Master's thesis

NTNU
Norwegian University of Science and Technology
Faculty of Natural Sciences
Department of Biology

Abiola Opeyemi Aremu

Herbarium metagenomics to determine the effects of preservation methods on potato plant microbial communities

Master's thesis in Biology
Supervisor: Michael D. Martin
Co-supervisor: Vanessa C. Bieker, Jaelle C. Brealey
May 2023

NTNU
Norwegian University of
Science and Technology

Abiola Opeyemi Aremu

Herbarium metagenomics to determine the effects of preservation methods on potato plant microbial communities

Master's thesis in Biology
Supervisor: Michael D. Martin
Co-supervisor: Vanessa C. Bieker, Jaelle C. Brealey
May 2023

Norwegian University of Science and Technology
Faculty of Natural Sciences
Department of Biology

NTNU
Norwegian University of
Science and Technology

ABSTRACT

Herbarium specimens are preserved plants kept for varying purposes, for example, botanical records, hobbies, referencing, research, economic and medicinal purposes. There are more than 3000 herbaria in the world that house approximately 350 million specimens from the last ca. 400 years. Due to advances in DNA extraction and next-generation sequencing (NGS) approaches, these specimens are now used for genetic studies and can be used to track genetic changes over time. Herbarium specimens do not only contain DNA from the plant itself, but also from associated microbes and pathogens that were present on the living plant and can thus be used to study the microbial community. However, due to different procedures they undergo during the process of preservation and storage of the specimens, herbarium specimens are also prone to microbial contamination.

In this project, plant samples were preserved with different methods including standard procedures for herbarium specimens preservation. Samples were taken at different steps of preservation in order to track changes in microbial abundance and identify possible contaminating microbial taxa. When these microbes are identified, herbarium specimens could be useful in carrying out metagenomic studies in the future. Two main potato *(Solanum tuberosum)* plant tissues (leaves and roots) were used for this project. Samples were collected from different individuals from two different farms in Trondheim. Samples were preserved at the Trondheim herbarium (TRH). Identification of the first 20 most abundant microbial taxa were done using a simple bar plot. Principal Coordinate Analysis (PCoA) was used to evaluate the similarities and variation between samples collected from different sites and preserved under varying conditions.

MaAslin 2.0 was used to identify whether changes in taxa abundance were significantly different between different preservation methods in the same plant tissue. Two genera were the most prominent in the plant tissues studied. No taxa abundance comparison is statistically significant in leaf samples while few microbial taxa have higher abundance in root samples upon comparisons. *Alternaria alternata* (common herbarium contaminant fungus) was not identified in any of the herbarium stored samples. We hypothesized that the absence of common herbarium contaminants could be due to high standards of modern herbarium practice at Trondheim Herbarium (TRH). Proper identification and removal of contaminants from herbarium specimens can help in making predictions about evolutionary trends of metagenomic communities on potato plants.

ACKNOWLEDGEMENTS

Firstly, I would like to say a big thank you to my supervisors Mike Martin, Vanessa Bieker and Jaelle Brealey. You provided all the support, advice and encouragement needed for me to pull through with this project. Thanks to Vanessa and Jaelle for the bi-weekly meetings in the last one year. They were quite helpful and encouraging. I greatly appreciate your persistence.

This project would not have been successful without the efforts of the farmers at Voll farm and Sverresborg farm. Thanks for the opportunity granted to collect our samples from your farm. To the collections manager of Trondheim Herbarium (TRH), Tommy Prestø, thanks for teaching me the standard procedures for preserving plant specimens. I also want to appreciate Nellie Henriette Nilsen for taking me through the process of mounting the specimens on the mounting sheets. A big thanks to Novogene UK for building the libraries and carrying out the next-generation sequencing.

In addition, I would like to thank the Almighty God for keeping me from the beginning to the end of this project. For impacting me with wisdom and strength for completion of this project. A big thanks to my mother (Mojirola), brothers (Olabode and Olaniyi) for encouraging me to always be a better version of myself every day. Thanks to all my friends from the Museum, Norway and abroad. I want to appreciate the Department of Biology, Institute of Natural history for their contribution in various ways.

Table of Contents

INTRODUCTION

LEAF AND ROOT MICROBIOMES

Plants harbour a diverse assembly of microbial communities within and outside their tissues (Brader et al., 2017; Vandenkoornhuyse et al., 2015). These microbes can either be beneficial or harmful to their host plants (Gupta et al., 2017). Endophytes are plant microbes such as fungi and bacteria that cause no harm to host plants and live inside plant tissues (Fadiji & Babalola, 2020). They play important roles in plant growth and health improvement by suppressing plant pathogens (Turner et al., 2013). They are distinguished from epiphytes, which live on plant surfaces. Epiphytes are capable of making their own food, without relying on host plants (Lüttge & Scarano, 2004; Zotz & Hietz, 2001). A third class of microbes are found living in the surrounding environment of the plants, such as soil microbes in the rhizosphere. They interact with plants without touching them in a variety of ways, such as facilitating gas exchange between plants and soil and assisting plants in defense against harmful pathogens (Hardoim et al., 2008).

With regards to the beneficiary effects of microorganisms, some microbes help in fixing nitrogen in root nodules of plants (Etesami, 2022). Nitrogen is essential for growth and development in plants, as it serves as the building block of proteins (Sharma et al., 2003). In human health and agriculture, plant microbes have a number of uses. Endophytic microbes such as bacteria, fungi and actinomycetes form associations with medicinal plants. These microbes secrete hormones that are capable of fighting against human and animal diseases. Hence, medicinal plants harbour these microbes which have antibiotic, antimalarial, anticancer and antifungal properties (Akerele et al., 1991), which are helpful in curing diseases. Microorganisms have the potential for being used as biofertilizers in agricultural practices, e.g. as an alternative to chemical fertilizers that pose several risks (Fadiji & Babalola, 2020). In addition, plant microbes can be good biological control agents against plant pathogens in farming (Turner et al., 2013). A review was carried out by (Wei et al., 2017) on the potential of using leaf and associated microbial communities to reduce the toxicity level of different air pollutants to the environment. It was revealed that plant leaves absorb air pollutants, and leaf endophytes are responsible for the biodegradation and modification of the pollutants into less toxic molecules. Therefore, leaf-based remediation technologies could be developed to mitigate the effect of air pollution via leaves of plants. Interestingly, the importance and benefits of endophytes have not been fully investigated.

On the other hand, harmful microbes (pathogens) can be the root of diseases or infections in their host organism (Kenneth Horst, 2013). Some pathogens cause diseases in plants by releasing harmful substances into the cell of the host plants, thereby suppressing their immunity and making it difficult for the host to fight against pathogens. Also, pathogens alter the rate of metabolism and physiology of their host plant which have impacts on the growth and toughness of plants (Burdon, 1987). Because pathogenic organisms would have utilized resources away, affected plants experience a reduction in the resources needed for their development and reproduction (Berger et al., 2007; Denancé et al., 2013; van Dijk et al., 2021), which may lead to crop yield loss. It has been reported that about 16% of global yield losses are caused by plant pathogens (Ficke et al., 2018). However, the use of fungicides and pesticides to fight against pathogens may alter natural microbial communities, including beneficial microbes of plants (Abdelfattah et al., 2016; Berlec, 2012). Therefore, the use of old herbarium specimens for metagenomic studies are useful in evaluating changes that have occurred over time in microbial communities before the introduction of chemicals and fertilizers (Délye et al., 2013).

The great diversity of microorganisms poses the problem of under exploration of their taxonomy, composition, structure and molecular mechanisms (Cowan et al., 2005). The classical approach for studying novel microbes involves culturing them in a liquid or solid growth medium. However, several microbes are not cultivable in the laboratory, which has made the classical approach inappropriate for the growth of many microbial taxa (Staley & Konopka, 1985). With the advent of recent technology, host organisms can be collected from their natural environment and be studied with all associated microorganisms. The genomic analysis of microorganisms through the process of extraction of DNA from their natural environment with no requirement to obtain pure cultures of the microorganisms is known as metagenomics (J. Singh et al., 2009). It allows nearly complete characterization of DNA of almost the entire microbial community within, on, and surrounding host organisms, many of which might be impossible to culture (Langille et al., 2013).

HERBARIUM SPECIMENS

Herbarium specimens are collections of pressed, dried, mounted and labelled plant materials stored in herbaria. They can consist of any plant parts (e.g. seeds, pollen, wood sections, flowers, stems, branches, roots), dried and pressed for future reference (Willis et al., 2017). Collections of various kinds of herbarium specimens originated in Europe between the 15th and 16th centuries (Lane,

1996). Today, there are about 3400 herbaria all around the world (Besnard, 2018). Approximately 350 million herbarium specimens have been collected globally during the last ~400 years (Soltis, 2017). For hundreds of years, naturalists and botanists have collected and preserved plants for solving the problems of taxonomy, economic uses, documenting regional floras, studying plant diversification, and simply as a hobby (Balick, 1996; Palmer et al., 1995; Sprague & Nelmes, 1931).

Photographs and digital herbaria are also used in preserving specimens as digitized databases. Many herbaria are digitalizing their collections, which make them more accessible for different kinds of research. For example, specimens no longer need to be physically posted between different herbaria to examine certain morphological characters. Moreover, researchers can easily search the database to see available information for project planning (Lane, 1996; Nelson et al., 2015).

Herbarium specimens contain huge amounts of information about morphology, phenology and health of plants of the past. This information, combined with the metadata found on the specimen labels (e.g. collection date, location, habitat) can be useful in studying the ecological and evolutionary processes in those plants. For example, the effects of changing climate on cyclic response of plants can be known using herbarium specimens (Calinger et al., 2013; Primack et al., 2004). From a phylogenetic perspective, historical herbarium collections can be used for studying threatened and extinct species. For instance, genome skimming information from a single specimen of the now-extinct monotypic Oleaceae genus *Hesperelaea* collected 140 years ago, was used to determine the phylogenetic rank of the genus (Zedane et al., 2015).

Also, herbarium specimens are quite useful for research purposes when faced with studying extinct or endangered populations, or plant taxa that cannot be reached under challenging field conditions, or when there are inadequate funds to make new collections (Przelomska et al., 2020). Herbarium specimens can also be useful in inferring the effects of pesticides on plants over a given period within a given geographical range (Lang et al., 2019). The use of chemicals such as pesticides, herbicides or fungicides on plants can be traced back to the twentieth century, and these chemicals have significant effects on the plants and their microbial communities (Hahn, 2014; Vats, 2015).

Although herbarium specimens are good sources of genetic information, the quality and quantity of their genetic information may be affected by the particular methods of specimen preservation (Adams & Sharma, 2010; Pyle & Adams, 1989; Staats et al., 2011). However, the effects of different preservation methods have yet to be studied systematically. Typically, it is unclear which kinds of treatments were used for preserving some historic plant samples. Some

historic samples were preserved with toxic substances such as formalin (Yadav, 2020) which affect the DNA quality and quantity of herbarium specimens (Särkinen et al., 2012). Modern collections are often preserved with milder methods, such as the use of warm air and wooden plant presses. However, tropical collections are dried at high temperature (fast drying method) (Yadav, 2020).

A review carried out by (Bieker & Martin, 2018) on the indication and future prospects of DNA in historical samples for evolutionary analyses showed that the quality of DNA in a specimen is dependent on the age of the specimen. Genetic materials in herbarium specimens decompose six times faster than in ancient bones (Weiß et al., 2016). Therefore, genetic material derived from herbarium samples are often fragmented, and DNA concentration and endogenous contents are low (Adams and Sharma 2010; Staats et al., 2011). These findings have called for gently following all necessary measures when handling herbarium specimens that are used as truly ancient samples to prevent contamination with modern DNA (Shepherd & Perrie, 2014).

Most of the time, it is challenging to disentangle when microbes colonise the plants. Microbial DNA on a historic herbarium sample could have been there when the plant specimen was alive (*antemortem*), or arrived around the time of death of the specimen (*perimortem*), when the dead plant was being dried and mounted on sheets in the herbarium, or during storage in the herbarium (*postmortem*). Furthermore, reagent contamination in the laboratory and DNA extraction kits used can also introduce contaminants into metagenomic communities of interest (Salter et al., 2014). Herbarium specimens can be further contaminated with microbes during preservation, storage, and specimen handling. For example, a fungus called *Alternaria alternata*, was found to be present only in herbarium specimens and is thus likely a herbarium contamination (Bieker et al., 2020).

Recently, the advent of culture-independent, high-throughput next-generation sequencing, and improvements in DNA extraction methods have made it feasible to include historical herbarium specimens in genetic studies (Pääbo et al., 2004; Yoshida et al., 2014). These advances make it possible to not only study the genetic material from the plant specimens themselves but also from their associated microbes (Ames & Spooner, 2008; Bieker et al., 2020; Ristaino, 1998). The potential of using herbarium specimens for studying microbial communities was examined in a study by (Bieker et al., 2020). They discovered that certain microbes were identified as potential herbarium contaminants. It was concluded that herbarium specimens can be used for metagenomic analyses, but they must be treated with extra care to be sure that identified taxa are representative of the host plants' natural communities, rather than communities that developed after the plant specimen was collected for herbarium storage.

In this project, I used potato (*Solanum tuberosum* L.) as a model to study the microbial communities associated with plant tissues during important stages of herbarium specimen preservation. Potato (*S. tuberosum*) is one of the world's most important staple foods (Mu & Sun, 2017; Salmensuu, 2021) and plays a key role in global food security (Wijesinha-Bettoni & Mouillé, 2019). Its vulnerability to pests and diseases have contributed to reduction in potato production. One of such pathogens is *Phytophthora infestans*, an oomycete that causes late blight disease in potato (Haas et al., 2009). It originated from Central Mexico (Austin Bourke, 1964; Martin et al., 2014) and, upon introduction to Europe, caused great losses to potato production e.g., it caused the Irish potato famine of the 1840s (Austin Bourke, 1964; Savary et al., 2017). The first historic herbarium specimens ever examined in a metagenomic investigation were potato plants infected by *P. infestans* (Schubert et al., 2014).

STUDY AIMS AND HYPOTHESIS

The aim of this study is to identify microbial contamination during various preservation procedures and methods, including standard herbarium procedures. The motivation for this work is to facilitate the future use of herbarium specimens in metagenomic studies of *antemortem* microbes by allowing the identification and removal of *perimortem* and *postmortem* microbes from consideration. For this study, potato plants were collected from two different locations, and these samples were preserved in different ways. Firstly, I test the prediction that varying the methods of preserving different plant tissues (leaf and root) within and outside of the herbarium would result in the identification of different microbial taxa. Also, I predicted that geographical proximity and microbial community similarity would be correlated. I predicted that there would be significant changes in microbial taxa abundance when comparing two different methods of preservation within the same tissue type. Lastly, I predicted that microbial species richness in the root tissue would be higher than in the leaf tissue.

To achieve this, I (1) identified the top 20 taxa present in leaf and root tissues considering the varying methods of preservation; (2) compared all microbial taxa according to different preservation methods in each tissue type; (3) analyzed the statistically significant changes in microbial taxa abundance between two preservation methods within the same plant tissue to determine whether some taxa are differently represented under a particular preservation method; and (4) compared species richness in leaf and root tissues to identify which tissues have higher taxa richness.

MATERIALS AND METHODS

SAMPLE COLLECTION AND METHODS OF PRESERVATION

Whole potato plants (*Solanum tuberosum*) were collected from two different farms in Trondheim: Vollgård farm in Moholt (Latitude 63.411°, Longitude 10.449°) and Sverresborg folkmuseum (Latitude 63.419°, Longitude 10.362°). The individual potato plant collected from Moholt farm was labelled as M1, and the individuals from Sverresborg were named S1, S2, S3, and S4. These individual plants or samples thereof were preserved under varying methods of preservation that are described in detail later (Table 1). Root and leaf samples were collected for each preservation step. For each plant, fresh leaf and root samples were taken, directly frozen and stored at -20 ˚C until DNA extraction (preservation method, P6). Another set of fresh leaf and root samples was dried on silica-gel until DNA extraction (preservation method, P1). The remaining plant material was dried using a plant press (Fig. 2a). For that, the plant was put into newspaper and cardboard sheets were placed between plants. The cardboard was changed at an interval of 3 days (for preservation methods P2 and P4). For one individual (M1), the plant was split, and one part was placed in another plant press where the cardboard was left throughout the drying processes without changing (for preservation methods P3 and P5). After drying in the plant press, specimens were placed in a -20˚C freezer for one week before they were brought to the herbarium and mounted on herbarium sheets, a standard practice to destroy any insects that may exist on the plants (Gilberg & Brokerhof, 1991). Another set of samples were taken directly after mounting (for preservation methods P4 and P5). The mounted specimens were then placed in the herbarium for about 4 months before the last set of samples were taken (for preservation methods P2 and P3). No root samples were collected for P3 and P5 preservation methods.

Table 1. Provenance and treatment of study samples. Collection sites were represented as 1 and 2. All 1 are individuals from Voll fam and 2 are individuals from Sverresborg farm

Sample ID	Plant ID	Tissue type	Collection sites	Preservation method	Preservation code
M1-1	M1	root	1	Freshly frozen and stored at -20 ˚C	P6
M1-2		leaf	1	Freshly frozen and stored at -20 ˚C	P6
M1-3		leaf	1	Silica dried	P1
M1-4		root	1	Silica dried	P1
M1-5		leaf	1	Directly after storing in herbarium (cardboard changed)	P4
M1-6		root	1	Directly after storing in herbarium (cardboard changed)	P4
M1-7		leaf	1	Directly after storing in herbarium (cardboard not changed)	P5
M1-8		leaf	1	After storing in herbarium (cardboard changed)	P2
M1-9		root	1	After storing in herbarium (cardboard changed)	P2
M1-10		leaf	1	After storing in herbarium (cardboard not changed)	P3
S1-1	S1	leaf	2	Freshly frozen and stored at -20 ˚C	P6
S1-2		leaf	2	Silica dried	P1
S1-3		root	2	Freshly frozen and stored at -20 ˚C	P6
S1-4		root	2	Silica dried	P1

S1-5		leaf	2	Directly after drying in the press (cardboard changed)	P4
S1-6		root	2	Directly after drying in the press (cardboard changed)	P4
S1-7		leaf	2	After storing in herbarium (cardboard changed)	P2
S1-8		leaf	2	After storing in herbarium (cardboard changed)	P2
S2-1	S2	leaf	2	Freshly frozen and stored at -20 ˚C	P6
S2-2		leaf	2	Silica dried	P2
S2-3		root	2	Freshly frozen and stored at -20 ˚C	P6
S2-4		root	2	Silica dried	P1
S2-5		leaf	2	Directly after drying in the press (cardboard changed)	P4
S2-6		root	2	Directly after drying in the press (cardboard changed)	P4
S2-7		root	2	After storing in herbarium (cardboard changed)	P2
S2-8		root	2	After storing in herbarium (cardboard changed)	P2
S3-1	S3	leaf	2	Freshly frozen and stored at -20 ˚C	P6
S3-2		leaf	2	Silica dried	P1
S3-3		root	2	Freshly frozen and stored at -20 ˚C	P6
S3-4		leaf	2	Silica dried	P1
S3-5		leaf	2	Directly after drying in plant press (cardboard changed)	P4
S3-6		root	2	Directly after drying in plant press (cardboard changed)	P4

S3-7		leaf	2	After storing in herbarium (cardboard changed)	P2
S3-8		root	2	After storing in herbarium (cardboard changed)	P2
S4-1	S4	leaf	2	Freshly frozen and stored at -20 ˚C	P6
S4-2		leaf	2	Silica dried	P1
S4-3		root	2	Freshly frozen and stored at -20 ˚C	P6
S4-4		root	2	Silica dried	P1
S4-5		leaf	2	Directly after drying in plant press (cardboard changed)	P4
S4-6		root	2	Directly after drying in plant press (cardboard changed)	P4
S4-7		leaf	2	After storing in herbarium (cardboard changed)	P2
S4-8		root	2	After storing in herbarium (cardboard changed)	P2

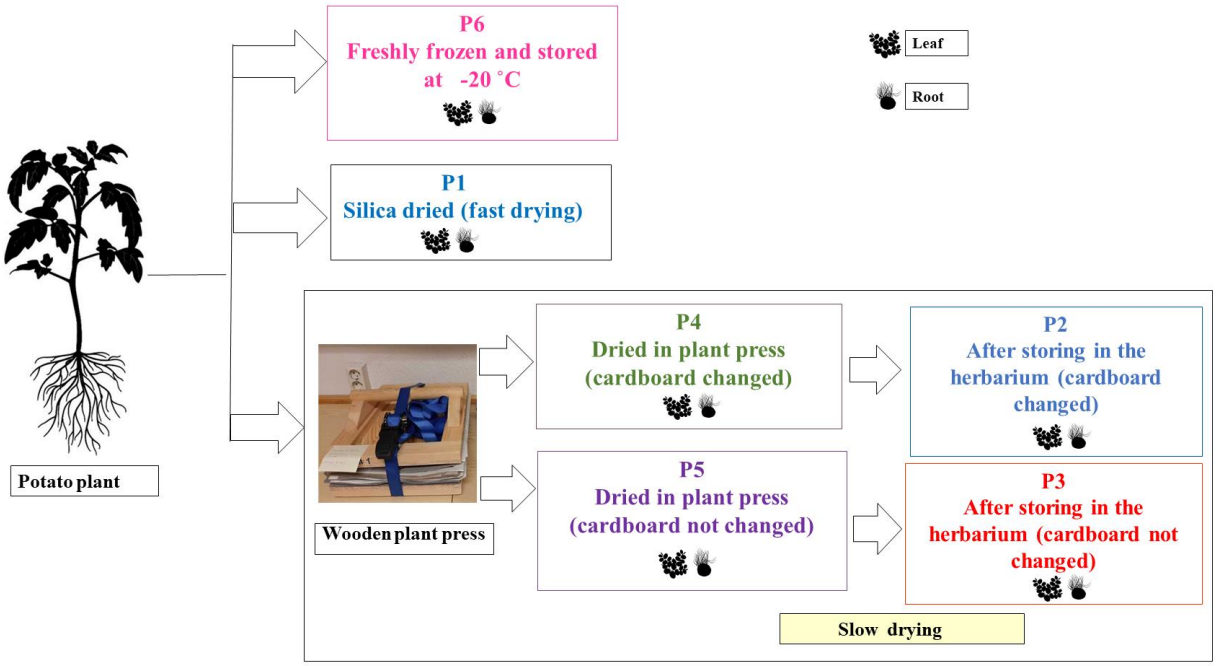


Figure 1. Schematic of the different preservation methods used in this study.

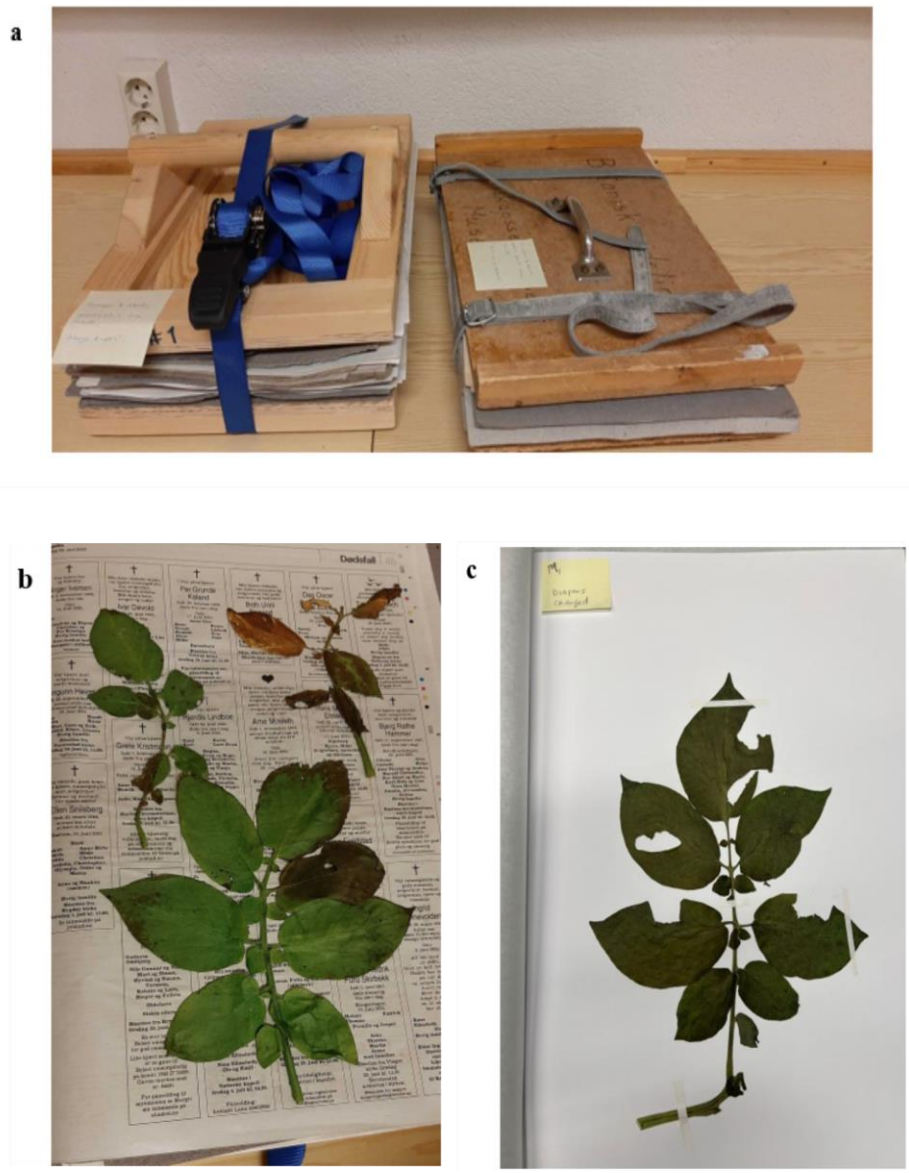


Figure 2. Images to illustrate the herbarium preservation process. a) Two wooden plant presses. b) Drying plant. c) Mounted specimen.

DNA EXTRACTION, LIBRARY PREPARATION AND SEQUENCING

DNA was extracted in the NTNU Museum molecular laboratories using a DNeasy Minikit (QIAGEN) following the manufacturer's instructions. DNA was extracted from 45 samples, including leaf and root tissues and negative control blanks, which were included throughout DNA extractions and library build procedures to monitor background reagent and laboratory microbial contamination. DNA was quantified using a Qubit fluorometer (Invitrogen) with the dsDNA BR kit. Novogene UK (United Kingdom), a commercial genomics service provider, sheared the DNA, built and quality-controlled dsDNA libraries, and conducted next-generation sequencing on the Illumina platform.

ALIGNMENT OF SEQUENCED READS TO REFERENCE GENOME

The bioinformatic software paleomix v.1.2.13.8 (Schubert et al., 2014) was used to map the data against the potato reference genome SolTub_3.0 (GCF_000226075.1). Firstly, AdapterRemoval v.2.3.1 (Schubert et al., 2016) was used for trimming adapters and paired reads with an overlap of at least 11 base pairs were collapsed into one read and are treated as single-end reads during the mapping. The trimmed reads were then mapped against the potato reference genome using bwa mem v.0.7.17 (Winter et al., 2023). The reads that mapped to the reference genome were filtered out using samtools v.1.16.1 bam2fq (Wood & Salzberg, 2014) with the *-f 4* flag, leaving the unmapped reads to be used for metagenomic classification. The bam files containing unmapped reads were converted into FASTQ files for downstream analysis. Polymerase chain reaction (PCR) duplicates were removed from the unmapped reads using dedupe.sh from the bbmap toolkit (sourceforge.net/projects/bbmap/), leaving only the unique sequences for further analysis. Also, raw reads were mapped against the *P. infestans* nuclear genome (PRJNA17665) (Haas et al., 2009) to determine the relative abundance of *P. infestans* in all samples.

CHARACTERIZATION OF MICROBIAL COMMUNITIES

To assign taxonomic names (labels) to metagenomic DNA sequences, the software Kraken v2.0.9 (Wood & Salzberg, 2014) was used. Kraken classifies more accurately and works quickly in comparison to BLAST, which can only classify subsets of metagenomic data (Wood and Salzberg, 2014). Kraken was run on the paired reads and collapsed reads separately. A database of k-mers used to classify the taxonomy of sequenced DNA sequences is called the Kraken database (Wood & Salzberg, 2014). The reference genomes of thousands of species are used to obtain it, and it is

regularly updated with additional genomes (Tessler et al., 2017). It has been extensively used to rank microbiomes and metagenomic materials in many genomic investigations. A minimum base quality of 20 for the kraken classification step was used. A database, ncbi_ntDB built in October 2022 was used. The resulting classification files were afterwards combined using KrakenTools (Lu et al., 2022). Intermediate ranks (such as subphyla, subclasses, subspecies, etc.) were removed for all classified organisms. Counts for reads classified at the genus and species level were extracted for this study. All animals (Metazoans), plants (Viridiplantae) and viruses (Viruses) were filtered out because we are mainly interested in microbes (especially bacteria). The software kraken-biom v1.2 (available at https://github.com/smdabdoub/kraken-biom) was used to combine kraken reports from individual samples into one tab-separated file. One file containing all samples, one file containing only root samples and one file containing only leaf samples were created. Output from the classification and metadata sets are put into R version 4.2.2 for further analyses. Microbial contamination introduced from laboratory reagents and DNA kits used were controlled for by comparing the taxa identified in the plant samples with those in the negative control samples via the R package decontam, using the 'prevalence' method with the threshold 0.5, which classifies all taxa that are more prevalent in blanks than in the plant samples as contaminants (Davis et al., 2018). Relative abundance was calculated after filtering of the assigned reads to normalize for different numbers of total classified reads among samples. Abundance filtering was done to remove all taxa with < 0.05% relative abundance in a sample.

Overall species richness was estimated to compare taxonomic diversity between leaf and root samples. This was done without accounting for the relative abundance of the species, i.e. using the number of species detected as the species richness metric. A rank-sum Wilcoxon test, which compares two groups of independent variables (leaf and root tissues) that are not normally distributed within a functional community, was carried out to test the statistical significance of higher species richness in a particular tissue.

PRINCIPAL COORDINATE ANALYSIS

After Kraken results and metadata were imported in R, principal coordinate analysis (PCoA) was conducted in R on the relative abundance data to estimate proximity matrices between different methods of preservation and varying collection sites. The PCoA was used to visualize whether the metagenomic communities amongst studied samples were more dissimilar (points ordinated farther from one another differ) or similar (points ordinated closer to another tend to be similar). Principal components (eigenvectors) were plotted, and ellipses based on 95% confidence intervals were estimated. PCoA showing the metagenomic communities present was carried out on all methods of preservation and sites of collections, performed for leaf and roots samples separately. Based on initial exploration of microbial community composition, all eight S2 leaf samples and two S3 samples (S3-5, S3-7) were found to be infected by *Phytophthora infestans*. These infected samples were removed from the leaf datasets for subsequent analyses.

DETERMINATION OF MOST ABUNDANT TAXA

MaAsLin 2.0 was used to identify taxa with different relative abundances as a result of different preservation conditions, while controlling for plant and region variation as random effects (Mallick et al., 2021). Each comparison was carried out between two kinds of preservation methods (e.g. P1 and P2, P2 and P6, etc.) to determine significant differences of any identified microbial taxa. MaAsLin 2.0 was run in R version 4.2.2, with no filtering of taxa by abundance (min_abundance = 0) and excluding taxa not detected in < 0.1% of samples (min_prevalance = 0.1), using the default linear model analysis (analysis_method = "LM"), normalisation of taxa counts via the centred log ratio transformation (normalisation = "CLR") and transformation of the normalized counts to $\log_2$ space (transform = "LOG"). A negative coefficient from MaAslin 2.0 indicates an increased abundance in the first preservation method of the comparison. A positive coefficient indicates higher abundance in the latter preservation method. Comparisons with q-values (false discovery rates) < 0.05 were regarded as significantly different. Comparisons were made to identify differences in taxa abundance between P1 and P2 (to identify the effect of fast compared to slow drying processes), P1 and P6 (to identify the effect of drying tissue in silica gel compared to collecting freshly frozen samples), P2 and P6 (to identify the effect of introducing samples to the herbarium) and P2 and P4 (to identify the effect of drying before introducing them to the herbarium and after introducing to the herbarium). While not the main purpose of this study, comparisons were also made between leaf samples infected and uninfected by *P. infestans* in order to identify changes in the microbial community during *P. infestans* infection.

RESULTS

SPECIES RICHNESS

Root samples have a mean species richness of 298.13 (standard deviation =49.7) and leaf samples have a mean species richness of 190.90 (standard deviation =130.26). This shows there is a significantly higher number of taxa per sample identified in root compared to leaf samples (Wilcoxon test, p-value = 0.00078). However, some leaf samples are outliers (Figure 3). The 20 most abundant species-level taxa in leaf and root samples were determined, presented by each preservation method and sites of collection (Figures 4, 5, 6 & 7). *Phytophthora infestans* was identified from all S2 leaf samples as well as some S3 leaf samples, thus these leaf samples were excluded from further analysis. In these, up to 6% of the raw reads mapped against the nuclear and nuclear genome of *P. infestans.*

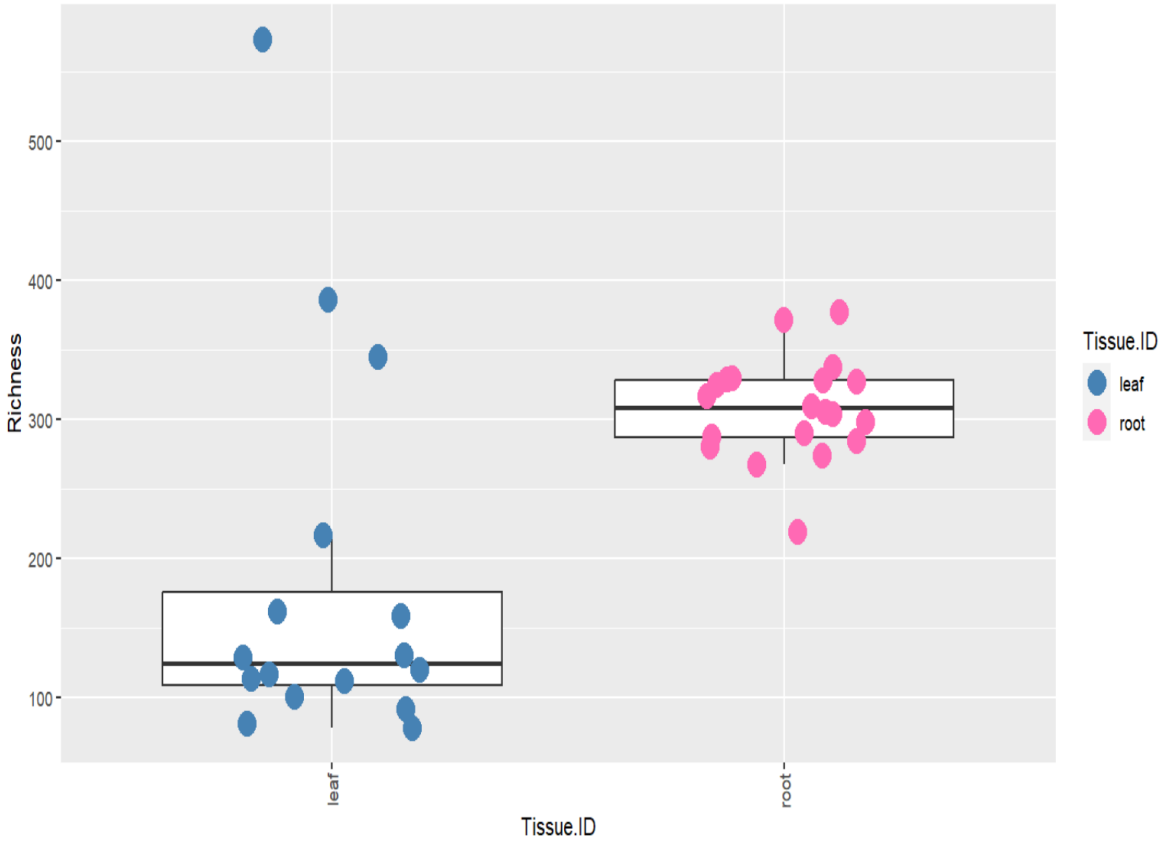


Figure 3. Overall species richness between leaf and root tissues studied under different preservation methods.

LEAF TAXA IDENTIFICATION ANALYSIS

For the leaf analysis, *Sphingomonas sp. PAMC26645, Sphingomonas sp. HMP9*, *Sphingomonas aerolata* and *Botrytis cinerea* are present in all P1 (silica-dried) leaf samples. *Methylobacterium bullatum, Methylobacterium duran* and *Methylobacterium SP WL1* are present in other leaf samples of P1 except M1-3. However, *Blumeria graminis* is detected only in M1-3 of P1 (Figure 4a). In P2 (after herbarium storage), three species of *Sphingomonas* and *Botrytis cinerea* are present in all leaf samples. Four species of *Methylobacterium, Cutibacterium acnes* and *Hymenobacter sp. PAMC26554* are present in all P2 leaf samples except M1-8. Also, *Blumeria graminis* (fungus causing powdery mildew in cereals especially wheat) is found only in M1-8 (Figure 4b). For leaf preservation method P3 (after herbarium storage with no cardboard change), only one sample was collected (M1-10), and the most prominent taxa in this sample are three species of *Sphingomonas, Methylobacterium bullatum, Botrytis cinerea* and *Blumeria graminis* (Figure 4c).

In P4 (after drying in plant press with cardboard change), M*ethylobacterium. sp. OTU13CASTA1,* three species of *Sphingomonas, Botrytis cinerea,* and *Frondihabitans sp.* 762G35 are found in all the leaf samples. However, *Hymenobacter sp. BRD128* and *Blumeria graminis* are found only in M1-5, but *Methylobacterium. sp. WL1* was absent in M1-5 (See Figure 4d below). Only one leaf sample (M1-7) was subjected to preservation method P5 (after drying in plant press with no cardboard change), and in this sample, three species of *Sphingomonas*, *Methylobacterium bullatum,* and *Botrytis cinerea* are the most abundant among many others (Figure 4e). Three species of *Sphingomonas, Methylobacterium ballatum, Botrytis cinerea* and *Comamonadaceae bacterium* OTU4NAUVB1 are detected in all samples of P6 (freshly frozen, Figure 4f). M*ethylobacterium sp WL1, Methylobacterium durans* and *Sphingomonas sp. AASP5* are found in other samples of P6, except in M1-2. *Frondihabitans sp. 762G35* is found only in S3-1 of P6 (Figure 4f).

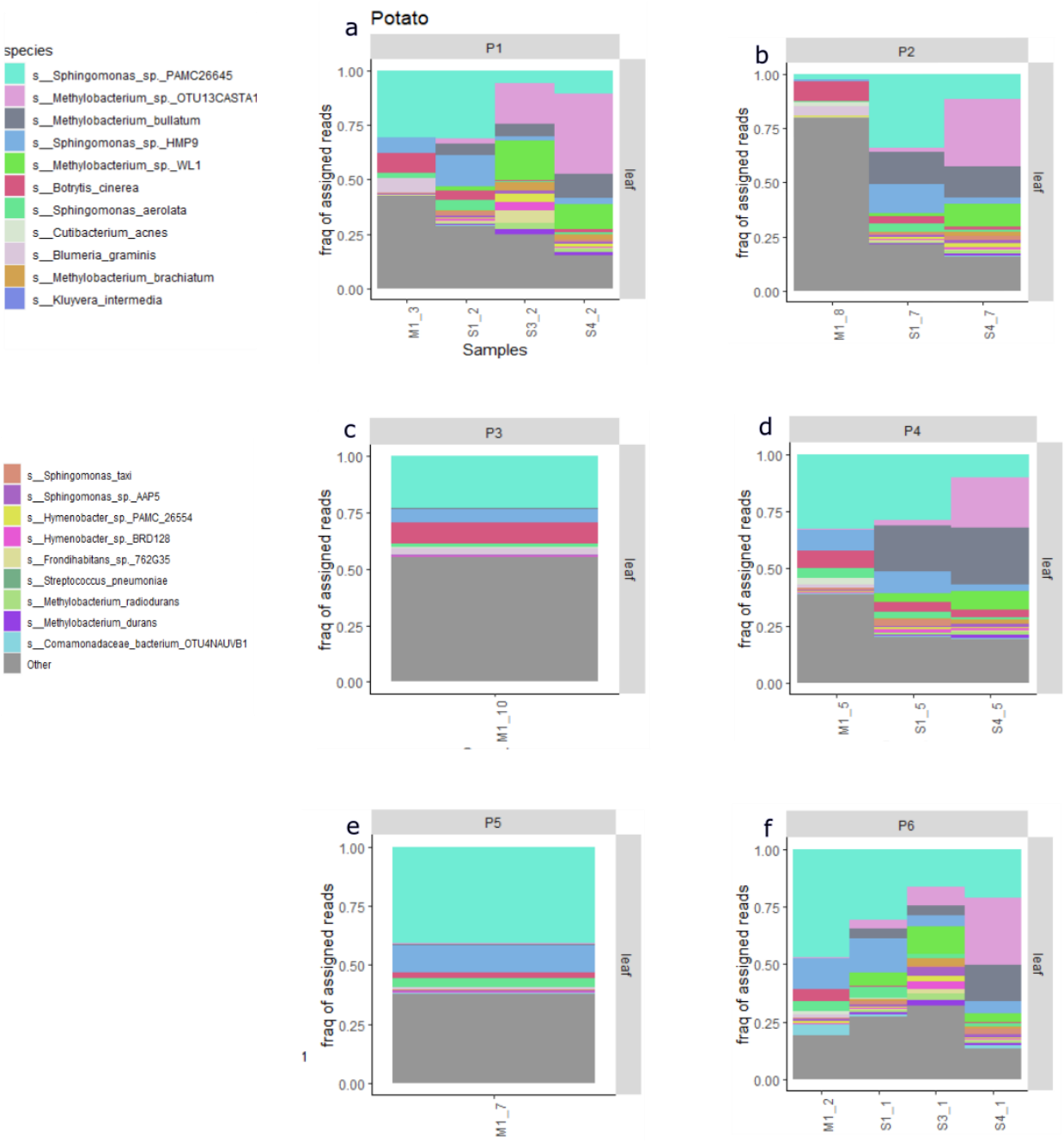


Figure 4. Top 20 most abundant species-level taxa identified for leaf samples. a) Silica dried (P1); b) Directly after storing in herbarium (cardboard changed) (P2); c) Directly after storing in herbarium (cardboards not changed) (P3); d) Directly after drying with plant presses (cardboard changed) (P4); e) Directly after drying with plant press (cardboards not changed) (P5); f) Freshly frozen at -20 degree celsius (P6). Species-level taxa are represented by different colours (see legend) and all species-level taxa not included in the top 20 are grouped as 'Other'.

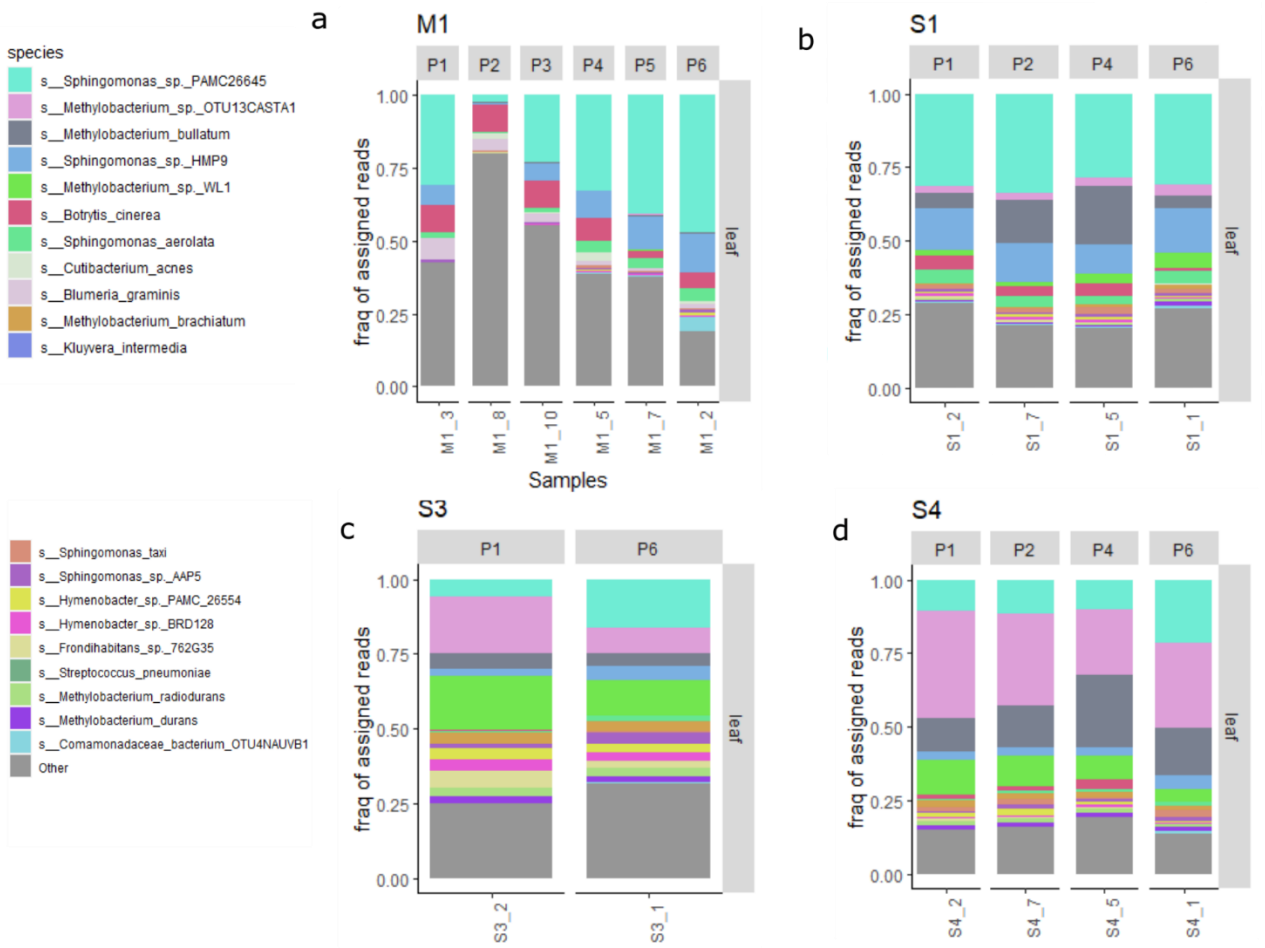


Figure 5. Top 20 most abundant taxa found on leaf samples collected from. a) Voll farm plant M1. b) Sverresborg folkemuseum plant S1. c) Sverresborg folkemuseum plant S3. d) Sverresborg folkemuseum plant S4. Preservation methods are as following: P1: Silica-dried; P2: Directly after storing in herbarium (cardboard changed); P3: Directly after storing in herbarium (cardboard NOT changed); P4: Directly after drying in the press (cardboard changed); P5: Directly after drying in the press (cardboard not changed); P6: Freshly frozen and stored at -20 ˚C. Species-level taxa are represented by different colours (see legend) and all species-level taxa not included in the top 20 are grouped as 'Other'.

Considering varying collection sites for leaf tissue, all plants collected from site M1 have similar metagenomic communities across all leaf samples (See figure 5 above). Three species of *Sphingomonas*, *Botrytis cinerea* and *Blumeria graminis* are present in all M1 samples. Additionally, *Cutibacterium acnes* are found in M1-8, M1-5 and M1-2. *Comamonadaceae bacterium* OTU4NAUVB1 is present only in M1-2 (Figure 5a). All S1 samples have five species of *Sphingomonas*, and two species of *Methylobacterium*, *Botrytis cinerea* and *Cutibacterium acnes* (Figure 5b). Three species of *Sphingomonas*, five species of *Methylobacterium* and *Frondihabitans. sp.* 762G35 are present in the examine S3 samples, because some were excluded due to infection (Figure 5c). The microbial communities in S4 leaf samples are very similar, with four species of

Sphingomonas, five species of *Methylobacterium* and *Botrytis cinerea* found at high abundances (Figure 5d).

ROOT TAXA IDENTIFICATION ANALYSIS

Different taxa are present in the root in comparison to the leaf samples, and *Phytophthora infestans* was not identified in any of them (See figure 6 below). Within the preservation method P1 of root samples, *Lentzea sp. HUAS12* is only present in M1-4. *Pseudomonas sp. OE 28.3, Rhodanobacter canariense,* and *Pseudomonas flourescens* are identified only in S4-4 of P1. *Variovax paradoxus* was identified in all P1 root samples, except for S4-4. *Niastella korensis, Bradyrhizobium erythrophlei, Bradyrhizobium sp. 170, Bradyrhizobium sediminis, Bradyrhizobium labiabi* and *Rhizophagus irregularis* are found in P1 root samples, except in M1-4 and S4-4 (Figure 6a). Taxa found in preservation method P2 are alike to taxa identified in P1, except *Bradyrhizobium sp. CCBAU051011* which is absent in M1-9 of P2. Only M1-9 in P2 has *Lentzea sp HUAS12* and *Pseudomonas flourescens* (Figure 6b). For both P3 and P5, no root sample was taken. The microbial communities in P4 are similar to P1 with the addition of an 'uncharacterised bacterium' across all P4 root samples. *Lentzea guizhouensis* is present only in M1-6 of P4. *Pseudomonas flourescens* was recognized only in S3-6 of P4 (Figure 6c). Metagenomic communities in P6-preserved root samples are highly similar to microbial communities in P1-preserved root samples (Figure 6a and d).

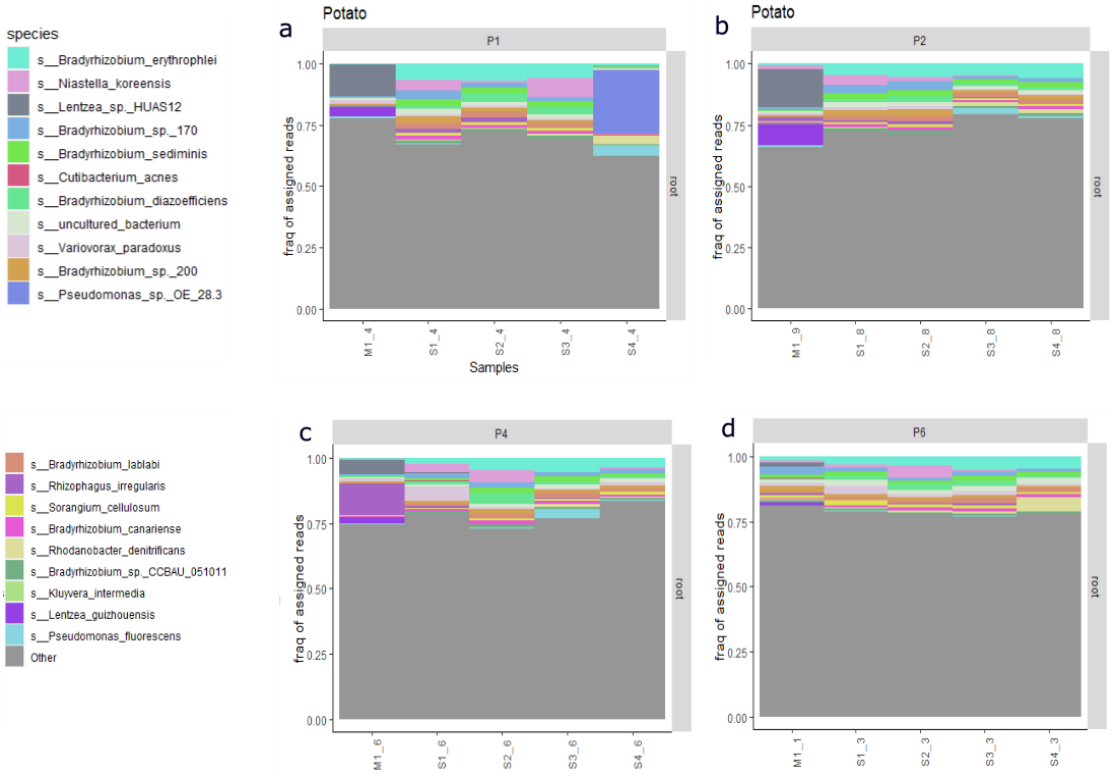


Figure 6. Top 20 most abundant taxa identified for root samples. a) Dried with silica gel (P1); b) Stored in the herbarium after drying with plant press (cardboards changed) (P2); c) After drying with the plant press (cardboards changed) (P4); d) Stored as freshly frozen at -20 degree celsius (P6). Species-level taxa are represented by different colours (see legend) and all species-level taxa not included in the top 20 are grouped as 'Other'.

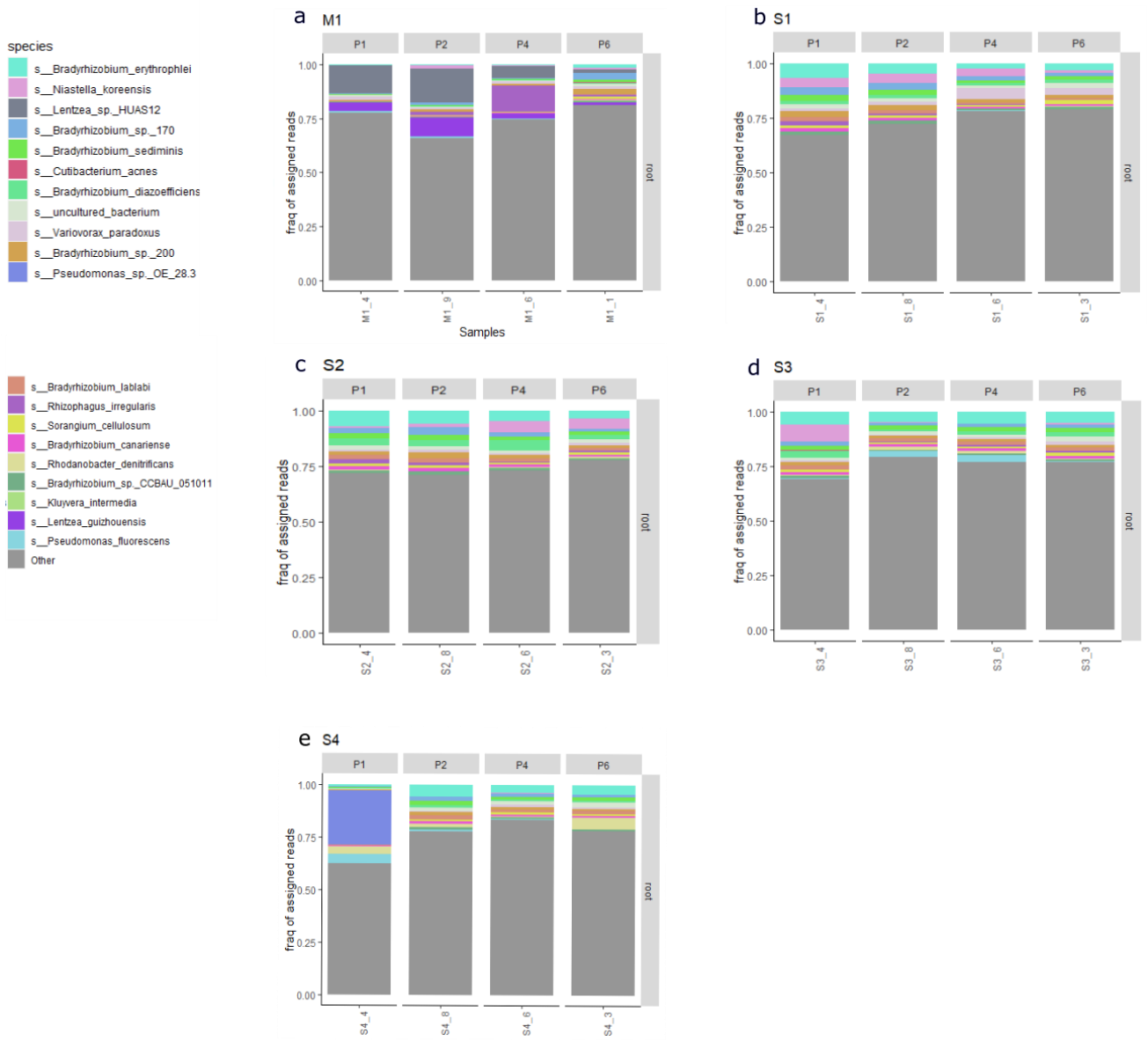


Figure 7. Top 20 most abundant taxa found on root samples collected from a) Voll farm. b) Sverresborg folkemuseum 1. c) Sverresborg folkemuseum samples 2. d) Sverresborg folkemuseum samples 3. e) Sverresborg folkemuseum samples 4. Preservation methods are as following: P1: Silica-dried; P2: Directly after storing in herbarium (cardboard changed); P3: Directly after storing in herbarium (cardboard NOT changed); P4: Directly after drying in the press (cardboard changed); P5: Directly after drying in the press (cardboard not changed); P6: Freshly frozen and stored at -20 ˚C. Species-level taxa are represented by different colours (see legend) and all species-level taxa not included in the top 20 are grouped as 'Other'.

When comparing root specimens from different collections sites (See figure 7 above), *Lentzea sp. HUAS12, B. sp. 170, Bradyrhizobium erythrophlei, Variovax paradoxus, Bradyrhizobium sp. 200* and *Pseudomonas fluorescens* are taxa present in all samples of M1. The metagenomic communities in S1 are similar to the communities found in M1 with additional taxa such as an 'uncharacterised bacterium', *Bradyrhizobium sp. CCBAU051011* and *Sorangium cellulosum* (Figure 7b). S2 has

metagenomic communities similar to S1 (Figure 7b & 7c). Five species of *Bradyrhizobium, Variovorax paradoxus, Niastella koreensis* and *Rhodanobacter dentrificans* are present in all samples of S4. However, *Pseudomonas sp. OE 28.3* is found only in S4-4 while *Bradyrhizobium diazoefficiens* is only absent in S4-4 (Figure 7d).

PRINCIPAL COORDINATE ANALYSIS

Considering the similarity matrix for leaf tissues, P1 (silica dried) of S3 and S4 plants are closer to each other compared to M1 and S1 (Figure 8). This indicates that microbial communities present in S3 and S4 for silica dried samples are similar while S1 and M1 have more variation (Figure 8a**).** P2 (after storing in herbarium, cardboard changed) samples have varying microbial communities across all sites of collection. From Figure 8b, the point representing M1 is found outside the ellipse, S1 and S4 are far apart as well. These results imply that different microbial communities are present in different collection sites though they were preserved in a similar manner. In addition, microbial communities found in S4, S1 and M1 plants of P4 are different (Figure 8c). Interestingly, there is close clustering (slight variation) among S1, S3 and S4 samples in P6, while the M1 sample is far apart from them. This shows that microbes present in S1, S3 and S4 collection sites are highly similar compared to M1 (a different farm) (Figure 8d).

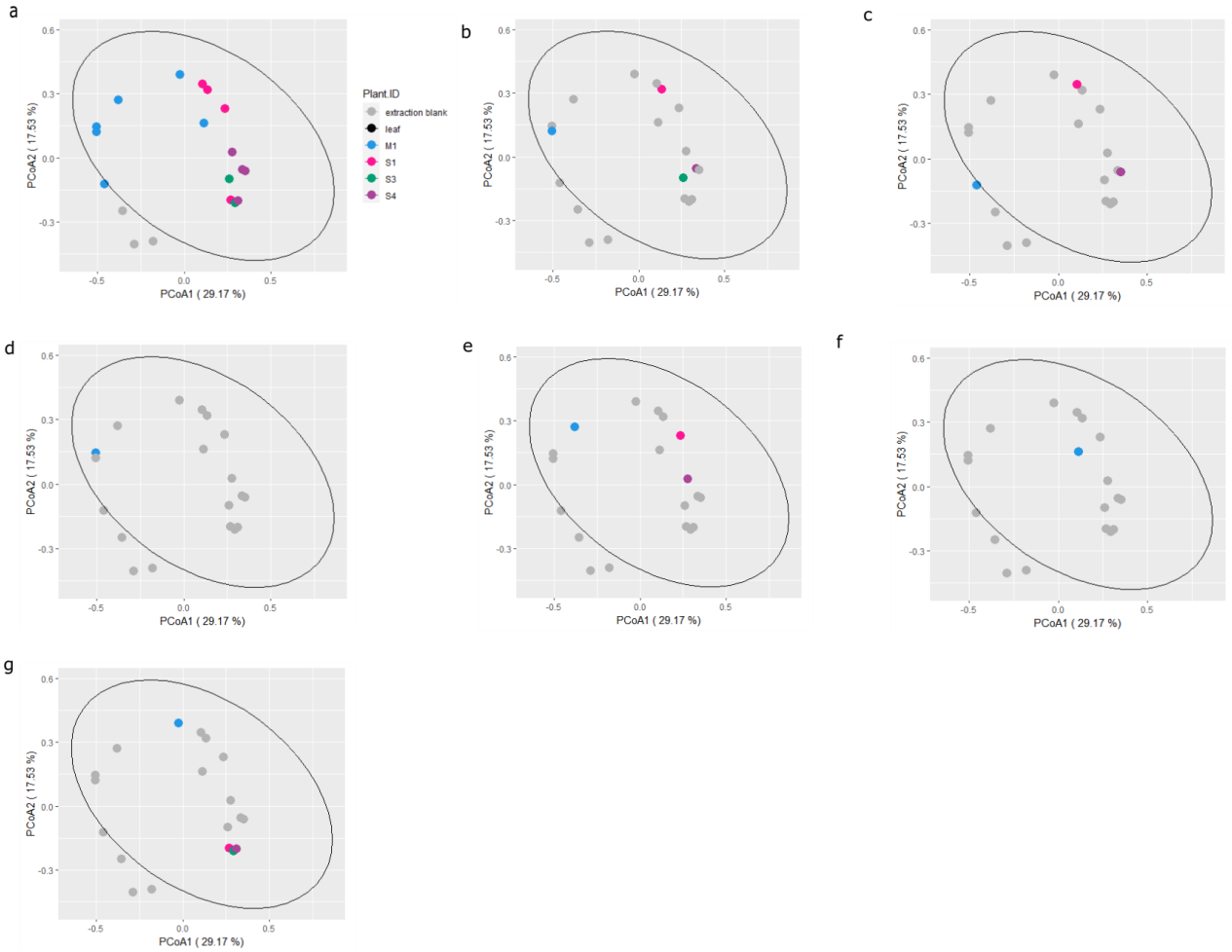


Figure 8. Ordinate of variation in leaf samples preserved using different methods in different collection sites. a) All sites of collection and all preservation methods b) Silica dried (P1); c) Directly after storing in herbarium (cardboard changed) (P2); d) Directly after storing in herbarium (cardboards not changed) (P3); e) Directly after drying with plant presses (cardboard changed) (P4); f) Directly after drying with plant press (cardboards not changed) (P5); g) Freshly frozen at -20 degree celsius (P6). In b-g, points in grey colours are samples that are not preserved by the specified preservation method and extraction blank. Ellipses represent tissue type (leaf). Points are colored based on the individual they were taken from.

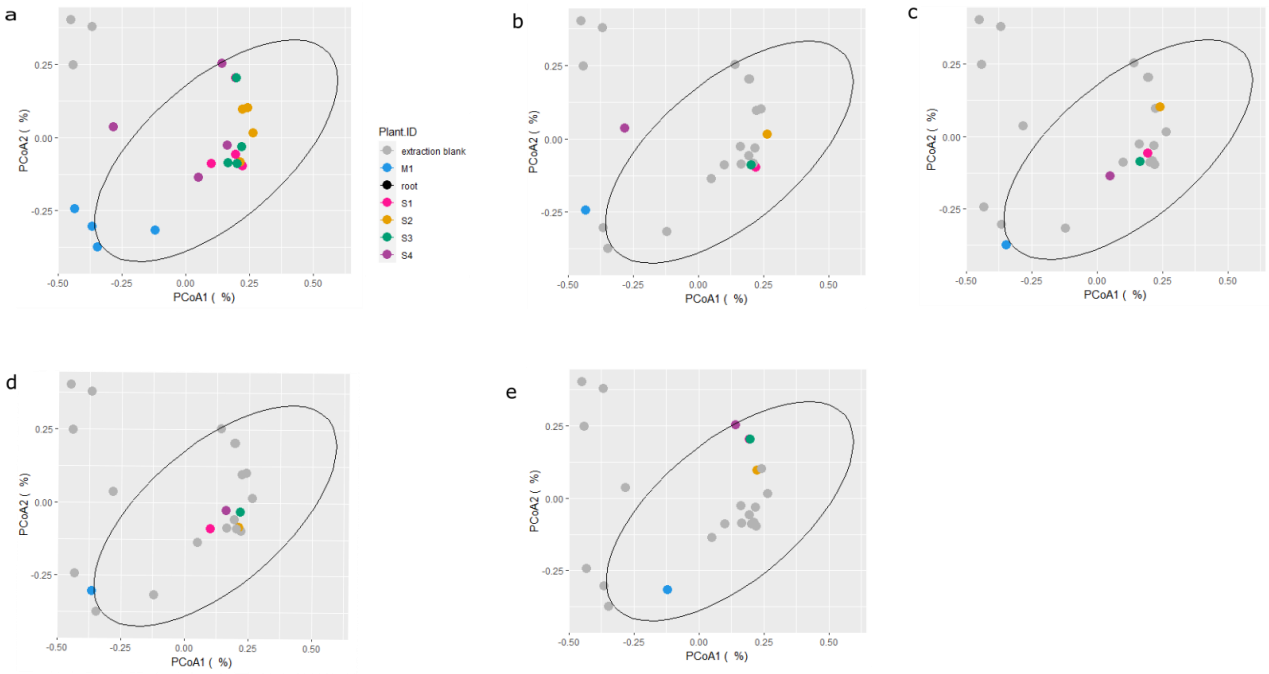


Figure 9. Ordinate of variation in root samples preserved using different methods. a) All sites of collection and all preservation methods b) Silica dried (P1); c) Directly after storing in herbarium (cardboard changed) (P2); d) Directly after drying with plant presses (cardboard changed) (P4); e) Freshly frozen at -20 degree celsius (P6). In b-e, points in grey colours are samples that are not preserved by the specified preservation method and extraction blank. Ellipses represent tissue type (root). Points are colored based on the individual they were taken from.

For the root analysis, P1 preserved samples of S1 and S3 plants are in close proximity (See figure 9 above). The S2 sample is a bit closer to S1 and S3 samples, while M1 and S4 samples are widely apart from others. This indicates that the metagenomic communities present in S1 and S3 are highly similar. The microbial communities found in S2 have some similarities with S1 and S3, but greatly different in M1 and S4 (see Figure 9a). For P2, the M1 samples are far apart from others. S1 and S3 are highly similar, while S4 and S2 are far apart from S1 and S3. This means that S1 and S3 have similar metagenomic communities compared to others. Also, the microbes found in M1 have great disparities amidst others (Figure 9)**.** In P4, only M1 is greatly separated from others. It has a coordinate outside the ellipse, whereas other collection sites have proximity. (Figure 9c). For P6, S1 and S3 samples have slightly overlapping coordinates. S4 and S2 are a bit far from S1 and S3, but M1 is far apart from other ordinates. It implies that there is little or no variation in the microbial communities of S1 and S3. S2 and S4 have different microbes but may share some similar microbes with S1 and S3. However, M1 is far apart from others, indicating the presence of different metagenomic communities (Figure 9d).

TAXA ABUNDANCE ANALYSIS

For leaf tissues, no taxa are significantly different in abundance (q-value < 0.05) in all compared methods of preservation. Analysis carried out between *Phytophthora infestans* infected and non-infected leaf samples reveal that *Phytophthora infestans*, *Phytophthora parasitica*, *Bacillus licheniformis* and *Bradyrhizobium erythrophlei* have higher abundance in the infected samples (Figure 10a–d and Table S3). In contrast, *Xylophilus rhododendri* and *Methylobacterium spp*. AMS5 have significantly lower abundance in infected samples (Figure 11a–b).

For the root analysis, comparison between P1 and P2 reveals that *Granulicella sp. WH15* has higher abundance in P2 (after storing in herbarium) than P1 (silica-dried). *Granulicella sp. WH15* has been previously isolated from decaying wood (Costa et al., 2020). *Acidovorax avenae* has higher abundance in P2 compare to P1. Comparison between P2 and P6 found that *Massilia forsythiae*, has higher abundance in P2 than in P6. Comparisons between P1 and P6 reveals that *Streptacidiphilus sp. P02.A3a*, a bacteria isolated from decaying pinewood has higher abundance in P1 than in P6. Also, *Frateuria sp. 5GH9.34* has higher abundance in P6 than P1. (further details are in Supplementary tables).

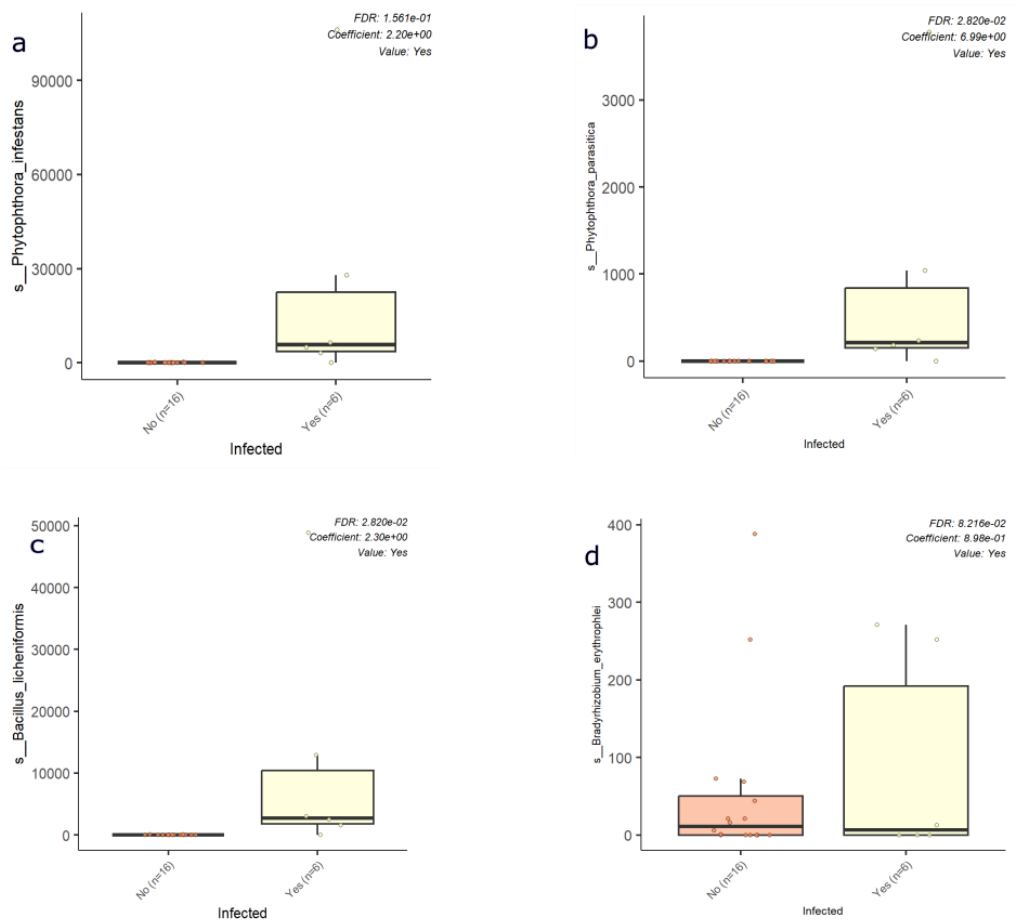


Figure 10. Differential abundance of microbial taxa on infected leaf samples by *Phytophthora infestans*. a) *P. infestans*. b) *P. parasitica*. c) *B. erythrophlei*. d) *B. licheniformis.* Yes represent the number of samples infected with *Phytophthora infestans*. No represent the number of samples uninfected with *Phytophthora infestans*.

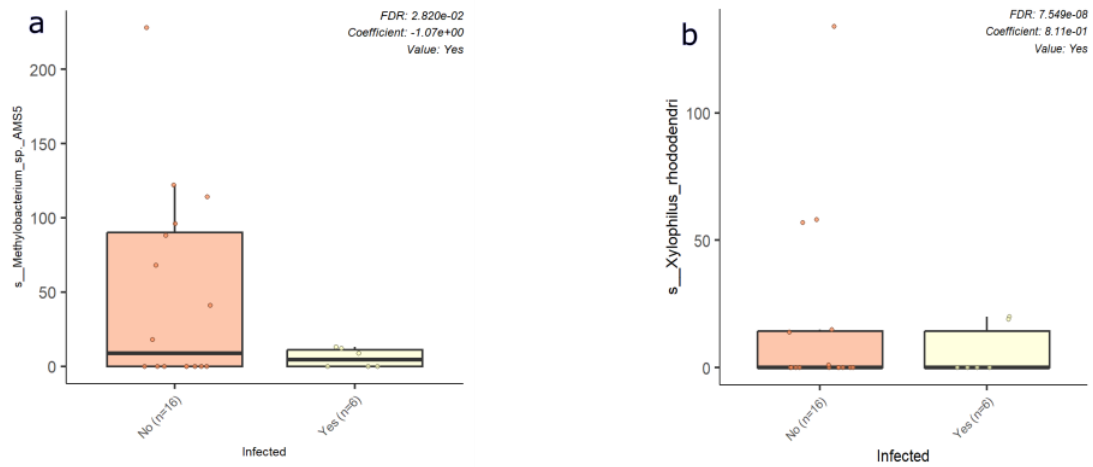


Figure 11. Differential abundance of microbial taxa on uninfected leaf samples. a) *Methylobacterium sp.* AMS5. b) *Xylophilus rhododendri*. Yes represents the number of samples infected with *Phytophthora infestans*. No represents the number of samples uninfected with *Phytophthora infestans*; n indicates the number of observations.

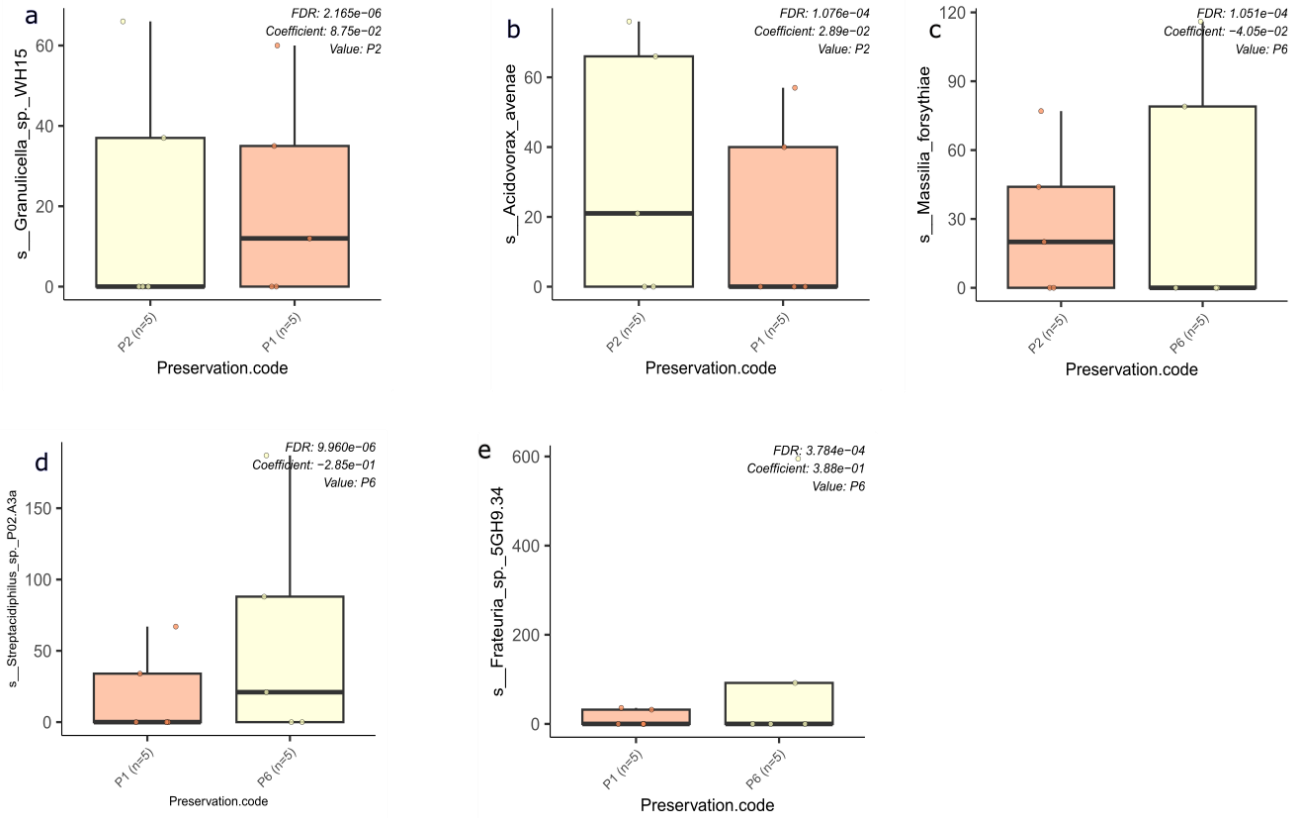


Figure 12. Differential abundance of microbial taxa in root samples. a) *Granulicella sp. WH15* shows higher abundance in P2 when comparing P1 and P2. b) *Acidovorax avenae* shows higher abundance in P2 when comparing P1 and P2. c) *Massilia forsythia* shows higher abundance in P2 when comparing P2 and P6. d) *Streptacidiphilus sp. P02.A3a* shows higher abundance in P1 when comparing P1 and P6. e) *Frateuria sp. 5GH9.34* shows higher abundance in P6 when comparing P1 and P6; n indicates number of observations. P1: Silica dried; P2: Directly after storing in herbarium (cardboard changed); P6: Freshly frozen at -20 degree celsius.

DISCUSSION

This pilot study aimed to track changes in microbial abundance and identify microbial taxa throughout the different stages of herbarium specimen preservation and thus contaminate metagenomic datasets with their DNA. These microbial contaminants are evidenced by DNA sequence reads that were not truly present in the living plant. They potentially come from various sources such as laboratory reagents or environments, or they could be introduced during the process of drying, when mounting samples in the herbarium etc. (Davis et al., 2018). Generally, in this study, fewer highly abundant microbial taxa (contaminants) were detected than was expected. A key reason for this may be that there was only little human contact with the specimens. For instance, human contact with specimens could lead to infection with skin microbes like *Cutibacterium acnes* (Mayslich et al., 2021). From the figure 4b & 4d, high qualitative representation of *Cutibacterium acnes* was found in P2 and P4 (samples where cardboard was changed and thus had more contact with humans) than in samples without changed cardboards (P3 and P5). However, the sample size was too low to carry out a test of statistical significance. For historic samples, collectors, herbarium staff and visitors are more likely to have had several contacts with samples over the long period of storage in the herbarium. Therefore, specimens kept in the herbarium for a longer time are likely to have had occasions for various microbes to colonize their tissues (Bieker & Martin, 2018). It is common in morphological studies, especially of type specimens, that specimens are loaned from one herbarium to another for research reasons. The process of packing and moving the specimens from one herbarium to another, microbes can also be introduced in them. Also, sending specimens from one herbarium to another has the tendency of changing temperature and humidity conditions of the specimens and may favor the growth of contaminants such as mould fungi. Thus, more loans and examinations could provide more opportunities for microbial colonization.

A comparison of root and leaf samples in this study found that generally, microbial taxa identified on leaf tissues are different from microbes on root tissues. Also, microbial taxa identified in the root samples were significantly more diverse in terms of species richness than in the leaf samples. It has been reported by several authors in the last one decade that there are different microbial collections found on different plant parts (Hardoim et al., 2015; Reinhold-Hurek et al., 2015; Vorholt, 2012). Several studies have shown that root samples tend to have higher species richness than leaves (Bulgarelli et al., 2012; Yeoh et al., 2017). These variations in their species richness can be due to many unique conditions (such as nutrients available for use and

microenvironments) in the soil microbiomes that play a key role in the formation and nurturing of diversified microbial communities in the roots (Lundberg et al., 2012).

The different preservation methods led to significant differences in microbial taxa abundances among root samples. However, no significant changes in microbial taxa abundances were found in leaf samples. From the leaf sample results, it can be inferred that all preservation methods used are good enough to prevent the *antemortem*/*postmortem* growth of microbial taxa, provided samples are kept under stable conditions both inside and outside the herbarium. This is because the Trondheim Herbarium was recently renovated, the temperature and humidity are controlled and thus stable, which may have contributed to the unexpectedly limited *antemortem*/*postmortem* growth of microbes. In a study by (Forrest et al., 2019), the impact of preservation techniques on *Begonia* plants was evaluated, revealed that specimens dried with silica gel (fast drying) and all slow drying methods for the study have better host plant DNA quality and quantity than other preservation methods such as drying with hair dryer, alcohol, RNAlater and pickling. We hypothesize that when the genetic materials of plants are well preserved, the genetic information of the microbes on the plants would equally be preserved. The findings by (Forrest et al., 2019) supports our findings that no or less contaminating microbial DNA can be found when specimens are well preserved.

Furthermore, *Alternaria alternata* (a common herbarium contaminant) was completely absent in all studied samples. It has been previously reported by (Bieker et al., 2020) that infection rate of *A. alternata* on historical herbarium specimens of the plants *Arabidopsis thaliana* and *Ambrosia artemisiifolia* vary between 33% to 100% among tested herbaria. The differences between herbaria may be due to their different indoor environmental conditions. It is possible that less stable conditions or higher humidity favour infection with *A. alternata*. We hypothesize that the unexpectedly few microbial contaminants identified in this study may be due to the standard good practices followed during the preservation manipulations as well as the stable, well-controlled, dry conditions in the recently renovated Trondheim Herbarium (TRH). Whether an unstable and/or humid indoor environment is more beneficial for the growth of microbes like *Alternaria alternata* should be further investigated in future studies.

LEAF SAMPLE MICROBES CHARACTERIZATION, ABUNDANCE AND RICHNESS

Due to the overall major differences in microbial communities between leaf and root samples, they were analyzed separately in this study. However, interpretation of the leaf sample results was complicated by the molecular identification of *Phytophthora infestans* in some individuals. Leaf samples infected with *P. infestans* also carried other associated taxa (*Phytophthora parasitica, Bacillus licheniformis* and *Bradyrhizobium erythrophlei*) with high abundance whereas *Xylophilus rhododendri* and *Methylobacterium sp. AMS5* have a relatively low abundance in both uninfected and infected samples. However, it is not clearly known whether these microbes are highly abundant in infected samples because of the compromised immune systems caused by *P. infestans* or whether the presence of any of these bacteria increased the plant's susceptibility to infection with *P. infestans*. Several studies had revealed that some plant microbial interactions are detrimental to these interactions, especially fitness (Bordenstein & Theis, 2015; Moran & Sloan, 2015; Vandenkoornhuyse et al., 2015). Future studies should investigate whether *P. parasitica, B. licheniformis* and *B. erythrophlei* are interacting with *P. infestans* or contributing to disease. *Methylobacterium sp. AMS5* was found to be at higher relative abundance in uninfected samples. Genus *Methylobacterium* is a common bacterial taxon that dominates leaf surfaces. These bacteria are beneficial to plants as good consumers of plant-secreted methanol and by producing plant-growth-promoting metabolites (Sanjenbam et al., 2022). Moreso, if *Methylobacterium* is generally beneficial to the plant, and it decreases in relative abundance during infection, it may be a sign that the plant is stressed, or that *P. infestans* infection disrupts the healthy plant leaf microbial community. A research by (Schubert et al., 2014) revealed that genus *Methylobacterium* was not found or only in low relative abundance when taxonomic profile of microbial genera was conducted on three historical samples infected by *Phytophthora infestans*. *Methylobacterium* was identified only in one of the three historical infected samples studied by (Schubert et al., 2014). This supports my finding that *Methylobacterium* has significantly reduced abundance in infected leaf samples compared to uninfected ones (Fig. 11a). However, their discoveries cannot be directly comparable to my findings, because they studied the metagenome of historic infected samples at the genus level, whereas my research examined specimens at species level. In addition, I used Kraken tools and ncbi_ntDB database for my taxa classification while (Schubert et al., 2014) used Metagenomic phylogenetic analysis database of genomic markers. The difference in the database used may have caused some disparities in our findings. It would thus be good to analyze these samples with the same pipeline used here gain comparable results. The effect of *Xylophilus rhododendri* cannot be

clearly explained on infected and non-infected samples because there would not be a clear-cut difference when three points serving as outliers are removed in the uninfected samples. It would be interesting in the future to examine the effects of *P. infestans* on *Xylophilus rhododendri* with a larger sample size. Finally, due to some biases in the database used, it could be that some sequences assigned as *P. parasitica* actually originate from *P. infestans*, which could be responsible for the higher abundance of *P. parasitica*. A research by Huson et al.,(2007) revealed that reads can be assigned to taxon closely related to the exact taxon when true taxon is absent in the database.

Two prominent microbial taxa, *Sphingomonas* and *Methylobacterium* were identified across all kinds of preservation methods in all leaf samples studied. A previous study using metagenomic analysis (Sumbula et al., 2020) identified the genus *Sphingomonas* amongst the bacterial diversity existing on tomato leaves in high abundance (*Solanum lycopersicum*). So, the findings from (Sumbula et al., 2020) supports my findings that *Sphingomonas* are common microbial taxa on leaf samples of genus *Solanum.*

We observed considerable variation in the leaf microbial communities between sample sites. In leaf samples from site M1, *Blumeria graminis* was present across multiple preservation stages. *B. graminis* is a fungus responsible for powdery mildew in wheat all over the world (Costamilan, 2005). We hypothesize that the presence of this fungus in M1 samples could be dependent solely on the collection site (Voll farm). Such variation in fungal communities among sampling sites has been observed in plants previously. (Lee & Hawkes, 2021) evaluated the variation in fungi on leaf of switch grass from coastal to mountains from 14 sites in North Carolina. Different fungal communities were identified on the plant leaves across different sites. Fan et al. (2019) revealed changes in microbial communities (especially bacteria and fungi) under varying vegetation restoration patterns in different ecosystems (karst and non-karst ecosystem).

There was variation in microbial communities found in all preservation methods across different sites of collection, except in P6 (freshly frozen samples at -20 ˚C). In P6, an overlap of S1, S3 and S4 samples was recorded i.e. microbial communities in these collection sites. It depicts great similarities between the microbial communities preserved as freshly frozen samples and some variations among metagenomic communities found in other methods of preservation. Changes (reduction) in DNA quality were recorded in ethanol-preserved samples versus fresh-frozen samples in selected tropical plants (Bressan et al., 2014). If the freshly frozen method preserves DNA better, it might also better preserve the microbial community DNA, resulting in more similar metagenomic results among individuals from the same collection site.

No microbial taxa have significant changes in abundance when comparisons were made between two preservation methods for leaf samples. It could be that all methods of preservation

used for leaf samples were too good to prevent the further growth of microbes, at least to the extent that we could detect no statistically significant changes with MaAslin 2.0 and with the study's small sample size.

ROOT SAMPLE MICROBES CHARACTERIZATION, ABUNDANCE AND RICHNESS

Various species of *Bradyrhizobium* were the most abundant taxa in the root samples. Species of the genus *Bradyrhizobium* are common root endophyte bacteria, and they are either symbiotic or free-living (Ramírez-Puebla et al., 2019; Schneijderberg et al., 2018). Hence, the high prevalence of *Bradyrhizobium* in the root samples studied was expected. There were variations in the identified root microbial communities based on different collection sites and preservation methods. For example, all root samples of M1 have their ordination outside of the ellipse, except P6. Thus, it is likely that all other methods of preservation have altered the microbial communities of M1 samples except for the freshly frozen samples (P6). Also, S4 of silica dried (P1) showed great variation when compared with other preservation levels. The ordination representing S4 of P1 was found outside the ellipse. It is likely that the fast-drying method has altered the microbial communities present in S4. Forrest et al. (2019) reported that different preservation methods may influence the length of DNA fragments and shorter fragments are harder to classify with Kraken. It was concluded that the process of drying can fragment plant DNA and can provide growth opportunities for microorganisms that are not naturally found on host plants, thus making microbial DNA more abundant relative to the host DNA.

Few taxa show significant changes in relative abundance when comparisons were made between two preservation methods for root samples. *Acidovorax avenae* has a high relative abundance on root samples in P2 compared to P1. *A. avenae* causes red stripe disease in sugarcane (Fontana et al., 2013). It could be that *A. avenae* can survive and grow more on dead plant material under a slow drying process (P2) than drying with silica gel (P1). Also, this taxon is not known to occur on potato but it occurs on *Cucurbitaceae, Cattleya, Phalaenopsis*, and *Poaceae* (Willems & Gillis, 2015). We hypothesize that this taxon could have also been from nearby grasses (*Poaceae*). If any of these assumptions are not true, it could depict that a new host (potato) has been discovered for *A. avenae.* In addition, *Streptacidiphilus sp. P02.A3a* has higher abundance in P1 than P6. *Streptacidiphilus sp. P02.A3a* is a bacterium isolated from decaying pinewood (Elsayed et al., 2020), suggesting that *Streptacidiphilus. sp. P02.A3a* is better preserved using the fast-drying method (P1) than keeping the host plant fresh (P6). In other words, it may be that *S. sp. P02.A3a* grows better on silica gel rather than on living plant material. *Frateuria sp. 5GH9.34* was found

with a high abundance in P6 than P1. *Frateuria sp. 5GH9.34* was originally isolated from the plant (*Lilium auratum*). It implies that *Frateuria* bacteria are better preserved in freshly frozen samples than during the fast-drying procedure.

LIMITATIONS

This pilot study explored differences in microbial plant communities during various specimen preservation procedures and methods. As it was a pilot study, only a small sample size was possible and the study would have benefited from including more samples among the different preservation methods tested. For example, an interesting comparison to consider would have been to compare taxa abundances between leaf samples of P2 (after storing in herbarium with changing cardboard) and P3 (after storing in herbarium without cardboard changed). It would allow testing whether less contaminants would be introduced in P3 due to less contact with humans, since cardboard used for P3 will not be changed from time to time. Unfortunately, our small sample size presents no statistical power to execute this test as only one sample was collected for P3. In the future, it would be exciting to investigate the effects of cardboard changing during the drying processes, as it is a common act today to replace cardboards during the process of drying to allow faster drying. A large sample size would probably need to be included for P3 and P5 samples.

The study was also limited by the microbial taxonomic identification method used. Taxonomic identification of short metagenomic reads is known to be affected by biases in the database used. Velsko et al. (2018) carried out a test of a program on ancient metagenomic communities. It was discovered that amongst other factors, the reference database affected the results. Kraken may have classified some taxa as absent because they were not present in the database used. Some of the reads may have been classified to a close taxon but not to the specific taxon to which they belong. A study by (Huson et al., 2007) where explicit investigation was done on the use of MEGAN for metagenomic data identification mentioned that reads can be assigned to the lowest common ancestors or closely related taxa in a database when the exact taxa is not represented in the database. For instance, only a few fungi were identified, possibly because the database used had insufficient fungi taxonomic ranking features. In the future, data can be consolidated from several sources to generate extensive and diversified databases. Researchers can foster collaborations to understand challenges and needs in improving databases.

CONCLUSIONS AND FUTURE PROSPECTS

We inferred a higher number of taxa per sample in root compared to the leaf samples. Great variations were found between identified microbial communities of potato leaf and root tissues studied. The most common species found in leaf samples are species from genera *Sphingomonas* and *Methylobacterium*, while genus *Bradyrhizobium* is the most abundant in root samples. Microbial communities are better preserved in freshly frozen samples in leaf and root tissues than other methods of preservation. Other preservation methods used in this study altered the microbial communities to different degrees. No significant changes in microbial taxa abundance in leaf samples were observed while few changes in taxa abundance were recorded in root samples. TRH has stable conditions that favour proper preservation of specimens, hence lesser contaminants were found on all samples studied than expected. For further research, specimens can be kept for a longer time under the same drying procedures and herbarium conditions used in this study to test whether long duration of preservation can bring drastic change to microbial communities present on the samples. In addition, large sample sizes of specimens after storing in herbarium (cardboard not changed) and after drying with press (cardboard not changed) can be created to be able to test statistically for the effect of human contacts on specimens with P2 and P4 respectively. Proper identification and removal of contaminants from present day herbarium specimens can allow scientists to know the effects of chemicals on plants metagenomic communities upon comparisons with old herbarium samples. In addition, inferences can be drawn about future evolutionary trends of metagenomic communities present on potato plants.

REFERENCES

Abdelfattah, A., Wisniewski, M., Li Destri Nicosia, M. G., Cacciola, S. O., & Schena, L. (2016). Metagenomic Analysis of Fungal Diversity on Strawberry Plants and the Effect of Management Practices on the Fungal Community Structure of Aerial Organs. *PloS One*, *11*(8), e0160470.

Adams, R. P., & Sharma, L. N. (2010). DNA from herbarium specimens: I. Correlation of DNA size with specimen age. *Phytologia*, *92*, 346–353.

Akerele, O., Heywood, V., & Synge, H. (1991). *Conservation of Medicinal Plants*. Cambridge University Press.

Ames, M., & Spooner, D. M. (2008). DNA from herbarium specimens settles a controversy about origins of the European potato. *American Journal of Botany*, *95*(2), 252–257.

Austin Bourke, P. M. (1964). Emergence of potato blight, 1843–46. *Nature*, *203*(4947), 805–808.

Balick, M. J. (1996). *Plants, People, and Culture: The Science of Ethnobotany*. Scientific American Library.

Berger, S., Sinha, A. K., & Roitsch, T. (2007). Plant physiology meets phytopathology: plant primary metabolism and plant–pathogen interactions. *Journal of Experimental Botany*, *58*(15-16), 4019–4026.

Berlec, A. (2012). Novel techniques and findings in the study of plant microbiota: search for plant probiotics. *Plant Science: An International Journal of Experimental Plant Biology*, *193-194*, 96–102.

Besnard, G. (2018). *Special Issue: Herbarium-based Science in the Twenty-first Century*. Taylor et Francis.

Bieker, V. C., & Martin, M. D. (2018). Implications and future prospects for evolutionary analyses of DNA in historical herbarium collections. *Botany Letters*, *165*(3-4), 409–418.

Bieker, V. C., Sánchez Barreiro, F., Rasmussen, J. A., Brunier, M., Wales, N., & Martin, M. D. (2020). Metagenomic analysis of historical herbarium specimens reveals a postmortem microbial community. *Molecular Ecology Resources*, *20*(5), 1206–1219.

Bordenstein, S. R., & Theis, K. R. (2015). Host Biology in Light of the Microbiome: Ten Principles of Holobionts and Hologenomes. *PLoS Biology*, *13*(8), e1002226.

Brader, G., Compant, S., Vescio, K., Mitter, B., Trognitz, F., Ma, L.-J., & Sessitsch, A. (2017). Ecology and Genomic Insights into Plant-Pathogenic and Plant-Nonpathogenic Endophytes. *Annual Review of Phytopathology*, *55*, 61–83.

Bressan, E. A., Rossi, M. L., Gerald, L. T. S., & Figueira, A. (2014). Extraction of high-quality DNA from ethanol-preserved tropical plant tissues. *BMC Research Notes*, *7*, 268.

Bulgarelli, D., Rott, M., Schlaeppi, K., Ver Loren van Themaat, E., Ahmadinejad, N., Assenza, F., Rauf, P., Huettel, B., Reinhardt, R., Schmelzer, E., Peplies, J., Gloeckner, F. O., Amann, R., Eickhorst, T., & Schulze-Lefert, P. (2012). Revealing structure and assembly cues for Arabidopsis root-inhabiting bacterial microbiota. *Nature*, *488*(7409), 91–95.

Burdon, J. J. (1987). *Diseases and Plant Population Biology*. CUP Archive.

Calinger, K. M., Queenborough, S., & Curtis, P. S. (2013). Herbarium specimens reveal the footprint of climate change on flowering trends across north-central North America. *Ecology Letters*, *16*(8), 1037–1044.

Costa, O. Y. A., Zerillo, M. M., Zühlke, D., Kielak, A. M., Pijl, A., Riedel, K., & Kuramae, E. E. (2020). Responses of Acidobacteria Granulicella sp. WH15 to High Carbon Revealed by Integrated Omics Analyses. *Microorganisms*, *8*(2). https://doi.org/10.3390/microorganisms8020244

Cowan, D., Meyer, Q., Stafford, W., Muyanga, S., Cameron, R., & Wittwer, P. (2005). Metagenomic gene discovery: past, present and future. *Trends in Biotechnology*, *23*(6), 321–329.

Davis, N. M., Proctor, D. M., Holmes, S. P., Relman, D. A., & Callahan, B. J. (2018). Simple statistical identification and removal of contaminant sequences in marker-gene and metagenomics data. *Microbiome*, *6*(1), 226.

Délye, C., Deulvot, C., & Chauvel, B. (2013). DNA analysis of herbarium Specimens of the grass weed Alopecurus myosuroides reveals herbicide resistance pre-dated herbicides. *PloS One*, *8*(10), e75117.

Denancé, N., Sánchez-Vallet, A., Goffner, D., & Molina, A. (2013). Disease resistance or growth: the role of plant hormones in balancing immune responses and fitness costs. *Frontiers in Plant Science*, *4*, 155.

Elsayed, S. S., Genta-Jouve, G., Carrión, V. J., Nibbering, P. H., Siegler, M. A., de Boer, W., Hankemeier, T., & van Wezel, G. P. (2020). Atypical Spirotetronate Polyketides Identified in the Underexplored Genus Streptacidiphilus. *The Journal of Organic Chemistry*, *85*(16), 10648–10657.

Etesami, H. (2022). Root nodules of legumes: A suitable ecological niche for isolating non-rhizobial

bacteria with biotechnological potential in agriculture. *Current Research in Biotechnology*, *4*, 78–86.

Fadiji, A. E., & Babalola, O. O. (2020). Metagenomics methods for the study of plant-associated microbial communities: A review. *Journal of Microbiological Methods*, *170*, 105860.

Fan, Z., Lu, S., Liu, S., Guo, H., Wang, T., Zhou, J., & Peng, X. (2019). Changes in Plant Rhizosphere Microbial Communities under Different Vegetation Restoration Patterns in Karst and Non-Karst Ecosystems. *Scientific Reports*, *9*(1), 8761.

Ficke, A., Cowger, C., Bergstrom, G., & Brodal, G. (2018). Understanding Yield Loss and Pathogen Biology to Improve Disease Management: Septoria Nodorum Blotch - A Case Study in Wheat. *Plant Disease*, *102*(4), 696–707.

Fontana, P. D., Rago, A. M., Fontana, C. A., Vignolo, G. M., Cocconcelli, P. S., & Mariotti, J. A. (2013). Isolation and genetic characterization of Acidovorax avenae from red stripe infected sugarcane in Northwestern Argentina. *European Journal of Plant Pathology / European Foundation for Plant Pathology*, *137*(3), 525–534.

Forrest, L. L., Hart, M. L., Hughes, M., Wilson, H. P., Chung, K.-F., Tseng, Y.-H., & Kidner, C. A. (2019). The Limits of Hyb-Seq for Herbarium Specimens: Impact of Preservation Techniques. *Frontiers in Ecology and Evolution*, *7*. https://doi.org/10.3389/fevo.2019.00439

Gilberg, M., & Brokerhof, A. (1991). The Control of Insect Pests in Museum Collections: The Effects of Low Temperature on *Stegobium Paniceum* (Linneaus), the Drugstore Beetle. *Journal of the American Institute for Conservation*, *30*(2), 197–201.

Gupta, A., Gupta, R., & Singh, R. L. (2017). Microbes and Environment. In R. L. Singh (Ed.), *Principles and Applications of Environmental Biotechnology for a Sustainable Future* (pp. 43–84). Springer Singapore.

Haas, B. J., Kamoun, S., Zody, M. C., Jiang, R. H. Y., Handsaker, R. E., Cano, L. M., Grabherr, M., Kodira, C. D., Raffaele, S., Torto-Alalibo, T., Bozkurt, T. O., Ah-Fong, A. M. V., Alvarado, L., Anderson, V. L., Armstrong, M. R., Avrova, A., Baxter, L., Beynon, J., Boevink, P. C., … Nusbaum, C. (2009). Genome sequence and analysis of the Irish potato famine pathogen Phytophthora infestans. *Nature*, *461*(7262), 393–398.

Hahn, M. (2014). The rising threat of fungicide resistance in plant pathogenic fungi: Botrytis as a case study. *Journal of Chemical Biology*, *7*(4), 133–141.

Hardoim, P. R., van Overbeek, L. S., Berg, G., Pirttilä, A. M., Compant, S., Campisano, A., Döring, M., & Sessitsch, A. (2015). The Hidden World within Plants: Ecological and Evolutionary Considerations for Defining Functioning of Microbial Endophytes. *Microbiology and Molecular Biology Reviews: MMBR*, *79*(3), 293–320.

Hardoim, P. R., van Overbeek, L. S., & van Elsas, J. D. (2008). Properties of bacterial endophytes and their proposed role in plant growth. *Trends in Microbiology*, *16*(10), 463–471.

Huson, D. H., Auch, A. F., Qi, J., & Schuster, S. C. (2007). MEGAN analysis of metagenomic data. *Genome Research*, *17*(3), 377–386.

Kenneth Horst, R. (2013). *Westcott's Plant Disease Handbook*. Springer Science & Business Media.

Lane, M. A. (1996). Roles of Natural History Collections. *Annals of the Missouri Botanical Garden. Missouri Botanical Garden*, *83*(4), 536–545.

Langille, M. G. I., Zaneveld, J., Caporaso, J. G., McDonald, D., Knights, D., Reyes, J. A., Clemente, J. C., Burkepile, D. E., Vega Thurber, R. L., Knight, R., Beiko, R. G., & Huttenhower, C. (2013). Predictive functional profiling of microbial communities using 16S rRNA marker gene sequences. *Nature Biotechnology*, *31*(9), 814–821.

Lang, P. L. M., Willems, F. M., Scheepens, J. F., Burbano, H. A., & Bossdorf, O. (2019). Using herbaria to study global environmental change. *The New Phytologist*, *221*(1), 110–122.

Lee, M. R., & Hawkes, C. V. (2021). Plant and Soil Drivers of Whole-Plant Microbiomes: Variation in Switchgrass Fungi from Coastal to Mountain Sites. *Phytobiomes Journal*, *5*(1), 69–79.

Lu, J., Rincon, N., Wood, D. E., Breitwieser, F. P., Pockrandt, C., Langmead, B., Salzberg, S. L., & Steinegger, M. (2022). Metagenome analysis using the Kraken software suite. *Nature Protocols*, *17*(12), 2815–2839.

Lundberg, D. S., Lebeis, S. L., Paredes, S. H., Yourstone, S., Gehring, J., Malfatti, S., Tremblay, J., Engelbrektson, A., Kunin, V., Del Rio, T. G., Edgar, R. C., Eickhorst, T., Ley, R. E., Hugenholtz, P., Tringe, S. G., & Dangl, J. L. (2012). Defining the core *Arabidopsis thaliana* root microbiome. *Nature*, *488*(7409), 86–90.

Lüttge, U., & Scarano, F. R. (2004). Ecophysiology. *Brazilian Journal of Botany*, *27*(1), 1–10.

Mallick, H., Rahnavard, A., McIver, L. J., Ma, S., Zhang, Y., Nguyen, L. H., Tickle, T. L., Weingart, G., Ren, B., Schwager, E. H., Chatterjee, S., Thompson, K. N., Wilkinson, J. E., Subramanian, A., Lu, Y., Waldron, L., Paulson, J. N., Franzosa, E. A., Bravo, H. C., & Huttenhower, C. (2021). Multivariable association discovery in population-scale meta-omics studies. *PLoS Computational Biology*, *17*(11), e1009442.

Martin, M. D., Ho, S. Y. W., Wales, N., Ristaino, J. B., & Gilbert, M. T. P. (2014). Persistence of the mitochondrial lineage responsible for the Irish potato famine in extant new world *Phytophthora infestans*. *Molecular Biology and Evolution*, *31*(6), 1414–1420.

Mayslich, C., Grange, P. A., & Dupin, N. (2021). Cutibacterium acnes as an Opportunistic Pathogen: An Update of Its Virulence-Associated Factors. *Microorganisms*, *9*(2). https://doi.org/10.3390/microorganisms9020303

Moran, N. A., & Sloan, D. B. (2015). The Hologenome Concept: Helpful or Hollow? *PLoS Biology*, *13*(12), e1002311.

Mu, T., & Sun, H. (2017). Progress in Research and Development of Potato Staple Food Processing Technology. *Journal of Applied Glycoscience* , *64*(3), 51–64.

Nelson, G., Sweeney, P., Wallace, L. E., Rabeler, R. K., Allard, D., Brown, H., Carter, J. R., Denslow, M. W., Ellwood, E. R., Germain-Aubrey, C. C., Gilbert, E., Gillespie, E., Goertzen, L. R., Legler, B., Marchant, D. B., Marsico, T. D., Morris, A. B., Murrell, Z., Nazaire, M., … Mast, A. R. (2015). Digitization workflows for flat sheets and packets of plants, algae, and fungi. *Applications in Plant Sciences*, *3*(9). https://doi.org/10.3732/apps.1500065

Pääbo, S., Poinar, H., Serre, D., Jaenicke-Despres, V., Hebler, J., Rohland, N., Kuch, M., Krause, J., Vigilant, L., & Hofreiter, M. (2004). Genetic analyses from ancient DNA. *Annual Review of Genetics*, *38*, 645–679.

Palmer, M. W., Wade, G. L., & Neal, P. (1995). Standards for the Writing of Floras. *Bioscience*, *45*(5), 339–345.

Primack, D., Imbres, C., Primack, R. B., Miller-Rushing, A. J., & Del Tredici, P. (2004). Herbarium specimens demonstrate earlier flowering times in response to warming in Boston. *American Journal of Botany*, *91*(8), 1260–1264.

Przelomska, N. A. S., Armstrong, C. G., & Kistler, L. (2020). Ancient Plant DNA as a Window Into the Cultural Heritage and Biodiversity of Our Food System. *Frontiers in Ecology and Evolution*, *8*. https://doi.org/10.3389/fevo.2020.00074

Pyle, M. M., & Adams, R. P. (1989). In situ preservation of DNA in plant specimens. *Taxon*, *38*(4), 576–581.

Ramírez-Puebla, S. T., Hernández, M. A. R., Guerrero Ruiz, G., Ormeño-Orrillo, E., Martinez-Romero, J. C., Servín-Garcidueñas, L. E., Núñez-de la Mora, A., Amescua-Villela, G., Negrete-Yankelevich, S., & Martínez-Romero, E. (2019). Nodule bacteria from the cultured legume *Phaseolus dumosus* (belonging to the Phaseolus vulgaris cross-inoculation group) with common tropical phenotypic characteristics and symbiovar but distinctive phylogenomic position and chromid. *Systematic and Applied Microbiology*, *42*(3), 373–382.

Reinhold-Hurek, B., Bünger, W., Burbano, C. S., Sabale, M., & Hurek, T. (2015). Roots shaping their microbiome: global hotspots for microbial activity. *Annual Review of Phytopathology*, *53*, 403–424.

Ristaino, J. B. (1998). The importance of archival and herbarium materials in understanding the role of oospores in late blight epidemics of the past. *Phytopathology*, *88*(11), 1120–1130.

Salmensuu, O. (2021). Potato Importance for Development Focusing on Prices. *Journal of Risk and*

Financial Management, *14*(3), 137.

Salter, S. J., Cox, M. J., Turek, E. M., Calus, S. T., Cookson, W. O., Moffatt, M. F., Turner, P., Parkhill, J., Loman, N. J., & Walker, A. W. (2014). Reagent and laboratory contamination can critically impact sequence-based microbiome analyses. *BMC Biology*, *12*, 87.

Sanjenbam, P., Shivaprasad, P. V., & Agashe, D. (2022). Impact of Phyllosphere *Methylobacterium* on Host Rice Landraces. *Microbiology Spectrum*, *10*(4), e0081022.

Särkinen, T., Staats, M., Richardson, J. E., Cowan, R. S., & Bakker, F. T. (2012). How to open the treasure chest? Optimising DNA extraction from herbarium specimens. *PloS One*, *7*(8), e43808.

Savary, S., Bregaglio, S., Willocquet, L., Gustafson, D., Mason D'Croz, D., Sparks, A., Castilla, N., Djurle, A., Allinne, C., Sharma, M., Rossi, V., Amorim, L., Bergamin, A., Yuen, J., Esker, P., McRoberts, N., Avelino, J., Duveiller, E., Koo, J., & Garrett, K. (2017). Crop health and its global impacts on the components of food security. *Food Security*, *9*(2), 311–327.

Schneijderberg, M., Schmitz, L., Cheng, X., Polman, S., Franken, C., Geurts, R., & Bisseling, T. (2018). A genetically and functionally diverse group of non-diazotrophic *Bradyrhizobium spp.* colonizes the root endophytic compartment of Arabidopsis thaliana. *BMC Plant Biology*, *18*(1), 61.

Schubert, M., Ermini, L., Der Sarkissian, C., Jónsson, H., Ginolhac, A., Schaefer, R., Martin, M. D., Fernández, R., Kircher, M., McCue, M., Willerslev, E., & Orlando, L. (2014). Characterization of ancient and modern genomes by SNP detection and phylogenomic and metagenomic analysis using PALEOMIX. *Nature Protocols*, *9*(5), 1056–1082.

Schubert, M., Lindgreen, S., & Orlando, L. (2016). AdapterRemoval v2: rapid adapter trimming, identification, and read merging. *BMC Research Notes*, *9*, 88.

Sharma, T., Kaur, A., Saajan, S., & Thakur, R. (2023.). *Effect of nitrogen on growth and yield of medicinal plants: A review paper*. Retrieved May 1, 2023, from https://ejmcm.com/article_4981_a9472d567dec0dbdecc7c43d21743a18.pdf

Shepherd, L., & Perrie, L. (2014). Genetic analyses of herbarium material: Is more care required? *Taxon*, *63*(5), 972–973.

Singh, J., Behal, A., Singla, N., Joshi, A., Birbian, N., Singh, S., Bali, V., & Batra, N. (2009). Metagenomics: Concept, methodology, ecological inference and recent advances. *Biotechnology Journal*, *4*(4), 480–494.

Soltis, P. S. (2017). Digitization of herbaria enables novel research. *American Journal of Botany*, *104*(9), 1281–1284.

Sprague, T. A., & Nelmes, E. (1931). The Herbal of Leonhart Fuchs. *Botanical Journal of the Linnean Society. Linnean Society of London*, *48*(325), 545–642.

Staats, M., Cuenca, A., Richardson, J. E., Vrielink-van Ginkel, R., Petersen, G., Seberg, O., & Bakker, F. T. (2011). DNA damage in plant herbarium tissue. *PloS One*, *6*(12), e28448.

Staley, J. T., & Konopka, A. (1985). Measurement of in situ activities of non-photosynthetic microorganisms in aquatic and terrestrial habitats. *Annual Review of Microbiology*, *39*, 321–346.

Sumbula, V., Kurian, P. S., Girija, D., & Cherian, K. A. (2020). Metagenomic analysis of bacterial diversity on tomato (*Solanum lycopersicum* L.) leaves. *International Journal of Current Microbiology and Applied Sciences*, *9*(10), 2162–2173.

Tessler, M., Neumann, J. S., Afshinnekoo, E., Pineda, M., Hersch, R., Velho, L. F. M., Segovia, B. T., Lansac-Toha, F. A., Lemke, M., DeSalle, R., Mason, C. E., & Brugler, M. R. (2017). Large-scale differences in microbial biodiversity discovery between 16S amplicon and shotgun sequencing. *Scientific Reports*, *7*(1), 6589.

Turner, T. R., James, E. K., & Poole, P. S. (2013). The plant microbiome. *Genome Biology*, *14*(6), 209.

Vandenkoornhuyse, P., Quaiser, A., Duhamel, M., Le Van, A., & Dufresne, A. (2015). The importance of the microbiome of the plant holobiont. *The New Phytologist*, *206*(4), 1196–1206.

van Dijk, L. J. A., Ehrlén, J., & Tack, A. J. M. (2021). Direct and insect-mediated effects of pathogens on plant growth and fitness. *The Journal of Ecology*, *109*(7), 2769–2779.

Vats, S. (2015). Herbicides: History, Classification and Genetic Manipulation of Plants for Herbicide Resistance. In E. Lichtfouse (Ed.), *Sustainable Agriculture Reviews: Volume 15* (pp. 153–192). Springer International Publishing.

Velsko, I. M., Frantz, L. A. F., Herbig, A., Larson, G., & Warinner, C. (2018). Selection of Appropriate Metagenome Taxonomic Classifiers for Ancient Microbiome Research. *mSystems*, *3*(4). https://doi.org/10.1128/mSystems.00080-18

Vorholt, J. A. (2012). Microbial life in the Phyllosphere. *Nature Reviews. Microbiology*, *10*(12), 828–840.

Weiß, C. L., Schuenemann, V. J., Devos, J., Shirsekar, G., Reiter, E., Gould, B. A., Stinchcombe, J. R., Krause, J., & Burbano, H. A. (2016). Temporal patterns of damage and decay kinetics of DNA retrieved from plant herbarium specimens. *Royal Society Open Science*, *3*(6), 160239.

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*, *8*, 1318.

Wijesinha-Bettoni, R., & Mouillé, B. (2019). The Contribution of Potatoes to Global Food Security, Nutrition and Healthy Diets. *American Journal of Potato Research: An Official Publication of*

the Potato Association of America, *96*(2), 139–149.

Willems, A., & Gillis, M. (2015). Acidovorax. In *Bergey's Manual of Systematics of Archaea and Bacteria* (pp. 1–16). Wiley. https://doi.org/10.1002/9781118960608.gbm00943

Willis, C. G., Ellwood, E. R., Primack, R. B., Davis, C. C., Pearson, K. D., Gallinat, A. S., Yost, J. M., Nelson, G., Mazer, S. J., Rossington, N. L., Sparks, T. H., & Soltis, P. S. (2017). Old Plants, New Tricks: Phenological Research Using Herbarium Specimens. *Trends in Ecology & Evolution*, *32*(7), 531–546.

Winter, S., Meißner, R., Greve, C., Ben Hamadou, A., Horin, P., Prost, S., & Burger, P. A. (2023). A chromosome-scale high-contiguity genome assembly of the cheetah (Acinonyx jubatus). *The Journal of Heredity*. https://doi.org/10.1093/jhered/esad015

Wood, D. E., & Salzberg, S. L. (2014). Kraken: ultrafast metagenomic sequence classification using exact alignments. *Genome Biology*, *15*(3), R46.

Yadav, S. S. (2020). Herbarium: Historical account, significance, preparation techniques and management issues. *Plant Archives 20(1),* 2915-2926.

Yeoh, Y. K., Dennis, P. G., Paungfoo-Lonhienne, C., Weber, L., Brackin, R., Ragan, M. A., Schmidt, S., & Hugenholtz, P. (2017). Evolutionary conservation of a core root microbiome across plant phyla along a tropical soil chronosequence. *Nature Communications*, *8*(1), 215.

Yoshida, K., Burbano, H. A., Krause, J., Thines, M., Weigel, D., & Kamoun, S. (2014). Mining herbaria for plant pathogen genomes: back to the future. *PLoS Pathogens*, *10*(4), e1004028.

Zedane, L., Hong-Wa, C., Murienne, J., Jeziorski, C., Baldwin, B. G., & Besnard, G. (2015). Museomics illuminate the history of an extinct, paleoendemic plant lineage (Hesperelaea, Oleaceae) known from an 1875 collection from Guadalupe Island, Mexico. *Biological Journal of the Linnean Society. Linnean Society of London*, *117*(1), 44–57.

Zotz, G., & Hietz, P. (2001). The physiological ecology of vascular epiphytes: current knowledge, open questions. *Journal of Experimental Botany*, *52*(364), 2067–2078.

SUPPLEMENTARY TABLES

P1 - Silica-dried
P2 - After storing in the herbarium (cardboard changed)
P3 - After storing in the herbarium (cardboard not changed)
P4 - Directly after drying in plant press (cardboard changed)
P5 - Directly after drying in plant press (cardboard not changed)
P6 - Freshly collected, frozen at -20 degree celsius

Supplementary table (ST.) 1. Taxa that shows significant difference in abundance in leaf infected with *Phytophthora infestans* between two preservation methods using MaAslin 2.0. S/N is the serial numbers assigned to each taxa. Negative coefficient means species was at increased abundance in the former preservation method while positive coefficient means species was at increased abundance in the latter preservation method. 3 s.f.g means three significant figures. q-value is the false discovery rates at < 0.05

S/N	Taxa	Coefficient (3 s.f.g)	q- value	Species abundance
P1 and P2 (to compare fast and slow drying, the effect of preserving in herbarium)				
1	*Methylobacterium sp.* 17Sr1.1	-0.2763756	1.114799e-06	High in P1
2	*Methylorubrum sp. B1.46*	0.1169114	8.545745e-05	High in P2
3	*Sphingomonas morindae*	0.2070247	7.147694e-04	High in P2
2 out of 3 microbes are higher in abundance in herbarium specimens compared to fresh				
P2 and P6 (After storing in herbarium vs fresh)				
1	*Sphingomonadaceae bacterium OTU29MARTA1*	1.2510555	1.758272e-09	High in P6
2	*Sphingomonas sp. LM7*	0.7481917	1.758272e-09	High in P6
3	*Sphingomonas sanxanigenens*	0.9693739	5.377931e-09	High in P6
4	*Hymenobacter sedentarius*	0.9693739	1.363647e-08	High in P6
5	*Marmoricola scoriae*	0.9693739	3.527227e-05	High in P6
6	*Methylorubrum sp. B1.46*	0.9693739	8.321645e-03	High in P6
7	*Sphingomonas panacis*	0.7059952	1.527320e-02	High in P6
8	*Sphingomonas sp. HMP6*	0.6055478	1.698671e-02	High in P6
9	*Sphingomonas sp. RMG20*	0.7662427	3.997756e-02	High in P6
P1 & P6 (Silica dried vs freshly frozen)				

1	*Sphingomonas panacis*	0.6277792	0.008405928	High in P6
P2 & P4 (after storing in herbarium vs directly after drying with press)				
1	*Methylobacterium sp. 17Sr1.1*	-0.1180619	2.999501e-06	High in P2
2	*Tilletiopsis washingtonensis*	-0.3487889	2.999501e-06	High in P2
3	*Curtobacterium sp. BH.2.1.1*	-0.7190065	1.000001e-04	High in P2

ST. 2. Taxa that shows significant difference in abundance in uninfected leaf tissue between two preservation methods using MaAslin 2.0. S/N is the serial numbers assigned to each taxa. 3 s.f.g means three significant figures. Negative coefficient means species was at increased abundance in uninfected samples while positive coefficient means species was at increased abundance infected samples. q-value is the false discovery rates at < 0.05

No statistically significant taxa (No comparisons have association)

ST. 3. Taxa that show significant difference in abundance between leaf infected with *Phytophthora infestans* and uninfected leaf tissues using MaAslin 2.0. S/N is the serial numbers assigned to each taxa. 3 s.f.g means three significant figures. Negative coefficient means species was at increased abundance in uninfected samples while positive coefficient means species was at increased abundance infected samples. q-value is the false discovery rates at < 0.05

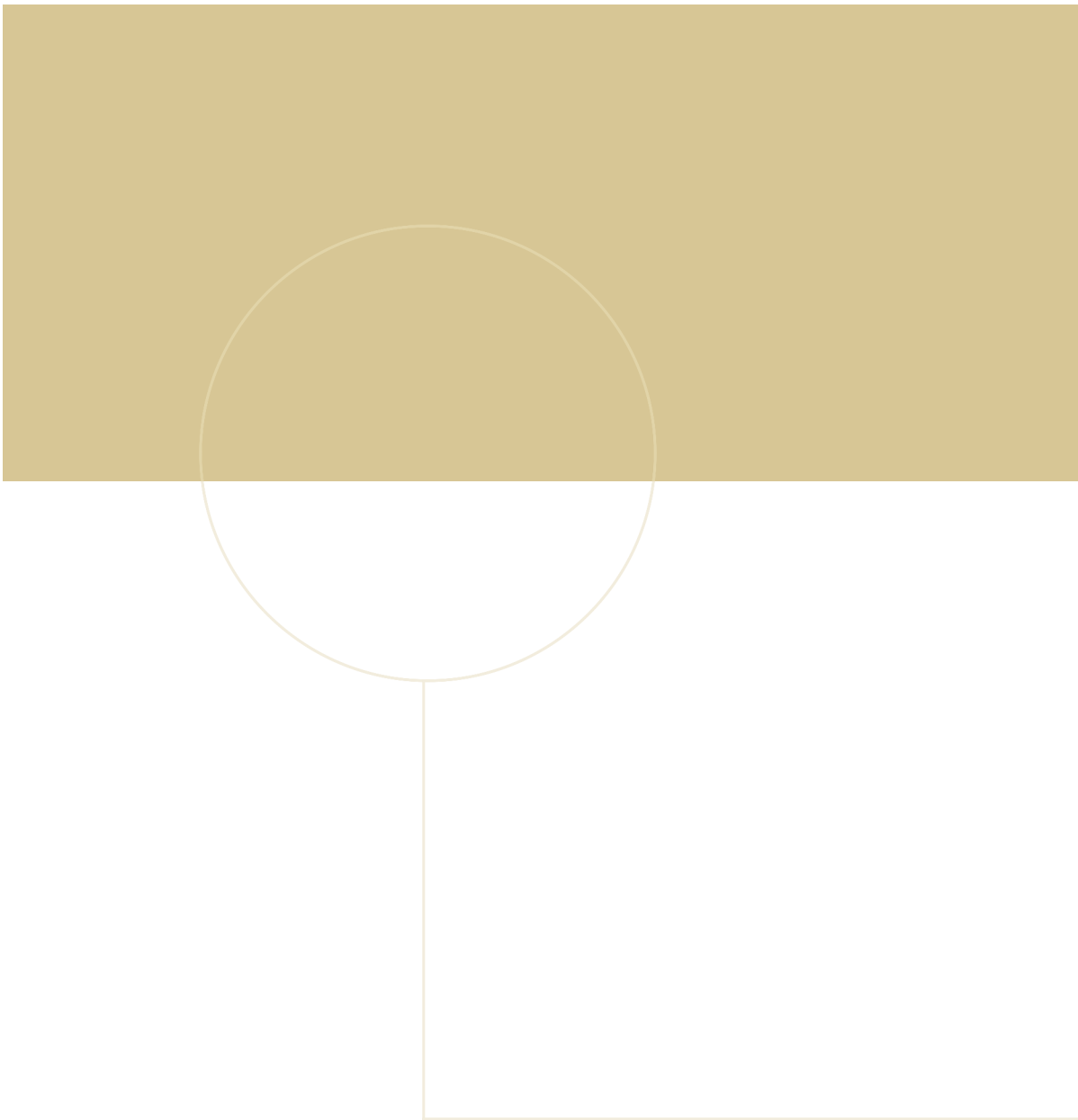
S/N	**Taxa**	**Coefficient (3 s.f.g)**	**q- value**	**Species abundance**
1	*Xylophilus rhododendri*	0.810752766	7.54899E-08	High in uninfected specimens
2	*Phytophthora parasitica*	6.994546929	0.028196194	High in infected specimens
3	*Bacillus licheniformis*	2.29652021	0.028196194	High in infected specimens
4	*Methylobacterium spp*. AMS5	-1.06721	0.0281961940	High in uninfected samples

ST. 4. Taxa that shows significant difference in abundance in root tissues between two preservation methods using MaAslin 2.0. Negative coefficient means species was at increased abundance in the former preservation method while positive coefficient means species was at increased abundance in the latter preservation method. S/N is the serial numbers assigned to each taxa. 3 s.f.g means three significant figures. q-value is the false discovery rates at < 0.05.

S/N	Taxa	Coefficient (3 s.f.g)	q- value	Species abundance
P1 and P2 (to compare fast and slow drying, effect of preserving herbarium)				
1	*Pedococcus dokdonensis*	-0.82245867	1.598847e-08	High in P1
2	*Actinoplanes sp.* NBRC 14428	-0.44167368	2.074483e-06	High in P1
3	*Saccharopolyspora erythrae*	0.05760993	2.074483e-06	High in P2
4	*Bradyrhizobiaceae bacterium*	-0.21219177	2.074483e-06	High in P1
5	*Granulicella* sp.WH15	0.08748534	2.165455e-06	High in P2
6	*Mollisia scopiformis*	-0.62323896	3.403764e-06	High in P1
7	*Streptomyces sp*. SN.593	0.19357762	3.403764e-06	High in P2
8	*Bradyrhizobium sp*. 4	-0.52875530	3.403764e-06	High in P1
9	*Terriglobus roseus*	-0.08759340	9.774567e-06	High in P1
10	*Sphingomonas psychrotolerans*	-0.93724645	2.768681e-05	High in P1
11	*Bacillus mycoides*	-0.36211220	8.054878e-05	High in P1
12	*Hyphomicrobium sp.* MC1	0.03492272	9.550386e-05	High in P2
13	*Labrys sp.* KNU.23	-0.02457972	9.987153e-05	High in P1
14	*Acidovorax avenae*	0.02894196	1.076122e-04	High in P2
15	*Nordella sp. HKS 07*	-0.16497890	6.537984e-04	High in P1
16	*Bradyrhizobium zhanjiangense*	-0.05778333	3.024533e-03	High in P1
17	*Caulobacter soli*	-0.70468404	3.903399e-03	High in P1
18	*Kribbella flavida*	-0.03196026	6.111259e-03	High in P1
P2 & P6 (Herbarium vs freshly frozen)				
1	*Acidobacterium capsulatum*	0.71744129	2.403929e-06	High in P6
2	*Pseudolysobacter antarcticus*	0.13208600	2.825010e-05	High in P6
3	*Mycolicibacterium aubagnense*	0.18301442	3.898839e-05	High in P6
4	*Bradyrhizobium sp. 4*	-0.16891388	5.025562e-05	High in P2

5	*Massilia forsythiae*	-0.04054495	1.050868e-04	High in P2
6	*Xanthobacter autotrophicus*	0.48220882	2.446210e-04	High in P6
7	*Alicycliphilus denitrificans*	-0.17767886	9.663548e-04	High in P6
8	*Microbacterium azadirachtae*	0.56737399	2.481132e-03	High in P6
P1 & P6 (Silica dried vs freshly frozen)				
1	*Bradyrhizobium sp. 4*	-0.27429974	6.280490e-08	High in P1
2	*Phytohabitans suffuscus*	0.06037671	2.297454e-06	High in P6
3	*Stutzerimonas stutzeri*	0.18633563	2.297454e-06	High in P6
4	*Sphingomonas panacis*	0.12230819	4.856658e-06	High in P6
5	*Streptomyces sp. SN.593*	0.1625939	7.125086e-06	High in P6
6	*Streptacidiphilus sp. P02.A3a*	-0.28522196	9.959938e-06	High in P1
7	*Mucilaginibacter xinganensis*	-0.42610879	1.362502e-05	High in P1
8	*Nocardia brasiliensis*	-0.17196812	1.946514e-05	High in P1
9	*Corallococcus coralloides*	0.10321003	1.001938e-04	High in P6
10	*Arthrobacter sp. FW306.2.2C.D06B*	-0.26914750	2.932879e-04	High in P1
11	*Asticcacaulis excentricus*	0.55095785	2.932879e-04	High in P6
12	*Frateuria sp. 5GH9.34*	0.38770044	3.783502e-04	High in P6
13	*Collimonas arenae*	-0.18790458	3.797614e-04	High in P1
14	*Streptomyces sp. So13.3*	0.15130362	6.588541e-04	High in P6
15	*Variovorax sp. WDL1*	-0.08698365	9.303951e-04	High in P1
16	*Leifsonia sp. PS1209*	-0.03547171	5.052510e-03	High in P1
17	*Caulobacter mirabilis*	0.30480394	4.917212e-02	High in P6
P2 & P4 (after storing in herbarium vs directly after drying with press)				
1	*Streptomyces sp. CB01881*	-0.14268616	8.987204e-08	High in P2
2	*Saccharothrix sp. 6.C*	-0.31170490	1.358026e-07	High in P2
3	*Aquabacterium sp. J223*	-0.60915559	1.468039e-07	High in P2
4	*Lacunisphaera_limnophila*	0.18759057	2.149925e-07	High in P4
5	*Streptomyces_sp._SN.593*	-0.04915779	2.149925e-07	High in P2
6	*Bradyrhizobium_sp._C.145*	0.29592987	2.366133e-06	High in P4

7	*Amycolatopsis_pretoriensis*	0.30012849	4.694164e-06	High in P4
8	*Chondromyces_crocatus*	-0.17830487	4.694164e-06	High in P2
9	*Bradyrhizobium_sp._144S4*	-0.14460530	4.694164e-06	High in P2
10	*Marmoricola_scoriae*	0.63864345	5.205428e-06	High in P4
11	*Ralstonia_solanacearum*	-0.15242132	6.253790e-06	High in P2
12	*Schlegelella_thermodepolymerans*	-0.24492213	7.580678e-06	High in P2
13	*Bradyrhizobium_sp._CCBAU_51765*	0.15702256	7.580678e-06	High in P4
14	*Kribbella flavida*	0.13720923	1.841289e-05	High in P4
15	*Deltaproteobacteria bacterium*	0.30679961	1.841289e-05	High in P4
16	*Mycolicibacterium rhodesiae*	0.06381610	1.841289e-05	High in P4
17	*Cupriavidus necator*	0.05271528	4.931463e-05	High in P4
18	*Mucilaginibacter gotjawali*	-0.24614979	1.002205e-04	High in P2
19	*Methylocella tundrae*	0.36126835	1.186314e-04	High in P4
20	*Rhodopseudomonas sp._SK50.23*	0.18930542	2.082673e-04	High in P4
21	*Massilia forsythiae*	-0.28104767	2.342277e-04	High in P2
22	*Paraburkholderia aromaticivorans*	0.27575043	2.360463e-04	High in P4
23	*Massilia sp. HC52*	0.07826113	2.556328e-04	High in P4
24	*Paraburkholderia phytofirmans*	0.22626062	2.811076e-04	High in P4
25	*Bradyrhizobium zhanjiangense*	-0.19738014	2.914110e-04	High in P2
26	*Amycolatopsis sp. CA.230715*	0.10492361	1.016282e-03	High in P4
27	*Actinomadur sp. WMMB_499*	0.01644033	1.016282e-03	High in P4
28	*Nocardioides anomalus*	0.31373851	2.571274e-0	High in P4



NTNU

Norwegian University of
Science and Technology