Environmental and biogeographical drivers of the *Leptospermum scoparium* (mānuka) phyllosphere microbiome

A thesis

submitted in partial fulfilment

of the requirements for the degree

of

Master of Science (Research)

at

The University of Waikato

by

Anya Sophia Noble

2018

THE UNIVERSITY OF
WAIKATO
Te Whare Wānanga o Waikato

Abstract

A substantial body of research exists on the physiology and genetics of *Leptospermum scoparium* (mānuka); however, the mānuka phyllosphere has not yet been investigated. This research, therefore, provides the first exploration of the bacterial communities comprising the mānuka phyllosphere. In total, 89 samples were collected from five native mānuka forests during the November 2016 – January 2017 mānuka flowering season, and examined alongside spatial, environmental, and host-tree related metadata. Cultivation-independent methods, including DNA extraction and 16S rRNA gene PCR amplicon sequencing analysis, were used to characterise phyllosphere communities. Using a combination of statistical tests, the BLAST algorithm, and a PICRUSt analysis, a habitat-specific core microbiome comprising members of *Acidobacteria*, *Alphaproteobacteria*, *Bacteroidetes*, and *Verrucomicrobia*, was present in all samples. Distinct microbial biogeographical patterning was also identified in the mānuka phyllosphere whereby spatially structured gradients in environmental variables, such as temperature and precipitation, correlated with bacterial community structure. The results from this study provide new insight pertaining to the structure, composition, and driving factors of the bacterial communities of the mānuka phyllosphere and shed light on theories of microbial ecology, such as microbial biogeography and core microbiomes. This research has the potential to open new avenues in both mānuka and microbial ecology fields of investigation within New Zealand.

Acknowledgements

I would like to thank my supervisor Dr Charles Lee for the opportunity to complete this research project at Waikato University.

In addition, I would like to thank Dr Michael Clearwater and Stevie Noe for arranging and managing sampling trips, as well as Steens Honey, Timberlands ltd, East Taupō Lands Trust, and Waipa District Council for providing sample sites.

I would also like to thank my amazing TRU-family; Shaun, Shelly, Georgia, Maria, Mafalda, and Hugo, for their support, inspiration, and the endless adventures and laughs that we shared both in and out of the lab.

Last, but not least, I would like to thank Liam and my family (my mum, my dad, and my brother) for their unconditional love, support, patience, and humour during this challenging but rewarding year.

Table of Contents

List of Figures

List of Tables

1. Chapter One:
Introduction

1.1. Introduction to the phyllosphere

Plants offer diverse habitats for microbial life on Earth whereby the surface of above- and below-ground plant structures, as well as internal plant tissues, are colonised by epiphytic and endophytic microorganisms, respectively (Vandenkoornhuyse *et al*, 2015, Vorholt 2012). Leaves are multi-layered plant organs characterised by photosynthetic mesophyll cells, upper and lower epidermal cells, and an outer waxy cuticle (Figure 1.1). Remarkably, leaves provide the largest known biological surface on Earth, representing a total surface area of 1.02×10^9 km^2, approximately twice the size of the Earth's surface (Woodward & Lomas 2004). The leaf surface offers a ubiquitous habitat for unique communities of epiphytic microorganisms and, collectively, the leaf surface microhabitat and its microbial colonisers are termed the 'phyllosphere' (Kim *et al*, 2012, Ruinen 1956). At the micron-scale, the phyllosphere is a complex and heterogenous environment with a mountainous topography. Bulges and troughs produced by epidermal cell arrangements, veins, trichomes and stomata, as well as the chemical composition of plant secretory compounds, affect the availability of this microhabitat for microbial colonisation (Baldotto & Olivares 2008, Mechaber *et al*, 1996, Remus-Emsermann *et al*, 2012). Moreover, differences in leaf size, morphology, longevity, as well as position and height in the tree canopy, creates large-scale heterogeneity whereby no two leaves are the same (Lindow & Brandl 2003). The phyllosphere is a hostile environment for microbiological life. The waxy leaf cuticle prevents water and solute loss from inner plant tissues which renders the leaf surface desiccated and nutrient-poor (Serrano *et al*, 2014). Furthermore, diurnal environmental variation and stochastic weather events drive rapid and extreme fluctuations of physicochemical factors on the leaf surface, such as ultraviolet (UV) radiation, humidity, temperature, and pH (Vorholt 2012).

The plant microbiome comprises microorganisms from all three primary domains of life: Bacteria, Archaea and Eucarya (Woese *et al*, 1990). Whilst a variety of epiphytic yeasts and fungi also inhabit the phyllosphere, epiphytic bacteria (often termed phylloepiphytic bacteria) are the most abundant and diverse coloniser, predicted

to colonise 6.4×10^8 km^2 of terrestrial leaf surface area at densities of up to 10^8 cells per cm^2 (Andrews & Harris 2000, Lindow & Brandl 2003, Redford *et al*, 2010). In accordance with these figures, the global bacterial population present in the phyllosphere has the potential to reach up to 10^{26} cells, likely representing a vast number of unique and undescribed species (Vorholt 2012). Owing to their numerical dominance, bacteria have been the primary focus of microbiological phyllosphere investigations to date and are therefore the primary focus of this thesis.

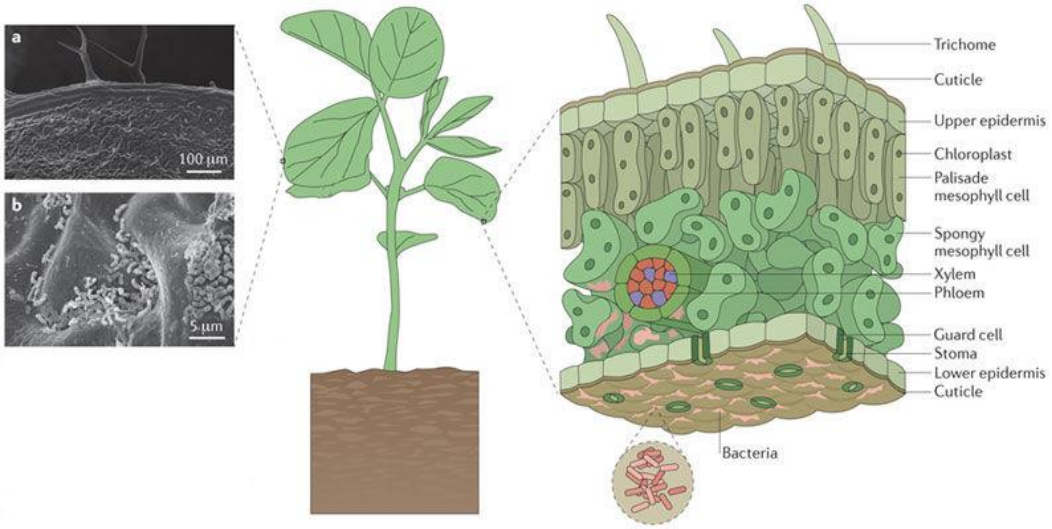

Figure 1.1: Schematic of an *Arabidopsis thaliana* leaf cross section displaying the characteristic epidermal and mesophyll cell layers. Scanning electron micrograph (A) shows a leaf edge with visible trichomes and (B) shows phylloepiphytic bacteria located in the epidermal cell grooves. Image sourced and modified from Vorholt (2012).

1.2. Early phyllosphere microbiology

Microbiological investigations of the phyllosphere date back to the mid-1900s and were traditionally established using cultivation-dependent methodology (Andrews *et al*, 1980, Ercolani 1991, Ruinen 1956, Ruinen 1961). Such cultivation-dependent methods involved enriching and isolating microorganisms on or in growth media whereby quantitative and qualitative assessments relied on plate counts and morphological or biochemical analyses, respectively (Wilson & Lindow 2000, Yang *et al*, 2001). However, a substantial proportion (97 – 99.9%) of environmental bacteria, such as those that occupy plant-associated habitats, are viable or metabolically active yet nonculturable in laboratory media (Amann *et al*, 1995, Oliver 1993, Xu *et al*, 1982). Furthermore, owing to the heterogeneity and complexity of the leaf microhabitat, culturing technologies have been unable to adequately replicate the natural environment in which phylloepiphytic bacteria normally thrive (Müller & Ruppel 2014). With this growing realisation, cultivation-independent methodology was introduced into phyllosphere research in 2001 (Andrews & Harris 2000, Wilson & Lindow 2000). This pioneering study immediately revealed bacterial community diversity on the leaf surface was far greater than previously depicted with cultivation-dependent methods (Yang *et al*, 2001). Large differences in microbial diversity and abundance have since been reported by studies that have comparatively utilised both cultivation-dependent and cultivation-independent techniques (Rastogi *et al*, 2010, Yashiro *et al*, 2011). However, the phyllosphere remains largely understudied and published literature is still predominantly comprised of outdated research which utilises cultivation-dependent methodology. For this reason, microbiological knowledge pertaining to the phyllosphere has historically lagged behind knowledge of other microbial habitats and, despite the enormity and potential significance of the phyllosphere, many knowledge gaps exist (Vorholt 2012). Rigorous cultivation-independent research in this field is therefore timely and only cultivation-independent research will be discussed for the remainder of this thesis.

1.3. Bacteria in the phyllosphere

Host-associated microbiomes are essential for the health, fitness, development and evolution of their hosts (Robinson *et al*, 2010). For this reason, the composition and function of such microbial communities can be collectively considered an extended

phenotype of the host organism (Richardson 2017). Accumulating evidence suggests phylloepiphytic bacteria are able to mediate the health and fitness of their host plant by promoting plant growth and development (Abanda-Nkpwatt *et al,* 2006, Gourion *et al,* 2006), modifying plant hormone production (Brandl & Lindow 1998), and increasing plant resistance to both abiotic and biotic stress via pathogen exclusion and transcriptional activation of plant defence genes (Innerebner *et al,* 2011, Kürkcüoglu *et al,* 2007, Ortega *et al,* 2016, Vogel *et al,* 2016). In addition, the phyllosphere of healthy host plants contains a greater diversity of phylloepiphytic bacteria compared to diseased plant hosts (Agler *et al,* 2016). Recent evidence proposes the phyllosphere microbiome may promote important ecological interactions by participating in many global biogeochemical processes, such as nitrogen fixation, nitrification, carbon cycles, and methanol degradation (Delmotte *et al,* 2009, Knief *et al,* 2012, Vorholt 2012). Such ecological interactions have the potential to influence plant biogeography as well as facilitate positive ecosystem function and productivity (Kembel *et al,* 2014, Laforest-Lapointe *et al,* 2017). Given their ubiquity and potential significance, phylloepiphytic bacteria have, therefore, been suggested to represent an ancient symbiosis with their host plant (Kembel *et al,* 2014, Vandenkoornhuyse *et al,* 2015). Thus far, investigations of this plant-bacterial symbiosis on the leaf surface have been primarily conducted under laboratory-controlled conditions, whereby individual host plant species are inoculated with individual bacterial strains (Innerebner *et al,* 2011). Consequently, the full significance of phylloepiphytic community function on host plants, terrestrial ecosystems, and global biogeochemical processes remains poorly understood. Owing to the importance of host microbiomes, the ability to model and manipulate host-microbiome interactions would be inherently beneficial. To achieve this, the structure, composition and driving factors that shape these host-associated microbial communities must be fully understood (Robinson *et al,* 2010). For example, the structure and composition of the mammalian gut microbiome is known to be primarily driven by ecologically important host attributes, such as diet (Groussin *et al,* 2017). A similar understanding of the drivers of variation in leaf-associated bacterial communities is lacking, however, and will require further study as well as the application of ecological principles (Costello *et al,* 2012, Kembel *et al,* 2014, Robinson *et al,* 2010).

1.4. Structure and composition of bacterial communities

Microbial colonisation patterns in the phyllosphere are non-uniform, likely reflecting the topographic heterogeneity of the leaf surface (Monier & Lindow 2004). Phylloepiphytic bacteria colonise discrete sites, such as epidermal cell wall surfaces and junctions, trichomes, veins, and stomata (Figure 1.1 & 1.2). These preferential colonisation areas offer bacteria increased nutrient availability as well as protection against abiotic and biotic stresses (Baldotto & Olivares 2008, Beattie & Lindow 1999). Within the phyllosphere, bacteria exist as solitary cells, multicellular aggregates and biofilms (Jacques *et al,* 2005). Biofilms are defined as large matrices of structured microbial cell communities and their extracellular polysaccharide (EPS) products (Costerton *et al*, 1995). In contrast, multicellular aggregates are dense patches of microbial cells observed in the absence of an EPS matrix. Ruinen (1961) provided the first description of a multi-layered epiphytic biofilm in the phyllosphere of a South American tropical plant using light microscopy. However, owing to the discontinuity of phylloepiphytic colonisation and the ephemeral nature of the phyllosphere habitat, discrepancy has arisen as to whether the leaf surface is, in fact, capable of supporting 'true' biofilm formation (Monier & Lindow 2004). Recent studies suggest biofilm formation is more likely on long-lived, evergreen leaves where free moisture availability is high (Jacques *et al*, 2005). That the leaf surface can support dense patches of localized microbial cell packing has several interesting implications for epiphytic survival. Dense cellular aggregates provide individual cells with enhanced protection from desiccation, UV, bacteriocides, and predation (Beattie & Lindow 1999). Bacterial cell aggregates also enhance genetic exchange through horizontal gene transfer and facilitate coordinated community behaviour through cell-cell signalling and quorum sensing (Morris & Monier 2003). For example, the accumulation of diffusible molecules, such as N-acyl homoserine lactones, effects numerous aspects of microbial behaviour in the phyllosphere, such as the production of EPS, antimicrobial compounds, pathogenicity traits, and the expression of genes involved in adhesion (Costerton *et al,* 1995, Morris *et al*, 1997). However, physiological properties, including those that facilitate the survival of epiphytes on the leaf surface, are not well understood and require further investigation.

The composition of the phyllosphere microbiome reflects immigration, survival and the growth of individual bacterial taxa on the leaf surface. These important community dynamics are presumably shaped by specific selective pressures pertaining to environmental conditions, leaf physicochemical properties, host-plant life history, and microbe-microbe interactions (Whipps *et al,* 2008). Recently, urban land use has also been found to influence the phyllosphere, suggesting that anthropomorphic changes are implicating these microbial communities (Smets *et al,* 2016). At the phylum level, phyllosphere bacterial communities across a range of host plants are largely composed of *Proteobacteria*, *Firmicutes*, *Bacteroidetes, Actinobacteria* and *Acidobacteria*. *Cyanobacteria* have also been reported on the leaf surface, albeit less abundant (Bulgarelli *et al,* 2013, Kim *et al,* 2012, Knief *et al,* 2012). Within these phyla, bacterial communities on tree leaves are diverse and species-rich, with single leaves hosting up to 800 bacterial taxa (Finkel *et al,* 2011, Kembel *et al,* 2014, Lambais *et al,* 2006). A proportion of bacteria found on the leaf surface are likely transient taxa, deposited from the surrounding environment via stochastic processes. However, it is becoming increasingly apparent that the phyllosphere microbiome also comprises a specialised consortium of bacterial taxa that are unique relative to the rest of the host plant and surrounding environment. Profound dissimilarities have been found when comparing phyllosphere communities to other plant-associated bacterial communities, such as in the rhizosphere (Bodenhausen *et al,* 2013, Vokou *et al,* 2012). Moreover, the leaf surface has been found to share very few bacterial taxa with the atmosphere and soil (Bowers *et al,* 2009, Copeland *et al,* 2015, Kim *et al,* 2012, Lauber *et al,* 2009).

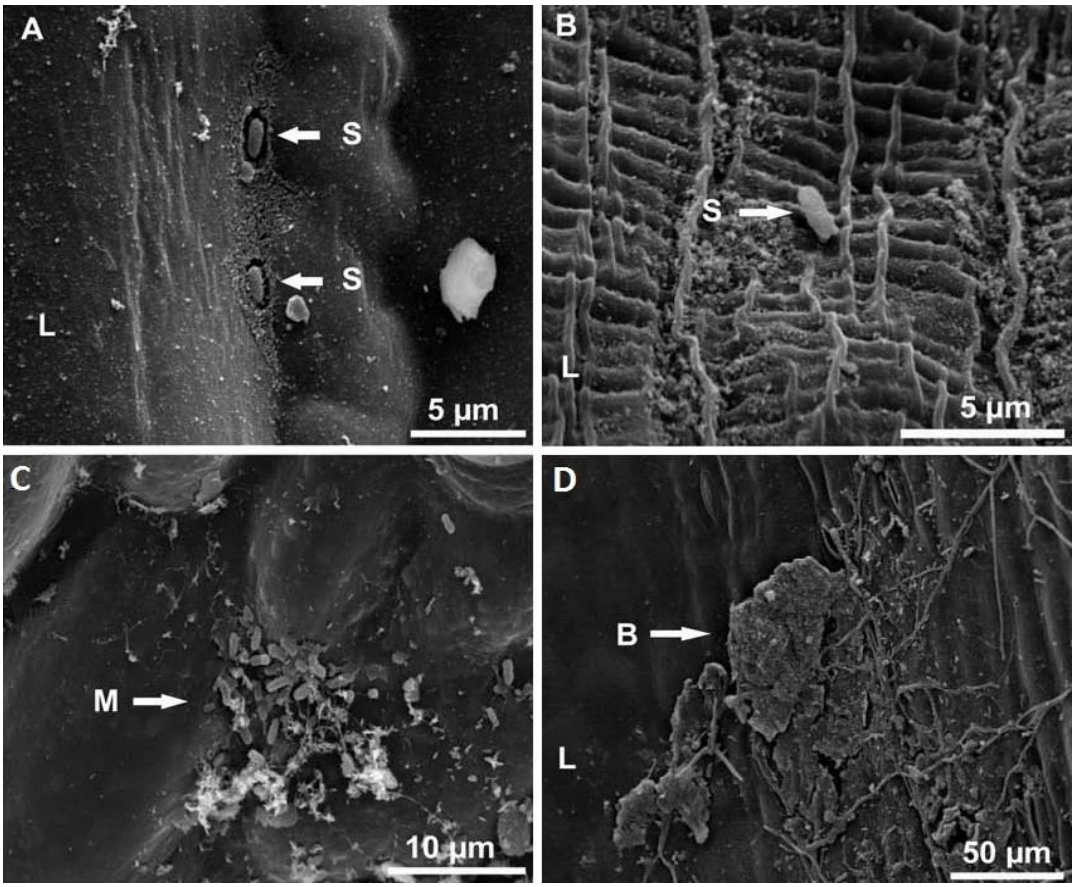

Figure 1.2: Scanning electron micrographs depicting the colonisation patterns of phylloepiphytic bacteria. Colonisation patterns include solitary cells (A & B), multicellular aggregates at epidermal cell wall junctions (C), and biofilms (D). Symbol representation: L, lower leaf surface; S, solitary; M, microcolony; B, biofilm. Image sourced and modified from Baldotto and Olivares (2008).

1.5. Microbial Biogeography

Biogeography is the study of the distribution of biodiversity over geographic space and time (Andrews & Harris 2000, Lomolino *et al*, 2010). Investigations pertaining to the geographic distribution of macroorganisms, such as plants and animals, date back to the 18th century. In contrast, microbial biogeography has only been recently explored and can be contentious. Since the early 20th century an old microbiological tenet, *"everything is everywhere, but, the environment selects"*, posited the view that, unlike macroorganisms, microorganisms do not possess dispersal limitations, become geographically isolated, nor undergo allopatric speciation (De Wit & Bouvier 2006). These views primarily originated from cultivation-dependent investigations of microbial communities across essentially continuous habitats, such as air, soil and large bodies of water (Glöckner *et al*, 2000, Massana *et al*, 2000). That microorganisms demonstrate an advanced ability to disperse, produce spores, and rapidly adapt to their environment also provided support in favour of these old views (Finlay 2002). However, examination of a wider range of environments, alongside the advent of new analytical methods and higher resolution genetic markers, ignited a recent resurgence of interest in microbial biogeography and provided the first validation of biogeographical patterning in microorganisms (Papke *et al*, 2003, Whitaker *et al*, 2003). Current research in microbial biogeography continues to explore whether microbial assemblages differ in different locations, the relative influence of historical contingencies and contemporary environmental factors on the spatial and temporal variation of these microbial assemblages, and the extent to which microbial biogeography resembles that of macroorganisms (Martiny *et al*, 2006).

The phyllosphere provides researchers with the unique opportunity to sample bacterial communities in a discrete and hierarchical manner, such as within and between individual leaves, trees, and host plant species across time and space (Redford *et al*, 2010). The phyllosphere, therefore, represents a fascinating environment to test ecological principles, such as microbial biogeography and community assembly theories (Kinkel *et al*, 1995, Meyer & Leveau 2012, Redford *et al*, 2010). Although phyllosphere microbial biogeography has been investigated by few studies, significant temporal and spatial variation has been reported. Moreover, recent research has sought to identify and quantify the relative importance of temporal, dispersal-related, and

niche-based drivers of microbial community dynamics on the leaf surface. This will be reviewed in more detail in the following sections.

1.6. Temporal variation in the phyllosphere microbiome

One basic question rarely addressed by phyllosphere research is whether there are seasonal or annual changes in community structure. The first cultivation-independent research to investigate short-term change in phyllosphere community size and composition was conducted in 2006. This study examined phyllosphere community structure on fronds of the resurrection fern *Polypodium polypodioides* following exposure to rainfall (Jackson *et al*, 2006). 16S rRNA gene PCR amplicon sequencing revealed a reduction in diversity following wetting. Sequences obtained prior to rainfall were affiliated with nine major lineages of bacteria whereas sequences obtained after rainfall were affiliated with only four lineages. Most sequences were taxonomically assigned to the alpha class of the phylum *Proteobacteria*, and as well as simplifying the overall community, wetting increased the dominance of this phylum group from 55 to 75%. This observation is analogous to enrichment, whereby the increased availability of soluble organic substrates selectively facilitates the growth of specific bacterial taxa.

Early cultivation-dependent investigations identified predictable changes in phyllosphere community composition depending on when in the growing season leaves were sampled (Ercolani 1991, Thompson *et al*, 1993). However, it was not until 2009 that the first cultivation-independent investigation of successional change in the phyllosphere was performed. Redford and Fierer (2009) explored bacterial community assembly patterns on the eastern cottonwood, *Populus deltoids*, across the deciduous growing season from leaf emergence to leaf fall. Intriguingly, the temporal variability of phyllosphere community composition within a single tree exceeded the variability between trees on any given day, independent of weather events and seasonal climatic differences. 16S rRNA gene PCR amplicon sequencing indicated early- and late-season communities were dominated by *Proteobacteria* whilst mid-season communities were dominated by *Acidobacteria*. This clustering pattern appeared predictable from year to year with the turnover of bacterial taxa resembling the species-time relationship of eukaryotic communities. Similar seasonal community changes have also been reported in the evergreen magnolia (*Magnolia grandiflora*) (Jackson & Denney 2011). 16S

rRNA gene sequencing identified *Proteobacteria* as the dominant phyla in the magnolia phyllosphere with distinct increases of *Acidobacteria* and *Actinobacteria*. Interestingly, the *Methylobacteriaceae* family was atypically abundant in the August phyllosphere while the *Beijerinckiaceae* family was unusually reduced. *Methylobacteriaceae* and *Beijerinckiaceae* are members of *Alphaproteobacteria*, and show contrasting capacities to grow at high temperatures (Sneath *et al*, 1986). The observed shift in dominant phylloepiphytic bacterial populations may have, therefore, reflected high summer temperatures prior to the August sampling. Jackson and Denney (2011) also collected samples yearly in order to explore annual patterns in magnolia phyllosphere community composition. No consistent annual patterns were found, however. Plants of different ages have also been suggested to harbour different microbial communities on their leaf surfaces. Examinations at the cellular level have revealed the leaf surface undergoes significant chemical, physiological, and topographic alterations over time as leaves mature (Kenzo *et al*, 2015, Reich *et al*, 1991, Steppe *et al*, 2011). Age-dependent changes on the leaf surface may drive long-term temporal compositional changes in the phyllosphere (Mechaber *et al*, 1996).

In summary, short-term fluctuations in phyllosphere community composition may be driven by stochastic weather events, such as rainfall and high winds, that directly influence selective growth and the immigration of bacteria from the surrounding environment to the leaf surface (Andrews & Harris 2000). Long-term temporal changes may meanwhile reflect seasonal climatic differences, microbial succession, and aging plant hosts (Jackson & Denney 2011, Redford & Fierer 2009). Temporal variability in phyllosphere communities may highlight the need to consider the time of sampling and tree age as a confounding variable in future microbial ecology investigations on the leaf surface. However, since current knowledge about temporal change in the phyllosphere stems primarily from two cultivation-independent phyllosphere investigations, no generalisable conclusions can be drawn (Table 1.1).

1.7. Dispersal-related drivers in the phyllosphere microbiome

One of the fundamental questions pertaining to microbial biogeography is whether microbial assemblages differ in different locations (Martiny *et al*, 2006). To address this question, several research groups have examined phyllosphere communities across a range of regional to continental spatial scales (Finkel *et al*, 2011,

Redford *et al*, 2010, Stone & Jackson 2016). Finkel *et al*, (2011) investigated the relative influence of geographic location and host tree species on the microbial community structure of three salt-secreting *Tamarix* desert tree species in the Mediterranean and Dead Sea regions in Israel as well as the United States. These authors report bacterial species richness and diversity correlated with geographical location; different *Tamarix* species grown in the same locale hosted similar communities whereas the same *Tamarix* tree species in different geographical regions hosted distinct phylloepiphytic communities. For example, *Tamarix* species in the Mediterranean region were dominated by *Halomonas* and *Halobacteria*, whereas trees from the Dead Sea region were dominated by *Actinomycetales* and *Bacillales*.

Controlling for the effect of abiotic and biotic environmental heterogeneity is important when attempting to elucidate the relative influence of historical contingencies, such as dispersal limitation, and contemporary environmental conditions on microbial biogeographical patterns (Martiny *et al*, 2006, Wiens 2011). Moreover, this has not been thoroughly implemented in phyllosphere research (Knief *et al*, 2010). One study that did control for environmental variation was Finkel *et al*, (2012). These authors measured the phyllosphere bacterial communities of *T. aphylla* trees that were spatially dispersed across the Sonoran Desert in the Southwestern United States. The climatic conditions of the Sonoran Desert are relatively uniform and the authors described the spread of *Tamarix* trees as analogous to an archipelago of islands. Interestingly, Finkel *et al*, (2012) identified a positive correlation between distance and bacterial community dissimilarity within the beta class of the phylum *Proteobacteria*. More recently, evidence of a distance-decay relationship has also been reported in the magnolia (*M. grandiflora)* phyllosphere at local scales (<500m) (Stone & Jackson 2016). The decay of community similarity with distance is a prediction of neutral ecological theories in studies of macroorganisms, which asserts that biodiversity arises at random through the interplay of immigration/emigration and speciation/extinction (Rosindell *et al*, 2011, Soininen *et al*, 2007, Wiens 2011). This evidence of a significant decay-distance relationship provided the first insight of the potential importance of dispersal histories and limitations in phyllosphere bacterial community assembly.

In summary, phylogenetically distinct phyllosphere communities have been identified from the same host tree species in different geographical locations (Finkel *et*

al, 2011). These findings highlight the potential of the phyllosphere to exhibit microbial biogeography. However, since only two investigations offer current evidence in support of a distance-decay relationship in the phyllosphere, further investigation is necessary (Table 1.1) (Finkel *et al,* 2012, Stone & Jackson 2016).

1.8. Niche-based drivers in the phyllosphere microbiome

Host plant genotype, irrespective of geographic location or environmental selective pressures, has also been identified as an important driver of bacterial community composition in the phyllosphere (Whipps *et al,* 2008). Contrasting with investigations that report geographical community patterning, different plant species have been reported to host distinctive assemblages of bacteria that remain constant across different geographical scales (Lambais *et al,* 2006). Redford *et al,* (2010) quantified the intra- and inter-individual variability of phyllosphere communities across 56 tree species in Boulder, Colorado. Interspecific variation between different host plant species significantly exceeded variation within and across individual plants of the same species. Furthermore, the degree of similarity between different phyllosphere communities was found to reflect host tree phylogenetic relatedness. In general, *Bacteroidetes* and *Betaproteobacteria* were more common on gymnosperms, while *Actinobacteria* and *Gammaproteobacteria* were more common on angiosperms. These authors also examined phyllosphere communities on *Pinus ponderosa* needles across the globe, testing the null hypothesis central to neutral theories of ecology that plants in close proximity harbour more similar microbial communities than those geographically separated. Unlike Finkel *et al,* (2011), however, minimal geographic differentiation was found.

A relationship between phylloepiphytic bacterial community phylogeny and host tree phylogeny has been reported by a number of other studies since Redford *et al,* (2010). This relationship between phyllosphere and host tree phylogeny may reflect phylogenetic variation in ecologically important plant functional traits, which influence the colonisation and survival of leaf-associated bacterial communities (Kembel *et al,* 2014, Kim *et al,* 2012). Recent research efforts have, therefore, sought to investigate the relationship between bacterial community structure and important functional traits of the host tree. Kembel *et al,* (2014) quantified the bacterial community structure on the leaves of 57 tree species in a Panama neotropical forest and found the majority

(51%) of variance in microbial community structure was explained by host plant taxonomy as well as ecological strategies for resource uptake and mortality-growth trade-offs. For example, the relative abundance of *Betaproteobacteria* and *Elusimicrobia* was significantly correlated with plant traits linked to resource uptake and retention, such as leaf thickness and mass per area. Meanwhile, the relative abundance of *Alphaproteobacteria* and *Thermomicrobia* was significantly associated with plant traits such as wood density and rate of growth. The results of this study were the first to provide evidence in support of the arising suspicion that the host plant's ecological strategy is an important driving factor of phyllosphere bacterial communities, likely through impacting leaf morphology and physiology. Using a 'host-parasite association test', Kembel *et al,* (2014) also report a significant overall evolutionary association between host plant species and phyllosphere bacterial taxa. In contrast to the predictions made by previous studies, such as Finkel *et al,* (2012), evidence for an evolutionary association between host plant and phyllosphere taxa provides support in favour of niche-based community assembly theories and further suggests that conditions on the leaves act as strong ecological filters on phyllosphere communities (Finkel *et al,* 2012, Kembel *et al,* 2014).

While many previous phyllosphere investigations have evaluated drivers of phyllosphere community structure one at a time, Laforest-Lapointe *et al,* (2016a) were the first to evaluate the relative influence of multiple driving factors simultaneously. These authors examined host species identity (phylogeny, functional traits, coevolution), geographical location (dispersal history, abiotic conditions), and time of sampling, on phyllosphere bacterial communities of five temperate tree species in Quebec. Interestingly, the variance in phyllosphere community structure explained by climatic differences and sampling time was small relative to host species, with each tree species harbouring a distinct community of phyllosphere epiphytes as observed in previous studies (Kembel *et al,* 2014, Lambais *et al,* 2006, Redford *et al,* 2010). Moreover, in accordance with the findings of Kembel *et al,* (2014), the epiphytic bacterial community structure in the temperate phyllosphere also correlated with plant traits linked to resource uptake and growth/mortality trade-off ecological strategies (Laforest-Lapointe *et al,* 2016a).

In summary, recent investigations across different plant species have observed distinct microbial assemblages in the phyllosphere associate with host plant genotype

(Redford *et al,* 2010). The phyllosphere of different host plant species may represent different ecological niches for bacteria due to divergent ecological strategies and phylogenetic differences in leaf physiology. However, the relationship between phyllosphere community structure and host tree functional traits has only been investigated on two occasions (Table 1.1) (Kembel *et al,* 2014, Laforest-Lapointe *et al,* 2016a). Therefore, whether plant species have evolved to actively recruit or filter for specific bacterial communities in the phyllosphere remains to be answered.

1.9. The core microbiome concept

A 'core microbiome' is defined as the taxa shared by all microbial assemblages associated with a specific habitat. These microbiome cores are hypothesised to represent the functionally critical members of a microbial community and have captured increasing interest within the field of microbial ecology (Shade & Handelsman 2012). Characterising core taxa may provide the first step towards identifying a 'healthy' community and predicting the response of microbial assemblages to environmental change. The core microbiome concept can be extended to an array of microbial assemblages, ranging from free-living communities to complex host-associated microbiomes. For example, identifying a core microbiome of the human gut microbiota has been the goal of the US National Institutes of Health Microbiome Project for several years (Cho & Blaser 2012). Interestingly, a core microbiome has been detected in the human gut, and deviations from this core have been associated with different physiological states (Turnbaugh *et al,* 2008). Core microbiomes have also been detected in coral, the evolutionarily ancient animals of the *Cnidaria* phylum (Reveillaud *et al,* 2014). Analyses of plant core microbiota are still in their infancy, although an initial attempt to define the core root microbiome of *A. thaliana* was published in 2012 (Lundberg *et al,* 2012). Moreover, recent studies have successfully identified a core microbiome in the phyllosphere. For example, Kembel *et al,* (2014) identified 104 core taxa on 95% of all trees sampled within a neotropical forest. This neotropical core microbiome represented 73% of all sequences and consisted of members of *Alphaproteobacteria*, *Sphingobacteria*, *Gammaproteobacteria*, *Betaproteobacteria,* and *Actinobacteria*. Laforest-Lapointe *et al,* (2016a) meanwhile identified 19 core taxa present in more than 99% of all trees sampled in a temperate forest. This temperate core microbiome represented 42.7% of

all sequences and was comprised of members of *Alphaproteobacteria*, *Acidobacteria*, and *Betaproteobacteria*. The core microbiome in the phyllosphere of a single host plant species is yet to be investigated; moreover, the function of a core microbiome in the phyllosphere is unknown. The core microbiome concept therefore provides an interesting foundation for future research in all environments, including the phyllosphere.

1.10. Summary of recent cultivation-independent phyllosphere research

Comparatively few cultivation-independent investigations of the phyllosphere microbiome have been conducted (Table 1.1). Therefore, the driving factors of community composition and diversity in the phyllosphere remain poorly understood. Bacterial communities on the leaf surface are complex and perhaps non-uniformly influenced by dispersal history, host selection, and extreme environmental fluctuations (Stone & Jackson 2016). Conflicting evidence across different studies may reflect the susceptibility of different bacterial taxa to be influenced by different selection pressures as well as methodological inconsistencies and pitfalls (which will be reviewed in Chapter 2). Furthermore, many fundamental questions regarding the spatial and temporal variability of phyllosphere communities have yet to be addressed. Fully disentangling the relative importance of driving factors on bacterial community structure and composition in the phyllosphere will require more cultivation-independent research and experimental manipulations of plant-microbe associations.

Table 1.1 Summary of cultivation-dependent phyllosphere research

Investigation aims	Objects	Main conclusions	Reference
Bacterial community successional patterns	Cottonwood (*Populus deltoids*), Colorado, USA.	Temporal variability is not random, phyllosphere bacterial communities follow predictable seasonal patterns.	Redford and Fierer (2009)
Influence of host tree species and geographic location on phyllosphere bacterial community composition	56 tree species, needles of *Pinus ponderosa* from various locations around the world.	Community composition reflected host tree phylogenetic relatedness while geographic differences were minimal.	Redford *et al*, (2010)
Spatial variability of the microbial phyllosphere	Three *Tamarix* tree species in the Mediterranean and Dead Sea regions as well as two locations in the USA.	Phyllosphere microbial assemblages of the same host tree species differ in different geographical locations.	Finkel *et al*, (2011)
Phyllosphere bacterial community annual and seasonal patterns	Single *Magnolia grandiflora* tree, Mississippi, USA.	Distinct seasonal patterns of bacterial communities. No predictable annual patterns.	Jackson and Denney (2011)
Role of geographical distance in phyllosphere community assembly	*Tamarix* tree species in the Sonoran Desert, USA.	Community similarity declines with geographic distance.	Finkel *et al*, (2012)
Phyllosphere composition and diversity in tropical trees	Six tropical tree species, Malaysia.	Tropical trees harbour distinctive phyllosphere bacterial communities similar to temperate and subtropical trees.	Kim *et al*, (2012)

Relationship between bacterial communities on tree leaves and the functional traits, taxonomy, and phylogeny of their plant hosts	57 tree species in a neotropical forest, Panama.	Core microbiome dominated phyllosphere communities. Community structure correlated with host evolutionary relatedness and suites of plant functional traits.	Kembel *et al*, (2014)
Phyllosphere communities of temperate forests with respect to host functional traits, taxonomy and phylogeny	Five temperate forest tree species, Quebec, Canada.	Host species identity is a stronger driver of temperate tree phyllosphere bacterial communities than site or time. Core microbiome detected.	Laforest-Lapointe *et al*, (2016a)
Intra-individual and inter-individual variation in phyllosphere communities	Five temperate forest tree species, Quebec, Canada.	Considerable intra-individual variation within a tree canopy. Host species is the strongest driver of inter-individual variation. Core microbiome detected.	Laforest-Lapointe *et al*, (2016b)
Phyllosphere community distance-decay relationship at local scales	91 *Magnolia grandiflora* trees, Mississippi, USA.	Phyllosphere communities closer together had more similar bacterial communities	Stone and Jackson (2016)
Relationship between phyllosphere biodiversity and ecosystem function	19 temperate tree species, Quebec, Canada.	A positive correlation between plant-associated microbial diversity and terrestrial ecosystem productivity.	Laforest-Lapointe *et al*, (2017)

1.11. Mānuka (*Leptospermum scoparium*)

Mānuka (*Leptospermum scoparium*) is a flowering, evergreen member of the Myrtaceae (myrtle) family and native to New Zealand (Figure 1.3 – 1.4). Mānuka is fast growing and exhibits excellent establishment in zones of impaired soil quality and drainage. Furthermore, mānuka plays an ecologically important role in erosion control, carbon sequestration and the restoration of vegetation (Prosser 2011). Mānuka honey, produced from mānuka flower nectar by the introduced honeybee (*Apis mellifera*), is renowned for its unique antifungal and antibacterial properties and is expected to reach an annual sales of $1.2 billion by 2028 (Cooper *et al,* 2002, Stephens *et al,* 2005, Sykes *et al,* 1988). Methylglyoxal (MGO) is a unique antibacterial compound found in mānuka honey and directly relates to the therapeutic bioactivity of mānuka honey. The amount of MGO in mānuka is therefore used to infer honey quality through the development of the 'unique mānuka factor' (UMF) rating (Hamilton *et al,* 2013). MGO forms in mānuka honey through the chemical transformation of dihydroxyacetone (DHA), which is present in the nectar of mānuka flowers (Atrott *et al,* 2012). The level of DHA found in the nectar of mānuka flowers varies significantly, however. Variable DHA levels drive fluctuations in MGO and thus mānuka honey UMF ratings. Furthermore, the reason for this variability is not yet well understood (Atrott *et al,* 2012, Mavric *et al,* 2008). A substantial body of research exists on the physiology and genetics of the mānuka plant; however, attempts to elucidate the driving factors of honey quality and forest productivity have so far proven unsuccessful. Despite the emerging ecological significance of such epiphytic communities, the phyllosphere of many New Zealand natives, including *Leptospermum scoparium* (mānuka), have not yet been investigated from a microbiological perspective.

Figure 1.3: Mānuka (*Leptospermum scoparium*) plant (top) and flowers (bottom). Sourced from www.cabi.org/isc/datasheet/30097. ©Trevor James/Hamilton, New Zealand

1.12. Phyllosphere research in New Zealand

New Zealand is a geographically isolated country abundant in unique flora such as pōhutukawa, totara and mānuka (Atrott *et al,* 2012). Whilst phyllosphere microbiology outside New Zealand has progressed over the last two decades, the scope of phyllosphere research within New Zealand remains incredibly primitive. Of particular note was a Master's thesis published through the University of Waikato investigating the diversity of *Methylobacterium* species associated with the phyllosphere of native plants (Jahan 2013). This research project, however, utilised cultivation-dependent methodologies and did not examine the composition of the entire phyllosphere community. To date, mānuka has featured in 781 journal publications, with a large portion of studies focussing on the bioactive properties and compounds of mānuka honey. Two studies in 1998 and 2016 investigated endophytic bacterial communities inside mānuka leaves; however, no study has yet investigated the epiphytic bacterial communities on the mānuka leaf surface (Johnston 1998, Wicaksono *et al,* 2016). Furthermore, many important questions pertaining to phyllosphere microbiology remain to be clarified, such as: what is the most important driver of microbial community structure? – and what is the impact of the phyllosphere on economically important host plants? Undoubtedly, therefore, rigorous studies directed at answering these questions are necessary to improve upon our current knowledge of microbial phyllosphere ecology for the future. Future research in phyllosphere microbiology can be set out to encompass four overarching objectives: (i) to identify the taxonomy and biodiversity of phyllosphere communities, (ii) classify the drivers of community assembly on the leaf surface, (iii) link microbial diversity with the function of individual bacteria and communities, and (iv) elucidate the impact of phylloepiphytic communities on their host plants, global biogeochemical processes, and terrestrial ecosystems.

1.13. Research aims

The primary objective of this Master's research was to provide the first exploration and characterisation of the mānuka phyllosphere microbiome. Specifically, the aims of this research are to investigate the environmental and biogeographical driving factors of the phyllosphere bacterial community composition. In addition, particular interest is directed towards exploring whether a core microbiome is present

in the mānuka phyllosphere. The milestones of this Master's research include; firstly, the review and development of a robust methodology to isolate phylloepiphytic bacteria off the surface of mānuka leaves (Chapter 2); secondly, the sampling of native mānuka forests and implementation of developed methodology (Chapter 3); and lastly, diversity and phylogenetic analysis of the mānuka phyllosphere microbial community (Chapter 4).

2. Chapter Two:
Review of molecular techniques

2.1. Introduction

Understanding the extent of microbial diversity in any environment is challenging. A host-associated community from the surface of a human tooth was one of the first microbial communities to be discovered. However, despite significant advancements in molecular techniques, understanding the composition and function of host-associated microbial assemblages continues to challenge microbiologists more than 340 years later (Robinson *et al,* 2010). With the growing realisation that phyllosphere microorganisms are mostly nonculturable, cultivation-dependent techniques have been steadily replaced by cultivation-independent analyses (Amann *et al,* 1995, Müller & Ruppel 2014). Consequently, the description of microbial community structure and the quantification of microbial diversity associated with leaf surfaces has improved. However, methodological differences make it challenging to draw generalisable comparisons from recent cultivation-independent research (Müller & Ruppel 2014). Owing to the novelty of the current research project and diversity of methodological approaches utilised by previous phyllosphere investigations, the establishment of a reliable method to isolate bacteria from the mānuka leaf surface was required. The current chapter will review the challenges associated with studying phylloepiphytic bacterial communities and the methodology employed by recent cultivation-independent phyllosphere research (Table 2.1). Throughout this review, favourable techniques were identified and implemented in preliminary trials using mānuka leaves collected locally from Waikato University campus. Successful techniques were thereafter integrated into the final methodology of the current research project.

2.2. 16S rRNA

16S ribosomal ribonucleic acid (rRNA) is a component of the prokaryotic 30S small ribosomal subunit and essential for protein synthesis in all living organisms (Woese *et al,* 1990). The 16S rRNA gene is approximately 1,550 base pairs long and comprises an alternating pattern of nine hypervariable and conserved regions (Petrosino *et al,* 2009). The 16S rRNA gene is present in all bacteria, and 16S rRNA

gene sequence and function are unaffected by environmental selective pressures and conserved across organisms (Woese & Fox 1977). Random sequence changes in the variable regions of the 16S rRNA gene occur as a function of time across bacterial evolution and are therefore considered a robust genetic marker of bacterial phylogenetic relatedness (Janda & Abbott 2007). The use of the 16S rRNA gene sequence as a molecular marker of bacterial phylogeny and taxonomy was pioneered in 1970 by Carl Woese (Woese *et al,* 1990). Universal polymerase chain reaction (PCR) primers bind conserved regions and amplify variable regions of the 16S rRNA gene. Sequence analysis of these variable 16S rRNA gene amplicons permits the identification of different bacterial taxa (Winker & Woese 1991). The relative ease of 16S rRNA gene sequencing compared to previous cumbersome manipulations, such as DNA-DNA hybridization, alongside the development of high throughput sequencing technologies has exponentially increased the number of recognised taxa and advanced our understanding of microbial ecology (Claesson *et al,* 2010, Laforest-Lapointe *et al,* 2017, Woo *et al,* 2008). Moreover, there has been increased use of next-generation sequencing (NGS) methods to describe phyllosphere bacterial communities (Barriuso *et al,* 2011, Petrosino *et al,* 2009, Stone & Jackson 2016).

Although cultivation-independent methods have the advantage of assessing a broader spectrum of microbial taxa, they too are subject to several limitations that can impact the reliability of research data (Janda & Abbott 2007). One disadvantage of such cultivation-independent methodologies is the inability to discriminate between live, dormant, or dead microorganisms as any DNA-based sequence analysis detects OTUs that may be inactive, deceased, or transient (Degefu *et al,* 2009, Shade & Handelsman 2012). In addition, differences in 16S rRNA gene copy numbers may lead to the overrepresentation of some taxa (Kembel *et al,* 2012, Větrovský & Baldrian 2013). For this reason, DNA-based sequence analysis may also provide results that are not entirely reflective of the true community of interacting species within a locality. However, owing to the benefits of cultivation-independent techniques, 16S rRNA gene amplicon sequencing was the chosen method for this study.

2.3. The chloroplast contamination conundrum

Reliable and accurate separation of microorganisms from the leaf surface poses numerous challenges. One of the fundamental challenges presented to those who wish

to study the phylloepiphytic microbial communities is ensuring maximum isolation of bacteria from the leaf surface whilst minimising plant organelle contamination (Rastogi *et al*, 2010). Exclusive amplification and sequencing of the bacterial 16S rRNA gene in plant microbiome research is difficult as chloroplasts are of cyanobacterial origin and exhibit striking homology with bacterial 16S rRNA genes (Dams *et al*, 1988). To minimise chloroplast contamination, physical damage to the leaf and carryover of plant debris must be prevented when isolating phyllosphere bacteria; minimising leaf damage is also important in order to prevent inclusion of endophyte bacteria. In some circumstances, phyllosphere investigations have utilised severe methods to isolate bacteria from leaf surfaces, such as cutting and macerating leaves. Maceration has been found to increase DNA yield by approximately 70% (O'brien & Lindow 1989). Although this may be due to increased availability of strongly-adhering phylloepiphytic bacteria, undoubtedly it is also the result of increased chloroplast and endophyte DNA contamination. As an alternative to maceration, scientists have physically swabbed the leaf surface. Whilst this approach is gentler than maceration, such physical manipulation will also cause a large amount of chloroplast contamination; therefore, more conservative approaches are essential (Müller & Ruppel 2014). The necessity of using chloroplast-excluding PCR primers to prevent plant organelle contamination has been stressed by several publications (Redford *et al*, 2010, Redford & Fierer 2009). However, the use of such primers also results in the exclusion of *Cyanobacteria*. Metagenomic studies have reported *Cyanobacteria* are rare colonisers of the vascular plant phyllosphere (Delmotte *et al*, 2009). This statement is challenged, however, by other research groups that describe *Cyanobacteria* as principal nitrogen-fixing epiphytes on the leaf surface that require further investigation (Fürnkranz *et al*, 2008, Rigonato *et al*, 2012, Rigonato *et al*, 2016). Therefore, the utilisation of chloroplast-excluding primers inevitably projects taxonomic bias in the bacterial community composition analysis (Laforest-Lapointe *et al*, 2016a). Moreover, the use of different primers may explain the differences in clade abundances among studies. To minimise taxonomic bias and permit the detection of *Cyanobacteria*, universal bacterial 16S rRNA primers of the extended v4-v5 region were chosen for the current project.

2.4. Isolating phylloepiphytic bacteria

Plant-associated microbial habitats, such as the phyllosphere, capture interest from both plant and microbial ecologists. As a result, a number of cultivation-independent phyllosphere studies use techniques that are neither microbiologically appropriate nor reproducible. Such techniques include the manual swirling of leaves in solution with 'gloved fingers', brushing the leaf surface with 'sterile toothbrushes', and filtration through 'cheese cloths' (Kadivar & Stapleton 2003, Rastogi *et al,* 2010, Stone & Jackson 2016). Moreover, the definition of the phyllosphere is often misconstrued. For example, a recent communication by Copeland *et al,* (2015) is titled *"Seasonal community succession of the phyllosphere microbiome"*, these authors, however, sampled leaf cores and collectively recovered both endophytic and epiphytic bacterial communities. This does not provide an accurate reflection of the phyllosphere microbiome and, therefore, careful attention needs to be paid to the methodologies and experimental design implemented by researchers.

2.4.1. Sonication

Sonication is the act of applying sound energy vibrations to agitate particles (Suslick 1990). The use of sonication in an ultrasonic cleaning bath is perhaps the most conservative, reproducible, and successful method of isolating phylloepiphytic bacteria from the leaf surface implemented in phyllosphere research to date (Müller & Ruppel 2014, Yang *et al,* 2001). In such cases, leaves are submerged in a leaf wash buffer within an ultrasonic cleaning bath and bacteria are dislodged from the leaf surface and into the wash buffer via ultrasonic waves (Suda *et al,* 2008). Isolating microorganisms from the leaf surface via sonication in an ultrasonic cleaning bath was the method of choice for the current research project, owing to its gentle manipulation of plant tissue and thus minimal contamination of plant-organelles. However, studies that have utilised this technique have used various wash buffer constituents, leaf mass (g) to wash buffer volume (mL) ratio, and different ultrasonic frequencies (Kinkel *et al,* 1995) (Table 2.1). The use of a gentle detergent in the wash buffer has been documented with success on a number of occasions (Yashiro *et al,* 2011). Therefore, phosphate buffer solution (PBS) and Tween 20 were the chosen constituents of the leaf wash buffer for this study. Different concentrations of Tween 20 were trialled, ranging from 0.5 – 5%. Whilst 0.5% Tween 20 was too weak to implement microbial cell isolation, 5% was

too strong and resulted in large bubble formation during downstream DNA extraction steps. The ideal Tween 20 concentration was 1% and this was implemented as part of the final methodology. The efficacy of the sonication process was thereafter trialled in plastic Falcon tubes and glass flasks. Plastic is known to absorb ultrasonic waves and, accordingly, leaves sonicated in plastic Falcon tubes did not yield sufficient bacterial DNA quantities. In comparison, glass is more capable of transmitting ultrasonic waves and glass flasks consistently enabled a larger quantity of DNA to be extracted (Capelo-Martínez 2009). The larger surface area of the glass flasks may have also been another factor permitting better sonication compared to the plastic falcon tubes. In addition, the mass of mānuka leaves required to yield sufficient DNA was trialled, ranging from 0.5 – 2.0 g. The amount of DNA recovered increased linearly with sample mass. Due to the small size of mānuka leaves, 1.0 g of mānuka leaves were chosen as the ideal sampling weight, offering sufficient DNA while compromising with practicality and time restraints for aseptic removal of leaves from branches. Lastly, the volume of wash buffer was trialled, ranging from 5 – 20mL per 1.0 g of mānuka leaves. For 1.0 g of mānuka leaf sample mass, 10 mL of leaf wash buffer was found sufficient for completely submerging all the leaves and yielding sufficient DNA concentrations. This was the same leaf-to-buffer ratio as used by Finkel *et al,* (2011) and Bodenhausen *et al,* (2013) (Table 2.1).

2.4.2. Centrifugation versus filtration

To date, two methods have been used to successfully and reliably isolate DNA from the leaf wash buffer solution following sonication in preparation of downstream DNA extraction protocols (Suda *et al,* 2008). With one method, bacterial cells are pelleted via centrifugation and thereafter resuspended for downstream DNA extraction and PCR protocols (Suda *et al,* 2008, Yang *et al,* 2001). With the other method, bacterial cells from the leaf wash buffer are collected on a 0.2 µm filter and DNA extraction is performed directly on the filtration apparatus (Table 2.1) (Bodenhausen *et al,* 2013, Finkel *et al,* 2011, Ruiz-Pérez *et al,* 2016). Moreover, within these two methodological designs, different filter sizes and centrifugation settings have been used (Table 2.1) (Finkel *et al,* 2012). To establish the most successful means of capturing the entire bacterial content from the mānuka leaf wash buffer, a series of trials were conducted. Firstly, the efficacy of pelleting bacterial cells via centrifugation at 3,200 *g*

for 30 minutes was comparatively assessed against capturing bacterial cells on a 0.2 µm filter (Millipore). DNA extraction performed on pelleted bacterial cells consistently yielded sufficient DNA quantities (9.46 – 28.8 ng) whilst DNA extraction from filters did not (0.088 – 1.37 ng). Discrepancies exist as to whether smaller bacteria are able to be successfully pelleted and the extent to which DNA is damaged during centrifugal compaction (Peterson *et al,* 2012). Therefore, a series of trials were conducted to investigate whether centrifugation was successfully pelleting the entire bacterial content from the wash buffer. Leaves were sonicated, and leaf wash buffer was centrifuged. After careful removal of the supernatant from the bacterial pellet via pipetting, the supernatant was then filtered through a 0.2 µm filter. DNA extraction was performed directly on the filter, however, no quantifiable DNA was recovered (<0.05 ng/µl). A second trial was also conducted whereby leaf wash buffer supernatant was subject to a secondary centrifugation (6000 rpm, 90 minutes). No bacterial pellet was observed and no quantifiable DNA was extracted after this secondary spin. These trials confirmed centrifugation was successfully pelleting the entire bacterial content in the leaf wash buffer, thus providing validation for the centrifugation technique. Centrifugation was therefore implemented into the final methodology of the current research project.

2.5. Cross-contamination

The use of the same equipment for all samples during the isolation of bacterial cells raised the concern of potential cross-contamination. Therefore, a thorough investigation of efficient cleaning procedures was carried out. Bleach is a potent cleaning agent containing sodium hypochlorite which is commonly used in forensics to eliminate DNA contamination from the surface of bones and crime scene investigations (Kemp & Smith 2005, Rennick *et al,* 2005). Sodium hypochlorite causes biosynthetic alterations in cellular metabolism, phospholipid destruction, and irreversible enzymatic inactivation in bacteria (Estrela *et al,* 2002). The use of bleach as a cleaning agent for the current study was therefore trialled, following protocols outlined in previous research (Thomasma & Foran 2013). Briefly, surfaces, glass flasks, tweezers, and filter apparatus were soaked in disinfectant (Virkon) for 30 minutes and then rinsed in Milli-Q water. Equipment thereafter underwent a secondary soak in bleach (10% v/v bleach/water) for 60 minutes and thoroughly rinsed in Milli-

Q water. No quantifiable DNA (< 0.05 ng/µl) was recovered from equipment cleaned via this protocol. Furthermore, PCR amplification of these negative control trials was unsuccessful. This cleaning procedure was therefore implemented into the final methodology of this research project.

In summary, a review of recent cultivation-independent phyllosphere research was used to select and develop methods for the current research project (Figure 2.1). Owing to the large differences in methodological design across phyllosphere literature, controlling for these differences through meta-analyses may enable a better understanding of plant-microbe associations in the phyllosphere throughout future research efforts (Laforest-Lapointe *et al,* 2016a).

Table 2.1: A chronological summary of cultivation-independent experimental designs implemented across previous phyllosphere research

Leaf sample: wash buffer volume	Wash buffer constituents	Bacterial Isolation	Centrifugation	DNA extraction	Reference
15 g: 40mL	0.1M potassium phosphate (pH 7.0)	Ultrasonic cleaning bath 7 mins	30,000 g 15 mins	FastDNA kit	**Yang *et al,* (2001)**
unspecified	1:50 diluted enzymatic lysis (1M TrisCl,500mM EDTA 1.2% Triton)	Manually swirled leaves with gloved fingers	3000 g 10 mins	DNAeasy protocol DNA-prep Qiagen Kit.	**Kadivar and Stapleton (2003)**
12g: 50mL	0.1 M potassium phosphate buffer (pH 7.0)	Ultrasonic cell disrupter 10 mins 22.5kHz	30,000 g 15 mins 4°C	FastDNA kit	**Lambais *et al,* (2006)**
30-35 leaves: 500 mL	1:50 diluted enzymatic lysis (1 M Tris-HCL, 500 mM EDTA, 1.2% Triton)	Shook at 350 rpm 5 mins	2200 g 15 mins 4°C	MoBio PowerSoil DNA kit	**Redford and Fierer (2009)**
16 leaves: 3 L	20 mM Tris-HCL, 10mM EDTA & 0.024% Triton	Manual swirling with glass pipette 15 mins	4000 rpm 10 mins 4°C	Power Soil DNA extraction kit	**Rastogi *et al,* (2010)**
4 g: 40 mL	Sterile phosphate buffered (pH 7.4)	Ultrasonic cleaning bath 3 mins Recovered on 0.22 µm filter	n/a	Soil Microbial DNA extraction kit	**Finkel *et al,* (2011)**
3 leaves: 50 mL	Sterile dH_2O, 2-3 drops of Tween 20 per litre	Ultrasonic cleaning bath 5 mins	5500 g 20 mins 4°C	FastDNA kit	**Yashiro *et al,* (2011)**
40 g: 40mL	0.1 M phosphate buffered saline (pH 7)	Sonicated & vortexed 30 mins Decanted & recovered on 0.45µm filter	n/a	DNA extracted directly from frozen filter	**Finkel *et al,* (2012)**

40 g: 30 mL	TE buffer (10 mM Tris, 1 mM EDTA) pH 7.5, 0.2% Silwet L-77	Shook at 350 rpm vortexed & sonicated 3 mins	3,150 g 15mins 4°C	Power Soil DNA extraction kit	**Kim *et al,* (2012)**
4.5 g: 50 mL	PBS (137 mM NaCl, 10 mM,phospha te, 2.7 mM KCl, pH 7.4)	Ultrasonic cleaning bath 10 mins < 20°C	9500 rpm 20 mins 4°C	ZR soil microbe DNA kit	**Vokou *et al,* (2012)**
1 g: 10 mL	0.1 M potassium phosphate buffer (pH 8.0)	Sonicated & vortexed (2 mins) Recovered on 0.2 µm filter	n/a	Power Soil DNA isolation kit	**Bodenhausen *et al,* (2013)**
50-100g: 100mL	1:50 diluted wash solution (1 M Tris HCl, 0.5 M Na EDTA and 1.2%CTAB)	Manually agitated 5 mins	4000 g 20 mins	MoBio PowerSoil kit	**Kembel *et al,* (2014)**
10-20g: 100mL	0.1 M potassium phosphate 0.1% glycerol, 0.15% Tween 80 (pH7.0)	Vortexed & swabbed 7 mins Recovered on 0.2 µm filter	n/a	PowerSoil DNA isolation kit	**Ruiz-Pérez *et al,* (2016)**
720cm^2: 50 mL	0.85% NaCl, Tween 80	Sonication 60 Hz 6 mins	6169 g 20 mins	FastDNA SPIN kit for soil	**Ortega *et al,* (2016)**
200cm^2: 20mL	TE buffer (10mM Tris, 1 mM EDTA, pH 8)	Vortexed & manually shaken (8 mins)	4000 g 10 mins 8000 g 10 mins	PowerFecal DNA isolation kit	**Smets *et al,* (2016)**
1 leaf: unspecified volume	TE buffer (pH 8.0)	Leaf surface brushed with sterile toothbrush (2 mins)	10,000 g 2 mins	PowerSoil DNA isolation kit	**Stone and Jackson (2016)**
20 leaves: unspecified volume	100mM Tris-HCl, 2% Percoll (pH7.5)	Sonicated 22.5 kHz (1.5 mins) Agitated 100 rpm (10mins)	3000 g 15 mins	Fast DNA kit	**Lambais *et al,* (2017)**

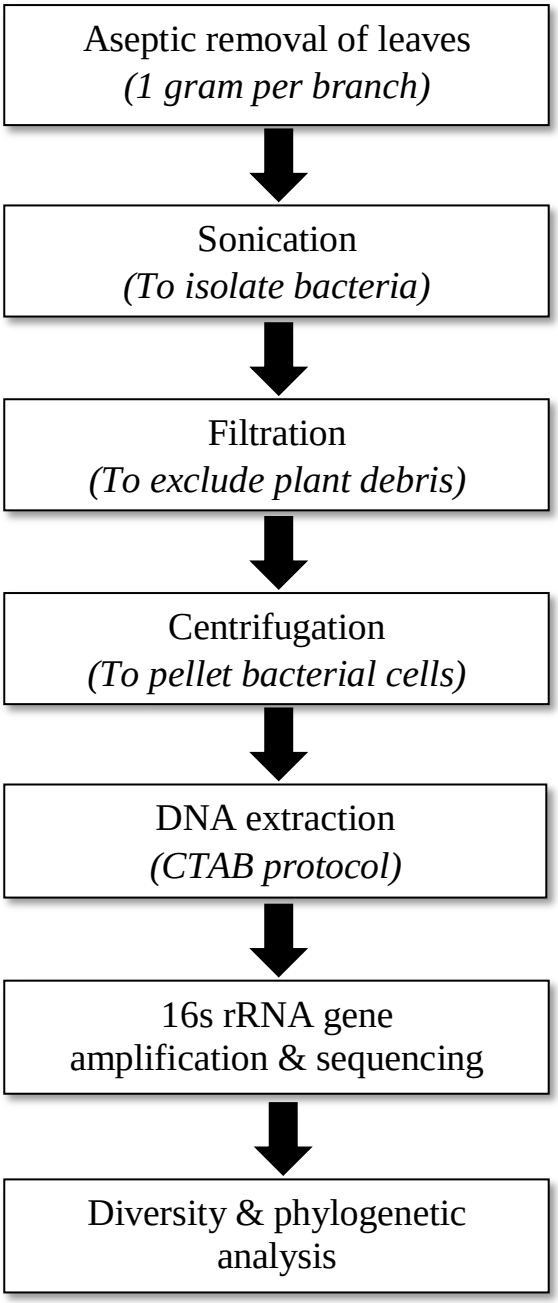

Figure 2.1: The methodological pipeline developed to analyse the bacterial communities of the mānuka phyllosphere

3. Chapter Three:
Materials and Methods

3.1. Study sites

Five discrete, naturally growing mānuka forests across the North Island were chosen for study in collaboration with Assoc. Prof Michael Clearwater and Steens Mānuka Honey Ltd. Native mānuka forests chosen for study were located in: Mohaka, Hawkes Bay (MK) (39°01 S; 177°08 E), Serpentine Lake, Ohoupo (SL) (37°56 S, 175°19 E), Mangatarere Valley, Wellington (MV) (40°57 S; 175°26 E), Mamaku, Bay of Plenty (KU) (38°02 S; 176°03 E), and North-eastern Kaimanawas, Taupo (HT) (39°06 S; 176°21 E) (Figure 3.1). The linear distance between the centres of each sample site ranged from 65 – 333 km for SL – KU and SL – MV, respectively (Figure 3.2).

3.2. Sample and metadata collection

Samples were collected during the spring and summer mānuka flowering season from November 2016 – January 2017.

- MK, Tuesday November 22nd, 2016
- SL, Saturday December 3rd, 2016
- MV, Sunday December 18th, 2016
- KU, Wednesday December 28th, 2016
- HT, Monday January 2nd, 2017

Upon arrival of each sample site, an environmental datalogger was set up and left to measure and record metadata every 15 minutes across the 24 hours prior to sampling. These environmental parameters included: air temperature (°C), relative humidity (%), and photosynthetically active radiation (μMm^{-2}). At each site, six mānuka trees were selected for sampling. All trees were approximately 15 m apart and seemingly healthy with no visible signals or symptoms of disease or damage. Three mānuka branches were randomly selected from each tree, providing a total of 90 samples. Metadata was thereafter collected for each mānuka tree selected for sampling, these included: GPS coordinates (WGS84 (G1762); degree minutes), elevation (m),

tree height (m), and the basal diameter of the tree base at 10 cm off the ground (cm). For multi-stemmed trees, the basal diameter was determined by measuring all the trunks and then adding the total diameter of the largest trunk to one-half the diameter of each additional trunk. For each branch sampled, collected metadata included branch height (m) and branch aspect (degrees). During sample collection, branches were handled aseptically with gloves, cut with clippers sterilised on site using 70% v/v ethanol/water in between each sampling to minimise cross-contamination, put in individual sterile zip lock bags and immediately placed on ice. Branches were collected at approximately the same time on each sampling day. Surface soil from multiple positions around the base of each mānuka tree were also collected into sterile 50 mL Falcon tubes using a spatula sterilised on site using 70% v/v ethanol/water and immediately placed on ice. Upon return to the Thermophile Research Unit at the University of Waikato, branch and soil samples were frozen at -20 °C and -80 °C, respectively, until further analysis. Additional metadata was collected from The National Climate Database (NIWA) after sampling (https://cliflo.niwa.co.nz/). These included monthly weather averages of: precipitation (mm), temperature (°C), wind speed (mph), pressure (mb), cloud cover (%), humidity (%), and sun (hrs) (NIWA 2017).

Figure 3.1: North Eastern Kaimanawa (HT) sampling site

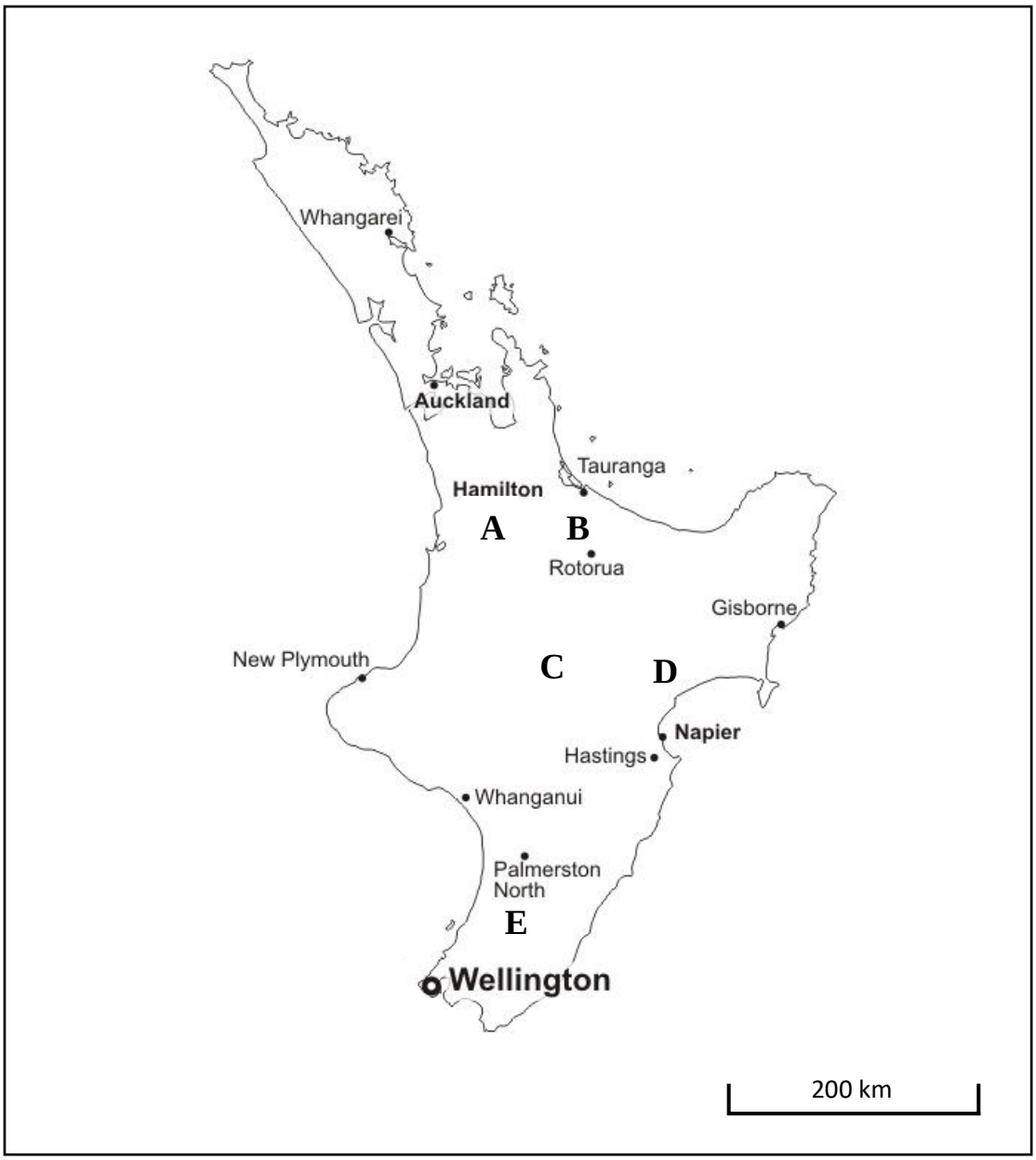

Figure 3.2: Sample mānuka populations across the North Island of New Zealand. Serpentine Lake, Ohaupo (SL) (A); Mamaku, Bay of Plenty (KU) (B); North-eastern Kaimanawas, Taupo (HT) (C); Mohaka, Hawkes Bay (MK) (D); and Mangatarere Valley, Wellington (MV) (E).

3.3. Isolation of microbial cells from the mānuka leaf surface

Bacterial cells were isolated from the mānuka leaf surface using a methodological approach developed in Chapter 2. One gram of predominantly healthy, undamaged green mānuka leaves were aseptically excised from each mānuka branch (Figure 3.3). Great care was taken to avoid inflicting physical damage to the leaf surface and ensure minimal petiole was collected. By not pooling leaves on a per-tree basis, this study design permitted the detection of intra-individual variation in phyllosphere community structure. Moreover, leaf age was not taken into consideration due to the difficulties associated with directly determining the ages of individual leaves.

The leaves from each branch were placed in individual sterile glass flasks, submerged in 10 mL of wash buffer [PBS, 1% T20] and sealed with parafilm. Flasks were carefully lowered into an ultrasonic cleaning bath and sonicated for 20 minutes at 60Hz (Figure 3.4). After sonicating, the leaf wash buffer was decanted from the leaves into a sterile 50 mL falcon tube. An extra 10 mL of PBS was used to further wash the leaves and decanted again. This washing step was repeated twice. The leaf wash buffer was thereafter syringe filtered through 90 µm nylon mesh into a second sterile 50 mL falcon tube in order to remove finer plant debris. The filtered leaf wash was centrifuged at 3,200 *g* for 30 minutes to pellet bacterial cells. After centrifugation, the supernatant was carefully removed via pipetting and the bacterial cell pellet was resuspended in 270 µl of PBS (100mM NaH_2PO_4), transferred to a 1.5-2.0 mL screw-capped conical bottomed polypropylene tube containing 0.5 g each of 0.1 mm and 2.5 mm silica-zirconia beads, and frozen at -80°C until DNA extraction. All equipment and vessels that contact and contain leaves were subject to stringent washing procedures whereby equipment was soaked in disinfectant (Virkon) and bleach (10% v/v bleach/water), rinsed in milli Q water, and left to dry.

Figure 3.3: Mānuka branch sample prior to aseptic leaf removal

Figure 3.4: Manuka leaves submerged in 10 mL leaf wash buffer and placed in an ultrasonic cleaning bath

3.4. DNA extraction

Total genomic DNA was extracted from the microbial cell suspension of each mānuka sample using the CTAB extraction protocol as outlined in Dreesens *et al*, (2014). Briefly, microbial cell suspensions were thawed on ice and 270 µL sodium dodecyl sulfate (SDS) lysis buffer (100mM NaCl, 500mM Tris pH 8.0, 10% SDS) added to each. Samples underwent bead beating for 15 seconds and horizontal shaking on a vortex genie for 10 minutes. Samples were then compacted by centrifugation at 10,000 rpm for 30 seconds. A working solution of CTAB buffer containing 2% cetyl trimethylammonium bromide, 1% polyvinyl pyrrolidone, 100 mM Tris-HCl, 1.4 M NaCl and 20 mM EDTA and 0.5% β-mercaptoethanol (BME) was made and 180 µl added to each sample. Samples were vortexed for 10 seconds and followed by a 30 minute incubation at 300 rpm and 60 °C. A second centrifugation at 10,000 rpm for 30 seconds was performed to reduce bubbles. Following centrifugation, 350 µl chloroform:isoamyl alcohol (24:1) was added and samples were vortexed for 15 seconds and centrifuged at 10,000 rpm for 5 minutes. The upper aqueous layer was collected into a new 1.5 mL microfuge tube to which 500 µl chloroform:isoamyl alcohol (24:1) was added. Samples were vortexed again for 15 seconds and placed on a rocking bed for 20 minutes. Another centrifugation at 10,000 rpm for 5 minutes was performed and the resulting upper aqueous layer was collected into a new 1.5 mL microcentrifuge tube. Ammonium acetate (2.5 M) was then added, following which samples were vortexed for an additional 10 seconds and centrifuged at 10,000 rpm for 5 minutes. The upper aqueous layer was transferred to a new 1.5 mL microcentrifuge tube. To precipitate the DNA, isopropanol (C_3H_8O) was added, mixed by 20 inversions, and incubated at -20 °C overnight. The next day, samples were centrifuged at 13,200 rpm for 20 minutes at 4 °C, the supernatant was discarded, and the pellet was washed with 1 mL cold 70% AR grade ethanol by centrifugation for a further minute at 10,000 rpm. The ethanol was carefully pipetted off and the pellet was dried in a speed vacuum (Thermo Fisher Scientific, USA) for 4-15 minutes during which the samples were checked every two minutes. Finally, DNA pellets were re-suspended in 20 µl sterile 1 X TE (10 mM Tris, 1mM EDTA) by pipetting up and down and vortexing for 10 seconds. Extracted DNA was quantified using the Qubit Flurometer High sensitivity double stranded DNA protocol (Life Technologies, Auckland) and stored at -20 °C

until further analysis. The Power Soil DNA Extraction kit (MoBio) was meanwhile used to extract DNA from 0.5 g of each soil sample collected from the base of each mānuka tree.

3.5. PCR amplification of 16s rRNA genes

Polymerase chain reaction (PCR) was used to amplify the V4 region of the 16S rRNA gene from extracted DNA using the primer set F515 (5'-GTGCCAGCMGCCGCGGTAA-3') and R926 (5'- CCACTACGCCTCCGCTTTCC TCTCTATGGGCAGTCGGTGATCCGYCAATTYMTTTRAGTTT-3'). These primers generate an amplicon approximately 500 base pairs in length with adapters suited for Ion Torrent sequencing. PCR reactions (20 µl) were performed in a 200 µL thin walled PCR reaction tube. Each PCR reaction consisted of 0.8 µL bovine serum albumin (BSA) (Promega Corporation, USA), 2.4 µL dNTPs (2mM each) (Invitrogen Ltd, New Zealand), 2.4 µL 10x PCR buffer (Invitrogen), 2.4 µL $MgCl_2$ (50 mM) (Invitrogen), 0.4 µL each primer (10 mM) (Integrated DNA Technologies, Inc), 0.096 µL Taq DNA polymerase (Invitrogen), 2 µL of genomic DNA (2.5 ng/µl), and 9.104 µL molecular-grade ultrapure water. Reactions were performed in triplicate for each sample with the following thermocycler conditions: 3-min initial denaturation at 94°C, followed by 30 cycles of 45 s at 94°C, 1-min at 50°C, and 1.5-min at 72°C, with a final 10-min elongation at 72°C. A positive and a negative control was run in every PCR. Triplicate reactions were pooled and stored at -20°C. Agarose gel electrophoresis was used to confirm PCR amplification. Briefly, 5 µL microliters of pooled PCR product was combined with 2 µL of 10X loading dye (Invitrogen) and loaded into the wells at the top of a 1% agarose gel. To estimate the size of the PCR products, 7 µL of each 50 bp and 1 KB+ ladder was also loaded into each gel. Gels were run for 25-30 min at 70 V after which they were removed from the gel comb and visualised using an Alpha Imager. Consistent failure to amplify lead to the exclusion of one mānuka sample (MV1.5) and one soil sample (SL3).

3.6. Next Generation Sequencing

The concentration of the PCR amplicons for each sample was normalized and purified using a SequelPrep Normalization Kit, according to manufacturer's instructions (Life Technologies, Auckland). Briefly, 25 µl of pooled PCR product was

added to a plate well with 25 µl of binding buffer and incubated at room temperature for 1 h. This amplicon/binding buffer mixture was thereafter removed, and the plate wells were washed by pipetting 50 µl of wash buffer up and down to remove residual PCR product. Amplicons were then eluted with 20 µl of normalized elution buffer and incubated at room temperature for 5 min. From the plate wells, samples were transferred into sterile microcentrifuge tubes and 2 µl of each sample was pooled and stored at 4 °C prior to commencing sequencing. DNA sequencing was undertaken at the Waikato DNA Sequencing Facility at the University of Waikato using an Ion Torrent PGM DNA sequencer with an Ion 318v2 chip (Life Technologies). Raw sequences in FASTQ format were first filtered in mothur to remove short reads, long reads, and reads with excessive homopolymers (Schloss *et al,* 2009). Thereafter, sequences were quality filtered using USEARCH (ver 9) (Edgar 2010). A total of 1,890,959 high quality sequence reads were obtained and clustered into operational taxonomic units' (OTUs) at a percent sequence similarity threshold of 97%. After filtering for chloroplast OTUs, a total 928,317 sequence reads and an average 10,431 (mean) reads per mānuka sample remained (n = 89). These sequence reads mapped to 1384 bacterial OTUs (97%). Meanwhile, a total 736, 849 reads and an average 25,409 (mean) reads per soil sample remained (n = 29). These sequence reads mapped to 6905 bacterial OTUs (97%). From this initial processing, a BIOM file, FASTA file and OTU table were generated (McDonald *et al,* 2012).

3.7. Data analyses

The raw FASTA file was run through the Michigan State University Ribosomal Database Project (RDP) Classifier whereby taxonomy was assigned to the 16S rRNA sequences (Wang *et al,* 2007). Taxonomic assignments were provided at Kingdom, Phylum, Class, Order, Family, and Genus level. No sequence reads were assigned at the Species level. Taxonomic assignments with estimated confidence less than 80% were classified as 'unknown'. Taxonomy data was merged with the BIOM file using the 'biom add-metadata command' in The BIOM file format (ver 2.1). Sequence data were statistically analysed using R (ver 3.4.3) (R Core Team 2017) and a number of R packages including: phyloseq (McMurdie & Holmes 2013), ggplot2 (Wickham 2016), vegan (Oksanen *et al,* 2007), dplyr (Wickham & Francois 2015), magritter (Bache & Wickham 2014), reshape2 (Wickham 2007), randomForest (Liaw & Wiener 2002),

knitr (Xie 2013), tibble (Wickham *et al,* 2017), and tidyverse (Wickham 2017). Sequence reads were rarefied on a per sample basis to minimum library sequencing read depth (3192) prior to alpha diversity analyses. During rarefaction, one mānuka sample (HT2.4_17) was excluded due to unusually low sequence reads (1375) (n = 88). The Bray-Curtis dissimilarity index was chosen for computing dissimilarity between bacterial communities because this distance accommodates for the presence of double zeros (absence) and enables the results of this study to be compared against many other studies (Chao *et al,* 2006). The Euclidean index was meanwhile chosen for computing the straight-line distance between spatial and environmental parameters (Danielsson 1980). Mantel tests using the Pearson product-moment correlation coefficient were completed to assess the correlation between community and environmental distance matrices (Guillot & Rousset 2013). Correlational analyses were performed on the relative abundance of main phyla as well as at the OTU level. Variance partitioning was used to resolve the explanatory power of different explanatory matrices. Moreover, communities were segregated into abundant (> 1% relative abundance) and rare (< 1% relative abundance) taxa, in order to investigate different community members. The Basic Local Alignment Search Tool (BLAST) algorithm was used to compare specific 16S rRNA nucleotide sequences obtained from NGS to available online sequence databases. Metabolic pathways were also predicted using the PICRUSt software (Langille *et al,* 2013). After performing an additional closed-reference OTU picking pipeline, KEGG orthologies (KO) and main metabolic pathways, categorised by the KEGG database, were obtained and qualitatively analysed.

4. Chapter Four:
Results and Discussion

4.1. General characteristics of bacterial communities

A total 1382 bacterial OTUs were identified from the leaves of 89 branches at an average 260 OTUs per sample. A considerable proportion of bacterial diversity was considered sample-specific, with 312 bacterial OTUs (22.6%) only occurring within a single leaf sample. The 312 OTUs that were only present in a single leaf sample represented 0.12% of sequence reads. In total, 1012 (73.5%) bacterial taxa were found at a relative abundance less than 1%, which represented 2.2% of sequence reads; 513 (37.1%) bacterial taxa were found at a relative abundance less than 0.1%; and 68 (4.9%) bacterial taxa were found at a relative abundance less than 0.01%. Although only 371 (26.9%) bacterial taxa in the mānuka phyllosphere microbiome were found at a relative abundance greater than 1%, these constituted 97.8% of sequence reads. Each sample was rarefied to an even sequencing depth (3192 reads) and alpha diversity of each sample was analysed using richness, Shannon, Simpson, Chao1, and ACE indices. These diversity measures did not vary significantly between different branches and were therefore averaged for each sample tree (PERMANOVA, $P > 0.05$) (Table 4.1). Moreover, diversity measures did not vary significantly across different sample trees or sites (PERMANOVA, $P > 0.05$) (Table 4.1). Sample locations situated at higher elevations generally supported a greater richness of bacterial taxa, except for the highest elevated site (HT), which contained the lowest average richness of bacterial taxa (160 ± 40.4) ($n=17$). The average taxon richness of samples from the lowest elevation (SL) was 162 ± 26.5 ($n = 18$), compared with 190 ± 47.9 ($n = 18$) for samples from the higher elevations (KU). However, the relationship between species richness and elevation was not statistically significant ($P > 0.05$). Interestingly, species richness did show a significant and positive linear relationship with tree height, whereby taller trees tended to support a greater richness of bacterial taxa ($R = 0.36$, $P < 0.001$).

Table 4.1 Alpha diversity estimation of the mānuka phyllosphere microbiome

Estimations (± standard deviation) were calculated for a randomized subset of 3192 reads per sample and averaged for each sample tree (continued over page).

Tree	Observed	Shannon	Simpson	Chao1	ACE
HT2	103 ± 19	2.61 ± 0.38	0.85 ± 0.05	130.2 ± 5.3	140.9 ± 1.6
HT3	135 ± 33	3.24 ± 0.48	0.92 ± 0.04	180.5 ± 43.6	173.7 ± 42.7
HT5	148 ± 33	3.28 ± 0.4	0.92 ± 0.03	206 ± 39.1	207.2 ± 49.1
HT6	185 ± 37	3.81 ± 0.12	0.96 ± 0.002	240 ± 52.6	247.7 ± 56.2
HT7	193 ± 41	3.79 ± 0.32	0.95 ± 0.009	237.2 ± 37.9	243.5 ± 35
HT9	179 ± 3	3.52 ± 0.1	0.94 ± 0.006	244.8 ± 17.5	253.5 ± 13.6
KU1	170 ± 46	3.53 ± 0.36	0.94 ± 0.02	218.8 ± 80.5	224.2 ± 75
KU2	240 ± 38	4.01 ± 0.18	0.95 ± 0.005	299.2 ± 53.8	306.4 ± 55
KU3	139 ± 9	3.16 ± 0.17	0.9 ± 0.009	179.5 ± 11.6	192 ± 9.8
KU4	225 ± 43	3.74 ± 0.44	0.93 ± 0.03	304.1 ± 68.3	306 ± 69
KU5	194 ± 32	3.65 ± 0.2	0.94 ± 0.007	252.8 ± 44.4	260.7 ± 42
KU6	157 ± 31	3.4 ± 0.38	0.93 ± 0.02	213.9 ± 59.6	220.1 ± 47.3
MK2	172 ± 32	3.45 ± 0.42	0.92 ± 0.03	210.1 ± 31.6	210.3 ± 30.1
MK3	195 ± 18	3.77 ± 0.19	0.95 ± 0.005	243.8 ± 45.1	248.7 ± 40.1
MK4	187 ± 35	3.76 ± 0.34	0.95 ± 0.01	246.1 ± 52.6	246.8 ± 51

MK5	149 ± 17	3.26 ± 0.21	0.91 ± 0.01	173.5 ± 21.4	177.9 ± 20.5
MK6	149 ± 26	3.49 ± 0.11	0.94 ± 0.005	186.7 ± 37	188.1 ± 37.5
MK9	170 ± 7	3.37 ± 0.06	0.92 ± 0.003	228.4 ± 16.6	219.2 ± 12.2
MV1	210 ± 6	3.81 ± 0.06	0.95 ± 0.001	274.1 ± 9.6	272 ± 0.3
MV2	171 ± 11	3.52 ± 0.08	0.93 ± 0.003	217.6 ± 7.7	224.3 ± 5.2
MV3	181 ± 9	3.61 ± 0.28	0.94 ± 0.01	252.9 ± 30.2	252.8 ± 30.4
MV4	138 ± 16	3.06 ± 0.17	0.89 ± 0.02	179.5 ± 26.3	189.6 ± 17.5
MV5	198 ± 17	3.61 ± 0.32	0.93 ± 0.008	288 ± 22	269.8 ± 14.8
MV6	185 ± 34	3.66 ± 0.3	0.94 ± 0.02	226.3 ± 37.2	232.3 ± 36.7
SL1	164 ± 36	3.47 ± 0.24	0.94 ± 0.01	212.8 ± 51.9	221.4 ± 45.8
SL2	157 ± 21	3.36 ± 0.4	0.92 ± 0.01	217.3 ± 38.8	225.3 ± 46.1
SL3	174 ± 40	3.40 ± 0.4	0.91 ± 0.05	219.2 ± 56.2	224.4 ± 57.6
SL4	167 ± 6	3.50 ± 0.1	0.94 ± 0.005	200.4 ± 15.4	209.4 ± 16
SL7	159 ± 28	3.44 ± 0.27	0.93 ± 0.02	203.7 ± 53.8	211.9 ± 53.6
SL8	148 ± 33	3.22 ± 0.57	0.91 ± 0.05	200.5 ± 50.2	208.1 ± 37

4.2. The mānuka phyllosphere core microbiome

Whether a core microbiome exists in the mānuka phyllosphere was one of the central questions of this current research. In total, 16S rRNA gene PCR amplicon sequencing of communities from the mānuka phyllosphere detected 21 bacterial phyla (for full phylum-level analysis see Supplementary Table 6.1). Taxa belonging to *Proteobacteria*, *Acidobacteria*, *Bacteriodetes*, *Firmicutes*, and *Verrucomicrobia* consistently represented the largest proportion of sequence reads across all samples, of which *Proteobacteria* showed the greatest relative abundance. Within *Proteobacteria*, *Alphaproteobacteria* was the most abundant bacterial class, representing 46.5% of sequence reads, followed by *Gammaproteobacteria* (3.2%), *Betaproteobacteria* (0.75%) and *Deltaproteobacteria* (0.34%). Taxa belonging to the *Methylobacterium* genus were also relatively abundant across all samples (2.7%). After manually removing chloroplast sequences, taxa belonging to the phylum *Cyanobacteria* were detected in 21 samples (0.15%). Intriguingly, a core microbiome, consisting of 10 bacterial OTUs, was identified (Figure 4.1). Core taxa were present in all leaf samples and belong to four phyla and five classes: *Alphaproteobacteria* [*Rhizobiales* (29.3% of all sequences), *Sphingomonadales* (4.9%)], *Bacteroidetes* (2.5%), *Verrucomicrobia* (2.4%), and *Acidobacteria* (1.0%). These were relatively abundant and represented 40.1% of all 16S rRNA gene sequences. The most abundant of these 10 core taxa was an unclassified member of the *Alphaproteobacteria* class in the order *Rhizobiales* (OTU 2) and accounted for 107,066 (11.15%) sequence reads.

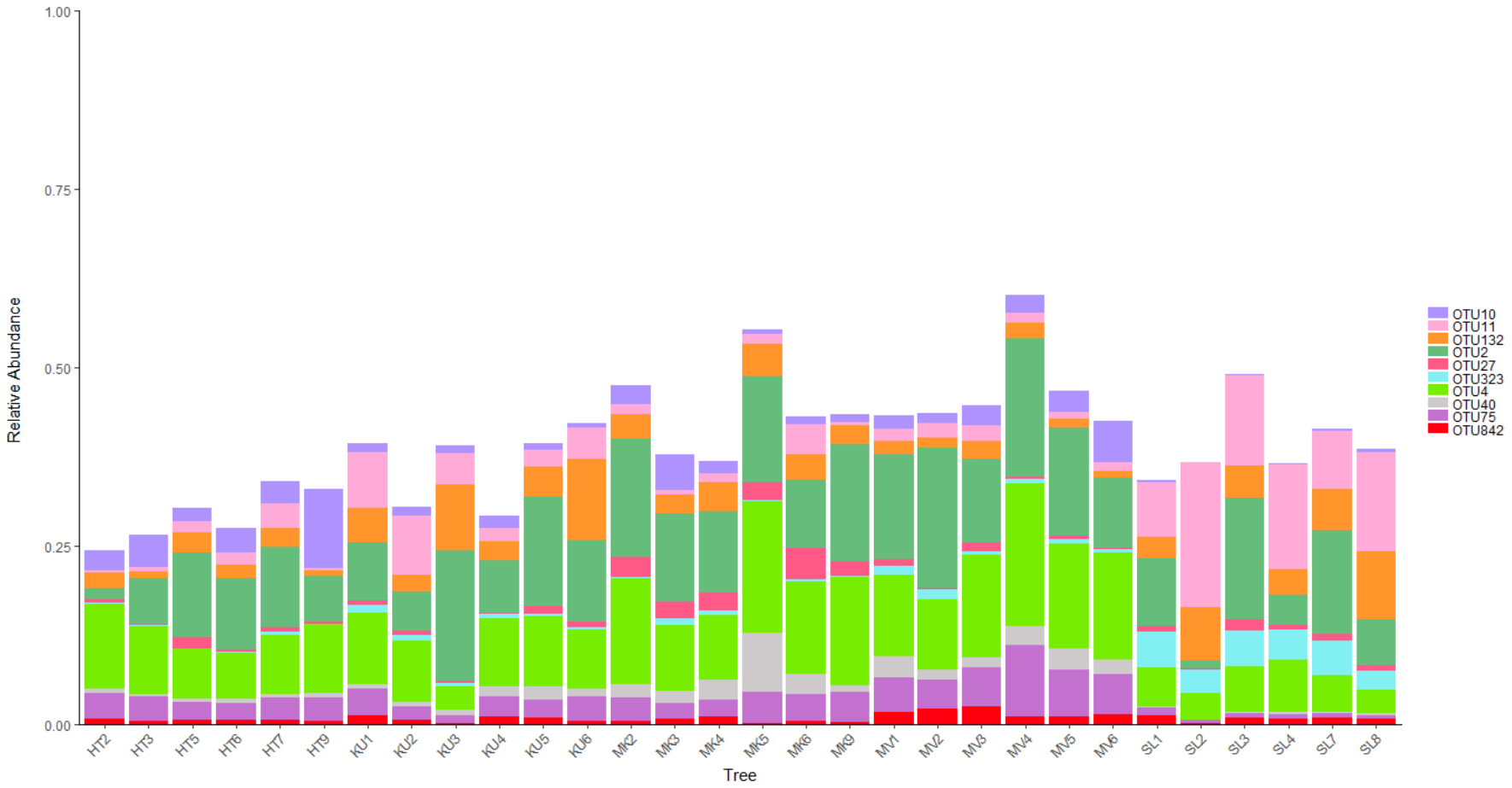

Figure 4.1: Relative abundance (%) of 10 OTUs that comprise the manuka phyllosphere core microbiome.
Relative abundance is averaged per sample tree.

To ascertain the habitat-specificity of core taxa detected in the mānuka phyllosphere, the microbial communities in the surface soil surrounding each mānuka tree were characterised. Across the soil samples, 16S rRNA gene PCR amplicon sequencing identified 27 bacterial phyla and four archaeal phyla. Taxa belonging to nine bacterial phyla represented the largest proportion of sequence reads: *Proteobacteria* (29.6%) *Acidobacteria* (19.7%), *Bacteriodetes* (13.1%), *Verrucomicrobia* (9.6%), *Actinobacteria* (5.5%), *Planctomycetes* (4.6%), *Chloroflexi* (2.3%) and *Bacteria candidate division WPS2* (1.6%). In total, 6905 bacterial OTUs were detected across all soil samples at an average 1820 bacterial OTUs per sample. Of these, 6433 (93.2%) bacterial taxa were found at a relative abundance less than 1%, 4460 (64.6%) taxa were found at a relative abundance less than 0.1%, and 634 (9.2%) bacterial taxa were found at a relative abundance less than 0.01%. Only 471 taxa were detected at an abundance greater than 1% (6.8%). A small proportion of bacterial taxa were sample-specific, with 1339 bacterial OTUs (19.4%) occurring within a single soil sample. The 1339 OTUs that were only present in a single soil sample represented 0.8% of sequence reads. Moreover, unlike the phyllosphere, a core microbiome was not detected across surface soil samples. The structure of phyllosphere and soil bacterial communities were compared at the OTU level using the Bray-Curtis dissimilarity index and significant differences were detected (PERMANOVA, $P < 0.001$). Overall, 599 bacterial OTUs were shared by at least one soil and leaf sample, rendering 6306 and 783 OTUs exclusively habitat-specific to the leaf surface and surrounding soil environment, respectively (Figure 4.2). Of the ten OTUs comprising the mānuka phyllosphere core microbiome, nine were present in at least one soil sample (Table 4.2). However, while these OTUs were relatively abundant across all mānuka samples, they were scarce within soil samples (Figure 4.1 & Table 4.2). Six OTUs (OTU323, OTU27, OTU2, OTU132, OTU75, and OTU10) were detected at an average relative abundance of 0.001% in 1-2 sample sites (Table 4.2). OTU4 was found at an average relative abundance of 0.001 – 0.003% across four of the five sample sites (Table 4.2). OTU842 and OTU11 were meanwhile detected at an average relative abundance of 0.001 – 0.005% and 0.032 – 0.162% across all five sites, respectively (Table 4.2).

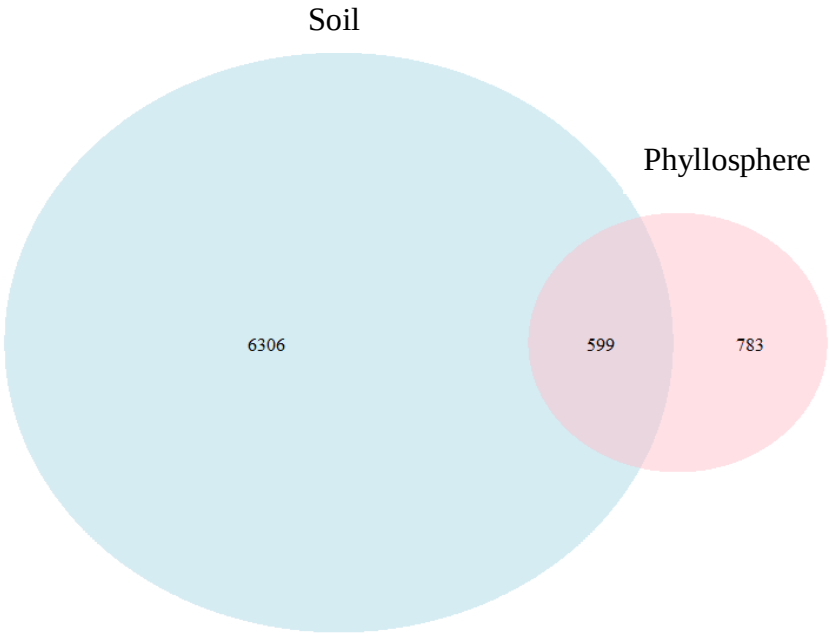

Figure 4.2: Venn diagram showing the number of shared and unique OTUs in soil and phyllosphere sample groups. An OTU is considered shared if it is present in at least one leaf and soil sample. Circles size is proportional to the OTU numbers detected in each sample group.

Table 4.2: Relative abundance (%) of 10 mānuka phyllosphere core OTUs in soil. Relative abundance is averaged per site. OTU absence is represented with '-'.

	HT	KU	MK	MV	SL
OTU323	-	-	0.001%	-	-
OTU27	-	0.001%	-	0.001%	-
OTU2	-	-	0.001%	0.001%	-
OTU132	0.001%	-	-	-	-
OTU11	0.054%	0.032%	0.162%	0.154%	0.038%
OTU75	-	-	-	0.001%	-
OTU4	0.002%	0.001%	0.003%	0.002%	-
OTU40	-	-	-	-	-
OTU842	0.004%	0.001%	0.005%	0.001%	0.001%
OTU10	0.001%	-	-	-	-

4.2.1. Putative identities of the core microbiome

Nucleotide BLAST analysis was conducted on the 10 core OTUs via blastn through the National Centre for Biotechnology Information (NCBI) (http://www.ncbi.nlm.nih.gov) to assess the nucleotide sequence similarity of core taxa to the NCBI 16S rRNA sequence database collection (Table 4.3). The taxonomy and physiology of closely related isolates were sourced from the associated primary literature and are described for each of the 10 core OTUs below. Sequences of core taxa are presented in Supplementary Table 6.2.

Two OTUs (OTU323 and OTU40) are most closely related to members of the genus *Hymenobacter* (Table 4.3). OTU 323 showed 99% 16S rRNA gene sequence similarity with *H. ginsengisoli*, which was isolated from soil of a ginseng field in a mountainous region of Chungnam province in South Korea (Hoang *et al,* 2013); 98% 16S rRNA gene sequence similarity with *H. metalli*, which was isolated from a uranium mine waste water treatment system in Urgeirica¸ Portugal (Chung *et al,* 2010); and 98% 16S rRNA gene sequence similarity with *H. marinus*, which was isolated from coastal seawater of the East Sea, South Korea (Kang *et al,* 2016). OTU 40 showed 96% 16S rRNA gene sequence similarity with *H. marinus*; 96% 16S rRNA gene sequence similarity with *H. flocculans*, which was isolated from a uranium mine wastewater treatment system in Urgeirica¸ Portugal (Chung *et al,* 2010); 95% 16S rRNA gene sequence similarity with *H. ginsengisoli*; and 95% 16S rRNA gene sequence similarity with *H. metalli*. These four *Hymenobacter* bacteria are non-motile, red pigmented, and exhibit optimum growth between 25 – 30 °C and pH 6.0 – 7.0.

Two OTUs (OTU27 and OTU11) are most closely related to members of the *Sphingomonas* genus (Table 4.3). OTU27 showed 98% 16S rRNA gene sequence similarity with *S. cynarae*, which was isolated from the phyllosphere of *Cynara cardunculus*, a Mediterranean native plant considered to be the wild ancestor of the globe artichoke and cultivated cardoon (Tala *et al,* 2013). Interesting physiological properties pertaining to *S. cynarae* include the copious secretion of extracellular polysaccharides (EPS) and the formation of slimy, viscous, orange-pigmented colonies that grow optimally at pH 6.0 – 6.5 and temperatures 26 – 30 °C. A component analysis of the exopolysaccharide produced by *S. cynarae* suggests the bacterium produces a novel type of sphingan. Sphingans are a group of structurally related bacterial exopolysaccharides and are useful for a number of large industrial applications (Tala

et al, 2013). OTU27 also showed 98% 16S rRNA gene sequence similarity with *S. xinjiangensis*, an aerobic, motile taxon which was isolated from desert sand of Xinjiang, China (An *et al,* 2011); and 98% 16S rRNA gene sequence similarity with *S. insulae*, which was isolated from the soil of Dokdo, South Korea (Yoon *et al,* 2008). OTU11 showed 99% 16S rRNA gene sequence similarity with *S. kyeonggiensis*, which was isolated from soil within a field of slow-growing perennial plants in the Pocheon province, South Korea (Son *et al,* 2014, Son *et al,* 2013); 99% 16S rRNA gene sequence similarity with *S. insulae*; and 99% 16S rRNA gene sequence similarity with *S. sanxanigenens*, which was isolated from cornfield topsoil in Xinhe County, China (Huang *et al,* 2009). *S. xinjiangensis*, *S. insulae* and *S. kyeonggiensis* grow optimally at temperatures ranging from 25 – 28 °C and pH 7.0. In addition, *S. xinjiangensis* and *S. kyeonggiensis* are both catalase- and oxidase-positive, and exhibit β-glucosidase activity. Catalase and oxidase enzymes protect aerobically-respiring bacteria from ROS (Sinha & Hader 2002).

Two OTUs (OTU75 and OTU4) are most closely related to members of the *Bradyrhizobium* genus (Table 4.3). The *Bradyrhizobium* genus belongs to the *Bradyrhizobiaceae* family and *Rhizobiales* order. These taxa were typically isolated from root nodules of various plants across the world, such as the peanut plant *(Arachis hypogaea),* legume species of the genus *Lespedeza* (i.e. pea plants), and flowering plants of the genus *Cytisus* (Chahboune *et al,* 2011). For example, OTU75 showed 92% 16S rRNA gene sequence similarity with *B. vignae*, which was isolated from root nodules of cowpea from farming fields in the Kavango region of Namibia, Africa (Gronemeyer *et al,* 2016). *B. vignae* is a motile, aerobic bacterium that demonstrated growth at 40 ℃ and contained the nifH gene. OTU4 exhibited 92% 16S rRNA gene sequence similarity with *B. ganzhouense*, which was isolated from the Australian blackwood, *Acacia melanoxylon*, an Acacia species native to Eastern Australia (Lu *et al,* 2014). *B. ganzhouense* is an aerobic bacterium that displays optimum growth at 28 °C and pH 7.0 as well as positive catalase and oxidase activity.

OTU2 is most closely related to *Beijerinckiaceae* and *Methylocystaceae*, families of the *Rhizobiales* order (Table 4.3). Interestingly, *Beijerinckiaceae* forms a monophyletic cluster with the *Methylocystaceae*, the only other family of alphaproteobacterial methanotrophs (Tamas *et al,* 2013). OTU2 showed 96% 16S rRNA gene sequence similarity with *Methylocapsa palsarum*, which was isolated from

a sub-arctic mossy soil in northern Norway (Dedysh *et al*, 2015); 96% 16S rRNA gene sequence similarity with *Beijerinckia doebereinerae*, which was isolated from soil samples collected in various regions of Brazil (Oggerin *et al*, 2009); and 96% 16S rRNA gene sequence similarity with *Methylocystis bryophila*, which was isolated from an acidic peat-bog lake in Teufelssee, Germany (Belova *et al*, 2013). *M. palsarum* is a non-motile, non-pigmented bacterium and possesses a particulate methane monooxygenase (MMO) enzyme. *M. palsarum* utilises methane and methanol and grows in a wide pH range of 4.1 – 8.0 (optimum pH 5.2 – 6.5) and at temperatures between 6 – 32 °C (optimum 18 – 25 °C). *B. doebereinerae* demonstrates optimal growth at 30 °C and pH 6.5. *M. bryophila* is facultatively methanotrophic, possesses both a soluble and a particulate MMO, and grows between pH 4.2 – 7.6 (optimum pH 6.0 – 6.5) and at 8 – 37 °C (optimum 25 – 30 °C). *M. palsarum* and *M. bryophila* are capable of atmospheric nitrogen fixation and *M. bryophila* also demonstrates slow growth on acetate in the absence of C1 substrates.

OTU132 is most closely related to bacteria of the genera *Bradyrhizobium* and *Camelimonas* (Table 4.3). The *Camelimonas* genus belongs to the *Beijerinckiaceae* family and *Rhizobiales* order. OTU132 showed 93% 16S rRNA gene sequence similarity with *C. fluminis*, which was isolated from the surface water of Hanjiang River in Wuhan, China (Zhang *et al*, 2015); 92% 16S rRNA gene sequence similarity with *B. cytisi*, which was isolated from *Cytisus villosus* nodules (Chahboune *et al*, 2011); and 92% 16S rRNA gene sequence similarity with *B. vignae*, which has been previously described for OTU2 (Gronemeyer *et al*, 2016). Interestingly, *C. fluminis* is capable of degrading cyhalothrin, an organic compound commonly found in pesticides, and *B. cytisi* is a nitrogen-fixing symbiont. These taxa grew optimally at 30 °C and pH 6.0. There were also 16S rRNA gene sequence similarities with many other taxa isolated from plant nodules; however, these were omitted due to low sequence similarity (90%).

OTU842 is most closely related to members of the *Terriglobus* genus and showed 98% 16S rRNA gene sequence similarity with *T. roseus*, which was isolated from the soil of a successional plant community (Eichorst *et al*, 2007); 98% 16S rRNA gene sequence similarity with *T. aquaticus*, which was isolated from an artificial reservoir in South Korea (Baik *et al*, 2013); 98% 16S rRNA gene sequence similarity with *T. albidus*, which was isolated from Namibian semiarid savannah soil (Pascual *et*

al, 2015); and 98% 16S rRNA gene sequence similarity with *T. tenax*, which was isolated from rhizosphere soil of a medicinal plant in Geumsan, Korea (Table 4.3) (Whang *et al*, 2014). *T. roseus* and *T. aquaticus* are aerobic, pink-pigmented, chemo-organotrophic bacteria that grow optimally at pH 6 and 25 °C. These bacteria are non-motile rods and produce EPS. *T. albidus* is the closest relative to *T. aquaticus* and is an aerobic, chemo-organoheterotrophic bacterium that is oxidase-positive, unlike other described species of the genus *Terriglobus*, and demonstrates weak catalase activity. Compared with the already established species of the genus *Terriglobus*, *T. albidus* uses a larger range of sugars and sugar alcohols for growth, lacks α-mannosidase activity, and exhibits optimum growth at a higher temperature range of 28 – 37 °C (Pascual *et al*, 2015). *T. tenax* is a non-pigmented, EPS-producing bacterium, and grows at 15 – 45 °C and pH 3.5 – 7.0, with no pigment nor growth at pH less than 4 or temperatures exceeding 30 °C. Interestingly, these four *Terriglobus* isolates utilise a large range of sugars and sugar alcohols for growth, contain either one or two copies of the 16S rRNA encoding gene, and exhibit a relatively slow doubling time (10 to 15 h at ca. 23°C). Collectively, these characteristics are indicative of an oligotrophic lifestyle.

OTU10 was only distantly related to a single member of the *Spartobacteria* class subdivision 2, showing 90% 16S rRNA gene sequence similarity to *Chthoniobacter flavus*, which was isolated from ryegrass and clover pasture soil in Victoria, Australia (Table 4.3) (Sangwan *et al*, 2004). This taxon is aerobic, yellow-pigmented, non-motile, and grows optimally at 25 °C with many of the saccharide components of plant biomass. *C. flavus* exhibits no growth with glucose under anaerobic conditions. However, low sequence similarity between OTU10 and *C. flavus*, warrants discretion when drawing inferences from the physiology of this isolate.

In summary, taxa comprising the mānuka phyllosphere core microbiome are closely related to bacteria that have been isolated from various habitats around the world, such as waste water, sea water, fresh water, soil, plant root nodules, and the phyllosphere of the Mediterranean *C. cardunculus*. Closely related isolates are members of the *Hymenobacter*, *Sphingomonas*, *Bradyrhizobium*, and *Terriglobas* genera, as well as other families of the *Rhizobiales* order and *Spartobacteria* class (Table 4.3). Closely related isolates are typically pigmented and grow through a wide range of temperatures and pH. Aerobic cellular respiration, positive catalase and

oxidase activity, nitrogen fixation, and methanotrophy are also common characteristics.

Table 4.3: BLAST analysis of 10 core OTU sequences derived from 16S rRNA gene PCR amplicon sequencing.
Relative abundance is averaged across all leaf samples.

OTUID	Taxonomy	Relative abundance (%)	Closest species	Accession number	Identity (%)	Number of matched bp
323	*k_Bacteria;p_Bacteroidetes; c_Cytophagia;o_Cytophagales; f_Cytophagaceae;g_Hymenobacter*	1.1	*Hymenobacter ginsengisoli*	NR_109449.1	99	347/351
			Hymenobacter metalli	NR_108905.1	98	344/350
			Hymenobacter marinus	NR_149279.1	98	343/350
27	*k_Bacteria;p_Proteobacteria; c_Alphaproteobacteria; o_ Sphingomonadales; f_Sphingomonadaceae; g_Sphingomonas*	0.8	*Sphingomonas cynarae*	NR_109167.1	98	343/350
			Sphingomonas xinjiangensis	NR_108386.1	98	342/350
			Sphingomonas insulae	NR_044187.1	98	342/350
2	*k_Bacteria;p_Proteobacteria; c_Alphaproteobacteria; o_ Rhizobiales; f_unknown; g_unknown*	11.5	*Methylocapsa palsarum*	NR_137418.1	96	336/350
			Beijerinckia doebereinerae	NR_116304.1	96	337/352
			Methylocystis bryophila	NR_108474.1	96	335/350
132	*k_Bacteria;p_Proteobacteria; c_Alphaproteobacteria; o_ Rhizobiales;f_unknown; g_unknown*	3.5	*Camelimonas fluminis*	NR_137267.1	93	310/335
			Bradyrhizobium cytisi	NR_116360.2	92	325/352
			Bradyrhizobium vignae	NR_147716.1	92	325/352
11	*k_Bacteria;p_Proteobacteria; c_Alphaproteobacteria;*	4.1	*Sphingomonas kyeonggiensis*	NR_134182.1	99	347/350
			Sphingomonas sanxanigenens	NR_121736.1	99	347/350

o_ Sphingomonadales; f_Sphingomonadaceae; g_ Sphingomonas		*Sphingomonas insulae*	NR_044187.1	99	347/350
k_Bacteria;p_Proteobacteria; c_Alphaproteobacteria; o_ Rhizobiales;f_unknown;g_unknown	3.6	*Bradyrhizobium vignae*	NR_147716.1	92	325/352
k_Bacteria;p_Proteobacteria; c_Alphaproteobacteria; o_ Rhizobiales;f_unknown; g_unknown	10.6	*Bradyrhizobium ganzhouense*	NR_133706.1	93	329/352
		Bradyrhizobium cytisi	NR_116360.2	93	329/352
k_Bacteria;p_Bacteroidetes; c_Cytophagia;o_Cytophagales; f_Cytophagaceae; g_Hymenobacter	1.4	*Hymenobacter marinus*	NR_149279.1	96	337/351
		Hymenobacter flocculans	NR_108904.1	96	335/350
		Hymenobacter ginsengisoli	NR_109449.1	95	335/351
		Hymenobacter metalli	NR_108905.1	95	333/351
k_Bacteria;p_Acidobacteria; c_ Acidobacteria_Gp1; o_Terriglobus;f_Terriglobus; g_Terriglobus	1.0	*Terriglobus roseus*	NR_043918.1	98	343/350
		Terriglobus aquaticus	NR_135733.1	98	342/350
		Terriglobus albidus	NR_137410.1	97	340/350
		Terriglobus tenax	NR_133877.1	97	340/350
k_Bacteria;p_Verrucomicrobia; c_Spartobacteria;o_Spartobacteria_ genera_incertae_sedis; f_Spartobacteria_genera_ incertae_sedis;g_Spartobacteria_ genera_incertae_sedis	2.4	*Chthoniobacter flavus*	NR_115225.1	90	318/352

4.3. Exploring biogeographical patterns in the mānuka phyllosphere microbiome

To address the first question pertaining to biogeography ("Do microbial assemblages differ in different locations, and does biogeography exist in the phyllosphere?") spatial variation in community structure and composition of the mānuka phyllosphere microbiome was explored between each of the sample locations (Table 4.4). First, spatial variation was explored at the phylum level by assessing the relative contribution (%) of each dominant phylum to the overall phyllosphere community among the five sample sites (Figure 4.3) (for full phylum-level analysis, see Supplementary Table 6.1). *Proteobacteria* and *Bacteriodetes* were most abundant at site SL, representing 61.5% and 23.2% sequence reads, respectively (Figure 4.3). *Acidobacteria* were less abundant at sites SL (1.49%) and KU (11.7%), compared to sites HT (26.4%), MK (23.5%), and MV (18.6%) (Figure 4.3). *Verrucomicrobia* and *Firmicutes* were abundant at HT, representing 6.2% and 11.1% sequence reads, respectively (Figure 4.3). *Actinobacteria* were more abundant at sites MV (1.1%) and SL (1.5%), compared to sites HT (0.86%), KU (0.5%), and MK (0.35%) (Figure 4.3). *Planctomycetes* were also most abundant at site MV (1.0%), compared to all other sites (0.4 – 0.8%) (Figure 4.3). To further ascertain the spatial variation between different sample sites in the mānuka phyllosphere microbiome, bacterial communities from all regions were further analysed at the OTU level using the Bray-Curtis dissimilarity index (based on relative abundance) and an unconstrained NMDS ordination analysis. Significant differences in bacterial community structure were detected between samples collected from different regions of New Zealand (Figure 4.4, pairwise PERMANOVA comparing data from all regions, $P < 0.001$). Permutational analyses of Bray-Curtis dissimilarities within individual sample regions further highlighted significant variability in bacterial community structure between individual host trees (PERMANOVA, $P < 0.001$). This inter-individual variability was significant in all regions, with the exception of site HT (PERMANOVA, $P > 0.05$). Spatial variation in bacterial community structure and composition was also detected between branches collected from the same individual sample tree. This intra-individual variation was on average 10% less than inter-individual variation, however. Moreover, both intra- and inter-individual variability in phyllosphere communities was less than the variability

between samples originating from different regions, with samples from the same location generally clustering together in ordinational space (Figure 4.4). A betadisper test was used to further examine the variation of phyllosphere communities originating from different regions. Moreover, the betadisper test was insignificant ($P > 0.05$), highlighting the presence of significant and non-random regional clustering.

Table 4.4: Sample site data. UTM grid coordinates are located in zone 60 and rounded to the nearest 10,000 to conceal exact sample location.

Sample Site	Location	UTM (z60) Northing	UTM (z60) Easting	Elevation (m)
MK	Mohaka, Hawkes Bay	5680000	510000	354.2 ± 5.6
SL	Ohoupo, Waikato	5800000	350000	66 ± 0
MV	Masterton, Wellington	5470000	370000	242.8 ± 3.9
KU	Mamaku, Bay of Plenty	5790000	420000	558.3 ± 3.7
HT	North Eastern Kaimanawa, Taupo	5670000	440000	634.7 ± 2.8

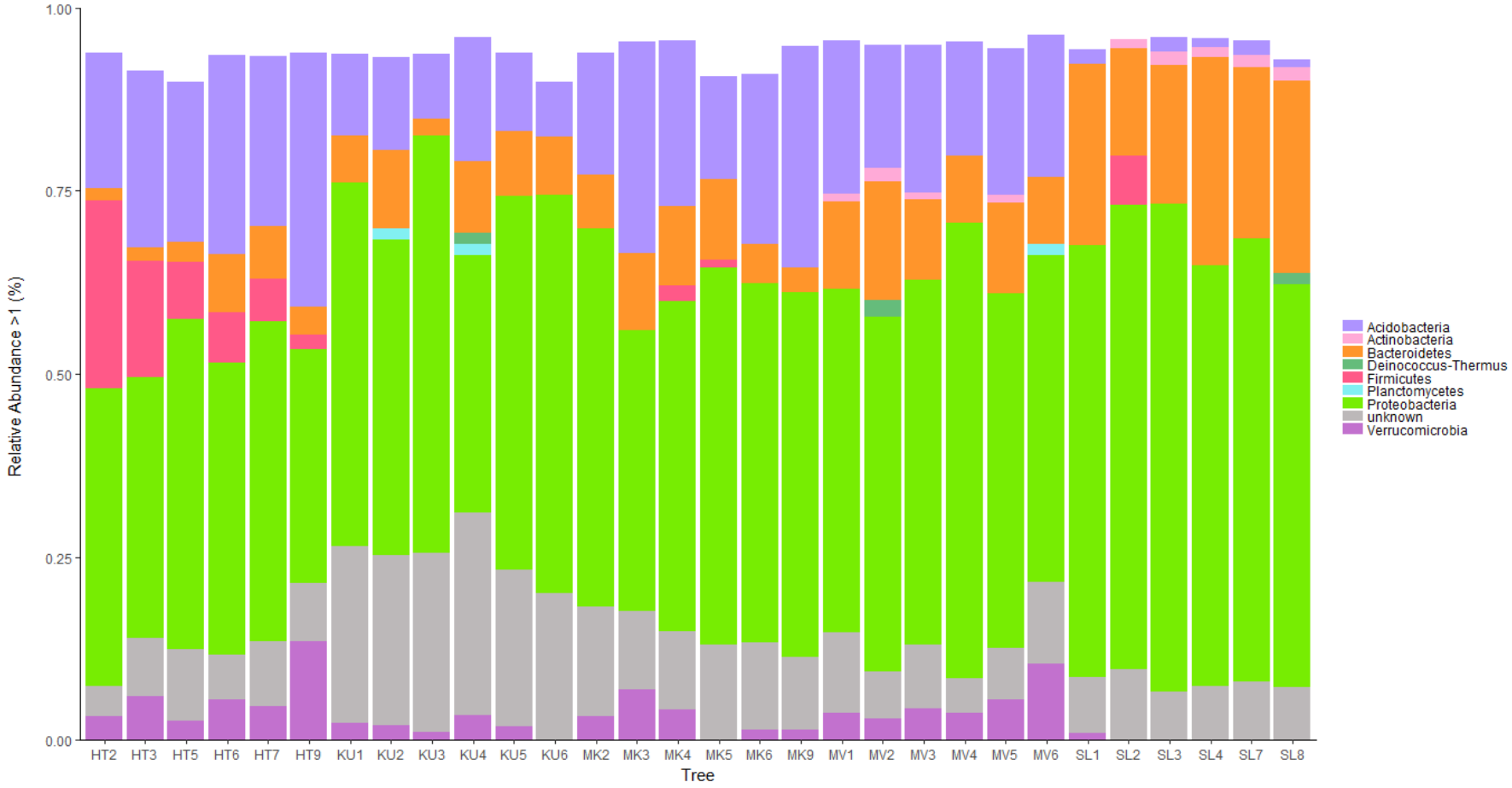

Figure 4.3: Phylum-level composition of the manuka phyllosphere based on the relative abundance (>1%) of 16S rRNA gene sequence reads.

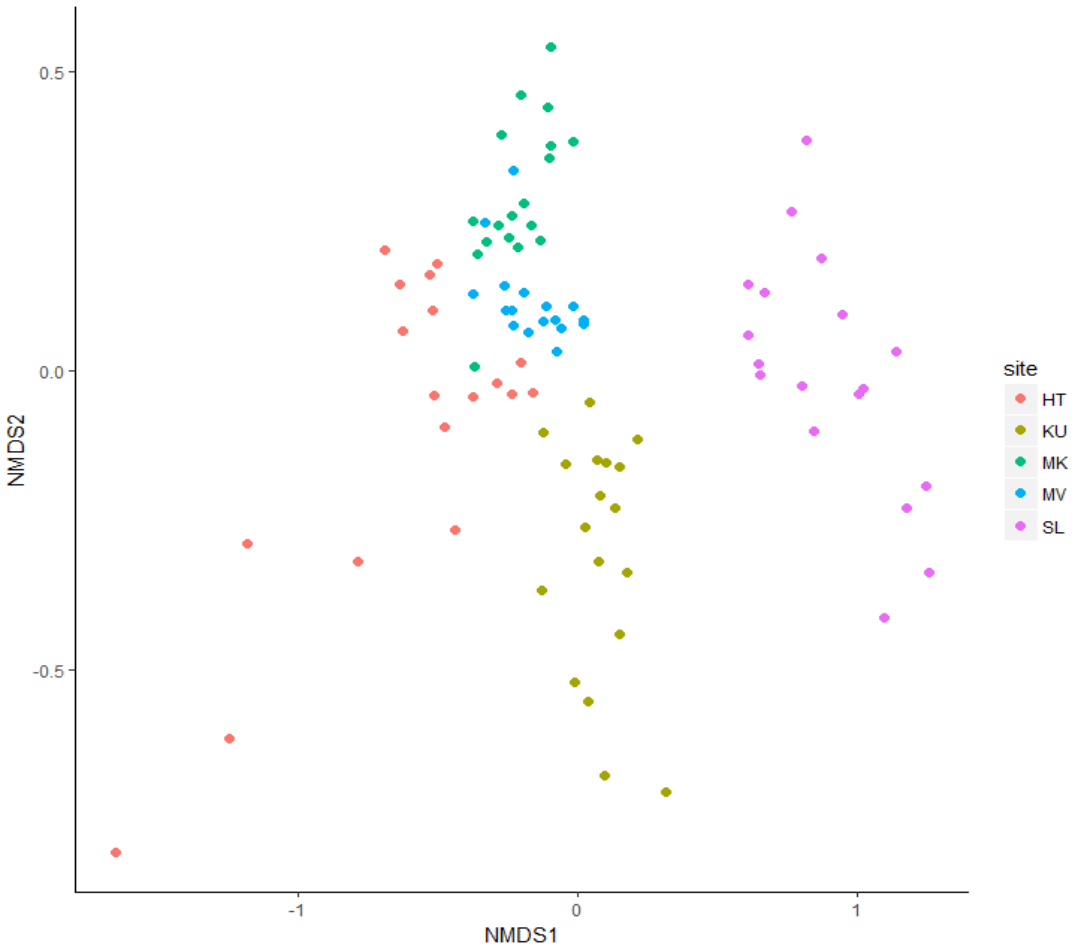

Figure 4.4: Non-metric multidimensional scaling (NMDS) ordination of variation in bacterial community structure in the mānuka phyllosphere. Ordination based on Bray-Curtis dissimilarities among 89 samples. Data points represent individual samples and are coloured according to site. Stress = 0.147. Adonis, P< 0.0001. Betadisper, P > 0.05.

To address the second research question related to microbial biogeography ("To what extent is spatial variation attributable to environmental and biogeographical driving factors?"), analyses were conducted to elucidate the relative influence of geographical separation, climatic heterogeneity, and host tree-related explanatory variables on the regional variation in the mānuka phyllosphere microbiome. The results of these analyses are presented below in their respective sections. See Supplementary Table 6.3 – 6.4 for raw spatial, environmental, and sample tree metadata.

4.3.1. Distance, latitude, longitude, and elevation (space)

The effect of distance on the mānuka phyllosphere microbiome was first assessed. The WGS84 coordinates of each sample site were converted into the universal transverse mercator (UTM) coordinate format, as conducted by recent biogeographical investigations in the magnolia phyllosphere (USA) and fresh water streams (NZ) (Table 4.4) (Lear *et al*, 2013, Stone & Jackson 2016). The UTM coordinate system is a two-dimensional grid-based mapping system which permits easier calculation of the straight-line distance between two points, compared to coordinate systems based on ellipsoidal models of the Earth (Williams 1995). The UTM coordinates were then used to calculate the Euclidean (straight-line) distance between each sample site. A correlational analysis using a Mantel test with the Pearson product-moment correlation coefficient (r_M) was then performed between these Euclidean distances and the Bray-Curtis community dissimilarity index of the entire phyllosphere community. Overall, Mantel analysis revealed community dissimilarity between two sample sites (based on relative abundance) increased when sites are separated by increasing geographical distance ($r_M = 0.19$, $P = 0.01$) (Table 4.5). This linear relationship between geographical distance and community dissimilarity was significant, albeit weak. On average, samples within the same sample site were 34% dissimilar. Sites located 50 – 150 km apart were on average 55% dissimilar. Sites located 150 – 200 km apart were on average 71.9% dissimilar (Figure 4.5). This monotonic relationship was, however, variable when sites were separated by distances exceeding 200 km (Figure 4.5). A second Mantel analysis was then performed between these Euclidean distances and a Bray-Curtis community dissimilarity index based on presence/absence data. Interestingly, phyllosphere community dissimilarity based on the presence and absence of bacterial OTUs showed a clearer relationship with

geographic distance and a slightly higher Mantel correlation coefficient compared to community dissimilarity based on the relative abundance of bacterial OTUs (r_M = 0.37, P = 0.001 and r_M = 0.19, P = 0.01; respectively) (Table 4.5). To further investigate the relative effect of distance on different members of the phyllosphere community, additional Mantel analyses were performed on abundant (>1%) and rare (<1%) taxa (Table 4.5). Distance was found to exhibit a significant correlation with rare taxa (r_M = 0.50, P = 0.001), whereas distance did not exhibit a significant correlation with abundant taxa (r_M = 0.14, P = 0.064) (Table 4.5).

Northing UTM coordinates of each sample site were used to calculate the latitudinal Euclidean distance between each sample site (Table 4.4). A Mantel analysis was performed between these latitudinal Euclidean distances and a Bray-Curtis community dissimilarity index of the entire community. Overall, Mantel analysis revealed a significant relationship, whereby community dissimilarity between two sample sites (based on relative abundance) increased when separated by increasing latitudinal distance (r_M = 0.24, P = 0.002) (Table 4.5). Community dissimilarity between two sample sites (based on presence/absence) also increased when separated by increasing latitudinal distance (r_M = 0.16, P = 0.001) (Table 4.5). Furthermore, community dissimilarity of rare taxa showed a stronger correlation with latitude (r_M = 0.30, P = 0.001) compared to abundant taxa (r_M = 0.23, P = 0.011) (Table 4.5). The longitudinal Euclidean distance between each region was then calculated using the easting UTM coordinates of each sample site (Table 4.4). Mantel analysis revealed a significant relationship, whereby community dissimilarity between two sample sites, based on relative abundance and presence/absence, increased when separated by increasing longitudinal distance (r_M = 0.32, P = 0.002 and r_M = 0.47, P = 0.001, respectively) (Table 4.5). Community dissimilarity of rare taxa also showed a stronger correlation with longitude (r_M = 0.51, P = 0.001) compared to abundant taxa (r_M = 0.3, P = 0.01) (Table 4.5).

Pearson's correlation coefficient tests were implemented between the relative abundance of dominant phyla and spatial parameters, such as latitude and longitude (See Supplementary Table 6.5 for full analysis). The relative abundance of *Bacteroidetes* showed a significant negative correlation with longitude (R = -0.58, P = 0.001). Meanwhile, the relative abundance of *Acidobacteria* showed a significant positive correlation with longitude (R = 0.57, P = 0.001). Although no strong

correlation (R > 0.5, R < 0.5) was found between the relative abundance of dominant phyla and latitude, the relative abundance of *Acidobacteria* and *Verrucomicrobia* did exhibit slightly weaker correlations with latitude (R = -0.43, P = 0.001 and R = -0.46, P = 0.001; respectively).

Table 4.5: Mantel Pearson's product-moment correlation coefficient test performed on Bray-Curtis dissimilarities of the total, abundant, and rare members of the phyllosphere bacterial community and spatial gradients
Total community correlations are based on relative abundance *(A)* and presence/absence *(B)*. Positive correlations greater than 0.5 are underlined. P values less than 0.05 are in bold text.

	Total *(A)*		**Total** *(B)*		**Abundant (> 1%)**		**Rare (< 1%)**	
	Corr	*P test*	*Corr*	*P test*	*Corr*	*P test*	*Corr*	*P test*
Distance	0.19	**0.01**	0.37	**0.001**	0.12	0.06	<u>0.50</u>	**0.001**
Latitude	0.24	**0.002**	0.16	**0.001**	0.24	**0.011**	0.30	**0.001**
Longitude	0.32	**0.002**	0.47	**0.001**	0.31	**0.001**	<u>0.51</u>	**0.001**
Elevation	<u>0.52</u>	**0.001**	0.46	**0.001**	<u>0.58</u>	**0.001**	<u>0.51</u>	**0.001**

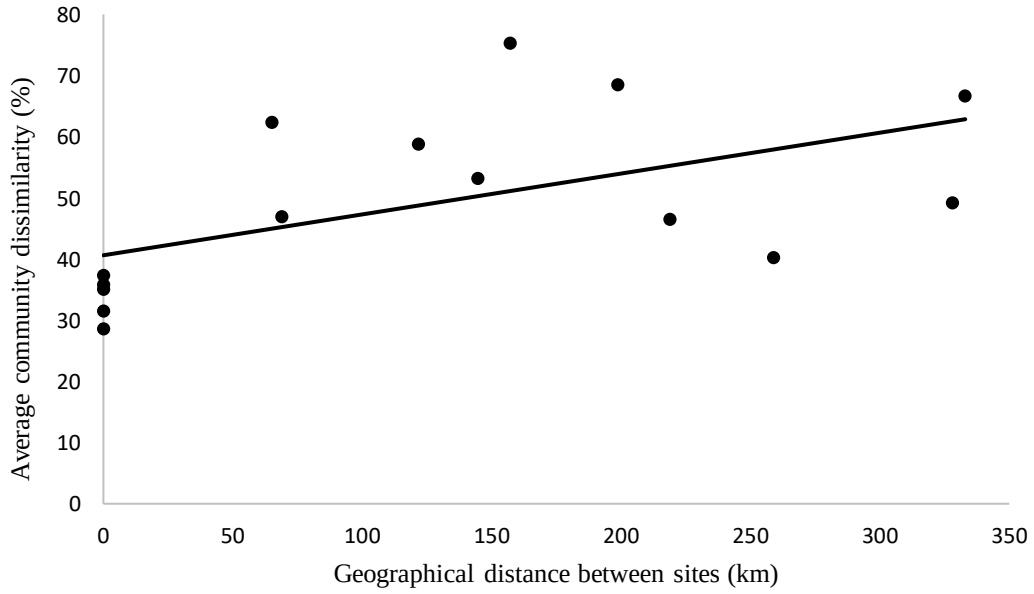

Figure 4.5: Average community dissimilarity based on Bray Curtis index versus geographical distance between samples (km).

The native mānuka forests sampled spanned a range of elevations from 66 m to 634 m above sea level (Table 4.4). Therefore, the effect of elevation on the mānuka phyllosphere microbiome was also assessed. Euclidean distance was calculated from the difference in elevation between each sample site. A correlational analysis using a Mantel test and the Pearson product-moment correlation coefficient (r_M) was then performed between these Euclidean distances and Bray-Curtis community dissimilarity of the entire phyllosphere community (based on relative abundance and presence/absence data). Mantel analysis detected a significant relationship between elevation and community dissimilarity based on relative abundance ($r_M = 0.52$, $P = 0.001$) (Table 4.5) and presence/absence ($r_M = 0.46$, $P = 0.001$). On average, samples that differed in elevation by less than 100 m exhibited 38% dissimilarity. Sites that differed in elevation by 200 – 400 m were on average 55% dissimilar. Meanwhile, sites that differed in elevation by more than 500 m were on average 75% dissimilar (Figure 4.6). In order to further investigate the relative effect of elevation gradients on different members of the phyllosphere community, additional Mantel analyses were performed on abundant (> 1%) and rare (< 1%) taxa. Both rare taxa and abundant taxa significantly correlated with elevation, with abundant taxa showing a slightly stronger correlation ($r_M = 0.58$, $P = 0.001$) compared to rare taxa ($r_M = 0.51$, $P = 0.001$) (Table 4.5). Pearson's correlation coefficient tests were also implemented between the relative abundance of dominant phyla and elevation (See Supplementary Table 6.5 for full analysis). The relative abundance of *Bacteroidetes* and *Proteobacteria* showed a significant negative correlation with elevation ($R = -0.76$, $P = 0.001$; $R = -0.54$, $P = 0.001$, respectively). The relative abundance of *Acidobacteria* meanwhile showed a significant positive correlation with elevation ($R = 0.51$, $P = 0.001$).

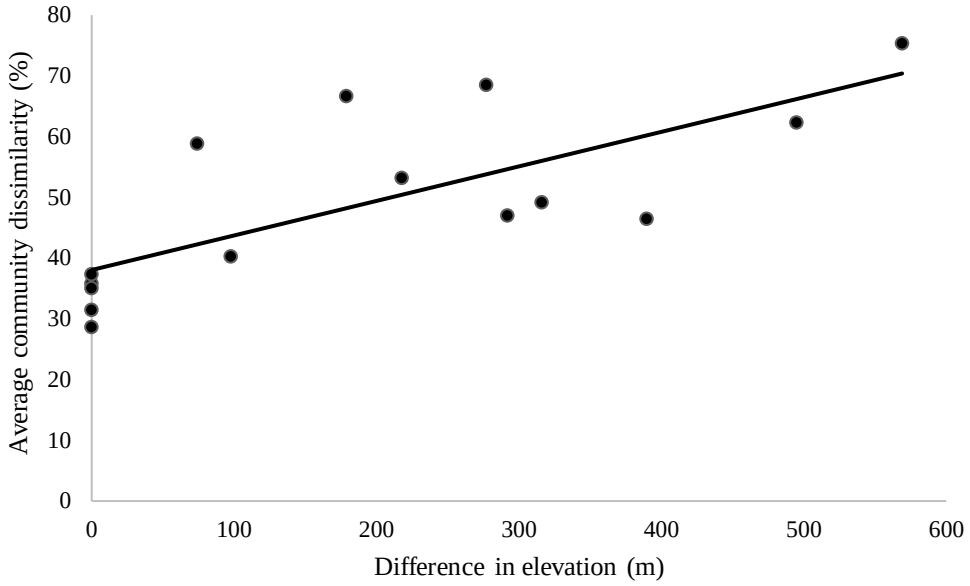

Figure 4.6: Average community dissimilarity based on Bray Curtis index (relative abundance) versus difference in elevation between sample sites (m).

4.3.2. Climatic heterogeneity

The influence of climate on phyllosphere bacterial communities was assessed. Contemporary environmental variables were measured across the 24 hours prior to sampling, and monthly averages were retrieved from the NIWA database (NIWA 2017). For a full list of environmental measurements, see Supplementary Table 6.3. The effect of these environmental variables on the phyllosphere was first examined using a canonical analysis of principal coordinates (CAP). CAP results confirmed the differences in bacterial community structure between samples originating from different geographical locations (Figure 4.7). Permutation tests on CAP revealed differences in community structure between geographical locations were significant (P = 0.001), and permutation pairwise comparisons showed significant differences between all regions. Correlational analyses using Mantel tests and the Pearson product-moment correlation coefficient (r_M) with Bray-Curtis community dissimilarity metrices (based on relative abundance) and Euclidean distance metrices of environmental variables, were then used to ascertain the relative influence of environmental variables on community dissimilarity at the OTU level (Table 4.6). Significant correlations were found between community dissimilarity and contemporary environmental variables

measured across the 24 hours prior to sampling. Of these, average night temperature and day-night temperature differential were the environmental variables most strongly correlated with community dissimilarity (r_M = 0.67, P = 0.001; r_M = 0.59, P = 0.001, respectively) (Table 4.6). Significant, albeit weaker, correlations were found between community dissimilarity and average day temperature (r_M = 0.24, P = 0.001), relative humidity (r_M = 0.20, P = 0.001), and photosynthetically active radiation (r_M = 0.12, P = 0.007) (Table 4.6). Significant correlations were also found between community dissimilarity and monthly averages of environmental variables pertaining to each sample location, of which monthly precipitation was the most strongly correlated (r_M = 0.52, P = 0.001) (Table 4.6). Monthly cloud cover and sun hours also showed significant correlations with community dissimilarity (r_M = 0.48, P = 0.001; r_M = 0.45, P = 0.001, respectively) (Table 4.6). Much weaker correlations were detected with monthly humidity (r_M = 0.33, P = 0.001) and pressure (r_M = 0.18, P = 0.001) (Table 4.6). Monthly temperature and wind showed no correlation (P > 0.05) (Table 4.6).

Pearson's correlation coefficient tests were implemented between the relative abundance of dominant phyla and environmental parameters (See supplementary Table 6.5 for full analysis). The relative abundance of *Proteobacteria* showed a significant positive correlation with day-night temperature differential (R = 0.56, P = 0.001), precipitation (R = 0.50, P = 0.001), and a significant negative correlation with night temperature (R = -0.50, P = 0.001). The relative abundance of *Acidobacteria* showed a significant positive correlation with night temperature (R = 0.75, P = 0.001), monthly precipitation (R = 0.64, P = 0.001), and monthly sun (R = 0.67, P = 0.001), as well as a significant negative correlation with day-night temperature differential (R = -0.58, P = 0.001), monthly cloud (R = -0.67, P = 0.001) and monthly humidity (R = -0.57, P = 0.001). The relative abundance of *Bacteroidetes* showed a significant positive correlation with day temperature (R = 0.53, P = 0.001), day-night temperature differential (R = 0.71, P = 0.001), monthly precipitation (R = 0.76, P = 0.001), and monthly humidity (R = 0.53, P = 0.001), as well as a significant negative correlation with night temperature (R = -0.60, P = 0.001) and monthly sun (R = -0.53, P = 0.001). The relative abundance of *Verrucomicrobia* showed a significant positive correlation with night temperature (R= 0.57, P = 0.001) and a significant negative correlation with day-night temperature differential (R = -0.50, P = 0.001). There was no significant correlation between the relative abundance of *Firmicutes* and any environmental

parameters.

To further explore the differential effect of environmental variables on the different members of the phyllosphere community, Mantel analyses were performed on abundant (> 1%) and rare (< 1%) component taxa. Compared to rare taxa, abundant taxa were more strongly correlated to the day-night temperature differential (r_M = 0.65, P = 0.001), night temperature (r_M = 0.78, P = 0.001), and monthly precipitation (r_M = 0.58, P = 0.001) (Table 4.6). In contrast, rare taxa were more significantly correlated with relative humidity (r_M = 0.43, P = 0.001), day temperature (r_M = 0.39, P = 0.001), photosynthetically active radiation (r_M = 0.40, P = 0.001), as well as monthly cloud cover (r_M = 0.61, P = 0.001) and sun hours (r_M = 0.60, P = 0.001) (Table 4.6) (Figure 4.9). No significant correlation was found between monthly wind and rare or abundant taxa (P > 0.05) (Table 4.6).

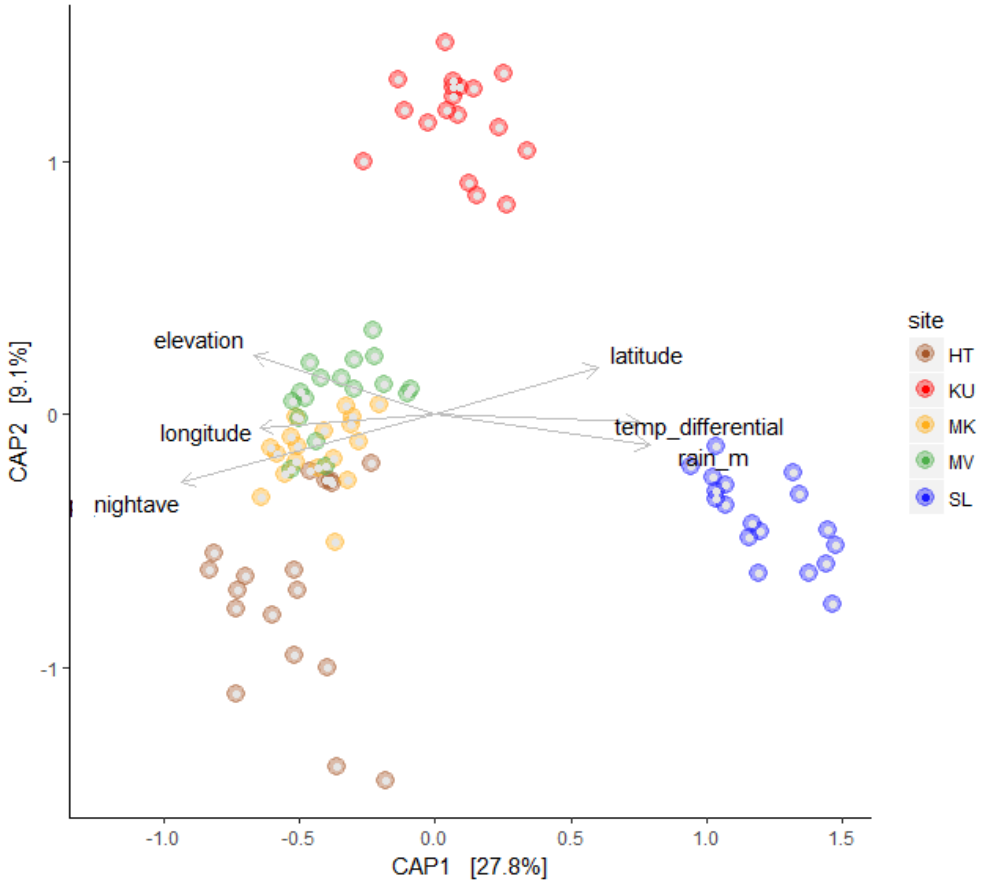

Figure 4.7: Canonical analysis of principal coordinates (CAP) of mānuka phyllosphere bacterial communities. CAP of bacterial communities from all regions were obtained with the vegan package using Bray-Curtis dissimilarity matrix based on relative abundance. The two axes collectively represent 36.9% of variation.

Table 4.6: Mantel Pearson's product-moment correlation coefficient test performed on Bray-Curtis dissimilarities of the total, abundant, and rare members of the phyllosphere bacterial community and environmental variables
Positive correlations greater than 0.5 are underlined. P values less than 0.05 are in bold text. The symbol (m) represents monthly averages obtained from the NIWA database (NIWA, 2017).

	Total		**Abundant (>1%)**		**Rare (< 1%)**	
	Corr	*P test*	*Corr*	*P test*	*Corr*	*P test*
Day temperature	0.24	**0.001**	0.24	**0.002**	0.39	**0.001**
Night Temperature	0.67	**0.001**	0.78	**0.001**	0.59	**0.001**
Day/night temperature differential	0.59	**0.001**	0.65	**0.001**	0.53	**0.001**
Relative humidity	0.20	**0.001**	0.19	**0.005**	0.43	**0.001**
Photosynthetically active radiation	0.12	**0.007**	0.08	**0.001**	0.40	**0.001**
Temperature (m)	0.06	0.09	0.05	0.24	0.30	**0.001**
Precipitation (m)	0.52	**0.001**	0.58	**0.001**	0.51	**0.001**
Wind (m)	-0.06	0.85	0.11	0.06	-0.05	0.67
Pressure (m)	0.18	**0.001**	0.17	**0.007**	0.39	**0.001**
Cloud (m)	0.48	**0.001**	0.54	**0.001**	0.61	**0.001**
Humidity (m)	0.33	**0.001**	0.34	**0.003**	0.50	**0.001**
Sun (m)	0.45	**0.001**	0.47	**0.001**	0.60	**0.001**

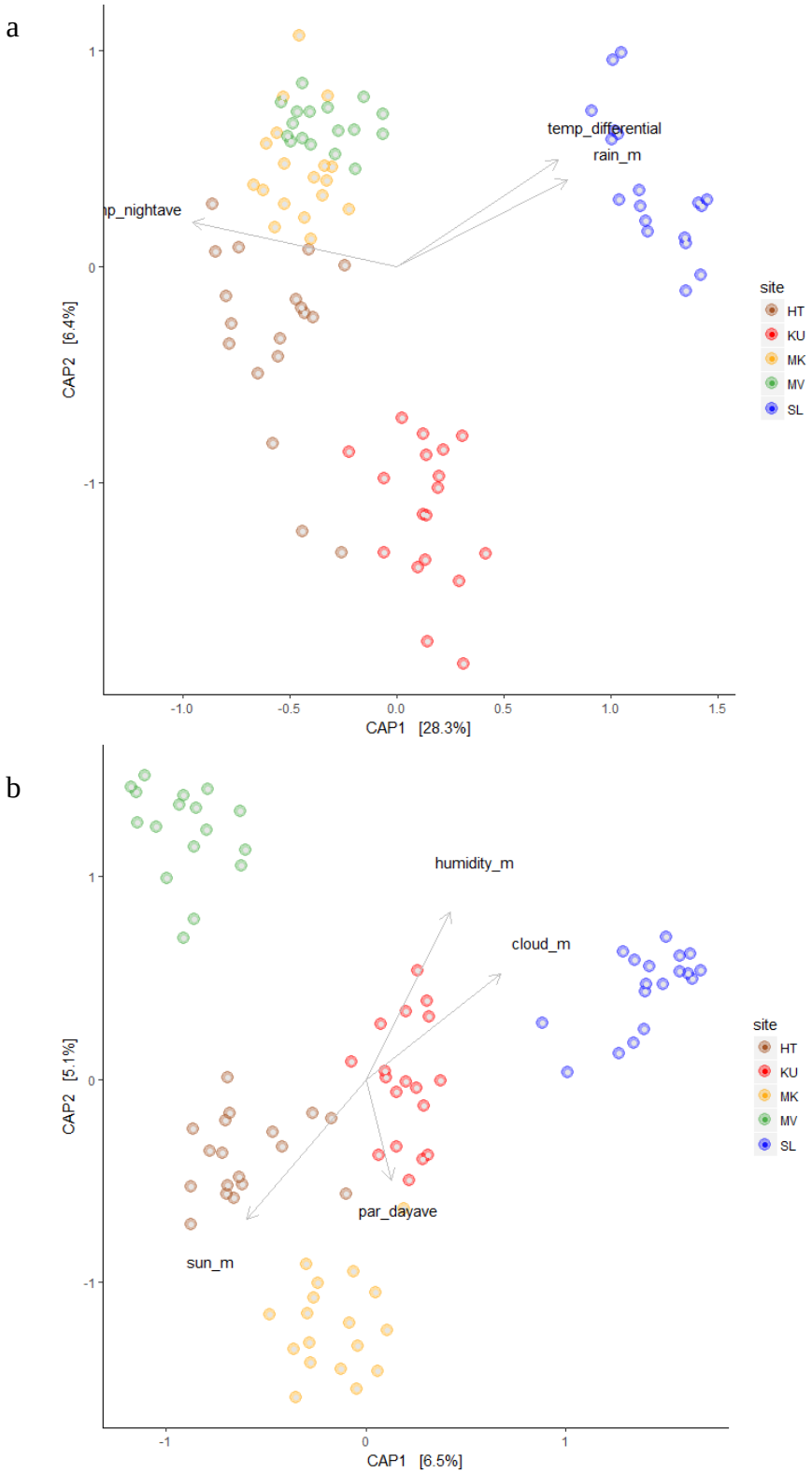

Figure 4.8: Canonical analysis of principal coordinates (CAP) of mānuka phyllosphere bacterial communities separated into abundant (> 1%) (a) and rare (< 1%) (b) component taxa. Obtained using Bray-Curtis dissimilarity matrix based on relative abundance. The two axes collectively represent 34.7% and 11.6% of variation for abundant and rare component taxa, respectively.

4.3.3. Host tree characteristics

Specific host tree-related characteristics have been previously reported to shape phyllosphere bacterial communities; therefore, the relative influence of such characteristics on community structure was also assessed (Kembel *et al*, 2014, Laforest-Lapointe *et al*, 2016a). The above-ground measurements collected in this study were largely related to plant stature-related traits (Table 4.7). Tree height and basal diameter were used as a rough estimate of overall tree size and age (Cochran 1979, Kenzo *et al*, 2015, Steppe *et al*, 2011). The height of sampled mānuka trees ranged from 1.8 m – 4.1 m and the basal diameter of sampled mānuka trees ranged from 1.5 cm – 24.8 cm (Table 4.7). Tree height is a stable, direct measurement, often used to predict the age of a given tree species (Cochran 1979). Although there are few records regarding the above-ground growth performance of mānuka, healthy mānuka is estimated 3.0 m tall at 4 – 5 years, 4.0 m tall at 10 years, 6.0 m tall at 20 years, and 8.0 – 10.0 m tall at 40 years (Boffa Miskell Ltd 2017). Branch height and aspect were also measured to permit further analysis of intra-individual spatial variation. The influence of these measurements on community structure was assessed at the OTU level using Mantel tests and the Pearson product-moment correlation coefficient (r_M) (Table 4.8). A significant correlation was found between Bray-Curtis community dissimilarity (based on relative abundance) and tree height, whereby trees of different heights tended to host different communities compared to trees of similar heights ($r_M = 0.19$, $P = 0.001$) (Table 4.8). Moreover, this correlation was significant for both rare ($r_M = 0.16$, $P = 0.001$) and abundant ($r_M = 0.18$, $P = 0.001$) taxa (Table 4.8). Rare taxa showed a significant correlation with branch height ($r_M = 0.13$, $P = 0.001$), which was not observed in the total community or in abundant taxa ($P > 0.05$) (Table 4.8). Tree diameter and branch aspect did not exhibit a significant correlation with total, rare, nor abundant phyllosphere community structure at the OTU level ($P > 0.5$) (Table 4.8). In addition, Pearson's and Spearman's correlation analyses were performed between these plant stature-related traits and relative abundance of dominant phyla (See Supplementary Table 6.5 for full analysis). Only weak correlations were identified, however. The relative abundance of *Bacteroidetes* showed a significant positive correlation with tree height ($R = 0.26$, $P = 0.15$). The relative abundance of *Proteobacteria* showed a significant positive and negative correlation with branch height ($R = 0.22$, $P = 0.04$) and tree diameter ($R = -0.22$, $P = 0.04$), respectively. The

relative abundance of *Acidobacteria* showed a significant negative correlation with tree height ($R = -0.22$, $P = 0.045$). The relative abundance of *Verrucomicrobia* showed a significant positive correlation with tree diameter ($R = 0.42$, $P = 0.001$). Lastly, the relative abundance of *Firmicutes* showed a significant negative correlation with branch height ($R = -0.28$, $P = 0.008$). No significant correlation was found between branch aspect and the relative abundance of dominant phyla ($P > 0.05$).

Table 4.7: Plant stature-related measurements of sampled mānuka trees

Tree ID	Elevation (m)	Tree height (m)	Tree diameter (cm)	Branch heights (m)	Branch aspect (°)
MK2	357	2.7	2.3	1.03, 1.55, 1.83	0, 135, 145
MK3	354	2.7	6.6	2.04, 2.25, 2.54	71, 111, 120
MK4	358	3.18	4.0	1.8, 2.2 2.8	0, 45, 80
MK5	356	1.82	1.5	1.28, 1.4, 1.62	0, 69, 128
MK6	357	2.52	3.0	1.8, 2.0, 2.2	0, 139, 359
MK9	343	2.86	3.55	2.25, 2.53, 2.6	55, 60, 75
SL1	66	4.0	8.9	1.75, 1.8, 2.4	91, 306, 339
SL2	66	3.9	8.5	2.1, 2.25, 2.35	110, 118, 160
SL3	66	4.1	12.8	2.0, 2.05, 2.15	93, 98, 123
SL4	66	3.3	4.4	1.25, 1.4, 1.7	130, 140, 150
SL7	66	3.3	4.5	1.9, 2.0, 2.2	12, 155, 173
SL8	66	2.2	3.7	1.4, 1.7, 1.8	9, 75, 313
MV1	245	3.3	9.7	1.85, 1.95	235, 270
MV2	237	2.65	3.9	2.3, 2.3, 2.35	10, 85, 110

MV3	240	4.0	4.4	2.6, 2.8, 3.3	195, 195, 195
MV4	247	3.4	7.6	2.6, 2.6, 2.7	0, 130, 160
MV5	246	2.85	3.5	2.05, 2.1, 2.35	280, 290, 300
MV6	242	2.9	4.1	2.0, 2.45, 2.7	0, 120, 140
KU1	561	2.3	1.5	0.8, 1.15, 1.5	300, 300, 310
KU2	562	3.8	3.6	1.7, 2.1, 3.1	10, 175, 260
KU3	562	2.5	2.3	1.9, 2.0, 2.1	65, 150, 210
KU4	555	4.0	7.35	2.1, 2.3, 2.5	70, 235, 330
KU5	555	3.7	15.75	1.9, 2.1, 2.4	120, 160, 210
KU6	555	2.5	2.4	1.55, 1.7, 2.0	15, 30, 280
HT2	638	2.8	3.5	1.2, 1.2, 1.8	25, 200, 205
HT3	633	2.7	6.3	1.1, 1.3, 1.5	210, 225, 250
HT5	632	2.8	7.2	0.9, 1.0, 1.6	10, 145, 230
HT6	632	2.6	8.8	1.1, 1.55, 1.7	45, 45, 115
HT7	638	3.1	17.6	1.0, 1.2, 1.5	75, 80, 205
HT9	635	3.3	24.75	0.6, 1.1, 1.2	25, 340, 345

Table 4.8: Mantel Pearson's product-moment correlation coefficient test performed on Bray-Curtis dissimilarities of the total, abundant, and rare members of the phyllosphere bacterial community and host tree-related characteristics. Total community Bray-Curtis dissimilarity index is based on relative abundance. P values less than 0.05 are in bold text.

	Total		**Abundant (> 1%)**		**Rare (< 1%)**	
	Corr	*P test*	*Corr*	*P test*	*Corr*	*P test*
Branch height	-0.03	0.71	0.04	0.28	0.13	**0.017**
Branch aspect	0.02	0.29	0.02	0.36	-0.03	0.68
Tree height	0.19	**0.001**	0.18	**0.02**	0.16	**0.01**
Tree diameter	-0.02	0.59	-0.04	0.60	-0.03	0.71

4.3.4. Relative importance of spatial, climatic, and host tree characteristics

Variation partitioning was used to further resolve the explanatory power of spatial, climatic, and host tree-related characteristics in the regional variation of the mānuka phyllosphere (Figure 4.9). A considerable proportion (58%) of heterogeneity in bacterial community structure could not be explained by any combination of explanatory variables; this may reflect the effect of unmeasured variables and highlights the need for future studies with a larger assortment of metadata. The remaining 42% of variation was explained by spatial, climatic, and host tree-related characteristics (Figure 4.9). Tree height was significantly correlated with variability in total bacterial community structure (Table 4.8). However, the relative effect of tree height, as well as other host tree-related characteristics, was very small, explaining only 1% of the observed variability (Figure 4.9). In support of previous correlational analyses (Figure 4.5 & Table 4.5), variation partitioning revealed geographical space had a significant and independent impact on the variability of bacterial community structure (Figure 4.9). However, as with host tree-related characteristics, the effect of geographical space was relatively small, explaining only 3% of the variability in bacterial community structure independent of other explanatory variables (Figure 4.9). Climatic variables had the greatest impact on the observed heterogeneity in

phyllosphere community structure, independently explaining 9% of the variability (Figure 4.9). Moreover, the combined effect of both climatic and spatial explanatory variables accounted for a further 21% of the total variation in bacterial community structure (Figure 4.9). Climatic and host tree explanatory variables accounted for a further 1% of bacterial community variability (Figure 4.9). Furthermore, host tree-related characteristics, climate and spatial explanatory factors collectively accounted for 8% of the total variation (Figure 4.9). Spatial gradients, such as elevation and latitude, tend to exhibit autocorrelation with other environmental gradients (Wang *et al*, 2011). Within the current research project, latitude correlated with temperature, wind, and cloud, while elevation correlated with temperature and monthly precipitation. Therefore, the independent impact of any one of these explanatory variables is difficult to accurately ascertain using this approach.

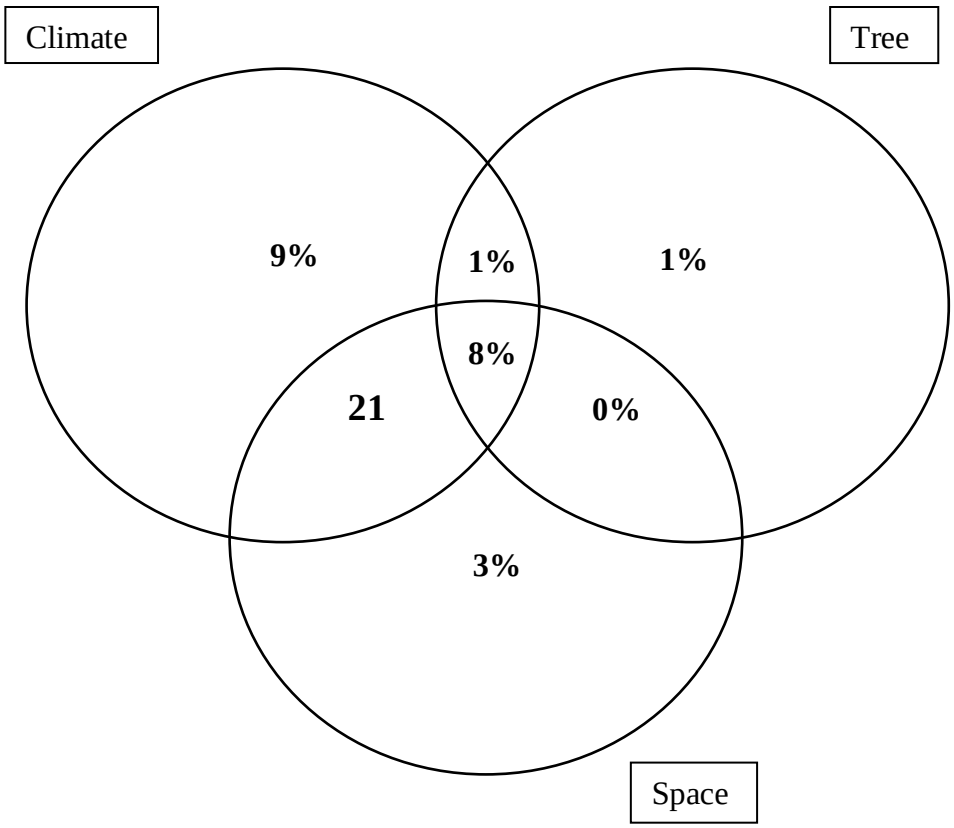

Figure 4.9: Venn diagram providing a visual representation of variance related to three groups of explanatory variables: geographic space, climate, and host tree attributes. Results of partial regression analysis show the proportion of variation attributable to each individual (A-C) and combination (D-G) of explanatory variables. Residuals = 0.58.

4.4. The predicted mānuka phyllosphere metagenome

In order to further ascertain the habitat-specificity of bacterial communities in the phyllosphere and surrounding soil, a PICRUSt analysis was conducted to infer general functionality of the entire phyllosphere and soil communities based on taxonomic identity (Langille *et al,* 2013). Using the relative abundance of KEGG orthologues, a Bray-Curtis dissimilarity index was calculated for the predicted metabolic profiles. An NMDS ordination analysis highlighted significant differences in the overall metabolic potential of soil and phyllosphere communities (Figure 4.10, PERMANOVA comparing data from all regions, $P < 0.001$). In addition, a betadisper test was used to further examine the similarity of metabolic profiles originating from each habitat. The betadisper test was insignificant ($P > 0.05$) and thus highlighted the distinct clustering of metabolic profiles was non-random and habitat-specific. This functional prediction highlighted the presence of genes involved in many different metabolic pathways, such as amino acid, cofactor, lipid, carbohydrate, and energy metabolism, as well as xenobiotic degradation, membrane transport, and biosynthesis of secondary metabolites. The relative abundance of functional pathways and enrichment ratios were calculated for phyllosphere and soil communities. Important gene KEGG orthologues (KO) are also described. See Supplementary Table 6.6 for full results.

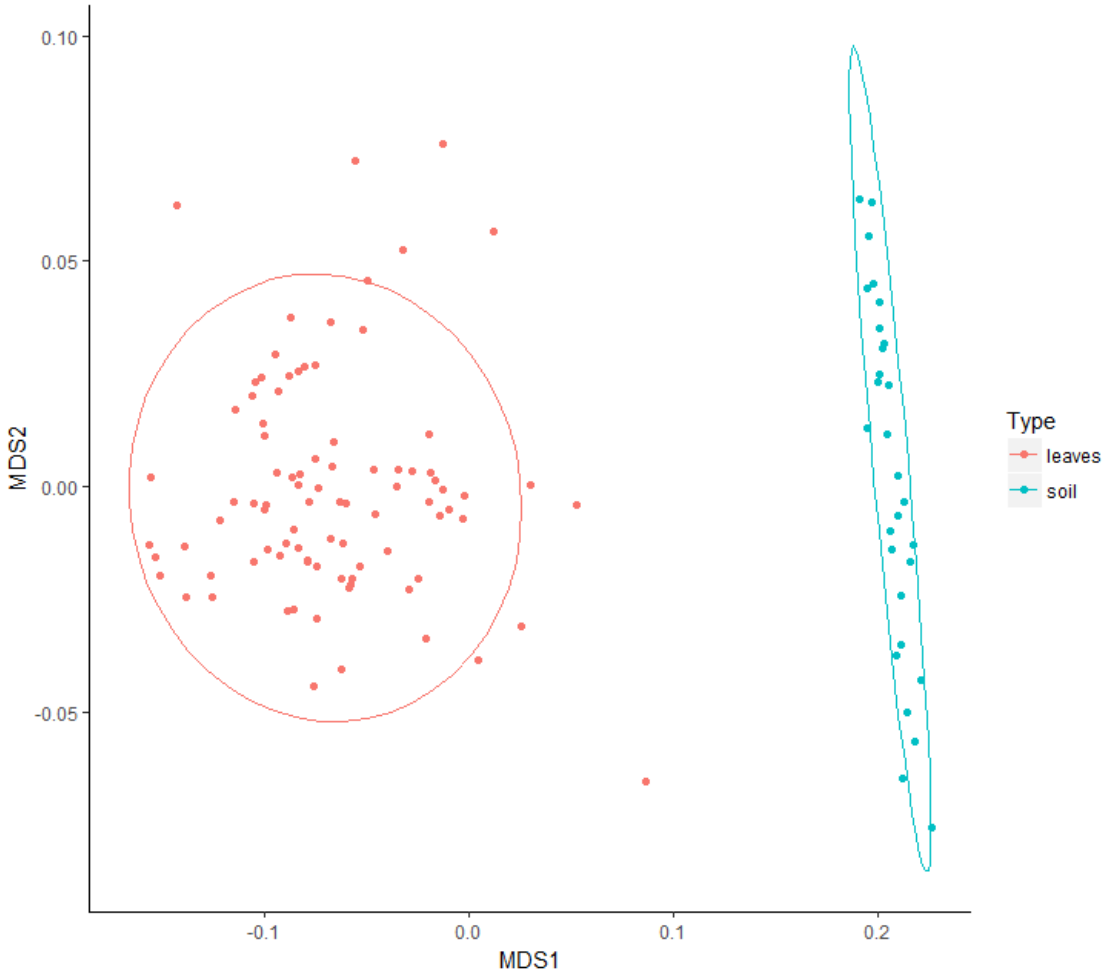

Figure 4.10: Analysis of microbial metabolic potential between different sample types - phyllosphere and soil. NMDS based on Bray-Curtis similarities calculated for metabolic profiles predicted with PICRUSt. Stress = 0. 067. Adonis, $P < 0.001$. Betadisper, $P > 0.05$.

4.4.1. Phyllosphere-enriched KEGG pathways

Genes involved in the biosynthesis of pigments, calcium signalling pathways, carbohydrate digestion and absorption, as well as the degradation of some xenobiotics, were enriched in the phyllosphere communities compared to soil (Figure 4.11).

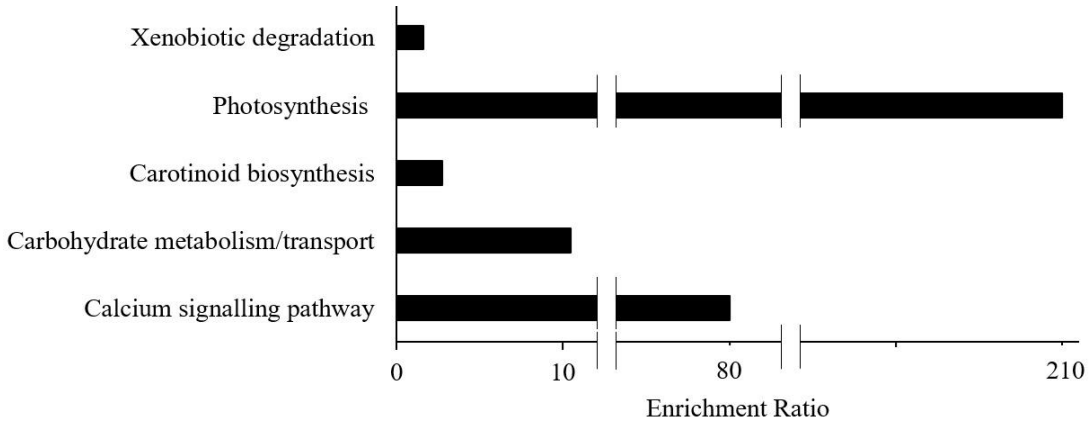

Figure 4.11: Enrichment ratio of KEGG pathways in the phyllosphere predicted metabolic profile. Enrichment ratio is calculated by dividing the relative abundance of the KEGG pathway in the phyllosphere by the relative abundance of the KEGG pathway in soil.

Calcium signalling – Genes involved in calcium signal transduction pathways were enriched 80-fold within the phyllosphere compared to soil (Figure 4.11). Calcium (Ca^{2+}) is an important secondary messenger in plants and has been implicated in a number of leaf processes, such as the regulation of stomatal guard cell closure (Quintana *et al,* 2001). Increasing calcium levels may provide the foundation for plant/microbe interactions and has previously been found to influence phyllosphere bacterial community structure (Hunter *et al,* 2010, Vadassery & Oelmüller 2009).

Energy metabolism- Genes involved in carbohydrate transport and metabolism were enriched 10-fold in the phyllosphere compared to soil (Figure 4.11). Moreover, within the carbohydrate metabolism pathway, the relative abundance of the pyruvate kinase gene (K00873) was 2-fold greater in the phyllosphere compared to soil (Figure 4.12). Pyruvate kinase catalyses the final step of glycolysis to yield a single molecule of pyruvate and ATP (Veith *et al,* 2013). The predicted presence of the pyruvate kinase

gene suggests the glycolysis metabolic pathway may be utilised in the phyllosphere through which glucose is anaerobically catabolised. In addition, the relative abundance of the pyruvate formate lyase gene (K00656), an enzyme that catalyses the anaerobic synthesis of pyruvate from formate and acetyl-CoA, was 12-fold greater in the phyllosphere compared to soil (Figure 4.12). Genes attributable to the utilisation of one-carbon (C1) compounds as alternative carbon and energy sources were also predicted in the phyllosphere. Methanol is a by-product of plant cell wall metabolism and is catalysed by methylotrophic bacteria (Fall & Benson 1996). The presence of such metabolic ability on the mānuka leaf surface was highlighted by the identification of the structural gene for the pyrroloquinoline quinone-linked enzyme methanol dehydrogenase (MxaF) (K14028). MxaF catalyses the oxidation of methanol to formaldehyde (McDonald & Murrell 1997). The relative abundance of MxaF was 20-fold greater in the phyllosphere compared to soil (Figure 4.12). The relative abundance of soluble methane monooxygenase (MMO) gene orthologue (K08684) was 8-fold greater in the phyllosphere compared to soil (Figure 4.12). MMO is an enzyme that is capable of oxidizing the C-H bond in methane and other one-carbon compounds (Feig & Lippard 1994). The predicted presence of the MMO gene in the phyllosphere is highly suggestive of methane oxidation (Delmotte *et al,* 2009). Large metabolic diversity is well known to permit survival in stressful oligotrophic environments where sources of energy are limited (Delmotte *et al,* 2009). A high degree of metabolic adaptiveness may therefore permit phylloepiphytic bacteria to utilise various, low-concentrated substrates on the leaf surface.

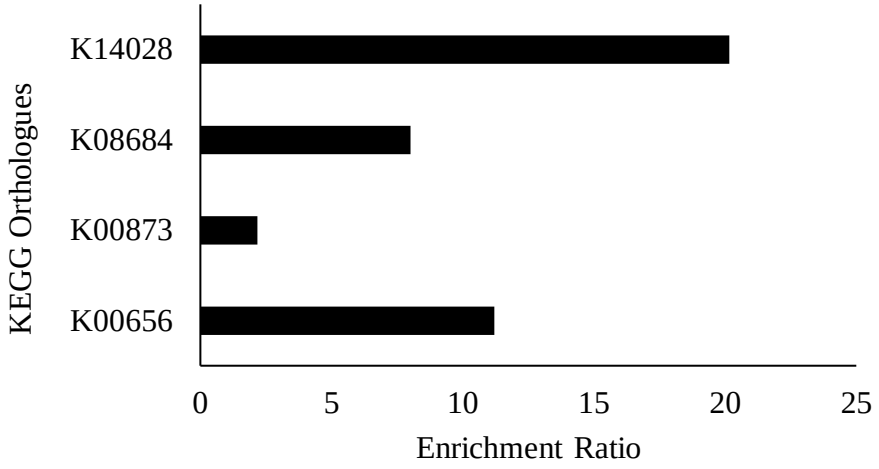

Figure 4.12: Enrichment ratio of KEGG orthologues (KO) in the predicted metabolic profile of the phyllosphere. Methanol dehydrogenase structural gene (MxaF) (K14028), methane monooxygenase gene (MMO) (K08684), pyruvate kinase gene (K00873), and the pyruvate formate lyase gene (K00656). Enrichment ratio is calculated by dividing the relative abundance of the KEGG orthologue in the phyllosphere by the relative abundance of the KEGG orthologue in soil.

Adaptation to UV exposure - Genes involved in photosynthesis were enriched 208-fold in the phyllosphere compared to soil (Figure 4.11). These predicted photosynthetic pathways included antenna complex proteins porphyrin and chlorophyll metabolism as well as, which are the main light-absorbing components in photosynthetic bacteria (Kühlbrandt 1995). Within the phyllosphere, there was also a 3-fold enrichment of genes involved in carotenoid biosynthesis (Figure 4.11). Carotenoids are a family of yellow to red pigments that assist light-harvesting complexes in photosynthetic bacteria by enhancing the absorption of radiant energy. Carotenoids also offer protection against UV radiation and cellular protection against reactive oxygen species for non-phototrophic bacteria (Armstrong 1997, Sundin & Jacobs 1999, Tian & Hua 2010).

Xenobiotics – Genes involved in the degradation of xenobiotics, such as atrazine and xylene, were enriched 2-fold in the phyllosphere compared to soil (Figure 4.11). Atrazine is a herbicide globally used to target broadleaf and grass weeds (Wenk *et al,* 1997). The ability of phyllosphere bacteria to degrade atrazine has not been previously reported. Xylene is a volatile pollutant with toxic characteristics. Xylene-degrading potential has been previously reported in phyllosphere bacteria (Sangthong *et al*, 2016).

In contrast, genes involved in the biosynthesis of type II polyketides and glycans, cell adhesion molecules, focal adhesion, and the regulation of actin cytoskeleton, cell cycle, endocytosis, and phagosomes were enriched in soil compared to the phyllosphere (Figure 4.13).

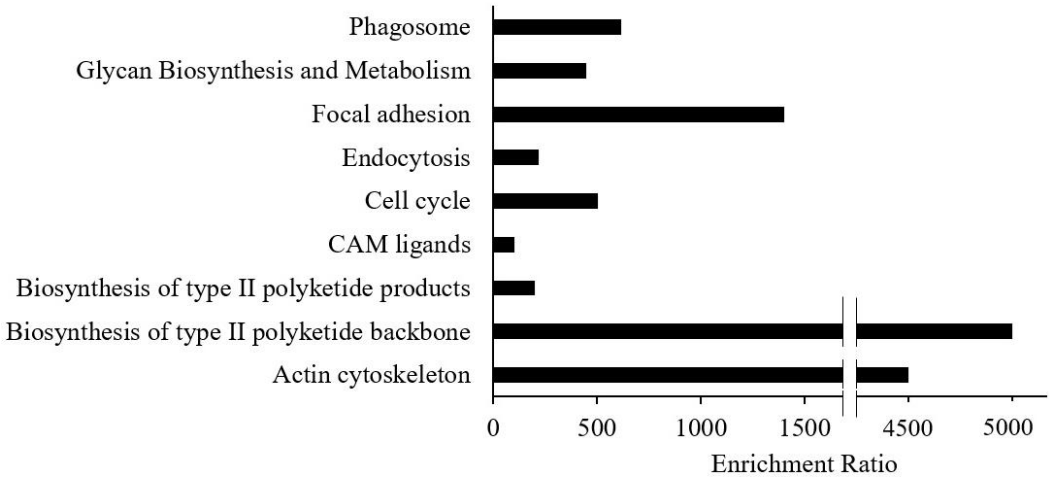

Figure 4.13: Enrichment ratio of KEGG pathways in soil predicted metabolic profile. Enrichment ratio is calculated by dividing the relative abundance of the KEGG pathway in soil by the relative abundance of the KEGG pathway in the phyllosphere.

4.4.2. KEGG pathways shared by phyllosphere and soil communities

Similarities in the inferred metagenomic composition of phyllosphere and soil bacterial communities were also detected. Genes belonging to membrane transport, oxidative phosphorylation, replication and repair, and two-component system signal transduction systems were among the most abundant functional pathways predicted in both phyllosphere and soil (Figure 4.14).

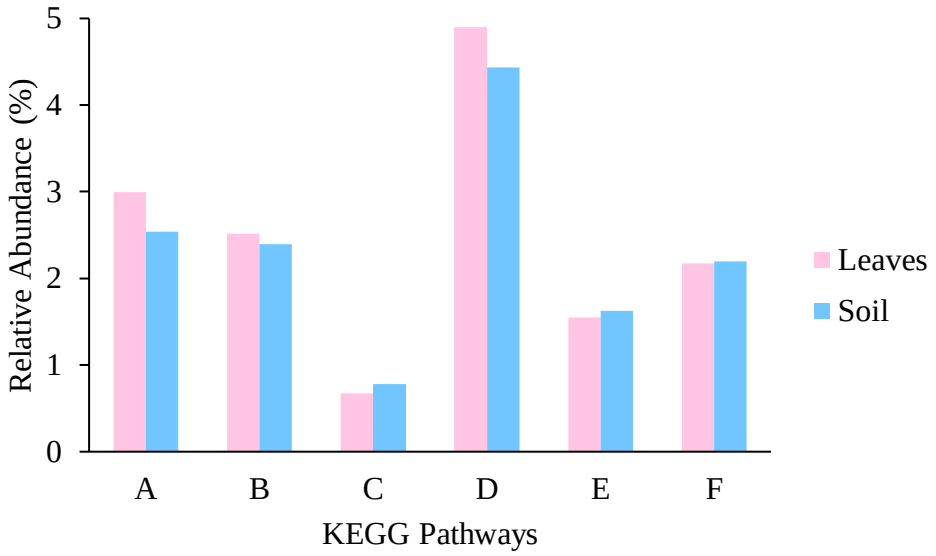

Figure 4.14: Relative abundance (%) of predicted KEGG pathways in phyllosphere and soil predicted metabolic profiles. A) ABC transporters, B) DNA replication and combination proteins, C) nitrogen fixation and metabolism, D) other membrane transporters, E) oxidative phosphorylation, F) two-component signal transduction system.

Membrane transport - Membrane transporter genes, such as those belonging to the OOP family, were predicted at a relative abundance of 4.4% and 5.0% in the soil and phyllosphere, respectively (Figure 4.14). Genes belonging to the ABC transporter superfamily were predicted at a relative abundance of 2.5% and 3.0% in the soil and phyllosphere, respectively (Figure 4.14). ABC transporters are involved in the export and import of many substrates, such as sugars, amino acids, lipids, and polysaccharides. Furthermore, ABC transporters may be involved in EPS secretion, epiphytic survival and the regulation of bacteria-plant interactions (Young & Holland 1999).

Replication and repair – Genes involved in DNA replication and repair pathways were predicted at a relative abundance of 2.4% and 2.5% in the soil and phyllosphere, respectively (Figure 4.14). These replication and repair pathways included DNA replication, base excision repair, homologous recombination, mismatch repair, and nucleotide excision repair systems (Sinha & Hader 2002). Gene orthologues for cold shock protein C (K03704), heat shock protein 60 (K04077), and DNA protection during starvation family protein (K04047) were identified in the soil and phyllosphere at equal relative abundances (Figure 4.15). These proteins function as

RNA chaperones, affecting the stability of stress-related gene transcripts under low- and high-temperature conditions and inhibits DNA replication under low nutrient stress conditions (Chaikam & Karlson 2010).

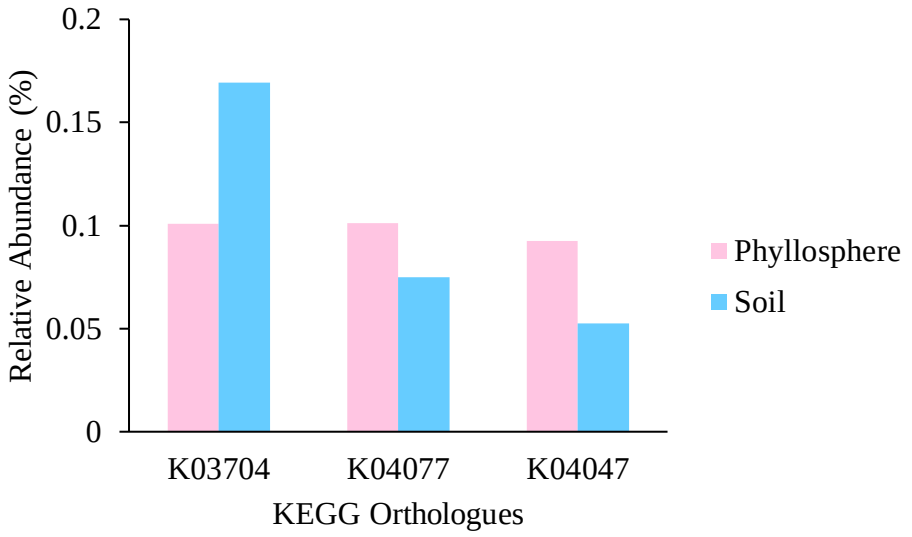

Figure 4.15: Relative abundance (%) of KEGG orthologues in phyllosphere and soil predicted metabolic profiles. The cold shock protein C gene (K03704), heat shock protein 60 gene (K04077), and DNA protection during starvation family protein gene (K04047).

Nitrogen fixation and metabolism – Genes involved in nitrogen fixation and metabolism were predicted at a relative abundance of 0.67% and 0.78% in phyllosphere and soil, respectively (Figure 4.14). Gene orthologues of glutamine synthetase (K01915), which catalyses the metabolism of nitrogen through the formation of glutamine from glutamate and ammonia, and nitrogenase iron protein (nifH) (K02588), an enzyme responsible for nitrogen fixation and regarded as a common marker of diazotrophic bacteria, were predicted in soil and phyllosphere metabolic profiles (Burbano *et al*, 2011, Delmotte *et al*, 2009, Fürnkranz *et al*, 2008, Knief *et al*, 2010, Lambais *et al*, 2017). Nitrogen fixation is an important part of the nitrogen cycle often performed in close association with plants; however, proteins indicative of nitrogen fixation have not always been consistently detected in the phyllosphere (Delmotte *et al*, 2009, Knief *et al*, 2012).

Oxidative phosphorylation – Genes involved in oxidative phosphorylation genes were predicted at a relative abundance of 1.5% and 1.6% in phyllosphere and soil, respectively (Figure 4.14) (Garvie 1980, Nakayama *et al,* 1971). An abundance of genes involved in the oxidative phosphorylation pathway is indicative of aerobic cellular respiration.

Two-component signal transduction – Genes involved in two-component signal transduction systems were predicted at a relative abundance of 2.2% in both phyllosphere and soil metabolic profiles (Figure 4.14). Two-component signal transduction systems are typically composed of a sensor histidine kinase, which receives input stimuli and phosphorylates response regulators; this system enables bacteria to alter their cellular physiology in response to external changes in the environment (Stock *et al,* 2000). Bacteria that occupy habitats with frequent environmental fluctuations possess more histidine kinases and response regulators (Capra & Laub 2012). It is therefore interesting that genes involved in two-component signal transduction systems were predicted in the phyllosphere, a habitat renowned for rapid physicochemical fluctuations.

Other interesting KO's - The most prevalent KO in the phyllosphere was the iron complex outer–membrane receptor protein (K02014), predicted at a relative abundance of 0.47% (Figure 4.16). In soil, K02014 was predicted at a relative abundance of 0.34% (Figure 4.16). In addition, the relative abundance of the gene orthologue for ferritin (K02217), an iron-storage protein, was greater in the phyllosphere than in soil (Figure 4.16). Iron is important for many biological processes, including photosynthesis, biological nitrogen fixation, methanogenesis, respiration and DNA synthesis. Ferritin regulates iron homeostasis and is therefore important for bacterial growth and controlling responses to oxidative stresses (Andrews *et al,* 2003). The OmpA porin gene (K03286) was also predicted in phyllosphere and soil (Figure 4.16). OmpA was the most abundant protein in a proteogenomic investigation of the *A.thaliana* phyllosphere. Moreover, OmpA may have important implications for epiphytic survival through involvement in bacterial recognition, adhesion, virulence and evasion of host plant defences as well as solute transport (Delmotte *et al,* 2009). Within this metabolic prediction, however, the OmpA gene was not common in the phyllosphere and was detected at a relative abundance of 0.00034% (Figure 4.16).

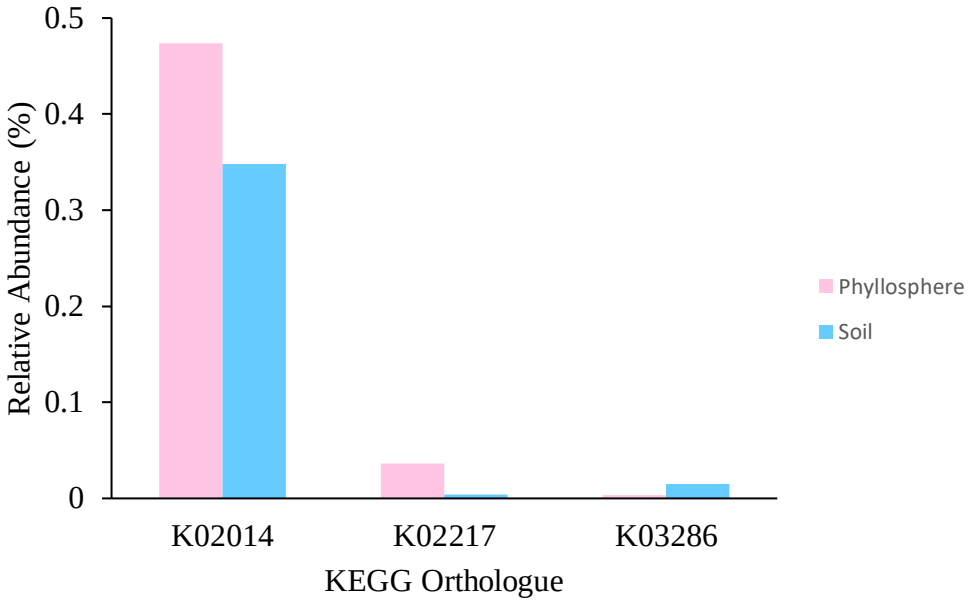

Figure 4.16 : Relative abundance (%) of KEGG orthologues in phyllosphere and soil predicted metabolic profiles. The iron complex outer–membrane receptor protein gene (K02014), ferritin gene (K02217), and OmpA porin gene (K03286).

In summary, the predicted metagenomic composition of entire phyllosphere and soil bacterial communities revealed differences and similarities in the relative abundance of KEGG pathways and KOs. Genes involved in the biosynthesis of carotenoids, carbohydrate metabolism and transport, degradation of xenobiotics, and calcium signalling were enriched in the phyllosphere compared to soil. Genes involved in the biosynthesis of type II polyketides and glycans as well as the regulation of actin cytoskeleton, CAM ligands, cell cycle, endocytosis, focal adhesion and phagosomes were meanwhile enriched in soil compared to the phyllosphere. Genes involved in membrane transporters, repair and replication pathways, nitrogen fixation, oxidative phosphorylation and two component signal transduction systems were common in both phyllosphere and soil predicted metabolic profiles.

4.5. Discussion

The phyllosphere is a unique and relatively understudied environment that is gaining increasing interest in recent years. Little is known regarding the assembly and driving factors of phyllosphere community structure and composition (Laforest-Lapointe *et al,* 2016a, Laforest-Lapointe *et al,* 2017, Vorholt 2012). Moreover, the phyllosphere presents a fascinating environment in which many theories pertaining to microbial ecology can be explored and tested, including the core microbiome concept and microbial biogeography. The focus of this explorative study was to provide the first characterisation of the mānuka phyllosphere microbiome.

The mānuka phyllosphere core microbiome

Over the last decade, an increasing number of investigations have sought to identify core microbiomes across various habitats (Cho & Blaser 2012, Reveillaud *et al,* 2014, Shade & Handelsman 2012, Turnbaugh *et al,* 2008). This study provides the first evidence that mānuka shares a unique core phyllosphere microbiome, comprising ca. 40% of the total phyllosphere community, even when separated by moderate geographical distances of up to 300 km (Figure 4.1). At the phylum level, the core microbiome consisted of *Alphaproteobacteria*, *Bacteriodetes*, *Acidobacteria,* and *Verrucomicrobia*. At the genus level, core taxa belonged to *Hymenobacter*, *Sphingomonas, Terriglobus, Spartobacteria genera incertae sedis*, and unclassified genera within the *Rhizobiales* order (Table 4.3). Core taxa were closely related to bacteria previously isolated from various habitats around the world, such as wastewater, sea water, fresh water, and soil (identified through BLAST) (Table 4.3) (Belova *et al,* 2013, Hoang *et al,* 2013, Kang *et al,* 2016). Interestingly, many of these relatives were also isolated from plant-associated environments, including the phyllosphere of the Mediterranean *C. cardunculus* and plant root nodules (Gronemeyer *et al,* 2016, Tala *et al,* 2013). Moreover, evidence for a core microbiome in the mānuka phyllosphere is in accordance with a recent publication that also detected a core microbiome which comprised 42% of the temperate phyllosphere communities in Quebec forests (Laforest-Lapointe *et al,* 2016a). These authors similarly identified members of *Hymenobacter* and *Sphingomonas* genera, as well as members of the *Acidobacteriaceae* family and *Rhizobiales* order.

According to the core microbiome concept, the observation that the mānuka phyllosphere is dominated by few phylloepiphytic bacteria strongly suggests these core taxa are the functional drivers of leaf-associated bacterial communities (Shade & Handelsman 2012, Vandenkoornhuyse *et al,* 2015). The putative identities of core taxa revealed physiological characteristics (sourced from the associated primary literature through BLAST) that may permit survival on the leaf surface and facilitate intimate host-plant associations. Numerous physiological adaptations that promote phylloepiphytic survival during extreme fluctuations of temperature, UV radiation, and nutrient availability were identified. These included catalase and oxidase enzymes, pigmentation, EPS secretion, and the formation of extracellular matrices (Lushchak 2001, Sinha & Hader 2002). EPS secretion, alongside the fact that mānuka leaves are evergreen and New Zealand's temperate climate may offer sufficient water availability year-round, suggests core phylloepiphytic bacteria may form biofilms on the surface of mānuka leaves. Biofilm formation offers numerous benefits to phylloepiphytic bacteria, such as protecting against abiotic stresses as well as enhancing adhesion and microbe-microbe interactions (Costerton *et al,* 1995, Morris *et al,* 1997). Large metabolic diversity is also well known to permit survival in oligotrophic environments, such as the phyllosphere (Delmotte *et al,* 2009). Interestingly, closely related isolates utilise a large range of sugars, sugar alcohols, and one-carbon compounds for growth (Belova *et al,* 2013, Pascual *et al,* 2015, Tamas *et al,* 2013). *Rhizobiales* isolates, related to the most abundant core taxa (OTU 2), are aerobic methanotrophs and metabolise methane. Proteins required for methylotrophy belong to the *Methylobacterium* genus. Although members of the *Methylobacterium* genus were not identified in the core microbiome, *Methylobacterium* are common colonisers of many phyllosphere communities and were relatively abundant in the phyllosphere of mānuka (Delmotte *et al,* 2009, Knief *et al,* 2010). Methanotrophy and methylotrophy may, therefore, be common metabolic functions in the mānuka phyllosphere microbiome. These findings suggest a high degree of metabolic adaptiveness may permit core phylloepiphytic bacteria to utilise various, low-concentrated substrates on the leaf surface.

The mānuka phyllosphere core microbiome may directly benefit the health of individual mānuka trees and higher ecosystem trophic levels. *Sphingomonas* have been shown to elicit various plant-protective mechanisms, such as directly promoting the

health of host plants by reducing the growth of pathogenic *Pseudomonas* strains (Innerebner *et al*, 2011). Furthermore, *Sphingomonas* are one of the most common colonisers in the phyllosphere of various host plants around the word and are a member of the mānuka phyllosphere core microbiome (Table 4.3) (Kim *et al*, 1998). Many closely related isolates also demonstrate the ability to fix nitrogen, produce novel sphingans, and degrade both herbicides and pollutants (Tala *et al*, 2013). The leaf surface is known to adsorb significant amounts of pollutants, therefore, the phyllosphere microbiome may play an important role in atmospheric phytoremediation (Nowak *et al*, 2006, Wei *et al*, 2017, Weyens *et al*, 2015).

Collectively, these findings suggest the mānuka phyllosphere microbiome comprises a specific microbial community that may have co-evolved with their host plant to survive on the leaf surface. Understanding the ecological attributes of core taxa will allow us to build a more predictive understanding of how phyllosphere bacterial communities vary across space, time, and in response to anthropogenic changes. Furthermore, the putative identities of the mānuka core microbiome reveal interesting physiological characteristics that may provide agricultural and industrial benefits in the future.

Biogeography of the mānuka phyllosphere microbiome

Detailed knowledge regarding the community assembly and driving factors of the phyllosphere is currently lacking (Kembel *et al*, 2014). If phyllosphere bacterial communities were primarily the result of passive or stochastic colonisation events from the surrounding environment, a significant overlap in shared bacterial taxa between the phyllosphere and the surrounding surface soil would be expected. In contrast, the results of this study highlight significant differences between phyllosphere and soil bacterial communities. Differences were detected in bacterial community composition and structure, whereby a large proportion of taxa were habitat-specific and found exclusively in either the phyllosphere or soil (Figure 4.2). This observation is further supported by the fact that core taxa are essentially absent in soil (Table 4.2). As suggested by previous studies, the few taxa that were shared between the mānuka phyllosphere and surface soil are likely the result of inconsequential events, such as leaf fall (Kim *et al*, 2012). Together, these results suggest that the mānuka phyllosphere microbiome is not a random consortium of microbial taxa.

The results of this study reveal that the mānuka phyllosphere microbiome exhibits significant variation in community structure between different native forests across the North Island of New Zealand. Regional variation in community structure was greater than inter-individual variation within any single region; moreover, no significant difference in alpha diversity (i.e. richness) was found between different regions. The detection of such spatial variation in the presence, absence, and relative abundance of dominant and rare component taxa at both course and fine taxonomic levels provides evidence in support of distinct microbial biogeographical patterning (Figure 4.3 – 4.4 & Table 4.5). Elucidating the primary driving factor of biogeography in microbial communities across various environments has been of increasing interest to microbial ecologists over the last decade (Martiny *et al*, 2006). Within the mānuka phyllosphere, biogeography was driven by a combination of spatial, climatic, and host tree-related explanatory variables (Figure 4.9). These findings indicate that phyllosphere microbial assemblages can exhibit both environmental segregation and biogeographic provincialism (Hanson *et al*, 2012, Rosindell *et al*, 2011).

The extent to which geographic distance and associated dispersal constraints impact the biogeography of microorganisms is largely unknown throughout many environments and conflicting results have emerged from the few biogeographical studies of the phyllosphere (Finkel *et al*, 2011, Redford *et al*, 2010). Moreover, evidence in support of a distance-decay phenomenon in the phyllosphere has only been previously reported on two separate occasions (Finkel *et al*, 2012, Stone & Jackson 2016). Interestingly, the results of the current study reveal the mānuka phyllosphere exhibits a general relationship with distance, whereby phyllosphere communities become increasingly dissimilar when separated by increasing geographical space (Figure 4.5). This indicates that the composition of the phyllosphere is not only different among locations but also that this variation is spatially autocorrelated (Soininen *et al*, 2007). The decay of community similarity with distance is a prediction of neutral ecological theories and emphasises the importance of dispersal in the assembly of bacterial communities (Rosindell *et al*, 2011, Soininen *et al*, 2007, Wiens 2011). Dispersal limitation can arise from past environmental conditions and geological events. Within New Zealand, volcanic activity in the Rotorua and Taupo region has formed a number of volcanic mountains, lakes and plateaus, such as the Mamaku plateau and Kaimanawa ranges where KU and HT sample sites are located,

respectively (Leonard *et al*, 2010); the Waikato region, where sample site SL is located, is a lowly-elevated basin enclosed by greywacke hills (Edbrooke 2005); the Hawkes Bay region, including sample site MK, is renowned for large limestone ridges (Lee *et al*, 2011); and a line of greywacke ranges lies northeast from Wellington and sample site MV (Begg & Johnston 2000). Interestingly, phyllosphere communities from site HT exhibited reduced inter-specific variation and diversity than expected. This observation may reflect large geographical isolation of the HT region. Diversification of mānuka trees and phyllosphere communities, owing to geographical isolation and genetic drift over the course of New Zealand's complex geological history, may have, therefore, played a role in the formation of such microbial biogeographical patterns in the mānuka phyllosphere microbiome.

The decay of community similarity with change in elevation was also detected within the mānuka phyllosphere, whereby bacterial communities sampled from similar elevations appeared to be phylogenetically clustered (Figure 4.6). As with distance-decay along a geographical distance, distance-decay with changing elevation likely reflects the importance of dispersal as well as different environmental selective pressures at each elevation. From assessing the relative abundance of major phyla, it was observed that *Bacteroidetes* and *Proteobacteria* were diminished in regions with high elevations while *Acidobacteria* were increased at high elevation regions (Supplementary Table 6.5). Although elevational patterns of diversity for plants and animals are well established, how microbial diversity and community structure varies across elevational gradients is still elusive. Moreover, how host-associated microbiomes vary with elevational gradients is even less understood as macroorganisms and microorganisms appear to exhibit contrasting patterns in elevational diversity (Wang *et al*, 2011). A number of studies have investigated elevational diversity gradients of fungal communities in the phyllosphere, however few have conducted similar studies for phylloepiphytic bacteria (Vacher *et al*, 2016). Previous research has so far demonstrated that microbial communities generally exhibit either a monotonic decrease in taxon richness with increasing elevation or that biodiversity peaks at intermediate elevations (Bryant *et al*, 2008, Enrico *et al*, 2016). Interestingly, the bacterial richness of the mānuka phyllosphere microbiome tended to increase with elevation from ca. sea level to 550 m, and then significantly diminished at ca. 640 m. This observation, although not statistically significant, could be indicative

of increased biodiversity at intermediate elevations. However, without rigorous systematic sampling along a continuous elevational gradient, it is difficult to ascertain the extent to which phyllosphere communities are affected by elevation solely with the results of the current study.

Environmental factors are known to affect phyllosphere microbial communities, these include UV exposure, temperature fluctuations, and water availability (Finkel *et al*, 2011, Kadivar & Stapleton 2003). Furthermore, taxonomic differences have been identified in the phyllosphere of 'tropical', 'subtropical', and 'temperate' trees (Kim *et al*, 2012, Lambais *et al*, 2006, Redford *et al*, 2010). Consistent with such evidence, many environmental variables in the current study significantly correlated with overall bacterial community dissimilarity based on the relative abundance of bacterial OTUs (Table 4.6). Of these, average night temperature, day-night temperature differential, and monthly precipitation appeared to elicit the largest influence (Figure 4.7). Sample site SL had the lowest average night temperature (9.7°C) and the greatest day-night temperature differential (12.6°C) compared to all other regions (11.3 – 14.5°C and 2.3 – 9.3 °C, respectively) (Supplementary Table 6.3). Bacteria vary considerably in their temperature responses; for each bacteria there is a minimum, maximum and optimum temperature at which cellular processes and growth can occur (Russell 2003). Temperature fluctuations may, therefore, influence phyllosphere community composition and structure between different climatic regions (Jackson & Denney 2011). Precipitation increases free moisture availability for phylloepiphytic bacteria (Lindow & Brandl 2003). The sample region with the greatest monthly precipitation was also SL (ca. 110 mm) (Supplementary Table 6.3). Interestingly, SL had the greatest average relative abundance of *Proteobacteria* (Supplementary Table 6.1). This observation is supported by a previous study that reported an increase of *Proteobacteria* in the phyllosphere after rainfall (Jackson *et al*, 2006). Furthermore, phyllosphere communities from site SL exhibited more distinct phylogenetic clustering compared to communities originating from other sites (Figure 4.4 & Figure 4.7). That site SL was also distinguishable from other sites based on temperature and water availability suggests these environmental variables may be important physicochemical drivers of bacterial community structure in the phyllosphere. However, as many environmental variables are irrevocably linked, such as precipitation and cloud cover, it is difficult to ascertain the individual impact of each

physicochemical parameter. Within the confines of this study, climate heterogeneity was most closely related to differences in bacterial community structure compared to spatial or host tree-related explanatory variables alone (Figure 4.9). Collectively, climatic and spatially determined gradients explained the largest proportion of variation observed in the phyllosphere. These findings suggest that spatially structured gradients in environmental variables, such as precipitation and temperature, are important in shaping the bacterial communities of the mānuka phyllosphere. However, owing to the large proportion of unexplained variance (58%), it remains to be seen whether spatially structured gradients in environmental variables are the most influential driver of the mānuka phyllosphere microbiome.

Resident versus transient bacterial members of the mānuka phyllosphere

Microbiomes are dynamic in time and space, and although investigating bacterial communities as a single entity is inherently important, recent studies suggest the need to consider such communities as assemblages of diverse bacteria that likely adopt different persistence strategies (Finkel *et al*, 2012, Stone & Jackson 2016). The mānuka phyllosphere communities sampled in this study likely represent a combination of resident taxa, which live permanently on leaves, as well as transient taxa, which are introduced temporarily to the phyllosphere from the atmosphere, rainwater, and contact with animal dispersal vectors (Lindow & Brandl 2003, Vorholt 2012, Whipps *et al*, 2008). Therefore, the mānuka phyllosphere was segregated into abundant (>1%) and rare (<1%) bacterial taxa as an attempt to separate established community members from transient taxa. It should be noted, however, that rare taxa may not necessarily represent transient taxa play and their role in phyllosphere microbiome function should not be underestimated (Jousset *et al*, 2017). Interestingly, abundant and rare phyllosphere community members were differentially affected by environmental and biogeographical driving factors (Table 4.5 – 4.6, Table 4.8, & Figure 4.8).

Rare taxa exhibited a significant relationship with spatial factors, such as geographical distance, longitude, and branch height, whilst abundant taxa did not (Table 4.5 & Table 4.8). In conjunction with previous studies, the observation of such spatial patterns suggests neutral forces, such as microbial dispersal, are more important in driving the communities of transient or rare taxa in the phyllosphere than abundant

taxa (Stone & Jackson 2016). This finding is supported by previous observations that microbial life history traits can create differences in dispersal potential between bacterial taxa (Nemergut *et al,* 2013). *Betaproteobacteria*, for example, have been previously reported to exhibit a potential dormancy mechanism that may increase resistance to environmental stressors and thus permit a highly successful dispersal capacity (Darcy *et al,* 2011). In accordance with this, the distance-decay relationship reported in the phyllosphere of *Tamarix* tree species by Finkel *et al,* (2012) was identified in rare *Betaproteobacteria* taxa only. In contrast, taxa with partial dependence on suitable hosts for completion of their life cycles are predicted to exhibit poorer dispersal capacities due to constraints on survival and population growth (Bissett *et al,* 2010). Bacteria that are identified to possess poorer dispersal capacities often include members of nitrogen-fixing *Rhizobiaceae* and *Bradyrhizobiaceae* families, which are abundant in the mānuka phyllosphere and members of the core microbiome (Table 4.3) (Bissett *et al,* 2010). Such differences in life history traits may consequently cause differences in microbial dispersal between abundant and rare taxa in the mānuka phyllosphere microbiome. The fact that core taxa are present across all mānuka samples suggests these taxa are likely selected for by other niche-based selective pressures.

Interestingly, the environmental parameters that elicited the largest correlation with overall community structure (night temperature, day-night temperature differential and monthly precipitation) were more strongly correlated with abundant taxa compared to rare taxa (Table 4.6). For example, the relative abundance of *Acidobacteria* was higher in regions with higher average night temperature and a decreased day-night temperature differential (Supplementary Table 6.5). This observation is supported by previous investigations that have identified high abundances of *Acidobacteria* in the phyllosphere of tropical trees (Kim *et al*, 2012). While rare taxa also correlated with temperature and precipitation, rare taxa were more strongly influenced by other environmental variables, such as relative humidity, monthly sun hours, and monthly cloud cover (Table 4.6). Given the possibility that abundant and rare taxa components may represent resident and transient community members, respectively, these findings suggest specific environmental variables may be more important in shaping the relative abundances of existing community members,

rather than affecting bacterial immigration and emigration (i.e. dispersal) of transient community members.

The mānuka trees sampled in this study varied in height from 1.8 – 4.0 m (Table 4.7). According to data collected by Boffa Miskell Ltd (2017), mānuka is a fast-growing tree species and the estimated age of the trees sampled in this study may range between 2-10 years. Abundant taxa showed a significant relationship with tree height while rare taxa did not (Table 4.8). For example, the relative abundance of *Bacteroidetes* was increased in the phyllosphere of taller trees, whilst *Acidobacteria* tended to be more abundant in the phyllosphere of shorter trees (Supplementary Table 6.5). Changes in total productivity and leaf physiology occur in angiosperms as host plants grow and age. For example, stomatal conductance to water vapour and carbon dioxide often declines, causing dramatic changes in the rate of net photosynthesis and overall carbon gain (Reich *et al*, 1991). Internal and external leaf morphological structures are also affected by alterations in nitrogen content, lignification, and stomatal density (Kenzo *et al*, 2015, Steppe *et al*, 2011). Therefore, a relationship with mānuka tree height could be the result of changing leaf properties or overall alterations in total tree physiology and productivity. However, it is difficult to ascertain whether such changes are age-dependent or the result of altered physiological and environmental stresses with increasing tree height. Moreover, a successional shift in the microbial communities of the phyllosphere, due to microorganism-mediated alterations of leaf surface properties, is also possible (Jackson & Denney 2011, Redford & Fierer 2009). This speculation is further supported by the observation that taller mānuka trees tended to harbour phyllosphere communities of increasing taxon richness, which may reflect the expansion of habitable zones in the phyllosphere as leaves grow and mature (Reich *et al*, 1991). Collectively, these findings suggest the ecological niche of the mānuka leaf surface has the potential to elicit a strong influence on composition, richness, and structure of abundant bacterial taxa in the phyllosphere.

Predicted mānuka phyllosphere metagenome

Although no direct inference can be made concerning the plant-specific adaptations or symbiont associations of the mānuka phyllosphere microbiome, general functionality of the entire phyllosphere community was predicted and compared to surface soil communities (PICRUSt). This metabolic prediction complemented the

inferred physiological characteristics of the core microbiome (Table 4.3). The metabolic profiles of phyllosphere and soil communities differed significantly, likely reflecting different selective pressures pertaining to each ecological niche (Figure 4.10). Compared to surface soil, the predicted metabolic profiles of the phyllosphere were more distinctly clustered, suggesting phyllosphere bacteria possess specific ecological attributes (Figure 4.10). Typically, genes involved in conferring phylloepiphytic survival during extreme fluctuations of temperature, UV radiation, and nutrient availability were enriched in the phyllosphere compared to soil (Figure 4.11). Moreover, there was also an enrichment of genes involved in calcium signalling and xenobiotic degradation, as well as an abundance of genes involved in signal transduction and iron-homeostasis (Figure 4.11 & Figure 4.14). PICRUSt is a robust bioinformatics software package often used to accurately and efficiently characterize microbial metabolic pathways (Abubucker *et al*, 2012, Mukherjee *et al*, 2017). However, owing to the inherently predictive nature of PICRUSt, these results must be primarily viewed as an interesting foundation for future metagenomic and metaproteogenomic studies of the phyllosphere.

In summary, a prominent core microbiome was identified across all samples of the mānuka phyllosphere and consisted of members of *Alphaproteobacteria*, *Bacteriodetes, Acidobacteria*, and *Verrucomicrobia*. Putative identities of core taxa revealed physiological adaptations that likely promote survival on the leaf surface and intimate associations with their plant host. Biogeographical patterning was also detected in the mānuka phyllosphere microbiome and appeared driven by a complex interplay of spatial, climatic, and host-tree related selection pressures. However, rare and abundant colonisers of the phyllosphere appear differentially affected by these selection pressures. Rare bacteria exhibit a greater propensity to disperse and display a distance-decay relationship. In contrast, abundant bacterial members, including the core microbiome, were more greatly affected by climatic selection pressures, such as temperature and precipitation, and plant-stature related characteristics.

5. Chapter Five:
Conclusions and future directions

The aim of this research was to explore the bacterial communities of the mānuka phyllosphere microbiome across five native forests in the North Island of New Zealand. The findings of this research revealed the mānuka phyllosphere is comprised of a habitat-specific bacterial community that is structurally and functionally distinct to the surrounding environment. Intriguingly, a mānuka phyllosphere core microbiome was identified, which may be responsible for the maintenance of a balanced phyllosphere community, health of individual mānuka trees, and ecosystem productivity. In addition, the mānuka phyllosphere was found to comprise a complex bacterial community controlled by both neutral and niche-based driving factors, whereby observed biogeographical patterning is primarily driven by spatially structured gradients in environmental variables, such as temperature and precipitation. The involvement of both stochastic and deterministic forces makes the phyllosphere an intriguing system in which to further explore microbial biogeography. It can be anticipated that these results will provide the foundation for confirming epiphytic physiology, identifying unique bacterial traits pertaining to the mānuka phyllosphere, and discovering new mechanisms of bacteria-plant relationships. Advancement of knowledge pertaining to the mānuka phyllosphere may have large implications for maintaining the health and productivity of native mānuka forests and mānuka plantations within New Zealand.

Owing to the largely explorative nature of the current research project, the findings of this study present an infinite number of avenues for future study. As only mānuka was sampled for this research, the analysis of additional tree species within the same sample regions would be useful in order to elucidate the species-specificity of the mānuka phyllosphere. Examining the phyllosphere of *Kunzea ericoides* (kānuka) would be of particular interest owing to its morphological and ecological resemblance yet genetic dissimilarity. In addition, exploring the bacterial communities from the surface of other aerial plant structures, such as the anthosphere (flowers), would also be useful. Conducting such studies would provide additional information regarding the

species- and habitat-specificity of the core microbiome identified in the current research project, to mānuka leaves. Furthermore, ecological insight into the core microbiome could be enriched with ‘omic approaches. Metagenomic sequencing would provide information on the entire genetic content of specific core taxa beyond that offered by 16S rRNA gene sequencing. Metagenomic library construction would allow the recovery of functional genes and improve our understanding of the physiological roles of this diverse but highly specialised phyllosphere microbiota. A metaproteogenomic investigation would meanwhile be inherently valuable to bridge the gap between genomic sequences and metabolic function. Research of this nature would directly extend the concept of the core microbiome beyond taxonomic identification to community function. From this knowledge we may be able to reveal whether or not metabolic interdependencies underpin the stability of plant-microbe associations in the mānuka phyllosphere. Structural and functional examination of the core microbiome from healthy and diseased mānuka trees would provide further information regarding the implication and involvement of core taxa on host tree health and vice versa. Furthermore, it would be interesting to compare the phyllosphere microbiome of different genotypes and phenotypes within the mānuka species, as well as between native forests and mānuka plantations with different honey productivities (UMF). The acquisition of a larger dataset could be used to create a nationwide inventory of dominant mānuka bacterial phylotypes.

To further knowledge regarding microbial biogeography, and the relative influence of abiotic and biotic factors in shaping phyllosphere community composition, a similar investigation to the current study with a larger number of sample sites and collection of metadata could be conducted. Specifically, variables directly related to leaf physiology, such as leaf nutrient concentrations, dry matter content, and mass per area, as well host tree ecological strategies, should be considered. As an example, the chemical composition of the waxy leaf cuticle could be measured by gas chromatography. Incorporating more intensive measurements of intra- and inter-individual variation in leaf functional traits may further elucidate relationships between mānuka and the bacterial communities in the phyllosphere. Moreover, temporal analyses of the mānuka phyllosphere microbiome over sequential years would be useful to ascertain successional shifts in bacterial communities. Lastly, oligotyping would serve as an inherently useful extension of the current research project if granted

more time. Oligotyping is a relatively new computational method that allows researchers to investigate the concealed diversity of bacterial taxa within the final operational units of classification (Eren *et al*, 2013). Oligotype analysis of sequences belonging to each core bacterial member would provide further information as to whether fine-scale biogeographical differences exist in the mānuka phyllosphere core microbiome. By discovering the driving factors of abundant and functionally important phyllosphere microorganisms it may become possible to predict how future environmental change will affect the spatial distribution of these taxa. The availability of such knowledge could be used to promote healthy mānuka populations in the face of ongoing global environmental change and would have huge applications for the growing mānuka honey industry within New Zealand.

References

Abanda-Nkpwatt, D., Müsch, M., Tschiersch, J. & Schwab, W. 2006, 'Molecular interaction between methylobacterium extorquens and seedlings: Growth promotion, methanol consumption, and localization of the methanol emission site', *Journal of Experimental Botany,* vol. 57**,** no. 15**,** pp. 4025-4032**.**

Abubucker, S., Segata, N., Goll, J., Schubert, A. M., Izard, J., Cantarel, B. L., Rodriguez-Mueller, B., Zucker, J., Thiagarajan, M. & Henrissat, B. 2012, 'Metabolic reconstruction for metagenomic data and its application to the human microbiome', *PLoS computational biology,* vol. 8**,** no. 6**,** pp. e1002358**.**

Agler, M., T. , Ruhe, J., Kroll, S., Morhenn, C., Kim, S., Weigel, D. & Kemen, E. M. 2016, 'Microbial hub taxa link host and abiotic factors to plant microbiome variation', *PLoS Biology,* vol. 14**,** no. 1**,** pp. e1002352**.**

Amann, R. I., Ludwig, W. & Schleifer, K. H. 1995, 'Phylogenetic identification and in situ detection of individual microbial cells without cultivation', *Microbiological Reviews,* vol. 59**,** no. 1**,** pp. 143-69**.**

An, H., Xu, M., Dai, J., Wang, Y., Cai, F., Qi, H., Peng, F. & Fang, C. 2011, 'Sphingomonas xinjiangensis sp. Nov., isolated from desert sand', *International Journal of Systematic and Evolutionary Microbiology,* vol. 61**,** no. 8**,** pp. 1865-1869**.**

Andrews, J. H. & Harris, R. F. 2000, 'The ecology and biogeography of microorganisms on plant surfaces', *Annual Review of Phytopathology,* vol. 38**,** no. 1**,** pp. 145-180**.**

Andrews, J. H., Kenerley, C. M. & Nordheim, E. V. 1980, 'Positional variation in phylloplane microbial populations within an apple tree canopy', *Microbial ecology,* vol. 6**,** no. 1**,** pp. 71-84**.**

Andrews, S. C., Robinson, A. K. & Rodríguez-Quiñones, F. 2003, 'Bacterial iron homeostasis', *FEMS Microbiology Reviews,* vol. 27**,** pp. 215-237**.**

Armstrong, G. A. 1997, 'Genetics of eubacterial carotenoid biosynthesis: A colorful tale', *Annual Review of Microbiology,* vol. 51**,** no. 1**,** pp. 629-659**.**

Atrott, J., Haberlau, S. & Henle, T. 2012, 'Studies on the formation of methylglyoxal from dihydroxyacetone in manuka (leptospermum scoparium) honey', *Carbohydrate Research,* vol. 361**,** pp. 7-11**.**

Bache, S. M. & Wickham, H. 2014, 'Magrittr: A forward-pipe operator for r', *R package version,* vol. 1**,** no. 1**.**

Baik, K. S., Choi, J. S., Kwon, J., Park, S. C., Hwang, Y. M., Kim, M. S., Kim, E. M., Seo, D. C., Cho, J. S. & Seong, C. N. 2013, 'Terriglobus aquaticus sp. Nov., isolated from an artificial reservoir', *International Journal of Systematic and Evolutionary Microbiology,* vol. 63**,** no. 12**,** pp. 4744-4749**.**

Baldotto, L. E. B. & Olivares, F. L. 2008, 'Phylloepiphytic interaction between bacteria and different plant species in a tropical agricultural system', *Canadian Journal of Microbiology,* vol. 54**,** no. 11**,** pp. 918-931**.**

Barriuso, J., Valverde, J. R. & Mellado, R. P. 2011, 'Estimation of bacterial diversity using next generation sequencing of 16s rdna: A comparison of different workflows', *BMC Bioinformatics,* vol. 12**,** pp. 473-473**.**

Beattie, G. A. & Lindow, S. E. 1999, 'Bacterial colonization of leaves: A spectrum of strategies', *Phytopathology,* vol. 89**,** no. 5**,** pp. 353-359**.**

Begg, J. & Johnston, M. 2000. *Geology of the wellington area: Institute of geological & nuclear sciences 1: 250 000 geological map 10*.

Belova, S. E., Kulichevskaya, I. S., Bodelier, P. L. & Dedysh, S. N. 2013, 'Methylocystis bryophila sp. Nov., a facultatively methanotrophic bacterium from acidic sphagnum peat, and emended description of the genus methylocystis (ex whittenbury et al. 1970) bowman et al. 1993', *International Journal of Systematic and Evolutionary Microbiology*, vol. 63**,** no. 3**,** pp. 1096-1104**.**

Bissett, A., Richardson, A. E., Baker, G., Wakelin, S. & Thrall, P. H. 2010, 'Life history determines biogeographical patterns of soil bacterial communities over multiple spatial scales', *Molecular Ecology*, vol. 19**,** no. 19**,** pp. 4315-4327**.**

Bodenhausen, N., Horton, M. W. & Bergelson, J. 2013, 'Bacterial communities associated with the leaves and the roots of arabidopsis thaliana', *PLoS One*, vol. 8**,** no. 2**,** pp. e56329**.**

Boffa Miskell Ltd 2017. The mānuka and kānuka plantation guide.

Bowers, R. M., Lauber, C. L., Wiedinmyer, C., Hamady, M., Hallar, A. G., Fall, R., Knight, R. & Fierer, N. 2009, 'Characterization of airborne microbial communities at a high-elevation site and their potential to act as atmospheric ice nuclei', *Applied and Environmental Microbiology*, vol. 75**,** no. 15**,** pp. 5121-5130**.**

Brandl, M. T. & Lindow, S. E. 1998, 'Contribution of indole-3-acetic acid production to the epiphytic fitness of erwinia herbicola', *Applied and Environmental Microbiology*, vol. 64**,** no. 9**,** pp. 3256-3263**.**

Bryant, J., Lamanna, C., Morlon, H., Kerkhoff, A., Enquist, B. & Green, J. 2008, 'Microbes on mountainsides: Contrasting elevational patterns of bacterial and plant diversity', *Proceedings of the National Academy of Sciences of the United States of America*, vol. 105**,** pp. 11505-11511**.**

Bulgarelli, D., Schlaeppi, K., Spaepen, S., Themaat, E. V. L. V. & Schulze-Lefert, P. 2013, 'Structure and functions of the bacterial microbiota of plants', *Annual Review of Plant Biology*, vol. 64**,** no. 1**,** pp. 807-838**.**

Burbano, C. S., Liu, Y., Rösner, K. L., Reis, V. M., Caballero-Mellado, J., Reinhold-Hurek, B. & Hurek, T. 2011, 'Predominant nifh transcript phylotypes related to rhizobium rosettiformans in field-grown sugarcane plants and in norway spruce', *Environmental microbiology reports*, vol. 3**,** no. 3**,** pp. 383-389**.**

Capelo-Martínez, J.-L. 2009. *Ultrasound in chemistry: Analytical applications*, John Wiley & Sons.

Capra, E. & Laub, M. 2012, 'Evolution of two-component signal transduction systems', *Annual Review of Microbiology*, vol. 66**,** pp. 325-347**.**

Chahboune, R., Carro, L., Peix, A., Barrijal, S., Velazquez, E. & Bedmar, E. J. 2011, 'Bradyrhizobium cytisi sp. Nov., isolated from effective nodules of cytisus villosus', *International Journal of Systematic and Evolutionary Microbiology*, vol. 61**,** no. 12**,** pp. 2922-2927**.**

Chaikam, V. & Karlson, D. T. 2010, 'Comparison of structure, function and regulation of plant cold shock domain proteins to bacterial and animal cold shock domain proteins', *BMB Reports*, vol. 43**,** no. 1**,** pp. 1-8**.**

Chao, A., Chazdon, R. L., Colwell, R. K. & Shen, T. J. 2006, 'Abundance-based similarity indices and their estimation when there are unseen species in samples', *Biometrics*, vol. 62**,** no. 2**,** pp. 361-371**.**

Cho, I. & Blaser, M. J. 2012, 'The human microbiome: At the interface of health and disease', *Nature reviews. Genetics*, vol. 13**,** no. 4**,** pp. 260-270**.**

Chung, A. P., Lopes, A., Nobre, M. F. & Morais, P. V. 2010, 'Hymenobacter perfusus sp. Nov., hymenobacter flocculans sp. Nov. And hymenobacter metalli sp. Nov. Three new species isolated from an uranium mine waste water treatment system', *Systematic and Applied Microbiology*, vol. 33**,** no. 8**,** pp. 436-443**.**

Claesson, M. J., Wang, Q., Sullivan, O., Greene-Diniz, R., Cole, J. R., Ross, R. P. & Toole, P. W. 2010, 'Comparison of two next-generation sequencing technologies for resolving highly complex microbiota composition using tandem variable 16s rrna gene regions', *Nucleic Acids Research*, vol. 38**,** no. 22**,** pp. e200**.**

Cochran, P. 1979. Site index and height growth curves for managed, even-aged stands of douglas-fir east of the cascades in oregon and washington. *PNW-RP-251, Portland, OR: US Department of Agriculture, Forest Service, Pacific Northwest Research Station. 16 p.*

Cooper, R. A., Molan, P. C. & Harding, K. G. 2002, 'The sensitivity to honey of gram-positive cocci of clinical significance isolated from wounds', *Journal of Applied Microbiology*, vol. 93**,** no. 5**,** pp. 857-863**.**

Copeland, J. K., Yuan, L., Layeghifard, M., Wang, P. W. & Guttman, D. S. 2015, 'Seasonal community succession of the phyllosphere microbiome', *Molecular plant-microbe interactions : MPMI*, vol. 28**,** no. 3**,** pp. 274-285**.**

Costello, E. K., Stagaman, K., Dethlefsen, L., Bohannan, B. J. M. & Relman, D. A. 2012, 'The application of ecological theory toward an understanding of the human microbiome', *Science*, vol. 336**,** no. 6086**,** pp. 1255-1262**.**

Costerton, J. W., Lewandowski, Z., Caldwell, D. E., Korber, D. R. & Lappin-Scott, H. M. 1995, 'Microbial biofilms', *Annual Reviews in Microbiology*, vol. 49**,** no. 1**,** pp. 711-745**.**

Dams, E., Hendriks, L., Van De Peer, Y., Neefs, J. M., Smits, G., Vandenbempt, I. & De Wachter, R. 1988, 'Compilation of small ribosomal subunit rna sequences', *Nucleic Acids Research*, vol. 16**,** no. suppl**,** pp. 87-173**.**

Danielsson, P.-E. 1980, 'Euclidean distance mapping', *Computer Graphics and image processing*, vol. 14**,** no. 3**,** pp. 227-248**.**

Darcy, J. L., Lynch, R. C., King, A. J., Robeson, M. S. & Schmidt, S. K. 2011, 'Global distribution of polaromonas phylotypes - evidence for a highly successful dispersal capacity', *PLoS ONE*, vol. 6**,** no. 8**,** pp. e23742**.**

De Wit, R. & Bouvier, T. 2006, '‘Everything is everywhere, but, the environment selects’; what did baas becking and beijerinck really say?', *Environmental Microbiology*, vol. 8**,** no. 4**,** pp. 755-758**.**

Dedysh, S. N., Didriksen, A., Danilova, O. V., Belova, S. E., Liebner, S. & Svenning, M. M. 2015, 'Methylocapsa palsarum sp. Nov., a methanotroph isolated from a subarctic discontinuous permafrost ecosystem', *International Journal of Systematic and Evolutionary Microbiology*, vol. 65**,** no. 10**,** pp. 3618-3624**.**

Degefu, Y., Virtanen, E. & Väyrynen, T. 2009, 'Pre-pcr processes in the molecular detection of blackleg and soft rot erwiniae in seed potatoes', *Journal of Phytopathology*, vol. 157**,** no. 6**,** pp. 370-378**.**

Delmotte, N., Knief, C., Chaffron, S., Innerebner, G., Roschitzki, B., Schlapbach, R., Von Mering, C. & Vorholt, J. A. 2009, 'Community proteogenomics reveals insights into the physiology of phyllosphere bacteria', *Proceedings of the National Academy of Sciences*, vol. 106**,** no. 38**,** pp. 16428-16433**.**

Dreesens, L. L., Lee, C. K. & Cary, S. C. 2014, 'The distribution and identity of edaphic fungi in the mcmurdo dry valleys', *Biology*, vol. 3**,** no. 3**,** pp. 466-483**.**

Edbrooke, S. 2005. *Geology of the waikato area: Scale 1:250,000 geological map 4. 68 p. + 1 folded map*. Lower Hutt, Wellington: Institute of Geological & Nuclear Sciences.

Edgar, R. C. 2010, 'Search and clustering orders of magnitude faster than blast', *Bioinformatics*, vol. 26**,** no. 19**,** pp. 2460-2461**.**

Eichorst, S. A., Breznak, J. A. & Schmidt, T. M. 2007, 'Isolation and characterization of soil bacteria that define terriglobus gen. Nov., in the phylum acidobacteria', *Applied and Environmental Microbiology*, vol. 73**,** no. 8**,** pp. 2708-2717**.**

Enrico, B., Francesco, C., Lorenzo, M., Florian, A., Ignacio, R.-I. & Andrea, R. 2016, 'Geomorphic controls on elevational gradients of species richness', *Proceedings of the National Academy of Sciences*, vol. 113**,** no. 7**,** pp. 1737**.**

Ercolani, G. 1991, 'Distribution of epiphytic bacteria on olive leaves and the influence of leaf age and sampling time', *Microbial Ecology*, vol. 21**,** no. 1**,** pp. 35-48**.**

Eren, A. M., Maignien, L., Sul, W. J., Murphy, L. G., Grim, S. L., Morrison, H. G. & Sogin, M. L. 2013, 'Oligotyping: Differentiating between closely related microbial taxa using 16s rrna gene data', *Methods in Ecology and Evolution*, vol. 4**,** no. 12**,** pp. 1111-1119**.**

Estrela, C., Estrela, C. R., Barbin, E. L., Spano, J. C., Marchesan, M. A. & Pecora, J. D. 2002, 'Mechanism of action of sodium hypochlorite', *Brazilian Dental Journal*, vol. 13**,** no. 2**,** pp. 113-117**.**

Fall, R. & Benson, A. A. 1996, 'Leaf methanol — the simplest natural product from plants', *Trends in Plant Science*, vol. 1**,** no. 9**,** pp. 296-301**.**

Feig, A. & Lippard, S. 1994, 'Reactions of non-heme iron(ii) centers with dioxygen in biology and chemistry', *Chemical Reviews*, vol. 94**,** no. 3**,** pp. 759-805**.**

Finkel, O. M., Burch, A. Y., Elad, T., Huse, S. M., Lindow, S. E., Post, A. F. & Belkin, S. 2012, 'Distance-decay relationships partially determine diversity patterns of phyllosphere bacteria on tamrix trees across the sonoran desert', *Applied and Environmental Microbiology*, vol. 78**,** no. 17**,** pp. 6187-6193**.**

Finkel, O. M., Burch, A. Y., Lindow, S. E., Post, A. F. & Belkin, S. 2011, 'Geographical location determines the population structure in phyllosphere microbial communities of a salt-excreting desert tree', *Applied and Environmental Microbiology*, vol. 77**,** no. 21**,** pp. 7647-7655**.**

Finlay, B. J. 2002, 'Global dispersal of free-living microbial eukaryote species', *Science*, vol. 296**,** no. 5570**,** pp. 1061-1063**.**

Fürnkranz, M., Wanek, W., Richter, A., Abell, G., Rasche, F. & Sessitsch, A. 2008, 'Nitrogen fixation by phyllosphere bacteria associated with higher plants and their colonizing epiphytes of a tropical lowland rainforest of costa rica', *The ISME Journal*, vol. 2**,** no. 5**,** pp. 561-570**.**

Garvie, E. I. 1980, 'Bacterial lactate dehydrogenases', *Microbiological Reviews*, vol. 44**,** no. 1**,** pp. 106-139**.**

Glöckner, F. O., Zaichikov, E., Belkova, N., Denissova, L., Pernthaler, J., Pernthaler, A. & Amann, R. 2000, 'Comparative 16s rrna analysis of lake bacterioplankton reveals globally distributed phylogenetic clusters including an abundant group of actinobacteria', *Applied and environmental microbiology*, vol. 66**,** no. 11**,** pp. 5053-5065**.**

Gourion, B., Rossignol, M. & Vorholt, J., A. 2006, 'A proteomic study of methylobacterium extorquens reveals a response regulator essential for epiphytic growth', *Proceedings of the National Academy of Sciences,* vol. 103**,** no. 35**,** pp. 13186-13191**.**

Gronemeyer, J. L., Hurek, T., Bunger, W. & Reinhold-Hurek, B. 2016, 'Bradyrhizobium vignae sp. Nov., a nitrogen-fixing symbiont isolated from effective nodules of vigna and arachis', *International journal of systematic and evolutionary microbiology,* vol. 66**,** no. 1**,** pp. 62-69**.**

Groussin, M., Mazel, F., Sanders, J. G., Smillie, C. S., Lavergne, S., Thuiller, W. & Alm, E. J. 2017, 'Unraveling the processes shaping mammalian gut microbiomes over evolutionary time', *Nature communications,* vol. 8**,** no. 14319**.**

Guillot, G. & Rousset, F. 2013, 'Dismantling the mantel tests', *Methods in Ecology and Evolution,* vol. 4**,** no. 4**,** pp. 336-344**.**

Hamilton, G., Millner, J., Robertson, A. & Stephens, J. 2013, 'Assessment of manuka provenances for production of high 'unique manuka factor'honey', *Agronomy New Zealand,* vol. 43**,** pp. 139-144**.**

Hanson, C. A., Fuhrman, J. A., Horner-Devine, M. C. & Martiny, J. B. 2012, 'Beyond biogeographic patterns: Processes shaping the microbial landscape', *Nature Reviews Microbiology,* vol. 10**,** no. 7**,** pp. 497-506**.**

Hoang, V. A., Kim, Y. J., Nguyen, N. L. & Yang, D. C. 2013, 'Hymenobacter ginsengisoli sp. Nov., isolated from soil of a ginseng field', *International journal of systematic and evolutionary microbiology,* vol. 63**,** no. 2**,** pp. 661-666**.**

Huang, H. D., Wang, W., Ma, T., Li, G. Q., Liang, F. L. & Liu, R. L. 2009, 'Sphingomonas sanxanigenens sp. Nov., isolated from soil', *International journal of systematic and evolutionary microbiology,* vol. 59**,** no. 4**,** pp. 719-723**.**

Hunter, P. J., Hand, P., Pink, D., Whipps, J. M. & Bending, G. D. 2010, 'Both leaf properties and microbe-microbe interactions influence within-species variation in bacterial population diversity and structure in the lettuce (lactuca species) phyllosphere', *Applied and Environmental Microbiology,* vol. 76**,** no. 24**,** pp. 8117-8125**.**

Innerebner, G., Knief, C. & Vorholt, J. A. 2011, 'Protection of arabidopsis thaliana against leaf-pathogenic pseudomonas syringae by sphingomonas strains in a controlled model system', *Applied and Environmental Microbiology,* vol. 77**,** no. 10**,** pp. 3202-3210**.**

Jackson, C. R. & Denney, W. C. 2011, 'Annual and seasonal variation in the phyllosphere bacterial community associated with leaves of the southern magnolia (magnolia grandiflora)', *Microbial Ecology,* vol. 61**,** no. 1**,** pp. 113-122**.**

Jackson, E. F., Echlin, H. L. & Jackson, C. R. 2006, 'Changes in the phyllosphere community of the resurrection fern, polypodium polypodioides, associated with rainfall and wetting', *FEMS Microbiology Ecology,* vol. 58**,** no. 2**,** pp. 236-246**.**

Jacques, M. A., Josi, K., Darrasse, A. & Samson, R. 2005, 'Xanthomonas axonopodis pv. Phaseoli var. Fuscans is aggregated in stable biofilm population sizes in the phyllosphere of field-grown beans', *Applied and Environmental Microbiology,* vol. 71**,** no. 4**,** pp. 2008-2015**.**

Jahan, R. 2013. *Diversity of methylobacterium species associated with new zealand native plants.* Master of Science, University of Waikato.

Janda, J. M. & Abbott, S. L. 2007, '16s rrna gene sequencing for bacterial identification in the diagnostic laboratory: Pluses, perils, and pitfalls', *Journal of Clinical Microbiology,* vol. 45**,** no. 9**,** pp. 2761-2764**.**

Johnston, P. R. 1998, 'Leaf endophytes of manuka (leptospermum scoparium)', *Mycological Research,* vol. 102**,** no. 8**,** pp. 1009-1016**.**

Jousset, A., Bienhold, C., Chatzinotas, A., Gallien, L., Gobet, A., Kurm, V., Küsel, K., Rillig, M. C., Rivett, D. W. & Salles, J. F. 2017, 'Where less may be more: How the rare biosphere pulls ecosystems strings', *The ISME journal,* vol. 11**,** no. 4**,** pp. 853-862**.**

Kadivar, H. & Stapleton, A. E. 2003, 'Ultraviolet radiation alters maize phyllosphere bacterial diversity', *Microbial Ecology,* vol. 45**,** no. 4**,** pp. 353-361**.**

Kang, H., Kim, H., Joung, Y., Kim, K. J. & Joh, K. 2016, 'Hymenobacter marinus sp. Nov., isolated from coastal seawater', *International Journal of Systematic and Evolutionary Microbiology,* vol. 66**,** no. 6**,** pp. 2212-7**.**

Kembel, S. W., O'connor, T. K., Arnold, H. K., Hubbell, S. P., Wright, S. J. & Green, J. L. 2014, 'Relationships between phyllosphere bacterial communities and plant functional traits in a neotropical forest', *Proceedings of the National Academy of Sciences,* vol. 111**,** no. 38**,** pp. 13715-13720**.**

Kembel, S. W., Wu, M., Eisen, J. A. & Green, J. L. 2012, 'Incorporating 16s gene copy number information improves estimates of microbial diversity and abundance (estimating 16s copy number and microbial abundance)', *PLoS Computational Biology,* vol. 8**,** no. 10**,** pp. e1002743**.**

Kemp, B. M. & Smith, D. G. 2005, 'Use of bleach to eliminate contaminating DNA from the surface of bones and teeth', *Forensic Science International,* vol. 154**,** no. 1**,** pp. 53-61**.**

Kenzo, T., Inoue, Y., Yoshimura, M., Yamashita, M., Tanaka-Oda, A. & Ichie, T. 2015, 'Height-related changes in leaf photosynthetic traits in diverse bornean tropical rain forest trees', *Oecologia,* vol. 177**,** no. 1**,** pp. 191-202**.**

Kim, H., Nishiyama, M., Kunito, T., Senoo, K., Kawahara, K., Murakami, K. & Oyaizu, H. 1998, 'High population of sphingomonas species on plant surface', *Journal of Applied Microbiology,* vol. 85**,** no. 4**,** pp. 731-736**.**

Kim, M., Singh, D., Lai-Hoe, A., Go, R., Abdul Rahim, R., A.N., A., Chun, J. & Adams, J. M. 2012, 'Distinctive phyllosphere bacterial communities in tropical trees', *Microbial Ecology,* vol. 63**,** no. 3**,** pp. 674-681**.**

Kinkel, L., Wilson, M. & Lindow, S. 1995, 'Effect of sampling scale on the assessment of epiphytic bacterial populations', *An International Journal,* vol. 29**,** no. 3**,** pp. 283-297**.**

Knief, C., Delmotte, N., Chaffron, S., Stark, M., Innerebner, G., Wassmann, R., Von Mering, C. & Vorholt, J. A. 2012, 'Metaproteogenomic analysis of microbial communities in the phyllosphere and rhizosphere of rice', *The ISME Journal,* vol. 6**,** no. 7**,** pp. 1378-1390**.**

Knief, C., Vorholt, J. A., Frances, L., Ramette, A. & Alonso-Blanco, C. 2010, 'Site and plant species are important determinants of the methylobacterium community composition in the plant phyllosphere', *The ISME Journal,* vol. 4**,** no. 6**,** pp. 719-728**.**

Kühlbrandt, W. 1995, 'Structure and function of bacterial light-harvesting complexes', *Structure,* vol. 3**,** no. 6**,** pp. 521-525**.**

Kürkcüoglu, S., Degenhardt, J., Lensing, J., Al-Masri, A. N. & Gau, A. E. 2007, 'Identification of differentially expressed genes in malus domestica after application of the non-pathogenic bacterium pseudomonas fluorescens bk3 to the phyllosphere', *Journal of Experimental Botany*, vol. 58**,** no. 3**,** pp. 733-741**.**

Laforest-Lapointe, I., Messier, C. & Kembel, S. 2016a, 'Host species identity, site and time drive temperate tree phyllosphere bacterial community structure', *Microbiome*, vol. 4**,** no. 27**.**

Laforest-Lapointe, I., Messier, C. & Kembel, S. W. 2016b, 'Tree phyllosphere bacterial communities: Exploring the magnitude of intra- and inter-individual variation among host species', *PeerJ*, vol. 4**,** pp. e2367**.**

Laforest-Lapointe, I., Paquette, A., Messier, C. & Kembel, S. W. 2017, 'Leaf bacterial diversity mediates plant diversity and ecosystem function relationships', *Nature*, vol. 546**,** no. 7656**,** pp. 145-147**.**

Lambais, M. R., Barrera, S. E., Santos, E. C., Crowley, D. E. & Jumpponen, A. 2017, 'Phyllosphere metaproteomes of trees from the brazilian atlantic forest show high levels of functional redundancy', *Microbial Ecology*, vol. 73**,** no. 1**,** pp. 123-134**.**

Lambais, M. R., Crowley, D. E., Cury, J. C., Büll, R. C. & Rodrigues, R. R. 2006, 'Bacterial diversity in tree canopies of the atlantic forest', *Science*, vol. 312**,** no. 5782**,** pp. 1917-1917**.**

Langille, M. G. I., Zaneveld, J., Caporaso, J. G., Mcdonald, D., Knights, D., Reyes, J. a. R., Clemente, J. C., Burkepile, D. E., Thurber, R. L. V. T., Knight, R., Beiko, R. G. & Huttenhower, C. 2013, 'Predictive functional profiling of microbial communities using 16s rrna marker gene sequences', *Nature Biotechnology*, vol. 31**,** no. 9**,** pp. 814–821**.**

Lauber, C. L., Hamady, M., Knight, R. & Fierer, N. 2009, 'Pyrosequencing- based assessment of soil ph as a predictor of soil bacterial community structure at the continental scale', *Applied and Environmental Microbiology*, vol. 75**,** no. 15**,** pp. 5111-5120**.**

Lear, G., Washington, V., Neale, M., Case, B., Buckley, H. & Lewis, G. 2013, 'The biogeography of stream bacteria', *Global Ecology and Biogeography*, vol. 22**,** no. 5**,** pp. 544-554**.**

Lee, J., Bland, K., Townsend, D. & Kamp, P. 2011. *Geology of the hawke's bay area: Scale 1:250,000 map 8. 86 p. + folded map*. Lower Hutt, Wellington.

Leonard, G., Begg, J. & Wilson, C. 2010. *Geology of the rotorua area: Scale 1:250,000 geological map 5. 99 p + 1 folded map*. Lower Hutt, New Zealand: GNS Science.

Liaw, A. & Wiener, M. 2002, 'Classification and regression by randomforest', *R news*, vol. 2**,** no. 3**,** pp. 18-22**.**

Lindow, S. E. & Brandl, M. T. 2003, 'Microbiology of the phyllosphere', *Applied and environmental microbiology*, vol. 69**,** no. 4**,** pp. 1875-1883**.**

Lomolino, M. V., Riddle, B., Whittaker, R. J., Brown, J. H. & Lomolino, M. V. 2010. *Biogeography*, Sunderland, Massachusetts, Sinauer Associates.

Lu, J. K., Dou, Y. J., Zhu, Y. J., Wang, S. K., Sui, X. H. & Kang, L. H. 2014, 'Bradyrhizobium ganzhouense sp. Nov., an effective symbiotic bacterium isolated from acacia melanoxylon r. Br. Nodules', *International Journal of Systematic and Evolutionary Microbiology*, vol. 64**,** no. 6**,** pp. 1900-1905**.**

Lundberg, D., Lebeis, S., Paredes, S., Yourstone, S., Gehring, J., Malfatti, S., Tremblay, J., Engelbrektson, A., Kunin, V., Del Rio, T., Edgar, R., Eickhorst, T., Ley, R., Hugenholtz, P., Tringe, S. & Dang, J. 2012, 'Defining the core arabidopsis thaliana root microbiome', *Nature,* vol. 487**,** no. 7409**,** pp. 86-90**.**

Lushchak, V. 2001, 'Oxidative stress and mechanisms of protection against it in bacteria', *Biochemistry (Moscow),* vol. 66**,** no. 5**,** pp. 476-489**.**

Martiny, J. B. H., Bohannan, B. J., Brown, J. H., Colwell, R. K., Fuhrman, J. A., Green, J. L., Horner-Devine, M. C., Kane, M., Krumins, J. A., Kuske, C. R., Morin, P. J., Naeem, S., Ovreas, L., Reysenbach, A. L., Smith, V. H. & Staley, J. T. 2006, 'Microbial biogeography: Putting microorganisms on the map', *Nature Reviews Microbiology,* vol. 4**,** no. 2**,** pp. 102-112**.**

Massana, R., Delong, E. F. & Pedrós-Alió, C. 2000, 'A few cosmopolitan phylotypes dominate planktonic archaeal assemblages in widely different oceanic provinces', *Applied and Environmental Microbiology,* vol. 66**,** no. 5**,** pp. 1777-1787**.**

Mavric, E., Wittmann, S., Barth, G. & Henle, T. 2008, 'Identification and quantification of methylglyoxal as the dominant antibacterial constituent of manuka (leptospermum scoparium) honeys from new zealand', *Molecular Nutrition and Food Research,* vol. 52**,** no. 4**,** pp. 483-489**.**

Mcdonald, D., Clemente, J. C., Kuczynski, J., Rideout, J. R., Stombaugh, J., Wendel, D., Wilke, A., Huse, S., Hufnagle, J. & Meyer, F. 2012, 'The biological observation matrix (biom) format or: How i learned to stop worrying and love the ome-ome', *GigaScience,* vol. 1**,** no. 1**,** pp. 7**.**

Mcdonald, I. R. & Murrell, J. C. 1997, 'The methanol dehydrogenase structural gene mxaf and its use as a functional gene probe for methanotrophs and methylotrophs', *Applied and Environmental Microbiology,* vol. 63**,** no. 8**,** pp. 3218-3224**.**

Mcmurdie, P. J. & Holmes, S. 2013, 'Phyloseq: An r package for reproducible interactive analysis and graphics of microbiome census data', *PloS one,* vol. 8**,** no. 4**,** pp. e61217**.**

Mechaber, W. L., Marshall, D. B., Mechaber, R. A., Jobe, R. T. & Chew, F. S. 1996, 'Mapping leaf surface landscapes', *Proceedings of the National Academy of Sciences of the United States of America,* vol. 93**,** no. 10**,** pp. 4600–4603**.**

Meyer, K. M. & Leveau, J. H. J. 2012, 'Microbiology of the phyllosphere: A playground for testing ecological concepts', *Oecologia,* vol. 168**,** no. 3**,** pp. 621-629**.**

Monier, J. M. & Lindow, S. E. 2004, 'Frequency, size, and localization of bacterial aggregates on bean leaf surfaces', *Applied and Environmental Microbiology,* vol. 70**,** no. 1**,** pp. 346-355**.**

Morris, C. E. & Monier, J.-M. 2003, 'The ecological significance of biofilm formation by plant-associated bacteria', *Annual Review of Phytopathology,* vol. 41**,** no. 1**,** pp. 429-453**.**

Morris, C. E., Monier, J. & Jacques, M. 1997, 'Methods for observing microbial biofilms directly on leaf surfaces and recovering them for isolation of culturable microorganisms', *Applied and Environmental Microbiology,* vol. 63**,** no. 4**,** pp. 1570-1576**.**

Mukherjee, A., Chettri, B., Langpoklakpam, J. S., Basak, P., Prasad, A., Mukherjee, A. K., Bhattacharyya, M., Singh, A. K. & Chattopadhyay, D. 2017, 'Bioinformatic approaches including predictive metagenomic profiling reveal

characteristics of bacterial response to petroleum hydrocarbon contamination in diverse environments', *Scientific reports*, vol. 7**,** no. 1108**.**
Müller, T. & Ruppel, S. 2014, 'Progress in cultivation-independent phyllosphere microbiology', *FEMS Microbiology Ecology*, vol. 87**,** no. 1**,** pp. 2-17**.**
Nakayama, H., Midwinter, G. & Krampitz, L. 1971, 'Properties of the pyruvate formate-lyase reaction', *Archives of Biochemistry and Biophysics*, vol. 143**,** no. 2**,** pp. 526-534**.**
Nemergut, D. R., Schmidt, S. K., Fukami, T., O'neill, S. P., Bilinski, T. M., Stanish, L. F., Knelman, J. E., Darcy, J. L., Lynch, R. C., Wickey, P. & Ferrenberg, S. 2013, 'Patterns and processes of microbial community assembly', *Microbiology and Molecular Biology Reviews : MMBR*, vol. 77**,** no. 3**,** pp. 342-356**.**
Niwa. 2017. *The national climate database* [Online]. Available: https://cliflo.niwa.co.nz/.
Nowak, D. J., Crane, D. E. & Stevens, J. C. 2006, 'Air pollution removal by urban trees and shrubs in the united states', *Urban forestry & urban greening*, vol. 4**,** no. 3-4**,** pp. 115-123**.**
O'brien, R. & Lindow, S. 1989, 'Effect of plant species and environmental conditions on epiphytic population sizes of pseudomonas syringae and other bacteria', *Phytopathology*, vol. 79**,** no. 5**,** pp. 619-627**.**
Oggerin, M., Arahal, D. R., Rubio, V. & Marin, I. 2009, 'Identification of beijerinckia fluminensis strains cip 106281t and uqm 1685t as rhizobium radiobacter strains, and proposal of beijerinckia doebereinerae sp. Nov. To accommodate beijerinckia fluminensis lmg 2819', *International Journal of Systematic and Evolutionary Microbiology*, vol. 59**,** no. 9**,** pp. 2323-2328**.**
Oksanen, J., Kindt, R., Legendre, P., O'hara, B., Stevens, M. H. H., Oksanen, M. J. & Suggests, M. 2007, 'The vegan package', *Community ecology package*, vol. 10**,** pp. 631-637**.**
Oliver, J. D. 1993. Formation of viable but nonculturable cells. *In:* KJELLEBERG, S. (ed.) *Starvation in bacteria.* Boston, MA: Springer US.
Ortega, R. A., Mahnert, A., Berg, C., Muller, H. & Berg, G. 2016, 'The plant is crucial: Specific composition and function of the phyllosphere microbiome of indoor ornamentals', *FEMS Microbiology Ecology*, vol. 92**,** no. 12**.**
Papke, R. T., Ramsing, N. B., Bateson, M. M. & Ward, D. M. 2003, 'Geographical isolation in hot spring cyanobacteria', *Environmental Microbiology*, vol. 5**,** no. 8**,** pp. 650-659**.**
Pascual, J., Wust, P. K., Geppert, A., Foesel, B. U., Huber, K. J. & Overmann, J. 2015, 'Terriglobus albidus sp. Nov., a member of the family acidobacteriaceae isolated from namibian semiarid savannah soil', *International Journal of Systematic and Evolutionary Microbiology*, vol. 65**,** no. 10**,** pp. 3297-3304**.**
Peterson, B. W., Sharma, P. K., Van Der Mei, H. C. & Busscher, H. J. 2012, 'Bacterial cell surface damage due to centrifugal compaction', *Applied and Environmental Microbiology*, vol. 78**,** no. 1**,** pp. 120-125**.**
Petrosino, J. F., Highlander, S., Luna, R. A., Gibbs, R. A. & Versalovic, J. 2009, 'Metagenomic pyrosequencing and microbial identification', *Clinical chemistry*, vol. 55**,** no. 5**,** pp. 856-866**.**
Prosser, J. A. 2011. *Manuka (leptospermum scoparium) as a remediation species for biosolids amended land.* Master of Science.

Quintana, J., Harrison, H., Palta, J., Nienhuis, J., Kmiecik, K. & Miglioranza, E. 2001, 'Stomatal density and calcium concentration of six snap bean cultivars', *Journal of the American Society for Horticultural Science,* vol. 126**,** no. 1**,** pp. 110-114**.**

R Core Team 2017. R: A language and environment for statistical computing. R foundation for statistical computing. Vienna, Austria.

Rastogi, G., Tech, J. J., Coaker, G. L. & Leveau, J. H. J. 2010, 'A pcr-based toolbox for the culture-independent quantification of total bacterial abundances in plant environments', *Journal of Microbiological Methods,* vol. 83**,** no. 2**,** pp. 127-132**.**

Redford, A. J., Bowers, R. M., Knight, R., Linhart, Y. & Fierer, N. 2010, 'The ecology of the phyllosphere: Geographic and phylogenetic variability in the distribution of bacteria on tree leaves: Biogeography of phyllosphere bacterial communities', *Environmental Microbiology,* vol. 12**,** no. 11**,** pp. 2885-2893**.**

Redford, A. J. & Fierer, N. 2009, 'Bacterial succession on the leaf surface: A novel system for studying successional dynamics', *Microbial Ecology,* vol. 58**,** no. 1**,** pp. 189-198**.**

Reich, P., Uhl, C., Walters, M. & Ellsworth, D. 1991, 'Leaf lifespan as a determinant of leaf structure and function among 23 amazonian tree species', *Oecologia,* vol. 86**,** no. 1**,** pp. 16-24**.**

Remus-Emsermann, M. N. P., Tecon, R., Kowalchuk, G. A. & Leveau, J. H. J. 2012, 'Variation in local carrying capacity and the individual fate of bacterial colonizers in the phyllosphere', *The ISME Journal,* vol. 6**,** no. 4**,** pp. 756-765**.**

Rennick, S. L., Fenton, T. W. & Foran, D. R. 2005, 'The effects of skeletal preparation techniques on DNA from human and non-human bone', *Journal of Forensic Science,* vol. 50**,** no. 5**,** pp. 1016-1019**.**

Reveillaud, J., Maignien, L., Eren, A. M., Huber, J. A., Apprill, A., Sogin, M. L. & Vanreusel, A. 2014, 'Host- specificity among abundant and rare taxa in the sponge microbiome', *The ISME Journal,* vol. 8**,** no. 6**,** pp. 1198-1209**.**

Richardson, L. A. 2017, 'Evolving as a holobiont', *PLoS Biology,* vol. 15**,** no. 2**,** pp. e2002168**.**

Rigonato, J., Alvarenga, D. O., Andreote, F. D., Dias, A. C., Melo, I. S., Kent, A. & Fiore, M. F. 2012, 'Cyanobacterial diversity in the phyllosphere of a mangrove forest', *FEMS Microbiology Ecology,* vol. 80**,** no. 2**,** pp. 312-322**.**

Rigonato, J., Gonçalves, N., Andreote, A. P. D., Lambais, M. R. & Fiore, M. F. 2016, 'Estimating genetic structure and diversity of cyanobacterial communities in atlantic forest phyllosphere', *Canadian Journal of Microbiology,* vol. 62**,** no. 11**,** pp. 953-960**.**

Robinson, C., Bohannan, B. & Young, V. 2010, 'From structure to function: The ecology of host-associated microbial communities', *Microbiology and Molecular Biology Reviews,* vol. 74**,** no. 3**,** pp. 453–476**.**

Rosindell, J., Hubbell, S. P. & Etienne, R. S. 2011, 'The unified neutral theory of biodiversity and biogeography at age ten', *Trends in Ecology & Evolution,* vol. 26**,** no. 7**,** pp. 340-348**.**

Ruinen, J. 1956, 'Occurrence of beijerinckia species in the ‘phyllosphere’', *Nature,* vol. 177**,** no. 4501**,** pp. 220–221**.**

Ruinen, J. 1961, 'The phyllosphere: I. An ecologically neglected milieu', *Plant and Soil,* vol. 15**,** no. 2**,** pp. 81-109**.**

Ruiz-Pérez, C. A., Restrepo, S. & Zambrano, M. M. 2016, 'Microbial and functional diversity within the phyllosphere of espeletia species in an andean high-mountain ecosystem', *Applied and Environmental Microbiology,* vol. 82**,** no. 6**,** pp. 1807-1817**.**

Russell, A. 2003, 'Lethal effects of heat on bacterial physiology and structure', *Science Progress,* vol. 86**,** no. 1-2**,** pp. 115-137**.**

Sangthong, S., Suksabye, P. & Thiravetyan, P. 2016, 'Air-borne xylene degradation by bougainvillea buttiana and the role of epiphytic bacteria in the degradation', *Ecotoxicology and Environmental Safety,* vol. 126**,** pp. 273-280**.**

Sangwan, P., Chen, X., Hugenholtz, P. & Janssen, P. H. 2004, 'Chthoniobacter flavus gen. Nov., sp. Nov., the first pure-culture representative of subdivision two, spartobacteria classis nov., of the phylum verrucomicrobia', *Applied and Environmental Microbiology,* vol. 70**,** no. 10**,** pp. 5875-5881**.**

Schloss, P. D., Westcott, S. L., Ryabin, T., Hall, J. R., Hartmann, M., Hollister, E. B., Lesniewski, R. A., Oakley, B. B., Parks, D. H. & Robinson, C. J. 2009, 'Introducing mothur: Open-source, platform-independent, community-supported software for describing and comparing microbial communities', *Applied and environmental microbiology,* vol. 75**,** no. 23**,** pp. 7537-7541**.**

Serrano, M., Coluccia, F., Torres, M., L'haridon, F. & Métraux, J.-P. 2014, 'The cuticle and plant defense to pathogens', *Frontiers in Plant Science,* vol. 5**.**

Shade, A. & Handelsman, J. 2012, 'Beyond the venn diagram: The hunt for a core microbiome', vol. 14**,** pp. 4-12**.**

Sinha, R. P. & Hader, D. P. 2002, 'Uv-induced DNA damage and repair: A review', *Photochemical and Photobiological Sciences,* vol. 1**,** no. 4**,** pp. 225-236**.**

Smets, W., Wuyts, K., Oerlemans, E., Wuyts, S., Denys, S., Samson, R. & Lebeer, S. 2016, 'Impact of urban land use on the bacterial phyllosphere of ivy (hedera sp.)', *Atmospheric Environment,* vol. 147**,** pp. 376-383**.**

Sneath, P. H., Mair, N. S., Sharpe, M. E. & Holt, J. G. 1986. *Bergey's manual of systematic bacteriology. Volume 2,* Baltimore, USA, Williams & Wilkins.

Soininen, J., Mcdonald, R. & Hillebrand, H. 2007, 'The distance decay of similarity in ecological communities', *Ecography,* vol. 30**,** no. 1**,** pp. 3-12**.**

Son, H. M., Kook, M., Tran, H. T., Kim, K. Y., Park, S. Y., Kim, J. H. & Yi, T. H. 2014, 'Sphingomonas kyeonggiense sp. Nov., isolated from soil of a ginseng field', *Antonie Van Leeuwenhoek,* vol. 105**,** no. 4**,** pp. 791-797**.**

Son, H. M., Yang, J. E., Park, Y., Han, C. K., Kim, S. G., Kook, M. & Yi, T. H. 2013, 'Sphingomonas kyungheensis sp. Nov., a bacterium with ginsenoside-converting activity isolated from soil of a ginseng field', *International Journal of Systematic and Evolutionary Microbiology,* vol. 63**,** no. 10**,** pp. 3848-3853**.**

Stephens, J. M. C., Molan, P. C. & Clarkson, B. D. 2005, 'A review of leptospermum scoparium (myrtaceae) in new zealand', *New Zealand Journal of Botany* vol. 43**,** pp. 431-449**.**

Steppe, K., Niinemets, Ü. & Teskey, R. O. 2011. Tree size- and age-related changes in leaf physiology and their influence on carbon gain. *In:* MEINZER, F. C., LACHENBRUCH, B. & DAWSON, T. E. (eds.) *Size- and age-related changes in tree structure and function.* Dordrecht: Springer Netherlands.

Stock, A. M., Robinson, V. L. & Goudreau, P. N. 2000, 'Two-component signal transduction', *Annual review of biochemistry,* vol. 69**,** no. 1**,** pp. 183-215**.**

Stone, B. W. G. & Jackson, C. R. 2016, 'Biogeographic patterns between bacterial phyllosphere communities of the southern magnolia (magnolia grandiflora) in a small forest', *Microbial Ecology*, vol. 71**,** no. 4**,** pp. 954-961**.**

Suda, W., Oto, M., Amachi, S., Shinoyama, H. & Shishido, M. 2008, 'A direct method to isolate DNA from phyllosphere microbial communities without disrupting leaf tissues', *Microbes and Environments*, vol. 23**,** no. 3**,** pp. 248-252**.**

Sundin, G. W. & Jacobs, J. L. 1999, 'Ultraviolet radiation (uvr) sensitivity analysis and uvr survival strategies of a bacterial community from the phyllosphere of field-grown peanut (arachis hypogeae l.)', *An International Journal*, vol. 38**,** no. 1**,** pp. 27-38**.**

Suslick, K. 1990, 'Sonochemistry', *Science*, vol. 247**,** no. 4949**,** pp. 1439-1445**.**

Sykes, W. R., Webb, C. J. & Garnock-Jones, P. J. 1988. *Flora of new zealand: Naturalised pteridophytes, gymnosperms, dicotyledons*, Botany Division, D.S.I.R.

Tala, A., Lenucci, M., Gaballo, A., Durante, M., Tredici, S. M., Debowles, D. A., Pizzolante, G., Marcuccio, C., Carata, E., Piro, G., Carpita, N. C., Mita, G. & Alifano, P. 2013, 'Sphingomonas cynarae sp. Nov., a proteobacterium that produces an unusual type of sphingan', *International Journal of Systematic and Evolutionary Microbiology*, vol. 63**,** no. 1**,** pp. 72-79**.**

Tamas, I., Smirnova, A. V., He, Z. & Dunfield, P. F. 2013, 'The (d)evolution of methanotrophy in the beijerinckiaceae—a comparative genomics analysis', *The ISME Journal*, vol. 8**,** pp. 369**.**

Thomasma, S. M. & Foran, D. R. 2013, 'The influence of swabbing solutions on DNA recovery from touch samples', *Journal of Forensic Sciences*, vol. 58**,** no. 2**,** pp. 465-469**.**

Thompson, I., Bailey, M., Fenlon, J., Fermor, T., Lilley, A., Lynch, J., Mccormack, P., Mcquilken, M., Purdy, K. & Rainey, P. 1993, 'Quantitative and qualitative seasonal changes in the microbial community from the phyllosphere of sugar beet (beta vulgaris)', *Plant and Soil*, vol. 150**,** no. 2**,** pp. 177-191**.**

Tian, B. & Hua, Y. 2010, 'Carotenoid biosynthesis in extremophilic deinococcus–thermus bacteria', *Trends in Microbiology*, vol. 18**,** no. 11**,** pp. 512-520**.**

Turnbaugh, P. J., Hamady, M., Yatsunenko, T., Cantarel, B. L., Duncan, A., Ley, R. E., Sogin, M. L., Jones, W. J., Roe, B. A., Affourtit, J. P., Egholm, M., Henrissat, B., Heath, A. C., Knight, R. & Gordon, J. I. 2008, 'A core gut microbiome in obese and lean twins', *Nature*, vol. 457**,** no. 7228**,** pp. 480–484**.**

Vacher, C., Cordier, T. & Vallance, J. 2016, 'Phyllosphere fungal communities differentiate more thoroughly than bacterial communities along an elevation gradient', *Microbial Ecology*, vol. 72**,** no. 1**,** pp. 1-3**.**

Vadassery, J. & Oelmüller, R. 2009, 'Calcium signaling in pathogenic and beneficial plant microbe interactions: What can we learn from the interaction between piriformospora indica and arabidopsis thaliana', *Plant Signaling & Behavior*, vol. 4**,** no. 11**,** pp. 1024-1027**.**

Vandenkoornhuyse, P., Quaiser, A., Duhamel, M., Le Van, A. & Dufresne, A. 2015, 'The importance of the microbiome of the plant holobiont', *The New Phytologist*, vol. 206**,** no. 4**,** pp. 1196-1206**.**

Veith, N., Feldman-Salit, A., Cojocaru, V., Henrich, S., Kummer, U. & Wade, R. C. 2013, 'Organism-adapted specificity of the allosteric regulation of pyruvate

kinase in lactic acid bacteria', *PLoS Computational Biology,* vol. 9**,** no. 7**,** pp. e1003159**.**

Větrovský, T. & Baldrian, P. 2013, 'The variability of the 16s rrna gene in bacterial genomes and its consequences for bacterial community analyses', *PLoS ONE,* vol. 8**,** no. 2**,** pp. e57923**.**

Vogel, C., Bodenhausen, N., Gruissem, W. & Vorholt, J. A. 2016, 'The arabidopsis leaf transcriptome reveals distinct but also overlapping responses to colonization by phyllosphere commensals and pathogen infection with impact on plant health', *New Phytologist,* vol. 212**,** no. 1**,** pp. 192-207**.**

Vokou, D., Vareli, K., Zarali, E., Karamanoli, K., Constantinidou, H. I., Monokrousos, N., Halley, J. M. & Sainis, I. 2012, 'Exploring biodiversity in the bacterial community of the mediterranean phyllosphere and its relationship with airborne bacteria', *Microbial Ecology,* vol. 64**,** no. 3**,** pp. 714-724**.**

Vorholt, J. A. 2012, 'Microbial life in the phyllosphere', *Nature Reviews Microbiology,* vol. 10**,** no. 12**,** pp. 828-840**.**

Wang, J., Zhang, Y., Yang, X., Shen, J., Soininen, J. & Wang, B. 2011, 'Contrasting patterns in elevational diversity between microorganisms and macroorganisms', *Journal of Biogeography,* vol. 38**,** no. 3**,** pp. 595-603**.**

Wang, Q., Garrity, G. M., Tiedje, J. M. & Cole, J. R. 2007, 'Naive bayesian classifier for rapid assignment of rrna sequences into the new bacterial taxonomy', *Applied and Environmental Microbiology,* vol. 73**,** no. 16**,** pp. 5261-5267**.**

Wei, X., Lyu, S., Yu, Y., Wang, Z., Liu, H., Pan, D. & Chen, J. 2017, 'Phylloremediation of air pollutants: Exploiting the potential of plant leaves and leaf-associated microbes', *Frontiers in Plant Science,* vol. 8**,** no. 1318**,** pp. 1-23**.**

Wenk, M., Bourgeois, M., Allen, J. & Stucki, G. 1997, 'Effects of atrazine-mineralizing microorganisms on weed growth in atrazine-treated soils', *Journal of Agricultural and Food Chemistry,* vol. 45**,** no. 11**,** pp. 4474-4480**.**

Weyens, N., Thijs, S., Popek, R., Witters, N., Przybysz, A., Espenshade, J., Gawronska, H., Vangronsveld, J. & Gawronski, S. W. 2015, 'The role of plant–microbe interactions and their exploitation for phytoremediation of air pollutants', *International Journal of Molecular Sciences,* vol. 16**,** no. 10**,** pp. 25576-25604**.**

Whang, K. S., Lee, J. C., Lee, H. R., Han, S. I. & Chung, S. H. 2014, 'Terriglobus tenax sp. Nov., an exopolysaccharide-producing acidobacterium isolated from rhizosphere soil of a medicinal plant', *International Journal of Systematic and Evolutionary Microbiology,* vol. 64**,** no. 2**,** pp. 431-437**.**

Whipps, J. M., Hand, P., Pink, D. & Bending, G. D. 2008, 'Phyllosphere microbiology with special reference to diversity and plant genotype', *Journal of Applied Microbiology,* vol. 105**,** no. 6**,** pp. 1744-1755**.**

Whitaker, R., Grogan, D. & Taylor, J. 2003, 'Geographic barriers isolate endemic populations of hyperthermophilic archaea', *Science,* vol. 301**,** no. 5635**,** pp. 976-978**.**

Wicaksono, W. A., Jones, E. E., Monk, J. & Ridgway, H. J. 2016, 'The bacterial signature of leptospermum scoparium (mānuka) reveals core and accessory communities with bioactive properties', *PLoS ONE,* vol. 11**,** no. 9**.**

Wickham, H. 2007, 'Reshaping data with the reshape package', *Journal of Statistical Software,* vol. 21**,** no. 12**,** pp. 1-20**.**

Wickham, H. 2016. *Ggplot2: Elegant graphics for data analysis,* Springer.
Wickham, H. 2017, 'Tidyverse: Easily install and load'tidyverse'packages', *R package version,* vol. 1**,** no. 1**.**
Wickham, H. & Francois, R. 2015, 'Dplyr: A grammar of data manipulation', *R package version 0.4,* vol. 3**.**
Wickham, H., Francois, R. & Müller, K. 2017. Tibble: Simple data frames.
Wiens, J. J. 2011, 'The niche, biogeography and species interactions', *Philosophical Transactions of the Royal Society B: Biological Sciences,* vol. 366**,** no. 1576**,** pp. 2336-2350**.**
Williams, R. T. 1995, 'Lambert and mercator map projections in geology and geophysics', *Computers and Geosciences,* vol. 21**,** no. 3**,** pp. 353-364**.**
Wilson, M. & Lindow, S. E. 2000. Viable but nonculturable cells in plant-associated bacterial populations. *In:* COLWELL, R. R. & GRIMES, D. J. (eds.) *Nonculturable microorganisms in the environment.* Boston, MA: Springer US.
Winker, S. & Woese, C. R. 1991, 'A definition of the domains archaea, bacteria and eucarya in terms of small subunit ribosomal rna characteristics', *Systematic and Applied Microbiology,* vol. 14**,** no. 4**,** pp. 305-310**.**
Woese, C. R. & Fox, G. E. 1977, 'Phylogenetic structure of the prokaryotic domain: The primary kingdoms', *Proceedings of the National Academy of Sciences of the United States of America,* vol. 74**,** no. 11**,** pp. 5088-5090**.**
Woese, C. R., Kandler, O. & Wheelis, M. L. 1990, 'Towards a natural system of organisms: Proposal for the domains archaea, bacteria, and eucarya', *Proceedings of the National Academy of Sciences,* vol. 87**,** no. 12**,** pp. 4576-4579**.**
Woo, P. C., Lau, S. K., Teng, J. L., Tse, H. & Yuen, K. Y. 2008, 'Then and now: Use of 16s rdna gene sequencing for bacterial identification and discovery of novel bacteria in clinical microbiology laboratories', *Clinical Microbiology and Infection,* vol. 14**,** no. 10**,** pp. 908-934**.**
Woodward, F. I. & Lomas, M. R. 2004, 'Vegetation dynamics - simulating responses to climatic change', *Biological Reviews,* vol. 79**,** no. 3**,** pp. 643-670**.**
Xie, Y. 2013, 'Knitr: A general-purpose package for dynamic report generation in r', *R package version,* vol. 1**,** no. 7**,** pp. 1**.**
Xu, H. S., Roberts, N., Singleton, F. L., Attwell, R. W., Grimes, D. J. & Colwell, R. R. 1982, 'Survival and viability of nonculturableescherichia coli andvibrio cholerae in the estuarine and marine environment', *Microbial Ecology,* vol. 8**,** no. 4**,** pp. 313-323**.**
Yang, C.-H., Crowley, D. E., Borneman, J. & Keen, N. T. 2001, 'Microbial phyllosphere populations are more complex than previously realized', *Proceedings of the National Academy of Sciences of the United States of America,* vol. 98**,** no. 7**,** pp. 3889-3894**.**
Yashiro, E., Spear, R. N. & Mcmanus, P. S. 2011, 'Culture-dependent and culture-independent assessment of bacteria in the apple phyllosphere', *Journal of Applied Microbiology,* vol. 110**,** no. 5**,** pp. 1284-1296**.**
Yoon, J. H., Kang, S. J., Lee, S. Y. & Oh, T. K. 2008, 'Sphingomonas insulae sp. Nov., isolated from soil', *International Journal of Systematic and Evolutionary Microbiology,* vol. 58**,** no. 1**,** pp. 231-236**.**

Young, J. & Holland, I. B. 1999, 'Abc transporters: Bacterial exporters-revisited five years on', *Biochimica et Biophysica Acta (BBA) - Biomembranes*, vol. 1461**,** no. 2**,** pp. 177-200**.**

Zhang, L., Song, M., Cao, Q., Wu, S., Zhao, Y., Huang, J. W., Chen, K., Li, S. P., Xia, Z. Y. & Jiang, J. D. 2015, 'Camelimonas fluminis sp. Nov., a cyhalothrin-degrading bacterium isolated from river water', *International Journal of Systematic and Evolutionary Microbiology*, vol. 65**,** no. 9**,** pp. 3109-3114**.**

6. Appendices

6.1. Supplementary material for Chapter Four

Supplementary Table 6.1: Relative abundance (%) of phyla in the phyllosphere microbiome. Relative abundance is averaged per sample site.

Phylum	HT	KU	MK	MV	SL
Acidobacteria	26.39	11.7	23.46	18.59	1.49
Actinobacteria	0.86	0.5	0.35	1.14	1.51
Armatimonadetes	0.12	0.56	0.19	0.38	0.54
Bacteroidetes	4.4	8.03	8.44	12.43	23.2
Candidate_division_WPS_1	0.14	0.13	0.2	0.27	0.16
candidate_division_WPS-2	0	0.04	0	0.21	0
Candidatus_saccharibacteria	0	0	0	0.01	0
Chlamydiae	0	0.03	0.06	0.02	0
Chloroflexi	0	0	0	0	0
Cyanobacteria	0	0.15	0.01	0	0
Deinococcus_Thermus	0	0.25	0	0.55	0.76
Firmicutes	11.11	0.19	0.8	0.04	1.35
Gemmatimonadetes	0	0	0	0	0
Nitrospirae	0	0	0	0	0
Planctomycetes	0.53	0.86	0.45	1.01	0.58
Proteobacteria	41.94	50.6	49.85	51.4	61.5
Synergistetes	0	0	0	0	0
Tenericutes	0.02	0.05	0.1	0.08	0.12
Thaumarchaeota	0	0	0	0	0
unknown	8.25	24.92	13.05	8.42	8.07
Verrucomicrobia	6.21	1.97	3.06	5.46	0.72

Supplementary Table 6.2: Core OTU 16S gene amplicon sequences

OTU2	TACGAAGGGGGCTAGCGTTGTTCGGATTTACTGGGCGTAAAG CGCACGTAGGCGGAAATTTAAGTCAGGGGTGAAATCCCGAGG CTCAACCTCGGAACTGCCTTTGATACTGGGTTTCTTGAGTCCG GAAGAGGTGAGTGGAACTGCGAGTGTAGAGGTGAAATTCGTA GATATTCGCAAGAACACCAGTGGCGAAGGCGGCTCACTGGTC CGGAACTGACGCTGAGGTGCGAAAGCGTGGGGAGCAAACAG GATTAGATACCCTGGTAGTCCACGCCGTAAACGATGGATGCTA GCCGTTGGATCGCTTGCGATTCAGTGGCGCAGTTAACGCTTTA AGCATCCCGCCT
OTU4	TACGAAGGGGGCTAGCGTTGTTCGGAATCACTGGGCGTAAAG GGTGCGTAGGCTGGTTCTTAAGTCAGGGGTGAAATCCCGAGG CTCAACCTCGGAACTGCCTTTGATACTGGGGATCTTGAGTCCG GAAGAGGTGAGTGGAACTGCGAGTGTAGAGGTGAAATTCGTA GATATTCGCAAGAACACCGGTGGCGAAGGCGGCTCACTGGTC CGGAACTGACGCTGAGGCACGACAGCGTGGGGAGCAAACAG GATTAGATACCCTGGTAGTCCACGCCGTAAACGATGAATGCC AGCCGTTGGGTGATTTATCACTGAGTGGCGCAGCTAACGCTTT GAGCATTCCGCCT
OTU10	TACAGAGGCCTCAAGCGTTGTTCGGATTCATTGGGCGTAAAGG GTGCGTAGGCGGTCGGGTAAGTCGGATGTGAAATCCTGGGGC TCAACCTCAGAACTGCATTCGATACTGCTTGGCTAGAGGACTG GAGAGGAGATCGGAATTCACGGTGTAGCAGTGAAATGCGTAG ATATCGTGAGGAAGACCAGTGGCGAAGGCGGATCTCTGGACA GTTCCTGACGCTGATGCACGAAGGCTAGGGGAGCAAACGGGA TTAGATACCCCGGTAGTCCTAGCAGTAAACGGTGCACGTTTGG TTTGGGAGGATTCGACCCCTTCCGAGCCGGAGCTAACGCGTTA AACGTGCCGC
OTU11	TACGGAGGGAGCTAGCGTTGTTCGGAATTACTGGGCGTAAAG CGCACGTAGGCGGCTTTGTAAGTTAGAGGTGAAAGCCTGGAG CTCAACTCCAGAATTGCCTTTAAGACTGCATCGCTTGAATCCA GGAGAGGTGAGTGGAATTCCGAGTGTAGAGGTGAAATTCGTA GATATTCGGAAGAACACCAGTGGCGAAGGCGGCTCACTGGAC TGGTATTGACGCTGAGGTGCGAAAGCGTGGGGAGCAAACAGG ATTAGATACCCTGGTAGTCCACGCCGTAAACGATGATAACTAG CTGTCCGGGCACTTGGTGCTTGGGTGGCGCAGCTAACGCATTA AGTTATCCGCC
OTU27	TACGGAGGGAGCTAGCGTTATTCGGAATTACTGGGCGTAAAG CGCACGTAGGCGGCTTTGTAAGTTAGAGGTGAAAGCCCAGAG CTCAACTTTGGAATTGCCTTTGATACTGCATCGCTTGAATCCA GGAGAGGTGAGTGGAATTCCGAGTGTAGAGGTGAAATTCGTA GATATTCGGAAGAACACCAGTGGCGAAGGCGGCTCACTGGAC

	TGGTATTGACGCTGAGGTGCGAAAGCGTGGGGAGCAAACAGG ATTAGATACCCTGGTAGTCCACGCCGTAAACGATGATAACTAG CTGTCCGGGGACTTGGTCTTTGGGTGGCGCAGCTAACGCATTA AGTTATCCGCC
OTU40	TACGGAGGGTGCGAGCGTTGTCCGGATTTATTGGGTTTAAAGG GTGCGTAGGCGGCCTTGTAAGTCCGGGGTGAAAGCCCGCTGC TCAACAGCGGAACTGCCCTGGATACTGCGGGGCTTGAATACA GTGGAGGTTGGCGGAATGGACCGAGTAGCGGTGAAATGCATA GATACGGTCCAGAACCCCGATTGCGAAGGCAGCTGACTACAC TGTGATTGACGCTGAGGCACGACAGCGTGGGGAGCGAACAGG ATTAGATACCCTGGTAGTCCACGCCGTAAACGATGGATACTCG CTGCGGGCTAAAGATTGTCCGTGGCTTAGGGAAACCGATAAG TATCCCACCTGG
OTU75	TACGAAGGGGGCTAGCGTTGTTCGGAATCACTGGGCGTAAAG GGTGCGTAGGCGGATCCATCAGTCAGGGGTGAAATCCCGAGG CTCAACCTCGGAACTGCCTTTGATACTGTGGATCTTGAGTCCG GAAGAGGTGAGTGGAACTGCGAGTGTAGAGGTGAAATTCGTA GATATTCGCAAGAACACCGGTGGCGAAGGCGGCTCACTGGTC CGGAACTGACGCTGAGGCACGACAGCGTGGGGAGCAAACAG GATTAGATACCCGTGGTAAGTCCACGCCGTAAACGATGAATG CCAGCCGTTGGGTGATTTATCACTGAGTGGCGCAGCTAACGCT TTGAGCATTCCGC
OTU132	TACGAAGGGGGCTAGCGTTGTTCGGAATCACTGGGCGTAAAG GGTGCGTAGGCGGGCATTTAAGTCAGGGGTGAAATCCCGAGG CTCAACCTCGGAACTGCCCTTGATACTGGATGTCTTGAGTCCG GAAGAGGTGAGTGGAACTGCGAGTGTAGAGGTGAAATTCGTA GATATTCGCAAGAACACCGGTGGCGAAGGCGGCTCACTGGTC CGGAACTGACGCTGAGGCACGACAGCGTGGGGAGCAAACAG GATTAGATACCCTGGTAGTCCACGCCGTAAACGATGAATGCC AGCCGTTGGGTGATTTATCACTGAGTGGCGCAGCTAACGCCTT TGGAGCATTCCGC
OTU323	TACGGAGGGTGCGAGCGTTGTCCGGATTTATTGGGTTTAAAGG GTGCGTAGGCGGCCCGTTAAGTCCGGGGTGAAAGCCCGTTGC TCAACAACGGAACTGCCCTGGATACTGACGGGCTTGAGTACA GACGAGGTTGGCGGAATGGACCGAGTAGCGGTGAAATGCATA GATACGGTCCAGAACCCCGATTGCGAAGGCAGCTGACTAGGC TGTTACTGACGCTGAGGCACGACAGCGTGGGGAGCGAACAGG ATTAGATACCCTGGTAGTCCACGCCGTAAACGATGGATACTCG CTGCGAGCTAAAGATTGTTCGTGGCTTAGGGAAACCGATAAG TATCCCACCTGG
OTU842	TACGAGGGGGGCAAGCGTTGTTCGGAATTATTGGGCGTAAAG GGCGCGTAGGCGGTTTGGCAAGTTTCGTGTGAAATCTTCGGGC TCAACTCGAAGTCTGCACGGAAAACTGCCGGGCTTGAGTATG

	GGAGAGGTGAGTGGAATTTCCGGTGTAGCGGTGAAATGCGTA GATATCGGAAGGAACACCTGTGGCGAAAGCGGCTCACTGGAC CATAACTGACGCTGAGGCGCGAAAAGCTAGGGGAGCAAACAG GATTAGATACCCTGGTAGTCCTAGCCCTAAACGATGATTGCTT GATGTGGCGGGTATCCAATCCCGCCGTGTCGAAGCTAACGCG ATAAGCAATCCG

Supplementary Table 6.3: Sample site spatial and environmental metadata

Site Information	**MK**	**SL**	**MV**	**KU**	**HT**
Sample date	22.11.16	03.12.16	18.12.16	28.12.16	02.01.17
GPS - WGS 1984 (G1762) (degree minutes) ** Rounded to minutes to obscure exact sample location*	39°01 S 177°08 E	37°56 S 175°19 E	40°57 S 175°26 E	38°02 S 176°03 E	39°06 S 176°21 E
Average temperature (°C)	18.29296	16.33686	17.87932	14.48121	15.5301
Average day temperature (°C)	22.72776	22.26717	21.1	17.46792	16.74974
Average night temperature (°C)	13.46	9.703659	14.35683	11.34	14.49146
Day-night temperature differential (°C)	9.267755	12.56352	6.743171	6.127917	2.258273
Relative humidity day average	43.72776	60.4737	52.16909	72.47458	78.6
Photosynthetically active radiation ($\mu mol\ m^{-2}\ s^{-1}$)	1375.684	903.7609	862.5023	827.5208	604.2105
Monthly average temperature (°C)	19.0	17.0	15.5	18.0	18.0
Monthly total precipitation (mm)	49.9	111.6	91.25	64.8	56.3
Monthly average wind speed (mph)	10.1	19.45	7.4	8.5	10.1
Monthly average pressure (mb)	1007.1	1009.8	1008.4	1014.3	1013.9
Monthly average cloud cover (%)	31.0	58.0	42.0	53.0	44.0
Monthly average humidity (%)	67.0	83.0	79.0	77.0	76.0
Monthly total sun (hr)	130.0	86.5	104.9	96.5	110.0

Supplementary Table 6.4: Mānuka sample tree and branch metadata

Site	Sample ID	Branch height (m)	Tree diameter (cm)	Branch aspect (°)	Tree ID	Elevation (m)	Tree height (m)
MK	MK2.4_03	1.03	2.3	135	MK2	357	2.7
	MK2.5_34	1.83	2.3	145			
	MK2.6_08	1.55	2.3	0			
	MK3.4_04	2.04	6.6	71	MK3	354	2.7
	MK3.5_36	2.25	6.6	120			
	MK3.6_37	2.54	6.6	111			
	MK4.4_06	2.2	4.0	80	MK4	358	3.18
	MK4.5_40	2.8	4.0	0			
	MK4.6_44	1.8	4.0	45			
	MK5.4_07	1.62	1.5	69	MK5	356	1.82
	MK5.5_41	1.28	1.5	128			
	MK5.6_45	1.4	1.5	0			
	MK6.4_86	1.8	3.0	139	MK6	357	2.52
	MK6.5_46	2.0	3.0	359			
	MK6.6_49	2.2	3.0	0			
	MK9.4_09	2.25	3.55	55	MK9	343	2.86
	MK9.5_50	2.6	3.55	60			
	MK9.6_51	2.53	3.55	75			
SL	SL1.4_10	2.4	8.9	91	SL1	66	4
	SL1.5_55	1.8	8.9	306			
	SL1.6_53	1.75	8.9	339			
	SL2.4_11	2.35	8.5	110	SL2	66	3.9
	SL2.5_58	2.1	8.5	118			
	SL2.6_56	2.25	8.5	160			
	SL3.4_12	2.15	12.8	93	SL3	66	4.1
	SL3.5_60	2.0	12.8	98			
	SL3.6_57	2.05	12.8	123			
	SL4.4_14	1.4	4.4	130	SL4	66	3.3
	SL4.5_61	1.7	4.4	140			
	SL4.6_05	1.25	4.4	150			
	SL7.4_13	1.9	4.5	12	SL7	66	3.3
	SL7.5_66	2.0	4.5	155			
	SL7.6_81	2.2	4.5	173			
	SL8.4_85	1.7	3.7	313	SL8	66	2.2
	SL8.5_03	1.8	3.7	9			
	SL8.6_34	1.4	3.7	75			
MV	MV1.4_04	1.85	9.7	270	MV1	245	3.3
	MV1.6_37	1.95	9.7	235			
	MV2.4_06	2.3	3.9	85	MV2	237	2.65
	MV2.5_40	2.3	3.9	10			
	MV2.6_44	2.35	3.9	110			
	MV3.4_07	2.6	4.4	195	MV3	240	4

	MV3.5_41	2.8	4.4	195			
	MV3.6_45	3.3	4.4	195			
	MV4.4_08	2.7	7.6	130	MV4	247	3.4
	MV4.5_46	2.6	7.6	160			
	MV4.6_49	2.6	7.6	0			
	MV5.4_09	2.05	3.5	280	MV5	246	2.85
	MV5.5_50	2.1	3.5	290			
	MV5.6_51	2.35	3.5	300			
	MV6.4_10	2.45	4.1	0	MV6	242	2.9
	MV6.5_55	2.7	4.1	120			
	MV6.6_53	2.0	4.1	145			
KU	KU1.4_11	1.15	1.5	300	KU1	561	2.3
	KU1.5_58	0.8	1.5	300			
	KU1.6_56	1.5	1.5	310			
	KU2.4_12	1.7	3.6	260	KU2	562	3.8
	KU2.5_60	2.1	3.6	10			
	KU2.6_57	3.1	3.6	175			
	KU3.4_14	1.9	2.3	65	KU3	562	2.5
	KU3.5_61	2.1	2.3	210			
	KU3.6_59	2.0	2.3	150			
	KU4.4_15	2.1	7.35	330	KU4	555	4
	KU4.5_63	2.5	7.35	235			
	KU4.6_25	2.3	7.35	70			
	KU5.4_22	1.9	15.75	120	KU5	555	3.7
	KU5.5_64	2.1	15.75	160			
	KU5.6_16	2.4	15.75	210			
	KU6.4_68	1.55	2.4	280	KU6	555	2.5
	KU6.5_18	1.7	2.4	15			
	KU6.6_65	2.0	2.4	30			
HT	HT2.4_17	1.2	3.5	205	HT2	638	2.8
	HT2.5_88	1.8	3.5	25			
	HT26_86	1.2	3.5	200			
	HT3.4_71	1.3	6.3	225	HT3	633	2.7
	HT3.5_19	1.5	6.3	250			
	HT3.6_89	1.1	6.3	210			
	HT5.4_72	1.0	7.2	230	HT5	632	2.8
	HT5.5_20	0.9	7.2	145			
	HT5.6_90	1.6	7.2	10			
	HT6.4_78	1.1	8.8	45	HT6	632	2.6
	HT6.5_21	1.55	8.8	115			
	HT6.6_91	1.7	8.8	45			
	HT7.4_80	1.2	17.6	205	HT7	638	3.1
	HT7.5_05	1.0	17.6	80			
	HT7.6_13	1.5	17.6	75			
	HT9.4_23	1.1	24.75	340	HT9	635	3.3
	HT9.5_24	1.2	24.75	25			
	HT9.6_48	0.6	24.75	345			

Supplementary Table 6.5: Summary table of Pearson's correlation coefficient association test between the relative abundance of dominant phyla in the mānuka phyllosphere and spatial, environmental, and sample tree metadata
Significant positive correlations greater than 0.5 are underlined. Significant negative correlations less than 0.5 are in italics. P values less than 0.05 are in bold text. The symbol (m) represents monthly averages obtained from the NIWA database (NIWA, 2017).

Phyla	**Bacteroidetes**		**Proteobacteria**		**Acidobacteria**		**Verrucomicrobia**		**Firmicutes**	
	Corr	*P test*	*Corr*	*P test*	*Corr*	*P test*	*Corr*	*P test*	*Corr*	*P test*
Elevation (m)	*-0.76*	**0.001**	*-0.54*	**0.001**	0.51	**0.001**	0.30	**0.004**	0.30	**0.004**
Latitude (UTM)	0.19	0.072	0.20	0.064	-0.43	**0.001**	-0.46	**0.001**	0.01	0.96
Longitude (UTM)	*-0.58*	**0.001**	-0.36	**0.001**	0.57	**0.001**	0.13	0.23	0.12	0.31
Average temperature (°C)	0.14	0.19	0.087	0.42	0.22	**0.035**	0.11	0.30	-0.17	0.18
Day temperature (°C)	0.53	**0.001**	0.41	**0.001**	-0.24	**0.024**	-0.25	**0.020**	-0.31	**0.003**
Night temperature (°C)	*-0.60*	**0.001**	*-0.50*	**0.001**	0.75	**0.001**	0.57	**0.001**	0.22	**0.001**
Day/night differential (°C)	0.71	**0.001**	0.56	**0.001**	*-0.58*	**0.001**	*-0.50*	**0.001**	-0.35	**0.001**
Relative humidity (%)	-0.25	**0.017**	-0.23	**0.027**	-0.014	0.90	0.12	0.32	0.33	**0.001**
PAR (µmol m^{-2} s^{-1})	0.094	0.38	0.17	0.12	0.044	0.68	-0.23	**0.032**	-0.30	**0.004**
Temperature (*m*) (°C)	-0.39	**0.0001**	-0.21	**0.05**	0.26	**0.016**	-0.12	0.26	0.12	0.26

Precipitation (*m*) (mm)	0.76	**0.001**	0.50	**0.001**	0.64	**0.001**	-0.24	**0.022**	-0.18	0.086
Wind (*m*) (mph)	0.18	0.087	0.088	0.42	0.070	0.51	0.24	**0.023**	-0.18	0.10
Pressure (*m*) (mb)	-0.33	**0.002**	-0.25	**0.02**	-0.01	0.92	0.061	0.57	0.27	**0.011**
Cloud (*m*) (%)	0.47	**0.001**	0.33	**0.002**	*-0.67*	**0.001**	-0.32	**0.003**	-0.031	0.77
Humidity (*m*) (%)	0.53	**0.001**	0.31	**0.003**	*-0.57*	**0.001**	-0.13	0.24	-0.01	0.92
Sun (*m*) (hrs)	*-0.53*	**0.001**	-0.35	**0.001**	0.67	**0.001**	0.27	0.012	0.069	0.52
Tree height (m)	0.26	**0.015**	0.086	0.42	-0.22	**0.045**	-0.021	0.85	-0.090	0.4
Branch height (m)	0.15	0.16	0.22	**0.04**	-0.11	0.33	-0.14	0.18	-0.28	**0.0088**
Tree diameter (cm)	0.57	-0.06	-0.22	**0.04**	0.20	0.055	0.42	**0.001**	0.025	0.82
Branch aspect (°)	0.035	0.75	-0.15	0.17	-0.050	0.65	0.12	0.29	0.056	0.61

Supplementary Table 6.6: Relative abundance and ratio of enrichment of KEGG pathways in phyllosphere and soil communities.

KEGG Pathway	Leaves	Soil	Ratio of enrichment
1,1,1-Trichloro-2,2-bis(4-chlorophenyl)ethane (DDT) degradation	0.002	0.003	0.663
ABC transporters	2.996	2.538	1.181
Adipocytokine signaling pathway	0.066	0.104	0.633
Alanine, aspartate and glutamate metabolism	0.872	0.984	0.886
alpha-Linolenic acid metabolism	0.03	0.033	0.914
Amino acid metabolism	0.266	0.267	0.994
Amino acid related enzymes	1.191	1.256	0.948
Amino sugar and nucleotide sugar metabolism	1.147	1.299	0.883
Aminoacyl-tRNA biosynthesis	1.02	0.998	1.022
Aminobenzoate degradation	0.264	0.441	0.598
Antigen processing and presentation	0.032	0.016	2.042
Apoptosis	0.017	0.018	0.934
Arachidonic acid metabolism	0.078	0.092	0.845
Arginine and proline metabolism	1.242	1.218	1.02
Ascorbate and aldarate metabolism	0.128	0.18	0.711
Atrazine degradation	0.072	0.046	1.567
Bacterial chemotaxis	0.34	0.592	0.575
Bacterial motility proteins	1.375	1.648	0.834
Bacterial secretion system	0.683	0.77	0.887
Bacterial toxins	0.093	0.111	0.839
Base excision repair	0.419	0.531	0.789
Benzoate degradation	0.219	0.455	0.48
beta-Alanine metabolism	0.261	0.405	0.645

beta-Lactam resistance	0.045	0.078	0.576
Betalain biosynthesis	0.004	0.014	0.304
Biosynthesis and biodegradation of secondary metabolites	0.061	0.055	1.101
Biosynthesis of ansamycins	0.043	0.064	0.675
Biosynthesis of siderophore group nonribosomal peptides	0.04	0.052	0.774
Biosynthesis of unsaturated fatty acids	0.255	0.337	0.758
Biosynthesis of vancomycin group antibiotics	0.074	0.072	1.016
Biotin metabolism	0.154	0.116	1.326
Bisphenol degradation	0.065	0.12	0.543
Butanoate metabolism	0.606	0.944	0.642
Butirosin and neomycin biosynthesis	0.061	0.073	0.839
C5-Branched dibasic acid metabolism	0.28	0.314	0.891
Caffeine metabolism	0.005	0.002	2.516
Calcium signaling pathway	0.021	0	79.803
Caprolactam degradation	0.089	0.215	0.415
Carbohydrate digestion and absorption	0.025	0.002	10.443
Carbohydrate metabolism	0.235	0.163	1.438
Carbon fixation in photosynthetic organisms	0.669	0.484	1.382
Carbon fixation pathways in prokaryotes	0.831	1.078	0.771
Carotenoid biosynthesis	0.2	0.073	2.756
Cell cycle - Caulobacter	0.57	0.488	1.167
Cell division	0.069	0.062	1.107
Cell motility and secretion	0.311	0.334	0.932
Cellular antigens	0.067	0.1	0.677
Chaperones and folding catalysts	1.073	0.903	1.188

Chloroalkane and chloroalkene degradation	0.285	0.298	0.957
Chlorocyclohexane and chlorobenzene degradation	0.119	0.134	0.887
Chromosome	1.468	1.161	1.264
Circadian rhythm - plant	0.014	0.022	0.647
Citrate cycle (TCA cycle)	0.536	0.807	0.664
Cyanoamino acid metabolism	0.253	0.291	0.867
Cysteine and methionine metabolism	0.749	0.846	0.885
Cytoskeleton proteins	0.328	0.196	1.676
D-Alanine metabolism	0.076	0.08	0.949
D-Arginine and D-ornithine metabolism	0.006	0.008	0.683
D-Glutamine and D-glutamate metabolism	0.137	0.13	1.057
Dioxin degradation	0.062	0.06	1.037
DNA repair and recombination proteins	2.516	2.395	1.05
DNA replication	0.469	0.456	1.03
DNA replication proteins	0.875	0.794	1.102
Drug metabolism - cytochrome P450	0.265	0.18	1.477
Drug metabolism - other enzymes	0.159	0.246	0.645
Electron transfer carriers	0.027	0.023	1.155
Energy metabolism	0.64	0.834	0.767
Ether lipid metabolism	0.033	0.051	0.645
Ethylbenzene degradation	0.048	0.061	0.774
Fatty acid biosynthesis	0.526	0.614	0.857
Fatty acid metabolism	0.46	0.667	0.689
Flagellar assembly	0.342	0.732	0.467
Flavone and flavonol biosynthesis	0.005	0.018	0.286
Flavonoid biosynthesis	0.088	0.046	1.905
Fluorobenzoate degradation	0.079	0.078	1.018

Folate biosynthesis	0.485	0.449	1.08
Fructose and mannose metabolism	0.658	0.608	1.084
Function unknown	1.541	1.631	0.945
Galactose metabolism	0.382	0.582	0.657
General function prediction only	3.855	3.698	1.043
Geraniol degradation	0.134	0.309	0.433
Glutamatergic synapse	0.086	0.088	0.973
Glutathione metabolism	0.527	0.386	1.363
Glycan biosynthesis and metabolism	0.068	0.054	1.258
Glycerolipid metabolism	0.329	0.336	0.981
Glycerophospholipid metabolism	0.479	0.564	0.848
Glycine, serine and threonine metabolism	0.755	0.983	0.768
Glycolysis / Gluconeogenesis	1.08	1.033	1.046
Glycosaminoglycan degradation	0.023	0.075	0.31
Glycosphingolipid biosynthesis - ganglio series	0.016	0.043	0.358
Glycosphingolipid biosynthesis - globo series	0.037	0.094	0.388
Glycosyltransferases	0.563	0.436	1.294
Glyoxylate and dicarboxylate metabolism	0.623	0.767	0.812
Histidine metabolism	0.508	0.598	0.85
Homologous recombination	0.7	0.609	1.149
Indole alkaloid biosynthesis	0.003	0.014	0.188
Inorganic ion transport and metabolism	0.241	0.422	0.57
Inositol phosphate metabolism	0.176	0.236	0.746
Insulin signaling pathway	0.153	0.074	2.063
Ion channels	0.037	0.036	1.031
Isoquinoline alkaloid biosynthesis	0.059	0.082	0.718
Limonene and pinene degradation	0.165	0.322	0.511

Linoleic acid metabolism	0.023	0.062	0.363
Lipid biosynthesis proteins	0.726	0.879	0.826
Lipid metabolism	0.157	0.131	1.193
Lipoic acid metabolism	0.11	0.072	1.522
Lipopolysaccharide biosynthesis	0.291	0.502	0.58
Lipopolysaccharide biosynthesis proteins	0.343	0.586	0.585
Lysine biosynthesis	0.562	0.61	0.921
Lysine degradation	0.23	0.432	0.532
Lysosome	0.038	0.142	0.268
MAPK signaling pathway - yeast	0.035	0.033	1.072
Meiosis - yeast	0.096	0.06	1.594
Melanogenesis	0.002	0.001	3.131
Membrane and intracellular structural molecules	0.673	0.767	0.877
Metabolism of cofactors and vitamins	0.178	0.202	0.882
Metabolism of xenobiotics by cytochrome P450	0.248	0.154	1.618
Methane metabolism	1.086	0.939	1.156
Mismatch repair	0.646	0.611	1.058
Naphthalene degradation	0.181	0.222	0.815
N-Glycan biosynthesis	0.074	0.054	1.357
Nicotinate and nicotinamide metabolism	0.41	0.397	1.034
Nitrogen metabolism	0.673	0.78	0.863
Nitrotoluene degradation	0.051	0.109	0.467
Non-homologous end-joining	0.04	0.106	0.378
Novobiocin biosynthesis	0.111	0.139	0.798
Nucleotide excision repair	0.307	0.365	0.843
Nucleotide metabolism	0.003	0.011	0.313
One carbon pool by folate	0.45	0.452	0.997

Other glycan degradation	0.11	0.242	0.455
Other ion-coupled transporters	0.886	0.915	0.969
Other transporters	0.279	0.226	1.239
Others	1.251	1.123	1.114
Oxidative phosphorylation	1.549	1.629	0.951
Pantothenate and CoA biosynthesis	0.598	0.564	1.061
Penicillin and cephalosporin biosynthesis	0.052	0.12	0.434
Pentose and glucuronate interconversions	0.347	0.558	0.621
Pentose phosphate pathway	0.788	0.794	0.993
Peptidases	1.938	1.761	1.1
Peptidoglycan biosynthesis	0.646	0.609	1.061
Peroxisome	0.226	0.281	0.805
Phenylalanine metabolism	0.268	0.409	0.654
Phenylalanine, tyrosine and tryptophan biosynthesis	0.713	0.717	0.995
Phenylpropanoid biosynthesis	0.155	0.17	0.912
Phosphatidylinositol signaling system	0.122	0.093	1.313
Phosphonate and phosphinate metabolism	0.018	0.047	0.381
Phosphotransferase system (PTS)	0.042	0.051	0.82
Photosynthesis	1.545	0.309	5.004
Photosynthesis - antenna proteins	0.349	0.002	208.077
Photosynthesis proteins	1.93	0.326	5.927
Plant-pathogen interaction	0.1	0.13	0.769
Polycyclic aromatic hydrocarbon degradation	0.128	0.14	0.916
Polyketide sugar unit biosynthesis	0.227	0.18	1.257
Pores ion channels	0.695	0.532	1.308
Porphyrin and chlorophyll metabolism	1.673	0.791	2.116
PPAR signaling pathway	0.162	0.204	0.792

Prenyltransferases	0.463	0.338	1.369
Propanoate metabolism	0.568	0.914	0.622
Proteasome	0.033	0.031	1.071
Protein export	0.526	0.522	1.008
Protein folding and associated processing	0.939	0.796	1.18
Protein kinases	0.592	0.432	1.368
Protein processing in endoplasmic reticulum	0.106	0.074	1.443
Proximal tubule bicarbonate reclamation	0.039	0.062	0.634
Purine metabolism	1.999	1.903	1.05
Pyrimidine metabolism	1.303	1.291	1.009
Pyruvate metabolism	1.064	1.018	1.046
Replication, recombination and repair proteins	0.949	0.735	1.292
Restriction enzyme	0.106	0.096	1.106
Retinol metabolism	0.116	0.084	1.378
Riboflavin metabolism	0.323	0.304	1.065
Ribosome	1.872	1.743	1.074
Ribosome Biogenesis	1.025	0.989	1.037
Ribosome biogenesis in eukaryotes	0.035	0.038	0.93
RNA degradation	0.555	0.436	1.273
RNA polymerase	0.156	0.13	1.202
RNA transport	0.086	0.08	1.084
Secretion system	1.632	1.625	1.005
Selenocompound metabolism	0.357	0.423	0.843
Signal transduction mechanisms	0.551	0.492	1.119
Sphingolipid metabolism	0.11	0.214	0.512
Sporulation	0.052	0.043	1.222
Starch and sucrose metabolism	0.868	0.846	1.026

Steroid biosynthesis	0.053	0.049	1.091
Steroid hormone biosynthesis	0.037	0.052	0.722
Stilbenoid, diarylheptanoid and gingerol biosynthesis	0.052	0.046	1.124
Streptomycin biosynthesis	0.398	0.351	1.133
Styrene degradation	0.058	0.116	0.502
Sulfur metabolism	0.374	0.37	1.011
Sulfur relay system	0.412	0.348	1.184
Synthesis and degradation of ketone bodies	0.066	0.151	0.437
Taurine and hypotaurine metabolism	0.115	0.136	0.843
Terpenoid backbone biosynthesis	0.524	0.49	1.07
Tetracycline biosynthesis	0.171	0.156	1.093
Thiamine metabolism	0.419	0.37	1.132
Toluene degradation	0.177	0.21	0.843
Transcription factors	1.103	1.481	0.745
Transcription machinery	0.778	1.14	0.683
Transcription related proteins	0.003	0.006	0.498
Translation factors	0.434	0.42	1.034
Translation proteins	0.808	0.729	1.108
Transporters	4.898	4.432	1.105
Tropane, piperidine and pyridine alkaloid biosynthesis	0.11	0.123	0.891
Tryptophan metabolism	0.353	0.606	0.582
Two-component system	2.174	2.197	0.989
Tyrosine metabolism	0.413	0.471	0.877
Ubiquinone and other terpenoid-quinone biosynthesis	0.525	0.307	1.709
Ubiquitin system	0.032	0.04	0.783
Valine, leucine and isoleucine biosynthesis	0.726	0.731	0.994

Valine, leucine and isoleucine degradation	0.478	0.829	0.577
Vitamin B6 metabolism	0.182	0.157	1.163
Xylene degradation	0.053	0.036	1.487