



UPPSALA
UNIVERSITET

UPTEC X 14 038

Examensarbete 30 hp
September 2014

The foliar bacterial endophyte community in native *Pinus radiata*: a role for protection against fungal disease?

Arielle Reivant Munters



UPPSALA
UNIVERSITET

Degree Project in Bioinformatics

Masters Programme in Molecular Biotechnology Engineering,
Uppsala University School of Engineering

UPTEC X 14 038		Date of issue 2014-09	
Author Arielle Reivant Munters			
Title (English) The foliar bacterial endophyte community in native <i>Pinus radiata</i>: a role for protection against fungal disease?			
Title (Swedish)			
Abstract <i>Pinus radiata</i> is the most planted tree in the southern hemisphere. The planted trees are especially susceptible to pathogens, but even the native population, nowadays limited to merely five locations, are threatened by diseases caused by arthropods, fungi and dehydration. Endophytes are bacteria or fungi that reside inside healthy plant tissue, and often have a beneficial effect on their hosts. Endophytes can help plants adapt to abiotic stress such as drought and protect them against pathogens and insect pests. Given the roles that endophytes play in host stress responses, it is possible that without studying endophytes we may not fully understand a plant's response to increased temperatures and climate-induced disease. Using Illumina-sequencing of the 16S rRNA-gene the bacterial endophyte community in 15 trees from three of the remaining native populations were studied. By investigating trees from several sites geographical community differences were discovered. The three overall most dominating bacterial taxa can all be connected with genera known to contain members with anti-fungal properties.			
Keywords Bacterial endophytes, 16S rRNA, conifers, <i>Pinus radiata</i> , Monterey pine			
Supervisors Carolyn Frank University of California Merced			
Scientific reviewer Jon Ågren Uppsala Universitet			
Project name		Sponsors	
Language English		Security	
ISSN 1401-2138		Classification	
Supplementary bibliographical information		Pages 31	
Biology Education Centre Box 592, S-751 24 Uppsala		Biomedical Center Tel +46 (0)18 4710000	Husargatan 3, Uppsala Fax +46 (0)18 471 4687

The foliar bacterial endophyte community in native *Pinus radiata*: a role for protection against fungal disease?

Arielle Reivant Munters

Populärvetenskaplig sammanfattning

Montereytallen är det mest planterade trädet i södra hemisfären. Planterade träd är särskilt känsliga för sjukdomsangrepp, men även ursprungspopulationen, begränsad till tre platser i centrala Kalifornien och två öar utanför Mexiko, är hotad av bl.a. sjukdomar orsakade av insekter, svampangrepp och torka. Liksom vi människor har en mag-tarmflora med tusentals bakteriearter som är viktiga för vår hälsa, så har växter också viktiga bakterier inuti och utanpå alla sina vävnader. Endofyter är bakterier eller svampar som huserar i alla växter, ofta med gynnsam effekt för värden. Endofyter kan hjälpa värdväxten med allt från kvävefixering och skydd mot sjukdomsframkallande svampar till att stimulera tillväxt genom produktion av växthormon. Endofyter spelar en avgörande roll i många växters stressrespons och näringsupptag. Det är därför möjligt att studier av endofyter kan bidra till en ökad förståelse av växters reaktioner vid stress och förhoppningsvis i framtiden skydda växter bättre mot torka, näringsbrist, svamp- och insektsangrepp.

Den bakteriella endofytpopulationen i 15 Montereytallar från de tre ursprungspopulationerna i Kalifornien fastställdes genom Illumina-sekvensering av 16S rRNA genen. Genom att undersöka träd från olika platser kunde vi urskilja geografiska skillnader mellan populationerna. Dessa skillnader kan bero på genetiska skillnader i värden, eller skillnader i den lokala miljön, t.ex. jordmånen. De tre mest dominerande taxa av bakterier i alla proverna är identiska eller nära släkt med bakterier som tidigare bevisats ha negativ effekt på sjukdomsframkallande svampar. Detta kan vara en viktig del av pusslet om hur ursprungspopulationerna klarat sig i förhållande till de planterade Montereytallarna.

Examensarbete 30 hp

Civilingenjörsprogrammet Molekylär bioteknik: inriktning bioinformatik

Uppsala Universitet, september 2014

Table of Contents

Introduction	7
Overview.....	7
Monterey Pine	8
Endophytes.....	8
To study uncultured microbial diversity.....	10
Methods.....	12
Sampling & Sterilization	12
DNA extraction.....	12
DNA amplification	13
Sequence analysis.....	13
Results	17
Recovered phylotypes	17
Difference in location.....	18
Dominating taxa	20
Discussion.....	24
Acknowledgement	26
References.....	27

Introduction

Overview

Today the biodiversity is changing at an unparalleled rate due to climate change. Large numbers of species are becoming extinct due to long term changes in environmental conditions that we as humans have created [1]. These changes induce an environmental stress on plants that they are not always equipped to handle. Endophytes—bacteria or fungi inside healthy plant tissue—can help plants adapt to abiotic stress such as drought [2]–[4] and protect them against pathogens and insect pests [2], [5], [6]. Given the roles that endophytes play in host stress responses, it is possible that without studying endophytes we may not fully understand a plant’s response to increased temperatures and climate-induced disease.

The majority of the research on endophytic bacteria has been focused on agronomical important plants [7] and invasive species [8]. Less is known of the endophytes that live inside our native plants. Studying endophytes in native plants, where plant-endophyte symbioses originated, will increase our knowledge about the symbiosis as well as plant physiology, including stress tolerance, transpiration, carbon- and nitrogen cycling throughout the plant. Even though little research has been done on endophytes in native system such as forest trees, it is recognised that conifers associate with mycorrhizal fungi (i.e. the symbiotic association between fungi and roots of certain plants) [9] and foliar endophytic fungi [10], [11]. In contrast, the knowledge of bacterial endophytes in conifers is still very limited [12].

Monterrey pine (*Pinus Radiata*) is the most widely planted conifer in the world but in its native habitat it is limited and threatened [13]–[15]. Monterey pine in the USA is restricted to three distinct areas of central-coastal California, and is currently threatened by loss of habitat, drought, fungal infections and insects [14], [15]. The native population has shown less susceptibility to certain fungal diseases than the planted trees [14]. Still, given the restricted habitat of Monterey pine, threats to the native stands in California jeopardize the genetic diversity of the species, and could decrease its ability to cope with climate change. It is possible that bacterial endophytes mediate how Monterey pine responds to environmental change. In order to understand their contribution, we need to first describe the endophytic community and identify potential symbionts.

Unfortunately, only a fraction of bacteria in a given community can be cultured [16], [17]. Instead, high-throughput sequencing of the 16S ribosomal RNA can be used to study the unknown biodiversity of bacteria in the environment. Such surveys can identify ‘core’ bacterial species that are consistently associated with a host species, e.g. across individuals, habitats, seasons, environments and developmental stages [18]–[21], and therefore likely to

provide specialized beneficial functions to the host. 16S rRNA pyrosequencing was recently used to characterize the endophytic community in limber pine (*Pinus flexilis*) needles, unravelling a potential nitrogen-fixing symbiosis [22].

This study analyses the microbial communities inside the foliage of Monterey pine from the three populations in central-coastal California. Illumina sequencing was used to identify symbiotic endophytic bacteria (i.e. those consistently associated with Monterey pine), and to identify differences in the endophytic community among the three populations.

Monterey Pine

P. radiata has a wide range of uses and while it is considered endangered in its native range [23] it is the most planted softwood tree in the world [15]. It is native to California, USA and Baja California, Mexico but is planted all over the southern hemisphere [13]. There are five small native habitats of *P. radiata*: three on the California mainland (Año Nuevo, Monterey, and Cambria) (**Figure 1**) and the two in the Mexican provenances on Guadalupe and Cedros islands. While the mainland provenances are all of the variety *radiata*, the var. *binata* originate from Guadalupe and var. *cedrosensis* grows on Cedros Island [15]. Even though the three mainland populations all belong to the same variety, the Cambria population has several phenotypic differences (e.g. difference in branch, cone and needle characteristics) compared to the population in Año Nuevo and Monterey [24]. In Australia stands that originate from the Cambria population have showed a diverging seasonal growth patterns and a slower growth rate [24]. These differences indicate that the stands in Cambria differs genetically from the other two mainland populations [24]. Several studies have tried to find the genetic relationship between the mainland stands but have gotten opposing results, and no clear relationship have been found [25]. Today the mainland provenances are highly threatened by loss of habitat, fire suppression, dehydration, fungal infections and insects [14], [15].

Monterey pine has suffered greatly by the fungi *Gibberella circinata* (anamorph = *Fusarium circinatum*) or pitch canker [14], [26]–[28]. *G. circinata* is a pathogenic fungi that increases tree mortality, generates deformed stems, reduced growth and a higher rate of seedling mortality [26]. The pathogen spreads between trees by insect vectors and airborne inoculums [26], [28]. Trees under stress have also show to have a greater susceptibility to the disease [26]. The native trees have a higher resistance than the planted populations [27]. There have also been evidence that an induced resistance now exists in the planted populaions [28]. The variance in resistance is a clear example of why it is important to study and conserve the natural habitats of *P. radiata*.

Endophytes

The word endophyte was first coined in 1866 by De Barry [7], [29] and is a combination between the two word ενδο (endo) and φυτό (phyte): endo means inside or within and

phyte means plant [29], [30]. 150 years after De Barry defined endophyte as an organism inside another organism, the definition have been specified into non-pathogenic bacteria and fungi that reside inside healthy plant tissue [2], [3]. Endophytes do not cause harm to its host, on the contrary bacterial endophytes are often beneficial for its host. Bacterial endophytes have been found in all plants studied to date [31], they have been found in pollen, root, stem, leaves, buds, flower and seed tissues [7]. In some plants, the root tissue contains the highest concentration of endophytes and the number of bacteria decreases with the distance from the root [7], [32].

Bacterial endophytes can be obligate or facultative [33]. The transmission of obligate endophytes is vertical since they are completely dependent on its host and can not survive outside their host. Vertical transmitted endophytes are transferred from parental-tissue, before germination into the new plant. The plant inherits (part of) their endophytic community [2]. Another source of endophytic bacteria is horizontal transfer from the host's surroundings. The main source of entry is thought to be the rhizosphere, the soil closest to the root of a plant that is affected by the plant's nutrients, hormones etc. [7], although endophyte transmission has not been studied in trees. Endophytic bacteria can also colonize the host horizontally through the stomata (i.e. the pores on the stem, leaves through which gas exchange occur), tissue wounds and abrasions or foliar damage from windblown soil particles [2], [7], [17], [34]. Facultative endophytes can survive outside their hosts and therefore and can be either horizontally or vertically transmitted [2], [3], [7].

Colonization success is determined by several factors, e.g. the mode of entry, intrinsic factors such as flagella [17] and membrane bound lipopolysaccharides [30]. The relative importance of host genotype and other factors differ among plants. The endophytic community can vary with plant age [35], [36], sampling season [36], tissue type [36], salinity [37] and host genotype [35], [36], [38].

When comparing aseptic and endophytic inoculated plants, large differences in growth rate and general health have been observed [7], [32]. Plant-endophyte co-cultures have shown to have growth promoting effect on plant length, biomass, lignification of xylem vessels, root and leaf-hair formations and nodules production [2]. Endophytic production of plant growth proteins, such as auxin, cytokinin and gibberellin, activates growth and increases the host's growth rate [2], [3]. Endophytes can also indirectly promote growth by stimulating plant production of growth hormones [2], [4], [6].

By protecting its host against phytopathogens (fungi, virus and bacteria), endophytes influence the host's health. Endophytes can physically hinder the pathogen by blocking otherwise vulnerable tissue, produce toxins to fight the pathogens and help produce structural compounds when its host are threatened [2], [6].

When plants are not optimally fitted to their environment, they may experience stress [2]. Increasing plants stress tolerance would increase the plants ability to survive despite high stress (e.g. drought, heat, and salinity) and losses would be reduced [2]. When a plant is exposed to stress, an increase of the stress hormone ethylene production occurs, which can inhibit plant growth[2], [3], [12]. Endophytes are known to hinder the plant's ethylene production by the endophytic produced 1-aminocyclopropane-1-carboxylate (ACC) deaminase activity. By suppressing ethylene production endophytes can increase plant stress tolerance [3], [4].

Endophytes can also assist the plant in nutrient acquisition. *Saccharum officinarum* (sugarcane) is one of very few species found to obtain a majority of their nitrogen from biological nitrogen fixation (BNF) without nodulation, [32]. The exact species which fix nitrogen have not been established but studies indicate that there are several endophytes with potential nitrogen fixation genes such as *nifH*, the gene encoding the enzyme nitrogenase that is responsible for nitrogen fixation[32].

Thus, endophytes are not only non-pathogenic for plants but essential to grow a healthy plant. Studying endophytic biodiversity increases our knowledge of not only the bacteria but also plant physiology.

To study uncultured microbial diversity

With next-generation-sequencing a microbial richness [39] and diversity [40] previously unknown have been discovered. The 16S ribosomal RNA (rRNA) gene have been used to map the human microbiome [39], the gut microbiome of insects [41] and many environmental samples [22], [36], [37], [42], [43]. It is the standard gene to use for phylogenetic studies and to taxonomically assign the bacteria [39].

The rRNA molecules consist of highly conserved regions interlaced with more variable stretches and are comprised of the small subunit (16S in Prokaryote, 18S in Eukaryote) and the large subunit (23S and 28S). By comparing the rRNA molecules, relationship between organisms can be deduced [44]. When sequencing the small subunit of ribosomal RNA (16S) large microbial communities can be identified[40], [44].

16S rRNA surveys using next-generation sequencing produces long “lists” of bacterial taxa present in samples, along with the relative abundance of each taxon. Still, given the high number of bacterial species present in most environments, such lists are only estimates of the true microbial diversity. Using methods to estimate how species diversity and richness, one can evaluate how well a sample represents reality. Species diversity estimates the number of species taking the abundances into account, while species richness simply represent how many different species a dataset/sample contains without abundances [40], [45]. Shannon index calculates an entropy that estimates species diversity. If an individual

is randomly chosen from the dataset, the Shannon index quantifies the uncertainty in species identity [45]. Chao1 estimates species richness by adding a correcting factor for observed number of species to the observed number of species [40]. Microbial data is often unequally sampled and rarefaction, randomly re-sampling to create subsets, makes it possible to compare diversity and richness despite the uneven sampling. By creating random subsets of a dataset a rarefied curve is the average of the observed accumulation curve, or number of species as the function of the number of samples. When the plot reaches an asymptote the full richness or diversity has been captured [40].

This study investigates the bacterial endophyte communities inside *P. radiata* var. *binata* from the three native mainland populations along the Californian coastline using Illumina sequencing of the 16S ribosomal RNA gene. The aim of the study was to investigate if there is a conserved bacterial community across individual trees or if the population varies within the native population. Conserved bacterial taxa may be symbionts with roles in the host's response to climate- and disease stress.

Methods

Sampling & Sterilization

Needles and buds from *P. radiata* were collected in February 2014 at the three native sites along California Pacific coast; Swanton Ranch in Año Nuevo, Point Lobos in Monterey and Cambria (SR, PL, CR) (**Figure 1**). From each site five trees were sampled. To assess differences in the endophytic community between tissue types, both needles and buds were collected. Not all trees had bud formation and this is likely due to the current drought situation in California [46], and therefore the bud samples for each location were pooled. In total 15 needle samples and three pooled bud samples, were collected. The tissue was removed with sterile razors and placed in sterile bags to be shipped to Merced, California at 4°C for processing. The samples were then surface sterilized to remove surface-bacteria. The samples were soaked in 30 % hydrogen peroxide for 3 min, then 10x washes in sterile autoclaved water, each 1 min and then stored at -20°C. The last rinse was saved to confirm sterility of samples.



Figure 1: Map over the native range of Monterey pine (green). The three sampling locations are indicated with arrows.

DNA extraction

Each sample was granulated into a fine powder with Fisher Scientific™ PowerGen™ Cryogenic Homogenizer with liquid nitrogen. All material used were autoclaved and UV-sterilized for 30 min prior to use. In addition to 0.6 g tissue powder, 800 µL CTAB solution (1 mL CTAB buffer, 0.04 g polyvinylpyrrolidone 5 µL 2-mercaptoethanol) were added into 2 mL screw cap tubes. The samples were then vortexed and placed in a dry bath at 60°C for 2h with intermittent vortexing. To homogenize the samples, 0.3 g of 0.11 mm sterile glass beads were added prior to bead beating for 3 min. After bead beating the samples were centrifuged for 2 min at 16000rcf and the supernatant were transferred with equal amount of chloroform into 2 mL snap cap tubes. The tubes were inverted and then centrifuged for 10 min at 16000rcf. The protein removal step was repeated to ensure that all protein was removed. To precipitate nucleic acids the aqueous top phase was transferred into sterile 2mL snap cap tubes together with 1/10 volume cold 3M sodium acetate. Before placing sample is -20°C for 12h an equal amount (as total volume in the tube) of cold isopropanol were added. The samples were centrifuged at 4°C for 30 min 16000rcf after 12h in -20°C. The supernatant were decanted and 700 µL of 70% ethanol were added. The samples were

centrifuged at 4°C for 15 min at 16000rcf before allowing the pellets to air-dry inside a sterile laminar flow PCR-hood. To resuspend the dry pellets 50µL of DNA resuspension fluid (1.0 M Tris-HCl, 0.1 M EDTA) was added and then the samples were stored in -20°C.

DNA amplification

Since the 16S ribosomal gene (16S) is homologous to the chloroplast 16S and the mitochondria 18S rRNA genes, a method to separate out chloroplast and mitochondria from the small fraction bacterial DNA was used [47]. The chloroplast excluding forward primer 799f (AACMGGATTAGATACCCCKG) together with the 16S reverse primer 1492r (TACGGHTACCTTGTTACGACTT) creates size difference between the mitochondrial and bacterial product, this can then be separated electrophoretically [42], [47]–[49]. The samples was amplified using a nested PCR with Golay-barcodes, forward primer 799f and reverse primer 1492r using the thermocycle profile described in [48]. Barcoded primers are used to be able to differentiate different samples when multiplex sequencing, parallel sequencing of several samples, is executed [50][50][48] and a nested PCR reduces the amount of chloroplast sequences and improves consistency [47]. To extract the bacterial product (750 bp) from mitochondrial product (1000 bp) [42] a 2% (w/v) agarose-gel stained with ethidium bromide. The bacterial extract were then thermo cycled using the profile described in [48] with the barcoded primer set 799f and the universal bacterial primer 1115r (AGGGTTGCGCTCGTTG). The amplicons were then cleaned, quantified using Nanodrop [51] and then pooled. The amplicons were sequenced at the University of California Davis Genome Center, Davis CA, USA on an Illumina MiSeq machine.

Sequence analysis

The sequences were analysed using the Qiime package [52]. The forward and reverse reads were joined using fastq-join [53]. To demultiplex, connect barcoded reads with the sample they derived from, the Illumina data split_libraries_fastq.py was used. The command split_libraries_fastq.py does a quality filtering were it removes reads that have less than 75 high-quality bases in a row by default. The reads were shorter than expected and the acceptance value was set to 50 high-quality bases in row instead. It resulted in 2263676 sequences matching the barcodes with a mean length of 301 after truncation.

The sequences were clustered into phylotypes against Greengenes 13.8 16S reference database [54] using UCLUST [55] with a minimum identity of 97 %. The sequences that did not get a hit in the reference database were clustered against each other without any external reference. Then the sequences were aligned using PyNAST [56] and columns containing only gaps were removed. Then taxonomic assignments were performed with the Ribosomal Database Project (RDP) classifier [57] against Greengenes database [54]. Contamination sequences classified as either “Chloroplast” or “Mitochondria” (0.9 %) were removed. To remove chimeras ChimeraSlayer [58] was used. An average of 9 % of the

sequences was removed from each sample after being classified as contaminants or chimeric sequences. An approximately maximum-likelihood master-tree was constructed using FastTree [59]. A phylotype-table was constructed containing the taxonomic identities and abundances. The resulting number of sequences are listed in **Table 1**.

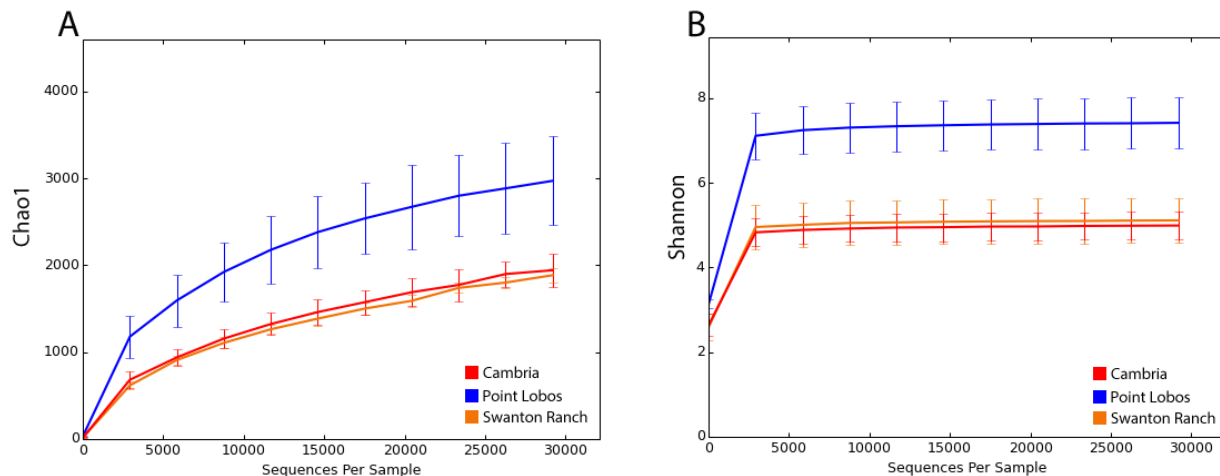


Figure 2: Alpha rarefaction curves for the three locations. A) Chao1 index: the lack of an asymptote indicates that the sequencing depth doesn't represent the full phylotype richness in the samples. The Point Lobos samples have higher species richness and differentiate significantly from lower richness of the similar Cambria and Swanton samples. B) Shannon: The curves reach an asymptote early and that indicates that the full diversity has been captured at all locations. Point Lobos have a higher Shannon index which indicates that the diversity is higher than in Cambria and Swanton samples.

To compare species richness and diversity levels, the phylotype-table was rarefied to a sequences depth of 29,190, to enable all samples to be able to be sampled equally. To estimate species richness Chao1 was used, since it is considered to be able to handle bias towards low abundances (i.e. microbial samples) especially good [40]. Chao1 and Shannon indexes were calculated on each phylotype-table (**Figure 2**). The relative abundance of bacterial classes in each sample (**Figure 5A**), in samples from the different location (**Figure 5B**), and in samples from the two tissue types (**Figure 5C**) was calculated using Qiime.

To compare the communities based on their composition, a random subsample of 29190 sequences were used to calculate beta diversity. Using the previously constructed master-tree, weighted and unweighted UniFrac [60] distance matrices were calculated. The unweighted UniFrac distance represent the fraction of total branch length that is unshared, while the weighted UniFrac distance takes the relative abundances into account when calculating the distances between two communities [61]. A weighted PCoA-plot is good to find underlying community differences that stem from difference in the founding community, while unweighted UniFrac distances can discern difference when the overall group of organisms present in each samples are static [61]. To compare the samples Principle Coordinate analysis (PCoA) plots (**Figure 3**) were constructed. To estimate the

uncertainty of the PCoA-plots, a workflow including jackknife replicates was used. A subsample of 21892 (75% of 29190) was used to create ten rarefied phylotype-tables. Using the master-tree and the rarefied phylotype-tables UniFrac [60] distance matrices (weighted and unweighted) were calculated, PCoA were then performed on each of the distance matrices. The PCoA-plots were then compared to each other (**Figure 4**).

The heatmap (**Figure 6**) illustrates the fraction of top ten phylotypes in each sample and was created by using the biom-table generated in Qiime inside RStudio [62]. First the absolute abundance was converted into relative abundances (per sample), then all sequences with abundances higher than 6.0 % was plotted into a heatmap using gplots [63]. The colour-breaks in the colour-key were determined based on the sample distribution at 25 %, 50 % and 75 % (1%, 4% and 9% of total abundance in each sample). From these ten OTUs, a representative sequence from each was extracted. To determine the similarity with known isolates, a homology search using BLAST against NCBI 16S rRNA database was conducted for each of the phylotypes (**Table 3**).

Table 1: Metadata of samples collected, and number of sequences in each sample after removal of plant DNA and chimeras used in downstream analysis.

Location	Tissue type	Sample	No. of sequences
Cambria Reserve	Buds	CR.bud	137387
	Needles	CR.T1	145241
		CR.T2	245482
		CR.T3	150452
		CR.T4	112137
Point Lobos	Buds	PL.bud	122496
	Needles	PL.T1	105240
		PL.T2	29190
		PL.T3	60192
		PL.T4	29929
Swanton Ranch	Buds	SR.bud	171470
	Needles	SR.T2	139513
		SR.T3	106460
		SR.T4	66649
		SR.T5	37199
		SR.T6	195284

Table 2: Average number of sequences recovered based on location and tissue type.

Average no. of sequences	
Cambria Reserve	151050
Point Lobos	74256
Swanton Ranch	119429
Buds	143784
Needles	109137

Results

Recovered phylotypes

From 18 samples a total of 2,068,413 sequences spread out across 8202 phylotypes were recovered after removing of chimeras and plant DNA. On average each sample yielded/generated 114911 sequences (**Table 1**). The number of sequences varied with tissue type. 38 % more sequences were generated from bud tissue samples (**Table 2**). The Point Lobos samples yielded the smallest amount of sequences but the highest amount of observed species. Swanton Ranch samples showed a 14 % increase in number of sequences compared to Point Lobos. Cambria Reserve samples yielded 103 % more sequences than Point Lobos (**Table 2**). On average Point Lobos- and needle-samples contains fewer sequences in comparison with the other two locations and buds (**Table 2**). While Point Lobos generated the smallest amount of sequences, the species richness curve (**Figure 2B**) showed that more species were observed. The rarefaction lacked an asymptote and therefore does these result show that the sequencing depth did not capture the full species richness (**Figure 2**).

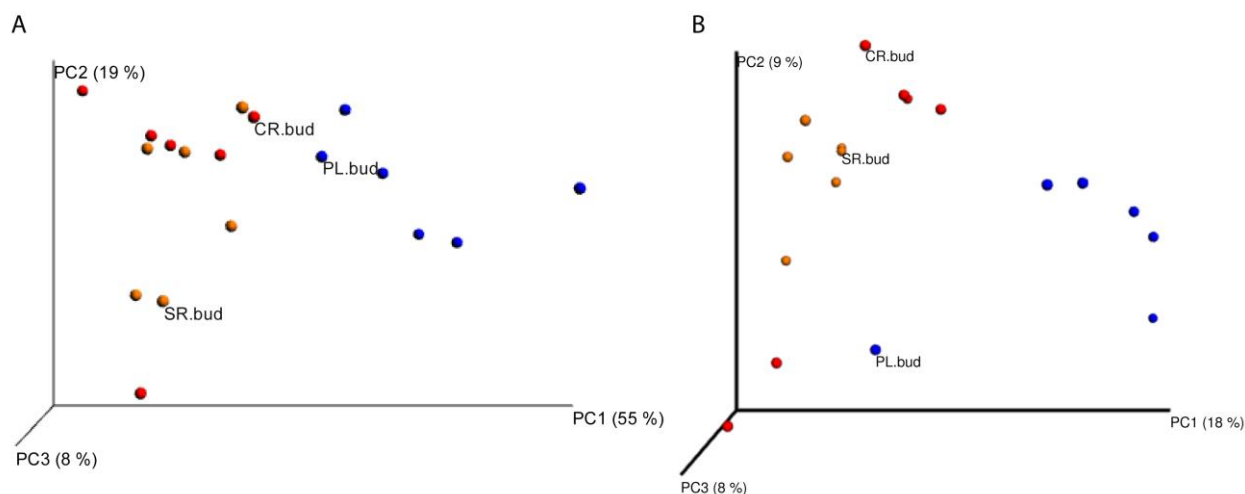


Figure 3: The weighted (A) and unweighted (B) PCoA plot represent distances between samples. Each colour corresponds to a sampling location, the red represent Cambria Reserve (CR) samples, blue Point Lobos (PL) samples and the oranges Swanton Ranch (SR) samples. The distance between points corresponds to similarity between bacterial endophyte communities, i.e. points close together have a similar community structure. Point Lobos samples create a separate cluster from the nested Cambria Reserve/Swanton Ranch samples. (A) The weighted PCoA show differences based on transitory changes in relative abundances. Swanton ranch and Cambria reserve samples are more nested than Point Lobos samples. (B) Unweighted PCoA plot exposes differences in communities based on differences in founding populations. Point Lobos bud samples are separated from the Point Lobos needle samples while the bud samples from Swanton and Cambria are more similar to their needle samples.

Difference in location

Location appears to contribute the endophytic community PCoA-plots (**Figure 3**). Point Lobos samples in the unweighted PCoA-plot (**Figure 3B**) shows a higher grade separation than in the weighted PCoA-plot (**Figure 3A**). But both PCoA-plots (**Figure 3**) show that the Point Lobos samples separate from Cambria and Swanton ranch samples. In the weighted PCoA-plot (**Figure 3A**) the Point Lobos samples all cluster together which indicates that part of what differentiates Point Lobos samples from Cambria reserve and Swanton ranch samples are due to differences in the founding communities, that the source of bacteria in Monterey pine at Point Lobos differs from both Swanton Ranch and Cambria Reserve. The Swanton Ranch and Cambria reserve samples are clustered together indicating that the community variances are not based on differences stemming from deviating sources of bacteria. In the unweighted PCoA-plot (**Figure 3B**) a separation of Point Lobos samples can still be distinguished, suggesting that part of the differences displayed is due to differences in what can survive in each sample. Ten jackknife-replicates were used to estimate the uncertainty of the UniFrac distances (**Figure 4**). The variances of the weighted samples were minimal (**Figure 4A**) indicating that the original PCoA-plot (**Figure 3A**) is correct. In the unweighted PCoA-plot (**Figure 4B**) the uncertainty is greater than the weighted, as expected. While the Point Lobos samples still are clustered together, the variance of the Cambria reserve and Swanton ranch bud-samples makes them hard to distinguish from each other.

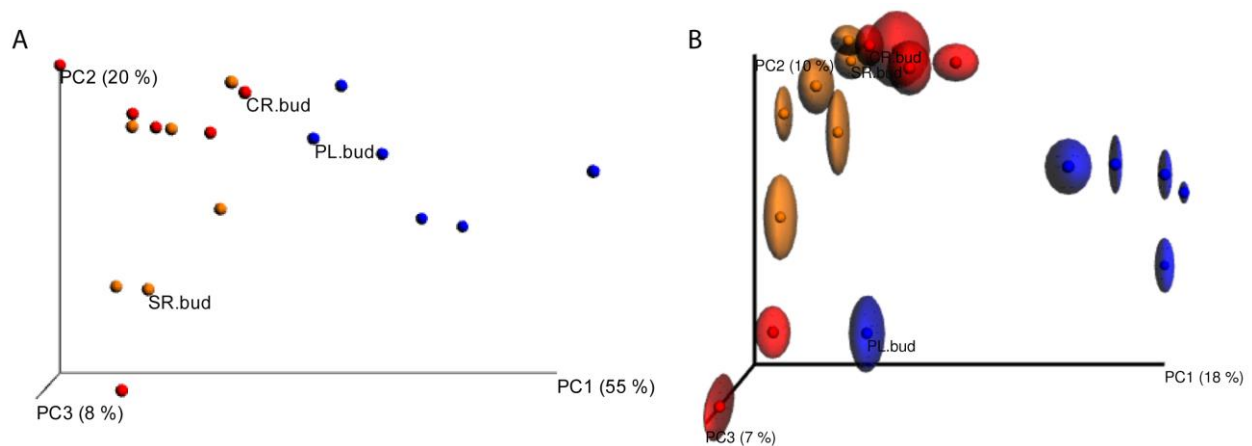


Figure 4: The variance of the PCoA-plots. The confident ellipsoids represent data variation of the UniFrac distances and were calculated by using ten jackknife replicates. (A) The variance is minimal in the weighted PCoA-plot, and no relevant uncertainty can be discerned. (B) The uncertainty is greater in the unweighted PCoA-plot. The variance of the Cambria reserve and Swanton ranch bud-samples overlaps each other and they cannot be distinguished from each other. While all samples have a larger variance than the weighted distances the Point Lobos samples are still separate from the other two locations. Cambria reserve = CR, Point Lobos = PL and Swanton ranch = SR.

Point Lobos samples diverge in both the PCoA-plot (**Figure 3**) and when comparing the relative abundances (**Figure 4**), this most likely due to the low abundance of the two

overall dominating phylotypes (**Figure 6**). There is a low similarity in the phylotype distribution between individual samples and between buds and needles (**figure 5A, C**), while a clearer difference can be seen between locations (**Figure 5B**). In particular, Point Lobos samples have a lower relative abundance of Betaproteobacterial sequences. This coincides with the PCoA result (**Figure 3**) where Point Lobos samples cluster separately from the remaining data samples.

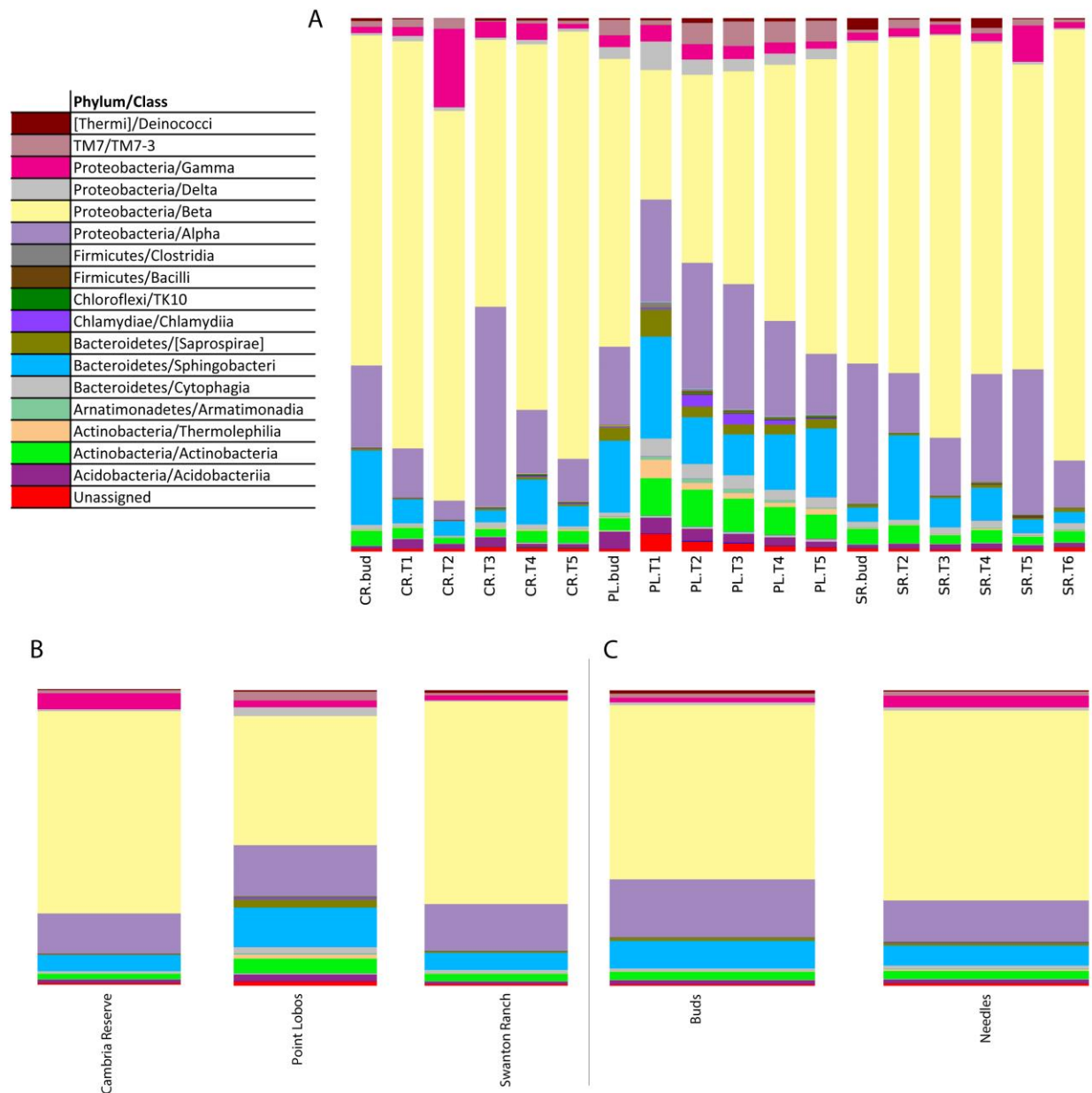


Figure 5: Relative abundance of major bacterial classes calculated as the percentage of sequences belonging to a particular lineage from each of (A) samples, (B) sampling location or (C) tissue type. The dominating phylum is Proteobacteria and a majority of Proteobacteria is Betaproteobacteria, followed by Alphaproteobacteria. Samples marked by location (CR = Cambria reserve, PL = Point Lobos, SR = Swanton ranch), all samples consists needle tissue except those marked bud. (A) Betaproteobacteria is dominating all samples. (B) Point Lobos bacterial community differs from Swanton Ranch and Cambria Reserve. (C) Slightly more Alphaproteobacteria was found in the bud samples compared to needle samples. Alpha = Alphaproteobacteria, Beta = Betaproteobacteria, Delta = Deltaproteobacteria and Gamma = Gammaproteobacteria.

Dominating taxa

The overall abundant phylum is Proteobacteria (73.1-92.0%), followed by Bacteroidetes (3.2-27.4%) and the dominating class at all locations is Betaproteobacteria (CR: 68.6%, PL:

43.7%, SR: 68.4%) closely followed by Alphaproteobacteria (CR: 15.7%, PL: 17.0%, SR: 13.5%) (**Figure 5A, B**). The heatmap displays the ten most abundant phylotypes in all samples (**Figure 6**). Eight out of the ten phylotypes belong to Proteobacteria and the remaining two are Bacteroidetes. The three most dominating phylotypes in the samples all belong to the bacterial family *Oxalobacteraceae* (**Figure 6 and Table 3**). *Oxalobacteraceae* have previously been isolated from the phyllosphere, rhizosphere and root from several agricultural plants and deciduous trees [64]. *Oxalobacteraceae* have been connected with several plant beneficial factors such as production of the plant hormone auxin, production of siderophores (used for the acquisition of iron), and production of antifungals [64]–[67]. While the three dominating phylotypes were present in all samples, their relative abundance differs among locations. Phylotype 200536 was most prominent in SR samples (CR 8.2%, PL 7.9% and SR 19.3% on average), phylotype 104023 was most prominent in CR samples (CR 18.9%, PL 7.5% and SR 10.0% on average), and phylotype 972341 was prominent at both SR and CR samples (CR 19.9%, PL 11.1% and SR 30.4% on average).

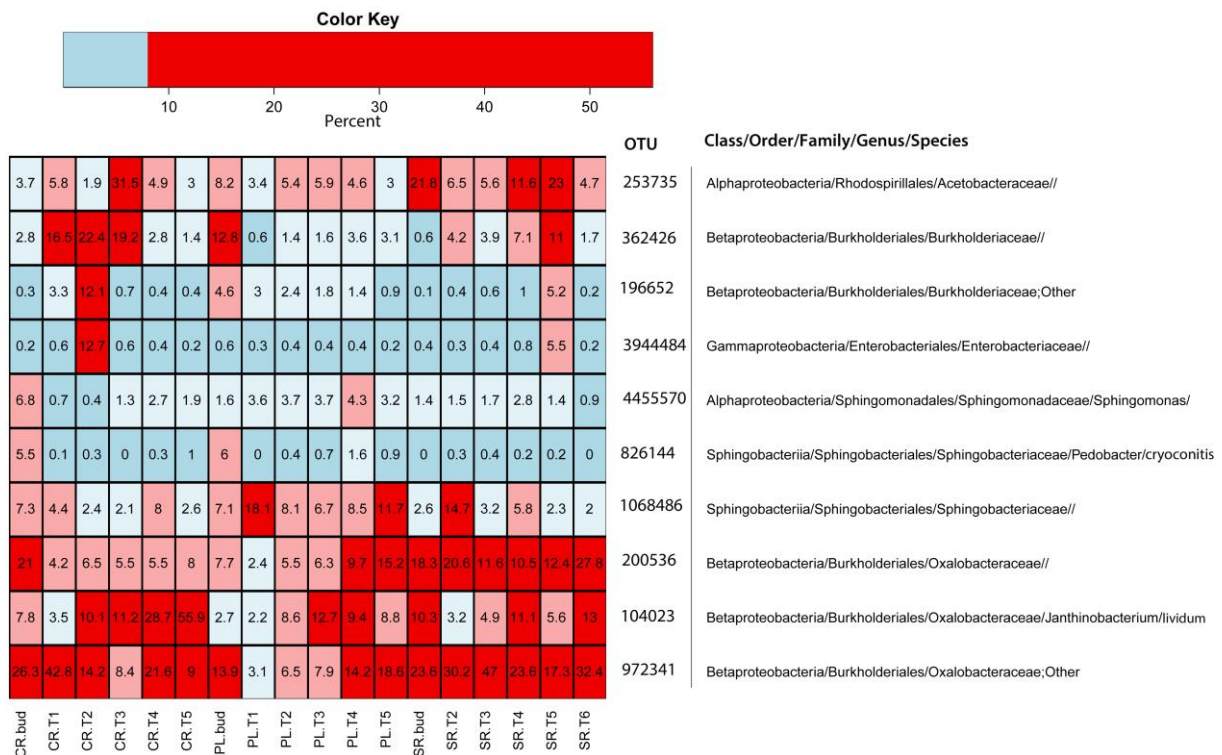


Figure 6: Heatmap over the top ten most abundant phylotypes in each samples. The numbers represent the fraction each p occurs in each sample, the darker red the higher abundance. OTU = phylotype cluster, Class/Order/Family/Genus/Species as defined by the taxonomic alignment done against Greengenes 16S database.

Phylotype 104023 makes up more than half of the recovered phylotypes in CR.T5 (**Figure 6**) and was 100% identical to *Janthinobacterium lividum*, identified as strain DSM1522 by Blast homology search (**Table 3**). *J. lividum* is a bacterial species with well-known

antifungal properties isolated from amphibian skin and in mycorrhiza [65], [66]. *J. lividum* produces violacenin, a compound that have been found to have antifungal, antibacterial, antiviral and antitumoral properties [65]. The samples isolated from mycorrhizal fungi spores have showed strong antifungal properties against several plant pathogenic-fungi and is capable of phosphorous solubilisation [66].

Phylotype 200536 and 97234 both have been identified as belonging to the genus *Massilia* (**Table 3**). *Massilia* have been characterized as a aerobic, gram-negative flagellated rod-shaped non-spore forming bacteria. Species in this genus may have a high tolerance of abiotic stress since many have been isolated from nutrient-poor environments, heavy metal contaminated sites and arid environments [64]. Phylotype 972341 had the overall highest relative abundance (3-47%) and was identified as *Massilia brevitalea* strain byr 23-80 with a 99% similarity. The phylotype 200536 is consistently high in Swanton Ranch samples and was 99% sequence identity to both *Massilia* and *Herbaspirillum* species(**Table 3 and Figure 6**).

Phylotype 253735 was identified as *Kozakia baliensis* (**Table 3**). This species have recently been found as a part of the core set of endophytic bacteria in both *P. flexilis* and *Pinus engelmannii* (Engelmann spruce) [22].

Table 3: The BLAST result of ten most dominating phylotypes with the sequence identity percent and e-value.

Phylotype	Strain	Accession	Ident%	e-value
253735	<i>Kozakia baliensis</i> strain Yo-3	NR_024773.1	97	9,00E-146
	<i>Kozakia baliensis</i> strain NBRC 16664	NR_113858.1	97	9,00E-146
362426	<i>Burkholderia sordidicola</i>	NR_041916.1	99	7,00E-152
196652	<i>Burkholderia phenazinium</i> strain LMG 2247	NR_118071.1	99	3,00E-155
	<i>Burkholderia phenazinium</i> strain A 1	NR_029212.1	99	3,00E-155
3944484	<i>Rahnella aquatilis</i> HX2 strain HX2	NR_074921.1	99	1,00E-154
	<i>Rahnella aquatilis</i> strain DSM 4594	NR_025337.1	99	1,00E-154
4455570	<i>Sphingomonas echinoides</i> strain NBRC 15742	NR_113806.1	99	2,00E-152
	<i>Sphingomonas glacialis</i> strain C16y	NR_117270.1	99	2,00E-152
826144	<i>Pedobacter piscium</i> strain NBRC 14985	NR_113717.1	100	3,00E-155
	<i>Sphingobacterium antarcticum</i> strain 4BY	NR_104917.1	100	3,00E-155
	<i>Pedobacter westerhofensis</i> strain WB 3.3-22	NR_042602.1	100	3,00E-155
	<i>Pedobacter piscium</i> strain AJ2502	NR_025536.1	100	3,00E-155
	<i>Pedobacter cryoconitis</i> strain A37	NR_025534.1	100	3,00E-155
1068486	<i>Mucilaginibacter calamicampi</i> strain WR-R1Y	NR_118563.1	99	1,00E-153
	<i>Mucilaginibacter lappiensis</i> strain ANJLI2	NR_115720.1	99	1,00E-153
200536	<i>Massilia oculi</i> strain CCUG 43427A	NR_117180.1	99	3,00E-155
	<i>Herbaspirillum rhizosphaerae</i> strain UMS-37	NR_043621.1	99	3,00E-155
	<i>Massilia aurea</i> strain AP13	NR_042502.1	99	3,00E-155
104023	<i>Janthinobacterium lividum</i> strain DSM 1522	NR_026365.1	100	2,00E-158
972341	<i>Massilia brevitalea</i> strain byr23-80	NR_044274.1	99	7,00E-157

Discussion

The alpha-rarefaction shows that while the full diversity (**Figure 2B**) was captured relatively fast, the sequencing depth was not able to cover the full phylotype richness (**Figure 2A**). Point Lobos samples contained fewer sequences (**Table 1 and 2**) but had a higher richness and diversity (**Figure 2**) than Cambria reserve and Swanton ranch samples. The lower abundance of the otherwise dominating 972341 phylotype could allow other phylotypes to surface and be sequenced to a fuller extent and therefore increase the species abundance.

No single phylotype was found to dominate all trees like the phylotype found in all studied samples of *P. flexilis* [22]. In all three mainland stands of Monterey pine there are individuals that contain a high relative abundance of phylotype 104023. This phylotype was 100% identical to *J. lividum*, a species that has previously been isolated from amphibian skin and mycorrhizal fungi. The strain isolated from amphibian skin prevents infection of chytrid fungus that threatens amphibian populations worldwide and has led to the extinction of many species [65]. It has been found to produce antifungal metabolites (e.g. violacenin) [65] and extracellular chitinases [66]. Chitinases breaks down chitin, an essential part of cell walls in fungi and some arthropods [68]. It is possible that the dominating phylotype 104023 has a roll in Monterey pine native trees' defence against fungi and/or insects.

Two of the dominating phylotypes (200536 and 972341) were both identified as belonging to the *Massilia* genus (**Table 3**). Bacteria in the *Massilia* genus have been found as one of the dominating endophytic genera in several agricultural crops throughout California [21], as well as in aerosols from close proximity to dairy farms in California [69]. Many of the *Massilia* species may have a high tolerance of abiotic stress [64] and it is a possibility that the *Massilia* genus could thrive inside the native stands of Monterey pine despite the currently harsh conditions. *In vitro* grown *Massilia* bacterium have also shown to have growth promoting effects as well as some antifungal properties [64]. Thus, it is possible that the secret to disease resistance in native *P. radiata* lies in the foliar endophyte community.

The results showed that the bacterial endophytic community in *P. radiata* sampled at Point Lobos differentiates from both Swanton Ranch and Cambria Reserve samples. Point Lobos samples had higher species richness, species diversity (**Figure 2**), and clustered together in PCoA-plots (**Figure 3**), the community composition diverges and the Point Lobos samples have a lower abundance of the overall dominating phylotype 972341 than the other two locations. Phylotype 972341 has been identified as belonging to the *Massilia* genus and the *Massilia* genus abundance have been show to fluctuate with soil composition [64]. The three sites' soil compositions differs some. Año Nuevo/Swanton ranch and Cambria primarily consists of smaller grains such as lime stone and sandstone, making it more clay-

like, the Monterey/Point Lobos soil also include granite [24] making it drier. This could be an explanation for the lower abundance of phylotype 972341 in the Point Lobos samples. Soil difference could also explain the high abundance of phylotype 200536 in Swanton ranch samples (**Table 3 and Figure 6**). The soil composition also likely affects soil moisture levels and could be one possible explanation for the deviating community. It is unlikely that Point Lobos distinctive community pattern is an effect of sampling and/or sequencing methods since the same method was used on all samples.

While genomic composition have been proven to affect the endophytic communities in many plants [35], [36], [38], [70] this does not seem to be the defining reason in the sampled Monterey pine. If genetic composition was the case, a greater separation would be seen in the weighted UniFrac distance PCoA-plot (**Figure 3A**). Swanton ranch and Cambria reserve samples are harder to differentiate from each other than Point Lobos, but in the unweighted PCoA-plot (**Figure 3B**) the samples separate. This indicates that the difference is due to transient factors and not the founding communities [61]. This is interesting since this difference then is likely not a result of the suggested genetic difference of the Cambria stands [24] but a more temporary factor such as nutrient availability.

To be able to significantly see any difference in endophytic communities between different tissue types, a more extensive sampling scheme would have been needed. When sampling for endophytic bacteria you can only sample healthy plants/trees to be able to guarantee that the microbes found are non-pathogenic endophytes, with the current drought situation in California[46] many trees did not have bud growth. This limited the study and with only three pooled bud samples no conclusive correlations based on tissue type can be done.

It is unlikely but possible that bias was created based on sample treatment (i.e. transportation, storage etc.), other conifer tissue that have been through the same treatment have not produced similar results [22]. Using 799f chloroplast-excluding primer likely removed other similar sequences such as Cyanobacteria and possible shifted the community structures. By using the primer-pair 799f-1115r created approximately 300 base pair long amplicons. Since the length of amplicons is positively correlated with species richness (i.e. the shorter amplicons the higher risk to underestimate species richness) it is a possibility that this study have underestimated the species richness.

In the future it would be interesting if one could optimize culture protocols to be able to grow the bacteria that dominate the Monterey pine endophyte community, and which are likely to play an important symbiotic role. Such isolates could be tested for antifungal properties. It would also be interesting to investigate how the endophytes colonize the native Monterey pine. If they are transmitted horizontally, could a “soil transplant” aid the threatened non-native stands?

Acknowledgement

I would like to thank Carolin Frank for welcoming me into her lab at UC Merced and introducing me to the world of endophytes. I would also like to thank Emily Wilson for allowing me to be a part of her collaboration with Megan Ruá at the University of Mississippi, as well as learning me the ropes around the lab and saving my sanity several times.

I would like to thank Uppsala University for five fantastic years, as well as my family and friends that have helped me thru these five amazing and sometimes challenging years.

References

- [1] O. E. Sala, F. S. Chapin, Iii, J. J. Armesto, E. Berlow, J. Bloomfield, R. Dirzo, E. Huber-Sanwald, L. F. Huenneke, R. B. Jackson, A. Kinzig, R. Leemans, D. M. Lodge, H. A. Mooney, M. Oesterheld, N. L. Poff, M. T. Sykes, B. H. Walker, M. Walker, and D. H. Wall, "Global Biodiversity Scenarios for the Year 2100," *Science*, vol. 287, no. 5459, pp. 1770–1774, Mar. 2000.
- [2] A. V. Sturz, B. R. Christie, and J. Nowak, "Bacterial Endophytes: Potential Role in Developing Sustainable Systems of Crop Production," *Critical Reviews in Plant Sciences*, vol. 19, no. 1, pp. 1–30, Jan. 2000.
- [3] B. Reinhold-Hurek and T. Hurek, "Living inside plants: bacterial endophytes," *Current Opinion in Plant Biology*, vol. 14, no. 4, pp. 435–443, Aug. 2011.
- [4] X. Liu, J. Jia, S. Atkinson, M. Cámara, K. Gao, H. Li, and J. Cao, "Biocontrol potential of an endophytic *Serratia* sp. G3 and its mode of action," *World J Microbiol Biotechnol*, vol. 26, no. 8, pp. 1465–1471, Aug. 2010.
- [5] M. Brownbridge, S. D. Reay, T. L. Nelson, and T. R. Glare, "Persistence of *Beauveria bassiana* (Ascomycota: Hypocreales) as an endophyte following inoculation of radiata pine seed and seedlings," *Biological Control*, vol. 61, no. 3, pp. 194–200, Jun. 2012.
- [6] B. B. Pageni, N. Z. Lupwayi, Z. Akter, F. J. Larney, L. M. Kawchuk, and Y. Gan, "Plant growth-promoting and phytopathogen-antagonistic properties of bacterial endophytes from potato (*solanum tuberosum* l.) cropping systems," *Can. J. Plant Sci.*, Mar. 2014.
- [7] J. Hallmann, A. Quadt-Hallmann, W. F. Mahaffee, and J. W. Kloepper, "Bacterial endophytes in agricultural crops," *Canadian Journal of Microbiology*, vol. 43, no. 10, pp. 895–914, 1997.
- [8] M. E. Rout, T. H. Chrzanowski, T. K. Westlie, T. H. DeLuca, R. M. Callaway, and W. E. Holben, "Bacterial endophytes enhance competition by invasive plants," *Am. J. Bot.*, vol. 100, no. 9, pp. 1726–1737, Sep. 2013.
- [9] S. E. Smith and D. J. Read, *Mycorrhizal Symbiosis*. Academic Press, 2010.
- [10] A. E. Arnold, L. C. Mejía, D. Kylo, E. I. Rojas, Z. Maynard, N. Robbins, and E. A. Herre, "Fungal endophytes limit pathogen damage in a tropical tree," *PNAS*, vol. 100, no. 26, pp. 15649–15654, Dec. 2003.
- [11] K. L. Higgins, P. D. Coley, T. A. Kursar, and A. E. Arnold, "Culturing and direct PCR suggest prevalent host generalism among diverse fungal endophytes of tropical forest grasses," *Mycologia*, vol. 103, no. 2, pp. 247–260, Mar. 2011.
- [12] A. M. Pirttilä and A. C. Frank, Eds., *Endophytes of forest trees: biology and applications*. Dordrecht ; New York: Springer, 2011.
- [13] USDA Natural Resources Conservation Service, "Plant Guide Monterey Pine." USDA NRCS National Plant Data Center, 29-May-2003.
- [14] T. I. Burgess, T. R. Gordon, M. J. Wingfield, and B. D. Wingfield, "Geographic isolation of *Diplodia scrobiculata* and its association with native *Pinus radiata*," *Mycological Research*, vol. 108, no. 12, pp. 1399–1406, 2004.
- [15] D. J. Mead, *Sustainable management of Pinus radiata plantations*. Rome: Food and agriculture organization of the United nations (FAO), 2013.
- [16] T. Woyke, H. Teeling, N. N. Ivanova, M. Huntemann, M. Richter, F. O. Gloeckner, D. Boffelli, I. J. Anderson, K. W. Barry, H. J. Shapiro, E. Szeto, N. C. Kyrpides, M. Mussmann, R. Amann, C. Bergin, C. Ruehland, E. M. Rubin, and N. Dubilier, "Symbiosis insights through metagenomic analysis of a microbial consortium," *Nature*, vol. 443, no. 7114, pp. 950–955, Oct. 2006.
- [17] F. D. Andreote, W. L. de Araújo, J. L. de Azevedo, J. D. van Elsas, U. N. da Rocha, and L. S. van Overbeek, "Endophytic Colonization of Potato (*Solanum tuberosum* L.) by a Novel Competent Bacterial Endophyte, *Pseudomonas putida* Strain P9, and Its Effect on Associated Bacterial Communities," *Appl. Environ. Microbiol.*, vol. 75, no. 11, pp. 3396–3406, Jun. 2009.

- [18] D. S. Lundberg, S. L. Lebeis, S. H. Paredes, S. Yourstone, J. Gehring, S. Malfatti, J. Tremblay, A. Engelbrektson, V. Kunin, T. G. del Rio, R. C. Edgar, T. Eickhorst, R. E. Ley, P. Hugenholtz, S. G. Tringe, and J. L. Dangl, "Defining the core *Arabidopsis thaliana* root microbiome," *Nature*, vol. 488, no. 7409, pp. 86–90, Aug. 2012.
- [19] K. H. Sharp, D. Distel, and V. J. Paul, "Diversity and dynamics of bacterial communities in early life stages of the Caribbean coral *Porites astreoides*," *ISME J*, vol. 6, no. 4, pp. 790–801, Apr. 2012.
- [20] J. Qin, R. Li, J. Raes, M. Arumugam, K. S. Burgdorf, C. Manichanh, T. Nielsen, N. Pons, F. Levenez, T. Yamada, D. R. Mende, J. Li, J. Xu, S. Li, D. Li, J. Cao, B. Wang, H. Liang, H. Zheng, Y. Xie, J. Tap, P. Lepage, M. Bertalan, J.-M. Batto, T. Hansen, D. Le Paslier, A. Linneberg, H. B. Nielsen, E. Pelletier, P. Renault, T. Sicheritz-Ponten, K. Turner, H. Zhu, C. Yu, S. Li, M. Jian, Y. Zhou, Y. Li, X. Zhang, S. Li, N. Qin, H. Yang, J. Wang, S. Brunak, J. Doré, F. Guarner, K. Kristiansen, O. Pedersen, J. Parkhill, J. Weissenbach, M. Antolin, F. Artiguenave, H. Blottiere, N. Borruel, T. Bruls, F. Casellas, C. Chervaux, A. Cultrone, C. Delorme, G. Denariáz, R. Dervyn, M. Forte, C. Friss, M. van de Guchte, E. Guedon, F. Haimet, A. Jamet, C. Juste, G. Kaci, M. Kleerebezem, J. Knol, M. Kristensen, S. Layec, K. L. Roux, M. Leclerc, E. Maguin, R. M. Minardi, R. Oozeer, M. Rescigno, N. Sanchez, S. Tims, T. Torrejon, E. Varela, W. de Vos, Y. Winogradsky, E. Zoetendal, P. Bork, S. D. Ehrlich, and J. Wang, "A human gut microbial gene catalogue established by metagenomic sequencing," *Nature*, vol. 464, no. 7285, pp. 59–65, Mar. 2010.
- [21] G. Rastogi, A. Sbodio, J. J. Tech, T. V. Suslow, G. L. Coaker, and J. H. J. Leveau, "Leaf microbiota in an agroecosystem: spatiotemporal variation in bacterial community composition on field-grown lettuce," *ISME J*, vol. 6, no. 10, pp. 1812–1822, Oct. 2012.
- [22] A. C. Frank and A. A. Carrell, "*Pinus flexilis* and *Picea engelmannii* share a simple and consistent needle endophyte microbiota with a potential role in nitrogen fixation," *Pinus flexilis*, vol. 5, p. 333, 2014.
- [23] A. Farjon, "IUCN Red List *Pinus Radiata*," *IUCN Red List of Threatened Species*. [Online]. Available: <http://www.iucnredlist.org/details/42408/0>. [Accessed: 31-Jul-2014].
- [24] P. M. McDonald and R. J. Laacke, "*Pinus Radiata*," in *Silvics of North America, Volume 1 Conifers*, vol. 1990, R. M. Burns and B. H. Honkala, Eds. Washington, D.C., U.S.A: Forest Service, United States Department of Agriculture.
- [25] "Signatures of adaptation and genetic structure among the mainland populations of *Pinus radiata* (D. Don) inferred from SNP loci - Springer."
- [26] J. C. Correll, T. R. Gordon, A. H. McCain, J. W. Fox, C. S. Koehler, D. L. Wood, and M. E. Schultz, "Pitch canker disease in California: pathogenicity, distribution, and canker development on Monterey pine (*Pinus radiata*)," *Plant Dis*, vol. 75, pp. 676–682, 1991.
- [27] A. J. Storer, P. Bonello, T. R. Gordon, and D. L. Wood, "Evidence of Resistance to the Pitch Canker Pathogen (*Fusarium circinatum*) in Native Stands of Monterey Pine (*Pinus radiata*)," *Forest Science*, vol. 45, no. 4, pp. 500–505, Nov. 1999.
- [28] T. R. Gordon, S. C. Kirkpatrick, B. J. Aegerter, A. J. Fisher, A. J. Storer, and D. L. Wood, "Evidence for the occurrence of induced resistance to pitch canker, caused by *Gibberella circinata* (anamorph *Fusarium circinatum*), in populations of *Pinus radiata*," *Forest Pathology*, vol. 41, no. 3, pp. 227–232, Jun. 2011.
- [29] D. Wilson, "Endophyte: The Evolution of a Term, and Clarification of Its Use and Definition," *Oikos*, vol. 73, no. 2, p. 274, Jun. 1995.
- [30] P. R. Hardoim, L. S. van Overbeek, and J. D. van Elsas, "Properties of bacterial endophytes and their proposed role in plant growth," *Trends in Microbiology*, vol. 16, no. 10, pp. 463–471, Oct. 2008.
- [31] H. Yu, L. Zhang, L. Li, C. Zheng, L. Guo, W. Li, P. Sun, and L. Qin, "Recent developments and future prospects of antimicrobial metabolites produced by endophytes," *Microbiological Research*, vol. 165, no. 6, pp. 437–449, Aug. 2010.

- [32] R. Anand, S. Grayston, and C. Chanway, "N₂-Fixation and Seedling Growth Promotion of Lodgepole Pine by Endophytic *Paenibacillus polymyxa*," *Microb Ecol*, vol. 66, no. 2, pp. 369–374, Aug. 2013.
- [33] J. L. Azevedo, W. Maccheroni Jr., J. O. Pereira, and W. L. De Araújo, "Endophytic microorganisms: a review on insect control and recent advances on tropical plants," *Electronic Journal of Biotechnology*, vol. 3, no. 1, Apr. 2000.
- [34] M. Rosenblueth and E. Martínez-Romero, "Bacterial Endophytes and Their Interactions with Hosts," *MPMI*, vol. 19, no. 8, pp. 827–837, Aug. 2006.
- [35] L. Van Overbeek and J. D. Van Elsas, "Effects of plant genotype and growth stage on the structure of bacterial communities associated with potato (*Solanum tuberosum* L.)," *FEMS Microbiology Ecology*, vol. 64, no. 2, pp. 283–296, May 2008.
- [36] J. Kuklinsky-Sobral, W. L. Araújo, R. Mendes, I. O. Geraldi, A. A. Pizzirani-Kleiner, and J. L. Azevedo, "Isolation and characterization of soybean-associated bacteria and their potential for plant growth promotion," *Environmental Microbiology*, vol. 6, no. 12, pp. 1244–1251, Dec. 2004.
- [37] B. Ma, X. Lv, A. Warren, and J. Gong, "Shifts in diversity and community structure of endophytic bacteria and archaea across root, stem and leaf tissues in the common reed, *Phragmites australis*, along a salinity gradient in a marine tidal wetland of northern China," *Antonie van Leeuwenhoek*, vol. 104, no. 5, pp. 759–768, Nov. 2013.
- [38] G. P. Cheplick and R. Cho, "Interactive effects of fungal endophyte infection and host genotype on growth and storage in *Lolium perenne*," *New Phytologist*, vol. 158, no. 1, pp. 183–191, 2003.
- [39] S. M. Huse, Y. Ye, Y. Zhou, and A. A. Fodor, "A Core Human Microbiome as Viewed through 16S rRNA Sequence Clusters," *PLoS ONE*, vol. 7, no. 6, p. e34242, Jun. 2012.
- [40] J. B. Hughes, J. J. Hellmann, T. H. Ricketts, and B. J. M. Bohannan, "Counting the Uncountable: Statistical Approaches to Estimating Microbial Diversity," *Appl Environ Microbiol*, vol. 67, no. 10, pp. 4399–4406, Oct. 2001.
- [41] P. Engel and N. A. Moran, "The gut microbiota of insects – diversity in structure and function," *FEMS Microbiol Rev*, vol. 37, no. 5, pp. 699–735, Sep. 2013.
- [42] M. K. Chelius and E. W. Triplett, "The Diversity of Archaea and Bacteria in Association with the Roots of *Zea mays* L.," *Microbial Ecology*, vol. 41, no. 3, pp. 252–263, Apr. 2001.
- [43] N. Fierer, C. M. McCain, P. Meir, M. Zimmermann, J. M. Rapp, M. R. Silman, and R. Knight, "Microbes do not follow the elevational diversity patterns of plants and animals," *Ecology*, vol. 92, no. 4, pp. 797–804, Nov. 2010.
- [44] I. M. Head, J. R. Saunders, and R. W. Pickup, "Microbial Evolution, Diversity, and Ecology: A Decade of Ribosomal RNA Analysis of Uncultivated Microorganisms," *Microbial Ecology*, vol. 35, no. 1, pp. 1–21, Jan. 1998.
- [45] H. Tuomisto, "A consistent terminology for quantifying species diversity? Yes, it does exist," *Oecologia*, vol. 164, no. 4, pp. 853–860, Dec. 2010.
- [46] "Office of Governor Edmund G. Brown Jr. - Newsroom," 17-Jan-2014. [Online]. Available: <http://gov.ca.gov/news.php?id=18368>. [Accessed: 01-Aug-2014].
- [47] A. S. Hanshew, C. J. Mason, K. F. Raffa, and C. R. Currie, "Minimization of chloroplast contamination in 16S rRNA gene pyrosequencing of insect herbivore bacterial communities," *Journal of Microbiological Methods*, vol. 95, no. 2, pp. 149–155, Nov. 2013.
- [48] J.-Y. Jiao, H.-X. Wang, Y. Zeng, and Y.-M. Shen, "Enrichment for microbes living in association with plant tissues," *Journal of Applied Microbiology*, vol. 100, no. 4, pp. 830–837, Apr. 2006.
- [49] A. J. Redford, R. M. Bowers, R. Knight, Y. Linhart, and N. Fierer, "The ecology of the phyllosphere: geographic and phylogenetic variability in the distribution of bacteria on tree leaves," *Environmental Microbiology*, vol. 12, no. 11, pp. 2885–2893, Nov. 2010.

- [50] D. Berry, K. B. Mahfoudh, M. Wagner, and A. Loy, "Barcoded Primers Used in Multiplex Amplicon Pyrosequencing Bias Amplification," *Appl. Environ. Microbiol.*, vol. 77, no. 21, pp. 7846–7849, Nov. 2011.
- [51] T. F. Scientific, "NanoDrop 1000 Spectrophotometer V3. 7 User's Manual," *Thermo Fisher Scientific*, pp. 1–105, 2008.
- [52] J. G. Caporaso, J. Kuczynski, J. Stombaugh, K. Bittinger, F. D. Bushman, E. K. Costello, N. Fierer, A. G. Peña, J. K. Goodrich, J. I. Gordon, G. A. Huttley, S. T. Kelley, D. Knights, J. E. Koenig, R. E. Ley, C. A. Lozupone, D. McDonald, B. D. Muegge, M. Pirrung, J. Reeder, J. R. Sevinsky, P. J. Turnbaugh, W. A. Walters, J. Widmann, T. Yatsunenko, J. Zaneveld, and R. Knight, "QIIME allows analysis of high-throughput community sequencing data," *Nature Methods*, vol. 7, no. 5, pp. 335–336, May 2010.
- [53] E. Aronesty, *ea-utils: Command-line tools for processing biological sequencing data*. 2011.
- [54] T. Z. DeSantis, P. Hugenholtz, N. Larsen, M. Rojas, E. L. Brodie, K. Keller, T. Huber, D. Dalevi, P. Hu, and G. L. Andersen, "Greengenes, a Chimera-Checked 16S rRNA Gene Database and Workbench Compatible with ARB," *Appl. Environ. Microbiol.*, vol. 72, no. 7, pp. 5069–5072, Jul. 2006.
- [55] R. C. Edgar, "Search and clustering orders of magnitude faster than BLAST," *Bioinformatics*, vol. 26, no. 19, pp. 2460–2461, Oct. 2010.
- [56] J. G. Caporaso, K. Bittinger, F. D. Bushman, T. Z. DeSantis, G. L. Andersen, and R. Knight, "PyNAST: a flexible tool for aligning sequences to a template alignment," *Bioinformatics*, vol. 26, no. 2, pp. 266–267, Jan. 2010.
- [57] Q. Wang, G. M. Garrity, J. M. Tiedje, and J. R. Cole, "Naïve Bayesian Classifier for Rapid Assignment of rRNA Sequences into the New Bacterial Taxonomy," *Appl Environ Microbiol*, vol. 73, no. 16, pp. 5261–5267, Aug. 2007.
- [58] B. J. Haas, D. Gevers, A. M. Earl, M. Feldgarden, D. V. Ward, G. Giannoukos, D. Ciulla, D. Tabbaa, S. K. Highlander, E. Sodergren, B. Methé, T. Z. DeSantis, J. F. Petrosino, R. Knight, and B. W. Birren, "Chimeric 16S rRNA sequence formation and detection in Sanger and 454-pyrosequenced PCR amplicons," *Genome Res.*, vol. 21, no. 3, pp. 494–504, Mar. 2011.
- [59] M. N. Price, P. S. Dehal, and A. P. Arkin, "FastTree 2 – Approximately Maximum-Likelihood Trees for Large Alignments," *PLoS ONE*, vol. 5, no. 3, p. e9490, Mar. 2010.
- [60] C. Lozupone and R. Knight, "UniFrac: a new phylogenetic method for comparing microbial communities," *Appl. Environ. Microbiol.*, vol. 71, no. 12, pp. 8228–8235, Dec. 2005.
- [61] C. A. Lozupone, M. Hamady, S. T. Kelley, and R. Knight, "Quantitative and Qualitative β Diversity Measures Lead to Different Insights into Factors That Structure Microbial Communities," *Appl Environ Microbiol*, vol. 73, no. 5, pp. 1576–1585, Mar. 2007.
- [62] R. Studio, *RStudio: integrated development environment for R*. Version 0.97, 2012.
- [63] G. R. Warnes, B. Bolker, L. Bonebakker, R. Gentleman, A. Liaw, W. Huber, T. Lumley, M. Maechler, A. Magnussin, S. Moeller, M. Schwartz, and B. Venables, *gplots: Various R programming tools for plotting data*. 2014.
- [64] M. Ofek, Y. Hadar, and D. Minz, "Ecology of Root Colonizing Massilia (Oxalobacteraceae)," *PLoS ONE*, vol. 7, no. 7, p. e40117, Jul. 2012.
- [65] R. M. Brucker, R. N. Harris, C. R. Schwantes, T. N. Gallaher, D. C. Flaherty, B. A. Lam, and K. P. C. Minbiole, "Amphibian Chemical Defense: Antifungal Metabolites of the Microsymbiont *Janthinobacterium lividum* on the Salamander *Plethodon cinereus*," *J Chem Ecol*, vol. 34, no. 11, pp. 1422–1429, Nov. 2008.
- [66] T. R. Scheublin, I. R. Sanders, C. Keel, and J. R. van der Meer, "Characterisation of microbial communities colonising the hyphal surfaces of arbuscular mycorrhizal fungi," *ISME J*, vol. 4, no. 6, pp. 752–763, Jun. 2010.
- [67] F. J. de Bruijn, *Molecular Microbial Ecology of the Rhizosphere, Two Volume Set*. John Wiley & Sons, 2013.

- [68] A. P. Gleave, R. K. Taylor, B. A. M. Morris, and D. R. Greenwood, "Cloning and sequencing of a gene encoding the 69-kDa extracellular chitinase of *Janthinobacterium lividum*," *FEMS Microbiology Letters*, vol. 131, no. 3, pp. 279–288, Sep. 1995.
- [69] S. V. Ravva, C. Z. Sarreal, and R. E. Mandrell, "Bacterial Communities in Aerosols and Manure Samples from Two Different Dairies in Central and Sonoma Valleys of California," *PLoS One*, vol. 6, no. 2, Feb. 2011.
- [70] D. Todd, "The effects of host genotype, growth rate, and needle age on the distribution of a mutualistic, endophytic fungus in Douglas-fir plantations," *Can. J. For. Res.*, vol. 18, no. 5, pp. 601–605, May 1988.