides with a maximum number of ambiguities set to 2. Reads shorter than 5 nucleotides in length were discarded.

The processed reads then were clustered into OTUs. Using the OTU Clustering tool, I chose to use the SILVA 16S v132 97% reference database, with the similarity percent specified by the OTU database option selected (Balvočiūtė and Huson 2017). 97% similarity is a standard value for microbial 16s analysis, although it should be noted that recent research has questioned the validity of this value (Stackebrandt and Ebers 2006). The creation of novel OTUs was enabled (Nguyen et al. 2015). An abundance table displaying the number of reads from each OTU discovered in each sample was generated by CLC Workbench and exported as a .csv file to R Studio for further examination. The R script for the following analysis can be found in Appendix 3.

Statistical Analysis

After processing the raw reads through the CLC Workbench genomic pipeline, I performed statistical analysis on my data. R Studio was used to perform the subsequent calculations, data transformations and statistical analysis. A file containing abundance data for each OTU present in each sample was imported to R Studio. An additional file containing metadata for each sample was also imported to R Studio. This file was examined by multiple parties for errors and was cleaned prior to analysis. While rarefaction has been used in previous microbiome studies to normalize abundance data, this technique is no longer recommended for use as reads, and the valuable information they contain, are lost in the process (McMurdie and Holmes 2014). Calle 2019 recommends analyzing microbiome abundance data alongside presence/absence of microbial OTUs within the same dataset, so the OTU abundance table generated in CLC Workbench was used to create a presence/absence table (Calle 2019; Quinn 2018; Weiss 2017).

Each batch of samples (samples that were extracted using the same Qiagen kit on the same day) is associated with a negative control (a “Blank”) that acts as a way to detect potentially contaminating bacterial DNA (Goodrich 2014). Bacterial DNA can contaminate samples by drifting from surfaces and into samples before or during processing. These blanks are processed alongside each batch in an attempt to capture OTU abundances that did not originate from a plant sample. Blanks 1 and 2 captured a large amount of contamination, likely due to improper sterilization techniques used on mortars and pestles. The process to sterilize mortars and pestles was adjusted after Blanks 1 and 2 revealed contamination; instead of autoclaving mortars and pestles in tin foil, they were instead autoclaved in autoclavable plastic bags that were sealed. Blank 1 contained 723 OTUs and 31,007 total reads, while Blank 2 contained 702 OTUs and 27,264 total reads. These values are remarkably high compared to Blanks 3-13 which contained an average of 58 OTUs and 3,354 total reads. The process to sterilize mortars and pestles was adjusted after Blanks 1 and 2 revealed contamination. Blanks 3-13 indicate that contamination was reduced as the total number of OTUs and total read abundance per blank decreased dramatically. Potentially contaminating bacterial OTUs and their respective abundances were used to filter contaminants from the batches of samples. Bacterial OTU reads recorded in each blank were subtracted from their respective batches; OTU abundances from Blank 1 were subtracted from Batch 1, OTU abundances from Blank 2 were subtracted from Batch 2, and so forth. Negative values, where more reads were detected from any particular OTU were discovered in a blank than in a plant sample, were set to zero. OTUs which were not present in a blank were unaffected, and OTUs were only subtracted using their respective blanks. Because of the high prevalence of contamination in Blanks 1 and 2, all samples that were

processed in batches 1 and 2 were excluded from all future analysis. No further transformations of the data were performed (Legendre and Gallagher 2001).

Several distance measures have been suggested for use on metagenomic data. The Bray-Curtis distance measure is commonly used with species composition data, however there are some noteworthy flaws in its application to microbiome data (Calle 2019). Microbiome abundance data is not strictly reflective of true species abundance, thus other distance measures such as the Aitchison distance measure and UniFrac distance measures are commonly recommended in scientific literature over the Bray-Curtis distance measure (Gloor et al. 2017). UniFrac measures have been used prolifically throughout the literature to calculate beta diversity. There are certain disadvantages to using UniFrac distance measures, however. Calle 2019 argues that Unifrac is inappropriate for microbiome data as these measures are not sub-compositionally dominant. Instead, Calle 2019 recommends the use of the Aitchison distance to analyze beta diversity. Given the advantages and disadvantages of these distance measures, the Bray-Curtis distance measure remains a robust statistical measure that continues to be applied in similar research endeavors and was thus chosen for use in this study (Maziarz et al. 2018).

Data characteristics were explored in R Studio using R base code. In some cases, supplementary tables were exported from R and organized in Excel for the production of visuals. Differences in bacterial OTU composition between plant species was tested using PERMANOVA. PERMANOVA can be implemented with multivariate data and in cases where normality cannot be assumed, which is appropriate for use on these microbiome abundance data. PERMANOVA was performed using the adonis() function in the vegan package (Oksanen et al. 2017). Negative controls (Blanks) were removed from the dataset prior to performing PERMANOVA. Number of permutations was set to 999, and the distance measure used to

implement PERMANOVA was the Bray-Curtis distance measure. Alpha was set to a = 0.05. After PERMANOVA, if a significant result was achieved, a pairwise test was performed to determine which plant species differ from other plant species in the composition of their bacterial OTUs. The pairwise test was performed using a modified function pairwise.adonis. The code for this function is included in the R script in Appendix 3.

Differences in bacterial OTU composition between plant species were then visualized using several Non-Metric Multidimensional Scaling (NMDS) ordinations. NMDS plots are recommended over PCoA plots by several review papers describing statistical methods for research on metagenomic data (Calle 2019; Gloor et al. 2017; Ramette 2007). NMDS plots are recommended as the analysis of PCoA plots can be driven by the presence or absence of taxa and by sparsity, which can be problematic when working with metagenomic data (Gloor et al. 2017). NMDS ordinations avoid these problems. Ordinations were created using the ggplot() function within the ggplot2 package (Wickham 2016). The settings for the Species NMDS were set to three dimensions (k) = 3, as the stress for two dimensions was too low and thus would not accurately reflect the data. For the Sites NMDS, CEAR Initial Disturbance NMDS and CEAR Disturbance Regime NMDS, stress was low enough at two dimensions to allow for an accurate reflection of the data, as well as allow for easier interpretation of the ordination. The other settings to generate the NMDS across all ordinations were set to the same parameters: maxit = 300, try = 40, trymax = 100. Ordinations were created using the metaMDS() function in the vegan package and visualized using the ggplot() function in the ggplot2 package (Oksanen et al. 2017; Wickham 2016).

Four potential sources of variation in the dataset are tested in the following analyses. First, differences in bacterial OTU composition between samples derived from different plant

species is investigated with a PERMANOVA test performed on the samples in Table 1.3. The following subset of samples were used to test hypothesis 1, where I expected to find differences in the bacterial OTU composition of samples taken from different plant species. For this analysis, samples derived from Glacial Heritage Preserve and Smith Prairie were both included. After PERMANOVA, pairwise tests were used to determine which species retained significant differences in their bacterial OTU composition. A 3-dimensional NMDS plot was generated to visualize the dataset.

Table 1.3: Total number of samples processed from each species after removal of potentially contaminated samples and outliers. The code translation for each species, as well as species taxonomic information, can be found in Appendix 1.C.

Species	Total
ACMI	27
AQFO	11
ASCU	11
BADE	14
CALE	43
CAQU	14
CEAR	26
DEME	13
ERLA	33
ERSP	7
FERO	19
LOTR	11
LOUT	21

LULE	15
POGR	14
SYAL	12
TOTAL	**292**

Second, differences in bacterial OTU composition within 4 species (*Achillea millefolium*, *Castilleja levisecta*, *Eriophyllum lanatum* and *Festuca roemeri*) taken from Glacial Heritage Preserve and Smith Prairie were compared with two PERMANOVA tests to determine if bacterial OTU composition differed between sites. The following subset of samples were used to test hypothesis 2a, where I expected to find differences in the bacterial OTU composition of samples taken from different sampling locations. The order of terms affects how variation is partitioned – the first term accounts for as much variation as possible and the second as much of the remaining variation as possible – so I conducted tests with terms in both orders. Testing site first and species second is less conservative, whereas testing species first and site second is more conservative. Only *A. millefolium*, *C. levisecta*, *E. lanatum* and *F. roemeri* were taken from Smith Prairie, thus only these four species were used in analysis. A summary of the samples used in this test is available in Table 1.4. A 2-dimensional NMDS plot was created to visualize this dataset.

Table 1.4: Number of samples processed from each species that were taken from either Glacial Heritage Preserve or Smith Prairie. Includes both the absolute number of plants taken from either Glacial Heritage Preserve or Smith Prairie as well as the percentage of the total number of plants that represents each site. The code translation for each species, as well as species taxonomic information, can be found in 1.C.

Species	**GHP**	**SM**	**Total**
ACMI	17 (63%)	10 (37%)	27
CALE	32	12	44

	(73%)	(27%)	
ERLA	25 (74%)	9 (26%)	34
FERO	8 (42%)	11 (58%)	19
TOTAL	**82 (66%)**	**42 (34%)**	**124**

Third, differences in bacterial OTU composition within species taken from plots with different initial disturbance treatments were compared with a series of PERMANOVA tests. The following subset of samples were used to test hypothesis 2b, where I expected to find differences in the bacterial OTU composition of samples taken from plots that had received different initial disturbance treatments. Only species taken from Glacial Heritage Preserve were used for these tests. Not all species were collected from all initial disturbance treatments; examine Table 1.5 below for a summary of the initial disturbance treatments associated with each species. Only *Aster curtisii*, *Castilleja levisecta*, *Cerastium arvense*, *Eriophyllum lanatum*, *Erigeron speciosus*, *Festuca roemeri* and *Lomatium triternatum* had enough representative samples from all three treatments to be suitable for use in these tests. Only samples taken from sites within GHP 2009, GHP 2010, and GHP 2011 were used for initial disturbance treatment tests; samples taken from scaled up plots and the mounded prairie were excluded from analysis.

Table 1.5: Number of samples derived from each species that was taken from plots that received one of three initial disturbance treatments. Includes both the absolute number of plants taken from plots with their respective disturbance treatments as well as the percentage of the total number of plants that represents each initial disturbance treatment. Only those species and samples in bold were used for this analysis. The code translation for each species, as well as species taxonomic information, can be found in Appendix 1.C.

Species	Solarize	Burn	Two Year	Total
ACMI	3 (75%)	0 (0%)	1 (25%)	4
AQFO	2 (67%)	0 (0%)	1 (33%)	3
ASCU	**6**	**3**	**1**	**10**

	(60%)	**(30%)**	**(10%)**	
BADE	3 (75%)	1 (1%)	0 (0%)	4
CALE	**5 (28%)**	**10 (56%)**	**3 (16%)**	**18**
CAQU	4 (57%)	3 (43%)	0 (0%)	7
CEAR	**7 (41%)**	**8 (47%)**	**2 (12%)**	**17**
DEME	1 (33%)	2 (67%)	0	3
ERLA	**5 (62%)**	**2 (25%)**	**1 (13%)**	**8**
ERSP	**3 (43%)**	**3 (43%)**	**1 (14%)**	**7**
FERO	**3 (37%)**	**3 (37%)**	**2 (26%)**	**8**
LOTR	**3 (50%)**	**2 (33%)**	**1 (17%)**	**6**
LOUT	5 (45%)	6 (55%)	0 (0%)	11
LULE	0 (0%)	4 (80%)	1 (20%)	5
POGR	3 (50%)	3 (50%)	0 (0%)	6
SYAL	1 (50%)	1 (50%)	0 (0%)	2
TOTAL	**54 (45%)**	**51 (43%)**	**14 (12%)**	**119**

Finally, differences in bacterial OTU composition within species taken from plots with different disturbance regime treatments were compared with a series of PERMANOVA tests. The following subset of samples were used to test hypothesis 2c, where I expected to find differences in the bacterial OTU composition of samples taken from plots that had received different disturbance regime treatments. Only species taken from Glacial Heritage Preserve were used for these tests. Not all species were collected from all disturbance regime treatments; examine Table 1.6 below for a summary of the initial disturbance treatments associated with

each species. Only *Castilleja levisecta*, *Cerastium arvense*, *Eriophyllum lanatum*, *Festuca roemeri* and *Lomatium utriculatum* had enough representative samples from all 5 treatments to be suitable for use in these tests. Only samples taken from sites within GHP 2009, GHP 2010, and GHP 2011 were used for disturbance regime treatment tests; samples taken from scaled up plots and the mounded prairie were excluded from analysis.

Table 1.6: Number of samples derived from each species that was taken from plots that received one of five disturbance regime treatments. Includes both the absolute number of plants taken from plots with their respective disturbance treatments as well as the percentage of the total number of plants that represents each disturbance regime treatment. Only those species and samples in bold were used for this analysis. The code translation for each species, as well as species taxonomic information, can be found in Appendix 1.C.

Species	Annual Early Burn	Annual Late Burn	Triannual Early Burn	Triannual Late Burn	Annual Mow	Total
ACMI	2 (50%)	0 (0%)	2 (50%)	0 (0%)	0 (0%)	4
AQFO	0 (0%)	1 (33%)	1 (33%)	0 (0%)	1 (33%)	3
ASCU	0 (0%)	3 (30%)	3 (30%)	0 (0%)	4 (40%)	10
BADE	0 (0%)	1 (25%)	3 (75%)	0 (0%)	0 (0%)	4
CALE	**4 (22%)**	**4 (22%)**	**4 (22%)**	**4 (22%)**	**2 (12%)**	**18**
CAQU	3 (44%)	2 (28%)	2 (28%)	0 (0%)	0 (0%)	7
CEAR	**2 (12%)**	**2 (12%)**	**6 (35%)**	**5 (29%)**	**2 (12%)**	**17**
DEME	0 (0%)	0 (0%)	1 (33%)	2 (67%)	0 (0%)	3
ERLA	**2 (25%)**	**1 (13%)**	**1 (13%)**	**3 (36%)**	**1 (13%)**	**8**
ERSP	3 (44%)	0 (0%)	3 (44%)	1 (22%)	0 (0%)	7
FERO	**2 (25%)**	**1 (13%)**	**1 (13%)**	**3 (36%)**	**1 (13%)**	**8**
LOTR	0 (0%)	2 (33%)	3 (50%)	1 (17%)	0 (0%)	6
LOUT	**1 (9%)**	**2 (18%)**	**4 (37%)**	**2 (18%)**	**2 (18%)**	**11**
LULE	1 (20%)	0 (0%)	0 (0%)	1 (20%)	3 (60%)	5
POGR	0	3	0	1	2	6

	(0%)	(50%)	(0%)	(17%)	(33%)	
SYAL	0 (0%)	2 (100%)	0 (0%)	0 (0%)	0 (0%)	2
TOTAL	**20 (17%)**	**24 (20%)**	**34 (29%)**	**23 (19%)**	**18 (15%)**	**119**

Results

Sample Summary

A total of 7,365 unique bacterial OTUs were generated from 335 samples and 13 negative controls. Of these bacterial OTUs, 3,374 were not listed in the SILVA database and were considered de-novo OTUs. The number of bacterial OTUs derived from non-replicate plant samples ranged from 23 OTUs to 590 OTUs per sample. The abundance of reads per sample for plant samples, after removal of potentially contaminating bacterial OTU abundances, ranged from 65 reads/sample to 32,597 reads/sample. The number of OTUs per sample within plant species varied widely, where a sample derived from one plant species could have up to 410 more OTUs than a sample derived from the same plant species. Sample 0258, a *Lomatium utriculatum* sample, had an extraordinarily low count of bacterial OTUs (13 total OTUs) and abundance (23 total reads) after removing potentially contaminating bacteria, and thus sample 0258 was removed from all analysis. Additionally, analysis of samples 0055, 0438 and 0439 identified these samples as outliers, and were thus also removed from the datasets used in all further analysis. A table displaying the total number of OTUs and read abundance per sample is available in Appendix 1.E.

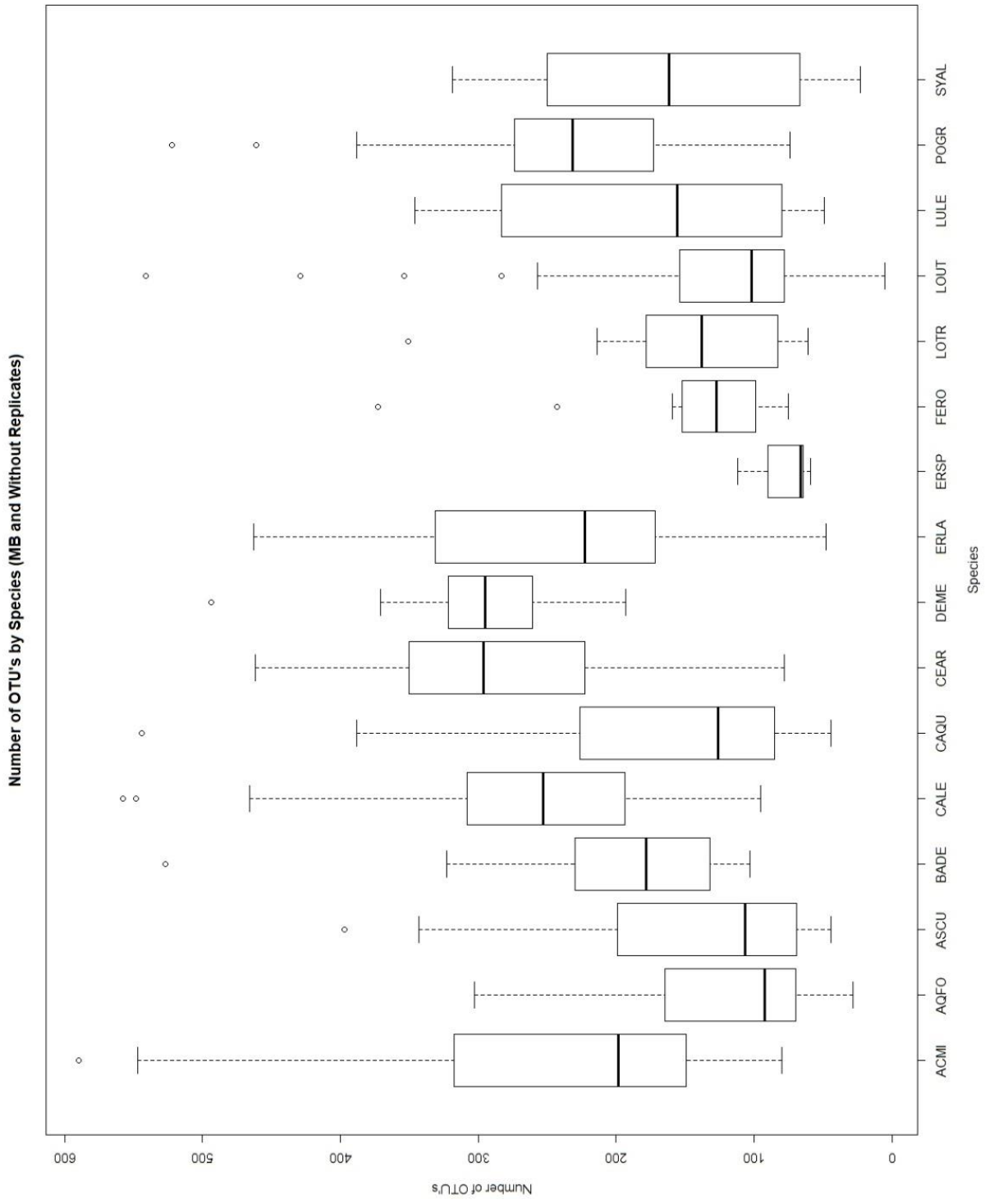


Figure 1.4: Box and whisker plot of the number of bacterial OTUs derived from each species, after removal of potentially contaminating bacterial OTUs. The median is marked by the vertical line inside the box, the upper and lower quartiles are the ends of the box, and the whiskers represent the range of non-outliers. Data points that lie above the whiskers are outliers.

Using the SILVA database, known bacterial OTUs can often be identified to the genus, and sometimes to the species taxonomic level (Balvočiūtė and Huson 2017). Bacterial OTUs that are classified as de-novo cannot be identified to any taxonomic level, as these OTUs cannot be identified to known bacterial 16s rRNA genes in the SILVA database. The OTUs in this dataset represented 26 phyla, 51 classes, 120 orders, 297 families, 694 genera, and 472 species. A count of the number of phyla, classes, orders, families, genera, and species represented by the OTUs is available in Appendix 1.E. The total abundance of bacterial OTUs were averaged across samples within species groups and compiled in a histogram to illustrate the average abundance of bacterial phylum present within each species (Figure 1.5). Actinobacteria, Cyanobacteria, and Proteobacteria contribute largely to the abundance of bacterial OTU reads within nearly all species. Firmicutes and Bacteroidetes occasionally also contribute largely to the abundance of bacterial OTU reads within select species.

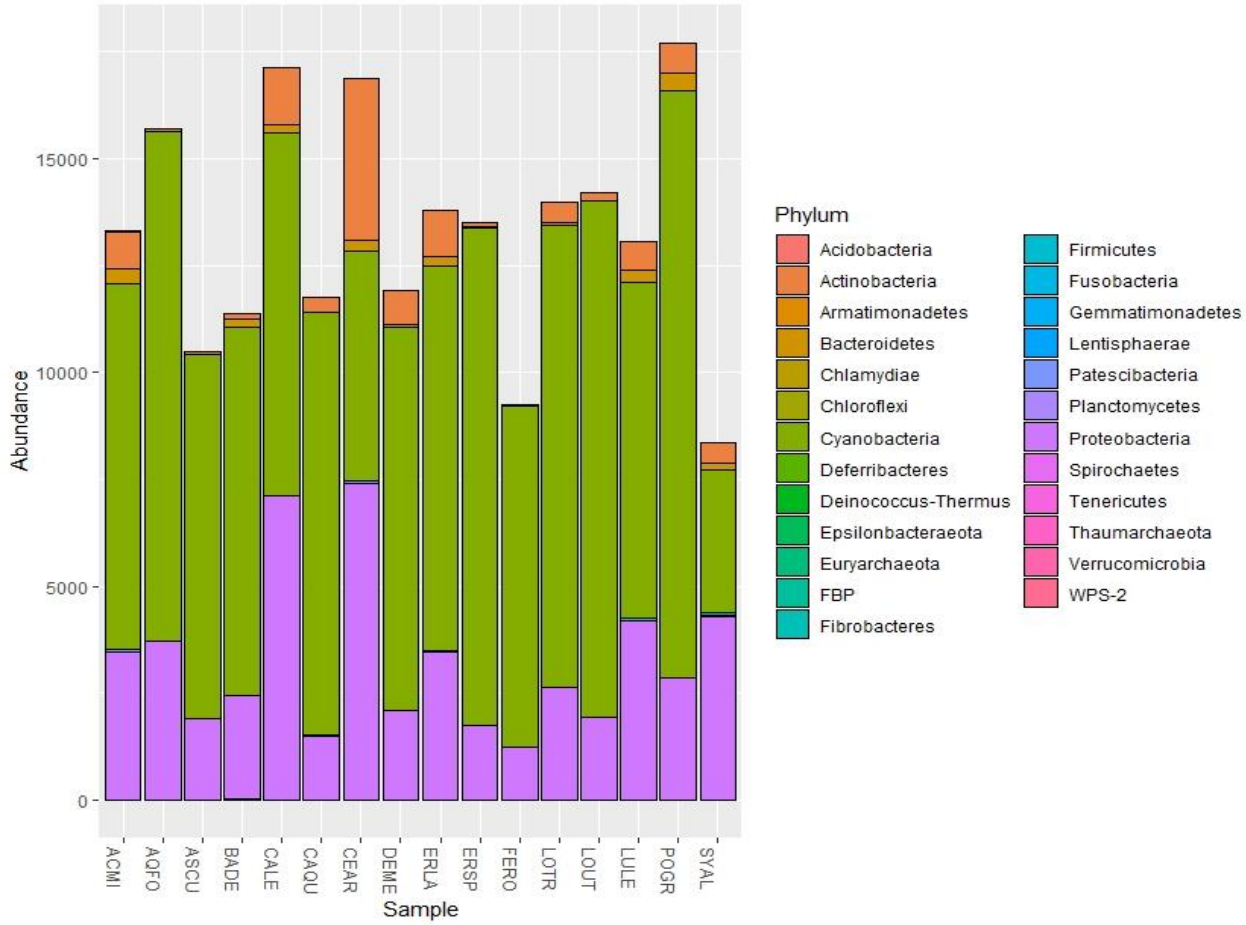


Figure 1.5: Average abundance of bacterial OTU reads representing each bacterial phyla within a plant species. Five bacterial Phyla appear to dominate the bacterial OTU abundance within these 16 plant species: Actinobacteria, Bacteroidetes, Cyanobacteria, Firmicutes and Proteobacteria.

Differences between Plant Species

The results of this test address Chapter 1 hypothesis 1, where it is expected that bacterial OTU composition will differ between Puget prairie plant species. Based on the results of the PERMANOVA test that examined the difference in OTU composition between different plant species, there are significant differences in bacterial OTU composition between different plant species, which supports my hypothesis. The p value of the PERMANOVA test was smaller than alpha ($p = 0.001$), indicating strong statistical significance. The pairwise test reveals that, while

the bacterial OTU composition of certain species does not differ significantly from the bacterial OTU composition of other species, some species have quite a divergent bacterial OTU composition from all other species. 78% of the variation in the dataset can be explained by differences in OTU composition between plant species. A summary of the PERMANOVA test is illustrated in Table 1.7.

Table 1.7: PERMANOVA- Difference in OTU Composition between Plant Species. Differences in bacterial OTU composition were tested on the basis of plant species groups. The total degrees of freedom in this PERMANOVA test were large, allowing for a small p value to be achieved. The p value was less than alpha ($p<0.05$), indicating that there were statistically significant differences in bacterial OTU composition between some plant species.

	DF	Sum of Squares	R2	F	PR (>F)
Species	15	90.145	0.78433	66.917	**0.001**
Residual	276	24.787	0.21567		
Total	291	114.932	1.00000		

Pairwise tests determined that OTU composition differed between almost all plant species: of the 121 plant species pairs examined, 115 differed in bacterial OTU composition. The six pairs that did not differ in composition included species that belong to the same plant family. *Aquilegia formosa* and *Delphinium menziesii* did not differ in their bacterial OTU composition, and belong to Ranunculaceae. *Lomatium triternatum* and *Lomatium utriculatum* did not differ in their bacterial OTU composition, and belong to Apiaceae. The other set of species that did not differ were in the Asteraceae. *Balsamorhiza deltoidea* did not differ from any other Asteraceae, and *Achillea millefolium*, *Aster curtisii*, *Erigeron speciosus*, and *Eriophyllum lanatum* did not often differ from other Asteraceae. However, *Achillea millefolium* and *Aster curtisii* differed from one another, as well as *Erigeron speciosus* and *Eriophyllum lanatum*,

despite these species belonging to Asteraceae. A summary of the pairwise test is illustrated in Appendix 1.G, Table 22.

A three-dimensional NMDS ordination was chosen for visualization of the data, where plant species was overlayed onto one plot as different colors, and plant family was overlayed onto the second plot as different colors (stress = 0.16). The differences in bacterial OTU composition as calculated in PERMANOVA are apparent in the ordination; samples derived from the same plant species appear to cluster into individual groups. *Lomatium triternatum* and *Lomatium utriculatum* do not form distinct groups, as indicated in PERMANOVA that the composition of their bacterial OTUs are not distinct (Figure 1.6). *Achillea millefolium*, *Aster curtisii*, *Balsamorhiza deltoidea*, *Erigeron speciosus* and *Eriophyllum lanatum* also do not form a distinct groups, as indicated in PERMANOVA that the composition of the bacterial OTUs between most members of the Asteraceae were not significantly different from one another (Figure 1.6). *Lomatium triternatum* and *Lomatium utriculatum* belong to the plant family Apiaceae, and *Achillea millefolium*, *Aster curtisii*, *Basamorhiza deltoidea*, *Erigeron speciosus* and *Eriophyllum lanatum* belong to the plant family Asteraceae, which likely explains the lack of significant differences in their bacterial OTU compositions (Figure 1.6). *Delphinium meziezii* and *Aquilegia formosa* both belong to the Ranunculaceae family; while they were found in the pairwise test to have significantly different bacterial OTU compositions, their sample groups are close in ordination space, indicating that while their bacterial OTU compositions are different enough to be distinct, they are not very different.

Several patterns observed in the NMDS indicate opportunities for further exploration. While both *Lomatium utriculatum* and *Lomatium triternatum* were found to have distinct bacterial OTU compositions from *Lupinus lepidus*, *L. lepidus* is observed in close proximity to

these two plant species. These plants have distant taxonomic relations, and it is unclear why these three plant species appear to share similar space in the following ordination. Two data points diverge from the cluster that typically defines the ordination space occupied by their species; a *Symphoricarpos albus* sample (0216) can be found occupying space similar to *Camassia quamash*, and an *Aster curtisii* (0280) is a fair distance from the group defined by the other *A. curtisii* samples. Interestingly, while the two monocots observed in this study -*C. quamash* and *Festuca roemeri*- maintain distinct clusters, they are close to one another in ordination space. *Castilleja levisecta* and *C. arvense* have significantly different bacterial OTU compositions, but the clusters for these species overlap in ordination space.

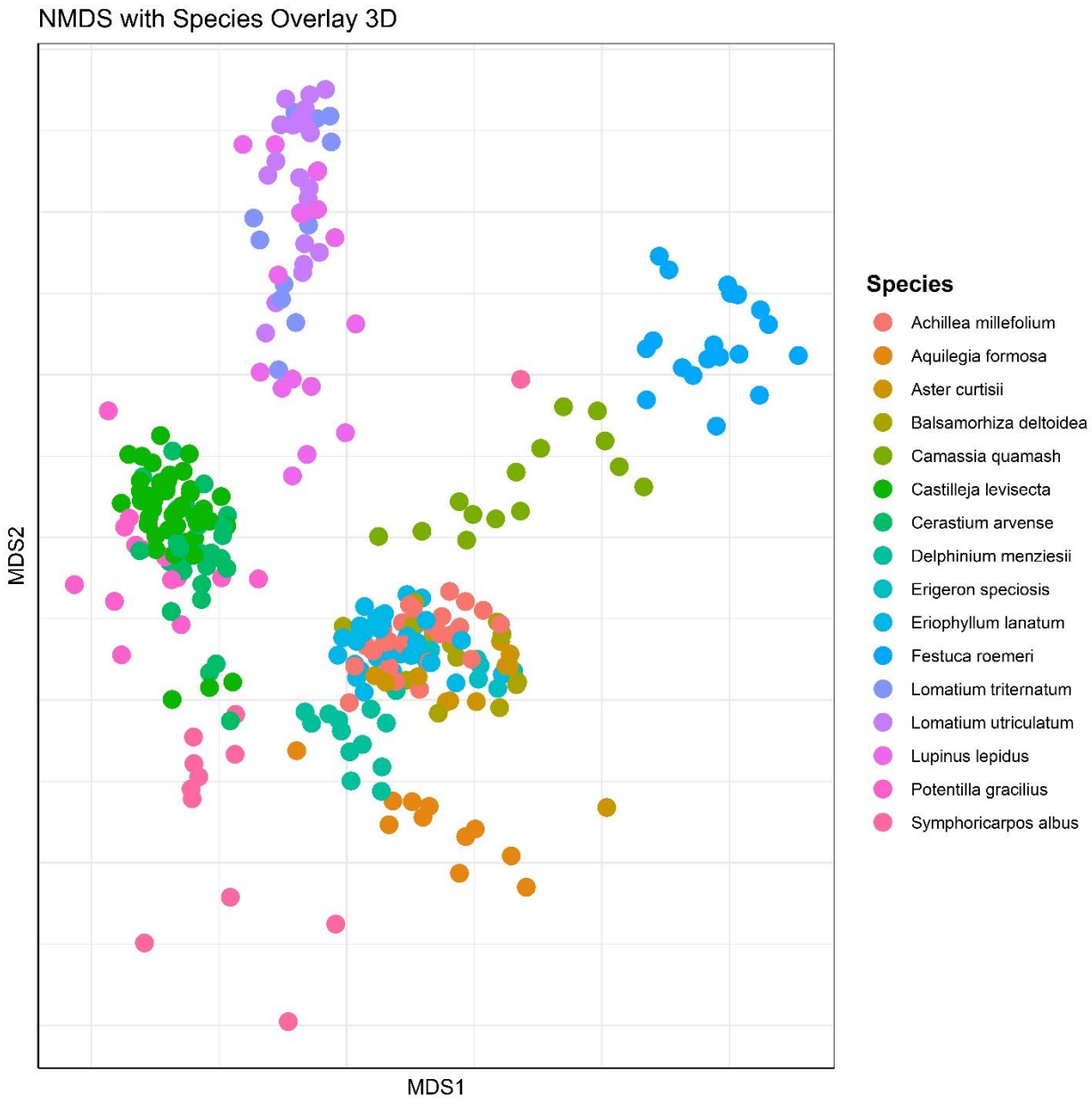


Figure 1.6: Three-dimensional NMDS ordination of bacterial OTU abundance with plant species overlay (stress = 0.16). Only two axis are displayed; MDS1 is the axis that explains the most variation across the dataset, and MDS2 is the axis that explains the second most variation across the dataset. Colors represent the plant species from which the sample was derived.

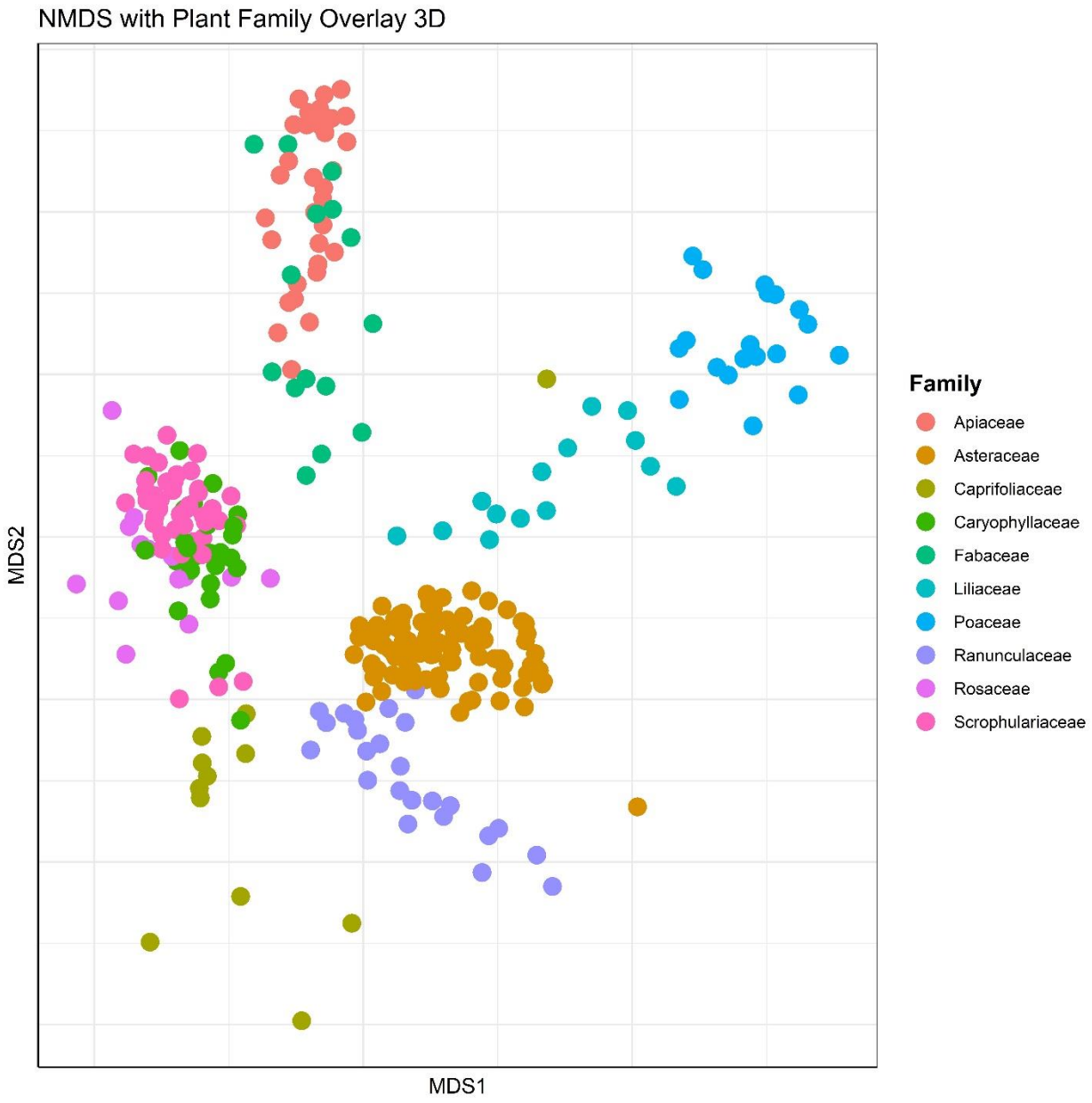


Figure 1.7: Three-dimensional NMDS ordination of bacterial OTU abundance with plant family overlay (stress = 0.16). Only two axis are displayed; MDS1 is the axis that explains the most variation across the dataset, and MDS2 is the axis that explains the second most variation across the dataset. Colors represent the plant family from which the sample was derived. Much of the overlap in the plant species ordination is explained by plant family, where in this ordination, fewer overlaps between groups occur. There appears to be a strong pattern of plant taxonomy, where samples belonging to the same species and plant family can be found occupying similar ordination space.

Effect of Site Location: Glacial Heritage Preserve and Smith Prairie

The results of this test address Chapter 1 hypothesis 2a, where it is expected that bacterial OTU composition will differ between Puget prairie plants taken from GHP and SM. Only four plant species were taken from both Glacial Heritage Preserve and Smith Prairie: *Achillea millefolium*, *Castilleja levisecta*, *Eriophyllum lanatum*, and *Festuca roemeri*. The conservative test for site location is summarized in Table 1.8, and the conservative test for species is summarized in Table 1.9.

Table 1.8: PERMANOVA- Difference in OTU Composition of *Achillea millefolium*, *Castilleja levisecta*, *Eriophyllum lanatum*, and *Festuca roemeri* based on Species and Site Location as crossed terms. With Species as the first term in analysis and GHP.SM as the second term, this is the more conservative test for site location. Site location accounted for less than 1% of variation in the dataset, while differences in species accounted for 72% of the variation in the dataset. The p values for species and site location ($p = 0.001$ and $p = 0.013$, respectively) were smaller than alpha for both, thus there are significant differences in bacterial OTU composition between samples derived from Glacial Heritage Preserve and Smith Prairie and from different species. Interaction between GHP.SM and Species generated a p value of 0.118 which is larger than alpha, thus there is no interaction effect between GHP.SM and Species; species did not differ in the effect of site location.

	DF	Sum of Squares	R2	F	PR (>F)
Species	3	27.548	0.72160	105.6168	**0.001**
GHP.SM	1	0.303	0.00794	3.4883	**0.013**
GHP.SM:Species	3	0.413	0.0108	1.5847	0.118
Residual	114	9.912	0.25963		
Total	121	38.177	1		

Table 1.9: PERMANOVA- Difference in OTU Composition of *Achillea millefolium*, *Castilleja levisecta*, *Eriophyllum lanatum*, and *Festuca roemeri* based on Species and Site Location as crossed terms. With GHP.SM as the first term in analysis and Species as the second term, this is the less conservative test for site location. Site location accounted for 2% of variation in the dataset, while differences in species accounted for 71% of the variation in the dataset. The p values for site location and species ($p = 0.001$) were smaller than alpha, thus there are significant differences in bacterial OTU composition between samples derived from Glacial Heritage Preserve and Smith Prairie and from different species. Interaction between GHP.SM and Species generated a p value of 0.125, larger than alpha, thus there is no interaction effect between GHP.SM and Species; species did not differ in the effect of site location.

	DF	Sum of Squares	R2	F	PR (>F)
GHP.SM	1	0.847	0.0222	9.7467	**0.001**

Species	3	27.004	0.70735	103.5306	**0.001**
GHP.SM:Species	3	0.413	0.0108	1.5847	0.125
Residual	114	9.912	0.25963		
Total	121	38.177	1		

Both tests indicated that bacterial OTU composition varies among sites. The p value for the conservative site location test is larger than for the non-conservative test (p = 0.013 and 0.001, respectively), and allows site location to account for a smaller amount of variation in the dataset (p = 0.8% and 2%, respectively). Even while allowing site location to account for as much variation as possible, site location still only accounts for a minimal amount of variation compared to species (2% compared to 71%, respectively). Although there are differences in bacterial OTU composition between samples taken from GHP and SM, these differences are largely eclipsed by the differences in species (Figure 1.8). The site x species interaction was not significant, indicating that compositional differences among sites were similar across all species.

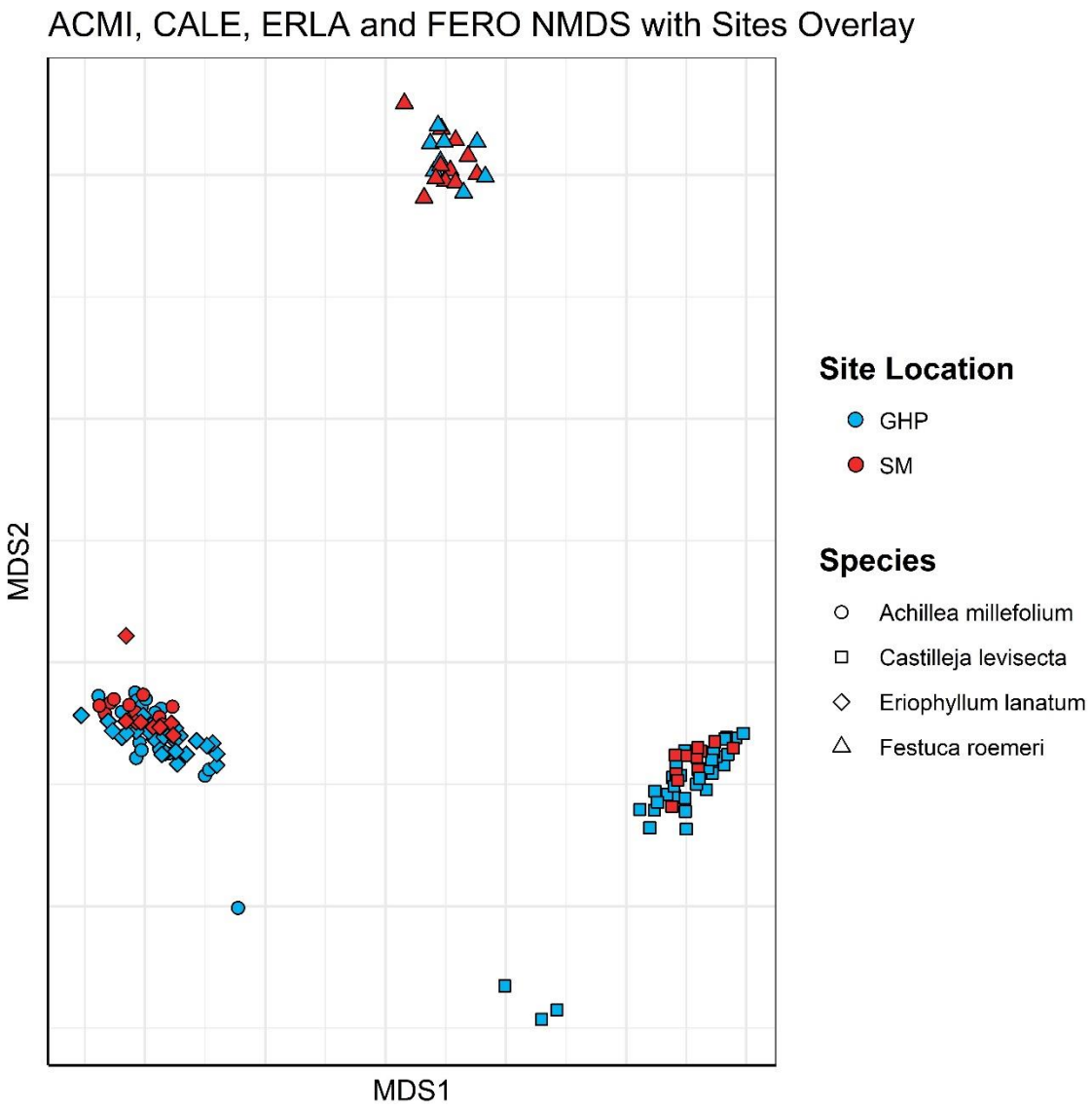


Figure 1.8: Two-dimensional NMDS ordination of bacterial OTU abundance of *Achillea millefolium, Castilleja levisecta, Eriophyllum lanatum,* and *Festuca Roemeri* samples with site location overlay (stress = 0.07). MDS1 is the axis that explains the most variation across the dataset, and MDS2 is the axis that explains the second most variation across the dataset. Colors represent the sample location from which the sample was derived, shapes represent the species from which the sample was derived. Samples appear to group strongly by species and weakly by site location, which supports the findings of the PERMANOVA test. *A. millefolium* and *E. lanatum* sample clusters overlap, likely due to these species belonging to the sample plant family, Asteraceae.

Initial Restoration Treatments

The results of this test address Chapter 1 hypothesis 2b, where it is expected that bacterial OTU composition will differ between Puget prairie plant species taken from plots that had received different initial disturbance treatments. The following PERMANOVA and pairwise tests were performed on samples taken from arrays within GHP. SM did not receive a disturbance regime fire/mow treatment, and thus was excluded from initial disturbance treatment analysis and disturbance regime analysis. Because statistically significant differences in bacterial OTU composition was found to be determined in large part based on plant species, further analysis was conducted on an individual plant species basis. Separate PERMANOVA tests were calculated for each plant species to examine the effect of initial disturbance treatment on bacterial OTU composition. Only 5 species, *Aster curtisii*, *Castilleja levisecta*, *Cerastium arvense*, *Eriophyllum lanatum*, and *Festuca roemeri* contained enough samples representative of each initial disturbance treatment type to perform the PERMANOVA tests; a summary of the samples used in these tests is available in the Methods section in Table 1.5 above. Summaries of the PERMANOVA tests are illustrated in the following tables (Tables 1.10-1.14).

Table 1.10: PERMANOVA- Difference in OTU Composition of *Aster curtisii* based on initial disturbance treatment. Differences in bacterial OTU composition of *Aster curtisii* samples were tested on the basis of initial disturbance treatment groups. Initial disturbance treatment accounted for 14.0% of variation in the dataset. The p value (p = 0.44) was larger than alpha, thus there is not a significant difference in bacterial OTU composition between samples derived from plots with different initial disturbance treatments.

	DF	Sum of Squares	R2	F	PR (>F)
Initial Treatment	2	0.13431	0.14025	0.571	0.44
Residual	7	0.82334	0.85975		
Total	9	0.95765	1.00000		

Table 1.11: PERMANOVA- Difference in OTU Composition of *Castilleja levisecta* based on initial disturbance treatment. Differences in bacterial OTU composition of *C. levisecta* samples were tested on the basis of initial disturbance treatment groups. Initial disturbance treatment accounted for 13.0% of variation in the dataset. The p value (p = 0.296) was larger than alpha, thus there is not a significant difference in bacterial OTU composition between samples derived from plots with different initial disturbance treatments.

	DF	Sum of Squares	R2	F	PR (>F)
Initial Treatment	2	0.20172	0.12978	1.1185	0.296
Residual	15	1.35262	0.87022		
Total	17	1.55435	1.00000		

Table 1.12: PERMANOVA- Difference in OTU Composition of *Cerastium arvense* based on initial disturbance treatment. Differences in bacterial OTU composition of *C. arvense* samples were tested on the basis of initial disturbance treatment groups. Initial disturbance treatment accounted for 21.3% of variation in the dataset. The p value (p = 0.025) was smaller than alpha, thus there is a significant difference in bacterial OTU composition between samples derived from plots with different initial disturbance treatments.

	DF	Sum of Squares	R2	F	PR (>F)
Initial Treatment	2	0.34435	0.21331	1.8981	**0.025**
Residual	14	1.26996	0.78669		
Total	16	1.61431	1.00000		

Table 1.13: PERMANOVA- Difference in OTU Composition of *Eriophyllum lanatum* based on initial disturbance treatment. Differences in bacterial OTU composition of *E. lanatum* samples were tested on the basis of initial disturbance treatment groups. Initial disturbance treatment accounted for 39.9% of variation in the dataset. The p value (p = 0.128) was larger than alpha, thus there is not a significant difference in bacterial OTU composition between samples derived from plots with different initial disturbance treatments.

	DF	Sum of Squares	R2	F	PR (>F)
Initial Treatment	2	0.20732	0.39945	1.6628	0.128
Residual	5	0.31169	0.60055		
Total	7	0.51901	1.00000		

Table 1.14: PERMANOVA- Difference in OTU Composition of *Festuca roemeri* based on initial disturbance treatment. Differences in bacterial OTU composition of *F. roemeri* were tested on the basis of initial disturbance treatment groups. Initial disturbance treatment accounted for 41.1% of variation in the dataset. The p value (p = 0.194) was larger than alpha, thus there is not a significant difference in bacterial OTU composition between samples derived from plots with different initial disturbance treatments.

	DF	Sum of Squares	R2	F	PR (>F)
Initial Treatment	2	0.20915	0.41058	1.7415	0.194
Residual	5	0.30025	0.58942		
Total	7	0.50940	1.00000		

Of these 5 tests, *Cerastium arvense* was the only species that indicated that there is a statistically significant difference in bacterial OTU composition based on initial disturbance treatment. A pairwise test was performed on the *C. arvense* dataset, where it was determined that there is a statistically significant difference in bacterial OTU composition between sites that received burn treatments versus solarize treatments, but no differences were detected between sites that received two year herbicide and burn treatments or differences between sites that received two year herbicide and solarization treatments. A two-dimensional NMDS plot was generated with just *C. arvense* bacterial OTU data, with an initial disturbance treatment and year of inception overlays (Figure 1.9). Summaries of the pairwise tests are illustrated in Table 1.22, Appendix 1.I. The differences between the burn treatments and solarize treatments are apparent. As the two-year herbicide treatment only had two representative samples, differences between two-year and the other treatments are not apparent in the NMDS nor statistically significant in the PERMANOVA and pairwise tests.

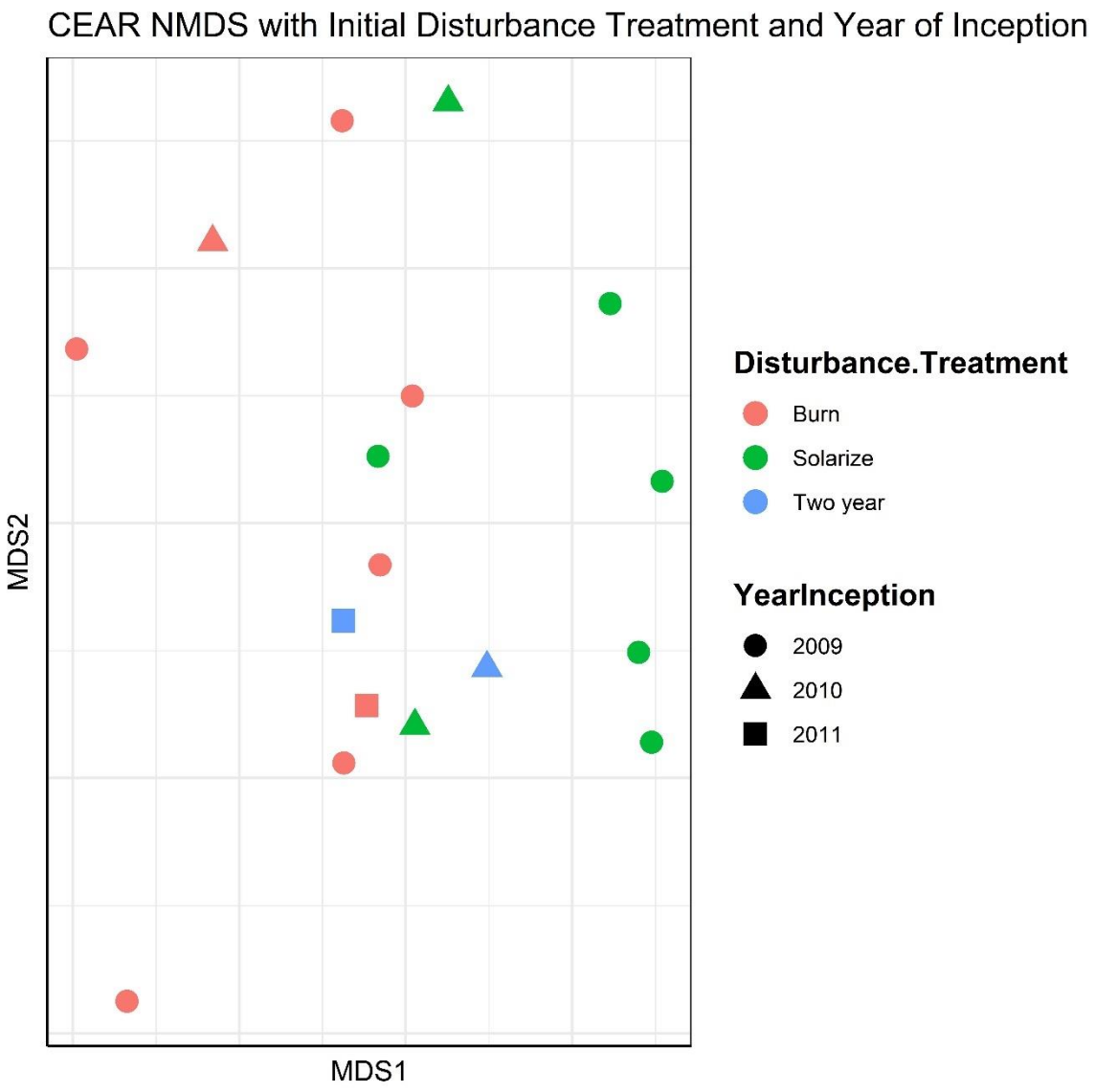


Figure 1.9: Two-dimensional NMDS ordination of bacterial OTU abundance of *Cerastium arvense* samples with initial disturbance treatment and Year of Inception overlays. MDS1 is the axis that explains the most variation across the dataset, and MDS2 is the axis that explains the second most variation across the dataset (stress = 0.125). Colors represent the initial disturbance treatment the plot that the *C. arvense* was taken from had received at the onset of the experiment. Shapes represent the year the intial disturbance treatment was applied to the plot. There is a statistically significant difference between the burn and solarize initial disturbance treatments, as illustrated in the pairwise test ($p = 0.006$). However, there is no statistically significant difference between two-year herbicide and burn treatments ($p = 0.45$), nor between two-year herbicide and solarize treatments ($p = 0.61$).

Disturbance Regime

The results of this test address Chapter 1 hypothesis 2c, where it is expected that bacterial OTU composition will differ between Puget prairie plant species taken from plots that had received different disturbance regime treatments. Because plant species accounts for considerable variation in bacterial OTU composition, this analysis was conducted separately for each of the three species present in all five disturbance regime treatments (Tables 15-17). Of the three species tested (*Castilleja levisecta*, *Cerastium arvense*, and *Lomatium utriculatum*), only *C. arvense* had a statistically significant difference in bacterial OTU composition based on disturbance regime (Table 16). A pairwise test identified differences in bacterial OTU composition between plots that were mowed annually and plots burned every three years in early summer. Additionally, statistical significance was determined between sites that were burned once every three years and later in the summer and sites that were burned every three years and early in the summer. No differences were detected between sites that received other disturbance regime treatments.

A follow-up PERMANOVA test was performed on *Cerastium arvense* to determine if bacterial OTU composition relates to time since last treatment. All site treatments were included in this analysis. The results of this test were also significant; plots treated in 2017 differed from plots treated in 2018 ($p = 0.044$) and from plots treated in 2014 ($p = 0.042$). A summary of the pairwise test is illustrated in Appendix 1.I, Table 23. A two-dimensional NMDS plot was generated with just *Cerastium arvense* bacterial OTU data, with disturbance regime and date of last Treatment overlays (Figure 1.10).

Table 1.15: PERMANOVA- Difference in OTU Composition of *Castilleja levisecta* based on disturbance regime. Differences in bacterial OTU composition of *C. levisecta* samples were tested on the basis of initial disturbance treatment groups. Initial disturbance treatment accounted for 26.1% of variation in the dataset. The p value (p = 0.3) was larger than alpha, thus there not a significant difference in bacterial OTU composition between samples derived from plots with different initial disturbance treatments.

	DF	Sum of Squares	R2	F	PR (>F)
Disturbance Regime	4	0.40543	0.26084	1.1469	0.3
Residual	13	1.14892	0.73916		
Total	17	1.55435	1.000000		

Table 1.16: PERMANOVA- Difference in OTU Composition of *Cerastium arvense* based on disturbance regime. Differences in bacterial OTU composition of *C. arvense* samples were tested on the basis of initial disturbance treatment groups. Initial disturbance treatment accounted for 36.5% of variation in the dataset. The p value (p = 0.014) was smaller than alpha, thus there is a significant difference in bacterial OTU composition between samples derived from plots with different initial disturbance treatments.

	DF	Sum of Squares	R2	F	PR (>F)
Disturbance Regime	4	0.58947	0.36515	1.7255	**0.014**
Residual	12	1.02484	0.63485		
Total	16	1.61431	1		

Table 1.17: PERMANOVA- Difference in OTU Composition of *Lomatium utriculatum* based on disturbance regime. Differences in bacterial OTU composition of *L. utriculatum* samples were tested on the basis of initial disturbance treatment groups. Initial disturbance treatment accounted for 19.4% of variation in the dataset. The p value (p = 0.959) was larger than alpha, thus there not a significant difference in bacterial OTU composition between samples derived from plots with different initial disturbance treatments.

	DF	Sum of Squares	R2	F	PR (>F)
Disturbance Regime	4	0.039915	0.19363	0.3602	0.959
Residual	6	0.166223	0.80637		
Total	10	0.206138	1		

Of these three tests, *Cerastium arvense* was the only that produced a statistically significant difference in bacterial OTU composition based on disturbance regime. A pairwise test

was performed on the *C. arvense* dataset, where it was determined that there is a statistically significant difference in bacterial OTU composition between sites that were mowed annually and sites that were burned every three years and in early summer. Additionally, statistical significance was determined between sites that were burned once every three years and later in the summer and sites that were burned every three years and early in the summer. No differences were detected between sites that received other disturbance regime treatments. A following PERMANOVA test was performed on *C. arvense* to test if differences in bacterial OTU composition are present between *C. arvense* plants taken from sites that last treated at different dates. The results of this test were also significant, where plots treated in 2017 differed from plots treated in 2018 ($p = 0.044$) and from plots treated in 2014 ($p = 0.042$). Summaries of the pairwise tests are illustrated in Table 1.23, Appendix 1.I. A two-dimensional NMDS plot was generated with just *C. arvense* bacterial OTU data, with disturbance regime and date of last treatment overlays (Figure 1.10).

CEAR NMDS with Disturbance Regime and Date of Last Treatment Overlays

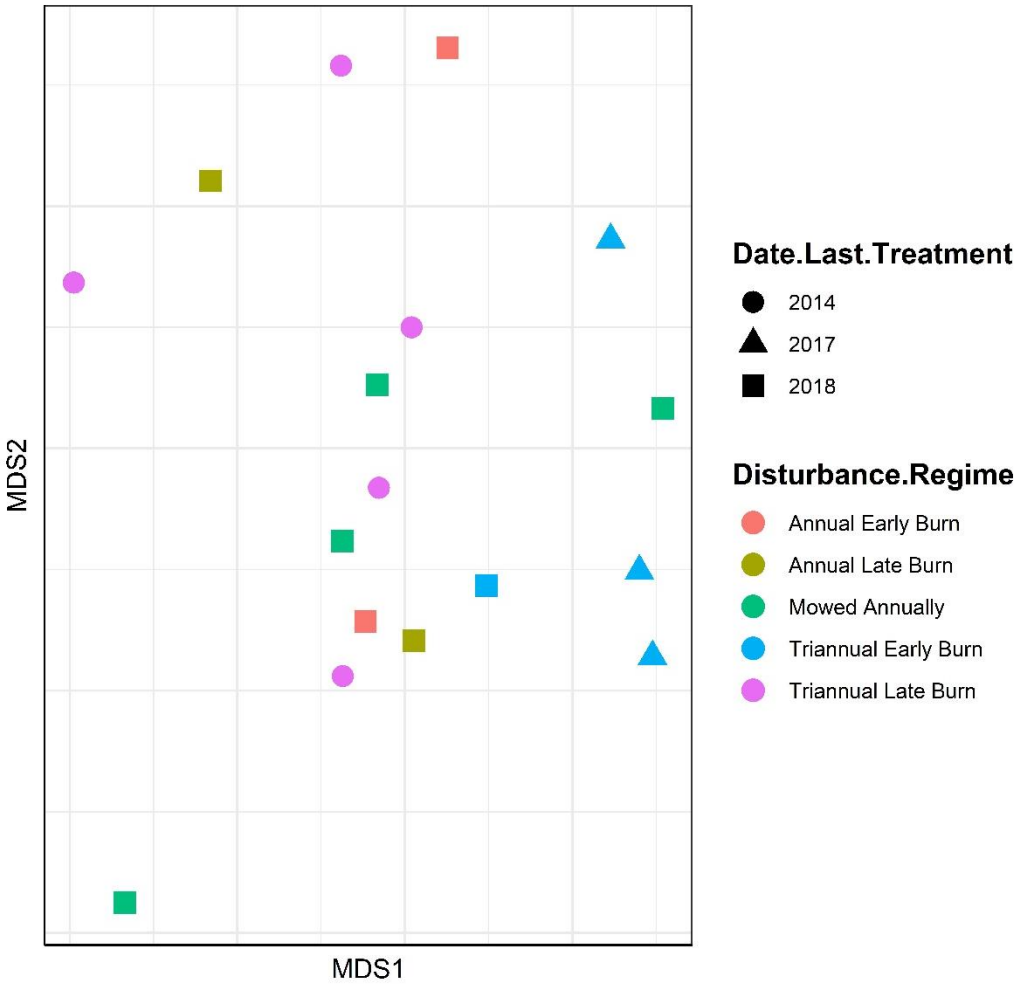


Figure 1.10: Two-dimensional NMDS ordination of bacterial OTU abundance of *Cerastium arvense* samples with disturbance regime and date of last treatment overlays. MDS1 is the axis that explains the most variation across the dataset, and MDS2 is the axis that explains the second most variation across the dataset (stress = 0.125). Colors represent the disturbance regime treatment the plot that the *C. arvense* sample was taken from had received throughout the experiment. Shapes represent the last year the disturbance regime treatment was applied to the plot. There is a statistically significant difference between the triannual early burn and the triannual late burn treatments, as illustrated in the pairwise test ($p = 0.007$). The date of last treatment differed for the triannual early burn and triannual late burn treatments, as these plots were located in different arrays (triannual late burn samples taken from plots treated exclusively in 2014, while triannual early burn plots were mostly treated in 2017 with one sample from a plot treated in 2018). There is marginal statistical significance between annual late burn and annual early burn ($p= 0.067$), between triannual early burn and the annual mow treatments ($p = 0.080$), and between annual late burn and triannual early burn (0.067). However, there are no statistically significant differences between other treatments. This indicates a weak trend that the type of treatment (burning vs. mowing) and timing of the treatment (early vs. late) may have an affect on the bacterial OTU composition of *C. arvense* samples.

Discussion

Previous studies have determined that plant species identity plays a major role in determining a plant's internal bacterial community (Ding and Melcher 2016; Ding et al. 2013; Graner et al. 2003). There is considerable evidence to suggest that plants and bacteria -both beneficial and pathogenic bacteria - coevolved as a result of bacterial and plant cohabitation (Levy et al. 2017; Hassani et al. 2018), and while some bacteria have a narrow host range, others have a more broad host range and are able to colonize and influence growth in many plant species (Afzal et al. 2019). Thus far, only a small fraction of the bacterial communities comprising the microbiome of the total diversity of plant taxa in the world have been studied. This paper contributes to the knowledge of the bacterial communities of 16 additional plant species: *Achillea millefolium, Aquilegia formosa, Aster curtisii, Balsamorhiza deltoidea, Castilleja levisecta, Camassia quamash, Cerastium arvense, Delphinium menziesii, Eriophyllum lanatum, Erigeron speciosus, Festuca roemeri, Lomatium utriculatum, Lomatium triternatum, Lupinus lepidus, Potentilla gracilis,* and *Symphoricarpos albus*.

The results of this study support previous research findings and my own hypothesis; most of the 16 Puget prairie species examined in this study retained distinct bacterial OTU communities, based on results of PERMANOVA and pairwise tests. Visualizations of the dataset with plant species and plant family overlay supports the results of the PERMANOVA and pairwise tests, where samples derived from the same species organized into semi-distinct groups. The plant species that did not retain distinct bacterial OTU communities were species that belonged to the same plant family, and thus share similar physiological and life history traits. It is possible that these shared traits between members of the same plant family create an interior environment of a plant that is suitable for not only bacteria able to colonize a wide diversity of

host plants, but also suitable for bacteria that coexist within a smaller range of host plants. Thus, plants belonging to the same plant family might be more capable of hosting similar bacteria than plants belonging to different families. With a larger sample size and fewer confounding variables (such as initial disturbance treatments and continuous disturbance regime treatments), it may be possible to further distinguish unique bacterial communities between species that belong to the same family.

While it is evident that the bacterial OTU composition derived from a sample is largely determined by plant species, it is not yet evident which bacterial OTUs may be driving differences between these Puget prairie plant species. Further investigation of these data could include Indicator Species Analysis, which would identify any bacterial OTUs responsible for driving disproportionate differences in the bacterial OTU composition between species or family groups. Indicator Species Analysis has been used to examine the composition of microbiomes in previous studies and could be applied to this dataset (Cariveau et al. 2014). Information on indicator bacterial OTUs could be used to examine the coevolutionary history of a particular bacterial endophyte and a host plant species, as well as other significant physiological interactions between indicator bacterial endophytes and their host species, such as nutrient exchange, hormone production, and pathogen resistance (Dufrene and Legendre 1997; Zilber-Rosenberg and Rosenberg 2008; Baltrus 2017).

I theorized that there may be differences in the bacterial microbiome of *Achillea millefolium*, *Castilleja levisecta*, *Festuca roemeri* and *Eriophyllum lanatum* collected at GHP and SM, and found that there were statistically significant differences in the bacterial OTU composition of samples taken from GHP and SM. Observing NMDS ordinations of these four species, samples are clustered much more strongly by species than by site location (Figure 1.8).

Disregarding grouping based on species, it appears that the SM samples cluster closer together than GHP samples and are nested within the GHP samples. Many more samples were collected from GHP than from SM for all species except *F. roemeri*; it is likely that a more balanced sampling approach, with more samples collected from SM, could reveal stronger differences in the bacterial OTU composition of plant samples taken from GHP and SM. Given the statistical significance of my PERMANOVA test and observations drawn from the NMDS ordination, I conclude that there is an effect of sampling location on bacterial OTU composition within species.

Many environmental factors play a role in determining the bacterial community composition within a plant. Soil chemistry is known to affect the soil microbial community, where pH favors certain bacterial species over others (Burns et al. 2015). As fire regimes alter soil chemistry, and many bacterial endophytes are derived from the soil, it follows that fire disturbance treatments applied to the sites may affect the composition of the bacterial endophyte community within plants. In a study of foliar endophytes residing in trees, wildfires were found to change the diversity, community structure and taxonomic composition of the bacterial endophyte community (Huang et al. 2016). This study only examined trees that were not killed after fire; their sampling pool included only trees that had survived surface defoliation and resprouted post-fire. Despite the differences in prairie and forested systems, it may be appropriate to compare these systems as prairie plant individuals often survive burns in a similar way, by relying on underground root systems to resprout after fire.

Non-fire disturbances may also have the ability to affect bacterial endophyte composition as well. Herbicides eliminate target plants without removing biomass from the surface, and leaves behind residual chemicals in the soil that can persist long after the initial treatment of the

herbicide. Herbicides have been known to decrease the diversity, richness, and evenness of fungal leaf endophytes within soy plants and may be true for bacterial endophytes as well (Stuart et al. 2018). Solarization eliminates all plants under a surface, typically made of plastic, that uses heat and soil humidity to disrupt biological processes (Elmore et al. 1997). Soil moisture and temperature have been known for decades to determine the survival of soil microorganisms (Dunn et al. 1985). Mowing removes surface material typically above 1 inch from the surface of the soil, leaving soil properties relatively unchanged. Mowing has been shown to have mixed impacts on bacterial endophyte communities; in a study of grassland management techniques, bacterial endophyte communities within grasses were sometimes found to be affected by mowing and fertilizer treatments, but other grass species were resistant to changes due to mowing and fertilizer treatments (Wemheuer et al. 2017). Fire treatments, herbicide treatments, solarize treatments and mow treatments differ in their impact on living plant matter and soil properties, and in their immediate and long-term impacts on the plot they are applied to (Dickenson 2019; Bahm et al. 2011; Elmore et al. 1997). As bacterial endophytes largely depend on living plants and particular soil properties, it follows that changes in living plant matter and soil properties induced by disturbance treatments could affect bacterial endophyte communities that colonize the plants that regenerate after disturbance.

As rhizosphere inhabiting bacteria are sensitive to soil and climate conditions presented by their immediate environment, it is of interest to observe if slight differences in the conditions of similar environments produce different soil bacteria communities. Glacial Heritage Preserve and Smith Prairie are approximately 95 miles apart, and the Puget Sound isolates Whidbey island from mainland Washington. While both Glacial Heritage Preserve and Smith Prairie host Puget prairie ecosystems, it is possible that their unique geographical locations and subsequent

differences in their soil, climate, and bacterial dispersal history could drive present day differences in the soil bacterial communities and thus differences in the bacterial communities comprising the interior microbiome of Puget prairie plant species.

One key variable used in the study of disturbance treatments on the Puget prairie research plots is initial disturbance treatment. Solarization, two-year herbicide, and broadcast burning were applied to plots at the Glacial Heritage Preserve in 2009, 2010, and 2011. Results from PERMANOVA and pairwise testing indicates that, for *Aster curtisii*, *Castilleja levisecta*, *Cerastium arvense*, *Eriophyllum lanatum*, and *Festuca roemeri* samples collected from these plots, there were no differences in the bacterial OTU composition of samples taken from plots treated with two year herbicide and burning, nor differences in bacterial OTU composition from samples taken from plots treated with two year herbicide and solarization. For *C. arvense* alone, there was a significant difference between samples taken from plots treated with solarization and burning. These results indicate that there is a very weak effect, if any effect, of initial disturbance treatment on bacterial community composition.

While the microenvironmental conditions within each plot were theorized to differ between plots due to initial treatment type, even when controlling for plant species, there was only one instance where initial disturbance treatment resulted in significantly different bacterial OTU composition for plants collected from different initial disturbance treatment plots. It is possible that initial disturbance treatment may only have a weak effect on bacterial community composition because of the large amount of time that has passed since the initial treatments were applied to the sites. Samples from these five species were collected in 2018, nearly a decade after some of the plots had been treated. It is likely that even long-lasting effects of these initial disturbance treatments, such as the impact of chemical herbicide and the alterations in soil

chemistry caused by broadcast burning, would have weakened with time. Given 7, 8 and 9 years for plants and bacteria to recover from the initial disturbance treatments, it is likely that initial differences in the bacterial communities between treatment plots have disappeared. Additionally, these treatments were only applied to each treatment site once throughout the course of the experiment, and while long term applications of herbicide treatments have been shown to impact the bacterial soil community, short term applications are likely to have a weaker effect (Seghers et al. 2003). Despite the significant differences in bacterial OTU composition observed between *Cerastium arvense* plants sampled from plots exposed to different initial disturbance treatments, I cannot conclude that there is either a consistent or strong effect of initial disturbance treatment on bacterial OTU composition several years after treatment application.

Other key variables that were used in the study of disturbance on Puget prairie research plots were the frequency and seasonality of disturbance regime applied continuously to the plots after initial disturbance treatments. Burning and mowing disturbance regimes were applied to experimental plots on different set schedules, where plots were either burned annually, burned triannually, or mowed annually. Additionally, plots that were burned annually or triannually were also either burned early in the fire season or late in the fire season. Results from PERMANOVA and pairwise testing reveal that for *Castilleja levisecta*, *Cerastium arvense*, *Eriophyllum lanatum*, *Festuca roemeri*, and *Lomatium utriculatum*, there were no differences in bacterial OTU composition between plots that were either burned annually or burned triannually. Neither were there differences in plots that were either burned annually or mowed annually. However, for *C. arvense* alone, there were differences in bacterial OTU composition between plants collected from plots that were burned triannually at different times in the season (late vs. early) and differences in bacterial OTU composition from plants collected in plots that were

burned triannually in early season and mowed annually. Within *C. arvense* plants collected at Glacial Heritage Preserve, plants in plots that received triannual early burn treatments and triannual late burn treatments retain statistically significant differences in bacterial OTU composition, and plants in plots that received triannual early burn treatments and annual mow treatments retain statistically significant differences in bacterial OTU composition.

Bacterial communities are known to have different compositions based on the season in which the community is sampled. Application of disturbance treatments during different seasons, therefor, could alter bacterial communities in varying degrees. Plots that are burned in early season, for example, could disproportionately impact bacterial organisms that depend on early season conditions to complete their life cycles as compared to bacteria that rely more on later season conditions for their biological processes. Thus, differences in the bacterial OTU composition of *Cerastium arvense* samples collected from plots treated with either triannual early burn or triannual late burn could be due to the interference of important bacterial biological processes that occur during different seasons. However, it is unclear why there were only differences observed in triannual burns and not also in annual burns which were also burned on different seasonal schedule. Marginal significance was achieved between annual burn sites treated in early season and late season, indicating that a greater sample size could generate a stronger effect of seasonal burning on the bacterial OTU composition of *C. arvense*. Despite the significant differences in bacterial OTU composition observed between *C. arvense* plants sampled from plots exposed to seasonal burning disturbance treatments, I cannot conclude that there is either a consistent or strong effect of seasonal burning treatment on bacterial OTU composition.

Fire treatment frequency was also predicted to affect bacterial communities of Puget prairie plants. A previous study on soil bacteria and fire regime revealed that long-term, repeated fire disturbance (applied once every two years, once every four years, or not at all) had an effect on bacterial community structure driven by soil pH and C:N ratio factors; however, the abundance of bacteria in this study were not changed due to fire regime (Shen et al. 2016). The results of my study indicate that none of the five Puget prairie species tested revealed a statistically significant difference in bacterial OTU composition from samples derived from plots treated annually and triannually. Puget prairie ecosystems have encountered fire as a part of their disturbance regime for thousands of years as a practice implemented and maintained by indigenous peoples, and these prairie communities are resilient to frequent fires (Boyd 1999). It is likely that, as the organisms that occupy this landscape are adapted to a frequent fire regime, both the plant and their associated bacterial communities are able to quickly recover from fire disturbances. I conclude that there is no effect of fire frequency on bacterial OTU composition in these five prairie plants.

Finally, it is unclear why only *Cerastium arvense* achieved statistically significant differences in bacterial community composition between any of the five treatments, while the other four plants did not. It is possible that the differences in bacterial OTU composition discovered within *C. arvense* samples taken from plots treated with different disturbance regimes are indicative of a larger trend, that disturbance regime, as a combination of the effects of seasonality, frequency of treatment, and fire vs. mow treatments, has an effect on bacterial community composition within plants. Although the sample sizes used in these tests were low (ranging from 7 to 17 degrees of freedom), the percentage of variation that the disturbance regime treatment accounted for in the data was notably high (ranging between 19% to 75% of the

variation). No other plants used in the study of disturbance treatments, besides *C. arvense*, were found to have significant differences in bacterial OTU composition of samples taken from plots that received different disturbance regime treatments. With a larger sample size, the effect of frequency of treatment (annual fire versus triannual fire), season of treatment (early burn versus late burn), and type of treatment (fire versus mow) would increase in accuracy and potentially illuminate statistically significant results in the bacterial composition of plants taken from plots receiving different treatments that were only previously observed in *C. arvense*. However, given the lack of consistent statistically significant differences in bacterial OTU composition on the basis of initial disturbance treatment and disturbance regime, I cannot conclude that there is an overall pattern of disturbance treatment driving bacterial OTU composition within plant species.

Current research suggests that species of plants have co-evolved alongside their root nodule or other tissue-inhabiting bacterial endophytes, and have thus developed complex host/symbiont relationships which benefit both the host plant and the endophytic bacteria (Clay and Schardl 2002; Brooks et al. 2016; Levy et al. 2018). However, it is of particular research interest to examine if bacterial endophytes are able to be isolated from their original host plants and applied to target plant species which were not originally recognized as natural host plants. Bacterial endophytes that are able to be taken from their original host plants and transferred to non-host plants continue to provide plant growth promoting substances in some cases (Afzal et al. 2019). Further analysis of this dataset could lead to the discovery of bacterial endophytes existing in native prairie plant species which could be used in other systems to promote plant growth. Substantial and consistent increases in plant growth would help determine if inoculating plants with plant growth promoting endophytes could act as a potential means to reduce disease, increase nutrient uptake, and generally enhance the growth of target plants (Doty 2017).

Endophytic bacteria acting as biological control agents have been investigated for industrial application in agriculture, agroforestry, and plant nurseries (Rabiey 2019). Using endophytic bacteria to control pests such as nematodes, insects, pathogenic bacteria, and pathogenic fungi makes it possible to control pest populations while reducing or eliminating the need for artificial pesticides (Alström and Van Vuurde 2001). Large scale use of artificial pesticides has been found to be problematic in many instances, causing extensive ecological damage across many systems (Edwards 1993). Alternative modes of controlling plant pest populations have the potential to mitigate damage from pests without the negative impacts of artificial pesticides. The application of endophytic bacteria to target plants has been proposed as a potential alternative to chemical pesticides and studies have investigated the application of bacterial endophytes as biological controls in experimental investigations (Melnick et al. 2005; Mmbaga et al. 2018; Etesami et al. 2019). Initial research studies are encouraging; the application of endophytes in many studies and in many species of plants have demonstrated positive results.

The data that has been collected to complete this thesis provides many more opportunities for investigation. Thus far, I have only examined large picture patterns in bacterial OTU composition that are present in these samples; moving forward, I could narrow my investigation to focus on individual bacterial OTUs using Indicator Species Analysis. In a brief exploration of the data, I found that several OTUs serve as “perfect indicators” for plant species, where an individual bacterial OTU would be found in every sample derived from a single plant species and not found once in any other sample. Further applications of Indicator Species Analysis on this dataset could be used highlight what bacterial OTUs are likely to drive differences in bacterial OTU composition, and conversely, which bacterial OTUs are likely to be found

homogeneously across samples. Other analysis and further research questions inspired by this research include but are not limited to: investigating if reads or OTU count correlate with plant sample biomass; exploring if bacterial OTU composition differs within these species taken from different ecosystems; examining for differences in bacterial OTU composition across these species using a plant tissue other than stem (leaf, root).

Our knowledge of native plant microbiomes may not only allow us to apply bacterial endophytes in more effective ways, but also to discover aspects of the earth's ecosystems that remain unknown. The Earth Microbiome Project, an organization that collects and analyzes microbial diversity across the globe, is one such organization that is leading the study of the earth's various microbiomes (Gilbert et al. 2014). The Earth Microbiome Project has led to the publication of at least 60 scientific papers across a various collection of fields, and is hailed as a source of data "predicated on the value of voyages of discovery" (Gilbert et al. 2011). Microbial data from an incredibly diverse set of systems, including the digestive system of carnivorous pitcher plants, ice-covered Antarctic lakes, the human gut microbiome, Komodo Dragon skin, and others provide only a small fraction of the data derived from the Earth Microbiome Project alone. With immense amounts of microbiome data available to scientists -now including my data on the microbiomes of 16 Puget prairie plant species- the possibilities for new research pursuits are expansive.

Literature Cited

Afzal, I., Z. K. Shinwari, S. Sikandar, S. Shahzad. 2019. Plant beneficial endophytic bacteria: Mechanisms, diversity, host range and genetic determinants. *Microbiological Research*, 221, 36–49. doi: 10.1016/j.micres.2019.02.001

Aktar, W., D. Sengupta, A. Chowdhury. 2009. Impact of pesticides use in agriculture: their benefits and hazards. *Interdisciplinary Toxicology*, 2(1), 1–12. doi: 10.2478/v10102-009-0001-7

Antos, J. A., B. McCune, C. Bara. 1983. The Effect of Fire on an Ungrazed Western Montana Grassland. The American Midland Naturalist, Vol. 110, No. 2.

Alström S., J. W. L. Van Vuurde. 2001. Endophytic Bacteria and Biocontrol of Plant Diseases. In: De Boer S.H. (eds) *Plant Pathogenic Bacteria*. Springer, Dordrecht

Bahm, M., T. G. Barnes, and K. C. Jensen. 2017. Herbicide and Fire Effects on Smooth Brome (Bromus inermis) and Kentucky Bluegrass (Poa pratensis) in Invaded Prairie Remnants. *Cambridge University Press*, Vol. 4 Iss. 2. pp. 189-197.

Bailes, G., D. Thomas, S. D. Bridgham, B. A. Roy. 2020. Drivers of grass endophyte communities in prairies of the Pacific Northwest, USA. BioRxiv, https://doi.org/10.1101/2020.02.23.953489

Balvočiūtė, M. and D. H. Huson. 2017. SILVA, RDP, Greengenes, NCBI and OTT — how do these taxonomies compare?. *BMC Genomics* volume 18, Article number: 114

Buee, M., W. D. Boer, F. Martin, L. van Overbeek, E. Jurkevitch. 2009. The rhizosphere zoo: An overview of plant-associated communities of microorganisms, including phages, bacteria, archaea, and fungi, and of some of their structuring factors. *Plant and Soil*, 321, 189-212.

Baltrus, D. A. 2017. Adaptation, specialization, and coevolution within phytobiomes. *Current Opinion in Plant Biology*, Volume 38, Pages 109-116, ISSN 1369-5266.

Berg, G. 2009. Plant–microbe interactions promoting plant growth and health: perspectives for controlled use of microorganisms in agriculture. *Applied Microbiology and Biotechnology*, 84(1), 11–18. doi: 10.1007/s00253-009-2092-7

Boyd, R. 1999. Indians, Fire, and the Land in the Pacific Northwest. Page 320. *Oregon State University Press*, Corvallis, OR.

Brooks, A. W., K. D. Kohl, R. M. Brucker, E. J. van Opstal, S. R. Bordenstein. 2016. Phylosymbiosis: relationships and functional effects of microbial communities across host evolutionary history. *PLoS Biol*., 14, p. e2000225

Burns, J. H., B. L. Anacker, S. Y. Strauss, D. J. Burke. 2015. Soil microbial community variation correlates most strongly with plant species identity, followed by soil chemistry, spatial location, and plant genus. *AoB PLANTS*, 7(0). doi: 10.1093/aobpla/plv030

Calle, M. L. 2019. Statistical Analysis of Metagenomics Data. *Genomics & Informatics*, 17(1). doi: 10.5808/gi.2019.17.1.e6

Cariveau, D. P., J. E. Powell, H. Koch, R. Winfree, N. A. Moran. 2014. Variation in gut microbial communities and its association with pathogen infection in wild bumble bees (Bombus). *The ISME Journal*, 8(12), 2369–2379. doi: 10.1038/ismej.2014.68

Carthey, A. J. R, D. T. Blumstein, R. V. Gallagher, S. G. Tetu, M. R. Gillings. 2020. Conserving the holobiont. *Functional Ecology*. 34: 764– 776. https://doi.org/10.1111/1365-2435.13504

Clay, K., and C. Schardl. 2002. Evolutionary origins and ecological consequences of endophyte symbiosis with grasses. *Am Nat.* 160:99–127.

Compant, S. B. Duffy, J. Nowak, C. Clement, E. A. Barka. 2005. Use of Plant Growth-Promoting Bacteria for Biocontrol of Plant Diseases: Principles, Mechanisms of Action, and Future Prospects. *Applied and Environmental Microbiology*, ;71(9):4951-9

Cox, S. E., and Stushnoff, C. (2001). Temperature-related shifts in soluble carbohydrate content during dormancy and cold acclimation in *Populus tremuloides*. *Can. J. For. Res*. 31, 730–737. doi: 10.1139/x00-206

Delvin, E. G. 2013. Restoring Abandoned Agricultural Lands in Puget Lowland Prairies: A New Approach. Delvin, Eric G. *University of Washington, ProQuest Dissertations Publishing*.

Dheilly, N. M. 2014. Holobiont–holobiont interactions: redefining host–parasite interactions. *PLoS Pathogens* 10, e1004093.

Dickenson, T. 2019. Burning and mowing similarly increase prairie plant production in the spring, but not due to increased soil temperatures. *Ecosphere*, vol. 10, iss 2. https://doi.org/10.1002/ecs2.2606

Ding, T., and U. Melcher. 2016. Influences of Plant Species, Season and Location on Leaf Endophytic Bacterial Communities of Non-Cultivated Plants. *Plos One*, 11(3). doi: 10.1371/journal.pone.0150895

Ding, T., M. W. Palmer, U. Melcher. 2013. Community terminal restriction fragment length polymorphisms reveal insights into the diversity and dynamics of leaf endophytic bacteria. *BMC Microbiology*, 13(1), 1. doi: 10.1186/1471-2180-13-1

Doty, S. L. 2017. Functional Importance of the Plant Endophytic Microbiome: Implications for Agriculture, Forestry, and Bioenergy. *Functional Importance of the Plant Microbiome*, 10.1007/978-3-319-65897-1_1, (1-5).

Dufrene, M. and P. Legendre. 1997. Species assemblages and indicator species: the need for a flexible asymmetrical approach. *Ecological Monographs* 67:345–366.

Dunn, P. H., S. C. Barro, M. Poth. 1985. Soil moisture affects survival of microorganisms in heated chaparral soil. Soil Biology and Biochemistry. 17: 143-148.

Dunwiddie, P. and J. Bakker. 2011. The Future of Restoration and Management of Prairie-Oak Ecosystems in the Pacific Northwest. *Northwest Science* 85:83–92.

Edwards, C.A. 1993. The Impact of Pesticides on the Environment. In: *The Pesticide Question*. Springer, Boston, MA

Elmore, C. L., J. J. Stapleton, C. E. Bell. 1997. Soil Solarization: A Non-pesticidal Method for Controlling Diseases, Nematodes, and Weeds. *University of California Division of Agriculture and Natural Resources*, Publication 21377.

Etesami, H. H., A. Alikhani, M. H. Hossein. 2019. Archives of Phytopathology and Plant Protection: Applications of Agriculturally Important Microorganisms for Management of Soiborne Phytopathogens, 09 May 2019, Vol.52(7-8), pp.560-581

Frank, A. C., J. P. Guzman, J.E. Shay. 2017. Transmission of Bacterial Endophytes. *Microorganisms*, 10;5(4).

Ferchichi, N., W. Toukabri, M. Boularess, A. Smaoui, R. Mhamdi, D. Trabelsi. 2019. Isolation, identification and plant growth promotion ability of endophytic bacteria associated with lupine root nodule grown in Tunisian soil. Archives of Microbiology. DOI:10.1007/s00203-019-01702-3

Granér, G., P. Persson, J. Meijer, S. Alström. 2013. A study on microbial diversity in different cultivars of Brassica napus in relation to its wilt pathogen, Verticillium longisporum. *FEMS Microbiol. Lett.*, 224, pp. 269-276

Gilbert, J. A., R. O'Dor, N. K, T. M. Vogel. 2011. The importance of metagenomic surveys to microbial ecology: or why Darwin would have been a metagenomic scientist. *Microbial Informatics and Experimentation*, 1(5).

Gilbert, J. A., J. K. Jansson, R. Knight. 2014. The Earth Microbiome Project: successes and aspirations. BMC Biology, 12(1). Doi: 10.1186/s12915-014-0069-1

Glick, B. R. 2012. Plant Growth-Promoting Bacteria: Mechanisms and Applications. *Scientifica*, vol. 2012, pp. 1–15., doi:10.6064/2012/963401.

Gloor, G. B., J. M. Macklaim, V. Pawlowsky-Glahn, J. J. Egozcue 2017. Microbiome Datasets Are Compositional: And This Is Not Optional. *Frontiers in microbiology*. 8:2224. Epub. pmid:29187837; PubMed Central PMCID: PMC5695134.

Goodrich, J. K., S. C. Di Rienzi, A. C. Poole, O. Koren, W. A. Walters, J. G. Caporaso, R. Knight, R. E. Ley. 2014. Conducting a microbiome study. *Cell*, *158*(2), 250–262. https://doi.org/10.1016/j.cell.2014.06.037

Haichar, F. Z., C. Marol, O. Berge, J. I. Rangel-Castro, J. I. Prosser, J. Balesdent, T. Heulin and W. Achouak. 2009. Plant host habitat and root exudates shape soil bacterial community structure. *The ISME Journal*, Vol 2, 1221–1230.

Hardoim, P. R., L. S. Van Overbeek, J. D. Elsas. 2008. Properties of bacterial endophytes and their proposed role in plant growth. *Trends Microbiol* **16**: 463– 471.

Harrison J. G. and E. A. Griffin. 2020. The diversity and distribution of endophytes across biomes, plant phylogeny, and host tissues—how far have we come and where do we go from here? *Environmental Microbiology*, Version 2. doi: 10.1111/1462-2920.14968

Hassani, M. A., P. Duran, S. Hacquard. 2018. Microbial interactions within the plant holobiont. *Microbiome, 6*(1), 58. https://doi.org/10.1186/s40168-018-0445-0

Huang Y. L., M. M. N. Devan, J. M. U'Ren, S. H. Furr, A. E. Arnold. 2016. Pervasive effects of wildfire on foliar endophyte communities in montane forest trees. Microbial Ecology 71:452–468.

Klausmeyer K. R. and M. R. Shaw. 2009. Climate change, habitat loss, protected areas, and the climate adaptation potential of species in Mediterranean ecosystems worldwide. PLoS ONE, 4(7), e6392

Lauber, C. L., M. Hamady, R. Knight, and N. Fierer. 2009. Pyrosequencing-based assessment of soil pH as a predictor of soil bacterial community structure at the continental scale. Appl. Environ. Microbiol. 75(15): 5111–5120. doi:10.1128/AEM. 00335-09.

Legendre P. and E. D. Gallagher. 2001. Ecologically meaningful transformations for ordination of species data. Oecologia. 129:271–280, doi:10.1007/s004420100716.

Levy, A., I. Salas Gonzalez, M. Mittelviefhaus, S. Clingenpeel, S. H. Paredes, J. Miao, K. Wang. 2017. Genomic features of bacterial adaptation to plants. Nat Genet 50:138-150. https://doi.org/10.1038/s41588-017-0012-9

Liu, H., L. C. Carvalhais, M. Crawford, E. Singh, P. G. Dennis, C. M. J. Pieterse and P. M. Schenk. 2017. Inner Plant Values: Diversity, Colonization and Benefits from Endophytic Bacteria, *Frontiers in Microbiology,* 10.3389

Martiny, J. B. H., B. J. M. Bohannan, J. H. Brown, R. K. Colwell, J. A. Fuhrman, J. L. Green, M. C. HornerDevine, M. Kane, J. A. Krumins, C. R. Kuske, P. J. Morin, S. Naeem, L. Øvreås, A. L. Reysenbach, V. H. Smith, J. T. Staley. 2006. Microbial biogeography: putting microorganisms on the map. *Nat Rev Microbiol*. 4:102–112, doi:10.1038/nrmicro1341.

Maziarz, M., R. M. Pfeiffer, Y. Wan, M. H. Gail. 2018. Using standard microbiome reference groups to simplify beta-diversity analyses and facilitate independent validation. *Bioinformatics*, 34(19), 3249–3257. doi: 10.1093/bioinformatics/bty297

Mcmurdie, P. J., S. Holmes. 2013. phyloseq : An R package for reproducible interactive analysis and graphics of microbiome census data. *PLoS One*. 8:1–11, doi:10.1371/journal.pone. 0061217.

Mei, C. and B. Flinn. 2010. The Use of Beneficial Microbial Endophytes for Plant Biomass and Stress Tolerance Improvement. *Recent Patents on Biotechnology*, vol. 4, no. 1, pp. 81–95., doi:10.2174/187220810790069523.

Melnick, R. l., P. A. Backman, B. Bailey, S. Maximova. 2005. Bacterial endophytes as biological control agents for black pod rot of cacao. *Phytopathology*, Vol.95(6) Suppl S, pp.S171-S171

Mmbaga, M., S. Gurung, A. Maheshwari. 2018. Screening of Plant Endophytes as Biological Control Agents against Root Rot Pathogens of Pepper (Capsicum annum L.) *Journal of Plant Pathology & Microbiology*, Vol.09(03)

Nguyen, N. H., D. P. Smith, K. G. Peay, P. G. Kennedy. 2015. Parsing ecological signal from noise in next generation amplicon sequencing. *New Phytol*. 205:1389–1393

Oksanen, J., F. G. Blanchet, M. Friendly, R. Kindt, P. Legendre, D. Mcglinn, P. R. Minchin, R. B. O. Hara, G. L Simpson, P. Solymos, M. H. H. Stevens, E. Szoecs 2017. vegan: Community ecology package. R package version 2.0-2.

Ou, T., W. Xu, F. Wang, G. Strobel, Z. Zhou, Z. Xiang, J. Liu, J. Xie. A Microbiome Study Reveals Seasonal Variation in Endophytic Bacteria Among Different Mulberry Cultivars. Computational and structural Biotechnology Journal. Vol 17, pp 1091-1100. https://doi.org/10.1016/j.csbj.2019.07.018

Patwardhan A., S. Ray, A. Roy. 2014. Molecular markers in phylogenetic studies—a review. *J Phylogen Evol Biol*. 2014; 2:131.

Peterson, E. A., J. W. Rouatt, H. Katznelson. 1965. Microorganisms in the root zone in relation to soil moisture. Canadian Journal of Microbiology, 11(3): 483-489. https://doi.org/10.1139/m65-063

Qiagen. 2019. CLC Microbial Genomics Module User Manual 20.0 for Windows, macOS, and Linux. Available online at: http://resources.qiagenbioinformatics.com/manuals/clcmgm/current/User_Manual.pdf

Quinn, T. P., I. Erb, M. F. Richardson, T. M. Crowley. 2018. Understanding sequencing data as compositions: an outlook and review. *Bioinformatics*, Volume 34, Issue 16, Pages 2870–2878, https://doi.org/10.1093/bioinformatics/bty175

R Core Team. 2020. R: A language and environment for statistical computing. *R Foundation for Statistical Computing*, Vienna, Austria. URL http://www.R-project.org/.

Rabiey, M., L. E. Hailey, S. R. Roy. 2019. Endophytes vs tree pathogens and pests: can they be used as biological control agents to improve tree health? *Eur J Plant Pathol* 155, 711–729. https://doi.org/10.1007/s10658-019-01814-y

Ramette, A. 2007. Multivariate analyses in microbial ecology. *FEMS Microbiology Ecology*, 62(2), 142–160. doi: 10.1111/j.1574-6941.2007.00375.x

Rodríguez-Blanco A., M. Sicardi, L. Frioni. 2015. Plant genotype and nitrogen fertilization effects on abundance and diversity of diazotrophic bacteria associated with maize (Zea mays L.). *Biol. Fertil. Soils*. 2015;51:391–402. doi: 10.1007/s00374-014-0986-8.

Santoyo, G., G. Moreno-Hagelsieb, M. D. C. Orozco-Mosqueda, and B. R. Glick. 2016. Plant growth-promoting bacterial endophytes. *Microbiological Research*, 183, 92–99. doi: 10.1016/j.micres.2015.11.008

Seghers, D., K. Verthe, D. Reheul, R. Bulcke, S. D. Siciliano, W. Verstraete, E. M. Top. 2003. Effect of long-term herbicide applications on the bacterial community structure and function in an agricultural soil. *FEMS Microbiology Ecology*, Vol. 46, Iss 2, pp 139-146. https://doi.org/10.1016/S0168-6496(03)00205-8

Shen, J., C. R. Chen, T. Lewis. 2016. Long term repeated fire disturbance alters soil bacterial diversity but not the abundance in an Australian wet sclerophyll forest. *Sci. Rep*. 19639. 10.1038/srep19639

Shen, S. Y. and R. Fulthorpe. 2015. Seasonal variation of bacterial endophytes in urban trees. Front. Microbiol., https://doi.org/10.3389/fmicb.2015.00427

Stackebrandt, E., J. Ebers. 2006. Taxonomic parameters revisited: tarnished gold standards. *Microbiol Today*. 33:152–1555.

Stanley A. G., P. W. Dunwiddie, T. N. Kaye. 2011. Restoring invaded pacific Northwest prairies: Management recommendations from a region-wide experiment. *Northwest Sci.* 85:233–246, doi:10.3955/046.085.0212.

Stone, J.K., C. W. Bacon & J. F. White. 2000. An overview of endophytic microbes: endophytism defined. *Microbial Endophytes*, pp. 3–29. Marcel Dekker, New York, NY, USA.

Strange, R. N. and P. R. Scott. 2005. Plant Disease: A Threat to Global Food Security. *Annual Review of Phytopathology*, 43(1), 83–116. doi: 10.1146/annurev.phyto. 43.113004.1 33839

Stuart, A., C. Kandd, R. M. Stuart, I. C. Pimentel. 2018. Effect of agrochemicals on endophytic fungi community associated with crops of organic and conventional soybean (Glycine max L. Merril). *Agriculture and Natural Resources* 52, 388–392.

U.S. Fish and Wildlife Service (UFWS). 2010. Recovery plan for the prairie species of western Oregon and southwestern Washington.

Vandenkoornhuyse, P., A. Quaiser, M. Duhamel, A. Le Van, A. Dufresne. 2015. The importance of the microbiome of the plant holobiont. *New Phytol*., 206, pp. 1196-1206

Walitang, D. I., C. G. Kim, S. Jeon, Y. Kang, T. Sa. 2018. Conservation and transmission of seed bacterial endophytes across generations following crossbreeding and repeated inbreeding of rice at different geographic locations. *MicrobiologyOpen*, 8(3). doi: 10.1002/mbo3.662

Wani, Z. A., N. Ashraf, T. Mohiuddin, S. Riyaz-Ul-Hassan. 2015. Plant-endophyte symbiosis, an ecological perspective. *Appl Microbiol Biotechnol*. 99:2955–2965, doi:10.1007/s00253-015-6487-3.

Wemheuer, F., K. Kaiser, P. Karlobsky, R. Daniel, S. Vidal, B. Wemheuer. 2017. Bacterial endophyte communities of three agricultural important grass species differ in their response towards management regimes. *Scientific reports* 7 (1), 1-13

Weiss, S., Z. Z. Xu, S. Peddada. 2017. Normalization and microbial differential abundance strategies depend upon data characteristics. *Microbiome* 5, 27

Wickham, H. 2016. ggplot2: Elegant Graphics for Data Analysis. Springer-Verlag New York. ISBN 978-3-319-24277-4, https://ggplot2.tidyverse.org.

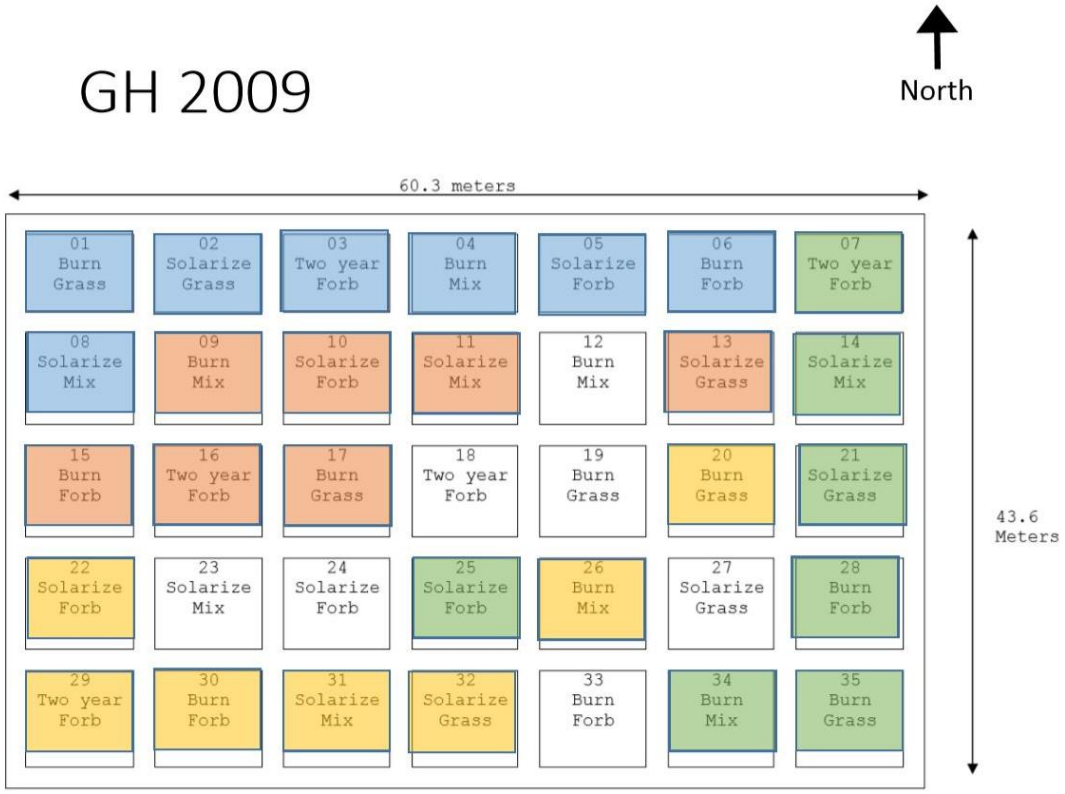
Yan Y., E. E. Kuramae, M. de Hollander, P. G. L. Klinkhamer, J. A. van Veen. 2016. Functional traits dominate the diversity-related selection of bacterial communities in the rhizosphere. *ISME J.*, 11 pp. 56-66.

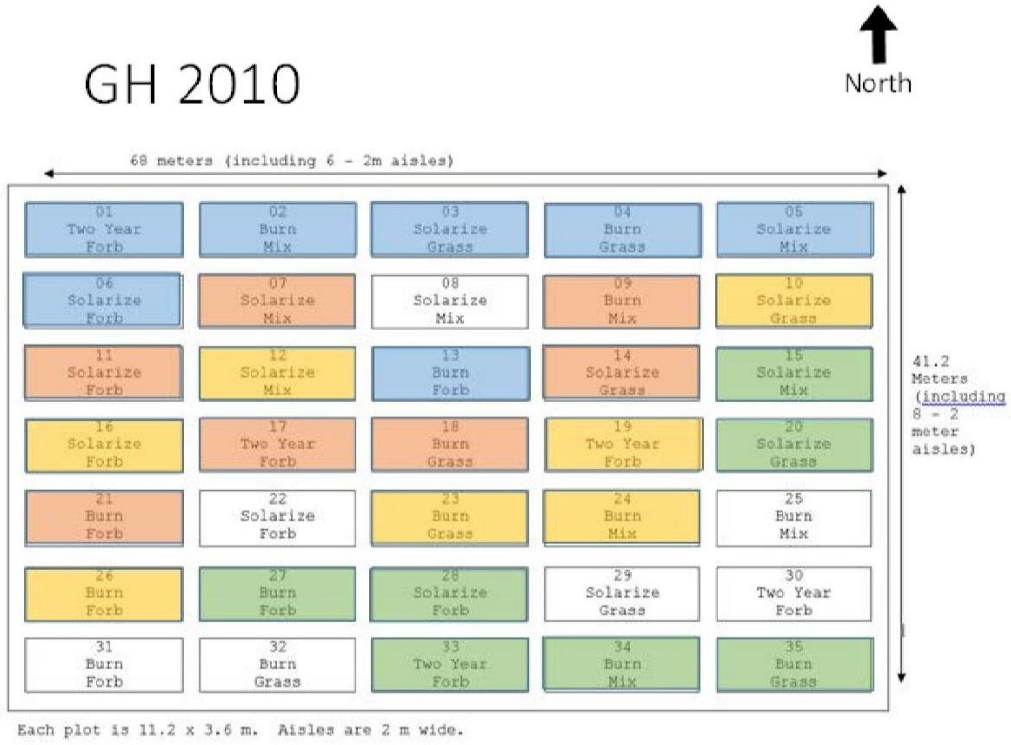
Zhang, Q., J. J. Acuña, N. G. Inostroza. 2019. Endophytic Bacterial Communities Associated with Roots and Leaves of Plants Growing in Chilean Extreme Environments. *Sci Rep* **9,** 4950

Zilber-Rosenberg I. and E. Rosenberg. 2008. Role of microorganisms in the evolution of animals and plants: The hologenome theory of evolution. *FEMS Microbiol Rev*.;32:723–35.

16s Illumina Amplicon Protocol, 2018. Retrieved from https://earthmicrobiome.org/protocols-and-standards/16s/.

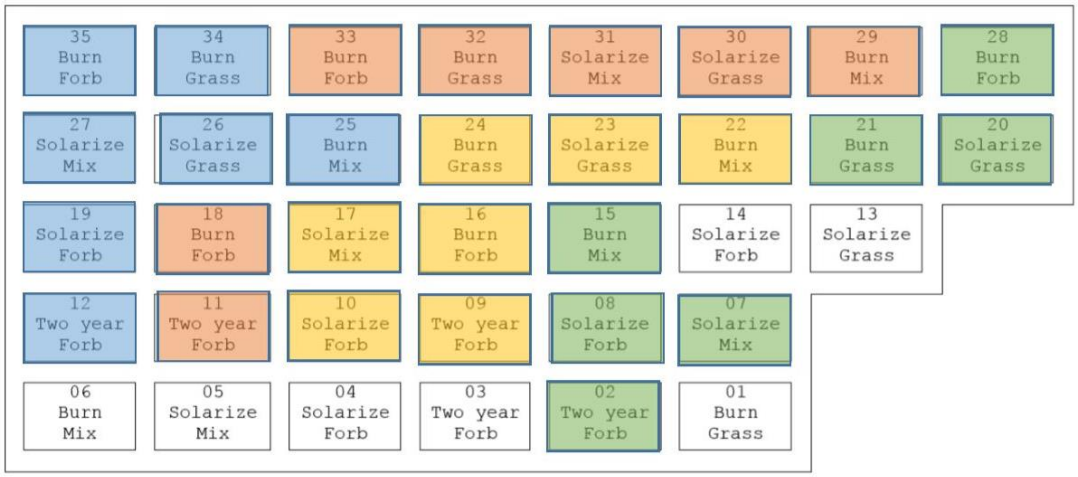
Appendix 1.A:


GH 2009
North
60.3 meters
01 Burn Grass
02 Solarize Grass
03 Two year Forb
04 Burn Mix
05 Solarize Forb
06 Burn Forb
07 Two year Forb
08 Solarize Mix
09 Burn Mix
10 Solarize Forb
11 Solarize Mix
12 Burn Mix
13 Solarize Grass
14 Solarize Mix
15 Burn Forb
16 Two year Forb
17 Burn Grass
18 Two year Forb
19 Burn Grass
20 Burn Grass
21 Solarize Grass
22 Solarize Forb
23 Solarize Mix
24 Solarize Forb
25 Solarize Forb
26 Burn Mix
27 Solarize Grass
28 Burn Forb
29 Two year Forb
30 Burn Forb
31 Solarize Mix
32 Solarize Grass
33 Burn Forb
34 Burn Mix
35 Burn Grass
43.6 Meters
Each plot is 6.32 x 6.32 m. Aisles are 2 m wide.


GH 2010
North
68 meters (including 6 - 2m aisles)
01 Two Year Forb
02 Burn Mix
03 Solarize Grass
04 Burn Grass
05 Solarize Mix
06 Solarize Forb
07 Solarize Mix
08 Solarize Mix
09 Burn Mix
10 Solarize Grass
11 Solarize Forb
12 Solarize Mix
13 Burn Forb
14 Solarize Grass
15 Solarize Mix
16 Solarize Forb
17 Two Year Forb
18 Burn Grass
19 Two Year Forb
20 Solarize Grass
21 Burn Forb
22 Solarize Forb
23 Burn Grass
24 Burn Mix
25 Burn Mix
26 Burn Forb
27 Burn Forb
28 Solarize Forb
29 Solarize Grass
30 Two Year Forb
31 Burn Forb
32 Burn Grass
33 Two Year Forb
34 Burn Mix
35 Burn Grass
41.2 Meters (including 8 - 2 meter aisles)
Each plot is 11.2 x 3.6 m. Aisles are 2 m wide.

North

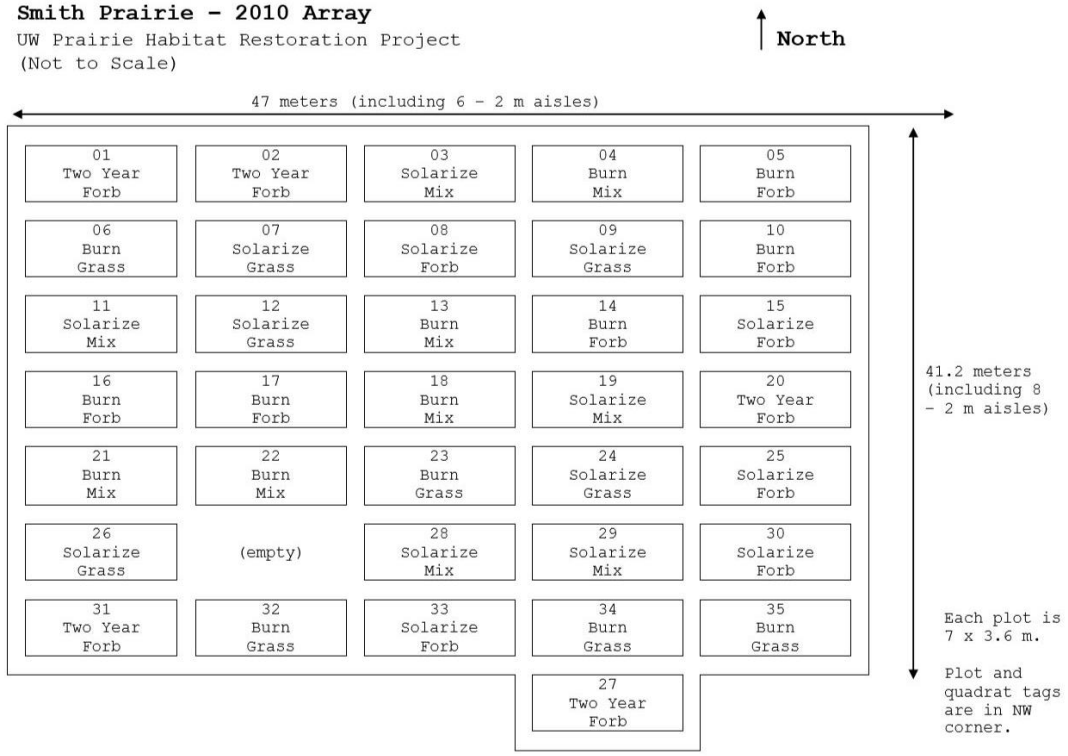
GH 2011



Each plot is 11.2 x 3.6 m. Aisles are 2 m wide.

Appendix 1.B:

Smith Prairie – 2010 Array
UW Prairie Habitat Restoration Project
(Not to Scale)



Smith Prairie – 2011 Array
UW Prairie Habitat Restoration Project
(Not to Scale)

Table 1.18: Summary of full scientific name, code, and taxonomy of the 16 Puget prairie plant species used in this study.

Plant Species	Code	Kingdom	Division	Class	Order	Family	Genus
Achillea millefolium	ACMI	Plantae	Magnoliophyta	Magnoliopsida	Asterales	Asteraceae	Achillea
Aquilegia formosa	AQFO	Plantae	Magnoliophyta	Magnoliopsida	Ranunculales	Ranunculaceae	Aquilegia
Aster curtisii	ASCU	Plantae	Tracheophyta	Magnoliopsida	Asterales	Asteraceae	Symphyotrichum
Balsamorhiza deltoidea	BADE	Plantae	Magnoliophyta	Magnoliopsida	Asterales	Asteraceae	Balsamorhiza
Camassia quamash	CAQU	Plantae	Magnoliophyta	Liliopsida	Liliales	Liliaceae	Camassia
Castilleja levisecta	CALE	Plantae	Magnoliophyta	Magnoliopsida	Scrophulariales	Orobanchaceae	Castilleja
Cerastium arvense	CEAR	Plantae	Magnoliophyta	Magnoliopsida	Caryophyllales	Caryophyllaceae	Cerastium
Delphinium menziesii	DEME	Plantae	Magnoliophyta	Magnoliopsida	Ranunculales	Ranunculaceae	Delphinium
Erigeron speciosus	ERSP	Plantae	Magnoliophyta	Magnoliopsida	Asterales	Asteraceae	Erigeron
Eriophyllum lanatum	ERLA	Plantae	Magnoliophyta	Magnoliopsida	Asterales	Asteraceae	Eriophyllum
Festuca idahoensis ssp. Roemeri	FERO	Plantae	Magnoliophyta	Liliopsida	Cyperales	Poaceae	Festuca
Lomatium triternatum	LOTR	Plantae	Magnoliophyta	Magnoliopsida	Apiales	Apiaceae	Lomatium
Lomatium utriculatum	LOUT	Plantae	Magnoliophyta	Magnoliopsida	Apiales	Apiaceae	Lomatium
Lupinus lepidus	LULE	Plantae	Magnoliophyta	Magnoliopsida	Fabales	Fabaceae	Lupinus
Potentilla gracilis	POGR	Plantae	Magnoliophyta	Magnoliopsida	Rosales	Rosaceae	Potentilla
Symphoricarpos albus	SYAL	Plantae	Magnoliophyta	Magnoliopsida	Dipsacales	Caprifoliaceae	Symphoricarpos

Appendix 1.D:

UMGC Illumina Sequencing Protocol

Three cycles of PCR are performed on the samples: qPCR, PCR1, and PCR2. qPCR was performed using the following cycling conditions:

95°C for 5 minutes

35 cycles:

- 98°C for 20 seconds
- 55°C for 15 seconds
- 72°C for 1 minute

72°C for 5 minutes

Hold at 4°C

After qPCR, samples are normalized to 167,000 molecules/ul. 3 ul is used during PCR1. The Meta_V4_515F/Meta_V4_806R primer pairs are incorporated into the samples during PCR1. PCR1 was performed using the following cycling conditions:

95°C for 5 minutes

25 cycles:

- 98°C for 20 seconds
- 55°C for 15 seconds
- 72°C for 1 minute

72°C for 5 minutes

Hold at 4°C

After PCR1, products are diluted 1:100 and 5 ul of this diluted product is used in PCR2. "Different combinations of forward and reverse indexing primers" is incorporated into the samples during PCR2. PCR2 was performed using the following cycling conditions:

95°C for 5 minutes

10 cycles:

- 98°C for 20 seconds
- 55°C for 15 seconds
- 72°C for 1 minute

72°C for 5 minutes

The UMGC Sequencing protocol is as follows:

"Pooled sample was denatured with NaOH, diluted to 8 pM in Illumina's HT1 buffer, spiked with 15% PhiX, and heat denatured at 96C for 2 minutes immediately prior to loading. A MiSeq 600 cycle v3 kit was used to sequence the sample."

Two fastq files were generated per sample, a pair of forward sequence and reverse sequence for each sample. Compressed fastq files (.gz) were made available for download from the UMGC.

Appendix 1.E:

Table 1.19: Total number of OTUs derived from all samples after removal of potentially contaminating OTUs. OTUs were commonly identified to the phyla, class, and order the OTU was assigned to. NA values indicate bacterial OTUs that could not be identified.

Phyla	Number of OTU's
NA	3390
Firmicutes	1231
Proteobacteria	1166
Bacteroidetes	593
Actinobacteria	484
Cyanobacteria	161
Patescibacteria	74
Verrucomicrobia	68
Acidobacteria	42
Planctomycetes	42
Chloroflexi	23
Armatimonadetes	18
Tenericutes	15
FBP	12
Epsilonbacteraeota	9
WPS-2	6
Spirochaetes	5
Chlamydiae	4
Gemmatimonadetes	4
Thaumarchaeota	4
Deinococcus-Thermus	3
Euryarchaeota	3
Fusobacteria	3
Deferribacteres	2
Fibrobacteres	2
Lentisphaerae	1

Class	Number of OTU's
NA	3398
Clostridia	928
Alphaproteobacteria	627
Bacteroidia	593
Gammaproteobacteria	449
Actinobacteria	379
Bacilli	240
Oxyphotobacteria	154
Deltaproteobacteria	90
Saccharimonadia	74
Verrucomicrobiae	68
Erysipelotrichia	56
Thermoleophilia	49
Planctomycetacia	36
Acidimicrobiia	33
Acidobacteriia	28
Coriobacteriia	23
Mollicutes	15
Armatimonadia	11
Blastocatellia (Subgroup 4)	10

Class	Number of OTU's
Campylobacteria	9
Chloroflexia	9
uncultured bacterium	9
Melainabacteria	6
Phycisphaerae	6
Fimbriimonadia	5
Negativicutes	5
Chlamydiae	4
KD4-96	4
Ktedonobacteria	4
Nitrososphaeria	4
Brachyspirae	3
Deinococci	3
Fusobacteriia	3
Methanobacteria	3
Subgroup 6	3
TK10	3
Anaerolineae	2
Deferribacteres	2
Fibrobacteria	2

Class	Number of OTU's
Gemmatimonadetes	2
Limnochordia	2
Longimicrobia	2
Spirochaetia	2
Chthonomonadetes	1
Dehalococcoidia	1
Lentisphaeria	1
metagenome	1
Sericytochromatia	1
Thermoanaerobaculia	1
uncultured	1

Order	Number of OTU's
NA	3403
Clostridiales	928
Bacteroidales	263
Rhizobiales	229
Betaproteobacteriales	172
Sphingomonadales	164
Chloroplast	148
Enterobacteriales	129
Micrococcales	126
Lactobacillales	121
Bacillales	118
Cytophagales	111
Rickettsiales	95
Sphingobacteriales	91
Propionibacteriales	74
Pseudomonadales	74
Saccharimonadales	74
Chitinophagales	68
Flavobacteriales	60
Erysipelotrichales	56

Order	Number of OTU's
Corynebacteriales	49
Acetobacterales	45
Frankiales	45
Verrucomicrobiales	39
Caulobacterales	36
Myxococcales	36
Solirubrobacterales	36
Microtrichales	30
Bdellovibrionales	28
Xanthomonadales	28
Chthoniobacterales	25
Isosphaerales	24
Coriobacteriales	23
Acidobacteriales	22
Micromonosporales	20
Rhodobacterales	20
Kineosporiales	17
Pseudonocardiales	17
Desulfovibrionales	14
Streptomycetales	14

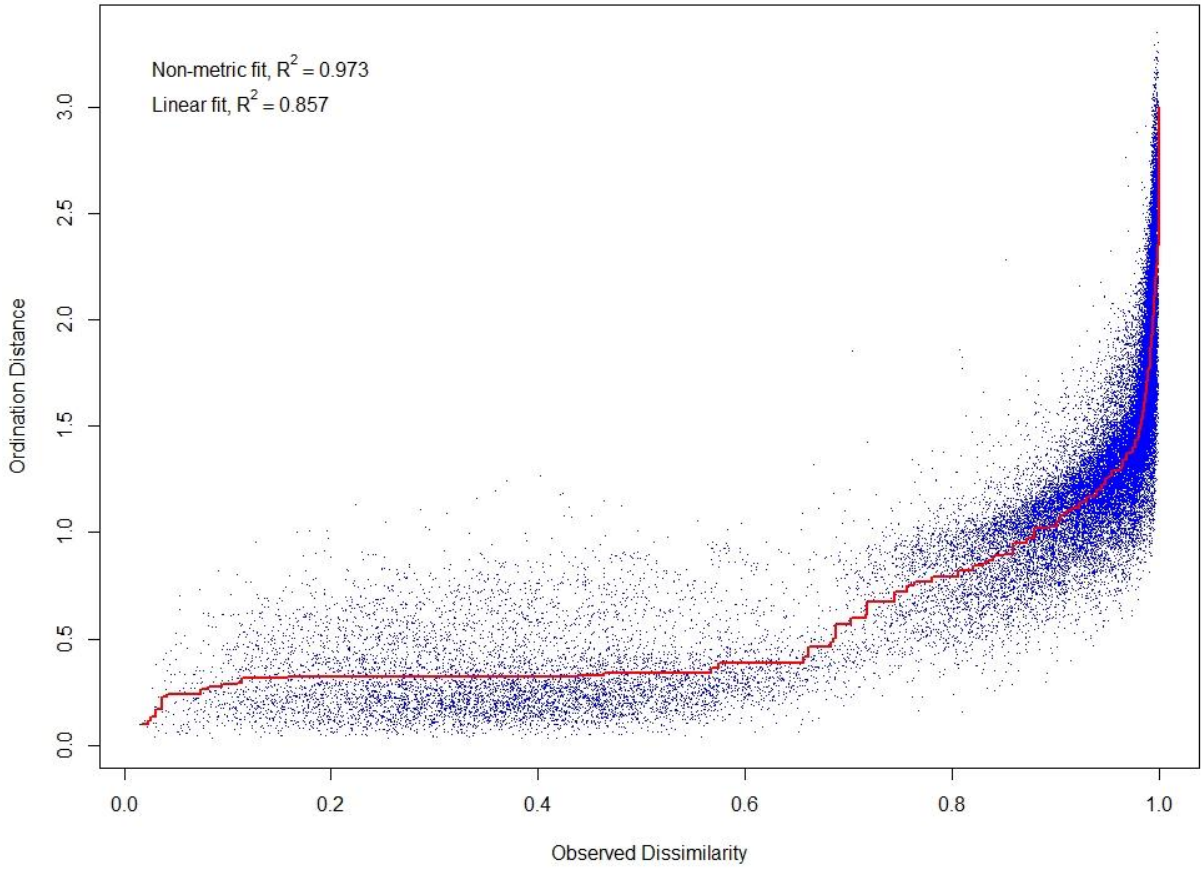
Order	Number of OTU's
Gaiellales	13
Legionellales	13
uncultured bacterium	13
Oligoflexales	12
Armatimonadales	11
Micavibrionales	11
Mollicutes RF39	10
Campylobacterales	9
Rhodospirillales	9
Bifidobacteriales	8
Blastocatellales	8
Pasteurellales	7
Thermomicrobiales	7
Gemmatales	6
Nostocales	6
Paracaedibacterales	6
Solibacterales	6
Tepidisphaerales	6
Fimbriimonadales	5
Gammaproteobacteria Incertae Sedis	5

Order	Number of OTU's
Gastranaerophilales	5
Pirellulales	5
Selenomonadales	5
Streptosporangiales	5
Actinomycetales	4
Chlamydiales	4
Diplorickettsiales	4
Mycoplasmatales	4
Nitrososphaerales	4
Aeromonadales	3
Brachyspirales	3
Fusobacteriales	3
IMCC26256	3
Methanobacteriales	3
R7C24	3
Tistrellales	3
Alteromonadales	2
C0119	2
Caedibacterales	2
Cellvibrionales	2

Order	Number of OTU's	Order	Number of OTU's
Deferribacterales	2	KF-JG30-C25	1
Deinococcales	2	metagenome	1
Elsterales	2	Micropepsales	1
Fibrobacterales	2	Opitutales	1
Gemmatimonadales	2	Piscirickettsiales	1
Kallotenuales	2	Planctomycetales	1
Ktedonobacterales	2	Pyrinomonadales	1
Limnochordales	2	Reyranellales	1
Longimicrobiales	2	S085	1
Methylacidiphilales	2	Salinisphaerales	1
Oceanospirillales	2	SAR11 clade	1
Spirochaetales	2	SBR1031	1
uncultured	2	Thermales	1
Alphaproteobacteria Incertae Sedis	1	Thermoanaerobaculales	1
Anaeroplasmatales	1	uncultured Acidobacteria bacterium	1
Ardenticatenales	1	uncultured beta proteobacterium	1
Azospirillales	1	Unknown Order	1
Cardiobacteriales	1	Vampirovibrionales	1
Chthonomonadales	1	Vibrionales	1
DS-100	1	Victivallales	1

Appendix 1.F:

Figure 1.11 Three-dimensional Stress Plot for Species and Family Ordinations



Legend	
Green	Significant Result
Red	Non-significant Result

	ACMI	AQFO	ASCU	BADE	CAQU	CALE	CEAR	DEME	ERSP	ERLA	FERO	LOTR	LOUT	LULE	POGR
AQFO	0.001														
ASCU	0.04	0.002													
BADE	0.077	0.001	0.173												
CAQU	0.001	0.001	0.001	0.001											
CALE	0.001	0.001	0.001	0.001	0.001										
CEAR	0.001	0.001	0.001	0.001	0.001	0.001									
DEME	0.001	0.002	0.001	0.001	0.001	0.001	0.001								
ERSP	0.054	0.001	0.144	0.113	0.001	0.001	0.001	0.001							
ERLA	0.199	0.001	0.055	0.125	0.001	0.001	0.001	0.001	0.04						
FERO	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001					
LOTR	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001				
LOUT	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.298			
LULE	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001		
POGR	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	
SYAL	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001

Table 1.20: Summary of pairwise test performed in R. Differences in bacterial OTU composition were tested on the basis of plant species groups, where statistically significant differences were found between some plant species. Green boxes indicate that the differences in bacterial OTU composition between two species were statistically significant. Red boxes indicates that the differences in bacterial OTU composition between two species were not statistically significant.

Appendix 1.H:

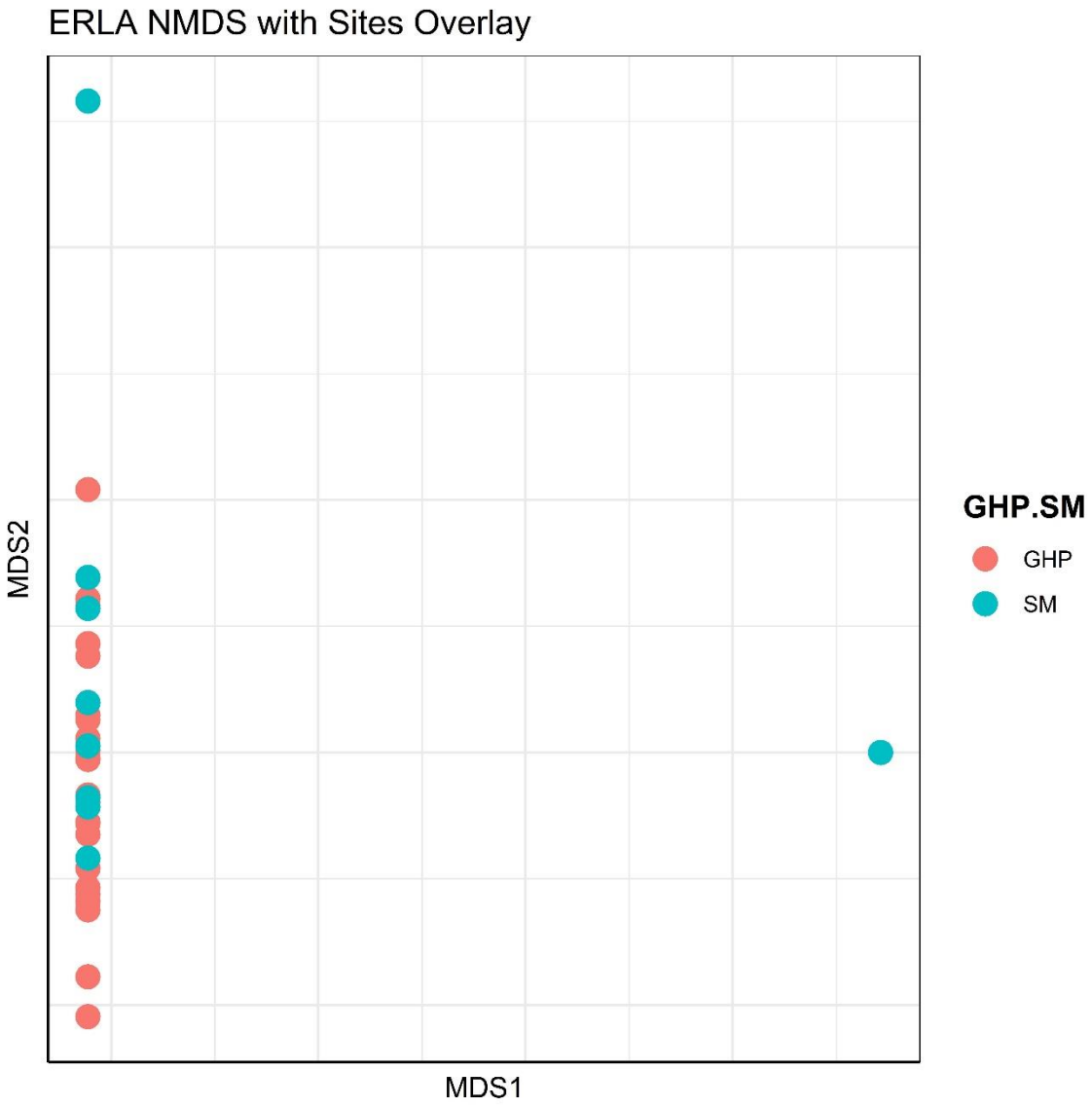


Figure 1.12: Two-dimensional NMDS ordination of bacterial OTU abundance of *Eriophyllum lanatum* samples with site location overlay prior to the removal of sample 0439 (stress = 8.520031e-05). MDS1 is the axis that explains the most variation across the dataset, and MDS2 is the axis that explains the second most variation across the dataset. Colors represent the sample location from which the sample was derived. Stress for *Eriophyllum lanatum* was concerningly low, and thus the NMDS graph was unable to capture the spread of the datapoints. This plot exists to demonstrate how sample 0439 was identified as an outlier and removed from analysis.

Appendix 1.I:

Table 1.21: Legend for Pairwise Tests

Legend	
Green	Significant Result
Red	Non-significant Result

Table 1.22: Initial Disturbance Regime Pairwise Test

PAIRWISE	Burn	Solarize
Solarize	0.006	N/A
Two Year	0.45	0.61

Table 1.23: Disturbance Regime Pairwise Test

PAIRWISE	Annual Early Burn	Annual Late Burn	Triannual Early Burn	Mowed Annually
Annual Late Burn	0.0667			
Triannual Early Burn	0.061	0.067		
Mowed Annually	1	0.667	0.03	
Triannual Late Burn	0.258	0.692	0.007	0.274

Table 1.24: Date of Last Treatment Pairwise Test

PAIRWISE	2014 Treatment	2017 Treatment
2017 Treatment	0.046	
2018 Treatment	0.319	0.009

Chapter 2: Transfer of Bacteria between *Castilleja Levisecta* and Host Plants

Abstract

The Puget prairie ecosystem is a charismatic and ecologically important feature of North America's Pacific Northwest ecosystems, but faces mounting threats from land use change, invasion of non-native plant species, and climate change. Solutions to these threats require enhanced knowledge of these systems, and novel approaches to the particular challenges that impede the recovery of prairie ecosystems. The interactions between plants and microbes are increasingly appreciated, as recent discoveries have placed a spotlight on the ways in which the bacterial microbiome influences plant growth. However, the bacteria community of many plant systems remains unexplored, as well as the ways in which bacteria may be traveling within these systems.

I performed an observational field study to determine if bacteria are able to travel between hemiparasitic plants and their host plants using haustorial root connections. I used Illumina sequencing, CLC Workbench, and R technologies to investigate similarities between the bacterial community profiles of hemiparasitic plants and their hosts. I compared differences in bacterial Operational Taxonomic Unit (OTU) composition between parasitic plants and their (assumed) hosts, and parasitic plants and non-hosts, to examine if the microbial community could be influenced by *Castilleja levisecta* parasitism. For all trios capable of testing, regardless of species sample size, there was a significant effect of parasitism on bacterial OTU composition. I then tested individual plant species with large sample sizes; for *Lomatium utriculatum* and *Eriophyllum lanatum*, Host.Parasite Bray-Curtis distances are significantly smaller than NonHost.Parasite Bray-Curtis distances, indicating that the bacterial OTU composition of *Castilleja levisecta* and the host samples that were parasitized by a *C. levisecta* plant were more

similar to each other than *C. levisecta* and non-host samples. Although *Balsamorhiza deltoidea* and *Festuca roemeri* had non-significant differences between the Host.Parasite Bray-Curtis distances and NonHost.Parasite Bray-Curtis distances, their p values were small, suggesting that parasitism may have an effect on OTU composition; larger sample sizes are needed to strengthen this test. Finally, for *Camassia quamash,* the differences between the Host.Parasite Bray-Curtis distances and NonHost.Parasite Bray-Curtis distances were not significant by a large margin. The effect of parasitism appears to depend on a species by species basis, likely due to the efficacy of *C. levisecta* to parasitize certain species. The knowledge gained through this research will directly benefit land managers assisting the recovery of ecosystems containing parasitic plants, as well enhance our understanding of the ways in which bacteria use their associations with plants to colonize new environments.

Introduction

Ecosystems that have been threatened and endangered by environmental degradation are of particular focus for ecological restoration. One such threatened system is the Puget prairie ecosystem, which exist in the Pacific Northwest region of the United States. These landscapes occur in Mediterranean climate systems, which experience hot, dry summers and mild, warm winters (Klausmeyer and Shaw 2009). Mediterranean prairies in the Pacific Northwest are rich in biodiversity, but have declined to less than 10% of their historical range (UFWS 2010). Altered fire regimes, land use change, climate change, invasions of non-native species, and habitat fragmentation, amongst other threats, imperil the survival of Puget prairie ecosystems. Without changes in management, it is likely that Puget prairie ecosystems will continue to decline and these systems may fail to persist into the future (Dunwiddie and Bakker 2011). As a result, natural resource management organizations across the Pacific Northwest have considered prairie ecosystems to be high priority areas for ecological restoration (UFWS 2010).

Ecologists and land managers aiming to restore Puget prairie ecosystems have conducted research and experiments on Puget prairies for decades, accruing a wealth of information on subjects such as floral and faunal species composition, applications of land management techniques, interspecific interactions, the effects of land use change, and future projections for Puget prairie ecosystems, among other studies (Bachelet et al 2001; Stanley et al. 2011; Delvin 2013; Klausmeyer & Shaw 2009; Dunwiddie and Bakker 2011). However, there remains a lack of knowledge of community interactions on smaller scales; the microbial ecology of Puget prairie ecosystems remains an understudied aspect of these systems. Bacterial endophytes and pathogens have been detected in the plant tissues of every plant ever surveyed for the presence of bacterial. These bacteria are known to interact with their hosts, and can have profound effects on

the health of individual plants. Thus, it is of critical importance to understand the microbial community of Puget prairie plants.

Bacterial endophytes are known to occur in an extensive number of plant species, and have been recorded living in the space between cells within plant stem, leaf, and root tissues. Many of these stem and leaf inhabiting bacterial endophyte species have been determined to have nutrient provisioning plant growth promoting traits (Hardoim et al. 2008). The microbiome that can be found within the stem and leaf tissue is typically less diverse than that of root tissue, and generally hosts a smaller abundance of bacteria than root tissue (Zhang et al. 2019; Liu et al. 2017). The majority of bacterial endophytes discovered within the plant tissue are derived from the surrounding environment, since vertical transmission (transmission of bacteria from parent plant to seed) is selective in the species of bacteria that colonize the seed (Walitang et al. 2018). The rhizosphere acts as a main contact zone for root inhabiting endophytes (Yan et al. 2016). Exposed entrances to inner plant tissues, such as stomata or wounds, allow both pathogenic and endophytic bacteria to colonize the intercellular space (Frank et al. 2017).

Current research suggests that species of plants have co-evolved alongside their root nodule or plant tissue inhabiting bacterial endophytes, and have thus developed complex host/symbiont relationships which benefit both the host plant and the endophytic bacteria (Clay and Schardl 2002). However, it is of particular research interest to examine if bacterial endophytes are able to be isolated from their original host plants and applied to target plant species which were not originally recognized as natural host plants. Bacterial endophytes that are able to be taken from their original host plants and transferred to plants used in restoration efforts may continue to provide plant growth promoting substances. Substantial and consistent increases in plant growth due to inoculation with bacterial endophytes will help determine if inoculating

plants with plant growth promoting endophytes could act as a potential means to reduce disease, increase nutrient uptake, and generally enhance the growth of target plants.

Bacterial endophytes use various methods to colonize host plants. Vertical transmission, where seeds, rhizomatous sprouts, and other forms of plant progeny are inoculated with a small number of endophytic species inherited by their parent plant, act as a plants first encounter with endophytic bacteria. While these first bacterial endophyte colonizers are able to assist plant growth to a certain degree, not many bacteria are transferred through vertical transmission compared to the number of bacterial endophytes acquired through entry points in the tissue that the plant develops later in life (Parke 1991). Horizontal transmission is the transfer of bacterial endophytes from the surrounding environment into plant tissues. Subsequent host plant inoculation of bacterial endophytes after first colonization (vertical transmission) typically occurs via recruitment of bacteria through entry points in the root, shoot and leaf as horizontal transmission (Bulgarelli 2013). Bacteria in the soil, and particularily those bacteria associated with the rhizosphere, are often the most common candidates for root colonization, as rhizosphere bacteria maintain positive symbiotic relationships with plants and occupy spaces closest to potential infection points in the root. As root hairs exit the epidermis of the main root, the hairs form gaps between the cell walls which allow opportunities for bacteria to enter the root tissue. From the root, bacteria can make their way into phloem or xylem and disperse further throughout the plant, making their way into stem and leaf tissues (Liu et al. 2017).

Another horizontal transfer mechanism that bacterial endophytes use to colonize plant tissues is via the phyllosphere. Above ground plant tissues, including leaves, stems, flowers, and trunks, among other tissue types, comprise the phyllosphere. Entrances from the phyllosphere into the intercellular space within the plant are exploited by endophytic bacteria and pathogens

alike to colonize plants from the exterior environment. Common entrances that facilitate bacterial colonization include the stomatal organs in the leaf and open wounds caused by herbivores (Santoyo 2016). There is considerable evidence to suggest that plants are able to identify bacterial species, and intentionally exclude some bacteria from passage through the stomata (Frank et al. 2017). However, passage of bacteria through wounds are less controlled and thus more vulnerable to bacteria the plant would otherwise exclude from its intercellular space. The function of leaf tissue is naturally different from root or stem tissues, and plant leaves thus generally host a different microbiome than can be found throughout the same plants' roots or stem tissues.

A mechanism that has recently become a research focus is the transfer of endophytes directly from plant tissue to plant tissue. Parasitic plants form physical connections to their host plants, absorbing nutrients, growth hormones, defense hormones, and other crucial compounds from their host plants. These connections are typically formed by the parasite as it penetrates the host plant's conductive system and absorbs compounds from the host plant phloem or xylem. Recent studies have discovered that bacterial endophytes may be among the compounds that a parasitic plant absorbs from its host plant, as endophytes have been discovered inhabiting plant phloem and xylem streams. In a recent study, *Orobanche hederae*, a holoparasite of *Hedera* species, was found to have a microbiome that was "derived but distinct" from the microbiome of its host plant, indicating that bacterial transfer between the plants had occurred during parasitism (Fitzpatrick and Schneider 2019). Additionally, in a study of holoparasitic *Phelipanche aegyptiaca* plants and its host *Solanum lycopersicum*, while the endophyte community of *P. aegyptiaca* was distinct from its host pre-parasitism, the *P. aegyptiaca* microbiome became indistinguishable from its host after parasitism, indicating that transfer of bacterial endophytes

was highly likely (Kruh et al. 2017). The pre-parasitism community of endophytes of *P. aegyptiaca* plants was determined by examining pre-haustorium stage seedlings.

Castilleja levisecta is a threatened hemiparasitic plant species that inhabits native Puget prairie ecosystems (Schmidt 2016). *C. levisecta* parasitizes a wide range of host plants, forming root connections to host plants tethered by haustoria (Schmidt 2016). Essential nutrients, plant defense compounds, and other beneficial supplements are drawn from the host plant into the parasitizing *C. levisecta* in a similar way to other parasitic plant species. As bacterial transfer as been found to occur between other parasitic plants and their host plants, I hypothesized that bacteria may also transfer between *C. levisecta* and its host plants in the field. To test this hypothesis, I collected 387 samples from Puget prairie plants and examined the bacterial OTU composition of 5 potential host plant species. I collected samples from *C. levisecta* parasitizing host plants, the parasitized host plants, and non- parasitized plants belonging to the same species as the host plant in close proximity to *C. levisecta* and host pairs. After reviewing the scientific literature, I hypothesized that the bacterial OTU composition of *C. levisecta* parasites would be more similar to the bacterial OTU composition of their host plants than to non-hosts of the same species collected from the same local area.

Methods

Study Area

I studied two locations in western Washington State. The primary study site, from which the majority of the samples were collected, are research plots that had already been established in the Glacial Heritage Preserve (Figure 1.1). GHP is owned by Thurston County and the Washington Department of Fish and Wildlife, and managed by the Center for Natural Lands Management. The second study site is at Smith Prairie (SM), on Whidbey Island in Island County. SM is owned and managed by the Pacific Rim Institute for Environmental Stewardship. Experimental restoration plots were established at both sites about a decade ago (Figures 1.2, 1.3). Research plots were established for use as restoration experiments in July 2008 and are a part of an ongoing study of Puget prairie restoration. Site preparation and seeding mix differ between plots within the prairie; data on the plot that each plant sample was collected was recorded in the metadata. Plants removed from these plots would not have a detrimental impact on one of the few remaining natural Puget prairies existing in Washington State.

Sample Collection and Selection

In May and June 2019, 328 prairie plant stem samples were collected from GHP and 59 samples were collected from SM. Each sample was either a leaf or a stem of a plant, but only stems were used in the set of samples submitted for sequencing. I recorded data on the date the sample was collected, its collection location (site, array, and plot number), and the taxonomic identity of the plant. Plant samples were taken from 16 different prairie plant species (Appendix 1.C). However, only 5 species were examined from this dataset.

The sampling process was as follows. Eight trips to the Glacial Heritage Preserve were made throughout the months of May and June. A healthy plant was identified and selected for

use in the field (plants with unknown identity were collected and preserved for later identification upon return to Seattle). Each sample was collected by taking a stem cutting of the plant with sterilized scissors, close to where the stem reaches the roots. As much stem material as could fit in one Eppendorf tube was collected. Samples were surface sterilized in the field to remove external bacteria that are present on the surface of the plant. Surface sterilization was performed by soaking the stem in 70% ethanol for 10 minutes then rinsing the plant in sterile water before placing the stem immediately in a sterile Eppendorf tube. Samples were temporarily preserved for transport in a cooler, and held for long term storage in -20^{o}C in an industrial freezer until they were processed. I attempted to collect at least 25 of samples from each plant species. However, due to the nature of the Puget prairie system, not all plant species occurred in the research plots in equal numbers. *Erigeron speciosus* and *Symphoricarpos albus* were among the species that were the most difficult to find, and thus I was unable to collect many samples from these species.

Each sample was either denoted as a parasite (*Castilleja levisecta* plant samples), a host plant (a plant sample taken from within a 4-inch radius of a parasite that was collected), or non-host plant (a plant sample taken further than 2 feet from any *C. levisecta* plant). Non-hosts were then assigned as "neighbors" to nearby host/parasite duo's belonging to the same species as the host if the non-host and host/parasite were collected in the same plot or in a plot immediately adjacent to one another. Due to the nature of the Puget prairie system, not all plant species occurred in the research plots in equal numbers.

Samples were screened for quality of preservation and relevance for the questions asked. Because there was a budgetary limit to the number of samples that I could sequence, I choose only to sequence samples that were well preserved in sterile conditions and that allowed me to

investigate my hypothesis. Plant samples that were stored in cracked Eppendorf tubes, samples that thawed before processing, and samples that were processed under questionably sterile conditions were not selected for sequencing by the UMGC. Additionally, samples from plants that had an abundance of replicates and samples from plant species that did not have enough replicates were not selected for processing or sequencing. The samples that were not selected for processing or sequencing, but that were still preserved in sterile conditions, were prepared for long term storage at -80^{o}C for potential use in future studies. Of the 328 samples that were collected from the Glacial Heritage Preserve, 293 were selected for processing and analysis. Of the 59 samples that were collected from the Smith Prairie, 42 were selected for processing and analysis. 13 negative controls ("Blanks") were also submitted for sequencing to check for sterility during processing and sequencing of the plant samples.

Sample Processing

Samples were processed between September and December 2019. Plant samples were ground into powder by immersing the stems in liquid nitrogen and crushed using sterilized mortars and pestles. Mortars and pestles were only on one sample per batch, and were washed in hot water and wiped with paper towels soaked in 70% ethanol before being placed in autoclavable plastic bags and sterilized via autoclave after each use. In batches 1 and 2 (removed from analysis due to contamination), mortars and pestles were not autoclaved in plastic bags, and were instead autoclaved with tin foil sealing the top of the mortars and pestles wrapped in tin foil. Mortars and pestles were autoclaved in plastic bags after batches 1 and 2 were found to be contaminated, as it was thought that permeations in the tin foil could have allowed bacteria to contaminate samples from the lab environment. Samples were processed using the Qiagen DNeasy PowerSoil Pro Kit. After all samples were processed, the DNA extracts were thawed

and DNA concentration was calculated using a NanoDrop Spectrophotometer (ND1000). The NanoDrop Spectrophotometer readouts provide data on the quantity and purity of the nucleic acids present in each sample. A concentration of 1-100 ng/ul was required for sequencing; all samples were quantified, and no samples with lower than 1 ng/ul were present.

30 ul of each extract was loaded into 5 96-well plates. Samples were submitted to the University of Minnesota Genomics Center for sequencing. Based on established protocols developed by the Earth Microbiome Project, the sample extractions were sequenced using primers 515F/806R, targeting the hypervariable V4 region of the conserved 16s bacterial ribosomal RNA. The University of Minnesota Genomics Center completed indexing, library preparation and Illumina protocols for sequencing. The Miseq v3 Chemistry 2x300 sequencing platform was used to sequence pooled DNA. mPNA and pPNA blockers were used during sequencing to prevent mitochondria and chloroplast from interfering with sequencing. The UMGC workflow for sample processing is available in Appendix 1.D.

Data Processing

The raw sequence reads were processed using a genomic pipeline generated in CLC Genomics Workbench 12.0.3, a data analysis package created by Qiagen. The Microbial Genomics Module for CLC Genomics Workbench software is designed to process and analyze "16s rRNA and other commonly used metagenome derived amplicon data." (CLC Microbial Genomics Module User Manual). The Microbial Genomics Module was used to trim, filter, and cluster reads into OTUs. The process for read editing is described below.

First, I uploaded the forward and reverse paired-end Illumina files to Workbench. In the Import wizard, the import type was set to Paired Reads, the minimum distance was set to 200,

the maximum distance was set to 550, and quality scores associated with the reads were imported as well. Then, reads with quality scores less than 0.05 were trimmed. The Trim Reads tool was also used to trim ambiguous nucleotides with a maximum number of ambiguities set to 2. Reads shorter than 5 nucleotides in length were discarded.

The processed reads then were clustered into OTUs. Using the OTU Clustering tool, I chose to use the SILVA 16S v132 97% reference database, with the similarity percent specified by the OTU database option selected (Balvočiūtė and Huson 2017). 97% similarity is a standard value for microbial 16s analysis, although it should be noted that recent research has questioned the validity of this value (Stackebrandt and Ebers 2006). The creation of novel OTUs was enabled. An abundance table displaying the number of reads from each OTU discovered in each sample was generated by CLC Workbench and exported as a .csv file to R Studio for further examination. The R script for the following analysis can be found in Appendix 3.

Statistical Analysis

After processing the raw reads through the CLC Workbench genomic pipeline, I performed statistical analysis on my data. R Studio was used to perform the subsequent calculations, data transformations and statistical analysis. A file containing abundance data for each OTU present in each sample was exported to R Studio. This file was examined by multiple parties for errors and was cleaned prior to analysis. While rarefaction has been used in previous microbiome studies to normalize abundance data, this technique is no longer recommended for use as the reads, and the valuable information they contain, are lost in the process (McMurdie and Holmes 2014). Calle 2019 recommends analyzing microbiome abundance data alongside presence/absence of microbial OTUs within the same dataset, so the OTU abundance table generated in CLC Workbench was used to create a presence/absence table (Calle 2019).

Each batch of samples (samples that were extracted using the same Qiagen kit on the same day) is associated with a negative control (a “Blank”) that acts as a way to detect potentially contaminating bacterial DNA. Bacterial DNA can contaminate samples by drifting from surfaces and into samples before or during processing. These blanks are processed alongside each batch in an attempt to capture OTUs that did not originate from a plant sample. Blanks 1 and 2 captured a large amount of contamination, likely due to improper sterilization techniques used on mortars and pestles. The process to sterilize mortars and pestles was adjusted after Blanks 1 and 2 revealed contamination; instead of autoclaving mortars and pestles in tin foil, they were instead autoclaved in autoclavable plastic bags that were sealed. Blank 1 contained 723 OTUs and 31,007 total reads, while Blank 2 contained 702 OTUs and 27,264 total reads. These values are remarkably high compared to Blanks 3-13 which contained an average of 58 OTUs and 3,354 total reads. The process to sterilize mortars and pestles was adjusted after Blanks 1 and 2 revealed contamination. Blanks 3-13 indicate that contamination was reduced as the total number of OTUs and total read abundance per blank decreased dramatically. Potentially contaminating bacterial OTUs and their respective abundances were used to filter contaminants from the batches of samples. Bacterial OTU reads recorded in each blank were subtracted from their respective batches; OTU abundances from Blank 1 were subtracted from Batch 1, OTU abundances from Blank 2 were subtracted from Batch 2, and so forth. Negative values, where more reads were detected from any particular OTU were discovered in a blank than in a plant sample, were set to zero. OTUs which were not present in a blank were unaffected, and OTUs were only subtracted using their respective blanks. Because of the high prevalence of contamination in Blanks 1 and 2, all samples that were processed in batches 1 and 2 were excluded from all future analysis.

Several distance measures have been suggested for use on metagenomic data. The Bray-Curtis distance measure is commonly used with species composition data, however there are some noteworthy flaws in its application to microbiome data (Calle 2019). Microbiome abundance data is not strictly reflective of true species abundance, thus other distance measures such as the Aitchison distance measure and UniFrac distance measures are commonly recommended in scientific literature over the Bray-Curtis distance measure (Gloor et al. 2017). UniFrac measures have been used prolifically throughout the literature to calculate beta diversity. There are certain disadvantages to using UniFrac distance measures, however. Calle 2019 argues that Unifrac is inappropriate for microbiome data as these measures are not sub-compositionally dominant. Instead, Calle 2019 recommends the use of the Aitchison distance to analyze beta diversity. Given the advantages and disadvantages of these distance measures, the Bray-Curtis distance measure remains a robust statistical measure that continues to be applied in similar research endeavors and was thus chosen for use in this study (Maziarz et al. 2018).

Data characteristics were explored in R Studio using R base code. Raw read data and data after removal of potentially contaminating bacterial OTUs were examined for potential outliers and problems. Because of the high prevalence of contamination in Blanks 1 and 2, all samples that were processed in batches 1 and 2 were excluded from all future analysis. Additionally, sample 0258 contained extraordinarily low OTU abundance and was thus excluded from all future analysis. Sites GHP- Mounded, GHP- Mounded 2, and GHP- Array were excluded from analysis because their locations within Glacial Heritage Preserve were not precise, and potential effects from spatial autocorrelation could not be controlled.

To examine the similarity in bacterial OTU composition between a host plant and its respective *Castilleja levisecta* parasite, samples were grouped into trio's: a host plant, it's

parasite, and a non-host plant collected from within the plot same plot as the host/parasite pair or a neighboring plot. This analysis differs from tests used to examine the bacterial OTU composition on a species basis or site treatment basis; I am not testing to see if bacterial OTU composition between the host plant, non-host plant and parasite are significantly different or not. Instead, I am testing to see if the Bray-Curtis distance calculated between a host plant and it's parasite are larger or smaller than the Bray-Curtis distance calculated between a non-host plant and it's parasite. If these Bray-Curtis distances between host plants and their respective parasites are smaller than the distances between non-host plants and the same parasite, this would indicate that the bacterial OTU composition of the host plants are more similar to the parasite bacterial OTU composition that it would be otherwise. This would provide evidence that the parasite status of the sample could affect bacterial OTU composition, suggesting that bacterial transfer may occur between parasitic *C. levisecta* plants and their hosts.

In tests performed in Chapter 1, it was determined that bacterial OTU composition within samples is largely shaped by the plant species the sample was derived from. Thus, further testing should incorporate methods to eliminate the effect the plant species has on bacterial OTU composition. To remove the element of plant species from the analysis of bacterial OTU composition between host plants, non-host plants, and parasitic plants, analyses were performed within species groups, with a Paired Sample T Test performed on all samples and separate Paired Sample T Tests performed on each species. Additionally, to eliminate the potential interference of spatial autocorrelation, each Host.Parasite pair was assigned a Non-host plant belonging to the same species as the host and either within the same plot or in a neighboring plot. Host.Parasite pairs that did not have a Non-host plant nearby (within the same plot or in a neighboring plot) were excluded from analysis.

In cases where multiple Non-host plants were classified as a neighbor of a Host.Parasite duo, the Host.Parasite pair was duplicated and paired with each Non-host plant. Bray-Curtis distances based on differences in bacterial OTU composition were calculated between hosts and their respective parasites (Host.Parasite) and between non-hosts and the parasite belonging to their respective trio (NonHost.Parasite). The Bray-Curtis distance measure was calculated using the vegdist() function in the vegan package (Oksanen et al. 2017). For host/parasite duo's that were matched with several non-hosts, the Bray-Curtis distances for NonHost.Parasite were averaged; as the distance between the host and parasite would remain the same between these trio's, the Host.Parasite value remained unchanged. Of the total 335 samples processed, 129 were selected for use as trio's. A summary of the Host.Parasite duo's and the average distance of their trio groupings is available in Table 2.1. An extended summary of the samples and their full trio groupings is available in Appendix 2.A, Table 2.3.

Table 2.1: Summary of Host.Parasite Duo's, Sample ID's, and Bray-Curtis Distances.

Host/Parasite Sample Duo ID	NonHost Sample ID(s)	Species	Host.Parasite Bray-Curtis Distance	NonHost.Parasite Bray-Curtis Distance
0057_0056	0059	ACMI	0.946	0.933
0124_0121	0167	ACMI	0.959	0.967
0417_0416	0429, 0453	ACMI	0.982	0.956
0092_0086	0102	AQFO	0.986	0.997
0327_0326	0102	AQFO	0.894	0.998
0139_0136	0130, 0140	ASCU	0.991	0.990
0269_0265	0281, 0282, 0283	ASCU	0.926	0.942
0094_0093	0101, 0244	BADE	0.952	0.978
0236_0235	0101, 0244	BADE	0.947	0.967
0243_0242	0101, 0244	BADE	0.880	0.961
0334_0332	0182	BADE	0.938	0.972
0077_0076	0080,	CAQU	0.996	0.995
0088_0086	0103, 0104	CAQU	0.982	0.968
0108_0105	0103, 0104	CAQU	0.982	0.971
0227_0225	0234	CAQU	0.968	0.988

0266_0265	0274	CEAR	0.501	0.524
0293_0290	0285	CEAR	0.469	0.478
0312_0309	0285	CEAR	0.584	0.502
0207_0201	0232	DEME	0.930	0.937
0222_0220	0232	DEME	0.960	0.973
0106_0076	0099	ERLA	0.873	0.956
0123_0121	0162, 0163, 0164, 0165, 0166	ERLA	0.954	0.957
0194_0191	0162, 0163, 0164, 0165, 0166	ERLA	0.810	0.811
0238_0235	0099	ERLA	0.844	0.964
0311_0309	0493	ERLA	0.799	0.849
0331_0326	0099	ERLA	0.874	0.967
0405_0403	0408	FERO	0.991	0.993
0421_0420	0424, 0430, 0431, 0432, 0454	FERO	0.994	0.995
0423_0422	0424, 0430, 0431, 0432, 0454	FERO	0.991	0.996
0426_0425	0424, 0430, 0431, 0432, 0454	FERO	0.995	0.995
0489_0488	0490	FERO	0.989	0.995
0091_0086	0100	LOUT	0.969	0.987
0097_0093	0100	LOUT	0.952	0.986
0109_0105	0100	LOUT	0.941	0.984
0137_0136	0160, 0161	LOUT	0.951	0.981
0089_0086	0135, 0302, 0303	LULE	0.984	0.898
0096_0093	0135, 0302, 0303	LULE	0.968	0.985
0125_0121	0120	POGR	0.944	0.988
0213_0210	0233	POGR	0.888	0.871
0202_0201	0215, 0216, 0217, 0218, 0219	SYAL	0.703	0.933
0212_0210	0215, 0216, 0217, 0218, 0219	SYAL	0.794	0.967

Data characteristics were explored in R Studio using R version 3.6.2 (R Core Team 2019). Differences in the Bray-Curtis distance between Host.Parasite and NonHost.Parasite trio's were tested using Paired Sample T Tests. Paired Sample T Tests were performed using the

t.test() function, which returns the T test statistic, degrees of freedom, p value, 95% confidence interval, and the mean of the differences. Alpha was set to a = 0.05. After the initial Paired Sample T Test performed on all Host.Parasite and NonHost.Parasite Bray-Curtis distances (including all species) was found to be significant, Paired Sample T Tests were performed on individual species groups. Species specific tests were only performed on plant species that had more than three trio's (DF > 2); only *Balsamorhiza deltoidea, Camassia quamash, Eriophyllum lanatum, Festuca roemeri* and *Lomatium utriculatum* met this constraint. *Achillea millefolium, Aster curtisii, Balsamorhiza deltoidea, Cerastium arvense, Erigeron speciosus, Eriophyllum lanatum, Festuca roemeri, Lomatium utriculatum, Lupinus lepidus, Potentilla gracilius* and *Symphoricarpos albus* were excluded from analysis too few trio's could be assembled. All 5 species that passed the DF minimum test were combined in an additional test. Potentially contaminating bacterial OTUs were removed from the abundance dataset. Alpha was set to a = 0.05. Differences in bacterial OTU composition between samples were then visualized using several variations of scatterplots.

Results

The overall test, 5 Combo test, and individual *Eriophyllum lanatum* and *Lomatium utriculatum* tests support my hypothesis that the difference in bacterial community composition will be larger between *Castilleja levisecta* and a nearby non-host than between *C. levisecta* and its respective host plant. A summary table of the Paired Sample T Test results is illustrated in Table 2.2.

Table 2.2: Paired Sample T Test- Difference in Bray-Curtis distance based on bacterial OTU composition between Host.Parasite and NonHost.Parasite groups.

Species	DF	Mean of the Differences	T Test Statistic	PR (>T)
All	42	-0.0232	-2.57	**0.014**
5 Combo	22	-0.0279	-3.75	**0.001**
BADE	3	-0.0404	-2.93	0.061
CAQU	3	0.0016	0.21	0.849
ERLA	5	-0.0581	-2.90	**0.033**
FERO	4	-0.0027	-2.28	0.085
LOUT	3	-0.0315	-6.12	**0.009**

Difference in the Bray-Curtis distances between Host.Parasite and NonHost.Parasite groups is the difference of interest, as this comparison could indicate that the composition of bacterial OTUs within a species is, in part, shaped by the connection or lack of connection to a parasite. Based on the results of the Paired Sample T Tests, there are significant differences in the Bray-Curtis distances between Host.Parasite and NonHost.Parasite groups when testing

differences within all species (Figure 2.3). Additionally, there are significant differences in the Bray-Curtis distances between Host.Parasite and NonHost.Parasite groups with *Lomatium utriculatum* as a host plant (Figure 2.2). For *Lomatium utriculatum*, Host.Parasite Bray-Curtis distances are significantly smaller than NonHost.Parasite Bray-Curtis distances. This suggests that, for *L. utriculatum*, the bacterial OTU composition of *Castilleja levisecta* and *L. utriculatum* samples that were parasitized by a *C. levisecta* plant were more similar to each other than *C. levisecta* and *L. utriculatum* samples that were not parasitized by a *C. levisecta* plant. The result is similar for *Eriophyllum lanatum*; the bacterial OTU composition of *C. levisecta* and *E. lanatum* samples that were parasitized by *C. levisecta* were more similar to each other than C. *levisecta* and *E. lanatum* samples that were not parasitized. Although *Balsamorhiza deltoidea* and *Festuca roemeri* did not have significant differences between the Host.Parasite Bray-Curtis distances and NonHost.Parasite Bray-Curtis distances, their p values were still quite small ($p = 0.061$ and $p = 0.085$, respectively). Finally, *Camassia quamash* had a large P value, indicating that the differences between the Host.Parasite Bray-Curtis distances and NonHost.Parasite Bray-Curtis distances were not significant by a large margin. With all five of these species trios combined in the 5 Combo test, I found a significant effect of parasitism (Figure 2.1).

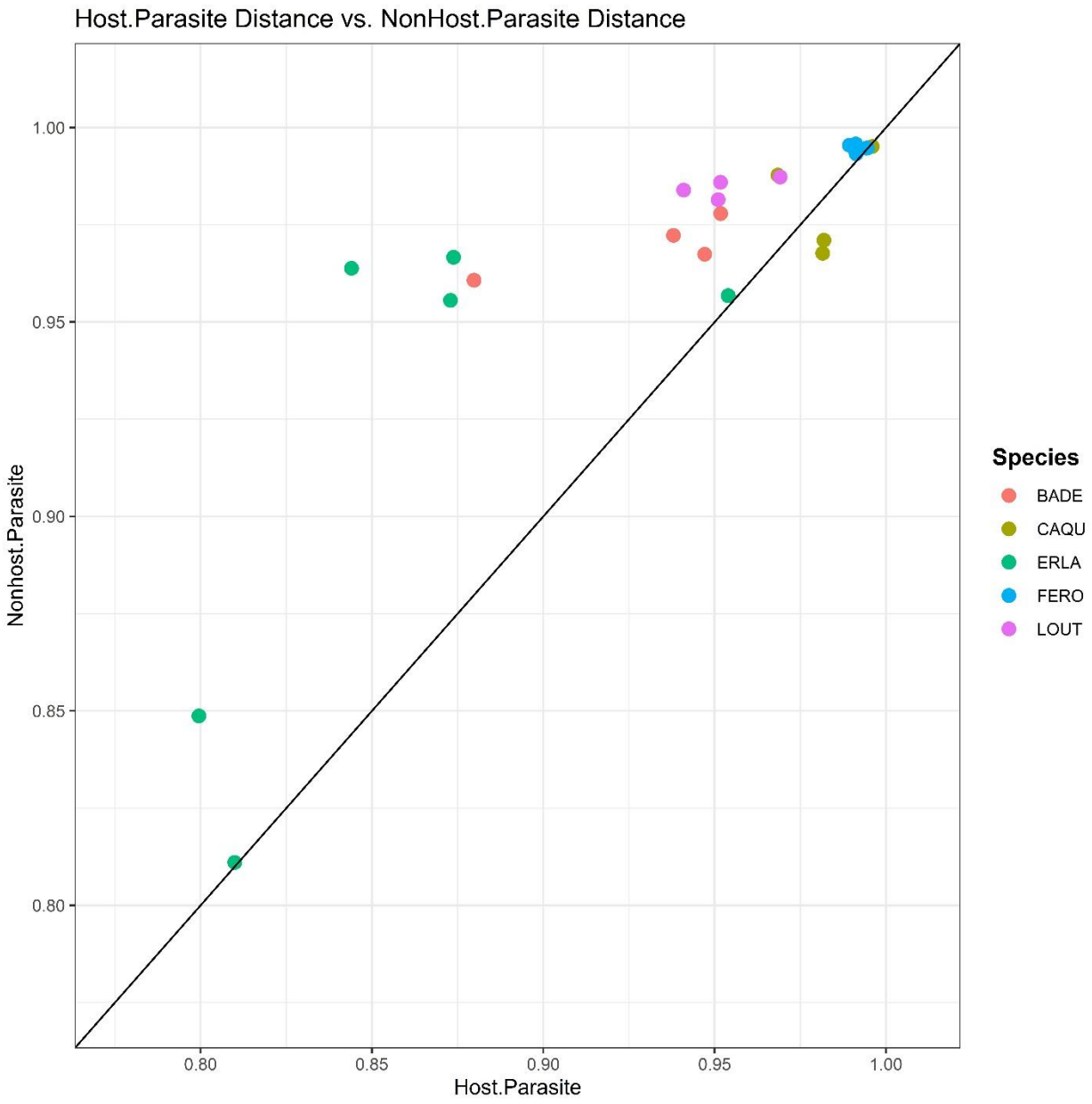


Figure 2.1: Scatterplot of Bray-Curtis distances. X axis: Bray-Curtis distances calculated between host plants and their respective parasites. Y axis: Bray-Curtis distances calculated between non-host plants and the parasite from their respective host/parasite duo. Data points are color coded by species. A 1x1 line was imposed onto the plot to assist with visualization of differences between Bray-Curtis distances; points that fall below the line are sample trios that have a larger difference in bacterial OTU composition between host plants and their respective parasites than between non-host plants and their respective parasites. Points that fall above the line are sample trios that have a larger difference in bacterial OTU composition between non-host plants and their respective parasites than host plants and their respective parasites. Figure illustrates data points used in the 5 Combo test (Table 2.2).

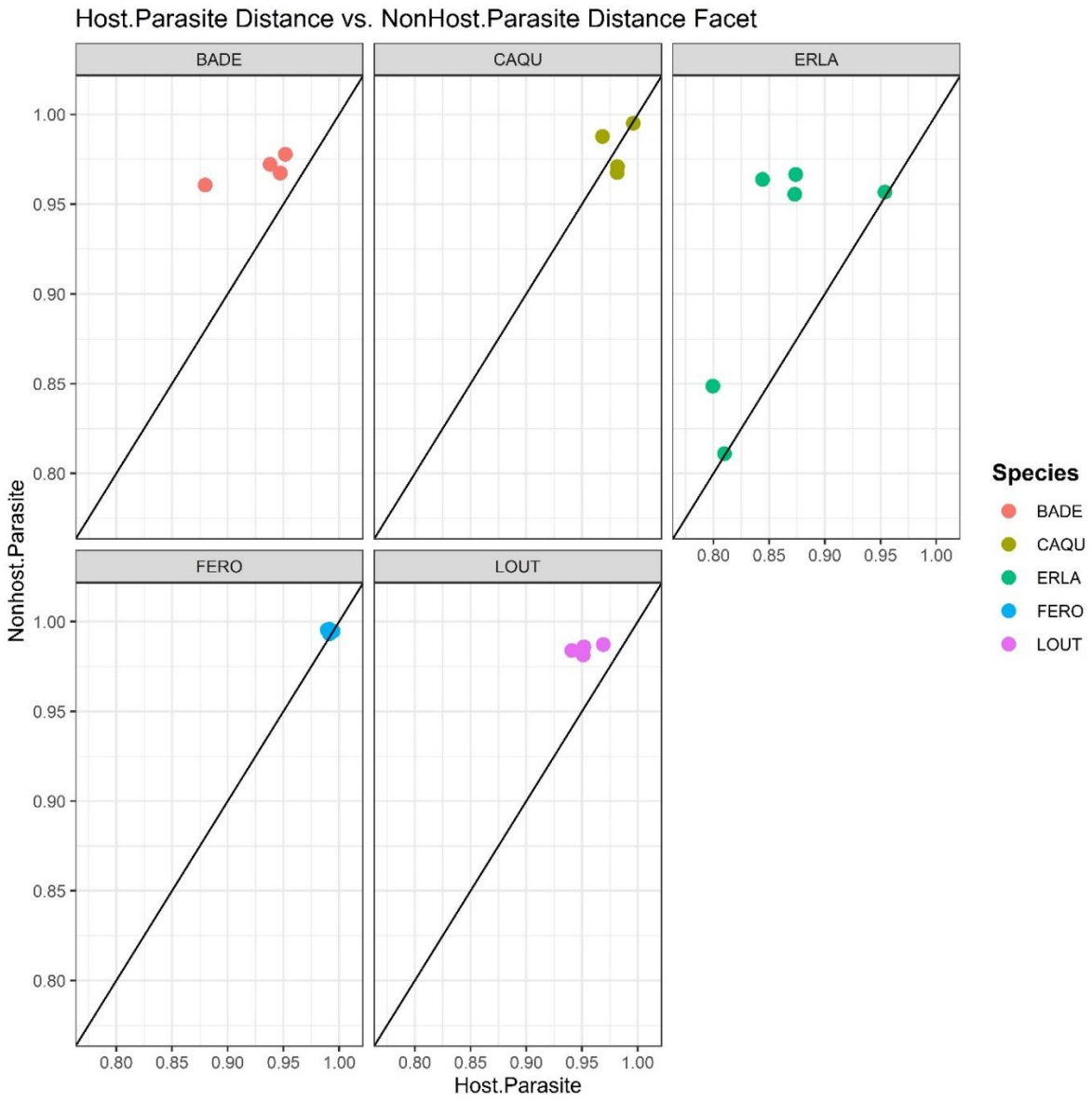


Figure 2.2: Faceted Scatterplot of Bray-Curtis distances. Plots were faceted for ease of visualization and interpretation. X axis: Bray-Curtis distances calculated between host plants and their respective parasites. Y axis: Bray-Curtis distances calculated between non-host plants and the parasite from their respective host/parasite duo. Data are faceted and color coded by species. A 1x1 line was imposed onto the plot to assist with visualization of differences between Bray-Curtis distances; points that fall below the line are sample trios that have a larger difference in bacterial OTU composition between host plants and their respective parasites than between non-host plants and their respective parasites. Points that fall above the line are sample trios that have a larger difference in bacterial OTU composition between non-host plants and their respective parasites than host plants and their respective parasites. All *Eriophyllum lanatum* samples fell close to or above the line, the differences in the bacterial OTU composition between host plants and their respective parasites are significantly smaller than the differences between non-host plants that their respective parasites. The *Eriophyllum lanatum* samples are not tightly clustered together, indicating that there is a small amount of variance in the difference in Bray-Curtis distances. All *Lomatium utriculatum* samples fell above the line; the differences in the bacterial OTU composition between host plants and their respective parasites are significantly smaller than the differences between non-host plants that their respective parasites. The *Lomatium utriculatum* samples cluster closely together, indicating that there is not much variance in the difference in Bray-Curtis distances. Figure illustrates data points used in the individual species tests (Table 2.2).

Visualization of the scatterplots reveal that for most of the sample trio's, the differences between the Bray-Curtis distances are small (clustered around the 1x1 line) and the Bray-Curtis distances themselves are relatively large (close to 1) (Figure 2.1). There are a few notable exceptions, however. *Lomatium utriculatum* trio data points, which has significantly smaller Host.Parasite Bray-Curtis values than NonHost.Parasite Bray-Curtis values, do not fall along the 1x1 line and instead all fall above the line. *Lomatium utriculatum* sample data points also have significantly smaller Host.Parasite Bray-Curtis values than NonHost.Parasite Bray-Curtis values and fall on or above the 1x1 line. Although *Balsamorhiza deltoidea* was not found to have a statistically significant difference in Host.Parasite Bray-Curtis and NonHost.Parasite Bray Curtis distances, this difference is marginally significant, all data points lie above the 1x1 line, and may indicate that a larger sample size could lead to significantly different values.

While *Cerastium arvense* did not have enough samples to investigate for an effect of parasitism (DF < 3), I discovered that *C. arvense* samples display an unusual pattern of Host.Parasite and NonHost.Parasite Bray-Curtis distances (Figure 2.3). In Chapter 1, it was determined that *C. arvense* and *Castilleja levisecta* have significantly different bacterial OTU compositions. In an NMDS analysis of the data, however, *C. levisecta* and *C. arvense* were observed to occupy a similar space in a three-dimensional ordination. Additional evidence that *C. arvense* and *C. levisecta* samples are similar was indicated in Chapter 2; a scatterplot of Bray-Curtis distances of all samples, regardless of the sample size for each species, revealed that *C. arvense* stands out from the other samples (Figure 2.3). The Bray-Curtis distances for both the Host.Parasite and NonHost.Parasite were both much smaller than Bray-Curtis distances for *C. arvense* trio's any other species trio's. Although *C. levisecta* and *C. arvense* have significantly distinct bacterial OTU compositions, the small Bray-Curtis distances calculated between both

Host.Parasite and NonHost.Parasite groups indicates that the bacterial OTU composition of *C. arvense* is especially similar to *C. levisecta*.

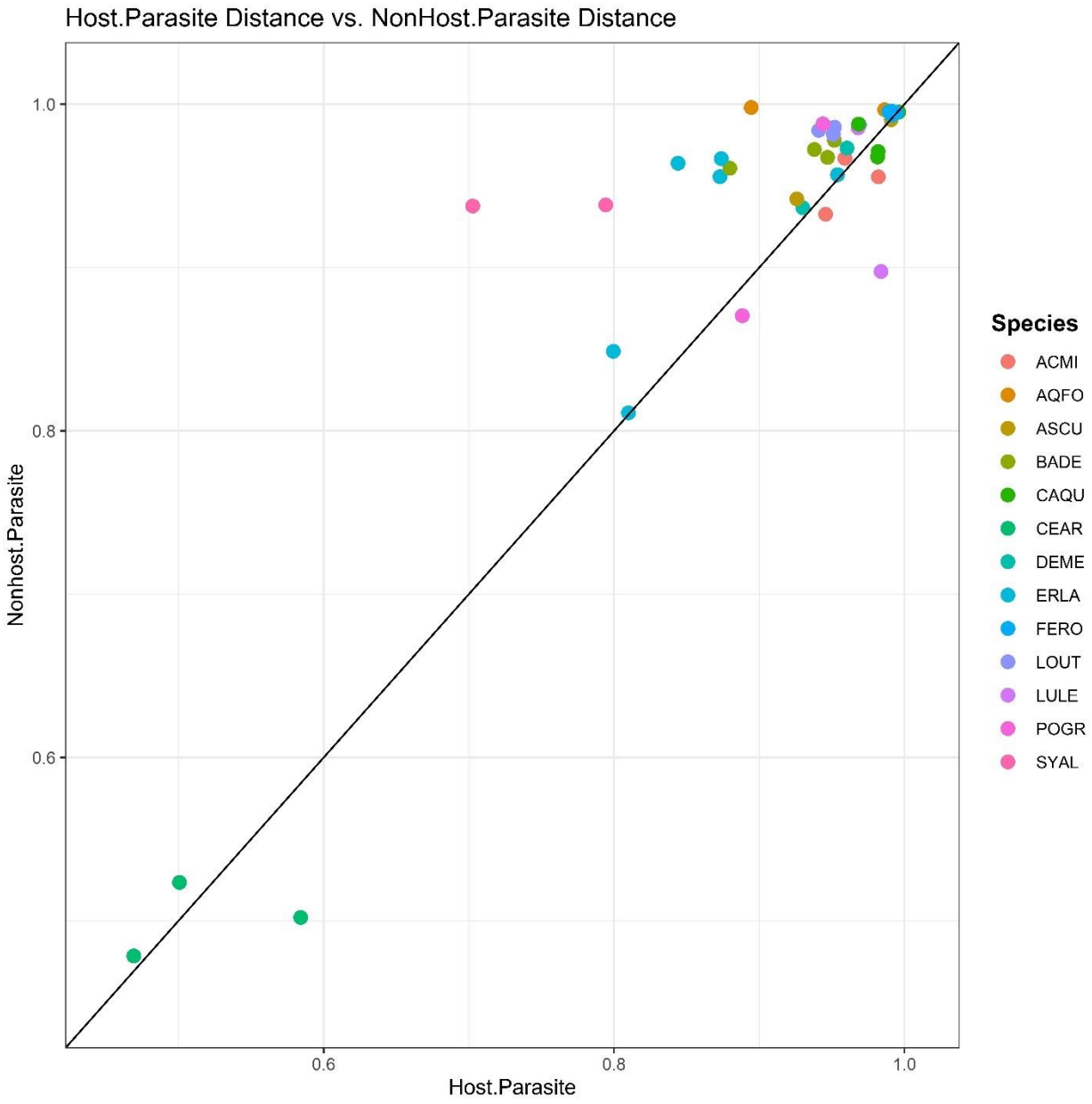


Figure 2.3: Scatterplot of Bray-Curtis distances, including untested trio data points from 13 species. X axis: Bray-Curtis distances calculated between host plants and their respective parasites. Y axis: Bray-Curtis distances calculated between non-host plants and the parasite from their respective host/parasite duo. Data points are color coded by species. Points that are closer to zero have both hosts and non-hosts with a similar bacterial OTU composition to their respective *Castilleja levisecta* parasite. Note that *C. arvense* data points fall much closer to zero than any other data point. Figure illustrates data points used in the All species trios test (Table 2.2).

Discussion

The study of the plant microbiome in relation to plant parasitism is a new and quickly developing field, and this study is perhaps the first to investigate the influence of a hemiparasitic plant on the plant microbiome. This work contributes to our understanding of the interactions between plants and the bacteria that reside within them. Vertical transmission -the transfer of bacteria from the surrounding environment to the plant interior- is theorized to occur between a plant parasite and a plant host as a result of parasitism. This theory was tested for the first time between hemiparasitic plant *Castilleja levisecta* and five of its potential host plants: *Balsamorhiza deltoidea, Camassia quamash, Eriophyllum lanatum, Festuca roemeri,* and *Lomatium triternatum.*

As the root connections between *Castilleja levisecta* and its host plants could provide a pathway for bacterial transfer, I theorized that the microbiomes of *C. levisecta* and plants that it had parasitized would be more similar than the microbiomes of *C. levisecta* and a nearby, non-parasitized plant of the same species. To test for similarity and dissimilarity, Bray-Curtis distance measures were used to calculate the differences in bacterial microbiome composition between host plants and parasites, and between non-host plants and parasites. As bacterial transfer was expected between host plants and plant parasites and not between plant parasites and non-host plants, the Bray-Curtis distance between host plants and parasitic plants were expected to be smaller than the Bray-Curtis distance between parasitic plants and non-host plants. However, after conducting Paired Sample T tests for each species, only *Lomatium utriculatum* and *Eriophyllum lanatum* were found to have statistically significant differences in the Bray-Curtis distances of host plants and parasitic plants and Bray-Curtis distances of non-host plants and parasitic plants.

Interestingly, *Cerastium arvense* samples have remarkably low Bray-Curtis distances for both host/parasite and non-host/parasite pairs. While all other species have Bray-Curtis distance values larger than 0.8, the *C. arvense* Bray-Curtis values are all smaller than 0.6. This indicates that the bacterial OTU composition of *C. arvense* samples are more similar to the *Castilleja levisecta* microbiome than the other species. In Chapter 1, it was determined that *C. arvense* and *C. levisecta* have significantly different microbiomes (Table 1.22). However, in an NMDS ordination of the data with species overlay, *C. arvense* and C. *levisecta* samples are clustered closely together (Figure 1.6). Both *C. levisecta* and *C. arvense* have a larger portion of their microbiome comprised by Proteobacteria, relative to bacteria of other phyla that comprise the microbiome of the other plant species. In previous analysis, it was determined that there are often similarities in the bacterial OTU composition between plants of different species that belonging to the same plant family; despite this observation, *C. arvense* belongs to the Caryophyllaceae family, and *C. levisecta* belongs to the Orobanchaceae family. The closes taxonomic classification that *C. levisecta* and *C. arvense* share is class, where both belong to Magnoliophyta. However, most other prairie species tested in this study (with exception to *Camassia quamash* and *Festuca roemeri*) belong to Magnoliophyta yet retain significant differences in their bacterial community composition; it is unlikely that traits shared across Magnoliophyta alone led to the similarities between *C. levisecta* and *C. arvense*.

Because many samples could not be used in analysis due to the contamination of batches 1 and 2, and because the scaled-up plots were excluded from analysis as the distance between parasite/host duo's and non-hosts were large, the sample size used in this analysis was small. Each *Castilleja levisecta* plant sampled had to have parasitized at least one plant, and only plants that were closer than 4 inches were considered to be host plants. Additionally, a non-host plant of

the same species had to be sampled from a nearby plot, but further than 2 feet from the *C. levisecta* and its host plant. These conditions are quite stringent, and could not be found naturally occurring in the study area with great frequency. The significant result of *Lomatium utriculatum* and *Eriophyllum lanatum* indicate that there may be an influence of parasitism on the bacterial microbiome of host plants, the parasite, or both. *Balsamorhiza deltoidea* and *Festuca roemeri* were marginally significant, with P values that fell just short of 0.05 (0.061 and 0.085, respectively); a larger sample size could strengthen this test and determine if there is an effect of parasitism for these species. In contrast, *Camassia quamash* had a high P value (0.89), indicating that there was not an effect of parasitism for this species.

Eriophyllum lanatum and *Festuca roemeri* are well established in the scientific literature as host plants for *Castilleja levisecta* (Schmidt 2016). However, *Balsamorhiza deltoidea*, *Camassia quamash*, and *Lomatium utriculatum* have not been examined for their compatibility to serve as host plants for *C. levisecta*. The ability for these plants to serve as host plants for *C. levisecta* is likely, in part, due to the morphology of their root systems (Demey et al. 2014). *E. lanatum* and *F. roemeri* have thin roots that spread widely close to the surface of the soil. In contrast, *B. deltoidea* has thick tap roots that deeply penetrate the soil. *Lomatium utriculatum* roots are somewhat thick and carrot like. *C. quamash* is a bulb-forming plant, with blubs forming several inches deep (typically 4-6 inches) in the soil and has thin, short roots that extend from the base of the root (Stevens et al. 2000). It is possible that *C. quamash* did not indicate an effect of parasitism in the Pairwise test because the *C. quamash* plants chosen as “host plants” weren’t actually parasitized; the root system of *C. quamash* is small and lies beneath the bulb, possibly impeding *C. levisecta* haustoria formation. *C. quamash* plants even within 4 inches of *C. levisecta* may not have been parasitized. *Achillea millefolium* performed well as a host plant in a

study of *C. levisecta* host plant interactions (Schmidt 2016); future studies on bacterial transfer between *C. levisecta* and host plants would be advised to use *A. millefolium* as a model host plant. Additionally, future studies in this system could expand on our understanding of which plants frequently act as host plants for *C. levisecta* and which do not, how root structures play a role in the ability of C. levisecta to form haustorial root connections, and subsequently which plants may be facilitating the transfer of bacteria between *C. levisecta* and host plants.

Although the study of the plant microbiome in relation to plant parasitism is a new and developing field, this study is not the first to investigate the influence of plant parasitism on host and parasite microbiomes. In a study of holoparasitic *Orobanche hederae* and its host species, *Hedera* spp., Fitzpatrick and Schneider 2019 discovered that although *O. hederae* parasites retain a distinct microbiome from their host plants, host plants and parasites shared many bacterial species (Fitzpatrick and Schneider 2019). Fitzpatrick and Schneider 2019 argue that only microbiome of the parasitic plant is significantly influenced by parasitism as opposed to a more equal sharing of bacteria between host and parasite, as haustoria facilitate unidirectional flow of phloem and thus a unidirectional flow of bacteria from host plant to parasitic plant.

While the obligate parasite in Fitzpatrick and Schneider's study was found to have a distinct microbiome from its host plant, this is not always the case. In a study of holoparasitic *Phelipanche aegyptiaca* and its host plant *Solanum lycopersicum*, the microbiomes of *P. aegyptiaca* and *S. lycopersicum* were found to be indistinguishable after parasitism (Kruh et al. 2017). The bacterial microbiome of the seed and prehaustorium stage of the parasite were distinct from the host plant, indicating that the microbiomes of the parasite and host plant shifted after parasitism, indicating that bacterial transfer is likely the cause of microbiome similarity and not due to natural similarity in microbiomes between *P. aegyptiaca* and *S. lycopersicum*.

Additionally, the community composition of both the host plant and parasite shifted after parasitism, indicating that bacterial exchange occurred between both the parasite and the host plant as a bidirectional exchange of bacteria.

A crucial aspect that makes this study difficult to compare with the findings of previous parasitic plant microbiome studies is that *Castilleja levisecta* is a hemiparasite: a parasitic plant that only derives part of its essential compounds from a host plant, but that retains its own root system and photosynthetic capabilities. Fitzpatrick and Schneider 2019 and Kruh et al. 2017 performed their studies on the transfer of bacteria between holoparasites and their hosts. These obligate parasitic plants are entirely dependent on their host plants for essential compounds and do not maintain their own root systems (Těšitel 2016). As *C. levisecta* maintains a root system that contacts the soil, it is likely that more of the microbiome of this hemiparasite is derived from the surrounding environment than the microbiome of a holoparasite, as the root/soil interface is a critical source of bacterial endophytes for plants with root systems.

There are elements of the study design and the biology of the plants investigated that make bacterial transfer between plants difficult to observe. First, it is difficult and impractical to confirm parasitism between *Castilleja levisecta* and its host plants in the field. Confirming parasitism would require digging up the *C. levisecta* and the host plant and observing the roots for haustorial connections, which are small and delicate. This process would damage the *C. levisecta* plant, which should be avoided unless absolutely necessary to protect this threatened species. *C. levisecta* roots are typically only a few inches in length, thus plants that are assumed to be hosts were collected from at most 4 inches from a *C. levisecta* plant, ideally much closer in all circumstances and especially closer for *C. levisecta* plants of smaller size. Additionally, plants labeled as hosts were done so based on the proximity of above ground vegetation, but below

ground root structures may be smaller or pointed in a different direction. While the collection method attempted to reduce the potential that plants near the *C. levisecta* were not parasitized, without directly confirming parasitism, it is still possible that samples identified as host plants were not actually parasitized. Mistaking nearby plants as parasitized hosts would confound the results of the Paired Sample T Tests and lead to incorrect conclusions about the transfer of bacteria between *C. levisecta* parasites and their host plants. A follow up laboratory test, with *C. levisecta* plants and its hosts in pots that could be directly observed for parasitism, would improve the accuracy of the findings.

As it is still possible that differences in soil condition and subsequent regional differences in soil bacteria communities could influence the microbiome of plants throughout the prairie study site, a laboratory experiment with controlled soil conditions would also control for differences in bacterial OTU composition based on microclimate (Martiny 2006). Arranging a laboratory experiment to examine bacterial exchange would eliminate the need to determine trio's based on their collection site, as they were in the field. A laboratory experiment to directly observe the transfer of bacteria from host plant to plant parasite was originally designed as a part of this thesis. A bacterial endophyte was derived from *Achillea millefolium* and electroporated with a gene to generate glowing compounds. Sterile *Castilleja levisecta* and *A. millefolium* were to be grown in magenta boxes in agar, and *A. millefolium* was to be inoculated with the fluorescent endophyte. After a span of a few weeks, the *C. levisecta* and *A. millefolium* were to be harvested and investigated for the presence of the bacterial endophyte using both genetic sequencing and observation of the fluorescent compounds within the bacteria. Complications and delays due to equipment malfunctions and coronavirus safety precautions prevented me from completing this experiment.

Although observing the similarities in bacterial OTU composition between hosts and parasites provides evidence to suggest that parasitism affects the microbiome of hosts or parasites, it does little to illuminate the processes that drive these shifts in the microbiome. It is possible that parasitism only has an effect on the microbiome of the parasite, as bacteria travel through the phloem and initiates a unidirectional transfer of bacteria. However, it is possible that parasitism affects both the host plant and the parasitic plant, with a bidirectional transfer of bacteria through the haustorial root organs that connect the plants. The laboratory experiment, where parasitism can be directly observed, can also be used to observe the direction of microbe transfer. *Castilleja levisecta* and host samples can be collected before and after parasitism, and shifts in the bacterial community composition of *C. levisecta*, host, neither or both microbiomes will indicate if bacterial transfer occurs and if this transfer is unidirectional or bidirectional.

An abundance of research opportunities relevant to this endeavor are worth further exploration. A follow up study examining these 16 plant species in a laboratory study where parasitism could be directly observed would also strengthen these research findings. Alternative *Castilleja levisecta* hosts may be worth investigation for the transfer of bacteria; *Danthonia californica* and *Deschampsia caespitosa* were found to act as excellent hosts for *C. levisecta* (Schmidt 2016), though they were not studied as a part of my thesis. It will be of great importance to establish the direction of bacterial transfer between host plants and parasites, as recent research findings contradict the theory that the unidirectional transfer of phloem from the host plant to the plant parasite also lends to a unidirectional transfer of bacteria in the same direction (Fitzpatrick and Schneider 2019; Kruh 2017). Laboratory studies may be able to not only determine the direction of bacterial transfer, but also determine the rate of bacterial transfer between *C. levisecta* and its host plants. Finally, it may be of interest to examine if certain

bacterial taxa are more prone to transfer and colonization in new plant hosts than other bacterial taxa; bacterial taxa with broader host plant ranges may successfully establish more often in plants connected via parasitism than bacterial taxa with smaller host plant ranges.

Literature Cited

Bachelet, D., B. R. Johnson, S. D. Bridgham, P. V. Dunn, H. E. Anderson, and B. M. Rogers. 2011. Climate change impacts on western Pacific Northwest prairies and savannas. Northwest Science 85:411-429.

Bulgarelli, D., K. Schlaeppi, S. Spaepen, E. V. L. Themaat, P. Schluze-Lefert. 2013. Structure and Functions of the Bacterial Microbiota of Plants. Annu Rev Plant Biol. 64:807-38. doi: 10.1146/annurev-arplant-050312-120106.

Calle, M. L. 2019. Statistical Analysis of Metagenomics Data. Genomics & Informatics, 17(1). doi: 10.5808/gi.2019.17.1.e6

Clay, K. and Schardl C. 2002. Evolutionary origins and ecological consequences of endophyte symbiosis with grasses. Am Nat. 160:99–127.

Delvin, E. G. 2013. Restoring Abandoned Agricultural Lands in Puget Lowland Prairies: A New Approach. Delvin, Eric G. University of Washington, ProQuest Dissertations Publishing.

Demey, A., P. D. Frenne, L. Baeten, G. Verstraeten, M. Hermy, P. Boeckx, K. Verheyen. 2014. The effects of hemiparasitic plant removal on community structure and seedling establishment in semi-natural grasslands. Journal of Vegetation Science. Vol 26, Issd 3. https://doi.org/10.1111/jvs.12262.

Dunwiddie, P. and J. Bakker. 2011. The Future of Restoration and Management of Prairie-Oak Ecosystems in the Pacific Northwest. Northwest Science 85:83–92.

Fitzpatrick, C. R. and A. C. Schneider. 2019. The bacterial microbiota of a parasitic plant and its host. BioRxiv. https://doi.org/10.1101/775155

Frank, A. C., J. P. Guzman, J.E. Shay. 2017. Transmission of Bacterial Endophytes. Microorganisms, 10;5(4).

Gloor G. B., J. M. Macklaim, V. Pawlowsky-Glahn, J. J. Egozcue. 2017. Microbiome Datasets Are Compositional: And This Is Not Optional. Frontiers in microbiology. 8:2224. Epub. pmid:29187837; PubMed Central PMCID: PMC5695134.

Hardoim, P. R., L. S. Van Overbeek, and J. D. Elsas. (2008) Properties of bacterial endophytes and their proposed role in plant growth. *Trends Microbiol* **16**: 463– 471.

Klausmeyer K. R. and Shaw M. R. 2009. Climate change, habitat loss, protected areas, and the climate adaptation potential of species in Mediterranean ecosystems worldwide. PLoS ONE, 4(7), e6392

Kruh, L. I., T. Lahav, J. Abu-Nassar, G. Achdari, R. Salami, S. Freilich, R. Aly. 2017. Host-Parasite-Bacteria Triangle: The Microbiome of the Parasitic Weed Phelipanche aegyptiaca and Tomato-Solanum lycopersicum (Mill.) as a Host. Front. Plant Sci., https://doi.org/10.3389/fpls.2017.00269

Liu, H., L. C. Carvalhais, M. Crawford, E. Singh, P. G. Dennis, C. M. J. Pieterse and P. M. Schenk. 2017. Inner Plant Values: Diversity, Colonization and Benefits from Endophytic Bacteria, *Frontiers in Microbiology,* 10.3389

Martiny J. B. H., B. J. M. Bohannan, J. H. Brown, R. K. Colwell, J. A. Fuhrman, J. L. Green, M. C. HornerDevine, M. Kane, J. A. Krumins, C. R. Kuske, P. J. Morin, S. Naeem, L. Øvreås, A-L. Reysenbach, V. H. Smith, J. T. Staley. 2006. Microbial biogeography: putting microorganisms on the map. Nat Rev Microbiol. 4:102–112, doi:10.1038/nrmicro1341.

Maziarz, M., R. M. Pfeiffer, Y. Wan, M. H. Gail. 2018. Using standard microbiome reference groups to simplify beta-diversity analyses and facilitate independent validation. Bioinformatics, 34(19), 3249–3257. doi: 10.1093/bioinformatics/bty297

Mcmurdie P. J and S. Holmes. 2013. phyloseq : An R package for reproducible interactive analysis and graphics of microbiome census data. PLoS One. 8:1–11, doi:10.1371/journal.pone.0061217.

Oksanen J., F. G. Blanchet, M. Friendly, R. Kindt, P. Legendre, D. Mcglinn, P. R. Minchin, R. B. O. Hara, G. L Simpson, P. Solymos, M. H. H. Stevens, E. Szoecs 2017. vegan: Community ecology package.

Parke J. L. 1991. Root colonization by indigenous and introduced microorganisms. The Rhizosphere and Plant Growth: Beltsville Symposia in Agricultural Research. D. L. Keister and P. B. Cregan, Eds), Vol. 14, pp. 3342. Kluwer, The Netherlands.

R Core Team (2020). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL http://www.R-project.org/.

Santoyo, G., G. Moreno-Hagelsieb, M. D. C. Orozco-Mosqueda, B. R. Glick. 2016. Plant growth-promoting bacterial endophytes. Microbiological Research, 183, 92–99. doi: 10.1016/j.micres.2015.11.008

Schmidt, N. R. 2016. Parasitic plants and community composition: how Castilleja levisecta affects, and is affected by, its community. Dissertation, University of Washington.

Stackebrandt, E. and J. Ebers. 2006. Taxonomic parameters revisited: tarnished gold standards. Microbiol Today. 33:152–1555.

Stanley A. G., P. W. Dunwiddie, T. N. Kaye. 2011. Restoring invaded pacific Northwest prairies: Management recommendations from a region-wide experiment. Northwest Sci. 85:233–246, doi:10.3955/046.085.0212.

Stevens, M., D.C. Darris, S.M. Lambert. 2000. Plant guide for common camas (Camassia quamash). USDA-Natural Resources Conservation Service, National Plant Data Center, Greensboro, NC, and Corvallis Plant Materials Center, Corvallis, OR.

Těšitel, J. 2016. Functional biology of parasitic plants: A review. Plant Ecology and Evolution, Vol. 149, 1. Pp. 5-20(16)

U.S. Fish and Wildlife Service (UFWS). 2010. Recovery plan for the prairie species of western Oregon and southwestern Washington.

Walitang, D. I., C. G. Kim, S. Jeon, Y. Kang, T. Sa. 2018. Conservation and transmission of seed bacterial endophytes across generations following crossbreeding and repeated inbreeding of rice at different geographic locations. MicrobiologyOpen, 8(3). doi: 10.1002/mbo3.662

Yan Y., E. E. Kuramae, M. de Hollander, P. G. L. Klinkhamer, J. A. van Veen. 2016. Functional traits dominate the diversity-related selection of bacterial communities in the rhizosphere. ISME J., 11 pp. 56-66.

Zhang, Q., J. J. Acuña, N. G. Inostroza. 2019. Endophytic Bacterial Communities Associated with Roots and Leaves of Plants Growing in Chilean Extreme Environments. *Sci Rep* **9,** 4950

Appendix 2

Table 2.3: Trio groups with their respective sample numbers derived from a host plant, non-host plant, and parasitic plant. Trio's were organized by host plants sampled within four inches of a parasitic *Castilleja levisecta* plant which was also sampled, and associated with a nearby host plant of the same species that was sampled no closer than 2 feet to the parasite and host plant and sampled either within the same plot as the parasite and host or from a neighboring plot.

Trio ID	Host	Non-host	Parasite
ACMI1	0124	0167	0121
ACMI2	0057	0059	0056
ACMI3	0417	0429	0416
ACMI4	0417	0453	0416
AQFO1	0092	0102	0086
AQFO2	0327	0102	0326
ASCU1	0139	0140	0136
ASCU3	0269	0282	0265
ASCU4	0269	0283	0265
ASCU5	0269	0281	0265
ASCU6	0139	0130	0136
BADE1	0094	0244	0093
BADE2	0094	0101	0093
BADE3	0236	0101	0235
BADE4	0236	0244	0235
BADE5	0243	0101	0242
BADE6	0243	0244	0242
BADE7	0334	0182	0332
CAQU4	0077	0080	0076
CAQU5	0088	0103	0086
CAQU6	0088	0104	0086
CAQU7	0108	0103	0105
CAQU8	0108	0104	0105
CAQU9	0227	0234	0225
CEAR3	0266	0274	0265
CEAR4	0266	0285	0265
CEAR5	0266	0285	0265
CEAR6	0312	0307	0309
CEAR7	0293	0285	0290
CEAR8	0293	0286	0290
DEME1	0207	0232	0201
DEME2	0222	0232	0220
ERLA4	0106	0099	0076
ERLA5	0123	0162	0086
ERLA6	0123	0163	0086

ERLA7	0123	0164	0105
ERLA8	0123	0165	0105
ERLA9	0123	0166	0225
ERLA10	0194	0162	0191
ERLA11	0194	0163	0191
ERLA12	0194	0164	0191
ERLA13	0194	0165	0191
ERLA14	0194	0166	0191
ERLA15	0238	0099	0235
ERLA16	0311	0493	0309
ERLA17	0311	0099	0326
FERO1	0421	0424	0420
FERO2	0421	0430	0420
FERO3	0421	0431	0420
FERO4	0421	0432	0420
FERO5	0421	0454	0420
FERO6	0423	0424	0422
FERO7	0423	0430	0422
FERO8	0423	0431	0422
FERO9	0423	0432	0422
FERO10	0423	0454	0422
FERO11	0426	0424	0425
FERO12	0426	0430	0425
FERO13	0426	0431	0425
FERO14	0426	0432	0425
FERO15	0426	0454	0425
FERO16	0489	0490	0488
FERO17	0405	0408	0403
LOUT1	0091	0100	0086
LOUT2	0097	0100	0093
LOUT3	0109	0100	0105
LOUT4	0137	0161	0136
LOUT5	0137	0160	0136
LULE1	0089	0135	0086
LULE2	0089	0302	0086
LULE3	0089	0303	0086
LULE4	0096	0135	0093
LULE5	0096	0305	0093
LULE6	0096	0135	0093
POGR1	0125	0120	0121
POGR2	0213	0233	0210
SYAL1	0202	0215	0201
SYAL2	0202	0216	0201
SYAL3	0202	0217	0201

SYAL4	0202	0218	0201
SYAL5	0202	0219	0201
SYAL6	0212	0215	0210
SYAL7	0212	0216	0210
SYAL8	0212	0217	0210
SYAL9	0212	0218	0210
SYAL10	0212	0219	0210

Conclusion

The study of the plant microbiome offers many fascinating research opportunities and pathways of exploration. My research has led me to explore the microbiome of 16 Puget plant species that had yet to be surveyed for their bacterial microbiome compositions, and to investigate the ways in which bacteria may be using a Puget prairie hemiparasitic plant as a "root connection highway" to travel across the prairie landscape. In Chapter 1, I discovered that plant species generally host different bacterial OTU compositions, which are distinct from the OTU compositions of many other species. However, I also discovered that plant species which did not have different bacterial OTU compositions belonged to the same plant family, indicating that similarities in traits shared within plant families could lead plants to host similar bacterial communities. Also in Chapter 1, I investigated the differences in bacterial OTU composition within plant species derived from two different study sites. I hypothesized that the bacterial OTU composition of plants derived from Glacial Heritage Preserve would differ from the bacterial OTU composition of plants derived from Smith Prairie, and found that there were statistically significant differences between the two study sites. While it was theorized that disturbance regimes would alter the soil conditions of research plots and consequently impact the soil microbiome, leading plants taken from different treatment plots to retain different bacterial OTU compositions, I found that neither initial disturbance treatments (applied in 2009, 2010, or 2011) nor disturbance regime treatments (applied on different, continuous schedules throughout the decade) had an effect on the bacterial OTU composition of plant species except for *Cerastium arvense*. For *Cerastium arvense* alone, there were differences in bacterial OTU composition between plants collected from burn and solarize initial disturbance treatments, and differences in bacterial OTU composition between plants collected from triannual early burn and triannual late burn disturbance regime treatments.

I further explored the interconnections of the plant microbiome by investigating the relationship between plant parasitism and bacterial OTU composition between host plants, parasitic plants, and non-host plants. I found that, for two of the five species tested, the differences in bacterial OTU composition between host plants and their respective parasites were smaller than the differences in bacterial OTU composition between non-host plants and their respective parasites. This provides some evidence to suggest that transfer of bacteria between parasitic plants and host plants via haustorial root connections may occur in the field. However, these results must be taken with a grain of salt as parasitism could not be confirmed in the field, and I was unable to directly observe the transfer of bacteria in a laboratory experiment I had planned to execute as a part of this thesis.

A research of this undertaking is not without its own unique suite of challenges. Students planning on pursuing similar studies may benefit from several key lessons I have learned throughout the process of my research. First, maintaining sterility at each step in the sample collection and DNA extraction process is crucial; sterility of sample collection equipment, lab surfaces, sample processing equipment, and of personal protective equipment require different methods of sterilization. Sterilization of mortars and pestles were a particular issue in this study, where autoclaving the equipment in autoclave safe bags proved more effective than sterilizing them in aluminum foil. As much effort as is poured into maintaining a sterile environment, it is almost inevitable that contamination will render some samples unusable; it is thus wise to gather and process a surplus of samples to avoid issues of sample size during analysis. Contamination will also affect samples that had remained largely sterile and thus usable, so it is also necessary to process negative controls alongside each batch of samples. Additionally, it is important to budget an appropriate amount of time for processing and sequencing. On average, I processed

approximately 50 samples using the Qiagen PowerSoil Pro kit in an 8-hour period, with a day between sampling batches to wash and sterilize mortars and pestles. Two months passed between sample submission and raw read retrieval to sequence 376 samples via the UMGC.

As metagenomic studies become cheaper to perform and sequencing technology more accessible to researchers, studies such as this are likely to be implemented on other target plants and throughout other ecosystem types. This data sheds light on how variable *in planta* bacterial communities are between plants belonging to the same ecosystem, and how the taxonomic relationships of plants generates similarities and differences in these bacterial communities. Additionally, this study contributes to the growing body of knowledge established by previous parasitic plant research to reveal the how the intricacies of hemiparasitic plant relationships may influence the plant microbiome. With this data on the microbiomes of 16 Puget prairie plants as a launching point, scientists can continue to pursue important research endeavors to enhance our understanding of plant and bacterial ecology.

Appendix 3

```
#Victoria Fox
#Prairie Microbiome Data

setwd("~/School/Thesis Work/Data and Data Analysis/R Thesis/")

library("vegan")
library("plyr")
library("tidyverse")
library("fossil")
library("phyloseq")
library("ggplot2")
library("ggordiplots")
library("ggrepel")

###DATA IMPORT

#WOB: Blanks are excluded
#WOR: Replicates are excluded
#WOBR: Blanks and replicates are excluded
#MB: OTU's found in blanks are subtracted from their corresponding samples

#Decided to remove CAHI and CALExCAHI from dataset
#Also decided to remove non-parasitizing CALE from dataset

#Metadata import
MetadataOriginal <- read.csv("Tables/Final Data Sheet.csv")
MetadataOriginal <- (MetadataOriginal[order(MetadataOriginal$Sample),])
Metadata <- MetadataOriginal[c(-grep("CAHI", MetadataOriginal$Plant.ID), -grep("Non-Parasite",
MetadataOriginal$ParasiteStatus), -grep("0055", MetadataOriginal$Sample), -grep("0258",
MetadataOriginal$Sample), -grep("0438", MetadataOriginal$Sample), -grep("0439", MetadataOriginal$Sample)), ]

#Creating dataframes
rownames(Metadata) <- c(1:412)
MetadataB <- Metadata[grep("Blank", Metadata$Plant.ID), ]
MetadataR <- Metadata[grep("Replicate", Metadata$Original.Replicate), ]
MetadataWOBR <- Metadata[c(-grep("Blank", Metadata$Plant.ID), -grep("Replicate",
Metadata$Original.Replicate), -grep("0055", Metadata$Sample), -grep("0258", Metadata$Sample), -grep("0438",
Metadata$Sample), -grep("0439", Metadata$Sample)), ]
MetadataWOB <- rbind(MetadataWOBR, MetadataR)
MetadataWOR <- rbind(MetadataWOBR, MetadataB)

### Greengenes data import
GenomicsData <- read.csv("Tables/SILVA 97% PERMANOVA (Edited).csv")

##Plant Taxonomy Import
PlantTaxonomy <- read.csv("Tables/Plant Taxonomy.csv")

###CREATING ABUNDANCE TABLE
## Separating the taxonomy column (which lists kingdom, plylum, class etc. in one column) into a column for each.
Taxa <- GenomicsData %>%
```

```
  select(OTUNumber, Taxonomy, Combined.Abundance, Min, Max, Mean, Median, Std, Sequence) %>%
  separate(Taxonomy, into = c("Kingdom", "Phylum", "Class", "Order", "Family", "Genus", "Species"), sep = ", ")

Taxa <- Taxa %>%
  mutate(Taxa$Kingdom, Kingdom=sapply(strsplit(Taxa$Kingdom, split="__", fixed = TRUE), function(x) (x[2]))) %>%
  mutate(Taxa$Phylum, Phylum=sapply(strsplit(Taxa$Phylum, split="__", fixed = TRUE), function(x) (x[2]))) %>%
  mutate(Taxa$Class, Class=sapply(strsplit(Taxa$Class, split="__", fixed = TRUE), function(x) (x[2]))) %>%
  mutate(Taxa$Order, Order=sapply(strsplit(Taxa$Order, split="__", fixed = TRUE), function(x) (x[2]))) %>%
  mutate(Taxa$Family, Family=sapply(strsplit(Taxa$Family, split="__", fixed = TRUE), function(x) (x[2]))) %>%
  mutate(Taxa$Genus, Genus=sapply(strsplit(Taxa$Genus, split="__", fixed = TRUE), function(x) (x[2]))) %>%
  mutate(Taxa$Species, Species=sapply(strsplit(Taxa$Species, split="__", fixed = TRUE), function(x) (x[2])))

## Removing non-relevant columns to create just an abundance table.
Abundance <- read.csv("Tables/SILVA 97% PERMANOVA (Simplified).csv")
OTUNumbers <- GenomicsData$OTUNumber
rownames(Abundance) <- Abundance$SampleNumber
Abundance <- Abundance[c(-grep("0027", Abundance$SampleNumber), -grep("0323",
Abundance$SampleNumber), -grep("1027", Abundance$SampleNumber), -grep("1323",
Abundance$SampleNumber), -grep("0250", Abundance$SampleNumber), -grep("1250",
Abundance$SampleNumber), -grep("0055", Abundance$SampleNumber), -grep("0258",
Abundance$SampleNumber), -grep("0438", Abundance$SampleNumber), -grep("0439",
Abundance$SampleNumber)), ]
Abundance <- Abundance %>%
  select(-SampleNumber)

### EDITING ABUNDANCE AND METADATA

#Removing sample numbers from Abundance
#Temporarily merge Metadata and Abundance

AbundanceMetadata <- cbind(Metadata, Abundance)
AbundanceMetadata <- AbundanceMetadata[-c(grep("Blank 01", AbundanceMetadata$Blank.Number),
grep("Blank 02", AbundanceMetadata$Blank.Number)) , ]
rownames(AbundanceMetadata) <- AbundanceMetadata$Sample
AbundanceMetadataWOBR <- AbundanceMetadata[c(-grep("Blank", AbundanceMetadata$Species), -
grep("Replicate", AbundanceMetadata$Original.Replicate)) , ]
AbundanceMetadataWOB <- AbundanceMetadata[c(-grep("Blank", AbundanceMetadata$Species)) , ]

SampleNumberWOBR <- AbundanceMetadataWOBR$Sample

#Removing Blanks and Replicates from Abundance
AbundanceWOBR <- AbundanceMetadataWOBR[ , colnames(AbundanceMetadataWOBR) %in% OTUNumbers]

#Removing Batch 1 and Batch 2 from Metadata
Metadata <- Metadata[-c(grep("Blank 01", Metadata$Blank.Number), grep("Blank 02", Metadata$Blank.Number)),
]
MetadataB <- MetadataB[-c(grep("Blank 01", MetadataB$Blank.Number), grep("Blank 02",
MetadataB$Blank.Number)), ]
MetadataR <- MetadataR[-c(grep("Blank 01", MetadataR$Blank.Number), grep("Blank 02",
MetadataR$Blank.Number)), ]
MetadataWOBR <- MetadataWOBR[-c(grep("Blank 01", MetadataWOBR$Blank.Number), grep("Blank 02",
MetadataWOBR$Blank.Number)), ]
```

```
MetadataWOB <- MetadataWOB[-c(grep("Blank 01", MetadataWOB$Blank.Number), grep("Blank 02",
MetadataWOB$Blank.Number)), ]
MetadataWOR <- MetadataWOR[-c(grep("Blank 01", MetadataWOR$Blank.Number), grep("Blank 02",
MetadataWOR$Blank.Number)), ]

### REMOVING POTENTIALLY CONTAMINATING BACTERIAL OTU's
Blank03 <- AbundanceMetadata[AbundanceMetadata$Sample == "B03",]
Blank04 <- AbundanceMetadata[AbundanceMetadata$Sample == "B04",]
Blank05 <- AbundanceMetadata[AbundanceMetadata$Sample == "B05",]
Blank06 <- AbundanceMetadata[AbundanceMetadata$Sample == "B06",]
Blank07 <- AbundanceMetadata[AbundanceMetadata$Sample == "B07",]
Blank08 <- AbundanceMetadata[AbundanceMetadata$Sample == "B08",]
Blank09 <- AbundanceMetadata[AbundanceMetadata$Sample == "B09",]
Blank10 <- AbundanceMetadata[AbundanceMetadata$Sample == "B10",]
Blank11 <- AbundanceMetadata[AbundanceMetadata$Sample == "B11",]
Blank12 <- AbundanceMetadata[AbundanceMetadata$Sample == "B12",]
Blank13 <- AbundanceMetadata[AbundanceMetadata$Sample == "B13",]

FirstOTU <- which(colnames(Blank03) == "OTU0001")
LastOTU <- which(colnames(Blank03) == "OTU7365")

Blank03 <- Blank03[,FirstOTU:LastOTU]
Blank04 <- Blank04[,FirstOTU:LastOTU]
Blank05 <- Blank05[,FirstOTU:LastOTU]
Blank06 <- Blank06[,FirstOTU:LastOTU]
Blank07 <- Blank07[,FirstOTU:LastOTU]
Blank08 <- Blank08[,FirstOTU:LastOTU]
Blank09 <- Blank09[,FirstOTU:LastOTU]
Blank10 <- Blank10[,FirstOTU:LastOTU]
Blank11 <- Blank11[,FirstOTU:LastOTU]
Blank12 <- Blank12[,FirstOTU:LastOTU]
Blank13 <- Blank13[,FirstOTU:LastOTU]

AbundanceBlank03 <- as.data.frame(AbundanceMetadataWOB[AbundanceMetadataWOB$Blank.Number ==
"Blank 03",])
AbundanceBlank04 <- as.data.frame(AbundanceMetadataWOB[AbundanceMetadataWOB$Blank.Number ==
"Blank 04",])
AbundanceBlank05 <- as.data.frame(AbundanceMetadataWOB[AbundanceMetadataWOB$Blank.Number ==
"Blank 05",])
AbundanceBlank06 <- as.data.frame(AbundanceMetadataWOB[AbundanceMetadataWOB$Blank.Number ==
"Blank 06",])
AbundanceBlank07 <- as.data.frame(AbundanceMetadataWOB[AbundanceMetadataWOB$Blank.Number ==
"Blank 07",])
AbundanceBlank08 <- as.data.frame(AbundanceMetadataWOB[AbundanceMetadataWOB$Blank.Number ==
"Blank 08",])
AbundanceBlank09 <- as.data.frame(AbundanceMetadataWOB[AbundanceMetadataWOB$Blank.Number ==
"Blank 09",])
AbundanceBlank10 <- as.data.frame(AbundanceMetadataWOB[AbundanceMetadataWOB$Blank.Number ==
"Blank 10",])
AbundanceBlank11 <- as.data.frame(AbundanceMetadataWOB[AbundanceMetadataWOB$Blank.Number ==
"Blank 11",])
AbundanceBlank12 <- as.data.frame(AbundanceMetadataWOB[AbundanceMetadataWOB$Blank.Number ==
"Blank 12",])
```

```
AbundanceBlank13 <- as.data.frame(AbundanceMetadataWOB[AbundanceMetadataWOB$Blank.Number ==
"Blank 13",])

FirstOTU <- which(colnames(AbundanceBlank03) == "OTU0001")
LastOTU <- which(colnames(AbundanceBlank03) == "OTU7365")

AbundanceBlank03 <- AbundanceBlank03[,FirstOTU:LastOTU]
AbundanceBlank04 <- AbundanceBlank04[,FirstOTU:LastOTU]
AbundanceBlank05 <- AbundanceBlank05[,FirstOTU:LastOTU]
AbundanceBlank06 <- AbundanceBlank06[,FirstOTU:LastOTU]
AbundanceBlank07 <- AbundanceBlank07[,FirstOTU:LastOTU]
AbundanceBlank08 <- AbundanceBlank08[,FirstOTU:LastOTU]
AbundanceBlank09 <- AbundanceBlank09[,FirstOTU:LastOTU]
AbundanceBlank10 <- AbundanceBlank10[,FirstOTU:LastOTU]
AbundanceBlank11 <- AbundanceBlank11[,FirstOTU:LastOTU]
AbundanceBlank12 <- AbundanceBlank12[,FirstOTU:LastOTU]
AbundanceBlank13 <- AbundanceBlank13[,FirstOTU:LastOTU]

AbundanceBlank03MB <- as.data.frame(sweep(as.matrix(AbundanceBlank03), 2, c(as.matrix(Blank03)), "-"))
AbundanceBlank04MB <- as.data.frame(sweep(as.matrix(AbundanceBlank04), 2, c(as.matrix(Blank04)), "-"))
AbundanceBlank05MB <- as.data.frame(sweep(as.matrix(AbundanceBlank05), 2, c(as.matrix(Blank05)), "-"))
AbundanceBlank06MB <- as.data.frame(sweep(as.matrix(AbundanceBlank06), 2, c(as.matrix(Blank06)), "-"))
AbundanceBlank07MB <- as.data.frame(sweep(as.matrix(AbundanceBlank07), 2, c(as.matrix(Blank07)), "-"))
AbundanceBlank08MB <- as.data.frame(sweep(as.matrix(AbundanceBlank08), 2, c(as.matrix(Blank08)), "-"))
AbundanceBlank09MB <- as.data.frame(sweep(as.matrix(AbundanceBlank09), 2, c(as.matrix(Blank09)), "-"))
AbundanceBlank10MB <- as.data.frame(sweep(as.matrix(AbundanceBlank10), 2, c(as.matrix(Blank10)), "-"))
AbundanceBlank11MB <- as.data.frame(sweep(as.matrix(AbundanceBlank11), 2, c(as.matrix(Blank11)), "-"))
AbundanceBlank12MB <- as.data.frame(sweep(as.matrix(AbundanceBlank12), 2, c(as.matrix(Blank12)), "-"))
AbundanceBlank13MB <- as.data.frame(sweep(as.matrix(AbundanceBlank13), 2, c(as.matrix(Blank13)), "-"))

OTUNumbersWOBR <- colnames(MetadataWOBR)

AbundanceMBWOB <- rbind(AbundanceBlank03MB, AbundanceBlank04MB, AbundanceBlank05MB,
AbundanceBlank06MB, AbundanceBlank07MB, AbundanceBlank08MB, AbundanceBlank09MB,
AbundanceBlank10MB, AbundanceBlank11MB, AbundanceBlank12MB, AbundanceBlank13MB)
AbundanceMBWOB <- AbundanceMBWOB[order(rownames(AbundanceMBWOB)),]
AbundanceMBWOB[AbundanceMBWOB < 0] <- 0
AbundanceMBWOB <- as.data.frame(AbundanceMBWOB)
AbundanceMetadataMBWOB <- cbind(MetadataWOB, AbundanceMBWOB)

AbundanceB <- rbind(Blank03, Blank04, Blank05, Blank06, Blank07, Blank08, Blank09, Blank10, Blank11, Blank12,
Blank13)

AbundanceMetadataMBR <- AbundanceMetadataMBWOB[grep("Replicate",
AbundanceMetadataMBWOB$Original.Replicate), ]
AbundanceMBR <- AbundanceMetadataMBR[ , colnames(AbundanceMetadataMBR) %in% OTUNumbers]

AbundanceMetadataMBWOBR <- AbundanceMetadataMBWOB[-c(grep("Replicate",
AbundanceMetadataMBWOB$Original.Replicate), grep("0055", AbundanceMetadataMBWOB$Sample),
grep("0258", AbundanceMetadataMBWOB$Sample), grep("0438", AbundanceMetadataMBWOB$Sample),
grep("0439", AbundanceMetadataMBWOB$Sample)),]
AbundanceMBWOBR <- AbundanceMetadataMBWOBR[ , colnames(AbundanceMetadataMBWOBR) %in%
OTUNumbers]
```

```
AbundanceMBWOR <- rbind(AbundanceMBWOBR, Blank03, Blank04, Blank05, Blank06, Blank07, Blank08, Blank09,
Blank10, Blank11, Blank12, Blank13)
AbundanceMetadataMBWOR <- cbind(MetadataWOR, AbundanceMBWOR)

AbundanceMB <- rbind(AbundanceMBWOBR, AbundanceMBR, AbundanceB)
AbundanceMetadataMB <- cbind(Metadata, AbundanceMB)

# Creating Presence/Absence Matrix
PAMatrix <- as.data.frame(ifelse(Abundance[,] > 0, 1, 0))
PAMatrixWOBR <- as.data.frame(ifelse(AbundanceWOBR[,] >0, 1, 0))
PAMatrixB <- as.data.frame(ifelse(AbundanceB[,] > 0, 1, 0))
PAMatrixMBR <- as.data.frame(ifelse(AbundanceMBR[,] > 0, 1, 0))
PAMatrixMBWOBR <- as.data.frame(ifelse(AbundanceMBWOBR[,] > 0, 1, 0))
PAMatrixMBWOB <- rbind(PAMatrixMBWOBR, PAMatrixMBR)
PAMatrixMBWOR <- rbind(PAMatrixMBWOBR, PAMatrixB)

PAMatrixMB <- as.data.frame(ifelse(AbundanceMB[,] > 0, 1, 0))

# Merging Data
AllDataPAMB <- cbind(Metadata, PAMatrixMB)
AllDataPAMBWOBR <- cbind(MetadataWOBR, PAMatrixMBWOBR)

AllDataABMB <- cbind(Metadata, AbundanceMB)
AllDataABMBWOR <- cbind(MetadataWOR, AbundanceMBWOR)
AllDataABMBWOBR <- cbind(MetadataWOBR, AbundanceMBWOBR)

#Count taxa levels
PhylumTable <- count(Taxa, vars=Taxa$Phylum)
colnames(PhylumTable) <- c("Phylum", "Count")
PhylumTable <- as.data.frame(PhylumTable)
write.csv(PhylumTable, "Tables/PhylumTableMB WOB1B2.csv")

###SIMPLE CALCULATIONS

#How many OTU's per Sample
OTUSample <- as.matrix(rowSums(PAMatrix[,]))
OTUSampleWOBR <- as.matrix(rowSums(PAMatrixWOBR[,]))

#How many OTU's per Sample MB
OTUSampleMB <- rowSums(PAMatrixMB[,])
OTUSampleMBWOBR <- rowSums(PAMatrixMBWOBR[,])

#Abundance per Sample
TotalReadAbundance <- as.matrix(rowSums(Abundance))
TotalReadAbundanceAbundanceWOBR <- as.matrix(rowSums(AbundanceWOBR))

#Abundance per Sample MB
TotalReadAbundanceMB <- rowSums(AbundanceMB)
TotalReadAbundanceMBWOBR <- rowSums(AbundanceMBWOBR)

#OTU ABUNDANCE AND OTU NUMBER MB
```

```
OTUAbundanceMBTable <- data.frame(Metadata$Sample, Metadata$Species, OTUSampleMB,
TotalReadAbundanceMB, Metadata$Original.Replicate, Metadata$SampleID, Metadata$Replicate.Pairs)
colnames(OTUAbundanceMBTable) <- c("SampleNumber", "Species", "NumOTU", "NumReads",
"Original.Replicate", "SampleID", "Replicate Pairs")
rownames(OTUAbundanceMBTable) <- c(1:359)
OTUAbundanceMBTable <- transform(OTUAbundanceMBTable, SpeciesCode = Species)
OTUAbundanceMBTable$SpeciesCode <- ifelse(OTUAbundanceMBTable$SpeciesCode == "Achillea millefolium",
"ACMI",
 ifelse(OTUAbundanceMBTable$SpeciesCode == "Aquilegia formosa", "AQFO",
 ifelse(OTUAbundanceMBTable$SpeciesCode == "Aster curtisii", "ASCU",
 ifelse(OTUAbundanceMBTable$SpeciesCode == "Balsamorhiza deltoidea", "BADE",
 ifelse(OTUAbundanceMBTable$SpeciesCode == "Blank", "Blank",
 ifelse(OTUAbundanceMBTable$SpeciesCode == "Blank 1", "B01",
 ifelse(OTUAbundanceMBTable$SpeciesCode == "Blank 2", "B02",
 ifelse(OTUAbundanceMBTable$SpeciesCode == "Blank 3", "B03",
 ifelse(OTUAbundanceMBTable$SpeciesCode == "Blank 4", "B04",
 ifelse(OTUAbundanceMBTable$SpeciesCode == "Blank 5", "B05",
 ifelse(OTUAbundanceMBTable$SpeciesCode == "Blank 6", "B06",
 ifelse(OTUAbundanceMBTable$SpeciesCode == "Blank 7", "B07",
 ifelse(OTUAbundanceMBTable$SpeciesCode == "Blank 8", "B08",
 ifelse(OTUAbundanceMBTable$SpeciesCode == "Blank 9", "B09",
 ifelse(OTUAbundanceMBTable$SpeciesCode == "Blank 10", "B10",
 ifelse(OTUAbundanceMBTable$SpeciesCode == "Blank 11", "B11",
 ifelse(OTUAbundanceMBTable$SpeciesCode == "Blank 12", "B12",
 ifelse(OTUAbundanceMBTable$SpeciesCode == "Blank 13", "B13",
 ifelse(OTUAbundanceMBTable$SpeciesCode == "Camassia quamash", "CAQU",
 ifelse(OTUAbundanceMBTable$SpeciesCode == "Castilleja levisecta", "CALE",
 ifelse(OTUAbundanceMBTable$SpeciesCode == "Cerastium arvense", "CEAR",
 ifelse(OTUAbundanceMBTable$SpeciesCode == "Delphinium menziesii", "DEME",
 ifelse(OTUAbundanceMBTable$SpeciesCode == "Eriophyllum lanatum", "ERLA",
 ifelse(OTUAbundanceMBTable$SpeciesCode == "Erigeron speciosis", "ERSP",
 ifelse(OTUAbundanceMBTable$SpeciesCode == "Festuca roemeri", "FERO",
 ifelse(OTUAbundanceMBTable$SpeciesCode == "Lomatium utriculatum", "LOUT",
 ifelse(OTUAbundanceMBTable$SpeciesCode == "Lomatium triternatum", "LOTR",
 ifelse(OTUAbundanceMBTable$SpeciesCode == "Symphoricarpos albus", "SYAL",
 ifelse(OTUAbundanceMBTable$SpeciesCode == "Lupinus lepidus", "LULE",
 ifelse(OTUAbundanceMBTable$SpeciesCode == "Potentilla gracilius", "POGR",
 "NA"))))))))))))))))))))))))))))))))
SpeciesCode <- OTUAbundanceMBTable$SpeciesCode
OTUAbundanceMBTable$SpeciesCode <- as.factor(OTUAbundanceMBTable$SpeciesCode)
OTUAbundanceMBTable <- as.data.frame(OTUAbundanceMBTable)

OTUAbundanceMBTableWOBR <- OTUAbundanceMBTable[c(-grep("Blank", OTUAbundanceMBTable$Species), -
grep("Replicate", OTUAbundanceMBTable$Original.Replicate)), ]
OTUAbundanceMBTableWOBR <- select(OTUAbundanceMBTableWOBR, -Original.Replicate)
OTUAbundanceMBTableWOR <- OTUAbundanceMBTable[c(-grep("Replicate",
OTUAbundanceMBTable$Original.Replicate)),]

OTUAbundanceMBTableR <- OTUAbundanceMBTable[grep("Replicate",
OTUAbundanceMBTable$Original.Replicate), ]
OTUAbundanceMBTableB <- OTUAbundanceMBTable[grep("Blank", OTUAbundanceMBTable$Species), ]
OTUAbundanceMBTableO <- OTUAbundanceMBTable[grep("Original",
OTUAbundanceMBTable$Original.Replicate),]
```

```
OTUAbundanceMBTableOandR <- rbind(OTUAbundanceMBTableO, OTUAbundanceMBTableR)
write.csv(OTUAbundanceMBTable, "Tables/OTUAbundanceMBTableWOB1B2.csv")
write.csv(OTUAbundanceMBTableWOBR, "Tables/OTUAbundanceMBTableWOBRWOB1B2.csv")

###AbundanceMB TABLES####
#Removing blanks. Have to use "factor" function to get rid of the blanks levels.
OTUAbundanceMBTableWOBR$Species <- factor(OTUAbundanceMBTableWOBR$Species)
OTUAbundanceMBTableWOBR$SpeciesCode <- factor(OTUAbundanceMBTableWOBR$SpeciesCode)

OTUAbundanceMBTableOandR$Species <- factor(OTUAbundanceMBTableOandR$Species)
OTUAbundanceMBTableOandR$SpeciesCode <- factor(OTUAbundanceMBTableOandR$SpeciesCode)


OTUperSpeciesTable <- OTUAbundanceMBTableWOBR %>%
  group_by(Species) %>%
  summarise(AverageNumOTU = round(mean(NumOTU),1)) %>%
  mutate(SpeciesCode = c("ACMI", "AQFO", "ASCU", "BADE", "CAQU", "CALE", "CEAR", "DEME", "ERSP", "ERLA",
"FERO", "LOTR", "LOUT", "LULE", "POGR", "SYAL"))

unique(OTUAbundanceMBTableWOBR$Species)

#Average AbundanceMB per Species
AveOTUAbundanceMBperSpeciesTable <- OTUAbundanceMBTableWOBR %>%
  group_by(Species) %>%
  summarise(Average = round(mean(TotalReadAbundanceMB),1)) %>%
  mutate(SpeciesCode = c("ACMI", "AQFO", "ASCU", "BADE", "CAQU", "CALE", "CEAR", "DEME", "ERSP", "ERLA",
"FERO", "LOTR", "LOUT", "LULE", "POGR", "SYAL"))

min(OTUAbundanceMBTableWOBR$NumOTU)

#Plotting the Average Number of OTU's by Species
#No Blanks or Replicates
OTUAbundanceMBPlotWOBR<- plot(x=OTUAbundanceMBTableWOBR$SpeciesCode,
    y=OTUAbundanceMBTableWOBR$NumOTU,
    xlab="Species",
    ylab="Number of OTU's",
    main = "Number of OTU's by Species")

#Scatter Plot
ggplot(OTUAbundanceMBTable, aes(x=SpeciesCode, y=NumOTU))+
  geom_point()+
  theme_bw()+
  labs(title = "Scatterplot",
       subtitle = "Number of OTU's by Species",
       y = "Number of OTU's",
       x = "Species Code",
       xlab("SpeciesCode"))+
  ggsave("Graphics/Number of OTU's by Species MB.jpg",
         width = 12, height = 10, units = "in", dpi = 1200)

#No Blanks or Replicates
ggplot(OTUAbundanceMBTableWOBR, aes(x=SpeciesCode, y=NumOTU))+
```

```
 geom_point()+
 theme_bw()+
 labs(title = "Scatterplot",
     subtitle = "Number of OTU's by Species MB WOBR",
     y = "Number of OTU's",
     x = "Species Code",
     xlab("SpeciesCode"))+
 ggsave("Graphics/Number of OTU's by Species MB WOBR WOB1B2.jpg",
       width = 10, height = 10, units = "in", dpi = 1200)

DeNovoOTU <- GenomicsData[GenomicsData$ID == "De-Novo OTU", "ID"]
length(DeNovoOTU)


###SPECIES PERMANOVA and PAIRWISE
#Bray-Curtis
SpeciesPERMANOVAMBWOBR <- adonis2(AbundanceMBWOBR ~ Species, data = AllDataABMBWOBR, method =
"bray")
SpeciesPERMANOVAMBWOBR
#Statistically significant result! p < 0.001.
#Percentage of the variation in the data is due to this factor : 78%

source("Scripts/pairwise.adonis.R")
SpeciesPairwiseMBWOBR <- pairwise.adonis(resp = vegdist(AbundanceMBWOBR), fact =
AllDataABMBWOBR$Species)
SpeciesPairwiseMBWOBR <- do.call(rbind.data.frame, SpeciesPairwiseMBWOBR)
SpeciesPairwiseMBWOBR <- SpeciesPairwiseMBWOBR[2:16,]
rownames(SpeciesPairwiseMBWOBR) <- c("AQFO", "ASCU", "BADE", "CAQU", "CALE", "CEAR", "DEME", "ERSP",
"ERLA", "FERO", "LOTR", "LOUT", "LULE", "POGR", "SYAL")
colnames(SpeciesPairwiseMBWOBR) <- c("ACMI", "AQFO", "ASCU", "BADE", "CAQU", "CALE", "CEAR", "DEME",
"ERSP", "ERLA", "FERO", "LOTR", "LOUT", "LULE", "POGR")
write.csv(SpeciesPairwiseMBWOBR, "Tables/SpeciesPairwiseWOBRWOB1B2.csv")

#Difference in OTU composition between GHP and SM (SITE LOCATION)

ACMI <- AllDataABMBWOBR[AllDataABMBWOBR$Species == "Achillea millefolium", ]
ACMIAbundance <- ACMI[ , colnames(ACMI) %in% OTUNumbers]

ACMI.GHP.SM.PERMANOVA <- adonis2(ACMIAbundance ~GHP.SM, data = ACMI, method = "bray")
ACMI.GHP.SM.PERMANOVA
CALE <- AllDataABMBWOBR[AllDataABMBWOBR$Species == "Castilleja levisecta", ]
CALEAbundance <- CALE[ , colnames(CALE) %in% OTUNumbers]
CALE.GHP.SM.PERMANOVA <- adonis2(CALEAbundance ~GHP.SM, data = CALE, method = "bray")
CALE.GHP.SM.PERMANOVA
ERLA <- AllDataABMBWOBR[AllDataABMBWOBR$Species == "Eriophyllum lanatum", ]
ERLAAbundance <- ERLA[ , colnames(ERLA) %in% OTUNumbers]

ERLA.GHP.SM.PERMANOVA <- adonis2(ERLAAbundance ~GHP.SM, data = ERLA, method = "bray")
ERLA.GHP.SM.PERMANOVA
FERO <- AllDataABMBWOBR[AllDataABMBWOBR$Species == "Festuca roemeri", ]
FEROAbundance <- FERO[ , colnames(FERO) %in% OTUNumbers]
FERO.GHP.SM.PERMANOVA <- adonis2(FEROAbundance ~GHP.SM, data = FERO, method = "bray")
FERO.GHP.SM.PERMANOVA
```

```
ALLSP <- rbind(ACMI, CALE, ERLA, FERO)
ALLSPAbundance <- rbind(ACMIAbundance, CALEAbundance, ERLAAbundance, FEROAbundance)
ALLSP.GHP.SM.PERMANOVA <- adonis2(ALLSPAbundance ~GHP.SM, data = ALLSP, method = "bray")
ALLSP.GHP.SM.PERMANOVA
ALLSP.GHP.SM.PERMANOVA2 <- adonis2(ALLSPAbundance ~GHP.SM+Species, data = ALLSP, method = "bray")
ALLSP.GHP.SM.PERMANOVA2
ALLSP.GHP.SM.PERMANOVA3 <- adonis2(ALLSPAbundance ~GHP.SM*Species, data = ALLSP, method = "bray")
ALLSP.GHP.SM.PERMANOVA3

#Site Treatment Tests
#For this set of tests, I only want to work with the data belonging to the arrays that are broken down into 35 sites.

SiteTreatmentMetadata <- read.csv("Tables/Site Treatment Metadata.csv")

AllDataABMBWOBR.TreatmentT <- merge(x = AllDataABMBWOBR, y = SiteTreatmentMetadata[ , c("Plot.Name",
"Disturbance.Treatment", "Disturbance.Regime", "Burn.Mow", "Date.Last.Treatment")],
                by.x = "Collection.Site", by.y = "Plot.Name", all.x = TRUE)
AllDataABMBWOBR.TreatmentT <- AllDataABMBWOBR.TreatmentT %>%
  select("Disturbance.Treatment", "Disturbance.Regime", "Burn.Mow", "Date.Last.Treatment", everything())

#Excluding certain plots from analysis
AllSitesData.TreatmentT <- AllDataABMBWOBR.TreatmentT[c(-grep("GHP-  2009 Array",
AllDataABMBWOBR.TreatmentT$Collection.Site), -grep("GHP- Mounded",
AllDataABMBWOBR.TreatmentT$Collection.Site), -grep("GHP- Mounded #2",
AllDataABMBWOBR.TreatmentT$Collection.Site), -grep("SM Fenced: East",
AllDataABMBWOBR.TreatmentT$Collection.Site),
                    -grep("GHP- 100X2 2011", AllDataABMBWOBR.TreatmentT$Collection.Site), -grep("GHP- 100X3
2012", AllDataABMBWOBR.TreatmentT$Collection.Site), -grep("GHP- 10X1 SF",
AllDataABMBWOBR.TreatmentT$Collection.Site), -grep("GHP- 10X2 BF",
AllDataABMBWOBR.TreatmentT$Collection.Site),
                    -grep("GHP- 10X5 HF", AllDataABMBWOBR.TreatmentT$Collection.Site), -grep("GHP- 10X6 HF",
AllDataABMBWOBR.TreatmentT$Collection.Site), -grep("SM 10X1 HM",
AllDataABMBWOBR.TreatmentT$Collection.Site)), ]

JustGHP.TreatmentT <- AllSitesData.TreatmentT[c((-grep("SM 10X1 HM",
AllSitesData.TreatmentT$Collection.Site)), (-grep("SM 2010 06", AllSitesData.TreatmentT$Collection.Site)), (-
grep("SM 2010 11",  AllSitesData.TreatmentT$Collection.Site)), (-grep("SM 2010 13",
AllSitesData.TreatmentT$Collection.Site)),
                (-grep("SM 2010 16",  AllSitesData.TreatmentT$Collection.Site)), (-grep("SM 2010 18",
AllSitesData.TreatmentT$Collection.Site)), (-grep("SM 2010 22",  AllSitesData.TreatmentT$Collection.Site)), (-
grep("SM 2010 28", AllSitesData.TreatmentT$Collection.Site)),
                (-grep("SM 2010 29",  AllSitesData.TreatmentT$Collection.Site)), (-grep("SM 2011 16",
AllSitesData.TreatmentT$Collection.Site)), (-grep("SM 2011 26",  AllSitesData.TreatmentT$Collection.Site)), (-
grep("SM 2011 28", AllSitesData.TreatmentT$Collection.Site))),]

JustSM.TreatmentT <-  AllSitesData.TreatmentT[c(grep("SM 2010 06", AllSitesData.TreatmentT$Collection.Site),
grep("SM 2010 11",  AllSitesData.TreatmentT$Collection.Site),  grep("SM 2010 13",
AllSitesData.TreatmentT$Collection.Site),
                grep("SM 2010 16",  AllSitesData.TreatmentT$Collection.Site),  grep("SM 2010 18",
AllSitesData.TreatmentT$Collection.Site),  grep("SM 2010 22",  AllSitesData.TreatmentT$Collection.Site),
grep("SM 2010 28", AllSitesData.TreatmentT$Collection.Site),
```

```
                grep("SM 2010 29",  AllSitesData.TreatmentT$Collection.Site),  grep("SM 2011 16",
AllSitesData.TreatmentT$Collection.Site),  grep("SM 2011 26",  AllSitesData.TreatmentT$Collection.Site),
grep("SM 2011 28", AllSitesData.TreatmentT$Collection.Site)),]

#BREAK IT DOWN TO A SPECIES BY SPECIES MATRIX
JustGHP.ACMI.TreatmentT <- JustGHP.TreatmentT[JustGHP.TreatmentT$Species == "Achillea millefolium", ]
JustGHP.AQFO.TreatmentT <- JustGHP.TreatmentT[JustGHP.TreatmentT$Species == "Aquilegia formosa", ]
JustGHP.ASCU.TreatmentT <- JustGHP.TreatmentT[JustGHP.TreatmentT$Species == "Aster curtisii", ]
JustGHP.BADE.TreatmentT <- JustGHP.TreatmentT[JustGHP.TreatmentT$Species == "Balsamorhiza deltoidea", ]
JustGHP.CAQU.TreatmentT <- JustGHP.TreatmentT[JustGHP.TreatmentT$Species == "Camassia quamash", ]
JustGHP.CALE.TreatmentT <- JustGHP.TreatmentT[JustGHP.TreatmentT$Species == "Castilleja levisecta", ]
JustGHP.CEAR.TreatmentT <- JustGHP.TreatmentT[JustGHP.TreatmentT$Species == "Cerastium arvense", ]
JustGHP.DEME.TreatmentT <- JustGHP.TreatmentT[JustGHP.TreatmentT$Species == "Delphinium menziesii", ]
JustGHP.ERSP.TreatmentT <- JustGHP.TreatmentT[JustGHP.TreatmentT$Species == "Erigeron speciosis", ]
JustGHP.ERLA.TreatmentT <- JustGHP.TreatmentT[JustGHP.TreatmentT$Species == "Eriophyllum lanatum", ]
JustGHP.FERO.TreatmentT <- JustGHP.TreatmentT[JustGHP.TreatmentT$Species == "Festuca roemeri", ]
JustGHP.LOTR.TreatmentT <- JustGHP.TreatmentT[JustGHP.TreatmentT$Species == "Lomatium triternatum", ]
JustGHP.LOUT.TreatmentT <- JustGHP.TreatmentT[JustGHP.TreatmentT$Species == "Lomatium utriculatum", ]
JustGHP.LULE.TreatmentT <- JustGHP.TreatmentT[JustGHP.TreatmentT$Species == "Lupinus lepidus", ]
JustGHP.POGR.TreatmentT <- JustGHP.TreatmentT[JustGHP.TreatmentT$Species == "Potentilla gracilius", ]
JustGHP.SYAL.TreatmentT <- JustGHP.TreatmentT[JustGHP.TreatmentT$Species == "Symphoricarpos albus", ]

#Removing Metadata
JustGHP.TreatmentT.Abundance <- JustGHP.TreatmentT[ , colnames(JustGHP.TreatmentT) %in% OTUNumbers]
JustGHP.ACMI.TreatmentT.Abundance <- JustGHP.ACMI.TreatmentT[ , colnames(JustGHP.ACMI.TreatmentT) %in%
OTUNumbers]
JustGHP.AQFO.TreatmentT.Abundance <- JustGHP.AQFO.TreatmentT[ , colnames(JustGHP.AQFO.TreatmentT)
%in% OTUNumbers]
JustGHP.ASCU.TreatmentT.Abundance <- JustGHP.ASCU.TreatmentT[ , colnames(JustGHP.ASCU.TreatmentT) %in%
OTUNumbers]
JustGHP.BADE.TreatmentT.Abundance <- JustGHP.BADE.TreatmentT[ , colnames(JustGHP.BADE.TreatmentT) %in%
OTUNumbers]
JustGHP.CALE.TreatmentT.Abundance <- JustGHP.CALE.TreatmentT[ , colnames(JustGHP.CALE.TreatmentT) %in%
OTUNumbers]
JustGHP.CAQU.TreatmentT.Abundance <- JustGHP.CAQU.TreatmentT[ , colnames(JustGHP.CAQU.TreatmentT)
%in% OTUNumbers]
JustGHP.CEAR.TreatmentT.Abundance <- JustGHP.CEAR.TreatmentT[ , colnames(JustGHP.CEAR.TreatmentT) %in%
OTUNumbers]
JustGHP.DEME.TreatmentT.Abundance <- JustGHP.DEME.TreatmentT[ , colnames(JustGHP.DEME.TreatmentT)
%in% OTUNumbers]
JustGHP.ERSP.TreatmentT.Abundance <- JustGHP.ERSP.TreatmentT[ , colnames(JustGHP.ERSP.TreatmentT) %in%
OTUNumbers]
JustGHP.ERLA.TreatmentT.Abundance <- JustGHP.ERLA.TreatmentT[ , colnames(JustGHP.ERLA.TreatmentT) %in%
OTUNumbers]
JustGHP.FERO.TreatmentT.Abundance <- JustGHP.FERO.TreatmentT[ , colnames(JustGHP.FERO.TreatmentT) %in%
OTUNumbers]
JustGHP.LOTR.TreatmentT.Abundance <- JustGHP.LOTR.TreatmentT[ , colnames(JustGHP.LOTR.TreatmentT) %in%
OTUNumbers]
JustGHP.LOUT.TreatmentT.Abundance <- JustGHP.LOUT.TreatmentT[ , colnames(JustGHP.LOUT.TreatmentT) %in%
OTUNumbers]
JustGHP.LULE.TreatmentT.Abundance <- JustGHP.LULE.TreatmentT[ , colnames(JustGHP.LULE.TreatmentT) %in%
OTUNumbers]
```

```
JustGHP.POGR.TreatmentT.Abundance <- JustGHP.POGR.TreatmentT[ , colnames(JustGHP.POGR.TreatmentT)
%in% OTUNumbers]
JustGHP.SYAL.TreatmentT.Abundance <- JustGHP.SYAL.TreatmentT[ , colnames(JustGHP.SYAL.TreatmentT) %in%
OTUNumbers]

#INITIAL DISTURBANCE TREATMENT PERMANOVA

DisturbanceTreatment.GHP.ACMI.PERMANOVA <- adonis2(JustGHP.ACMI.TreatmentT.Abundance
~Disturbance.Treatment, data = JustGHP.ACMI.TreatmentT, method = "bray")
DisturbanceTreatment.GHP.ACMI.PERMANOVA

DisturbanceTreatment.GHP.AQFO.PERMANOVA <- adonis2(JustGHP.AQFO.TreatmentT.Abundance
~Disturbance.Treatment, data = JustGHP.AQFO.TreatmentT, method = "bray")
DisturbanceTreatment.GHP.AQFO.PERMANOVA

DisturbanceTreatment.GHP.ASCU.PERMANOVA <- adonis2(JustGHP.ASCU.TreatmentT.Abundance
~Disturbance.Treatment, data = JustGHP.ASCU.TreatmentT, method = "bray")
DisturbanceTreatment.GHP.ASCU.PERMANOVA

DisturbanceTreatment.GHP.BADE.PERMANOVA <- adonis2(JustGHP.BADE.TreatmentT.Abundance
~Disturbance.Treatment, data = JustGHP.BADE.TreatmentT, method = "bray")
DisturbanceTreatment.GHP.BADE.PERMANOVA

DisturbanceTreatment.GHP.CALE.PERMANOVA <- adonis2(JustGHP.CALE.TreatmentT.Abundance
~Disturbance.Treatment, data = JustGHP.CALE.TreatmentT, method = "bray")
DisturbanceTreatment.GHP.CALE.PERMANOVA

DisturbanceTreatment.GHP.CAQU.PERMANOVA <- adonis2(JustGHP.CAQU.TreatmentT.Abundance
~Disturbance.Treatment, data = JustGHP.CAQU.TreatmentT, method = "bray")
DisturbanceTreatment.GHP.CAQU.PERMANOVA

DisturbanceTreatment.GHP.CEAR.PERMANOVA <- adonis2(JustGHP.CEAR.TreatmentT.Abundance
~Disturbance.Treatment, data = JustGHP.CEAR.TreatmentT, method = "bray")
DisturbanceTreatment.GHP.CEAR.PERMANOVA

DisturbanceTreatment.GHP.DEME.PERMANOVA <- adonis2(JustGHP.DEME.TreatmentT.Abundance
~Disturbance.Treatment, data = JustGHP.DEME.TreatmentT, method = "bray")
DisturbanceTreatment.GHP.DEME.PERMANOVA

DisturbanceTreatment.GHP.ERLA.PERMANOVA <- adonis2(JustGHP.ERLA.TreatmentT.Abundance
~Disturbance.Treatment, data = JustGHP.ERLA.TreatmentT, method = "bray")
DisturbanceTreatment.GHP.ERLA.PERMANOVA

DisturbanceTreatment.GHP.ERSP.PERMANOVA <- adonis2(JustGHP.ERSP.TreatmentT.Abundance
~Disturbance.Treatment, data = JustGHP.ERSP.TreatmentT, method = "bray")
DisturbanceTreatment.GHP.ERSP.PERMANOVA

DisturbanceTreatment.GHP.FERO.PERMANOVA <- adonis2(JustGHP.FERO.TreatmentT.Abundance
~Disturbance.Treatment, data = JustGHP.FERO.TreatmentT, method = "bray")
DisturbanceTreatment.GHP.FERO.PERMANOVA

DisturbanceTreatment.GHP.LOTR.PERMANOVA <- adonis2(JustGHP.LOTR.TreatmentT.Abundance
~Disturbance.Treatment, data = JustGHP.LOTR.TreatmentT, method = "bray")
```

```
DisturbanceTreatment.GHP.LOTR.PERMANOVA

DisturbanceTreatment.GHP.LOUT.PERMANOVA <- adonis2(JustGHP.LOUT.TreatmentT.Abundance
~Disturbance.Treatment, data = JustGHP.LOUT.TreatmentT, method = "bray")
DisturbanceTreatment.GHP.LOUT.PERMANOVA

DisturbanceTreatment.GHP.LULE.PERMANOVA <- adonis2(JustGHP.LULE.TreatmentT.Abundance
~Disturbance.Treatment, data = JustGHP.LULE.TreatmentT, method = "bray")
DisturbanceTreatment.GHP.LULE.PERMANOVA

DisturbanceTreatment.GHP.POGR.PERMANOVA <- adonis2(JustGHP.POGR.TreatmentT.Abundance
~Disturbance.Treatment, data = JustGHP.POGR.TreatmentT, method = "bray")
DisturbanceTreatment.GHP.POGR.PERMANOVA

DisturbanceTreatment.GHP.SYAL.PERMANOVA <- adonis2(JustGHP.SYAL.TreatmentT.Abundance
~Disturbance.Treatment, data = JustGHP.SYAL.TreatmentT, method = "bray")
DisturbanceTreatment.GHP.SYAL.PERMANOVA

DisturbanceTreatment.GHP.ALL.PERMANOVA <- adonis2(JustGHP.TreatmentT.Abundance ~Disturbance.Treatment
+ Species + Disturbance.Treatment:Species, data = JustGHP.TreatmentT, method = "bray")
DisturbanceTreatment.GHP.ALL.PERMANOVA

#Only CEAR was significant, so I will perform a pairwise test on it alone
CEAR.Initial.Disturbance.Pairwise <- pairwise.adonis(resp = vegdist(JustGHP.CEAR.TreatmentT.Abundance), fact =
JustGHP.CEAR.TreatmentT$Disturbance.Treatment)
CEAR.Initial.Disturbance.Pairwise <- do.call(rbind.data.frame, CEAR.Initial.Disturbance.Pairwise)

#CONTINUOUS DISTURBANCE REGIME TREATMENT ANOVA
source("Scripts/pairwise.adonis.R")

DisturbanceRegimeTreatment.GHP.ACMI.PERMANOVA <- adonis2(JustGHP.ACMI.TreatmentT.Abundance
~Disturbance.Regime, data = JustGHP.ACMI.TreatmentT, method = "bray")
DisturbanceRegimeTreatment.GHP.ACMI.PERMANOVA

DisturbanceRegimeTreatment.GHP.AQFO.PERMANOVA <- adonis2(JustGHP.AQFO.TreatmentT.Abundance
~Disturbance.Regime, data = JustGHP.AQFO.TreatmentT, method = "bray")
DisturbanceRegimeTreatment.GHP.AQFO.PERMANOVA

DisturbanceRegimeTreatment.GHP.ASCU.PERMANOVA <- adonis2(JustGHP.ASCU.TreatmentT.Abundance
~Disturbance.Regime, data = JustGHP.ASCU.TreatmentT, method = "bray")
DisturbanceRegimeTreatment.GHP.ASCU.PERMANOVA

DisturbanceRegimeTreatment.GHP.BADE.PERMANOVA <- adonis2(JustGHP.BADE.TreatmentT.Abundance
~Disturbance.Regime, data = JustGHP.BADE.TreatmentT, method = "bray")
DisturbanceRegimeTreatment.GHP.BADE.PERMANOVA

DisturbanceRegimeTreatment.GHP.CALE.PERMANOVA <- adonis2(JustGHP.CALE.TreatmentT.Abundance
~Disturbance.Regime, data = JustGHP.CALE.TreatmentT, method = "bray")
DisturbanceRegimeTreatment.GHP.CALE.PERMANOVA

DisturbanceRegimeTreatment.GHP.CAQU.PERMANOVA <- adonis2(JustGHP.CAQU.TreatmentT.Abundance
~Disturbance.Regime, data = JustGHP.CAQU.TreatmentT, method = "bray")
DisturbanceRegimeTreatment.GHP.CAQU.PERMANOVA
```

```
DisturbanceRegimeTreatment.GHP.CEAR.PERMANOVA <- adonis2(JustGHP.CEAR.TreatmentT.Abundance
~Disturbance.Regime, data = JustGHP.CEAR.TreatmentT, method = "bray")
DisturbanceRegimeTreatment.GHP.CEAR.PERMANOVA

DisturbanceRegimeTreatment.GHP.DEME.PERMANOVA <- adonis2(JustGHP.DEME.TreatmentT.Abundance
~Disturbance.Regime, data = JustGHP.DEME.TreatmentT, method = "bray")
DisturbanceRegimeTreatment.GHP.DEME.PERMANOVA

DisturbanceRegimeTreatment.GHP.ERLA.PERMANOVA <- adonis2(JustGHP.ERLA.TreatmentT.Abundance
~Disturbance.Regime, data = JustGHP.ERLA.TreatmentT, method = "bray")
DisturbanceRegimeTreatment.GHP.ERLA.PERMANOVA

DisturbanceRegimeTreatment.GHP.ERSP.PERMANOVA <- adonis2(JustGHP.ERSP.TreatmentT.Abundance
~Disturbance.Regime, data = JustGHP.ERSP.TreatmentT, method = "bray")
DisturbanceRegimeTreatment.GHP.ERSP.PERMANOVA

DisturbanceRegimeTreatment.GHP.FERO.PERMANOVA <- adonis2(JustGHP.FERO.TreatmentT.Abundance
~Disturbance.Regime, data = JustGHP.FERO.TreatmentT, method = "bray")
DisturbanceRegimeTreatment.GHP.FERO.PERMANOVA

DisturbanceRegimeTreatment.GHP.LOTR.PERMANOVA <- adonis2(JustGHP.LOTR.TreatmentT.Abundance
~Disturbance.Regime, data = JustGHP.LOTR.TreatmentT, method = "bray")
DisturbanceRegimeTreatment.GHP.LOTR.PERMANOVA

DisturbanceRegimeTreatment.GHP.LOUT.PERMANOVA <- adonis2(JustGHP.LOUT.TreatmentT.Abundance
~Disturbance.Regime, data = JustGHP.LOUT.TreatmentT, method = "bray")
DisturbanceRegimeTreatment.GHP.LOUT.PERMANOVA

DisturbanceRegimeTreatment.GHP.LULE.PERMANOVA <- adonis2(JustGHP.LULE.TreatmentT.Abundance
~Disturbance.Regime, data = JustGHP.LULE.TreatmentT, method = "bray")
DisturbanceRegimeTreatment.GHP.LULE.PERMANOVA

DisturbanceRegimeTreatment.GHP.POGR.PERMANOVA <- adonis2(JustGHP.POGR.TreatmentT.Abundance
~Disturbance.Regime, data = JustGHP.POGR.TreatmentT, method = "bray")
DisturbanceRegimeTreatment.GHP.POGR.PERMANOVA

DisturbanceRegimeTreatment.GHP.SYAL.PERMANOVA <- adonis2(JustGHP.SYAL.TreatmentT.Abundance
~Disturbance.Regime, data = JustGHP.SYAL.TreatmentT, method = "bray")
DisturbanceRegimeTreatment.GHP.SYAL.PERMANOVA

DisturbanceRegimeTreatment.GHP.ALL.PERMANOVA <- adonis2(JustGHP.TreatmentT.Abundance
~Disturbance.Regime + Species + Disturbance.Regime:Species, data = JustGHP.TreatmentT, method = "bray")
DisturbanceRegimeTreatment.GHP.ALL.PERMANOVA

#Only CEAR had a significant effect so I'll run pairwise
CEAR.Disturbance.Regime.Pairwise <- pairwise.adonis(resp = vegdist(JustGHP.CEAR.TreatmentT.Abundance), fact =
JustGHP.CEAR.TreatmentT$Disturbance.Regime)
CEAR.Disturbance.Regime.Pairwise <- do.call(rbind.data.frame, CEAR.Disturbance.Regime.Pairwise)
CEAR.Disturbance.Regime.Pairwise

#BURN vs MOW TREATMENT ANOVA
```

```
BurnMowTreatment.GHP.ACMI.PERMANOVA <- adonis2(JustGHP.ACMI.TreatmentT.Abundance ~Burn.Mow, data
= JustGHP.ACMI.TreatmentT, method = "bray")
BurnMowTreatment.GHP.ACMI.PERMANOVA

BurnMowTreatment.GHP.AQFO.PERMANOVA <- adonis2(JustGHP.AQFO.TreatmentT.Abundance ~Burn.Mow, data
= JustGHP.AQFO.TreatmentT, method = "bray")
BurnMowTreatment.GHP.AQFO.PERMANOVA

BurnMowTreatment.GHP.ASCU.PERMANOVA <- adonis2(JustGHP.ASCU.TreatmentT.Abundance ~Burn.Mow, data
= JustGHP.ASCU.TreatmentT, method = "bray")
BurnMowTreatment.GHP.ASCU.PERMANOVA

BurnMowTreatment.GHP.BADE.PERMANOVA <- adonis2(JustGHP.BADE.TreatmentT.Abundance ~Burn.Mow, data
= JustGHP.BADE.TreatmentT, method = "bray")
BurnMowTreatment.GHP.BADE.PERMANOVA

BurnMowTreatment.GHP.CALE.PERMANOVA <- adonis2(JustGHP.CALE.TreatmentT.Abundance ~Burn.Mow, data =
JustGHP.CALE.TreatmentT, method = "bray")
BurnMowTreatment.GHP.CALE.PERMANOVA

BurnMowTreatment.GHP.CAQU.PERMANOVA <- adonis2(JustGHP.CAQU.TreatmentT.Abundance ~Burn.Mow, data
= JustGHP.CAQU.TreatmentT, method = "bray")
BurnMowTreatment.GHP.CAQU.PERMANOVA

BurnMowTreatment.GHP.CEAR.PERMANOVA <- adonis2(JustGHP.CEAR.TreatmentT.Abundance ~Burn.Mow, data
= JustGHP.CEAR.TreatmentT, method = "bray")
BurnMowTreatment.GHP.CEAR.PERMANOVA

BurnMowTreatment.GHP.DEME.PERMANOVA <- adonis2(JustGHP.DEME.TreatmentT.Abundance ~Burn.Mow, data
= JustGHP.DEME.TreatmentT, method = "bray")
BurnMowTreatment.GHP.DEME.PERMANOVA

BurnMowTreatment.GHP.ERLA.PERMANOVA <- adonis2(JustGHP.ERLA.TreatmentT.Abundance ~Burn.Mow, data =
JustGHP.ERLA.TreatmentT, method = "bray")
BurnMowTreatment.GHP.ERLA.PERMANOVA

BurnMowTreatment.GHP.ERSP.PERMANOVA <- adonis2(JustGHP.ERSP.TreatmentT.Abundance ~Burn.Mow, data =
JustGHP.ERSP.TreatmentT, method = "bray")
BurnMowTreatment.GHP.ERSP.PERMANOVA

BurnMowTreatment.GHP.FERO.PERMANOVA <- adonis2(JustGHP.FERO.TreatmentT.Abundance ~Burn.Mow, data
= JustGHP.FERO.TreatmentT, method = "bray")
BurnMowTreatment.GHP.FERO.PERMANOVA

BurnMowTreatment.GHP.LOTR.PERMANOVA <- adonis2(JustGHP.LOTR.TreatmentT.Abundance ~Burn.Mow, data
= JustGHP.LOTR.TreatmentT, method = "bray")
BurnMowTreatment.GHP.LOTR.PERMANOVA

BurnMowTreatment.GHP.LOUT.PERMANOVA <- adonis2(JustGHP.LOUT.TreatmentT.Abundance ~Burn.Mow, data
= JustGHP.LOUT.TreatmentT, method = "bray")
BurnMowTreatment.GHP.LOUT.PERMANOVA
```

```
BurnMowTreatment.GHP.LULE.PERMANOVA <- adonis2(JustGHP.LULE.TreatmentT.Abundance ~Burn.Mow, data =
JustGHP.LULE.TreatmentT, method = "bray")
BurnMowTreatment.GHP.LULE.PERMANOVA

BurnMowTreatment.GHP.POGR.PERMANOVA <- adonis2(JustGHP.POGR.TreatmentT.Abundance ~Burn.Mow, data
= JustGHP.POGR.TreatmentT, method = "bray")
BurnMowTreatment.GHP.POGR.PERMANOVA

BurnMowTreatment.GHP.SYAL.PERMANOVA <- adonis2(JustGHP.SYAL.TreatmentT.Abundance ~Burn.Mow, data =
JustGHP.SYAL.TreatmentT, method = "bray")
BurnMowTreatment.GHP.SYAL.PERMANOVA

#Date Last Burned CEAR PERMANOVA

Last.Treatment.CEAR.PERMANOVA <- adonis2(JustGHP.CEAR.TreatmentT.Abundance ~Date.Last.Treatment, data =
JustGHP.CEAR.TreatmentT, method = "bray")
Last.Treatment.CEAR.PERMANOVA
#Date Last Burned Others
Date.Last.Treatment.GHP.ACMI.PERMANOVA <-
adonis2(JustGHP.ACMI.TreatmentT.Abundance[JustGHP.ACMI.TreatmentT$Burn.Mow == "Burn", ]
~Date.Last.Treatment, data = JustGHP.ACMI.TreatmentT[JustGHP.ACMI.TreatmentT$Burn.Mow == "Burn", ],
method = "bray")
Date.Last.Treatment.GHP.ACMI.PERMANOVA

Date.Last.Treatment.GHP.AQFO.PERMANOVA <-
adonis2(JustGHP.AQFO.TreatmentT.Abundance[JustGHP.AQFO.TreatmentT$Burn.Mow == "Burn", ]
~Date.Last.Treatment, data = JustGHP.AQFO.TreatmentT[JustGHP.AQFO.TreatmentT$Burn.Mow == "Burn", ],
method = "bray")
Date.Last.Treatment.GHP.AQFO.PERMANOVA

Date.Last.Treatment.GHP.ASCU.PERMANOVA <-
adonis2(JustGHP.ASCU.TreatmentT.Abundance[JustGHP.ASCU.TreatmentT$Burn.Mow == "Burn", ]
~Date.Last.Treatment, data = JustGHP.ASCU.TreatmentT[JustGHP.ASCU.TreatmentT$Burn.Mow == "Burn", ],
method = "bray")
Date.Last.Treatment.GHP.ASCU.PERMANOVA

Date.Last.Treatment.GHP.BADE.PERMANOVA <-
adonis2(JustGHP.BADE.TreatmentT.Abundance[JustGHP.BADE.TreatmentT$Burn.Mow == "Burn", ]
~Date.Last.Treatment, data = JustGHP.BADE.TreatmentT[JustGHP.BADE.TreatmentT$Burn.Mow == "Burn", ],
method = "bray")
Date.Last.Treatment.GHP.BADE.PERMANOVA

Date.Last.Treatment.GHP.CALE.PERMANOVA <-
adonis2(JustGHP.CALE.TreatmentT.Abundance[JustGHP.CALE.TreatmentT$Burn.Mow == "Burn", ]
~Date.Last.Treatment, data = JustGHP.CALE.TreatmentT[JustGHP.CALE.TreatmentT$Burn.Mow == "Burn", ],
method = "bray")
Date.Last.Treatment.GHP.CALE.PERMANOVA

Date.Last.Treatment.GHP.CAQU.PERMANOVA <-
adonis2(JustGHP.CAQU.TreatmentT.Abundance[JustGHP.CAQU.TreatmentT$Burn.Mow == "Burn", ]
~Date.Last.Treatment, data = JustGHP.CAQU.TreatmentT[JustGHP.CAQU.TreatmentT$Burn.Mow == "Burn", ],
method = "bray")
Date.Last.Treatment.GHP.CAQU.PERMANOVA
```

```
Date.Last.Treatment.GHP.CEAR.PERMANOVA <-
adonis2(JustGHP.CEAR.TreatmentT.Abundance[JustGHP.CEAR.TreatmentT$Burn.Mow == "Burn", ]
~Date.Last.Treatment, data = JustGHP.CEAR.TreatmentT[JustGHP.CEAR.TreatmentT$Burn.Mow == "Burn", ],
method = "bray")
Date.Last.Treatment.GHP.CEAR.PERMANOVA

Date.Last.Treatment.GHP.DEME.PERMANOVA <-
adonis2(JustGHP.DEME.TreatmentT.Abundance[JustGHP.DEME.TreatmentT$Burn.Mow == "Burn", ]
~Date.Last.Treatment, data = JustGHP.DEME.TreatmentT[JustGHP.DEME.TreatmentT$Burn.Mow == "Burn", ],
method = "bray")
Date.Last.Treatment.GHP.DEME.PERMANOVA

Date.Last.Treatment.GHP.ERLA.PERMANOVA <-
adonis2(JustGHP.ERLA.TreatmentT.Abundance[JustGHP.ERLA.TreatmentT$Burn.Mow == "Burn", ]
~Date.Last.Treatment, data = JustGHP.ERLA.TreatmentT[JustGHP.ERLA.TreatmentT$Burn.Mow == "Burn", ],
method = "bray")
Date.Last.Treatment.GHP.ERLA.PERMANOVA

Date.Last.Treatment.GHP.ERSP.PERMANOVA <-
adonis2(JustGHP.ERSP.TreatmentT.Abundance[JustGHP.ERSP.TreatmentT$Burn.Mow == "Burn", ]
~Date.Last.Treatment, data = JustGHP.ERSP.TreatmentT[JustGHP.ERSP.TreatmentT$Burn.Mow == "Burn", ],
method = "bray")
Date.Last.Treatment.GHP.ERSP.PERMANOVA

Date.Last.Treatment.GHP.FERO.PERMANOVA <-
adonis2(JustGHP.FERO.TreatmentT.Abundance[JustGHP.FERO.TreatmentT$Burn.Mow == "Burn", ]
~Date.Last.Treatment, data = JustGHP.FERO.TreatmentT[JustGHP.FERO.TreatmentT$Burn.Mow == "Burn", ],
method = "bray")
Date.Last.Treatment.GHP.FERO.PERMANOVA

Date.Last.Treatment.GHP.LOTR.PERMANOVA <-
adonis2(JustGHP.LOTR.TreatmentT.Abundance[JustGHP.LOTR.TreatmentT$Burn.Mow == "Burn", ]
~Date.Last.Treatment, data = JustGHP.LOTR.TreatmentT[JustGHP.LOTR.TreatmentT$Burn.Mow == "Burn", ],
method = "bray")
Date.Last.Treatment.GHP.LOTR.PERMANOVA

Date.Last.Treatment.GHP.LOUT.PERMANOVA <-
adonis2(JustGHP.LOUT.TreatmentT.Abundance[JustGHP.LOUT.TreatmentT$Burn.Mow == "Burn", ]
~Date.Last.Treatment, data = JustGHP.LOUT.TreatmentT[JustGHP.LOUT.TreatmentT$Burn.Mow == "Burn", ],
method = "bray")
Date.Last.Treatment.GHP.LOUT.PERMANOVA

Date.Last.Treatment.GHP.LULE.PERMANOVA <-
adonis2(JustGHP.LULE.TreatmentT.Abundance[JustGHP.LULE.TreatmentT$Burn.Mow == "Burn", ]
~Date.Last.Treatment, data = JustGHP.LULE.TreatmentT[JustGHP.LULE.TreatmentT$Burn.Mow == "Burn", ],
method = "bray")
Date.Last.Treatment.GHP.LULE.PERMANOVA

Date.Last.Treatment.GHP.POGR.PERMANOVA <-
adonis2(JustGHP.POGR.TreatmentT.Abundance[JustGHP.POGR.TreatmentT$Burn.Mow == "Burn", ]
~Date.Last.Treatment, data = JustGHP.POGR.TreatmentT[JustGHP.POGR.TreatmentT$Burn.Mow == "Burn", ],
method = "bray")
```

```
Date.Last.Treatment.GHP.POGR.PERMANOVA

Date.Last.Treatment.GHP.SYAL.PERMANOVA <-
adonis2(JustGHP.SYAL.TreatmentT.Abundance[JustGHP.SYAL.TreatmentT$Burn.Mow == "Burn", ]
~Date.Last.Treatment, data = JustGHP.SYAL.TreatmentT[JustGHP.SYAL.TreatmentT$Burn.Mow == "Burn", ],
method = "bray")
Date.Last.Treatment.GHP.SYAL.PERMANOVA

#CEAR had a significant effect so I'll run pairwise
#BURN AND MOW
CEAR.Last.Treatment.Pairwise <- pairwise.adonis(resp = vegdist(CEAR[ , colnames(JustGHP.ACMI.TreatmentT)
%in% OTUNumbers]), fact = CEAR$Date.Last.Treatment)
CEAR.Last.Treatment.Pairwise <- do.call(rbind.data.frame, CEAR.Last.Treatment.Pairwise)
CEAR.Last.Treatment.Pairwise

#PARASITISM
### PARASITISM TRIO'S

ACMI1 <- rbind(AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Sample == "0124", ],
AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Sample == "0167", ],
AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Sample == "0121", ])
ACMI2 <- rbind(AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Sample == "0057", ],
AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Sample == "0059", ],
AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Sample == "0056", ])
ACMI3 <- rbind(AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Sample == "0417", ],
AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Sample == "0429", ],
AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Sample == "0416", ])
ACMI4 <- rbind(AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Sample == "0417", ],
AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Sample == "0453", ],
AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Sample == "0416", ])

JustACMI <- rbind(ACMI1, ACMI2, ACMI3, ACMI4)

#AQFO Trio's

AQFO1 <- rbind(AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Sample == "0092", ],
AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Sample == "0102", ],
AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Sample == "0086", ])
AQFO2 <- rbind(AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Sample == "0327", ],
AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Sample == "0102", ],
AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Sample == "0326", ])

#ASCU Trio's

ASCU1 <- rbind(AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Sample == "0139", ],
AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Sample == "0140", ],
AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Sample == "0136", ])
#ASCU2 <- rbind(AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Sample == "0038", ],
AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Sample == "0280", ],
AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Sample == "0036", ])
ASCU3 <- rbind(AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Sample == "0269", ],
AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Sample == "0282", ],
AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Sample == "0265", ])
```

```
ASCU4 <- rbind(AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Sample == "0269", ],
AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Sample == "0283", ],
AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Sample == "0265", ])
ASCU5 <- rbind(AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Sample == "0269", ],
AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Sample == "0281", ],
AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Sample == "0265", ])
ASCU6 <- rbind(AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Sample == "0139", ],
AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Sample == "0130", ],
AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Sample == "0136", ])
#ASCU7 <- rbind(AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Sample == "0038", ],
AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Sample == "0284", ],
AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Sample == "0036", ])

#BADE Trio's

BADE1 <- rbind(AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Sample == "0094", ],
AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Sample == "0244", ],
AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Sample == "0093", ])
BADE2 <- rbind(AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Sample == "0094", ],
AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Sample == "0101", ],
AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Sample == "0093", ])
BADE3 <- rbind(AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Sample == "0236", ],
AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Sample == "0101", ],
AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Sample == "0235", ])
BADE4 <- rbind(AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Sample == "0236", ],
AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Sample == "0244", ],
AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Sample == "0235", ])
BADE5 <- rbind(AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Sample == "0243", ],
AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Sample == "0101", ],
AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Sample == "0242", ])
BADE6 <- rbind(AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Sample == "0243", ],
AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Sample == "0244", ],
AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Sample == "0242", ])
BADE7 <- rbind(AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Sample == "0334", ],
AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Sample == "0182", ],
AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Sample == "0332", ])

#CAQU Trio's

CAQU1 <- rbind(AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Sample == "0047", ],
AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Sample == "0048", ],
AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Sample == "0046", ])
#CAQU2 <- rbind(AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Sample == "0047", ],
AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Sample == "0049", ],
AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Sample == "0046", ])
#CAQU3 <- rbind(AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Sample == "0053", ],
AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Sample == "0055", ],
AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Sample == "0051", ])
CAQU4 <- rbind(AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Sample == "0077", ],
AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Sample == "0080", ],
AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Sample == "0076", ])
```

```
CAQU5 <- rbind(AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Sample == "0088", ],
AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Sample == "0103", ],
AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Sample == "0086", ])
CAQU6 <- rbind(AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Sample == "0088", ],
AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Sample == "0104", ],
AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Sample == "0086", ])
CAQU7 <- rbind(AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Sample == "0108", ],
AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Sample == "0103", ],
AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Sample == "0105", ])
CAQU8 <- rbind(AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Sample == "0108", ],
AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Sample == "0104", ],
AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Sample == "0105", ])
CAQU9 <- rbind(AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Sample == "0227", ],
AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Sample == "0234", ],
AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Sample == "0225", ])
#CAQU10 <- rbind(AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Sample == "0324", ],
AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Sample == "0103", ],
AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Sample == "0323", ])
CAQU11 <- rbind(AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Sample == "0324", ],
AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Sample == "0104", ],
AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Sample == "0323", ])

#CEAR Trio's'

#CEAR1 <- rbind(AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Sample == "0266", ],
AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Sample == "0044", ],
AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Sample == "0265", ])
#CEAR2 <- rbind(AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Sample == "0266", ],
AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Sample == "0273", ],
AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Sample == "0265", ])
CEAR3 <- rbind(AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Sample == "0266", ],
AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Sample == "0274", ],
AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Sample == "0265", ])
CEAR4 <- rbind(AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Sample == "0266", ],
AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Sample == "0285", ],
AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Sample == "0265", ])
CEAR5 <- rbind(AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Sample == "0266", ],
AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Sample == "0286", ],
AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Sample == "0265", ])
CEAR6 <- rbind(AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Sample == "0312", ],
AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Sample == "0307", ],
AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Sample == "0309", ])
CEAR7 <- rbind(AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Sample == "0293", ],
AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Sample == "0285", ],
AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Sample == "0290", ])
CEAR8 <- rbind(AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Sample == "0293", ],
AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Sample == "0286", ],
AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Sample == "0290", ])

#DEME Trio's
```

```
DEME1 <- rbind(AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Sample == "0207", ],
AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Sample == "0232", ],
AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Sample == "0201", ])
DEME2 <- rbind(AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Sample == "0222", ],
AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Sample == "0232", ],
AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Sample == "0220", ])

#ERLA Trio's

#ERLA1 <- rbind(AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Sample == "0058", ],
AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Sample == "0060", ],
AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Sample == "0046", ])
#ERLA2 <- rbind(AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Sample == "0090", ],
AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Sample == "0099", ],
AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Sample == "0046", ])
#ERLA3 <- rbind(AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Sample == "0098", ],
AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Sample == "0099", ],
AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Sample == "0051", ])
ERLA4 <- rbind(AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Sample == "0106", ],
AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Sample == "0099", ],
AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Sample == "0076", ])
ERLA5 <- rbind(AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Sample == "0123", ],
AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Sample == "0162", ],
AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Sample == "0086", ])
ERLA6 <- rbind(AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Sample == "0123", ],
AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Sample == "0163", ],
AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Sample == "0086", ])
ERLA7 <- rbind(AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Sample == "0123", ],
AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Sample == "0164", ],
AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Sample == "0105", ])
ERLA8 <- rbind(AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Sample == "0123", ],
AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Sample == "0165", ],
AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Sample == "0105", ])
ERLA9 <- rbind(AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Sample == "0123", ],
AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Sample == "0166", ],
AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Sample == "0225", ])
ERLA10 <- rbind(AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Sample == "0194", ],
AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Sample == "0162", ],
AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Sample == "0191", ])
ERLA11 <- rbind(AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Sample == "0194", ],
AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Sample == "0163", ],
AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Sample == "0191", ])
ERLA12 <- rbind(AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Sample == "0194", ],
AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Sample == "0164", ],
AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Sample == "0191", ])
ERLA13 <- rbind(AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Sample == "0194", ],
AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Sample == "0165", ],
AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Sample == "0191", ])
ERLA14 <- rbind(AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Sample == "0194", ],
AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Sample == "0166", ],
AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Sample == "0191", ])
```

```
ERLA15 <- rbind(AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Sample == "0238", ],
AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Sample == "0099", ],
AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Sample == "0235", ])
ERLA16 <- rbind(AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Sample == "0311", ],
AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Sample == "0493", ],
AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Sample == "0309", ])
ERLA17 <- rbind(AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Sample == "0331", ],
AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Sample == "0099", ],
AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Sample == "0326", ])

#FERO's Trio's

FERO1 <- rbind(AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Sample == "0421", ],
AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Sample == "0424", ],
AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Sample == "0420", ])
FERO2 <- rbind(AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Sample == "0421", ],
AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Sample == "0430", ],
AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Sample == "0420", ])
FERO3 <- rbind(AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Sample == "0421", ],
AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Sample == "0431", ],
AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Sample == "0420", ])
FERO4 <- rbind(AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Sample == "0421", ],
AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Sample == "0432", ],
AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Sample == "0420", ])
FERO5 <- rbind(AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Sample == "0421", ],
AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Sample == "0454", ],
AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Sample == "0420", ])
FERO6 <- rbind(AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Sample == "0423", ],
AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Sample == "0424", ],
AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Sample == "0422", ])
FERO7 <- rbind(AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Sample == "0423", ],
AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Sample == "0430", ],
AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Sample == "0422", ])
FERO8 <- rbind(AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Sample == "0423", ],
AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Sample == "0431", ],
AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Sample == "0422", ])
FERO9 <- rbind(AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Sample == "0423", ],
AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Sample == "0432", ],
AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Sample == "0422", ])
FERO10 <- rbind(AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Sample == "0423", ],
AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Sample == "0454", ],
AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Sample == "0422", ])
FERO11 <- rbind(AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Sample == "0426", ],
AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Sample == "0424", ],
AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Sample == "0425", ])
FERO12 <- rbind(AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Sample == "0426", ],
AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Sample == "0430", ],
AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Sample == "0425", ])
FERO13 <- rbind(AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Sample == "0426", ],
AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Sample == "0431", ],
AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Sample == "0425", ])
```

```
FERO14 <- rbind(AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Sample == "0426", ],
AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Sample == "0432", ],
AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Sample == "0425", ])
FERO15 <- rbind(AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Sample == "0426", ],
AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Sample == "0454", ],
AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Sample == "0425", ])
FERO16 <- rbind(AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Sample == "0489", ],
AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Sample == "0490", ],
AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Sample == "0488", ])
FERO17 <- rbind(AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Sample == "0405", ],
AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Sample == "0408", ],
AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Sample == "0403", ])

#LOUT Trio's

LOUT1 <- rbind(AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Sample == "0091", ],
AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Sample == "0100", ],
AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Sample == "0086", ])
LOUT2 <- rbind(AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Sample == "0097", ],
AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Sample == "0100", ],
AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Sample == "0093", ])
LOUT3 <- rbind(AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Sample == "0109", ],
AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Sample == "0100", ],
AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Sample == "0105", ])
LOUT4 <- rbind(AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Sample == "0137", ],
AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Sample == "0161", ],
AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Sample == "0136", ])
LOUT5 <- rbind(AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Sample == "0137", ],
AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Sample == "0160", ],
AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Sample == "0136", ])

#LULE Trio's'

LULE1 <- rbind(AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Sample == "0089", ],
AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Sample == "0135", ],
AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Sample == "0086", ])
LULE2 <- rbind(AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Sample == "0089", ],
AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Sample == "0302", ],
AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Sample == "0086", ])
LULE3 <- rbind(AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Sample == "0089", ],
AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Sample == "0303", ],
AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Sample == "0086", ])
LULE4 <- rbind(AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Sample == "0096", ],
AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Sample == "0135", ],
AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Sample == "0093", ])
LULE5 <- rbind(AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Sample == "0096", ],
AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Sample == "0305", ],
AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Sample == "0093", ])
LULE6 <- rbind(AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Sample == "0096", ],
AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Sample == "0135", ],
AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Sample == "0093", ])

#POGR Trio's'
```

```
POGR1 <- rbind(AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Sample == "0125", ],
AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Sample == "0120", ],
AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Sample == "0121", ])
POGR2 <- rbind(AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Sample == "0213", ],
AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Sample == "0233", ],
AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Sample == "0210", ])

#SYAL Trio's

SYAL1 <- rbind(AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Sample == "0202", ],
AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Sample == "0215", ],
AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Sample == "0201", ])
SYAL2 <- rbind(AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Sample == "0202", ],
AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Sample == "0216", ],
AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Sample == "0201", ])
SYAL3 <- rbind(AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Sample == "0202", ],
AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Sample == "0217", ],
AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Sample == "0201", ])
SYAL4 <- rbind(AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Sample == "0202", ],
AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Sample == "0218", ],
AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Sample == "0201", ])
SYAL5 <- rbind(AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Sample == "0202", ],
AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Sample == "0219", ],
AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Sample == "0201", ])
SYAL6 <- rbind(AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Sample == "0212", ],
AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Sample == "0215", ],
AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Sample == "0210", ])
SYAL7 <- rbind(AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Sample == "0212", ],
AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Sample == "0216", ],
AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Sample == "0210", ])
SYAL8 <- rbind(AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Sample == "0212", ],
AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Sample == "0217", ],
AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Sample == "0210", ])
SYAL9 <- rbind(AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Sample == "0212", ],
AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Sample == "0218", ],
AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Sample == "0210", ])
SYAL10 <- rbind(AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Sample == "0212", ],
AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Sample == "0219", ],
AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Sample == "0210", ])


#PARASITISM Paired T TEST
TrioID <- readxl::read_xlsx("Tables/Trio ID.xlsx")
Pair.Distances <- c()
for(i in 1:nrow(TrioID)) {
  trio.data <- AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Sample %in% TrioID[i, c("Host",
"NonHost", "Parasite")] , ]
  trio.data$Row <- rownames(trio.data)
  rownames(trio.data) <- trio.data$ParasiteStatus
  trio.dist <- as.matrix(vegdist(trio.data[ , colnames(trio.data) %in% OTUNumbers], method = "bray"))
  temp1 <- data.frame(From = "Host", To = "NonHost", Dist = trio.dist["Host", "Non-host"])
  temp2 <- data.frame(From = "Host", To = "Parasite", Dist = trio.dist["Host", "Parasite"])
```

```
 temp3 <- data.frame(From = "NonHost", To = "Parasite", Dist = trio.dist["Non-host", "Parasite"])
 temp <- rbind(temp1,temp2, temp3)
 temp$Trio.ID <- TrioID$`Trio ID`[i]
 Pair.Distances <- rbind(Pair.Distances, temp)
}
Pair.Distances$Pair <- paste(Pair.Distances$From, Pair.Distances$To, sep = ".")

#Merging Pair Distances and Trio ID Tables
TrioMerge <- merge(x = TrioID, y = Pair.Distances, by.x = "Trio ID", by.y = "Trio.ID")
write.csv(TrioMerge, "Trio.dist.csv")

#Averaging the distances between hosts/parasites that are paired with different non-hosts
Pair.Averages <- TrioMerge  %>%
 rename(Trio.ID = `Trio ID`) %>%
 separate(Trio.ID, into = c("Species", "Rep"), sep = 4) %>%
 mutate(Duo.ID = paste(Host, Parasite, sep = "_")) %>%
 group_by(Duo.ID, Pair, Species) %>%
 summarize(mean.Dist = mean(Dist), min.Dist = min(Dist), max.Dist = max(Dist), N = length(Dist))

#Grouping based on Duo ID
library(tidyverse)
Pair.Averages.Grouped <- Pair.Averages %>%
 dplyr::select(Pair, mean.Dist, Duo.ID, Species) %>%
 pivot_wider(names_from = Pair, values_from = mean.Dist)

write.csv(Pair.Averages.Grouped, "Pair.Averages.Grouped.csv")

#Paired T Test including All Species
t.test(x = Pair.Averages.Grouped$Host.Parasite,
     y = Pair.Averages.Grouped$NonHost.Parasite,
     paired = TRUE)

#Paired T Test for Individual Species
ACMI.Pair <- Pair.Averages.Grouped[Pair.Averages.Grouped$Species == "ACMI", ]
AQFO.Pair <- Pair.Averages.Grouped[Pair.Averages.Grouped$Species == "AQFO", ]
ASCU.Pair <- Pair.Averages.Grouped[Pair.Averages.Grouped$Species == "ASCU", ]
BADE.Pair <- Pair.Averages.Grouped[Pair.Averages.Grouped$Species == "BADE", ]
CAQU.Pair <- Pair.Averages.Grouped[Pair.Averages.Grouped$Species == "CAQU", ]
CEAR.Pair <- Pair.Averages.Grouped[Pair.Averages.Grouped$Species == "CEAR", ]
DEME.Pair <- Pair.Averages.Grouped[Pair.Averages.Grouped$Species == "DEME", ]
ERLA.Pair <- Pair.Averages.Grouped[Pair.Averages.Grouped$Species == "ERLA", ]
FERO.Pair <- Pair.Averages.Grouped[Pair.Averages.Grouped$Species == "FERO", ]
LOUT.Pair <- Pair.Averages.Grouped[Pair.Averages.Grouped$Species == "LOUT", ]
LULE.Pair <- Pair.Averages.Grouped[Pair.Averages.Grouped$Species == "LULE", ]
POGR.Pair <- Pair.Averages.Grouped[Pair.Averages.Grouped$Species == "POGR", ]
SYAL.Pair <- Pair.Averages.Grouped[Pair.Averages.Grouped$Species == "SYAL", ]

#ACMI Paired T Test
t.test(x = (ACMI.Pair$Host.Parasite),
     y = (ACMI.Pair$NonHost.Parasite),
     paired = TRUE)

#AQFO Paired T Test
```

```
t.test(x = (AQFO.Pair$Host.Parasite),
     y = (AQFO.Pair$NonHost.Parasite),
     paired = TRUE)

#ASCU Paired T Test
t.test(x = (ASCU.Pair$Host.Parasite),
     y = (ASCU.Pair$NonHost.Parasite),
     paired = TRUE)

#BADE Paired T Test
t.test(x = (BADE.Pair$Host.Parasite),
     y = (BADE.Pair$NonHost.Parasite),
     paired = TRUE)

#CAQU Paired T Test
t.test(x = (CAQU.Pair$Host.Parasite),
     y = (CAQU.Pair$NonHost.Parasite),
     paired = TRUE)

#CEAR Paired T Test
t.test(x = (CEAR.Pair$Host.Parasite),
     y = (CEAR.Pair$NonHost.Parasite),
     paired = TRUE)

#DEME Paired T Test
t.test(x = (DEME.Pair$Host.Parasite),
     y = (DEME.Pair$NonHost.Parasite),
     paired = TRUE)

#ERLA Paired T Test
t.test(x = (ERLA.Pair$Host.Parasite),
     y = (ERLA.Pair$NonHost.Parasite),
     paired = TRUE)

#FERO Paired T Test
t.test(x = (FERO.Pair$Host.Parasite),
     y = (FERO.Pair$NonHost.Parasite),
     paired = TRUE)

#LOUT Paired T Test
t.test(x = (LOUT.Pair$Host.Parasite),
     y = (LOUT.Pair$NonHost.Parasite),
     paired = TRUE)

#LULE Paired T Test
t.test(x = (LULE.Pair$Host.Parasite),
     y = (LULE.Pair$NonHost.Parasite),
     paired = TRUE)

#POGR Paired T Test
t.test(x = (POGR.Pair$Host.Parasite),
     y = (POGR.Pair$NonHost.Parasite),
     paired = TRUE)
```

```
#SYAL Paired T Test
t.test(x = (SYAL.Pair$Host.Parasite),
      y = (SYAL.Pair$NonHost.Parasite),
      paired = TRUE)

#Graphs of differences for visualizations
#(1x1)
#First with all species, even ones not included in the “official” testing with df > 2

Parasite.Diff.Plot <-
  ggplot(data = Pair.Averages.Grouped, aes(x=Host.Parasite, y=NonHost.Parasite))+
  geom_point(aes(col = Species), size = 3) +
  labs(title = "Host.Parasite Distance vs. NonHost.Parasite Distance",
      x = "Host.Parasite", y = "Nonhost.Parasite") +
  geom_abline(intercept = 0, slope = 1) +
  xlim(0.45, 1.01) +
  ylim(0.45, 1.01)+
  theme_bw() +
  theme(legend.title = element_text(color = "black", size = 12, face = "bold")) +
  guides(col = guide_legend(title = "Species"))+
  ggsave("Graphics/Parasite.Diff.Plot.jpg",
        width = 8, height = 8, units = "in", dpi = 1000)
Parasite.Diff.Plot

Parasite.Diff.Facet <-
  ggplot(data = Pair.Averages.Grouped, aes(x=Host.Parasite, y=NonHost.Parasite))+
  geom_point(aes(col = Species), size = 3) +
  facet_wrap(facets = ~Species) +
  labs(title = "Host.Parasite Distance vs. NonHost.Parasite Distance Facet",
      x = "Host.Parasite", y = "Nonhost.Parasite") +
  geom_abline(intercept = 0, slope = 1) +
  xlim(0.45, 1.01) +
  ylim(0.45, 1.01)+
  theme_bw() +
  theme(legend.title = element_text(color = "black", size = 12, face = "bold")) +
  guides(col = guide_legend(title = "Species"))+
  ggsave("Graphics/Parasite.Diff.Facet.jpg",
        width = 8, height = 8, units = "in", dpi = 1000)
Parasite.Diff.Facet

#Only the select species

Pair.Averages.Grouped.Select <- rbind(Pair.Averages.Grouped[Pair.Averages.Grouped$Species == "BADE",],
Pair.Averages.Grouped[Pair.Averages.Grouped$Species == "CAQU",],
                          Pair.Averages.Grouped[Pair.Averages.Grouped$Species == "CAQU",],
Pair.Averages.Grouped[Pair.Averages.Grouped$Species == "ERLA",],
                          Pair.Averages.Grouped[Pair.Averages.Grouped$Species == "FERO",],
Pair.Averages.Grouped[Pair.Averages.Grouped$Species == "LOUT",])

Parasite.Diff.Plot.Select <-
  ggplot(data = Pair.Averages.Grouped.Select, aes(x=Host.Parasite, y=NonHost.Parasite))+
  geom_point(aes(col = Species), size = 3) +
```

```
  labs(title = "Host.Parasite Distance vs. NonHost.Parasite Distance",
      x = "Host.Parasite", y = "Nonhost.Parasite") +
  geom_abline(intercept = 0, slope = 1) +
  xlim(0.775, 1.01) +
  ylim(0.775, 1.01)+
  theme_bw() +
  theme(legend.title = element_text(color = "black", size = 12, face = "bold")) +
  guides(col = guide_legend(title = "Species"))+
  ggsave("Graphics/Parasite.Diff.Plot.Select.jpg",
         width = 8, height = 8, units = "in", dpi = 1000)
Parasite.Diff.Plot.Select

Parasite.Diff.Facet.Select <-
  ggplot(data = Pair.Averages.Grouped.Select, aes(x=Host.Parasite, y=NonHost.Parasite))+
  geom_point(aes(col = Species), size = 3) +
  facet_wrap(facets = ~Species) +
  labs(title = "Host.Parasite Distance vs. NonHost.Parasite Distance Facet",
      x = "Host.Parasite", y = "Nonhost.Parasite") +
  geom_abline(intercept = 0, slope = 1) +
  xlim(0.775, 1.01) +
  ylim(0.775, 1.01)+
  theme_bw() +
  theme(legend.title = element_text(color = "black", size = 12, face = "bold")) +
  guides(col = guide_legend(title = "Species"))+
  ggsave("Graphics/Parasite.Diff.Facet.Select.jpg",
         width = 8, height = 8, units = "in", dpi = 1000)
Parasite.Diff.Facet.Select

###ORDINATIONS
#3 Dimensional Ordinations
#Species Ordination
SpeciesNMDS3D <- metaMDS(comm = AbundanceMBWOBR, autotransform = FALSE,
                         distance = "bray", engine = "monoMDS",
                         k = 3, weakties = TRUE, model = "global",
                         maxit = 300, try = 40, trymax = 100)
SpeciesNMDS3D$stress

SpeciesNMDS3DPoints <- data.frame(SpeciesNMDS3D$points)

AllDataABPointsMBWOBR3D <- data.frame(AllDataABMBWOBR, SpeciesNMDS3DPoints)
AllDataABPointsMBWOBR3D$SpeciesCode <- OTUAbundanceMBTableWOBR$SpeciesCode

#ORDINATION WITH SPECIES OVERLAY
SpeciesOrdination3D <-
  ggplot(data = AllDataABPointsMBWOBR3D, aes(x=MDS1, y=MDS2))+
  geom_point(aes(col = Species), size = 4) +
  labs(title = "NMDS with Species Overlay 3D",
      x = "MDS1", y = "MDS2") +
  theme_bw() +
  theme(axis.line = element_line(),
        axis.ticks = element_blank(),
        axis.text = element_blank(),
        legend.title = element_text(color = "black", size = 12, face = "bold")) +
```

```
  guides(col = guide_legend(title = "Species"))+
  ggsave("Graphics/PlantSpeciesOrdination3D WO438.jpg",
       width = 8, height = 8, units = "in", dpi = 1000)
SpeciesOrdination3D

SpeciesOrdination3D.GOF <- goodness(object = SpeciesNMDS3D)
SpeciesOrdinationShepard.3D <- plot(SpeciesNMDS3D$diss, SpeciesNMDS3D$dist)
SpeciesOrdinationShepard.3D <- stressplot(SpeciesNMDS3D, p.col = "blue", l.col = "red", lwd = 2)

#ORDINATION WITH FAMILY OVERLAY

AllDataABTaxonomy <- AllDataABMBWOBR
AllDataABTaxonomy$SpeciesCode <- as.character(OTUAbundanceMBTableWOBR$SpeciesCode)
PlantTaxonomy$Code <- as.character(PlantTaxonomy$Code)
PlantTaxonomy$Kingdom <- as.character(PlantTaxonomy$Kingdom)
PlantTaxonomy$Division <- as.character(PlantTaxonomy$Division)
PlantTaxonomy$Class <- as.character(PlantTaxonomy$Class)
PlantTaxonomy$Order <- as.character(PlantTaxonomy$Order)
PlantTaxonomy$Family <- as.character(PlantTaxonomy$Family)
PlantTaxonomy$Genus <- as.character(PlantTaxonomy$Genus)
PlantTaxonomy$Species <- as.character(PlantTaxonomy$Species)

AllDataABTaxonomy$Family <- ifelse(AllDataABTaxonomy$SpeciesCode == "ACMI",
PlantTaxonomy[PlantTaxonomy$Code=="ACMI","Family"],
                ifelse(AllDataABTaxonomy$SpeciesCode == "AQFO",
PlantTaxonomy[PlantTaxonomy$Code=="AQFO", "Family"],
                ifelse(AllDataABTaxonomy$SpeciesCode == "ASCU",
PlantTaxonomy[PlantTaxonomy$Code=="ASCU", "Family"],
                ifelse(AllDataABTaxonomy$SpeciesCode == "BADE",
PlantTaxonomy[PlantTaxonomy$Code=="BADE", "Family"],
                ifelse(AllDataABTaxonomy$SpeciesCode == "CAQU",
PlantTaxonomy[PlantTaxonomy$Code=="CAQU", "Family"],
                ifelse(AllDataABTaxonomy$SpeciesCode == "CALE",
PlantTaxonomy[PlantTaxonomy$Code=="CALE", "Family"],
                ifelse(AllDataABTaxonomy$SpeciesCode == "CEAR",
PlantTaxonomy[PlantTaxonomy$Code=="CEAR", "Family"],
                ifelse(AllDataABTaxonomy$SpeciesCode == "DEME",
PlantTaxonomy[PlantTaxonomy$Code=="DEME", "Family"],
                ifelse(AllDataABTaxonomy$SpeciesCode == "ERSP",
PlantTaxonomy[PlantTaxonomy$Code=="ERSP", "Family"],
                ifelse(AllDataABTaxonomy$SpeciesCode == "ERLA",
PlantTaxonomy[PlantTaxonomy$Code=="ERLA", "Family"],
                ifelse(AllDataABTaxonomy$SpeciesCode == "FERO",
PlantTaxonomy[PlantTaxonomy$Code=="FERO", "Family"],
                ifelse(AllDataABTaxonomy$SpeciesCode == "LOTR",
PlantTaxonomy[PlantTaxonomy$Code=="LOTR", "Family"],
                ifelse(AllDataABTaxonomy$SpeciesCode == "LOUT",
PlantTaxonomy[PlantTaxonomy$Code=="LOUT", "Family"],
                ifelse(AllDataABTaxonomy$SpeciesCode == "LULE",
PlantTaxonomy[PlantTaxonomy$Code=="LULE", "Family"],
                ifelse(AllDataABTaxonomy$SpeciesCode == "POGR",
PlantTaxonomy[PlantTaxonomy$Code=="POGR", "Family"],
```

```
                    ifelse(AllDataABTaxonomy$SpeciesCode == "SYAL",
PlantTaxonomy[PlantTaxonomy$Code=="SYAL", "Family"],
                    NA)))))))))))))))

AllDataABFamilyPoints3D <- data.frame(AllDataABTaxonomy, SpeciesNMDS3DPoints)

PlantTaxonomyOrdination3D <-
 ggplot(data = AllDataABFamilyPoints3D, aes(x=MDS1, y=MDS2))+
 geom_point(aes(col = Family), size = 4) +
 labs(title = "NMDS with Plant Family Overlay 3D",
     x = "MDS1", y = "MDS2") +
 theme_bw() +
 theme(axis.line = element_line(),
       axis.ticks = element_blank(),
       axis.text = element_blank(),
       legend.title = element_text(color = "black", size = 12, face = "bold")) +
 guides(col = guide_legend(title = "Family"))+
 ggsave("Graphics/PlantFamilyOrdination3D WO438.jpg",
        width = 8, height = 8, units = "in", dpi = 1000)
PlantTaxonomyOrdination3D

#Species Facet
SpeciesFacet <-
 ggplot(data = AllDataABPointsMBWOBR3D, aes(x=MDS1, y=MDS2))+
 geom_point(data = transform(AllDataABPointsMBWOBR3D, Species = NULL), colour = "grey85") +
 geom_point(aes(col = Species), size = 4) +
 facet_wrap(facets = ~Species) +
 labs(title = "NMDS and Species",
     x = "MDS1", y = "MDS2") +
#  geom_text_repel(aes(label = AllDataABPointsMBWOBR3D$Sample), point.padding = unit(0.1, "lines"), vjust = 0,
size = 3.5)+
 theme_bw() +
 theme(axis.line = element_line(),
       axis.ticks = element_blank(),
       axis.text = element_blank(),
       legend.title = element_text(color = "black", size = 12, face = "bold")) +
 guides(col = guide_legend(title = "Species"))+
 ggsave("Graphics/SpeciesFacet 3D.jpg",
        width = 8, height = 8, units = "in", dpi = 1000)
SpeciesFacet

#Facet with Initial Disturbance overlay
AllDataABPointsID <- merge(x = AllDataABPointsMBWOBR3D, y = SiteTreatmentMetadata[ , c("Plot.Name",
"Disturbance.Treatment", "Disturbance.Regime", "Burn.Mow", "Date.Last.Treatment")],
                           by.x = "Collection.Site", by.y = "Plot.Name", all.x = TRUE)
AllDataABPointsID <- AllDataABPointsID %>%
 select("Disturbance.Treatment", "Disturbance.Regime", "Burn.Mow", "Date.Last.Treatment", everything())

AllDataABPointsID <- AllDataABPointsID[c(-grep("GHP-  2009 Array", AllDataABPointsID$Collection.Site), -
grep("GHP- Mounded", AllDataABPointsID$Collection.Site), -grep("GHP- Mounded #2",
AllDataABPointsID$Collection.Site), -grep("SM Fenced: East", AllDataABPointsID$Collection.Site),
```

```
                                -grep("GHP- 100X2 2011", AllDataABPointsID$Collection.Site), -grep("GHP- 100X3
2012", AllDataABPointsID$Collection.Site), -grep("GHP- 10X1 SF", AllDataABPointsID$Collection.Site), -grep("GHP-
10X2 BF", AllDataABPointsID$Collection.Site),
                                -grep("GHP- 10X5 HF", AllDataABPointsID$Collection.Site), -grep("GHP- 10X6 HF",
AllDataABPointsID$Collection.Site), -grep("SM 10X1 HM", AllDataABPointsID$Collection.Site)), ]

InitialDisturbanceFacet <-
 ggplot(data = AllDataABPointsID, aes(x=MDS1, y=MDS2))+
 geom_point(data = AllDataABPointsID, colour = "grey85") +
 geom_point(aes(col = Disturbance.Treatment, shape = GHP.SM), size = 2) +
 facet_wrap(facets = ~Species) +
 labs(title = "NMDS with Initial Disturbance Overlay",
    x = "MDS1", y = "MDS2") +
 theme_bw() +
 theme(axis.line = element_line(),
      axis.ticks = element_blank(),
      axis.text = element_blank(),
      legend.title = element_text(color = "black", size = 12, face = "bold")) +
 guides(col = guide_legend(title = "Blank Batch"))+
 ggsave("Graphics/InitialDisturbanceFacetMBWOBR WOB1B2.jpg",
       width = 10, height = 10, units = "in", dpi = 1000)
InitialDisturbanceFacet

#Just GHP Initial Disturbance Ordination

JustGHP.PointsID <- AllDataABPointsID[c((-grep("SM 10X1 HM",  AllDataABPointsID$Collection.Site)), (-grep("SM
2010 06", AllDataABPointsID$Collection.Site)), (-grep("SM 2010 11",  AllDataABPointsID$Collection.Site)), (-
grep("SM 2010 13", AllDataABPointsID$Collection.Site)),
                                (-grep("SM 2010 16",  AllDataABPointsID$Collection.Site)), (-grep("SM 2010 18",
AllDataABPointsID$Collection.Site)), (-grep("SM 2010 22",  AllDataABPointsID$Collection.Site)), (-grep("SM 2010
28", AllDataABPointsID$Collection.Site)),
                                (-grep("SM 2010 29",  AllDataABPointsID$Collection.Site)), (-grep("SM 2011 16",
AllDataABPointsID$Collection.Site)), (-grep("SM 2011 26",  AllDataABPointsID$Collection.Site)), (-grep("SM 2011
28", AllDataABPointsID$Collection.Site))),]

InitialDisturbanceFacetGHP <-
 ggplot(data = JustGHP.PointsID, aes(x=MDS1, y=MDS2))+
 geom_point(data = JustGHP.PointsID, colour = "grey85") +
 geom_point(aes(col = Disturbance.Treatment), size = 2) +
 facet_wrap(facets = ~Species) +
 labs(title = "NMDS with Initial Disturbance Overlay (GHP)",
    x = "MDS1", y = "MDS2") +
 theme_bw() +
 theme(axis.line = element_line(),
      axis.ticks = element_blank(),
      axis.text = element_blank(),
      legend.title = element_text(color = "black", size = 12, face = "bold")) +
 guides(col = guide_legend(title = "Blank Batch"))+
 ggsave("Graphics/InitialDisturbanceFacetGHP WOB1B2.jpg",
       width = 10, height = 10, units = "in", dpi = 1000)
InitialDisturbanceFacetGHP

#Facet with Continuous Disturbance Overlay
```

```
ContinuousDisturbanceFacet <-
  ggplot(data = AllDataABPointsID, aes(x=MDS1, y=MDS2))+
  geom_point(data = AllDataABPointsID, colour = "grey85") +
  geom_point(aes(col = Disturbance.Regime, shape = Burn.Mow), size = 2) +
  facet_wrap(facets = ~Species) +
  labs(title = "NMDS with Continuous Disturbance Overlay",
       x = "MDS1", y = "MDS2") +
  theme_bw() +
  theme(axis.line = element_line(),
        axis.ticks = element_blank(),
        axis.text = element_blank(),
        legend.title = element_text(color = "black", size = 12, face = "bold")) +
  guides(col = guide_legend(title = "Blank Batch"))+
  ggsave("Graphics/DisturbanceRegimeFacetMBWOBR WOB1B2.jpg",
         width = 10, height = 10, units = "in", dpi = 1000)
ContinuousDisturbanceFacet

#JUSTGHP
InitialDisturbanceFacetGHP <-
  ggplot(data = JustGHP.PointsID, aes(x=MDS1, y=MDS2))+
  geom_point(data = JustGHP.PointsID, colour = "grey85") +
  geom_point(aes(col = Disturbance.Regime, shape = Burn.Mow), size = 2) +
  facet_wrap(facets = ~Species) +
  labs(title = "NMDS with Continuous Disturbance Overlay (GHP)",
       x = "MDS1", y = "MDS2") +
  theme_bw() +
  theme(axis.line = element_line(),
        axis.ticks = element_blank(),
        axis.text = element_blank(),
        legend.title = element_text(color = "black", size = 12, face = "bold")) +
  guides(col = guide_legend(title = "Blank Batch"))+
  ggsave("Graphics/DisturbanceRegimeFacetGHP WOB1B2.jpg",
         width = 10, height = 10, units = "in", dpi = 1000)
InitialDisturbanceFacetGHP

#SPECIES ORDINATION with GHP/SM SITES OVERLAY
SitesOrdination <-
  ggplot(data = AllDataABPointsMBWOBR3D, aes(x=MDS1, y=MDS2))+
  geom_point(aes(col = Species, shape = GHP.SM), size = 4) +
  labs(title = "NMDS with Species and Sites Overlay",
       x = "MDS1", y = "MDS2") +
  theme_bw() +
  theme(axis.line = element_line(),
        axis.ticks = element_blank(),
        axis.text = element_blank(),
        legend.title = element_text(color = "black", size = 12, face = "bold")) +
  guides(col = guide_legend(title = "Sites"))
SitesOrdination

#DIFFERENCE IN ACMI BETWEEN GHP and SM
Sites.ACMI.NMDS.2D <- metaMDS(comm = ACMIAbundance, autotransform = FALSE,
                    distance = "bray", engine = "monoMDS",
                    k = 2, weakties = TRUE, model = "global",
```

```
                maxit = 300, try = 40, trymax = 100)
Sites.ACMI.NMDS.2D$stress
Sites.ACMI.NMDS.2DPoints <- data.frame(Sites.ACMI.NMDS.2D$points)
AllDataABPoints.ACMI <- data.frame(ACMI, Sites.ACMI.NMDS.2DPoints)

SitesOrdination.ACMI <-
 ggplot(data = AllDataABPoints.ACMI, aes(x=MDS1, y=MDS2))+
 geom_point(aes(col = GHP.SM), size = 4) +
 labs(title = "ACMI NMDS with Sites Overlay",
    x = "MDS1", y = "MDS2") +
 theme_bw() +
 theme(axis.line = element_line(),
      axis.ticks = element_blank(),
      axis.text = element_blank(),
      legend.title = element_text(color = "black", size = 12, face = "bold")) +
 guides(col = guide_legend(title = "Sites"),
 ggsave("Graphics/SitesOrdination.ACMI.jpg",
 width = 6, height = 6, units = "in", dpi = 1000))
SitesOrdination.ACMI

#DIFFERENCE IN CALE BETWEEN GHP and SM
Sites.CALE.NMDS.2D <- metaMDS(comm = CALEAbundance, autotransform = FALSE,
                distance = "bray", engine = "monoMDS",
                k = 2, weakties = TRUE, model = "global",
                maxit = 300, try = 40, trymax = 100)
Sites.CALE.NMDS.2D$stress
Sites.CALE.NMDS.2DPoints <- data.frame(Sites.CALE.NMDS.2D$points)
AllDataABPoints.CALE <- data.frame(CALE, Sites.CALE.NMDS.2DPoints)

SitesOrdination.CALE <-
 ggplot(data = AllDataABPoints.CALE, aes(x=MDS1, y=MDS2))+
 geom_point(aes(col = GHP.SM), size = 4) +
 labs(title = "CALE NMDS with Sites Overlay",
    x = "MDS1", y = "MDS2") +
 theme_bw() +
 theme(axis.line = element_line(),
      axis.ticks = element_blank(),
      axis.text = element_blank(),
      legend.title = element_text(color = "black", size = 12, face = "bold")) +
 guides(col = guide_legend(title = "Sites"),
 ggsave("Graphics/SitesOrdination.CALE.jpg",
    width = 6, height = 6, units = "in", dpi = 1000))
SitesOrdination.CALE

#DIFFERENCE IN ERLA BETWEEN GHP and SM
Sites.ERLA.NMDS.2D <- metaMDS(comm = ERLAAbundance, autotransform = FALSE,
                distance = "bray", engine = "monoMDS",
                k = 2, weakties = TRUE, model = "global",
                maxit = 300, try = 40, trymax = 100)
Sites.ERLA.NMDS.2D$stress
Sites.ERLA.NMDS.2DPoints <- data.frame(Sites.ERLA.NMDS.2D$points)
AllDataABPoints.ERLA <- data.frame(ERLA, Sites.ERLA.NMDS.2DPoints)
```

```
SitesOrdination.ERLA <-
  ggplot(data = AllDataABPoints.ERLA, aes(x=MDS1, y=MDS2))+
  geom_point(aes(col = GHP.SM), size = 4) +
  labs(title = "ERLA NMDS with Sites Overlay",
      x = "MDS1", y = "MDS2") +
#  geom_text_repel(aes(label = AllDataABPoints.ERLA$Sample), point.padding = unit(0.1, "lines"), vjust = 0, size =
3.5)+
  theme_bw() +
  theme(axis.line = element_line(),
        axis.ticks = element_blank(),
        axis.text = element_blank(),
        legend.title = element_text(color = "black", size = 12, face = "bold")) +
  guides(col = guide_legend(title = "Sites"),
  ggsave("Graphics/SitesOrdination.ERLA.jpg",
  width = 6, height = 6, units = "in", dpi = 1000))
SitesOrdination.ERLA

#DIFFERENCE IN FERO BETWEEN GHP and SM
Sites.FERO.NMDS.2D <- metaMDS(comm = FEROAbundance, autotransform = FALSE,
                      distance = "bray", engine = "monoMDS",
                      k = 2, weakties = TRUE, model = "global",
                      maxit = 300, try = 40, trymax = 100)
Sites.FERO.NMDS.2D$stress
Sites.FERO.NMDS.2DPoints <- data.frame(Sites.FERO.NMDS.2D$points)
AllDataABPoints.FERO <- data.frame(FERO, Sites.FERO.NMDS.2DPoints)

SitesOrdination.FERO <-
  ggplot(data = AllDataABPoints.FERO, aes(x=MDS1, y=MDS2))+
  geom_point(aes(col = GHP.SM), size = 4) +
  labs(title = "FERO NMDS with Sites Overlay",
      x = "MDS1", y = "MDS2") +
  theme_bw() +
  theme(axis.line = element_line(),
        axis.ticks = element_blank(),
        axis.text = element_blank(),
        legend.title = element_text(color = "black", size = 12, face = "bold")) +
  guides(col = guide_legend(title = "Sites"),
  ggsave("Graphics/SitesOrdination.FERO.jpg",
  width = 6, height = 6, units = "in", dpi = 1000))
SitesOrdination.FERO

###COMBO ACMI, CALE, ERLA and FERO

Sites.ALLSP.NMDS.2D <- metaMDS(comm = ALLSPAbundance, autotransform = FALSE,
                      distance = "bray", engine = "monoMDS",
                      k = 2, weakties = TRUE, model = "global",
                      maxit = 300, try = 40, trymax = 100)
Sites.ALLSP.NMDS.2D$stress
Sites.ALLSP.NMDS.2DPoints <- data.frame(Sites.ALLSP.NMDS.2D$points)
AllDataABPoints.ALLSP <- data.frame(ALLSP, Sites.ALLSP.NMDS.2DPoints)

SitesOrdination.ALLSP <-
  ggplot(data = AllDataABPoints.ALLSP, aes(x=MDS1, y=MDS2))+
```

```
 geom_point(aes(col = GHP.SM, shape= Species, fill = GHP.SM), size = 2) +
 scale_shape_manual(values = c(21, 22, 23, 24)) +
 scale_fill_manual(values = c("deepskyblue2", "firebrick2")) +
 scale_color_manual(values = c("gray0", "gray0")) +
 labs(title = "ACMI, CALE, ERLA and FERO NMDS with Sites Overlay",
    x = "MDS1", y = "MDS2") +
 theme_bw() +
 theme(axis.line = element_line(),
     axis.ticks = element_blank(),
     axis.text = element_blank(),
     legend.title = element_text(color = "black", size = 12, face = "bold")) +
 guides(col = guide_legend(title = "Sites"),
      ggsave("Graphics/SitesOrdination.ALLSP.jpg",
          width = 6, height = 6, units = "in", dpi = 1000))
SitesOrdination.ALLSP

#Ordination for CEAR in Initial Disturbance Treatment
CEAR <- AllDataABMBWOBR.TreatmentT[AllDataABMBWOBR.TreatmentT$Species == "Cerastium arvense", ]
CEAR <- CEAR[-c(grep("N/A", CEAR$Disturbance.Treatment), grep("N/A", CEAR$Disturbance.Regime)), ]
CEARAbundance <- CEAR[ , colnames(CEAR) %in% OTUNumbers]

Disturbance.CEAR <- metaMDS(comm = CEARAbundance, autotransform = FALSE,
                    distance = "bray", engine = "monoMDS",
                    k = 2, weakties = TRUE, model = "global",
                    maxit = 300, try = 40, trymax = 100)
Disturbance.CEAR$stress
Disturbance.CEARPoints <- data.frame(Disturbance.CEAR$points)
AllDataABPoints.CEAR <- data.frame(CEAR, Disturbance.CEARPoints)

Initial.Disurbance.Ordination.CEAR <-
 ggplot(data = AllDataABPoints.CEAR, aes(x=MDS1, y=MDS2))+
 geom_point(aes(col = Disturbance.Treatment), size = 4) +
 labs(title = "CEAR NMDS with Initial Disturbance Treatment Overlay",
    x = "MDS1", y = "MDS2") +
 theme_bw() +
 theme(axis.line = element_line(),
     axis.ticks = element_blank(),
     axis.text = element_blank(),
     legend.title = element_text(color = "black", size = 12, face = "bold")) +
 guides(col = guide_legend(title = "Sites"),
 ggsave("Graphics/Initial.Disturbance.CEAR.jpg",
 width = 6, height = 6, units = "in", dpi = 1000))
Initial.Disurbance.Ordination.CEAR

#Ordination for CEAR with Initial Disturbance Treatment AND Year of Inception
AllDataABPoints.CEAR$YearInception <- c("2009", "2009", "2009", "2009", "2009", "2009", "2009", "2009", "2009",
"2009", "2009", "2010", "2010", "2010", "2010", "2011", "2011")

InitialDisurbance.and.YearInception.Ordination.CEAR <-
 ggplot(data = AllDataABPoints.CEAR, aes(x=MDS1, y=MDS2))+
 geom_point(aes(col = Disturbance.Treatment, shape = YearInception), size = 4) +
 labs(title = "CEAR NMDS with Initial Disturbance Treatment and Year of Inception Overlay",
    x = "MDS1", y = "MDS2") +
```

```
  theme_bw() +
  theme(axis.line = element_line(),
        axis.ticks = element_blank(),
        axis.text = element_blank(),
        legend.title = element_text(color = "black", size = 12, face = "bold")) +
  guides(col = guide_legend(title = "Sites"),
         ggsave("Graphics/InitialDisturbance.and.YearInception.CEAR.jpg",
          width = 6, height = 6, units = "in", dpi = 1000))
InitialDisurbance.and.YearInception.Ordination.CEAR

#Ordination for CEAR in Disturbance Regime Treatment
Disturbance.Regime.Ordination.CEAR <-
  ggplot(data = AllDataABPoints.CEAR, aes(x=MDS1, y=MDS2))+
  geom_point(aes(col = Disturbance.Regime), size = 4) +
  labs(title = "CEAR NMDS with Disturbance Regime Treatment Overlay",
       x = "MDS1", y = "MDS2") +
  theme_bw() +
  theme(axis.line = element_line(),
        axis.ticks = element_blank(),
        axis.text = element_blank(),
        legend.title = element_text(color = "black", size = 12, face = "bold")) +
  guides(col = guide_legend(title = "Sites"),
  ggsave("Graphics/Disturbance.Regime.CEAR.jpg",
  width = 6, height = 6, units = "in", dpi = 1000))
Disturbance.Regime.Ordination.CEAR

#Ordination for CEAR with Disturbance Treatment AND Date Last Treatment
Disturbance.and.Year.Ordination.CEAR <-
  ggplot(data = AllDataABPoints.CEAR, aes(x=MDS1, y=MDS2))+
  geom_point(aes(col = Disturbance.Regime, shape = Date.Last.Treatment), size = 4) +
  labs(title = "CEAR NMDS with Disturbance Regime and Date of Last Treatment Overlay",
       x = "MDS1", y = "MDS2") +
  theme_bw() +
  theme(axis.line = element_line(),
        axis.ticks = element_blank(),
        axis.text = element_blank(),
        legend.title = element_text(color = "black", size = 12, face = "bold")) +
  guides(col = guide_legend(title = "Sites"),
         ggsave("Graphics/DisturbanceRegime.and.DateofLastBurn.CEAR.jpg",
              width = 6, height = 6, units = "in", dpi = 1000))
Disturbance.and.Year.Ordination.CEAR

#ORDINATION WITH HOST/NON-HOST/PARASITE OVERLAY
HostOrdination <-
  ggplot(data = AllDataABPointsMBWOBR3D, aes(x=MDS1, y=MDS2))+
  geom_point(aes(col = Species, shape = ParasiteStatus), size = 4) +
  labs(title = "NMDS and Host/Non-host/Parasite Status",
       x = "MDS1", y = "MDS2") +
  theme_bw() +
  theme(axis.line = element_line(),
        axis.ticks = element_blank(),
        axis.text = element_blank(),
        legend.title = element_text(color = "black", size = 12, face = "bold")) +
```

```
  guides(col = guide_legend(title = "Host/Non-Host/Parasite Status"))
HostOrdination


#Facet with Host/Non-host/Parasite overlay
SpeciesParasiteFacet <-
  ggplot(data = AllDataABPointsMBWOBR3D, aes(x=MDS1, y=MDS2))+
  geom_point(data = AllDataABPointsMBWOBR3D[AllDataABPointsMBWOBR3D$Host.Parasite == "Parasite", ],
colour = "grey85") +
  geom_point(aes(col = ParasiteStatus, shape = ParasiteStatus), size = 2) +
  facet_wrap(facets = ~Species) +
  labs(title = "NMDS with Parasite Status Overlay",
      x = "MDS1", y = "MDS2") +
  theme_bw() +
  theme(axis.line = element_line(),
        axis.ticks = element_blank(),
        axis.text = element_blank(),
        legend.title = element_text(color = "black", size = 12, face = "bold")) +
  guides(col = guide_legend(title = "Parasite Status"))+
  ggsave("Graphics/SpeciesParasiteFacetMBWOBR WOB1B2.jpg",
         width = 10, height = 10, units = "in", dpi = 1000)
SpeciesParasiteFacet

###PHYLOGENETIC HISTOGRAMS####
#For the OTU Table: Taxa are Columns and Samples are Rows
OTUTableWOBR <- AbundanceMBWOBR
rownames(OTUTableWOBR) <- SampleNumberWOBR
colnames(OTUTableWOBR) <- OTUNumbersWOBR
OTUTableWOBR <- as.matrix(OTUTableWOBR)
OTUWOBR <- otu_table(OTUTableWOBR, taxa_are_rows = FALSE)

TAXATableWOBR <- as.data.frame(Taxa[,c("Kingdom", "Phylum", "Class", "Order", "Family", "Genus", "Species")])
rownames(TAXATableWOBR) <- (OTUNumbersWOBR)
colnames(TAXATableWOBR) <- c("Kingdom", "Phylum", "Class", "Order", "Family", "Genus", "Species")
TAXATableWOBR <- as.matrix(TAXATableWOBR)
TAXWOBR <- tax_table(TAXATableWOBR)
physeq <- phyloseq(OTUWOBR, TAXWOBR)
Phylum level
plot_bar(physeq, fill = "Phylum")
PhylumGlommed <- tax_glom(physeq, "Phylum")
plot_bar(PhylumGlommed, fill ="Phylum")

#Class level
#plot_bar(physeq, fill = "Class")
#ClassGlommed <- tax_glom(physeq, "Class")
#plot_bar(ClassGlommed, fill ="Class")

#Family level
#plot_bar(physeq, fill = "Family")
#FamilyGlommed <- tax_glom(physeq, "Family")
#plot_bar(FamilyGlommed, fill ="Family")

TaxPhylumTable <- count(Taxa, vars=Taxa$Phylum)
```

```
TaxPhylumTable <- TaxPhylumTable[order(TaxPhylumTable$n, decreasing = TRUE),]
colnames(TaxPhylumTable) <- c("Phyla", "OTUs")
TaxClassTable <- count(Taxa, vars = Taxa$Class)
TaxClassTable <- TaxClassTable[order(TaxClassTable$n, decreasing = TRUE),]
colnames(TaxClassTable) <- c("Phyla", "OTUs")
TaxOrderTable <- count(Taxa, vars = Taxa$Order)
TaxOrderTable <- TaxOrderTable[order(TaxOrderTable$n, decreasing = TRUE),]
colnames(TaxOrderTable) <- c("Phyla", "OTUs")
TaxFamilyTable <- count(Taxa, vars = Taxa$Family)
TaxFamilyTable <- TaxFamilyTable[order(TaxFamilyTable$n, decreasing = TRUE),]
colnames(TaxFamilyTable) <- c("Phyla", "OTUs")
TaxGenusTable <- count(Taxa, vars = Taxa$Genus)
TaxGenusTable <- TaxGenusTable[order(TaxGenusTable$n, decreasing = TRUE),]
colnames(TaxGenusTable) <- c("Phyla", "OTUs")
TaxSpeciesTable <- count(Taxa, vars = Taxa$Species)
TaxSpeciesTable <- TaxSpeciesTable[order(TaxSpeciesTable$n, decreasing = TRUE),]
colnames(TaxSpeciesTable) <- c("Phyla", "OTUs")

write.csv(TaxPhylumTable, "Tables/Tables for Graphics/TaxPhylumTable.csv")
write.csv(TaxClassTable, "Tables/Tables for Graphics/TaxClassTable.csv")
write.csv(TaxOrderTable, "Tables/Tables for Graphics/TaxOrderTable.csv")
write.csv(TaxFamilyTable, "Tables/Tables for Graphics/TaxFamilyTable.csv")
write.csv(TaxGenusTable, "Tables/Tables for Graphics/TaxGenusTable.csv")

#For the OTU Table: Taxa are Columns and Species are Rows

AveACMIAbundance <- AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Species == "Achillea
millefolium", colnames(AbundanceMetadataMBWOBR) %in% OTUNumbers]
AveACMIAbundance <- colMeans(AveACMIAbundance)
AveAQFOAbundance <- AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Species == "Aquilegia
formosa", colnames(AbundanceMetadataMBWOBR) %in% OTUNumbers]
AveAQFOAbundance <- colMeans(AveAQFOAbundance)
AveASCUAbundance <- AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Species == "Aster curtisii",
colnames(AbundanceMetadataMBWOBR) %in% OTUNumbers]
AveASCUAbundance <- colMeans(AveASCUAbundance)
AveBADEAbundance <- AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Species == "Balsamorhiza
deltoidea", colnames(AbundanceMetadataMBWOBR) %in% OTUNumbers]
AveBADEAbundance <- colMeans(AveBADEAbundance)
AveCALEAbundance <- AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Species == "Castilleja
levisecta", colnames(AbundanceMetadataMBWOBR) %in% OTUNumbers]
AveCALEAbundance <- colMeans(AveCALEAbundance)
AveCAQUAbundance <- AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Species == "Camassia
quamash", colnames(AbundanceMetadataMBWOBR) %in% OTUNumbers]
AveCAQUAbundance <- colMeans(AveCAQUAbundance)
AveCEARAbundance <- AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Species == "Cerastium
arvense", colnames(AbundanceMetadataMBWOBR) %in% OTUNumbers]
AveCEARAbundance <- colMeans(AveCEARAbundance)
AveDEMEAbundance <- AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Species == "Delphinium
menziesii", colnames(AbundanceMetadataMBWOBR) %in% OTUNumbers]
AveDEMEAbundance <- colMeans(AveDEMEAbundance)
AveERLAAbundance <- AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Species == "Eriophyllum
lanatum", colnames(AbundanceMetadataMBWOBR) %in% OTUNumbers]
AveERLAAbundance <- colMeans(AveERLAAbundance)
```

```
AveERSPAbundance <- AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Species == "Erigeron
speciosus", colnames(AbundanceMetadataMBWOBR) %in% OTUNumbers]
AveERSPAbundance <- colMeans(AveERSPAbundance)
AveFEROAbundance <- AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Species == "Festuca
roemeri", colnames(AbundanceMetadataMBWOBR) %in% OTUNumbers]
AveFEROAbundance <- colMeans(AveFEROAbundance)
AveLOTRAbundance <- AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Species == "Lomatium
triternatum", colnames(AbundanceMetadataMBWOBR) %in% OTUNumbers]
AveLOTRAbundance <- colMeans(AveLOTRAbundance)
AveLOUTAbundance <- AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Species == "Lomatium
utriculatum", colnames(AbundanceMetadataMBWOBR) %in% OTUNumbers]
AveLOUTAbundance <- colMeans(AveLOUTAbundance)
AveLULEAbundance <- AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Species == "Lupinus
lepidus", colnames(AbundanceMetadataMBWOBR) %in% OTUNumbers]
AveLULEAbundance <- colMeans(AveLULEAbundance)
AvePOGRAbundance <- AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Species == "Potentilla
gracilius", colnames(AbundanceMetadataMBWOBR) %in% OTUNumbers]
AvePOGRAbundance <- colMeans(AvePOGRAbundance)
AveSYALAbundance <- AbundanceMetadataMBWOBR[AbundanceMetadataMBWOBR$Species ==
"Symphoricarpos albus", colnames(AbundanceMetadataMBWOBR) %in% OTUNumbers]
AveSYALAbundance <- colMeans(AveSYALAbundance)
AverageAbundanceTable <- cbind(AveACMIAbundance, AveAQFOAbundance, AveASCUAbundance,
AveBADEAbundance, AveCALEAbundance, AveCAQUAbundance, AveCEARAbundance, AveDEMEAbundance,
AveERLAAbundance, AveERSPAbundance, AveFEROAbundance, AveLOTRAbundance, AveLOUTAbundance,
AveLULEAbundance, AvePOGRAbundance, AveSYALAbundance)
AverageAbundanceTable <- t(round(AverageAbundanceTable, digits = 0))
rownames(AverageAbundanceTable) <- c("ACMI", "AQFO", "ASCU", "BADE", "CALE", "CAQU", "CEAR", "DEME",
"ERLA", "ERSP", "FERO", "LOTR", "LOUT", "LULE", "POGR", "SYAL")
as.data.frame(AverageAbundanceTable)
write.csv(AverageAbundanceTable, "Tables/AverageAbundanceTable.csv")
OTUTableSPECIESWOBR <- AverageAbundanceTable
OTUTableSPECIESWOBR <- as.matrix(OTUTableSPECIESWOBR)
OTUSPECIESWOBR <- otu_table(OTUTableSPECIESWOBR, taxa_are_rows = FALSE)

TAXATableWOBR <- as.data.frame(Taxa[,c("Kingdom", "Phylum", "Class", "Order", "Family", "Genus", "Species")])
rownames(TAXATableWOBR) <- (OTUNumbersWOBR)
colnames(TAXATableWOBR) <- c("Kingdom", "Phylum", "Class", "Order", "Family", "Genus", "Species")
TAXATableWOBR <- as.matrix(TAXATableWOBR)
TAXWOBR <- tax_table(TAXATableWOBR)
write.csv(TAXATableWOBR, "Tables/TAXATableWOBR.csv")

physeqSPECIES <- phyloseq(OTUSPECIESWOBR, TAXWOBR)
PhylumGlommedSPECIES <- tax_glom(physeqSPECIES, "Phylum")
plot_bar(PhylumGlommedSPECIES, fill ="Phylum")

#What comprises each species community profile?
FirstOTU <- which(colnames(ACMIGroupMB) == "OTU0001")
LastOTU <- which(colnames(ACMIGroupMB) == "OTU7365")

ACMITAX <- ACMIGroupMB[c(ACMIGroupMB$Species == "Achillea millefolium"), FirstOTU:LastOTU]
AQFOTAX <- AQFOGroupMB[c(AQFOGroupMB$Species == "Aquilegia formosa"), FirstOTU:LastOTU]
ASCUTAX <- ASCUGroupMB[c(ASCUGroupMB$Species == "Aster curtisii"), FirstOTU:LastOTU]
BADETAX <- BADEGroupMB[c(BADEGroupMB$Species == "Balsamorhiza deltoidea"), FirstOTU:LastOTU]
```

```
CALETAX <- AbundanceMetadataMBWOBR[c(AbundanceMetadataMBWOBR$Species == "Castilleja levisecta"),
FirstOTU:LastOTU]
CAQUTAX <- CAQUGroupMB[c(CAQUGroupMB$Species == "Camassia quamash"), FirstOTU:LastOTU]
CEARTAX <- CEARGroupMB[c(CEARGroupMB$Species == "Cerastium arvense"), FirstOTU:LastOTU]
DEMETAX <- DEMEGroupMB[c(DEMEGroupMB$Species == "Delphinium menziesii"), FirstOTU:LastOTU]
ERLATAX <- ERLAGroupMB[c(ERLAGroupMB$Species == "Eriophyllum lanatum"), FirstOTU:LastOTU]
ERSPTAX <- AbundanceMetadataMBWOBR[c(AbundanceMetadataMBWOBR$Species == "Erigeron speciosis"),
FirstOTU:LastOTU]
FEROTAX <- FEROGroupMB[c(FEROGroupMB$Species == "Festuca roemeri"), FirstOTU:LastOTU]
LOTRTAX <- LOTRGroupMB[c(LOTRGroupMB$Species == "Lomatium triternatum"), FirstOTU:LastOTU]
LOUTTAX <- LOUTGroupMB[c(LOUTGroupMB$Species == "Lomatum utriculatum"), FirstOTU:LastOTU]
LULETAX <- LULEGroupMB[c(LULEGroupMB$Species == "Lupinus lepidus"), FirstOTU:LastOTU]
POGRTAX <- POGRGroupMB[c(POGRGroupMB$Species == "Potentilla gracilius"), FirstOTU:LastOTU]
SYALTAX <- SYALGroupMB[c(SYALGroupMB$Species == "Symphoricarpos albus"), FirstOTU:LastOTU]

ACMITAXSUMS <- as.data.frame(colSums(ACMITAX))
AQFOTAXSUMS <- as.data.frame(colSums(AQFOTAX))
ASCUTAXSUMS <- as.data.frame(colSums(ASCUTAX))
BADETAXSUMS <- as.data.frame(colSums(BADETAX))
CALETAXSUMS <- as.data.frame(colSums(CALETAX))
CAQUTAXSUMS <- as.data.frame(colSums(CAQUTAX))
CEARTAXSUMS <- as.data.frame(colSums(CEARTAX))
DEMETAXSUMS <- as.data.frame(colSums(DEMETAX))
ERLATAXSUMS <- as.data.frame(colSums(ERLATAX))
ERSPTAXSUMS <-as.data.frame(colSums(ERSPTAX))
FEROTAXSUMS <- as.data.frame(colSums(FEROTAX))
LOTRTAXSUMS <- as.data.frame(colSums(LOTRTAX))
LOUTTAXSUMS <- as.data.frame(colSums(LOUTTAX))
LULETAXSUMS <- as.data.frame(colSums(LULETAX))
POGRTAXSUMS <- as.data.frame(colSums(POGRTAX))
SYALTAXSUMS <- as.data.frame(colSums(SYALTAX))

TAXSpeciesSumTable <-
t(cbind(ACMITAXSUMS,AQFOTAXSUMS,ASCUTAXSUMS,BADETAXSUMS,CALETAXSUMS,CAQUTAXSUMS,CEARTAXS
UMS,DEMETAXSUMS,ERLATAXSUMS,

ERSPTAXSUMS,FEROTAXSUMS,LOTRTAXSUMS,LOUTTAXSUMS,LULETAXSUMS,POGRTAXSUMS,SYALTAXSUMS))
rownames(TAXSpeciesSumTable) <- c("ACMI", "AQFO", "ASCU", "BADE", "CALE", "CAQU", "CEAR", "DEME",
"ERLA", "ERSP", "FERO", "LOTR", "LOUT", "LULE", "POGR", "SYAL")

unique(Taxa$Phylum)

#INDICATOR SPECIES ANALYSIS

library(indicspecies)
#Practice: based on species
#Use as.character to remove blanks from level
Species.ISA <- multipatt(x = AbundanceMBWOBR, cluster = as.character(AllDataABMBWOBR$Species), duleg =
TRUE)

summary(Species.ISA)
str(Species.ISA)
Species.ISA$sign$stat[Species.ISA$sign$stat == 1 & !is.na(Species.ISA$sign$stat)]
```