



UNIVERSITÀ DEGLI STUDI DI PALERMO

Dottorato di Ricerca in Scienze Agrarie – Indirizzo “Gestione Fitosanitaria Ecocompatibile in Ambienti Agro-Forestali e Urbani”

Dipartimento di Agraria – Università Mediterranea di Reggio Calabria (sede consorziata)
Agr/12 – Patologia Vegetale

Molecular analysis of the fungal community associated with phyllosphere and carposphere of fruit crops

IL DOTTORE
Ahmed ABDELFAH

IL COORDINATORE
Prof. Stefano COLAZZA

IL TUTOR
Prof. Leonardo SCHENA

CO-TUTOR
Dr. Anna Maria D'ONGHIA
Dr. Michael WISNIEWSKI

ACKNOWLEDGEMENTS

I appreciate everyone's contributions of time, ideas, and funding to make my PhD possible. First I want to thank my advisor Leonardo Schena. He has taught me so many things in science and in life. It was a great honor to work with him during this PhD. It has been an amazing experience, not only for his tremendous academic support, but also for giving me so many wonderful opportunities.

I am greatly thankful to Dr. Michael Wisniewski for giving the opportunity to work in his lab at the USDA in West Virginia. Michael was a great advisor, I learnt so many things from working with him and it was a great honor to know him and his amazing family.

I am grateful to the University of Reggio Calabria for giving me the chance to work in their laboratories during my PhD. I'm sincerely grateful to the past and present group members of the Reggio Calabria University that I have had the pleasure to work with or alongside: Demetrio Serra, Antonio Biasi, Marisabel Prigigallo, Sonia Pangallo, David Ruano Rosa, Antonino Malacrinò, Orlando Campolo, Vincenzo Palmeri, and especially Saveria Mosca for being a great coworker and wonderful friend. Thanks to them, I had an incredible time in their amazing city. I would like also to thank the University of Palermo, and the coordinator of the PhD program, Prof. Stefano Colazza, for giving me this wonderful opportunity. Thanks also goes to the Mediterranean Agronomic Institute of Bari (MAIB) especially Anna Maria D'Onghia and Thaer Yaseen for their support.

I will be always thankful to the people in Shepherdstown for their hospitality, especially Katie Allen, Caitlin McAteer, Bob Keel and Richard Womeldorf for giving such a wonderful time during my visit to their beautiful town.

I would like to end this acknowledgment with thanking the most special people in my life, my family, for their limitless love and support all my life, keeping me full of enthusiasm and happiness. Same goes to a wonderful and beautiful person, Melissa Si Ammour, thank you for your support and encouragement.

Table of Contents

Chapter I. Introduction.....	4
Metabarcoding	4
1. ITS as a barcode gene for fungi.....	6
2. Challenges with the ITS as a DNA barcode.....	7
Metabarcoding Workflow	8
1. Sampling.....	8
2. DNA extraction and library preparations	9
3. Sequencing technologies	10
4. Quality and chimera filtering	13
5. Operational Taxonomic Units (OTUs).....	13
6. Normalizations	15
7. Taxonomy assignment.....	16
8. Methods in calculating microbial diversity	17
9. Statistical comparisons	20
Scope of the thesis.....	20
References.....	21
Chapter II. Metabarcoding analysis of fungal diversity in the phyllosphere and carposphere of olive (<i>Olea europaea</i>)	28
Abstract	28
Introduction.....	28
Materials and Methods.....	30
Sampling and DNA extractions	30
Fungal DNA amplification	31
Data analysis and statistics	32
Identification of fungal species.....	33
Results.....	34

Fungal diversity and richness	34
Olive fungal community structure	37
Identification of fungal species	40
Discussion	42
Conclusion	48
References	49
Chapter III. Metagenomic analysis of fungal diversity on strawberry plants and effect of management practices on the fungal community structure of aerial organs	54
Abstract	54
Introduction	54
Materials and Methods	56
Sampling and DNA extractions	56
Fungal DNA amplification and amplicon library preparation	57
Data analysis and statistics	58
Identification of fungal taxa	59
Results	59
General sequencing analysis	59
Alpha diversity and richness	60
Beta diversity analysis	61
Strawberry fungal community structure and identified taxa	63
Discussion	69
Conclusion	72
References	73
Chapter IV. Analysis of fungal diversity in orange (<i>Citrus sinensis</i>) leaves with and without greasy spot disease symptoms	77
Abstract	77
Introduction	77
Materials and Methods	80
Data analysis and statistics	81

Identification of fungal taxa.....	82
Results and Discussion.....	83
Conclusion	93
References.....	94
Chapter V. Metagenomic analysis of fungal populations in harvested organic and conventional apples and the impact of hot water treatments on fungal diversity.....	97
Abstract	97
Introduction.....	97
Materials and Methods.....	99
Experimental design	99
Hot water treatment	100
Sampling of apple tissues	100
DNA extraction, amplification and sequencing.....	100
Data Analysis.....	101
Results and Discussion.....	102
Impact of management practices (organic vs. conventional) on the fungal communities found on apple (“practices” effect)	106
Differences in fungal communities between apple stem end (SE) and calyx end (CE) samples (“location” effect);	107
Effect of hot water treatment on fungal communities (“treatment” effect)	109
Effect of time following hot water treatment on the fungal community (“time” effect)	109
Apple fungal communities.....	109
Identification of fungal species.....	115
Conclusion	122
References.....	123

Chapter I. Introduction

Microorganisms, such as fungi and bacteria, are distributed globally and can colonize any kind of matrix, including soil, water, air, plants, and animals. There are in total about 10^{30} prokaryotic cells on earth that play an essential role in almost all the biological processes occurring in any kind of ecosystem, including the human body. There are around 10^9 prokaryotic cells in a gram of soil, and 10^6 in every milliliter of sea or lake-water. Even the human body is believed to house 10 times more bacteria than human cells (Whitman et al. 1998, Sogin et al. 2006, Turnbaugh et al. 2007, Jorgensen and Fath 2008, Oakley et al. 2008). The diversity of fungi on earth is estimated to be about 1.5 million species (Hawksworth 2001). Identifying and characterizing microorganisms are important to understanding their role in any ecosystem. The fact that less than 1% of microbial organisms can be cultivated (Kellenberger 2001) indicates, however, that traditional culture-dependent microbial studies have greatly limited our understanding of the composition of microbial communities. In recent years, the development of next-generation sequencing technology, has made it possible to generate millions of sequences directly from any environment, revealing completely hidden microbial worlds (Pace 1997, Moreira and López-García 2002). While ribosomal RNA (16S rRNA) is the most widely used target gene for inferring phylogenetic relationships among prokaryotes (Lane et al. 1985), the internal transcribed spacer (ITS) of the rDNA represents the primary target gene in fungi utilized to discern taxonomic relationships, as well as for identification and detection purposes (Schoch et al. 2012). The genes are currently utilized in large-scale metagenomic projects to investigate various aspects of the composition of microbial communities, *e.g.* the Human Microbiome Project (<http://commonfund.nih.gov/hmp>), the International Census of Marine Microbes (<http://icomm.mbl.edu>), and the Earth Microbiome Project (<http://www.earthmicrobiome.org>). A multitude of 16S rRNA and ITS sequence datasets have been already generated and many others will be generated in the next future through individual and community projects.

Metabarcoding

DNA metabarcoding, also referred to as **targeted metagenomics** or **amplicon metagenomics** is defined as “the automated identification of multiple species from a single

bulk sample containing entire organisms or from a single environmental sample containing degraded DNA” (Taberlet et al. 2012). DNA metabarcoding relies on the amplification of a barcode gene from an environmental sample, which enables the simultaneous identification of many organisms. Environmental sample or environmental DNA (eDNA) refers to any DNA collected from soil, water, plant, air, or any other environment. Metabarcoding, however, should not be confused with “Metagenomics” described by Riesenfeld et al. (2004) as the functional and sequence-based analysis of the collective microbial genomes contained in an environmental sample. DNA metabarcoding does not provide genetic information beyond the genes that are amplified. Similarly, the terms microbiome and microbiota should not be confused. “Microbiota” should be used to refer to a study of the microbial taxa in a given environment, while “microbiome” is defined as the catalog of these microbes and their genes (Ursell et al. 2012).

The idea and the term “DNA barcoding” were first proposed by Hebert et al. (2003). The term “DNA Metabarcoding” was subsequently introduced to designate high-throughput multispecies (or higher-level taxon) identification, using both total and degraded DNA extracted from an environmental sample (Taberlet et al. 2012). Standardized criteria were also established for a DNA barcode indicating it must satisfy three aspects: (i) contain significant species-level genetic variability and divergence, (ii) possess conserved flanking sites for developing universal PCR primers for wide taxonomic application, and (iii) have a short sequence length so as to facilitate current capabilities of DNA extraction and amplification (Kress and Erickson 2008).

The first DNA barcodes to be proposed were for animal identification systems and based on the mitochondrial gene, cytochrome *c* oxidase I (*cox1*), (Hebert et al. 2003). This was followed by the development of DNA barcodes for plants (Hollingsworth et al. 2009). In contrast to higher plants and animals, that can be quite easily identified with conventional methods based on morphological or phenotypic methods, there is a stronger need to develop a universal barcode-based method to identify microorganisms such as bacteria and fungi since their identification by other means is often problematic. Most fungi and bacteria need to be cultured to be identified and this is not possible with many to most fungi and bacteria (Kellenberger 2001). Conventional techniques also fail in their ability to identify related species which are characterized by nearly identical morphological features. Finally, even when culturing is possible, the morphological identification of microorganisms is commonly based on a particular life or sexual stage,

organ or gender that may be difficult to obtain, making the analysis very time consuming and requiring a high level of expertise.

Culture-independent methods such as metabarcoding, have enabled us to answer the question “Who is there?”, providing a fundamental breakthrough in allowing researchers to compare microbial communities across environments (Ursell et al. 2012). In addition, the minimum information about a marker gene sequence (MIMARKS) standard and the minimum information about any (x) sequence (MIxS) have been developed by the Genomic Standards Consortium (Yilmaz et al. 2011), which is now considered as the standard for metadata by the International Nucleotide Sequence Database Consortium (INSDC), including GenBank, EBI and DNA Databank of Japan. Furthermore, the availability of unified DNA barcode databases for the identification of fungi and bacteria (such as UNITE, Greengene and RDP), has allowed research to be carried out in a more unified and accurate manner.

Metabarcoding analyses have been increasingly utilized in recent years to investigate microbial diversity in a number of different environments and, in most cases, these studies have revealed a vast, previously unknown level of microbial biodiversity that was entirely missed by conventional culture-based methods. However, a large number of these studies have focused on soils and rhizosphere (Buee et al. 2009, Stockinger et al. 2010, Orgiazzi et al. 2013). In contrast, the plant canopy has not been widely addressed, and relatively few studies have focused on phyllosphere endophytes (Unterseher et al. 2013). To the best of our knowledge, fungal biodiversity has been recently investigated only in tomato, grape leaves, and balsam poplar (Ottesen et al. 2013, Pinto et al. 2014).

1. ITS as a barcode gene for fungi

The internal transcribed spacer (ITS) regions of ribosomal DNA (rDNA), are the regions’ most commonly sequenced for studying fungi and oomycetes and have been widely utilized for phylogenetic studies and the development of diagnostic assays (Begerow et al. 2010, Schena et al. 2013). Although, these regions are now accepted as formal DNA barcode genes for molecular identification of fungi (Schoch et al. 2012), they were used for nearly two decades as an ‘unofficial fungal DNA barcode’ in both molecular systematics and ecological studies. Their suitability as a barcode gene is due to their high degree of interspecific variability, conserved flanking primer sites, and multicopy nature. Numerous studies have led to the accumulation of hundreds of thousands of fungal ITS

sequences in different databases (Karsch-Mizrachi et al. 2012). Fungal ITS regions vary in length and range between approximately 450 and 750 base pairs (bp) with some exceptions. The ITS includes the ITS1 and ITS2 regions, separated by the 5.8S gene, and is located between the 18S (SSU) and 28S (LSU) genes in rDNA (Blaalid et al. 2013) (Fig. I-1). The 18S, 5.8S, and 28S genes are highly conserved which has allowed the development of universal PCR primers. The sequences of the ITS1 and ITS2 spacers are highly variable and thus normally provide resolution at the species level, even though closely related species may be difficult to discriminate (Nilsson et al. 2008).



Figure I-1 Illustration of the eukaryotic ribosomal DNA (rDNA), including the ITS1 and ITS2 regions, separated by the 5.8S gene and located between the 18S (SSU) and 28S (LSU) genes

2. Challenges with the ITS as a DNA barcode

Although ITS regions are the barcode genes currently accepted for fungi, they still have some disadvantages. Uncertainties about the suitability of the ITS for phylogenetic comparisons are mainly due to the presence of hundreds of tandem copies of rRNA cistrons in a typical eukaryote genome, also known as “Intragenomic variation” (Coleman 2003). This variation could be in both copy number and the sequence itself. Copy number is strongly correlated with genome size, and can vary between 30 to 30,000 copies in most eukaryotes, including fungi. A heterogeneity of approximately 1.5% has been estimated in the ITS of some fungal species (Vilgalys and Gonzalez 1990, Horton 2002, Prokopowich et al. 2003, Izzo et al. 2005, Ganley and Kobayashi 2007, Nilsson et al. 2008, Kovács et al. 2011). Significant variation among ITS sequences in an organism, however, is only common within hybrid individuals, *i.e.* diploid or polyploids of disparate parents (Buckler et al. 1997). Indeed, it has been demonstrated that a process called “concerted evolution”, rapidly homogenizes the many copies of this multigene family, and as a consequence the ITS can be treated as a single gene (Hershkovitz et al. 1999).

At the genus and species level, most of the phylogenetic studies in diverse eukaryotes, that have utilized the ITS, have reported this region to be more informative than alternative genes (Coleman 2003). In some cases, however, insufficient genetic variability can limit the use of ITS regions as barcode genes where the discrimination of closely related taxa can be very difficult or impossible.

Metabarcoding Workflow

When conducting a metagenomic project, several steps need be followed, from the design of the experiment and sampling, to the analysis and interpretation of the results (Fig. I-2). In addition, multi-step, bioinformatic analytical workflows need to be pre-defined in order to access most of the information contained in such large datasets. In this section these steps are described in a pragmatic order.

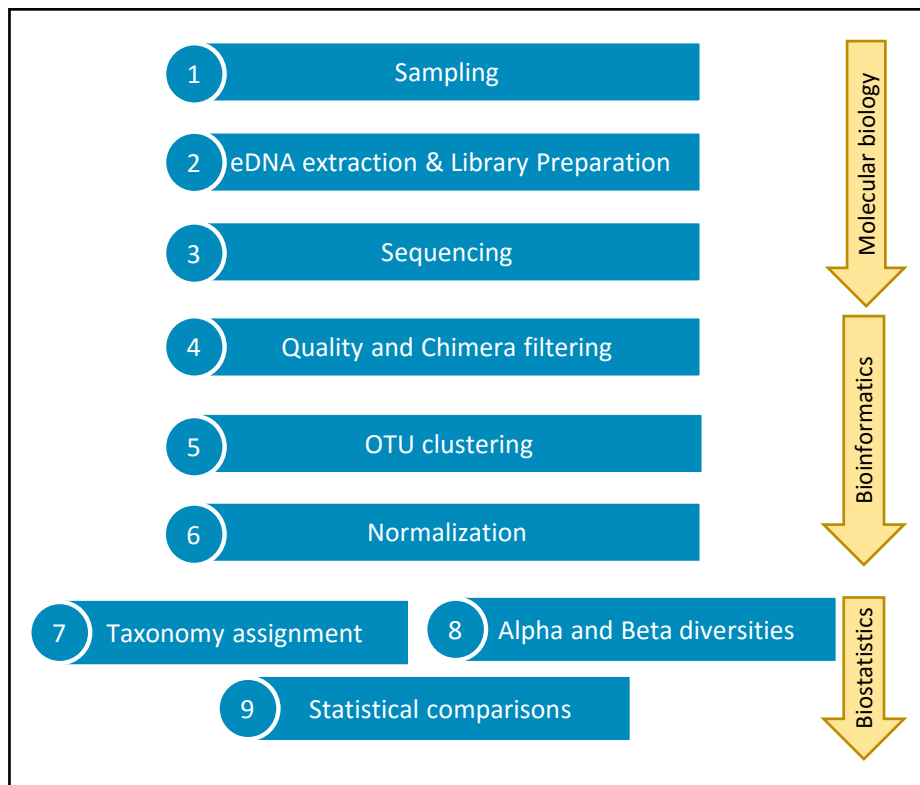


Figure I-2 Common steps used in metabarcoding projects. The arrows on the right indicate the type of analysis used, including: Molecular biology, Bioinformatics and Biostatistics analytical strategies.

1. Sampling

A sampling scheme relevant to the objectives of the metagenomic study is the first and one of the most important considerations (Thomas et al. 2012). A “good” sampling scheme is usually the result of a good experimental design. In order to have statistically robust conclusions, the number of collected samples should represent the environment of interest and be relevant to the study being conducted. Additionally, since environmental samples are usually complex, and the main objective of metabarcoding studies is to take a snapshot of who is there, it is important to arrest the microbial communities at the time of sampling so that further changes do not occur prior to DNA extraction. This could be done by

freezing the samples in liquid nitrogen at the sampling site. Moreover, since metabarcoding can theoretically detect all the targeted organisms that live everywhere, preventing contamination of the samples should be a logical precaution. This includes never handling samples with bare hands, mixing samples of different origins together, etc.

2. DNA extraction and library preparations

Many factors need to be taken into account when conducting an amplicon-based metagenomic study since there are a host of things that can greatly influence the results obtained and the subsequent interpretation of the results. These factors include, the DNA extraction method and the DNA template dilution, the choice of PCR primers and conditions, the number of cycles, and the fidelity of the enzyme used in the amplification process (Hong et al. 2009, Lindahl et al. 2013). Several DNA extraction methods are available as commercial kits or as published laboratory protocols. The choice of a method is mainly based on the type of sample and the targeted organisms. For example, the main concern when extracting DNA from water is the quantity, whereas in soil and plant samples, the quality of the DNA, in terms of the presence of PCR inhibitors *e.g.* humic acid, phenols and polysaccharides, is an important aspect to consider. Even though, PCR can amplify very low DNA concentrations, it is important that the DNA extraction method used should succeed in the recovery of as much DNA as possible of the targeted organisms. PCR artifacts are known to occur due to a high number of PCR cycles, and DNA polymerase errors. These PCR problems may affect both the diversity richness and community structure that is subsequently obtained, whereas, template dilution problems only affect diversity richness (Qiu et al. 2001, Acinas et al. 2005, Wu et al. 2010). Moreover, the use of high-fidelity DNA polymerases have produced a lower amount of PCR artifacts and resulted in lower taxa richness than when normal fidelity polymerases are used. The choice of primers for the barcode amplification is also crucial. Some of the ITS primers are known to introduce taxonomic biases in fungal samples during PCR. For instance, ITS1-F, ITS1, and ITS5 primers, are biased towards the amplification of basidiomycetes, whereas other primers such as ITS2, ITS3, and ITS4, are biased towards ascomycetes (Anderson et al. 2003). Once the barcode is amplified, the next step is to prepare the amplicons for sequencing.

The emergence of high-throughput sequencing technologies have enabled the processing and sequencing of samples from different origins in a single, pooled sample by multiplexing. Sequencing platforms and technology may change but the principle of

multiplexing is the same. Short sequences, known as molecular identifiers, usually 8 to 10bp, are added to the amplicons of each sample after it has initially amplified with barcode gene primers. The addition of these sequences can be done by either ligation or by a second PCR. In the case of 454 pyrosequencing, the short reads are called multiplex identifiers (MIDs), while in Illumina they are called indexes. After sequencing, molecular identifiers allow the separation of the multiplexed samples into data from each individual sample, a process called demultiplexing.

3. Sequencing technologies

DNA metabarcoding has relied heavily on improvements in DNA sequencing technology to achieve its increasing potential. The major events in the analysis of DNA are: i) the discovery of the DNA structure (Watson and Crick 1953); ii) the advent of Polymerase Chain Reaction (PCR) (Saiki et al. 1985, Mullis et al. 1987), iii) the development of Sanger sequencing (Sanger et al. 1977) and iii) the emergence of next generation sequencing (NGS) technologies. In 1996, the first commercial DNA sequencer, that used capillary electrophoresis rather than a slab gel (the ABI Prism 310), was commercialized. This technology allowed the rapid distribution and widespread use of DNA sequencing in a number of research fields and led to the sequencing of the first human genome. Sanger sequencing technology was the only technology available for a quite long period of time. It enabled, however, the sequencing of only one molecule at a time (either by direct sequencing of PCR products or after cloning the DNA fragments into a bacterial cell). However, the recent advent of NGS technologies, allowing for massively parallel sequencing of DNA, has revolutionized DNA sequencing.

Pyrosequencing

The first massively parallel sequencing method to become commercially available was developed by 454 Life Sciences and was based on pyrosequencing (Margulies et al. 2005). It differs from Sanger sequencing, in that it relies on the detection of pyrophosphate release upon nucleotide incorporation. More specifically, the technique utilizes enzyme cascades and charged coupled device (CCD) luminescence detection capabilities to measure the release of inorganic pyrophosphate with every nucleotide, rather than chain termination with dideoxynucleotides (Ronaghi et al. 1998). This technique allowed shotgun sequencing of whole genomes without cloning in *E. coli* or another host cell. In this technique, the DNA is first enzymatically or mechanically sheared and the fragments are then ligated to linker sequences that permit individual molecules, captured on the

surface of a bead, to be amplified while isolated within an emulsion droplet. Each incorporation event is accompanied by the release of pyrophosphate (PPi) in a quantity equimolar to the amount of incorporated nucleotide. Adenosine triphosphate (ATP) sulfurylase converts PPi to ATP in the presence of adenosine 5' phosphosulfate (APS). This ATP drives the luciferase-mediated conversion of luciferin to oxyluciferin which generates visible light in amounts that are proportional to the amount of ATP. The light produced in the luciferase-catalyzed reaction is detected by a CCD camera and seen as a peak in the raw data output (Pyrogram). The height of each peak (light signal) is proportional to the number of nucleotides incorporated.

Illumina

The second NGS technology to become commercially available in 2006 was Illumina (Solexa) sequencing (Bennett 2004). A key difference between this method and 454 pyrosequencing is that it uses chain-terminating nucleotides. The fluorescent label on the terminating base can be removed to leave an unblocked 3' terminus, making chain termination a reversible process. The method reads each base in a homopolymer run in a separate step and therefore does not produce as many indels as 454 pyrosequencing within a sequencing run. Since the reversible dye terminator nucleotides are not incorporated efficiently, the read length of the Solexa method is less than for 454 pyrosequencing. More base-substitution errors are also observed due to the use of modified polymerase and dye terminator nucleotides. The method sequences clusters of DNA molecules amplified from individual fragments attached randomly on the surface of a flow cell. Presently, the main three instruments produced by Illumina are MiSeq, NextSeq, and HiSeq.

SOLiD

Applied Biosystems also developed a massively parallel sequencer named Supported Oligonucleotide Ligation and Detection system (SOLiD) and made it commercially available in 2008. The technology is based on hybridization-ligation chemistry (Shendure et al. 2005). The sample preparation aspect of this technology, including library preparation and clonal amplification of the target DNA by emulsion PCR (emPCR) on beads, is very similar to 454 pyrosequencing in principle. The size of the beads used for emPCR (1 μm versus 26 μm), however, as well as the array format (random versus ordered) are different. These differences provide the SOLiD technology with the ability to generate a significantly higher density sequencing array (potentially over a few hundred fold higher), as well as more flexibility in terms of sample input format. The

sequence interrogation is done through repeated cycles of hybridization of a mixture of sequencing primers and fluorescently labeled probes, followed by ligation of the sequencing primers and the probes. This is followed by the detection of the fluorescent signals on the probes which encode the bases that are being interrogated. Although it can only provide a short read length of about 25 – 35, it can generate $\approx 2 - 3$ Gb of sequence per run.

Ion Torrent

Ion Torrent is also known also as Ion semiconductor sequencing, pH-mediated sequencing, silicon sequencing, or semiconductor sequencing. The technology is based on "sequencing by synthesis" meaning the detection of hydrogen ions that are generated and released during the polymerization of DNA which does not require any modified nucleotides or optics. When a single species of deoxyribonucleotide triphosphate (dNTP) are added to the template DNA strand, it becomes incorporated into the growing complementary strand. This causes the release of a hydrogen ion that changes the pH of the solution which indicates that a reaction has occurred.

Advantages and disadvantages of different sequencing platforms.

The SOLiD and Illumina HiSeq sequencers generate the largest amount of data per run at the lowest costs per base. However, they generate the shortest reads, particularly SOLiD, and both platforms are most suitable for studies such as transcriptomics. Illumina MiSeq produces up to 25 million reads with a maximum length of 300 bp when using the latest V3 chemistry and half of that number of reads when using older V2 chemistry. The longer read length in combination with the lower read number makes it the best fit for amplicon sequencing projects such as metabarcoding or amplicon metagenomics. It is also very suitable for small scale metagenomics projects or initial sample evaluation prior to deep sequencing on a HiSeq. In comparison to Illumina and SOLiD platforms, 454 pyrosequencing generates longer reads, up to 750bp and a much lower number of total reads. Although 454 has a higher cost when compared to Illumina, the fact that it was available a few years before the newer platforms made it the method of choice for several kinds of projects, including transcriptomics and metagenomic or amplicon sequencing projects. The Ion Torrent PGM sequencer provides up to 5 million reads with a length of 400 bp. Sequencing on Ion instruments is very fast, taking only a couple of hours.

4. Quality and chimera filtering

Filtering sequences based on quality score, read length and errors from amplification and sequencing is an important process to avoid overestimating microbial diversity (Bokulich et al. 2013). This step can be either done separately or combined with other steps such as assembling paired-end reads or demultiplexing sequences (Caporaso et al. 2010, Masella et al. 2012, Zhang et al. 2014). Data generated from most platforms will need to be filtered from chimeric sequences, a type of error introduced during amplification, defined as hybrid products between multiple parent sequences. Many software and packages that identify and eliminate this kind of erroneous reads are available but vary in performance and speed (Edgar et al. 2011, Rognes et al. 2015). In pyrosequencing however, an additional step, called denoising, is required to remove characteristic sequencing errors mostly imprecise signals for longer homopolymer runs (Reeder and Knight 2010).

5. Operational Taxonomic Units (OTUs)

Traditionally, DNA barcode sequences were used to identify species by alignment, phylogeny and/or BLAST searches. However, with the emergence of new sequencing technologies, the millions of produced sequences could not be elaborated manually anymore, even with the help of heavy duty computers. In addition to this issue, it was almost impossible for the traditional methods such as culturing to reach the same progress as the massive sequencing. This resulted in lack of specimens, hence lack of accurate identification particularly at species and subspecies levels. Therefore, there was a need to better classify organisms based on sequencing solely. In this context Blaxter et al. (2005) proposed the concept of Molecular Operational Taxonomic Unit (MOTU) or Operational Taxonomic Unit (OTU) for short. In a broader sense, OTU can be defined as clusters of sequences derived by grouping DNA sequences of a gene or gene fragment at certain similarity cutoff that is defined by the user. The use of OTUs has facilitated the analysis of large data produced by the new sequencing platforms and allowed scientists to avoid worrying about species concept and more importantly, reduced the computational resources necessary for such big data. OTUs can be created by comparing sequences to a reference database, sometimes called “closed reference”. In parallel, another approach clusters reads against one another without any external reference sequence *i.e. de novo* clustering. In this approach, the distance between sequences is used to cluster sequences into OTUs rather than the distance to a reference database. In addition to “closed

reference” and “*de novo*” methods, there is a method called “open reference” that is a combination of the previous methods; sequences are initially compared to a database (closed reference is performed first), but if the sequences are not similar to any reference sequence in the database, a *de novo* method is performed.

There are several tools for clustering sequences or picking OTUs. The most commonly used methods are USEARCH, UCLUST (Edgar 2010), BLAST (Altschul et al. 1990), Swarm (Mahé et al. 2014), Mothur (Schloss et al. 2009) and ESPRIT (Sun et al. 2009). These tools implements different algorithms and vary in their performances (Schloss and Westcott 2011).

The hierarchical approach is one way OTUs could be clustered. In this case, the distance between instances (sequences) is measured first, creating a distance matrix, then OTUs are created based on a threshold of similarity (T) defined by the user. Hierarchical clustering can be subdivided into single-linkage, complete linkage, weighted and unweighted average linkages. Single-linkage (*i.e.*, nearest neighbor algorithm) clustering defines the distance between two clusters as the minimum. It is calculated as $D(c_1, c_2) = \min_{x_1 \in c_1, x_2 \in c_2} D(x_1, x_2)$, where c is cluster, x is sequences and D is the distance. It searches for at least one sequence in two different clusters that are at least T similar. If true, the two clusters are grouped into one. One obvious disadvantage in this method is that it can end up creating long chains of sequences, clustering somehow unrelated sequences together. Unlike single linkage clustering, complete linkage (*i.e.*, furthest neighbor algorithm) method merges clusters of the distance D if the furthest instance is within or equal to the defined threshold (T). It is calculated as $D(c_1, c_2) = \max_{x_1 \in c_1, x_2 \in c_2} D(x_1, x_2)$. In addition, it ensures that all sequences clustered in one group are at least T similar. Arithmetic average clustering instead, calculates the average distance between all sequences in the two clusters, if the average distance is within the defined T, then clusters are grouped into one. The only difference between weighted and unweighted algorithms is that weighted algorithms take into consideration the abundance of the sequences, while unweighted algorithms do not.

Another line of research has developed greedy heuristic-clustering methods. Two well-known methods are CD-HIT (Li and Godzik 2006) and UCLUST (Edgar 2010). Both methods use pairwise sequence alignment and process input sequences sequentially. Given a predefined threshold, an input sequence is either assigned to an existing cluster if the

distance between the sequence and a seed is smaller than the threshold, or becomes a seed otherwise. Heuristic methods are known to reduce the search time and computational complexity at the cost of decreased accuracy. Such heuristics have a number of associated problems, including the tendency to create clusters that are more dissimilar than a specified threshold (Cai and Sun 2011)

6. Normalizations

The interpretation of the data produced by high-throughput sequencing are often normalized to account for the sampling process and to address variability in sampling depths. In other words, if no data normalization is applied, the samples can confound biological differences. This is because samples of lower sequencing depth could fail to detect rare taxa. Highly sequenced samples will thus appear more similar to each other than to shallow sequenced samples because they are scored as sharing the same rare taxa (Weiss et al. 2015). Rarefaction is one of the first and the most used normalization methods. In this method, samples are subsampled to have an equal number of sequences i.e. even depth, to avoid heterogeneity in sampling effort. To ease this process, a rarefaction curve is created, plotting the number of sequences vs. the value of species diversity e.g. Shannon, observed OTUs or Simpson. As illustrated in the examples (Fig. I-3), samples do not have the same number of sequences nor number of observed OTUs. To have an even sequencing depth a “vertical line” is drawn on the plot to indicate the maximum number of sequences in each sample. Deciding a high number of sequences, in this case more than 2000 sequences, will discard sample C, similarly, a very low number will result in ignoring a high number of the observed OTUs. Therefore, a combination of these two factors should be considered in deciding at which sequencing depth the samples should be rarefied. To have an even sequencing depth a “vertical line” is drawn on the plot to indicate the maximum number of sequences in each sample. Deciding a high number of sequences, in this case more than 2000 sequences, will discard sample C, similarly, a very low number will result in ignoring a high number of the observed OTUs. Therefore, a combination of these two factors should be considered in deciding at which sequencing depth the samples should be rarefied.

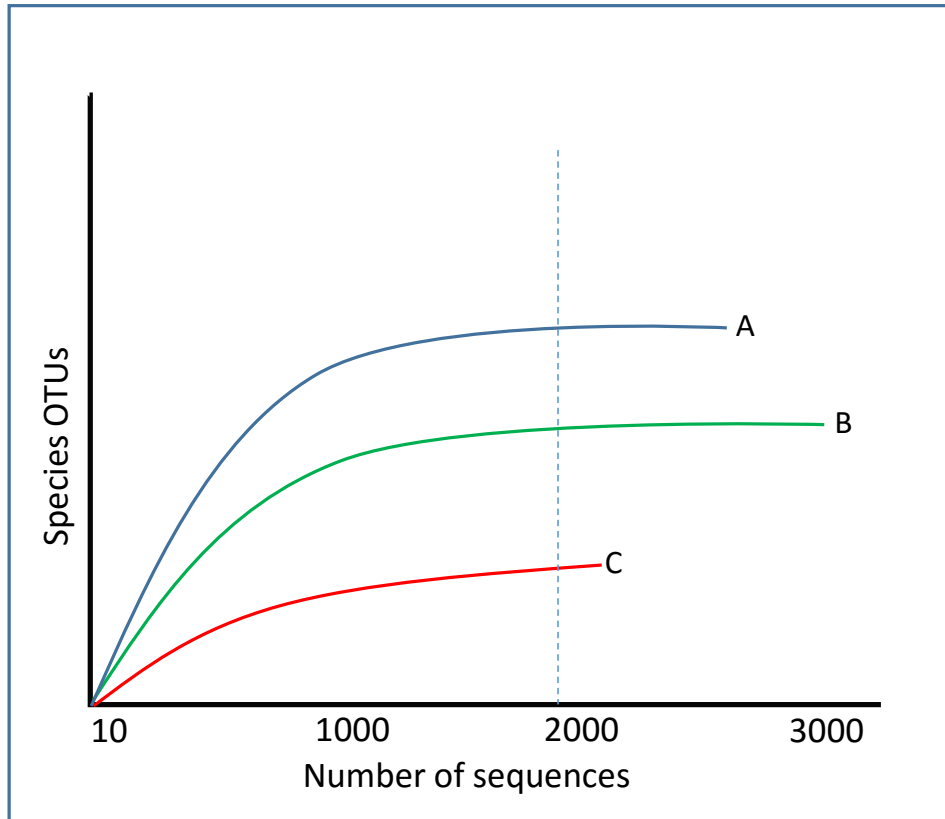


Figure I-3 Example of the rarefaction curve, plotting the number of sequences vs. the value of species diversity, demonstrating heterogeneous sequencing depth. Each curve represents a sample (A, B and C). A dashed line, nearly at 2000 sequences, indicates a “good” threshold to use as maximum number of sequences considered from each sample to have an even depth as method of normalization.

Even though, rarefying is not an ideal normalization method, as it reduces statistical power by removing some data, and was not designed to address compositionality, it still outperformed alternative methods that have not been developed until recently such as: MetagenomeSeq’s cumulative sum scaling (CSS) and DESeq’s variance-stabilizing transformation (Anders and Huber 2010, Paulson et al. 2013, Weiss et al. 2015).

7. Taxonomy assignment

Once sequences are clustered and OTUs are created, one sequence is used to represent each OTU, representative sequences are then assigned to the closest taxa possible. OTUs can be further clustered to represent any taxonomical level. Several different methods for taxonomic assignment have been proposed implementing different algorithms and for different types of data. They also vary in their performance, runtime, accuracy and suitability to the studied dataset. For instance, RTAX (Soergel et al. 2012) is more suitable for short sequences, especially paired-end reads such as produced by Illumina platforms, if they do not overlap. Whereas, taxonomy assigners such RDP (Wang et al. 2007) and Mothur (Schloss et al. 2009) are based on Naïve Bayesian Classifier algorithm, and were

developed originally for bacterial rRNA 16S sequences. Other assigners such as UCLUST (Edgar 2010) and BLAST (Altschul et al. 1990), although used for bacterial studies, they were the most used in classifying fungal ITS reads. These methods are available on most pipelines used to analyze environmental sequences. Regardless of the method, a database is essential for any taxonomy assigner to perform. Building a database however, is very challenging by itself.

Since the International Code of Botanical Nomenclature (McNeill et al. 2006) forbids the description of new species based solely on sequencing data, new fungal species identified by metabarcoding cannot have Latin names assigned to them. UNITE database (<http://unite.ut.ee>) provides informal, operational names or accession numbers until a formal description of the species. UNITE is a database for ITS sequences designed to provide a stable and reliable platform for sequence-borne identification of fungi, which includes only high-quality sequences of well identified fungi choosing quality over quantity. Other available databases based on the bacterial 16S rRNA are Silva (Pruesse et al. 2007), Greengene (DeSantis et al. 2006) and the International Nucleotide Sequence Database (INSD), which is the largest and the most popular database of DNA sequences worldwide (<http://www.insdc.org/>). However, this latter has increased the chance of misidentification because of the lower accuracy and reliability of the deposited sequences (Nilsson et al. 2006). For this reason, validated databases like UNITE have gained popularity in the last few years in studying fungal diversity and ecology from environmental samples. OTUs are also used for diversity assessments such as alpha and beta diversity, which are the main diversity measures used in the metabarcoding.

8. Methods in calculating microbial diversity

The main advantage of metabarcoding over traditional methods is its ability to theoretically detect and quantify all organisms existing in a given environment. However, results of metabarcoding analyses can be used to measure microbial diversity in terms of alpha and beta diversity, although these diversity indices have been defined and proposed more than 50 years ago (Whittaker 1960).

Alpha diversity

Alpha diversity is defined as the total diversity of a given sample. Alpha diversity is the result of the combination of two aspects in the species composition of a given community. The former is the number of species (species richness) usually referred to as

R and simply consider how many species are present in a given sample (more species are found in a sample the more R increases). The second aspect is the species evenness, which considers the relative abundance of the different species making up the richness of the sample. These aspects are commonly calculated together as an index of diversity. Several indexes have been developed to measure alpha diversity being Chao1, Shannon and Simpson the most common.

The Simpson's index is based on the probability that any two individuals randomly selected from a community belong to the same species and is calculated with the formula $D = \sum p_i^2$ where p_i is the proportion of individuals found in species, i for a finite community, this is $D = \sum \frac{n_i(n_i-1)}{N(N-1)}$ where n_i is the number of entities belonging to the i th type and N is the total number of entities in the dataset (Simpson 1949). The Simpson index is a dominance index because it gives more weight to common or dominant species and therefore rare taxa does not affect the diversity.

The Shannon index is also known as Shannon's diversity index, Shannon–Wiener index, Shannon–Weaver index or Shannon entropy. Like Simpson's index, Shannon's index accounts for both abundance and evenness of the species present in the analyzed sample. The index is calculated with the formula $H' = -\sum p_i \ln p_i$ where p_i is the proportion of individuals found in species i , \ln is the natural log, and Σ is the sum of the calculations (Shannon 2001). This index is relatively easy to calculate, yet sensitive to changes in rare species.

The Chao1 index estimates the total species richness as $S_{Chao1} = S_{obs} + \frac{n_1^2}{2n_2}$ where S_{obs} is the number of observed species, n_1 is the number of singletons *i.e.* only one observation, and n_2 is the number of doubletons *i.e.* only two observations (Chao 1984). Unlike The Simpson index, this method use information on the frequency of rare species in a sample to estimate the number of undetected species in an assemblage. Therefore, it is more sensitive to rare taxa than the other methods.

Beta diversity

Beta diversity measures the diversity between samples. It was originally defined as “*the extent of change of community composition, or degree of community differentiation, in relation to a complex gradient of environment, or a pattern of environments, which may be designated secondary*” (Whittaker 1960) or the ratio between

gamma (regional) and alpha (local) diversities $\beta = \gamma/\alpha$. The majority of the proposed methods to calculate beta diversity are based on similarity or distance/dissimilarity between samples which is calculated by either presence/absence of the species *i.e.* unweighted or take the species abundance into consideration *i.e.* weighted. More recently another method called UniFrac was introduced to consider the relative relatedness of community members by incorporating phylogenetic distances between observed organisms into the calculation (Lozupone and Knight 2005). UniFrac distance metrics can account for the species abundance “weighted” or considers only species presence or absence “unweighted”. UniFrac is calculated by dividing two separate equations (*u*) raw UniFrac on a scaling factor (*D*). $u = \sum_i^n b_i \times \left| \frac{A_i}{A_T} - \frac{B_i}{B_T} \right|$ where, *n* is the total number of branches in the tree, *b_i* is the length of branch *i*, *A_i* and *B_i* are the numbers of sequences that descend from branch *i* in communities *A* and *B*, respectively, and *A_T* and *B_T* are the total numbers of sequences in communities *A* and *B*, respectively. Whereas (*D*), measures the average distance of each sequence from the root as follows: $D = \sum_j^n d_j \times \left(\frac{A_j}{A_T} + \frac{B_j}{B_T} \right)$ where, *d_j* is the distance of sequence *j* from the root, *A_j* and *B_j* are the numbers of times the sequences were observed in communities *A* and *B*, respectively, and *A_T* and *B_T* are the total numbers of sequences from communities *A* and *B*, respectively.

Another index to evaluate the beta diversity is the Bray–Curtis index. It calculates the dissimilarity between samples. The dissimilarity metrics can vary only between 0 and 1, where 0 means samples are identical, sharing all species, and 1 means samples are completely different, and do not have any shared species (Bray and Curtis 1957). Unlike UniFrac, the Bray–Curtis accounts always for species abundance, and does not require information about the taxonomical relatedness of the observed species, hence no phylogenetic tree is required for the calculation. This makes this index useful when DNA alignments are not possible or not useful. The Bray-Curtis index is calculated using the following equation $BC_{ij} = 1 - \frac{2C_{ij}}{S_i + S_j}$ where *C_{ij}* is the sum of the lesser values for only those species in common between both sites. *S_i* and *S_j* are the total number of specimens counted at both sites. The index reduces to $2C/2 = C$ where abundances at each site are expressed as a percentage.

9. Statistical comparisons

Although useful, the interpretation of the metagenomic results based on only taxonomical assignments could be tricky. The amount of data is incredible and can be confusing especially when comparing group of samples or categories (treatment vs. control). Data need to be visualized and compared statistically to be correctly interpreted. One way to do that is by comparing alpha diversity (within-sample) of each sample or category. Regardless of the method used in calculating the alpha diversity, samples can be compared statistically by either the nonparametric two-sample t-test, which uses Monte Carlo permutations to calculate the *P*-value, or the parametric two-sample t-test, which uses the t-distribution instead of permutations. In either cases, the results can be visualized in a boxplots for further comparisons.

Another way to compare samples or categories statistically and test for trends, is by comparing beta diversity. Unlike alpha diversity, beta diversity comparison tests for relation between samples in their microbial communities. For instance, a comparison between two categories (healthy vs. diseased). When coupled with standard multivariate statistical techniques including principal coordinates analysis (PCoA), beta diversity can be easily visualized on PCA or PCoA plots. However, a statistical test is important to assess if significance exists or not. Several statistical methods are available, perhaps the most common ones used in metagenomic and microbial ecology studies are the nonparametric methods where the significance is determined through permutations, *e.g.* ADONIS, ANOSIM, PERMANOVA and Distance-based redundancy analysis (db-RDA).

In some cases, it is important to identify OTUs frequencies that differ between samples or categories, also known as differential abundance. OTU differential abundance can be applied by using parametric tests, such as Mann-Whitney test for tests of two classes, or nonparametric Kruskal-Wallis test for tests of multiple groups. In general, non-parametric tests are often preferred because most OTU counts are not normally distributed (Wagner et al. 2011).

Scope of the thesis

The main objective of this thesis was to investigate the total fungal diversity associated with the phyllosphere and/or carposphere of four major fruit species (olive, oranges, strawberries and apples), using a metabarcoding approach and universal fungal primers in conjunction with next generation sequencing (NGS). Two different NGS approaches (454

GS FLX Titanium and Illumina MiSeq) were utilized. The identification and quantification of endophytic and epiphytic microflora present in and on plants can provide important information that can be used to increase our understanding of the complex interactions that take place between plants and resident fungal microflora, including both pathogenic and beneficial species. Such data also have potential practical implications for disease management, particularly of economically important pathogens that can impact fruit yield and quality at both pre-and post-harvest stages of production. Documenting the relative abundance (RA) of fungal pathogens on different plant organs can be used to better understand the etiology of complex plant diseases and facilitate their management. Moreover, the presence of secondary pathogens and/or rare taxa on plants is not well documented due to either their low economic impact or that their presence has not been detected by conventional culturing methods. The impact of conventional chemical treatments on the composition and diversity of fungal communities on strawberries was also investigated. Since chemical treatments are intensively utilized in fruit production, it is very important to better understand the potential effect of these treatments on fungal communities. In addition, the fungal diversity of conventional and organic apple fruit was investigated. Finally, the effect of heat treatments, as an alternative control method to reduce postharvest rots, on apple fruit fungal diversity was evaluated.

References

- Acinas, S. G., R. Sarma-Rupavtarm, V. Klepac-Ceraj, and M. F. Polz. 2005. PCR-induced sequence artifacts and bias: insights from comparison of two 16S rRNA clone libraries constructed from the same sample. *Applied and Environmental Microbiology* **71**:8966-8969.
- Altschul, S. F., W. Gish, W. Miller, E. W. Myers, and D. J. Lipman. 1990. Basic local alignment search tool. *Journal of Molecular Biology* **215**:403-410.
- Anders, S., and W. Huber. 2010. Differential expression analysis for sequence count data. *Genome Biology* **11**:R106.
- Anderson, I. C., C. D. Campbell, and J. I. Prosser. 2003. Potential bias of fungal 18S rDNA and internal transcribed spacer polymerase chain reaction primers for estimating fungal biodiversity in soil. *Environmental Microbiology* **5**:36-47.
- Begerow, D., H. Nilsson, M. Unterseher, and W. Maier. 2010. Current state and perspectives of fungal DNA barcoding and rapid identification procedures. *Applied Microbiology and Biotechnology* **87**:99-108.
- Bennett, S. 2004. Solexa Ltd. *Pharmacogenomics* **5**:433-438.
- Blaalid, R., S. Kumar, R. H. Nilsson, K. Abarenkov, P. M. Kirk, and H. Kauserud. 2013. ITS1 versus ITS2 as DNA metabarcodes for fungi. *Molecular Ecology Resources* **13**:218-224.

- Blaxter, M., J. Mann, T. Chapman, F. Thomas, C. Whitton, R. Floyd, and E. Abebe. 2005. Defining operational taxonomic units using DNA barcode data. *Philosophical Transactions of the Royal Society B: Biological Sciences* **360**:1935-1943.
- Bokulich, N. A., S. Subramanian, J. J. Faith, D. Gevers, J. I. Gordon, R. Knight, D. A. Mills, and J. G. Caporaso. 2013. Quality-filtering vastly improves diversity estimates from Illumina amplicon sequencing. *Nature methods* **10**:57-59.
- Bray, J. R., and J. T. Curtis. 1957. An ordination of the upland forest communities of southern Wisconsin. *Ecological Monographs* **27**:325-349.
- Buckler, E. S., A. Ippolito, and T. P. Holtsford. 1997. The evolution of ribosomal DNA divergent paralogues and phylogenetic implications. *Genetics* **145**:821-832.
- Buee, M., M. Reich, C. Murat, E. Morin, R. H. Nilsson, S. Uroz, and F. Martin. 2009. 454 Pyrosequencing analyses of forest soils reveal an unexpectedly high fungal diversity. *New Phytologist* **184**:449-456.
- Cai, Y., and Y. Sun. 2011. ESPRIT-Tree: hierarchical clustering analysis of millions of 16S rRNA pyrosequences in quasilinear computational time. *Nucleic Acids Research*.
- Caporaso, J. G., J. Kuczynski, J. Stombaugh, K. Bittinger, F. D. Bushman, E. K. Costello, N. Fierer, A. G. Pena, J. K. Goodrich, J. I. Gordon, G. A. Huttley, S. T. Kelley, D. Knights, J. E. Koenig, R. E. Ley, C. A. Lozupone, D. McDonald, B. D. Muegge, M. Pirrung, J. Reeder, J. R. Sevinsky, P. J. Turnbaugh, W. A. Walters, J. Widmann, T. Yatsunencko, J. Zaneveld, and R. Knight. 2010. QIIME allows analysis of high-throughput community sequencing data. *Nature Methods* **7**:335-336.
- Chao, A. 1984. Nonparametric estimation of the number of classes in a population. *Scandinavian Journal of Statistics*:265-270.
- Coleman, A. W. 2003. ITS2 is a double-edged tool for eukaryote evolutionary comparisons. *Trends in Genetics* **19**:370-375.
- DeSantis, T. Z., P. Hugenholtz, N. Larsen, M. Rojas, E. L. Brodie, K. Keller, T. Huber, D. Dalevi, P. Hu, and G. L. Andersen. 2006. Greengenes, a Chimera-Checked 16S rRNA Gene Database and Workbench Compatible with ARB. *Applied and Environmental Microbiology* **72**:5069-5072.
- Edgar, R. C. 2010. Search and clustering orders of magnitude faster than BLAST. *Bioinformatics* **26**:2460-2461.
- Edgar, R. C., B. J. Haas, J. C. Clemente, C. Quince, and R. Knight. 2011. UCHIME improves sensitivity and speed of chimera detection. *Bioinformatics* **27**:2194-2200.
- Ganley, A. R., and T. Kobayashi. 2007. Highly efficient concerted evolution in the ribosomal DNA repeats: total rDNA repeat variation revealed by whole-genome shotgun sequence data. *Genome Research* **17**:184-191.
- Hawksworth, D. L. 2001. The magnitude of fungal diversity: the 1.5 million species estimate revisited. *Mycological Research* **105**:1422-1432.
- Hebert, P. D., A. Cywinska, and S. L. Ball. 2003. Biological identifications through DNA barcodes. *Proceedings of the Royal Society of London B: Biological Sciences* **270**:313-321.

- Hershkovitz, M., E. Zimmer, and W. Hahn. 1999. Ribosomal DNA sequences and angiosperm systematics. *Molecular systematics and plant evolution*. ((Systematics Association special volume **57**:268-326.
- Hollingsworth, P. M., L. L. Forrest, J. L. Spouge, M. Hajibabaei, S. Ratnasingham, M. van der Bank, M. W. Chase, R. S. Cowan, D. L. Erickson, and A. J. Fazekas. 2009. A DNA barcode for land plants. *Proceedings of the National Academy of Sciences* **106**:12794-12797.
- Hong, S., J. Bunge, C. Leslin, S. Jeon, and S. S. Epstein. 2009. Polymerase chain reaction primers miss half of rRNA microbial diversity. *ISME J* **3**:1365-1373.
- Horton, T. R. 2002. Molecular approaches to ectomycorrhizal diversity studies: variation in ITS at a local scale. *Plant and Soil* **244**:29-39.
- Izzo, A., J. Agbowo, and T. D. Bruns. 2005. Detection of plot-level changes in ectomycorrhizal communities across years in an old-growth mixed-conifer forest. *New Phytologist* **166**:619-629.
- Jorgensen, S. E., and B. Fath. 2008. *Encyclopedia of Ecology, Five-Volume Set: Online version*. Elsevier.
- Karsch-Mizrachi, I., Y. Nakamura, and G. Cochrane. 2012. The International Nucleotide Sequence Database Collaboration. *Nucleic Acids Research* **40**:D33-D37.
- Kellenberger, E. 2001. Exploring the unknown. *EMBO reports* **2**:5-7.
- Kovács, G. M., T. Jankovics, and L. Kiss. 2011. Variation in the nrDNA ITS sequences of some powdery mildew species: do routine molecular identification procedures hide valuable information? *European Journal of Plant Pathology* **131**:135-141.
- Kress, W. J., and D. L. Erickson. 2008. DNA barcodes: genes, genomics, and bioinformatics. *Proceedings of the National Academy of Sciences* **105**:2761-2762.
- Lane, D. J., B. Pace, G. J. Olsen, D. A. Stahl, M. L. Sogin, and N. R. Pace. 1985. Rapid determination of 16S ribosomal RNA sequences for phylogenetic analyses. *Proceedings of the National Academy of Sciences* **82**:6955-6959.
- Li, W., and A. Godzik. 2006. Cd-hit: a fast program for clustering and comparing large sets of protein or nucleotide sequences. *Bioinformatics* **22**:1658-1659.
- Lindahl, B. D., R. H. Nilsson, L. Tedersoo, K. Abarenkov, T. Carlsen, R. Kjølner, U. Kõljalg, T. Pennanen, S. Rosendahl, and J. Stenlid. 2013. Fungal community analysis by high-throughput sequencing of amplified markers—a user's guide. *New Phytologist* **199**:288-299.
- Lozupone, C., and R. Knight. 2005. UniFrac: a new phylogenetic method for comparing microbial communities. *Applied and Environmental Microbiology* **71**:8228-8235.
- Mahé, F., T. Rognes, C. Quince, C. de Vargas, and M. Dunthorn. 2014. Swarm: robust and fast clustering method for amplicon-based studies. *PeerJ* **2**:e593.
- Margulies, M., M. Egholm, W. E. Altman, S. Attiya, J. S. Bader, L. A. Bemben, J. Berka, M. S. Braverman, Y.-J. Chen, and Z. Chen. 2005. Genome sequencing in microfabricated high-density picolitre reactors. *Nature* **437**:376-380.
- Masella, A. P., A. K. Bartram, J. M. Truszkowski, D. G. Brown, and J. D. Neufeld. 2012. PANDAseq: paired-end assembler for Illumina sequences. *BMC bioinformatics* **13**:31.

- McNeill, J., F. R. Barrie, H. M. Burdet, V. Demoulin, D. L. Hawksworth, K. Marhold, D. H. Nicolson, J. Prado, P. C. Silva, and J. E. Skog. 2006. International Code of Botanical Nomenclature:(Vienna Code). ARG Ganter Verlag Ruggell, Liechtenstein.
- Moreira, D., and P. López-García. 2002. The molecular ecology of microbial eukaryotes unveils a hidden world. *Trends in microbiology* **10**:31-38.
- Mullis, K. B., H. A. Erlich, N. Arnheim, G. T. Horn, R. K. Saiki, and S. J. Scharf. 1987. One of the first Polymerase Chain Reaction (PCR) patents. U.S. Patent 4,683,195.
- Nilsson, R. H., E. Kristiansson, M. Ryberg, N. Hallenberg, and K.-H. Larsson. 2008. Intraspecific ITS variability in the kingdom Fungi as expressed in the international sequence databases and its implications for molecular species identification. *Evolutionary bioinformatics online* **4**:193.
- Nilsson, R. H., M. Ryberg, E. Kristiansson, K. Abarenkov, K.-H. Larsson, and U. Kõljalg. 2006. Taxonomic Reliability of DNA Sequences in Public Sequence Databases: A Fungal Perspective. *PloS one* **1**:e59.
- Oakley, B. B., T. L. Fiedler, J. M. Marrazzo, and D. N. Fredricks. 2008. Diversity of human vaginal bacterial communities and associations with clinically defined bacterial vaginosis. *Applied and environmental microbiology* **74**:4898-4909.
- Orgiazzi, A., V. Bianciotto, P. Bonfante, S. Daghino, S. Ghignone, A. Lazzari, E. Lumini, A. Mello, C. Napoli, S. Perotto, A. Vizzini, S. Bagella, C. Murat, and M. Girlanda. 2013. 454 pyrosequencing analysis of fungal assemblages from geographically distant, disparate soils reveals spatial patterning and a core mycobiome. *Diversity* **5**:73-98.
- Ottesen, A. R., A. G. Peña, J. R. White, J. B. Pettengill, C. Li, S. Allard, S. Rideout, M. Allard, T. Hill, and P. Evans. 2013. Baseline survey of the anatomical microbial ecology of an important food plant: *Solanum lycopersicum* (tomato). *BMC Microbiology* **13**:114.
- Pace, N. R. 1997. A molecular view of microbial diversity and the biosphere. *Science* **276**:734-740.
- Paulson, J. N., O. C. Stine, H. C. Bravo, and M. Pop. 2013. Differential abundance analysis for microbial marker-gene surveys. *Nature methods* **10**:1200-1202.
- Pinto, C., D. Pinho, S. Sousa, M. Pinheiro, C. Egas, and A. C. Gomes. 2014. Unravelling the diversity of grapevine microbiome. *PloS one* **9**:e85622.
- Prokopowich, C. D., T. R. Gregory, and T. J. Crease. 2003. The correlation between rDNA copy number and genome size in eukaryotes. *Genome* **46**:48-50.
- Pruesse, E., C. Quast, K. Knittel, B. M. Fuchs, W. Ludwig, J. Peplies, and F. O. Glöckner. 2007. SILVA: a comprehensive online resource for quality checked and aligned ribosomal RNA sequence data compatible with ARB. *Nucleic Acids Research* **35**:7188-7196.
- Qiu, X., L. Wu, H. Huang, P. E. McDonel, A. V. Palumbo, J. M. Tiedje, and J. Zhou. 2001. Evaluation of PCR-generated chimeras, mutations, and heteroduplexes with 16S rRNA gene-based cloning. *Applied and Environmental Microbiology* **67**:880-887.
- Reeder, J., and R. Knight. 2010. Rapidly denoising pyrosequencing amplicon reads by exploiting rank-abundance distributions. *Nature methods* **7**:668-669.

- Riesenfeld, C. S., P. D. Schloss, and J. Handelsman. 2004. Metagenomics: genomic analysis of microbial communities. *Annual Review of Genetics* **38**:525-552.
- Rognes, T., F. Mahé, and T. Flouri. 2015. vsearch
- Ronaghi, M., M. Uhlén, and P. Nyrén. 1998. A sequencing method based on real-time pyrophosphate. *Science* **281**:363-365.
- Saiki, R. K., S. Scharf, F. Faloona, K. B. Mullis, G. T. Horn, H. A. Erlich, and N. Arnheim. 1985. Enzymatic amplification of beta-globin genomic sequences and restriction site analysis for diagnosis of sickle cell anemia. *Science* **230**:1350-1354.
- Sanger, F., S. Nicklen, and A. R. Coulson. 1977. DNA sequencing with chain-terminating inhibitors. *Proceedings of the National Academy of Sciences* **74**:5463-5467.
- Schena, L., S. Mosca, S. O. Cacciola, R. Faedda, S. M. Sanzani, G. E. Agosteo, V. Sergeeva, and G. Magnano di San Lio. 2013. Species of the *Colletotrichum gloeosporioides* and *C. boninense* complexes associated with olive anthracnose. *Plant Pathology*: **63**(2), 437-446.
- Schloss, P. D., and S. L. Westcott. 2011. Assessing and improving methods used in operational taxonomic unit-based approaches for 16S rRNA gene sequence analysis. *Applied and Environmental Microbiology* **77**:3219-3226.
- Schloss, P. D., S. L. Westcott, T. Ryabin, J. R. Hall, M. Hartmann, E. B. Hollister, R. A. Lesniewski, B. B. Oakley, D. H. Parks, and C. J. Robinson. 2009. Introducing mothur: open-source, platform-independent, community-supported software for describing and comparing microbial communities. *Applied and Environmental Microbiology* **75**:7537-7541.
- Schoch, C. L., K. A. Seifert, S. Huhndorf, V. Robert, J. L. Spouge, C. A. Levesque, W. Chen, E. Bolchacova, K. Voigt, and P. W. Crous. 2012. Nuclear ribosomal internal transcribed spacer (ITS) region as a universal DNA barcode marker for Fungi. *Proceedings of the National Academy of Sciences* **109**:6241-6246.
- Shannon, C. E. 2001. A mathematical theory of communication. *ACM SIGMOBILE Mobile Computing and Communications Review* **5**:3-55.
- Shendure, J., G. J. Porreca, N. B. Reppas, X. Lin, J. P. McCutcheon, A. M. Rosenbaum, M. D. Wang, K. Zhang, R. D. Mitra, and G. M. Church. 2005. Accurate multiplex polony sequencing of an evolved bacterial genome. *Science* **309**:1728-1732.
- Simpson, E. H. 1949. Measurement of diversity. *Nature*.
- Soergel, D. A. W., N. Dey, R. Knight, and S. E. Brenner. 2012. Selection of primers for optimal taxonomic classification of environmental 16S rRNA gene sequences. *ISME J* **6**:1440-1444.
- Sogin, M. L., H. G. Morrison, J. A. Huber, D. M. Welch, S. M. Huse, P. R. Neal, J. M. Arrieta, and G. J. Herndl. 2006. Microbial diversity in the deep sea and the underexplored "rare biosphere". *Proceedings of the National Academy of Sciences* **103**:12115-12120.
- Stockinger, H., M. Kruger, and A. Schussler. 2010. DNA barcoding of arbuscular mycorrhizal fungi. *New Phytologist* **187**:461-474.
- Sun, Y., Y. Cai, L. Liu, F. Yu, M. L. Farrell, W. McKendree, and W. Farmerie. 2009. ESPRIT: estimating species richness using large collections of 16S rRNA pyrosequences. *Nucleic Acids Research* **37**:e76-e76.

- Taberlet, P., E. Coissac, F. Pompanon, C. Brochmann, and E. Willerslev. 2012. Towards next-generation biodiversity assessment using DNA metabarcoding. *Molecular Ecology* **21**:2045-2050.
- Thomas, T., J. Gilbert, and F. Meyer. 2012. Metagenomics - a guide from sampling to data analysis. *Microbial Informatics and Experimentation* **2**:3.
- Turnbaugh, P. J., R. E. Ley, M. Hamady, C. Fraser-Liggett, R. Knight, and J. I. Gordon. 2007. The human microbiome project: exploring the microbial part of ourselves in a changing world. *Nature* **449**:804.
- Unterseher, M., D. Peršoh, and M. Schnittler. 2013. Leaf-inhabiting endophytic fungi of European Beech (*Fagus sylvatica* L.) co-occur in leaf litter but are rare on decaying wood of the same host. *Fungal Diversity* **60**:43-54.
- Ursell, L. K., J. L. Metcalf, L. W. Parfrey, and R. Knight. 2012. Defining the human microbiome. *Nutrition Reviews* **70**:S38-S44.
- Vilgalys, R., and D. Gonzalez. 1990. Organization of ribosomal DNA in the basidiomycete *Thanatephorus praticola*. *Current Genetics* **18**:277-280.
- Wagner, B. D., C. E. Robertson, and J. K. Harris. 2011. Application of two-part statistics for comparison of sequence variant counts. *PloS one* **6**:e20296-e20296.
- Wang, Q., G. M. Garrity, J. M. Tiedje, and J. R. Cole. 2007. Naïve Bayesian Classifier for Rapid Assignment of rRNA Sequences into the New Bacterial Taxonomy. *Applied and Environmental Microbiology* **73**:5261-5267.
- Watson, J. D., and F. H. Crick. 1953. Molecular structure of nucleic acids. *Nature* **171**:737-738.
- Weiss, S. J., Z. Xu, A. Amir, S. Peddada, K. Bittinger, A. Gonzalez, C. Lozupone, J. R. Zaneveld, Y. Vazquez-Baeza, and A. Birmingham. 2015. Effects of library size variance, sparsity, and compositionality on the analysis of microbiome data. 2167-9843, PeerJ PrePrints.
- Whitman, W. B., D. C. Coleman, and W. J. Wiebe. 1998. Prokaryotes: the unseen majority. *Proceedings of the National Academy of Sciences* **95**:6578-6583.
- Whittaker, R. H. 1960. Vegetation of the Siskiyou mountains, Oregon and California. *Ecological Monographs* **30**:279-338.
- Wu, J. Y., X. T. Jiang, Y. X. Jiang, S. Y. Lu, F. Zou, and H. W. Zhou. 2010. Effects of polymerase, template dilution and cycle number on PCR based 16 S rRNA diversity analysis using the deep sequencing method. *BMC Microbiology* **10**:255.
- Yilmaz, P., R. Kottmann, D. Field, R. Knight, J. R. Cole, L. Amaral-Zettler, J. A. Gilbert, I. Karsch-Mizrachi, A. Johnston, G. Cochrane, R. Vaughan, C. Hunter, J. Park, N. Morrison, P. Rocca-Serra, P. Sterk, M. Arumugam, M. Bailey, L. Baumgartner, B. W. Birren, M. J. Blaser, V. Bonazzi, T. Booth, P. Bork, F. D. Bushman, P. L. Buttigieg, P. S. Chain, E. Charlson, E. K. Costello, H. Huot-Creasy, P. Dawyndt, T. DeSantis, N. Fierer, J. A. Fuhrman, R. E. Gallery, D. Gevers, R. A. Gibbs, I. San Gil, A. Gonzalez, J. I. Gordon, R. Guralnick, W. Hankeln, S. Highlander, P. Hugenholtz, J. Jansson, A. L. Kau, S. T. Kelley, J. Kennedy, D. Knights, O. Koren, J. Kuczynski, N. Kyrpides, R. Larsen, C. L. Lauber, T. Legg, R. E. Ley, C. A. Lozupone, W. Ludwig, D. Lyons, E. Maguire, B. A. Methe, F. Meyer, B. Muegge,

S. Nakielny, K. E. Nelson, D. Nemergut, J. D. Neufeld, L. K. Newbold, A. E. Oliver, N. R. Pace, G. Palanisamy, J. Peplies, J. Petrosino, L. Proctor, E. Pruesse, C. Quast, J. Raes, S. Ratnasingham, J. Ravel, D. A. Relman, S. Assunta-Sansone, P. D. Schloss, L. Schriml, R. Sinha, M. I. Smith, E. Sodergren, A. Spo, J. Stombaugh, J. M. Tiedje, D. V. Ward, G. M. Weinstock, D. Wendel, O. White, A. Whiteley, A. Wilke, J. R. Wortman, T. Yatsunenko, and F. O. Glockner. 2011. Minimum information about a marker gene sequence (MIMARKS) and minimum information about any (x) sequence (MIxS) specifications. *Nature Biotechnology* **29**:415-420.

Zhang, J., K. Kobert, T. Flouri, and A. Stamatakis. 2014. PEAR: a fast and accurate Illumina Paired-End reAd mergeR. *Bioinformatics* **30**:614-620.

Chapter II. Metabarcoding analysis of fungal diversity in the phyllosphere and carposphere of olive (*Olea europaea*)

Abstract

The fungal diversity associated with leaves, flowers and fruit of olive (*Olea europaea*) was investigated in different phenological stages (May, June, October and December) using an implemented metabarcoding approach. It consisted of the 454 pyrosequencing of the fungal ITS2 region and the subsequent phylogenetic analysis of relevant genera along with validated reference sequences. Most sequences were identified up to the species level or were associated with a restricted number of related *taxa* enabling supported speculations regarding their biological role. Analyses revealed a rich fungal community with 195 different OTUs. Ascomycota was the dominating phyla representing 93.6% of the total number of detected sequences followed by unidentified fungi (3.6%) and Basidiomycota (2.8%). A higher level of diversity was revealed for leaves compared to flowers and fruit. Among plant pathogens the genus *Colletotrichum* represented by three species (*C. godetiae* syn. *C. clavatum*, *C. acutatum* s.s and *C. karstii*) was the most abundant on ripe fruit but it was also detected in other organs. *Pseudocercospora cladosporioides* was detected with a high frequency in all leaf samples and to a less extent in ripe fruit. A much lower relative frequency was revealed for *Spilocaea oleagina* and for other putative pathogens including *Fusarium* spp., *Neofusicoccum* spp., and *Alternaria* spp. Among non-pathogen *taxa*, *Aureobasidium pullulans*, the species complex of *Cladosporium cladosporioides* and *Devriesia* spp. were the most represented. This study highlights the existence of a complex fungal consortium including both phytopathogenic and potentially antagonistic microorganisms that can have a significant impact on olive productions.

Introduction

Among the different fruit tree species, olive (*Olea europaea*) is one of the most important crops on a global scale and its cultivation is rapidly extending due to the growing awareness of the health benefits associated with olive oil consumption. Currently, the available data on fungal communities inhabiting olive phyllosphere and carposphere is very limited and mostly related to plant pathogens, while almost nothing is known about other fungi (Schena et al. 2011). Major fungal pathogens attacking olive canopy are

Spilocaea oleagina, *Colletotrichum* spp., and *Pseudocercospora cladosporioides*, responsible for peacock spot, anthracnose and cercosporiosis, respectively (Agosteo and Schena 2011). Other fungal pathogens belonging to the family Botryosphaeriaceae and to the genus *Fusarium* can occasionally cause rots on olive drupes (Frisullo and Carlucci 2011). Apart from plant pathogens, several other fungal species have been associated with the olive sooty mold - *i.e.* a consequence of the epiphytic growth of different saprophytic fungi forming a black film covering leaves, fruit, twigs and branches (Frisullo and Carlucci 2011).

Insights into the ecology, identification and quantification of olive pathogens as well as other epiphytic microflora will facilitate the understanding of the complex interactions between the plant and the resident fungal microflora, including potential pathogens and their antagonists. This information may have important practical implications on the management of phyllosphere and carposphere “microenvironments” since many microorganisms exert a competitive action against pathogens and/or induce mechanisms of resistance in the host. Therefore, accurate knowledge of the composition of fungal microbiota may be essential to develop effective biological disease control strategies that are environment-friendly and safe for consumers. These aspects are becoming more and more important in the public opinion and government policies as evidenced by the growing diffusion of organic agriculture.

The massive sequencing of polymerase chain reaction (PCR) amplicons from specific barcode genes represents a powerful culture-independent technique to investigate the whole microbial diversity and determine the relative quantity of community members in environmental samples (Lindahl et al. 2013). This technique may be properly referred to as amplicon metagenomics or metabarcoding and, in the case of fungal communities, is almost exclusively based on the amplification of the ITS regions of the ribosomal DNA (rDNA) (Bellemain et al. 2010, Eberhardt 2010, Schoch et al. 2012). A major drawback of the ITS regions concerns difficulties in discriminating phylogenetically related species that may have sequences which are identical or differ only by a few nucleotide positions. This weakness is emphasized in recent high-throughput sequencing approaches since they are commonly based either on the ITS1 or ITS2 regions due to restrictions in sequence read lengths (Blaalid et al. 2013). Consequently, utilized bioinformatics tools and genetic databases enable a good identification of microorganisms up to the level of the genus but they become less reliable when used to identify fungal species. This is particularly true for members of the phylum Ascomycota and becomes particularly relevant for fungal plant

pathogens since related species with very similar ITS sequences can be characterized by completely different pathogenetic behavior.

Despite the above mentioned drawbacks, the use of the ITS as barcode gene in fungal metabarcoding represents, at least for the moment, a forced choice since this gene can be easily amplified and sequenced with universal primers and largely prevails in GenBank and other databases (Kõljalg et al. 2005, Pruesse et al. 2007, Merget et al. 2012). The choice of either ITS1 or ITS2 is optional since these regions share many properties, and enable similar levels of discrimination (Bazzicalupo et al. 2013). However, ITS2 is generally easier to align because it is less variable in length and lacks the problem of co-amplification of a 5' SSU intron that is common in many Ascomycota (Lindahl et al. 2013). Furthermore, the ITS2 is better represented than ITS1 in databases (Nilsson et al. 2009).

In recent years, metagenomic analyses have been increasingly utilized to investigate microbial diversity in a number of different environments and, in most of the cases, they have revealed a vast previously unknown microbial biodiversity missed by conventional cultivation-based methods. A large part of these studies have focused on soils and rhizosphere environments (Buee et al. 2009, Stockinger et al. 2010, Orgiazzi et al. 2013). On the other hand, the plant canopy environment has not been widely investigated and only few studies have focused on phyllosphere endophytes (Unterseher et al. 2013) while, to the best of our knowledge, fungal biodiversity has been recently investigated only in tomato, grape leaves and balsam poplar (Ottesen et al. 2013, Pinto et al. 2014).

The aim of the present study was the use of the ITS2 region in a metabarcoding approach to investigate the fungal microbiota of the olive phyllosphere and carposphere in different phenological phases. Representative ITS2 sequences were phylogenetically analyzed along with selected reference sequences to enable the most accurate possible identification of putative species.

Materials and Methods

Sampling and DNA extractions

Samples of flowers, leaves, fertilized fruitlets and asymptomatic and symptomatic (rotted) fruit were collected in four phenological stages (May 28; June 29; October 17; and December 12) from nine different trees located in three different farms of the Gioia Tauro Plain (Calabria, Italy) with GPS coordinates 38°22'53.0"N and 15°56'27.5"E,

38°22'15.1"N and 15°55'38.3"E and 38°24'44.6"N and 15°56'23.1"E, respectively (Table II-1). Investigated farms shared the same olive cultivar (Ottobratica) and were representative of the Gioia Tauro plain, an olive-growing area with over 32,000 Ha, where epidemic outbreaks of anthracnose caused by *Colletotrichum* spp. occur on a yearly basis (Graniti et al. 1993). Symptoms on symptomatic fruit collected in December were considered to be specific of *Colletotrichum* infections, although the role of other fungal pathogens in fruit rots was not completely excluded.

A total of 72 samples (3 fields x 3 trees x 8 sample types) were collected during 2013 (Table II-1). Each sample consisted of 30-200 g of tissues according to the type of plant part and was harvested around the entire canopy at a height of approximately 2 m in order to cover all cardinal directions. Samples were collected in sterile plastic bags and maintained at 4°C for 4-5 hours before lyophilization. To facilitate the lyophilization, leaves were cut into small pieces of approximately 5-10 mm. Similarly, the flesh of drupes collected in December was cut and separated from the stone. Fertilized fruitlets were lyophilized without any additional treatment. Total DNA was extracted from lyophilized samples as described by Mosca et al. (2014) and purified using chromatography columns according to Ruano-Rosa et al. (2007). The quantity and quality of purified DNA extracts were determined using a Nanodrop 2000 spectrophotometer (Nano-drop Technologies, Wilmington, DE) and by electrophoresis on a 1.2 agarose gel.

Fungal DNA amplification

DNA extracts obtained from the collected samples (72) were amplified in triplicate, using the universal fungal primers ITS3-ITS4 to amplify the ITS2 region of the ribosomal DNA (Toju et al. 2012). Both primers were modified to construct fusion primers appropriate for 454 sequencing with adapter sequences A and B, a key sequence and multiplex identifiers (MIDs) (<http://www.454.com/>). Primers labelled with different MIDs were utilized during amplifications to identify the eight sample types collected (Table II-1). In all amplification set specific negative control reactions with water replacing template DNA were used.

PCR reactions were conducted in a total volume of 25 µl containing 2.5µl 10X of reaction buffer, 0.25 µl of each primer (10µM), 0.1µl of DNA of AccuPrime Taq DNA Polymerase High Fidelity (Invitrogen, CA, USA) and 1µl of DNA template (100 ng/µl). Reactions were incubated in an Eppendorf Mastercycler gradient (Hamburg, Germany) for

1 min at 94°C followed by 30 cycles of 30s at 94°C, 30 s at 55°C and 30 s at 68°C. All reactions ended with a final extension of 1 min at 68°C. Triplicate PCR products from each sample were combined and purified with agarose Gel Extract Mini (5 Prime-USA). After purification, concentration and quality were evaluated spectrophotometrically and by gel electrophoresis. Ten µl of each purified sample were pooled together and sequenced by Macrogen Inc. (Seoul, Korea) using 454 GS FLX+ System (Roche Diagnostics Corporation). Data were deposited in the BioProject database (NCBI) as PRJNA270912 (<http://www.ncbi.nlm.nih.gov/bioproject/270912>) (Table II-1).

Data analysis and statistics

Data analysis was conducted using the bioinformatics pipeline QIIME v. 1.8 (Caporaso et al. 2010). De-multiplexing and quality filtering analyses were done using a minimum quality score of 25, a minimum/maximum length ratio of 200/1000 and a maximum number of homopolymer bases of 6. Additionally, the sliding window test of quality scores (-w) was enabled with a value of 50 to discard sequences with bad windows according to the "-g" command. Sequences were denoised using the denoise wrapper (Reeder and Knight 2010) and the ITS2 region was extracted using ITSx software (Bengtsson-Palme et al. 2013). Chimeric sequences were identified and filtered using USEARCH 6.1 (Edgar 2010). The most abundant sequences were picked as representative sequences to be used in Operational taxonomical units (OTUs) picking and taxonomy assignments. OTUs were picked using the BLAST method (Altschul et al. 1990) and the UNITE dynamic database released on February 2, 2014 (<http://unite.ut.ee/>). The same database was also utilized for taxonomy assignments (Altschul et al. 1990) using a sequence similarity threshold of 0.97 and maximum e-values of 0.001 and 1e-10 in picking OTUs and in taxonomy assignments, respectively. The taxonomic assignments and the operational taxonomical unite map (OTU map) were used to create the OTU table needed to construct the heat-map and the taxa summaries.

Since the rarefaction plots of the entire OTU table as a function of the sequencing effort with a maximum of 6000 sequences per sample revealed heterogeneity in sampling, the OTU table was rarefied to even sequencing depth of 2800 sequences to remove sample heterogeneity. Weighted and unweighted UniFrac metrics were utilized to evaluate Beta diversity (Lozupone and Knight 2005). Alfa diversity was determined by Shannon's Diversity Index and Chao1 estimate. Beta diversity served to construct UPGMA trees and

PCoA plots. The uncertainty in the UPGMA tree was estimated by performing jack-knifing at a depth of 2000 sequences. Trees were visualized and edited in Mega6 (Tamura et al. 2013). To highlight shared phlotypes, Venn Diagrams were created using the OTUs table created in QIIME and visualized on the website <http://bioinfogp.cnb.csic.es/tools/venny/index.html> (Oliveros 2007).

Identification of fungal species

In order to confirm the accuracy of taxonomic assignments, sequences associated with each OTU within each identified fungal genus, were extracted and introduced in ElimDupes (<http://hcv.lanl.gov/content/sequence/ELIMDUPES/elimdupes.html>) to detect multiple identical sequences and determine their frequency. Unique representative sequences defined as sequence types (STs), *i.e.* distinct and reproducible ITS2 sequences recovered in this study, were then manually blasted to identify the closest available reference sequences in the complete NCBI nucleotide collection (<http://blast.ncbi.nlm.nih.gov/Blast>). Furthermore, ITS2 sequences of the most abundant fungal genera according to the QIIME taxonomic assignments (*Aureobasidium* spp., *Colletotrichum* spp., *Cladosporium* spp., *Pseudocercospora* spp., and *Devriesia* spp.) were phylogenetically analyzed. STs were analyzed along with genetically closely related reference sequences of the same genus to determine their phylogenetic collocation and enable their identification with the highest possible level of accuracy. Before analysis, validated panels of reference ITS2 sequences of *Colletotrichum acutatum* s.l. (Cannon et al. 2012, Damm et al. 2012a, Weir et al. 2012), *C. boninense* s.l. (Damm et al. 2012b), *Pseudocercospora* spp. (Crous et al. 2013), *Devriesia* spp. (Li et al. 2013), *Cladosporium* spp. (Bensch et al. 2012) and *Aureobasidium* spp. (Zalar et al. 2008) were analyzed with the software ElimDupes to delete multiple identical sequences. Some identical reference sequences were included in the panel because they were representative of different species. When none of the above-validated reference sequences was identical to sequences identified in the present study, eventual more closely related sequences were identified by BLAST analyses. Despite being low abundant, a similar analysis was also performed for the genus *Spilocaea* in light of its relevance as olive fungal pathogen (Agosteo and Schena 2011). In this case, reference sequences were downloaded from GenBank because of the lack of a validated panel of reference sequences. For each genus, STs identified in the present study and reference sequences were aligned using MUSCLE and introduced to

MEGA for phylogenetic analysis with the Maximum Likelihood method using the Tamura-Nei model (Tamura et al. 2013). Analyses were performed with 1000 bootstrap replications.

Results

Fungal diversity and richness

A fragment of the expected size ($\cong 340$ bp) was obtained from all olive samples investigated in the present study. The complete panel of amplicons obtained from 8 different olive sample types yielded a total of 151,671 reads with an average length of 337bp. After quality evaluations (length trimming, denoising, ITS2 extraction and chimeric sequence exclusion) a total of 58,245 high quality sequences were recovered and assigned to 195 OTUs. The number of detected sequences ranged from 14,610 in asymptomatic fruit collected in December (DAFr) and 2,801 in fertilized fruitlets collected in June (JFr) (Table II 1).

Table II-1 Summary of analyses and results of field surveys conducted with different olive tissues collected in four phenological phases from nine different trees located in three different farms.

Sampling date	Sample type	MIDs	Sequences	SRA*	OTUs (Total)**	OTUs (2800)***	Chao1 estimate	Shannon diversity
28.05.12	Leaves	MID3	3335	SRX821225	71	66	93.95	3.78
	Flowers	MID4	9551	SRX1025531	67	42	64.04	2.76
29.06.12	Leaves	MID5	4083	SRX1025537	72	66	95.70	3.63
	Fertilized fruit	MID2	2801	SRX1025554	44	44	91.50	2.163
17.10.12	Leaves	MID10	4816	SRX1025575	85	70	101.24	3.68
12.12.12	Leaves	MID16	12301	SRX1025576	117	88	109.52	3.61
	Asymptomatic fruit	MID19	14610	SRX1025577	88	49	78.22	1.76
	Symptomatic fruit	MID28	6748	SRX1025578	37	30	35.85	1.08

*Sequence Read Archive (SRA) accession numbers for sequences deposited in the BioProject database (NCBI) as PRJNA270912 (<http://www.ncbi.nlm.nih.gov/bioproject/270912>);

**Total number of detected OTUs;

***Number of OTUS detected with an even sequencing depth of 2800 sequences.

The rarefaction analysis assigned to 97% of OTUs similarity showed the achievement of the saturation zone for all samples, suggesting that the great majority of the OTUs was detected in the present study (Fig. II-1).

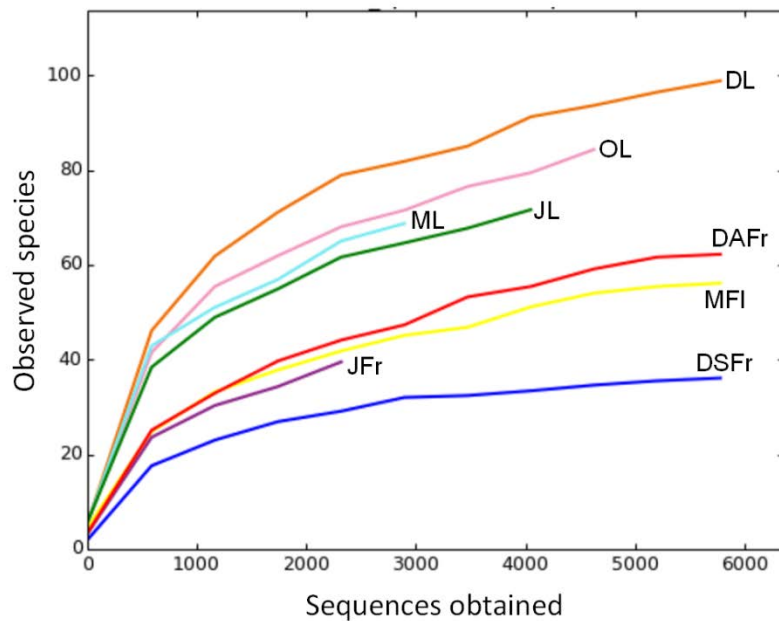


Figure II-1 Rarefaction curves at a genetic distance of 3% determined for May leaves (ML), June leaves (JL), October leaves (OL), December leaves (DL), May flowers (MFI), June fertilized fruitlets (JFr), December asymptomatic fruits (DAFr) and December symptomatic fruits (DSFr).

The number of detected OTUs varied between 37 in symptomatic fruit collected in December (DSFr) and 117 in December leaves (DL). A general higher number of OTUs was detected in leaves compared to May flowers (MFI), JFr and DSFr (Fig. II-2A, B, and C). Furthermore, leaves were characterized by a higher number of unique OTUs that were not recovered from the other samples of fruit and flowers (Fig. II-2C). On leaves, the number of OTUs progressively increased from May to December (Fig. II-2 E). With regard to ripe fruit a higher number of OTUs was revealed on DAFr compared to DSFr (Fig. II-2 F). In agreement with OTU numbers, Shannon's Diversity Index and the Chao1 estimate based on alpha diversity indexes, achieved their highest values for leaves against MFI, JFr, DSFr and DAFr (Table II-1). No big differences were detected for leaves according to the sampling period.

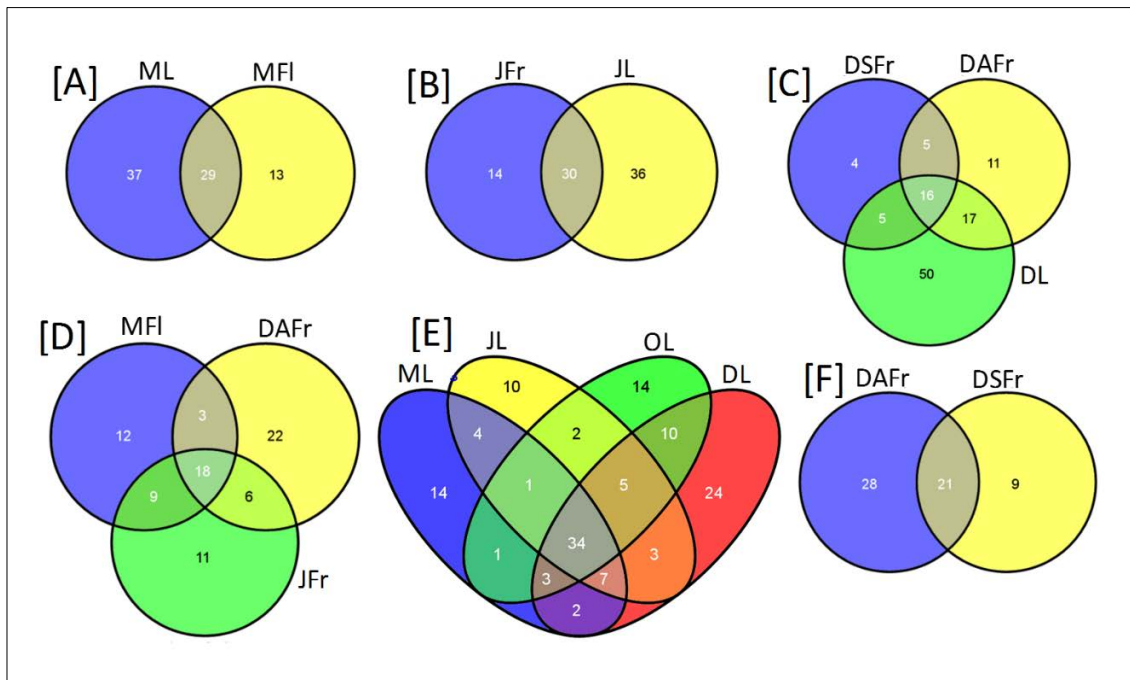


Figure II-2 Venn Diagrams reporting the number of OTUs shared among investigated olive sample types in different possible combinations. Analyzed samples included May leaves (ML), June leaves (JL), October leaves (OL), December leaves (DL), May flowers (MFI), June fertilized fruitlets (JFr), December asymptomatic fruits (DAFr) and December symptomatic fruits (DSFr).

As regards flowers evolving into fruit, higher Shannon’s Diversity values were determined for MFI compared to DSFr and DAFr, suggesting a uniform distribution of abundance amongst species. In particular, the lowest value was achieved for DSFr due to the large prevalence of the genus *Colletotrichum* (Table II-1).

Beta diversity analysis revealed a clear separation of samples according to sample type. Principal Coordinates Analysis (PCoA) showed a close association between leaf samples that were clearly segregated from all other samples. Interestingly, a close association was revealed between MFI and JFr (Fungal population on ripe fruit was quite different on DSFr and DAFr but, overall, clearly different compared to other sample types. Results of the PCoA analyses were confirmed by the UPGMA tree (Fig. II-3B).

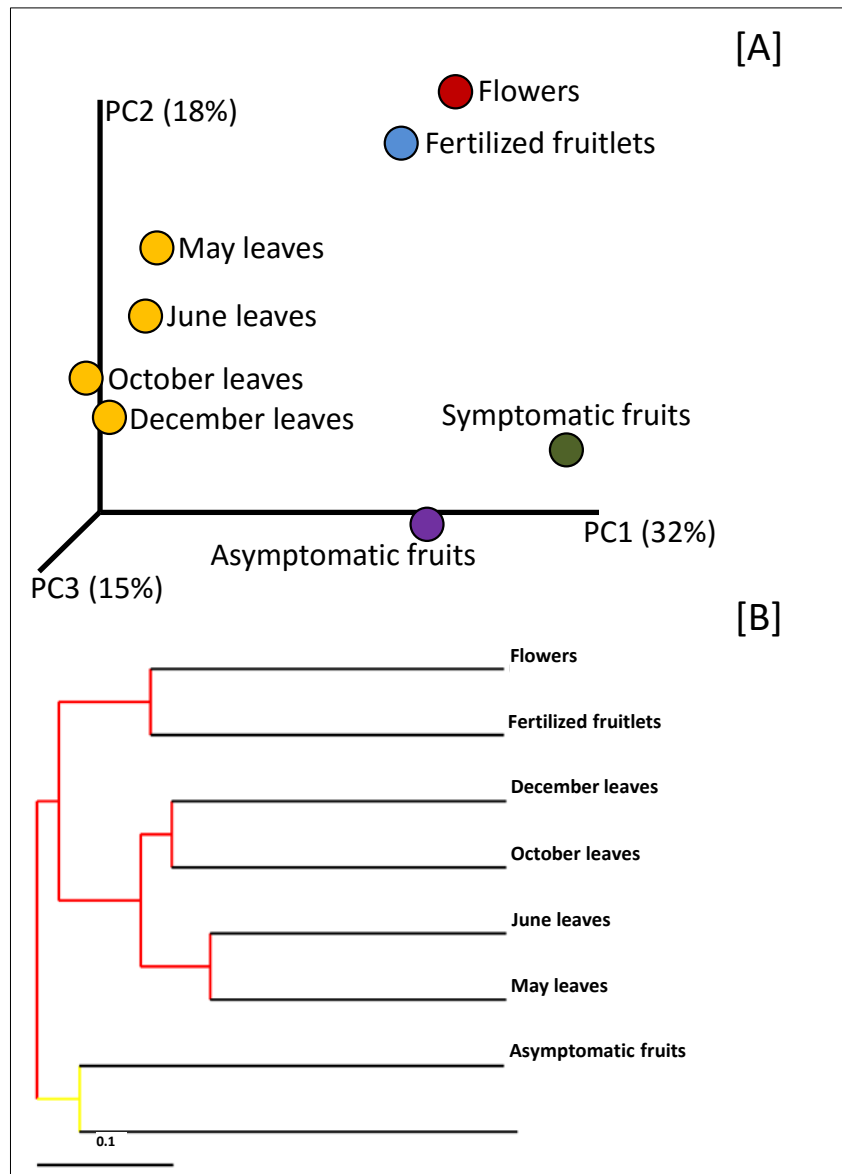


Figure II-3 Unweighted UniFrac distance PCoA (A) and Jackknife dendrogram (B) of fungal communities associated with May leaves (ML), June leaves (JL), October leaves (OL), December leaves (DL), May flowers (MFl), June fertilized fruitlets (JFr), December asymptomatic fruits (DAFr) and December symptomatic fruits (DSFr). In B red and yellow colors of nodes indicate 75-100% and 50-75% of bootstrap support, respectively.

Olive fungal community structure

According to the analysis of the complete ITS2 data set, members of the phylum Ascomycota dominated in all samples and accounted for 95.2% of the total number of detected sequences followed by unidentified fungi (3.6%) and Basidiomycota (1.5%).

In the different samples the incidence of Ascomycota varied between 88.6 (May Leaves, ML) and 99.3% (DSFr and DAFr) and was generally higher in fruit compared to flowers and leaves (Fig. II-4A). Basidiomycota accounted for 2.5-3.5% of the reads on leaves but were almost completely absent on flowers and fruit.

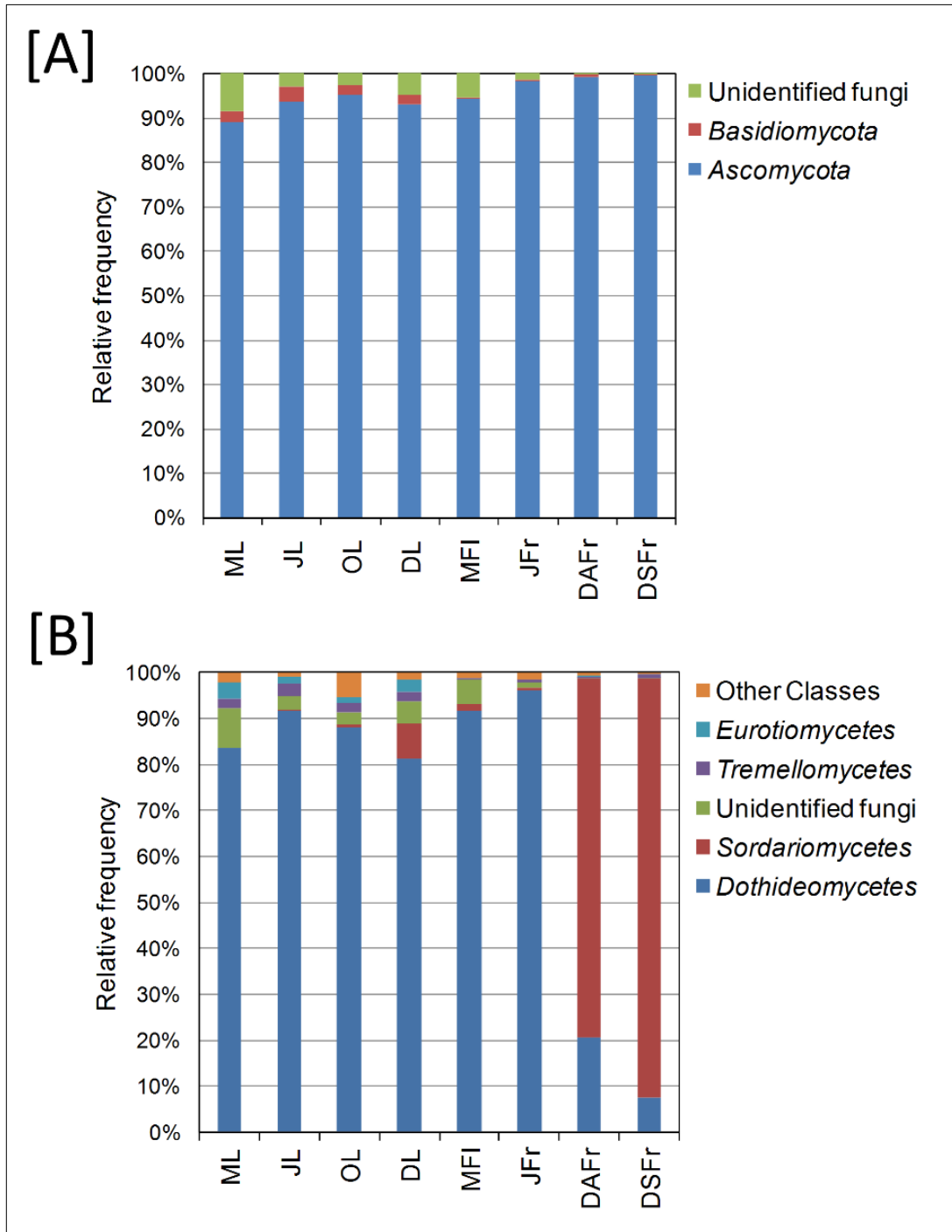


Figure II-4 Relative abundance of different fungal phyla (top) and classes (bottom) detected in May leaves (ML), June leaves (JL), October leaves (OL), December leaves (DL), May flowers (MFI), June fertilized fruitlets (JFr), December asymptomatic fruits (DAFr) and December symptomatic fruits (DSFr).

Similarly, sequences ascribed to non-identified fungi were more abundant on leaves and flowers (4.8-8.7%) and almost completely absent on fruit.

Ascomycota reads were largely members of the Dothideomycetes and Sordariomycetes classes (Fig. II-4B). The first class predominated on leaves (80.9-91.3%), MFI (91.5%) and JFr (96.4%) while members of the Sordariomycetes class were the most abundant on ripe fruit accounting for 76.8 and 91.1% of the reads on DAFr and DSFr, respectively (Fig. II-4B). Tremellomycetes (Basidiomycota) and Eurotiomycetes (Ascomycota) accounted for 0.2-2.7 and 0.1-3.5% of the total reads, respectively. All other detected fungal classes accounted for only 1.7% of the reads (Fig. II-4B).

Detected OTUs were associated with a total of 117 different *taxa*. Most of these (103) were identified at the level of genus. However, 1, 2, 2, 4, and 5 *taxa* were only associated with fungi at the level of Kingdom, Phylum, Class, Order and Family, respectively, due to the lack of closely related sequences in genetic databases. Despite the high number of *taxa* identified, few genera accounted for most reads. Looking at all analyzed samples, the most abundantly detected genus was *Aureobasidium* (31.70%) followed by *Colletotrichum* (19.10%), *Pseudocercospora* (12.70%), *Cladosporium* (12.70%), and *Devriesia* (9.50%) (Fig. I-5). However, the relative frequency of these genera was not consistent among samples nor during sampling periods. The genus *Aureobasidium* had its highest incidence (43.6-61.2%) in ML, JL and MFI and decreased to 5.3, 3.9 and 0.9% in DL, DSFr and DAFr, respectively (Fig. I-5). In contrast, *Colletotrichum*, which showed almost no presence until June, started to rise in OL (7.4%) and had its peak in DAFr (70.7%) and DSFr (87.6%). The genus *Pseudocercospora* showed a consistent presence in all leaf samples and its population reached the highest relative frequency in October (30.3%). This genus was also detected on DAFr (14.4%), but not on MFI, JFr and DSFr. The genus *Cladosporium* was detected in all samples but was more abundant in JFr (28.1%), JL (16.8%) and MFI (13.4%). On the contrary, it represented only 1.8 and 2.8% of the reads on DSFr and DAFr, respectively. Finally, the genus *Devriesia*, was almost exclusively detected on leaves and was particularly abundant in December and October, accounting for 35.7% and 18.2% of the total reads, respectively (Fig. II-5).

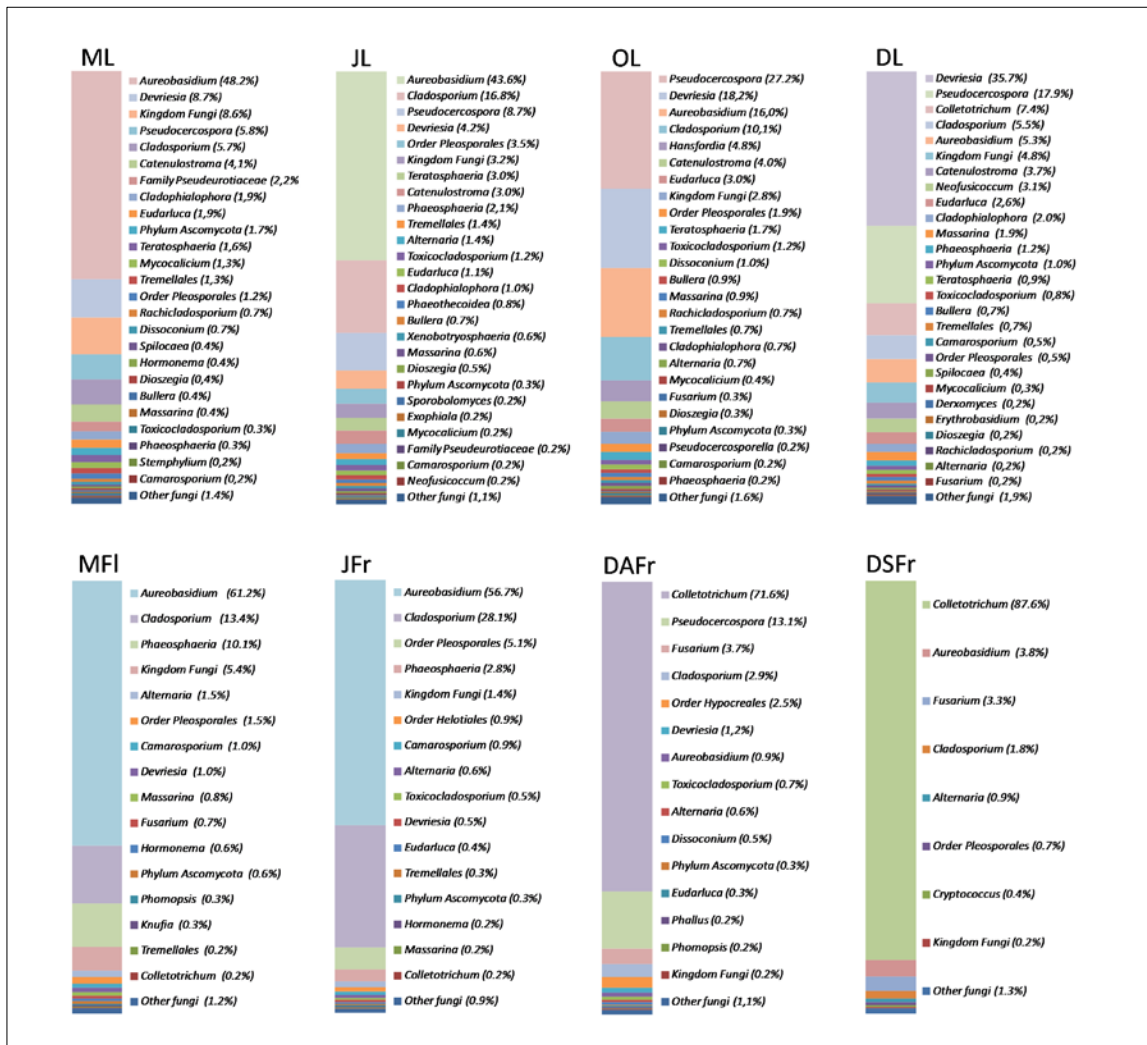


Figure II-5. Relative abundance of fungal genera detected in May leaves (ML), June leaves (JL), October leaves (OL), December leaves (DL), May flowers (MFI), June fertilized fruitlets (JFr), December asymptomatic fruits (DAFr) and December symptomatic fruits (DSFr). Fungal genera representing less than 0.2% of the total population in each sample, were considered as a single taxa and labeled as "Other fungi". In each column fungal genera are listed according to their abundance.

Identification of fungal species

The manual BLAST analysis of STs confirmed taxonomic assignments of detected OTUs. In addition, phylogenetic analysis of ITS2 sequences of *Aureobasidium* spp., *Colletotrichum* spp., *Cladosporium* spp., *Pseudocercospora* spp., *Devriesia* spp. and *Spilocaea* spp. enabled the identification of most detected STs at the level of species or their association with a restricted number of related species. Three different STs were associated with the genus *Aureobasidium* (Fig. I-6D). A largely prevalent ST (AUR 1)

accounting for 12,562 reads clustered with *A. pullulans* var. *pullunans* (Zalar et al. 2008). The same groups also contained a rare sequence type (AUR 2) differing in a single nucleotide. The other ST (AUR 3) was identified as *A. pullulans* var. *namibiae* (Fig. II-6D).

Five different STs of *Colletotrichum* spp. were detected and found to belong to *C. acutatum* s.l. (Fig. II-6E) and *C. boninense* s.l. (Fig. II-6F). Among *C. acutatum* s.l., an ST accounting for 141 reads (COLL 1) was identified as *C. acutatum* s.s. while the other two STs (COLL 2 and COLL 3) accounting for 7,929 and 9,393 reads, respectively, clustered with a group of genetically related species which also comprised *C. godetiae* (Fig. II-6E). The two STs clustering within *C. boninense* s.l. were associated with the species *C. karstii* and *C. phyllanthi* (Fig. II-6F).

Similarly, three STs were detected in the genus *Cladosporium* and according to their phylogenetic collocation were found to belong to the *C. cladosporioides* complex (Bensch et al. 2013). For the two largely prevalent sequences (CLA 1 and CLA 2) accounting for 2,200 and 24,476 reads, respectively, it was not possible to deepen the level of identification due to the high number of species having identical or very similar ITS2 sequences within the complex (Fig. II-6C). On the contrary, a third rare sequence (CLA 3) accounting for 7 reads was identified as *C. grevilleae* (Fig. II-6C).

Four different STs of *Pseudocercospora* differing for single indels' were identified as *P. cladosporioides* since they were clearly differentiated from all other recognized species of the genus (Fig. II-6A). A ST (PSEUD 1) was largely prevalent accounting for 5,892 reads while other sequences were rarely detected.

Similarly, four different STs were associated with the species *D. fraseriae* in light of a 98-99% sequence identity but, a single sequence (DEV 1) accounting for 5976 reads, was by large the most abundant (Fig. II-6G). Finally, a single ST was detected for the genus *Spilocaea* and was identical to reference sequences of *S. oleagina*, a major fungal pathogen of olive, and *Fusicladium phillyreae* (Fig. II-6B).

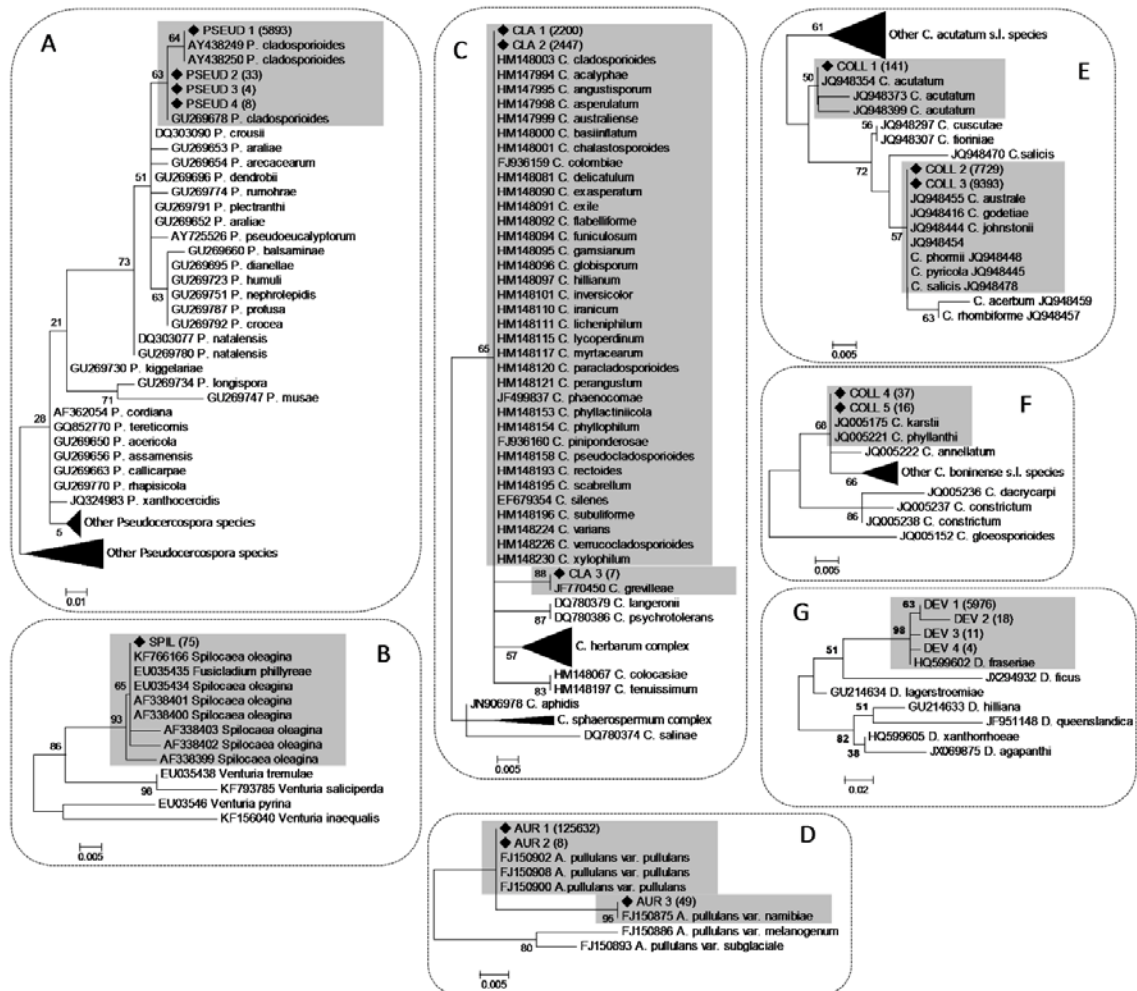


Figure II-6 Phylogenetic trees built using unique sequences representative of sequence types (STs) of the most relevant fungal genera in terms of both abundance and/or relevance as olive fungal pathogens, along with validated reference sequences. Reference sequences were from *Pseudocercospora* spp. (Crous et al. 2013) (A), *Cladosporium* spp. (Bensch et al., 2013) (C), *Aureobasidium* spp. (Zalar et al., 2008) (D), *Colletotrichum acutatum* s.l. (Damm et al., 2012a) (E), *Colletotrichum boninense* s.l. 2012 (Damm et al. 2012b) (F), and *Devriesia* spp. (Li et al. 2013) (B). Sequences of species closely related to *Spilocaea oleagina* were sourced in GenBank (B). Grey highlighted boxes contain STs identified in the present study (♦) and genetically close reference species with which they were associated. Numbers in parentheses along with STs indicate the number of identical sequences, represented by each ST. Numbers on nodes represent the posterior probabilities for the maximum likelihood method.

Discussion

In the present study, the fungal diversity associated with olive leaves, flowers and fruit was investigated at different phenological stages using a metabarcoding approach. Analyses revealed a rich fungal community with 195 different OTUs which were identified up to genus level in most of the cases. Furthermore, phylogenetic analyses permitted the identification of many STs at the level of species. A conspicuous number of OTUs, however, was only identified at the level of *Phylum*, *Order*, *Class* or *Family* suggesting the existence of several rare or still unknown fungal *taxa*. At least a part of these putative

species is likely to represent uncultivable *taxa* and their ecological role is currently unknown. These findings highlight the importance of the metabarcoding approaches in determining the complex microbial populations associated with plant surfaces since only a limited part of the available genetic variation can be investigated using traditional culturing methods.

Regardless of the sample type or period, Ascomycota was clearly the dominant phyla (93.6%) followed by unidentified fungi and Basidiomycota. Within Ascomycota, the class Dothideomycetes (70.1%) was the most abundant followed by Sordariomycetes (22.4%). The high incidence of the former class was primarily determined by the abundant detection of the genera *Aureobasidium* (Order Dothideales) and *Cladosporium*, *Pseudocercospora* and *Devriesia* (Order Capnodiales), while the class Sordariomycetes was almost completely represented by fungi of the genus *Colletotrichum* which were abundantly detected in both DAFr and DSFr. In agreement with our study, this fungal class was also predominant on grape leaves and grape must (Pinto et al. 2014). Indeed Dothideomycetes is one of the largest and most significant classes within Ascomycota, which also comprises thousands of plant pathogen species (Wijayawardene et al. 2014).

A higher level of diversity was revealed for leaves in comparison with flowers and fruit, regardless of the sampling period. In agreement with our results, fungal and bacterial populations on flowers and fruit were at least ten times less than those on leaves in *Sesamum* and *Gossypium* plants (Sharma and Mukerji 1976). This result may be related to the fact that leaves have a higher surface/volume ratio compared to fruit, cover most of the plant surface, are always present on the plant (olive is an evergreen tree), and each single leaf has a life cycle of approximately 18 months against the 5-6 months of the flowers evolving in fruit. Since leaves are always present on the tree, they can establish a sort of balanced community. In addition, leaves represent a different ecological niche in terms of surface composition and landscape (Mechaber et al. 1996, Lindow and Brandl 2003). Unlike leaves, the fungal community changed significantly from fruitlets to mature fruit and this was also associated with a sharp decline in fungal biodiversity (richness and evenness) mainly because of the epidemic outbreak of *Colletotrichum* spp. in mature symptomatic and asymptomatic fruit. In general, the olive fungal community on fruitlets was similar to that of the organs from which they originated (flowers) and progressively evolved in the fruit community. In agreement with our data, different anatomical parts of tomato were characterized by distinct microbial communities, but flowers and fruit shared a few bacterial *taxa* that were not detected in other parts of the plant (Ottesen et al. 2013).

It has been suggested that ovaries developing into the flesh of fruit represent a desired habitat for microbial colonization, potentially providing both long-term resources and shelter for the microbes able to enter (Alekkett et al. 2014). To some extent, a similar phenomenon is known to occur in humans since a new-born's gut microbial community is initially similar to that of the mother and then progressively evolves (Gonzalez et al. 2011).

In the present study, amplicons were analyzed with QIIME using a high quality filtering set up in order to minimize the impact of sequencing errors and achieve a reliable identification of OTUs. Indeed, the manual BLAST analysis of all detected STs confirmed the reliable identification of OTUs up to the level of genus. Conversely, standard QIIME analyses largely failed in the identification of fungal species. This result was partially expected considering that genetic variation within ITS regions may be very limited or inexistent among closely related species. Consequently, the level of homology commonly suggested in literature and utilized in the present study for picking OTUs (97%) did not enable an accurate identification of species. Furthermore, it is worth mentioning that unreliable annotations of sequences in public DNA repositories remain a serious obstacle to all sequence-based species identification (Nilsson et al. 2006). It should also be considered that a significant part of deposited ITS sequences are not updated and may not reflect recent advancements in fungal taxonomy (Crous et al. 2014, Schoch et al. 2014). However, despite the above limits and biases, the ITS regions are widely accepted as the official fungal DNA barcode marker because they can be easily amplified and sequenced using Sanger and second generation sequencing approaches (Schoch et al. 2012).

To exploit all available genetic variations within the ITS2 region and enable the identification of detected *taxa* with the highest possible level of accuracy, STs associated with the most relevant fungal genera in terms of both abundance and/or relevance as olive fungal pathogens, were subjected to specific phylogenetic analyses along with selected validated reference sequences. This approach proved to be reliable considering that the majority of the STs were identified at the level of species and that the remaining ones were associated with a restricted number of *taxa*. This achievement is very important for the analysis of plant-associated fungi with particular emphasis to pathogens, since related species with very similar ITS sequences can be characterized by completely different behaviors. Consequently, the determination of fungal diversity up to the level of species is essential to study the ecology and biology of fungal pathogens on their own hosts and in relation to other microorganisms. These data may be useful from a practical point of view since they may be used to evaluate the impact of control strategies on pathogens and non-

pathogens in the aerial plant surface with the aim of developing new more effective and less impacting means for disease management.

The reliability and accuracy of any phylogenetic approach used to identify fungal species are largely influenced by the comprehensiveness of data on the phylogenetic collocation of analyzed *taxa* and by the consequent availability of validated reference sequences. An accurately validated database has been recently released but it still does not cover all currently known genetic variations within the fungal kingdom (Schoch et al. 2014). However, several comprehensive studies have recently clarified the taxonomy in important fungal groups and many others are likely to come out in the near future. It is also possible to anticipate the future use of more variable markers as barcode genes in metabarcoding analyses to enable a higher level of discrimination among species (Cannon et al. 2012). However, the single copy nature of currently available marker genes is likely to provide lower levels of sensitivity compared to the multi-copy ITS regions. Furthermore, difficulties in designing reliable universal primers and the lower number of available reference sequences in genetic databases may represent an issue in species identification (Schoch et al. 2012).

A conspicuous part of sequences detected in the present study was associated with well-known olive fungal pathogens. The abundance detection of the genus *Colletotrichum* was expected considering that trials were conducted on the Gioia Tauro plain (southern Italy) which is characterized by recurrent anthracnose outbreaks. The detection of this genus in May, June and October with a low level of population suggests that flowers, leaves and unripe fruit may harbor latent infections. It has been suggested that latent infections in developing fruit during spring may favor the survival of the pathogen later in the hot and dry summer, but the relevance of these infections to *Colletotrichum* epidemic outbreaks is not yet well defined (Moral et al. 2009, Talhinas et al. 2011, Mosca et al. 2014). On the other hand, the high incidence of *Colletotrichum* species in DL represents a confirmation of reported data on the importance of fruit in the promotion of leaf infections (Moral et al. 2008, Moral et al. 2009). Among sequences associated with the species complex of *C. acutatum s.l.*, a ST was identified as *C. acutatum s.s.* while other two STs accounting for most of the detected reads clustered with a group of genetically related species which also comprised *C. godetiae*. Since this latter species is known as the causal agent of olive anthracnose in southern Italy it is likely that detected sequences belonged to it although the ITS2 region does not enable its differentiation from closely related species (Cacciola et al. 2012, Damm et al. 2012a). Similarly, STs clustering within the species

complex of *C. boninense s.l.* are likely to belong to *C. karstii* which has been recently reported on olives in southern Italy (Mosca et al. 2014). The detection of *C. acutatum s.s.* on olives in southern Italy confirms results of a recent metabarcoding investigation with *Colletotrichum* specific primers (Mosca et al. 2014). Interestingly, *C. acutatum s.s.* had been not associated with olive anthracnose in Italy until recently, suggesting an ongoing process of evolution in the populations of this important plant pathogen (Cacciola et al. 2012). The abundant detection of *Colletotrichum* spp. on ripe olives was expected for symptomatic fruit but it was to some extent surprising for the asymptomatic ones (71.6% of the total reads). This result may suggest a conspicuous colonization of olive tissues before the appearance of symptoms and the consequent competitive exclusion of other fungi.

Another widely detected plant pathogen was identified as *P. cladosporioides i.e.* the causal agent of olive cercosporiosis. The high incidence of *P. cladosporioides* on leaves in the investigated environment seems to be remarkable and may suggest a primary role of this pathogen in the defoliation of trees. By contrast and quite surprisingly the incidence of *S. oleagina*, commonly considered as the main causal agent of olive defoliation, was low (Nigro and Ferrara 2011). A possible explanation is the poor saprophytic and epiphytic ability of this fungal species; moreover, it is reported that the cultivar Ottobratica is highly resistant to the peacock spot disease. Infections by *P. cladosporioides* are generally considered to take place later in comparison with those caused by *S. oleagina* and therefore they should be limited to old, mature leaves of lower branches. However, in agreement with our data, field surveys conducted in southern Italy showed that the fungus can be active throughout the year (Nigro et al. 2002, Nigro and Ferrara 2011). The high incidence on DAFr (14.4% of the total reads) also suggests that infections on these organs may be quite common, although *P. cladosporioides* is mainly considered a leaf pathogen. In this regard, the sampling area (Gioia Tauro plain, southern Italy) is characterized by humid autumn conditions and it is reported that the severity of fruit infection is related to persistent humid and mild weather during the last three months before harvest (Pappas 1993).

At lower abundance, other olive pathogens represented by *Fusarium* spp. were primarily detected in asymptomatic and symptomatic ripe fruit. Although the analysis of sequences did not enable the identification of the species due to the complexity of the genus (Watanabe et al. 2011), it is reported that *Fusarium* species can be responsible for olive rots (Frisullo and Carlucci 2011). Interestingly, *Fusarium* spp. were also detected

with a lower incidence on flowers suggesting that the fungus may colonize and establish latent infection on these organs. Other detected fungi that are likely to act as plant pathogens were associated with the genera *Neofusicoccum* and *Alternaria*. Both genera comprise species that can cause leaf and fruit infections on olive (Frisullo and Carlucci 2011). In particular, according to BLAST analyses, two species of *Neofusicoccum*, *N. parvum* and *N. mediterraneum*, were detected on olive samples with the first accounting for a higher incidence. Additionally, *Stemphylium* spp. and *Phomopsis* spp., two genera comprising species that may be regarded as causal agents of minor olive diseases (Frisullo and Carlucci 2011) were detected in the current work with a very low incidence. Finally, some of the detected OTUs were associated with fungal genera (*Selenophoma* spp., *Teratosphaeria* spp. and *Cladophialophora* spp.) comprising relevant pathogens on other plant species, but their role on olive is completely unknown.

The genus *Aureobasidium* spp. was one of the most abundant on leaves, MFl and JFr and to a lesser extent was also detected in DAFr and DSFr. Detected STs were largely associated with *A. pullulans* var. *pullulans* and to a lesser extent with *A. pullulans* var. *namibiae*. The abundant presence of *A. pullulans* var. *pullulans* in the olive canopy was expected while, to the best of our knowledge, *A. pullulans* var. *namibiae* has never been reported as a colonizer of plant tissues (Zalar et al. 2008). *Aureobasidium pullulans* is a ubiquitous yeast-like fungus that can colonize almost all environmental niches including soil, water, air, and limestone. It has been reported as one of the most abundant fungal colonizers of phyllosphere and carposphere in a number of different plant species and may be present as both epiphyte and endophyte (Deshpande et al. 1992, Andrews et al. 1994). However, apart from a few reports in which it has been demonstrated to act as a pathogen on overripe fruit (Morgan and Michailides 2004, Kim 2014), it is generally considered a non-pathogen and has been widely exploited as a bio-control agent (Schena et al. 2003, Ippolito et al. 2005). On olives, it was mainly reported as a component of a complex of fungi causing sooty mold *i.e.* an epiphytic growth of different saprophytic fungi forming a black film which covers leaves, fruit, twigs and branches (Frisullo and Carlucci 2011).

Another widely detected genus in olive samples was *Cladosporium* spp. with three different STs. Two STs were largely prevalent and were associated with the *C. cladosporioides* species complex since available genetic variation within the ITS2 region does not enable the differentiation of most species within this complex (Dugan et al. 2013). A third ST accounting for few sequences was associated with the species *C. grevilleae*. The species complex of *C. cladosporioides* comprises very common cosmopolitan

saprophytic fungi. *C. cladosporioides* often occurs as a secondary invader of necrotic tissues or as a weak pathogen in many different host plants (Tashiro et al. 2013). However, some strains of *C. cladosporioides* have been also reported as effective bio-control agents (Wang et al. 2013). Together with *A. pullulans*, *C. cladosporioides* had been classified among sooty mold fungi that grow on the surface of leaves and other plant organs covered with insect or physiological honeydew in different plant species including olive (Nelson 2008, Frisullo and Carlucci 2011, Tashiro et al. 2013). In favorable conditions, both fungal species may produce a compact sooty thallus on the fruit surface even in the absence of honeydew (Grabowski 2007).

According to our data, the genus *Devriesia* is one of the most abundant fungal colonizers of olive canopy although it was almost exclusively detected on leaves and was particularly abundant in December and October. The Mediterranean climate characterized by hot and dry summers followed by mild temperatures and high rainfall incidence in autumn and winter, may have played a role in this temporal distribution. Interestingly, the genus *Devriesia* has never been reported as an olive fungal colonizer. It is possible that it had been formerly identified as a *Cladosporium* species since the genus *Devriesia* also contains species previously ascribed to this genus (Seifert et al. 2004). Sequences detected in the present study were associated with *D. fraseriae* a species described in 2010 to accommodate a fungus isolated from leaves of *Melaleuca* sp. in Australia. Considering its high incidence on olive leaves, it may be assumed to have a role in sooty mold along with other species of the Capnodiaceae family (Crous et al. 2009).

The high incidence of *Aureobasidium* spp., *Cladosporium* spp. and *Devriesia* spp. in olive phyllosphere and to a lesser extent in the olive carposphere suggests a primary role in the structure of fungal communities associated with olive trees and a possible competitive action against fungal plant pathogens. Furthermore, other genera detected with a low frequency (*Hormonema* spp., *Dissoconium* spp., *Hansfordia* spp., *Bullera* spp. and *Dioszegia* spp.) may have competitive actions since they comprise species with well documented antagonistic functions.

Conclusion

In conclusion results of the present study provide a comprehensive picture of the fungal diversity in the olive phyllosphere and carposphere. The majority of the detected sequences were identified up to the level of species making it possible to support assumptions regarding their role on the olive aerial plant surface. However, many others

were not associated with specific *taxa* and, even if identified, their significance remains to be interpreted. Altogether, this study reinforces the importance of investigating fungal biodiversity in olive culture and highlighting the need for more detailed analyses in this field. The olive fungal consortia showed to contain both beneficial and phytopathogenic microorganisms that can have a significant impact on olive productions. Beneficial colonizers, either epiphytes or entophytes, can be further explored as antagonists of the important pathogens of olive, and possibly developed as effective biocontrol agents.

References

- Agosteo, G. E., and L. Schena. 2011. Olive leaf spot. Pages 143-176 in L. Schena, G. E. Agosteo, and S. O. Cacciola, editors. Olive diseases and Disorders. Transworld Research Network, India
- Aleklett, K., M. Hart, and A. Shade. 2014. The microbial ecology of flowers: an emerging frontier in phyllosphere research 1. *Botany* **92**:253-266.
- Altschul, S. F., W. Gish, W. Miller, E. W. Myers, and D. J. Lipman. 1990. Basic local alignment search tool. *Journal of Molecular Biology* **215**:403-410.
- Andrews, J. H., R. F. Harris, R. N. Spear, G. W. Lau, and E. V. Nordheim. 1994. Morphogenesis and adhesion of *Aureobasidium pullulans*. *Canadian Journal of Microbiology* **40**:6-17.
- Bazzicalupo, A. L., M. Bálint, and I. Schmitt. 2013. Comparison of ITS1 and ITS2 rDNA in 454 sequencing of hyperdiverse fungal communities. *Fungal Ecology* **6**:102-109.
- Bellemain, E., T. Carlsen, C. Brochmann, E. Coissac, P. Taberlet, and H. Kauserud. 2010. ITS as an environmental DNA barcode for fungi: an *in silico* approach reveals potential PCR biases. *BMC Microbiology* **10**:189.
- Bengtsson-Palme, J., M. Ryberg, M. Hartmann, S. Branco, Z. Wang, A. Godhe, P. De Wit, M. Sánchez-García, I. Ebersberger, F. de Sousa, A. Amend, A. Jumpponen, M. Unterseher, E. Kristiansson, K. Abarenkov, Y. J. K. Bertrand, K. Sanli, K. M. Eriksson, U. Vik, V. Veldre, and R. H. Nilsson. 2013. Improved software detection and extraction of ITS1 and ITS2 from ribosomal ITS sequences of fungi and other eukaryotes for analysis of environmental sequencing data. *Methods in Ecology and Evolution* **4**:914-919.
- Bensch, K., U. Braun, J. Z. Groenewald, and P. W. Crous. 2012. The genus *Cladosporium*. *Studies in Mycology* **72**:1-401.
- Bensch, R., S. Song, O. Ronneberger, and W. Driever. 2013. Non-directional radial intercalation dominates deep cell behavior during zebrafish epiboly. *Biology open*:BIO20134614.
- Blaalid, R., S. Kumar, R. H. Nilsson, K. Abarenkov, P. M. Kirk, and H. Kauserud. 2013. ITS1 versus ITS2 as DNA metabarcodes for fungi. *Molecular Ecology Resources* **13**:218-224.

- Buee, M., M. Reich, C. Murat, E. Morin, R. H. Nilsson, S. Uroz, and F. Martin. 2009. 454 Pyrosequencing analyses of forest soils reveal an unexpectedly high fungal diversity. *New Phytologist* **184**:449-456.
- Cacciola, S. O., R. Faedda, F. Sinatra, G. E. Agosteo, L. Schena, S. Frisullo, and G. Magnano di San Lio. 2012. Olive anthracnose. *Journal of Plant Pathology* **94**:29-44.
- Cannon, P., U. Damm, P. Johnston, and B. Weir. 2012. *Colletotrichum* – current status and future directions. *Studies in Mycology* **73**:181-213.
- Caporaso, J. G., J. Kuczynski, J. Stombaugh, K. Bittinger, F. D. Bushman, E. K. Costello, N. Fierer, A. G. Pena, J. K. Goodrich, and J. I. Gordon. 2010. QIIME allows analysis of high-throughput community sequencing data. *Nature methods* **7**:335-336.
- Crous, P. W., U. Braun, G. C. Hunter, M. J. Wingfield, G. J. Verkley, H. D. Shin, C. Nakashima, and J. Z. Groenewald. 2013. Phylogenetic lineages in *Pseudocercospora*. *Studies in Mycology* **75**:37-114.
- Crous, P. W., A. Giraldo, D. L. Hawksworth, V. Robert, P. M. Kirk, J. Guarro, B. Robbertse, C. L. Schoch, U. Damm, and T. Trakunyingcharoen. 2014. The Genera of Fungi: fixing the application of type species of generic names. *IMA fungus* **5**:141.
- Crous, P. W., C. L. Schoch, K. D. Hyde, A. R. Wood, C. Gueidan, G. S. de Hoog, and J. Z. Groenewald. 2009. Phylogenetic lineages in the Capnodiales. *Studies in Mycology* **64**:17-47S17.
- Damm, U., P. Cannon, J. Woudenberg, and P. Crous. 2012a. The *Colletotrichum acutatum* species complex. *Studies in Mycology* **73**:37-113.
- Damm, U., P. Cannon, J. Woudenberg, P. Johnston, B. Weir, Y. Tan, R. Shivas, and P. Crous. 2012b. The *Colletotrichum boninense* species complex. *Studies in Mycology* **73**:1-36.
- Deshpande, M. S., V. B. Rale, and J. M. Lynch. 1992. *Aureobasidium pullulans* in applied microbiology: A status report. *Enzyme and Microbial Technology* **14**:514-527.
- Dugan, F. M., K. Schubert, and U. Braun. 2013. Check-list of *Cladosporium* names. *Schlechtendalia* **11**:1-103.
- Eberhardt, U. 2010. A constructive step towards selecting a DNA barcode for fungi. *New Phytologist* **187**:265-268.
- Edgar, R. C. 2010. Search and clustering orders of magnitude faster than BLAST. *Bioinformatics* **26**:2460-2461.
- Frisullo, S., and A. Carlucci. 2011. Minor fungal diseases of olive. Pages 291-304 in L. Schena, G. E. Agosteo, and S. O. Cacciola, editors. *Olive diseases and Disorders*. Transworld Research Network, India
- Gonzalez, A., J. C. Clemente, A. Shade, J. L. Metcalf, S. Song, B. Prithiviraj, B. E. Palmer, and R. Knight. 2011. Our microbial selves: what ecology can teach us. *EMBO reports* **12**:775-784.
- Grabowski, M. 2007. The study of new fungus species causing apple sooty blotch. *Folia Horticulturae* **19**.

- Graniti, A., S. Frisullo, A. M. Pennisi, and G. M. Di San Lio. 1993. Infections of *Glomerella cingulata* on olive in Italy. *EPPO Bulletin* **23**:457-465.
- Ippolito, A., L. Schena, I. Pentimone, and F. Nigro. 2005. Control of postharvest rots of sweet cherries by pre-and postharvest applications of *Aureobasidium pullulans* in combination with calcium chloride or sodium bicarbonate. *Postharvest Biology and Technology* **36**:245-252.
- Kim, Y. 2014. First report of a new postharvest rot in sweet cherries caused by *Aureobasidium pullulans*. *Plant disease* **98**:424-424.
- Kõljalg, U., K. H. Larsson, K. Abarenkov, R. H. Nilsson, I. J. Alexander, U. Eberhardt, S. Erland, K. Høiland, R. Kjølner, and E. Larsson. 2005. UNITE: a database providing web-based methods for the molecular identification of ectomycorrhizal fungi. *New Phytologist* **166**:1063-1068.
- Li, W., Y. Xiao, C. Wang, J. Dang, C. Chen, L. Gao, J. C. Batzer, G. Sun, and M. L. Gleason. 2013. A new species of *Devriesia* causing sooty blotch and flyspeck on Rubber Trees in China. *Mycological Progress* **12**:733-738.
- Lindahl, B. D., R. H. Nilsson, L. Tedersoo, K. Abarenkov, T. Carlsen, R. Kjølner, U. Kõljalg, T. Pennanen, S. Rosendahl, and J. Stenlid. 2013. Fungal community analysis by high-throughput sequencing of amplified markers—a user's guide. *New Phytologist* **199**:288-299.
- Lindow, S. E., and M. T. Brandl. 2003. Microbiology of the phyllosphere. *Applied and Environmental Microbiology* **69**:1875-1883.
- Lozupone, C., and R. Knight. 2005. UniFrac: a new phylogenetic method for comparing microbial communities. *Applied and Environmental Microbiology* **71**:8228-8235.
- Mechaber, W. L., D. B. Marshall, R. A. Mechaber, R. T. Jobe, and F. S. Chew. 1996. Mapping leaf surface landscapes. *Proceedings of the National Academy of Sciences* **93**:4600-4603.
- Merget, B., C. Koetschan, T. Hackl, F. Förster, T. Dandekar, T. Müller, J. Schultz, and M. Wolf. 2012. The ITS2 database. *Journal of Visualized Experiments* **61**:3806.
- Moral, J., K. Bouhmidi, and A. Trapero. 2008. Influence of fruit maturity, cultivar susceptibility, and inoculation method on infection of olive fruit by *Colletotrichum acutatum*. *Plant disease* **92**:1421-1426.
- Moral, J., R. de Oliveira, and A. Trapero. 2009. Elucidation of the disease cycle of olive anthracnose caused by *Colletotrichum acutatum*. *Phytopathology* **99**:548-556.
- Morgan, D., and T. Michailides. 2004. First report of melting decay of 'red globe' grapes in California. *Plant disease* **88**:1047-1047.
- Mosca, S., M. G. Li Destri Nicosia, S. O. Cacciola, and L. Schena. 2014. Molecular analysis of colletotrichum species in the carposphere and phyllosphere of olive. *PloS one* **9**:e114031.
- Nelson, S. (2008). Sooty molds. *Plant Disease* **52**: 1–6.
- Nigro, F., and M. Ferrara. 2011. Olive Cercosporiosis. Pages 247-258 in L. Schena, G. E. Agosteo, and S. O. Cacciola, editors. *Olive diseases and Disorders*. Transworld Research Network, India

- Nigro, F., A. Ippolito, P. Gallone, G. Romanazzi, P. Carmignano, and G. Laccone. 2002. Cercosporiosis of olive in Apulia and attempts to control the disease. Pages 773-776 in IV International Symposium on Olive Growing 586.
- Nilsson, R. H., M. Ryberg, K. Abarenkov, E. Sjökvist, and E. Kristiansson. 2009. The ITS region as a target for characterization of fungal communities using emerging sequencing technologies. *FEMS Microbiol Lett* **296**:97-101.
- Nilsson, R. H., M. Ryberg, E. Kristiansson, K. Abarenkov, K.-H. Larsson, and U. Kõljalg. 2006. Taxonomic reliability of DNA sequences in public sequence databases: a fungal perspective. *PLoS One* **1**:e59.
- Oliveros, J. C. 2007. VENNY. An interactive tool for comparing lists with Venn Diagrams.
- Orgiazzi, A., V. Bianciotto, P. Bonfante, S. Daghino, S. Ghignone, A. Lazzari, E. Lumini, A. Mello, C. Napoli, S. Perotto, A. Vizzini, S. Bagella, C. Murat, and M. Girlanda. 2013. 454 pyrosequencing analysis of fungal assemblages from geographically distant, disparate soils reveals spatial patterning and a core mycobiome. *Diversity* **5**:73-98.
- Ottesen, A. R., A. G. Peña, J. R. White, J. B. Pettengill, C. Li, S. Allard, S. Rideout, M. Allard, T. Hill, and P. Evans. 2013. Baseline survey of the anatomical microbial ecology of an important food plant: *Solanum lycopersicum* (tomato). *BMC Microbiology* **13**:114.
- Pappas, A. 1993. *Mycocentrospora cladosporioides* on olive in Greece. *EPPO Bulletin* **23**:405-409.
- Pinto, C., D. Pinho, S. Sousa, M. Pinheiro, C. Egas, and A. C. Gomes. 2014. Unravelling the diversity of grapevine microbiome. *PloS one* **9**:e85622.
- Pruesse, E., C. Quast, K. Knittel, B. M. Fuchs, W. Ludwig, J. Peplies, and F. O. Glöckner. 2007. SILVA: a comprehensive online resource for quality checked and aligned ribosomal RNA sequence data compatible with ARB. *Nucleic Acids Research* **35**:7188-7196.
- Reeder, J., and R. Knight. 2010. Rapid denoising of pyrosequencing amplicon data: exploiting the rank-abundance distribution. *Nature methods* **7**:668.
- Ruano-Rosa, D., L. Schena, A. Ippolito, and C. López-Herrera. 2007. Comparison of conventional and molecular methods for the detection of *Rosellinia necatrix* in avocado orchards in southern Spain. *Plant pathology* **56**:251-256.
- Schena, L., G. E. Agosteo, and S. O. Cacciola. 2011. Olive diseases and disorders. Transworld Research Network, Kerala, India
- Schena, L., F. Nigro, I. Pentimone, A. Ligorio, and A. Ippolito. 2003. Control of postharvest rots of sweet cherries and table grapes with endophytic isolates of *Aureobasidium pullulans*. *Postharvest Biology and Technology* **30**:209-220.
- Schoch, C. L., B. Robbertse, V. Robert, D. Vu, G. Cardinali, L. Irinyi, W. Meyer, R. H. Nilsson, K. Hughes, and A. N. Miller. 2014. Finding needles in haystacks: linking scientific names, reference specimens and molecular data for Fungi. *Database* **2014**:bau061.
- Schoch, C. L., K. A. Seifert, S. Huhndorf, V. Robert, J. L. Spouge, C. A. Levesque, W. Chen, C. Fungal Barcoding, and L. Fungal Barcoding Consortium Author. 2012. Nuclear ribosomal internal transcribed spacer (ITS) region as a universal DNA

- barcode marker for Fungi. Proceedings of the National Academy of Sciences **109**:6241-6246.
- Seifert, K. A., N. L. Nickerson, M. Corlett, E. D. Jackson, G. Louis-Seize, and R. J. Davies. 2004. *Devriesia*, a new hyphomycete genus to accommodate heat-resistant, cladosporium-like fungi. Canadian Journal of Botany **82**:914-926.
- Sharma, K., and K. Mukerji. 1976. Microbial ecology of *Sesamum orientale* L. and *Gossypium hirsutum* L. C. H. Dickinson and TF Preece. Academic Press, New York:375-390.
- Stockinger, H., M. Kruger, and A. Schussler. 2010. DNA barcoding of arbuscular mycorrhizal fungi. New Phytologist **187**:461-474.
- Talhinhas, P., C. Mota-Capitão, S. Martins, A. Ramos, J. Neves-Martins, L. Guerra-Guimarães, V. Várzea, M. d. C. Silva, S. Sreenivasaprasad, and H. Oliveira. 2011. Epidemiology, histopathology and aetiology of olive anthracnose caused by *Colletotrichum acutatum* and *C. gloeosporioides* in Portugal. Plant pathology **60**:483-495.
- Tamura, K., G. Stecher, D. Peterson, A. Filipowski, and S. Kumar. 2013. MEGA6: Molecular Evolutionary Genetics Analysis version 6.0. Molecular biology and evolution **30**:2725-2729.
- Tashiro, N., M. Noguchi, Y. Ide, and F. Kuchiki. 2013. Sooty spot caused by *Cladosporium cladosporioides* in postharvest Satsuma mandarin grown in heated greenhouses. Journal of General Plant Pathology **79**:158-161.
- Toju, H., A. S. Tanabe, S. Yamamoto, and H. Sato. 2012. High-coverage ITS primers for the DNA-based identification of ascomycetes and basidiomycetes in environmental samples. PloS one **7**:e40863.
- Unterseher, M., D. Peršoh, and M. Schnittler. 2013. Leaf-inhabiting endophytic fungi of European Beech (*Fagus sylvatica* L.) co-occur in leaf litter but are rare on decaying wood of the same host. Fungal Diversity **60**:43-54.
- Wang, X., M. M. Radwan, A. H. Taráwneh, J. Gao, D. E. Wedge, L. H. Rosa, H. G. Cutler, and S. J. Cutler. 2013. Antifungal activity against plant pathogens of metabolites from the endophytic fungus *Cladosporium cladosporioides*. Journal of agricultural and food chemistry **61**:4551-4555.
- Watanabe, M., T. Yonezawa, K.-i. Lee, S. Kumagai, Y. Sugita-Konishi, K. Goto, and Y. Hara-Kudo. 2011. Molecular phylogeny of the higher and lower taxonomy of the *Fusarium* genus and differences in the evolutionary histories of multiple genes. BMC evolutionary biology **11**:322.
- Weir, B., P. Johnston, and U. Damm. 2012. The *Colletotrichum gloeosporioides* species complex. Studies in Mycology **73**:115-180.
- Wijayawardene, N. N., P. W. Crous, P. M. Kirk, D. L. Hawksworth, S. Boonmee, U. Braun, D.-Q. Dai, M. J. D'souza, P. Diederich, and A. Dissanayake. 2014. Naming and outline of Dothideomycetes–2014 including proposals for the protection or suppression of generic names. Fungal Diversity:1-55.
- Zalar, P., C. Gostinčar, G. De Hoog, V. Uršič, M. Sudhaddham, and N. Gunde-Cimerman. 2008. Redefinition of *Aureobasidium pullulans* and its varieties. Studies in Mycology **61**:21-38.

Chapter III. Metagenomic analysis of fungal diversity on strawberry plants and effect of management practices on the fungal community structure of aerial organs

Abstract

Metabarcoding, defined as NGS sequencing of amplicons of the ITS2 region (DNA barcode), was used to identify the composition of the fungal community on different strawberry organs *i.e.* leaves, flowers, immature and mature fruit grown on a farm using disease and insect control practices that entailed the routine use of various chemical pesticides. Strawberry plants supported a high diversity of microbial organisms, although two genera, *Botrytis* and *Cladosporium*, were more highly abundant than all the other identified genera, representing 70-99% of the relative abundance (RA) of all detected sequences. Leaves supported the most diverse fungal community, followed by flowers. Some of the detected taxa are known to be resident on strawberries. On the other hand, several plant pathogens of other plant species, that would not be intuitively expected to be present on strawberry plants, were detected while some common strawberry pathogens were less evident or not detected. The study also indicated a high degree of diversity in the ITS sequence within certain taxonomic groups. The interruption of chemical treatments for one month resulted in a significant modification in the structure of the fungal community and the RA of specific genera. The overwhelming abundance of just two fungal genera may be a consequence of the intensive chemical pesticide schedule commonly utilized in the investigated farm. The differences observed between treated samples and samples untreated for a month indicate, however, that natural fungal communities exhibit a great plasticity whose composition can quickly change in response to modified environmental conditions.

Introduction

Strawberry (*Fragaria x ananassa*) is an important horticultural crop worldwide that is produced both conventionally and organically in open fields, greenhouses, and in plastic covered tunnels. World production of strawberry fruit reached 739,622.443 tons in 2013 (FAOSTAT Date: Mon Jun 01 20:45:19 CEST 2015). Strawberries are susceptible to several fungal pathogens and thus require the use of fungicide treatments to prevent

economic losses. A major disease of strawberry is grey mold incited by *Botrytis cinerea*. Other fungal pathogens affecting strawberries, such as *Podosphaera aphanis* (powdery mildew), *Colletotrichum acutatum* (anthracnose), *Phytophthora cactorum* (leather rot, crown rot), *Phytophthora fragariae var. fragariae* (red stele root rot), and *Verticillium dahliae*, *Verticillium albo-atrum* (Verticillium wilt) are also considered important diseases that require chemical control measures (Maas 1998, Agrios 2005).

Although studies of the strawberry fungal or bacterial community have been previously reported (Jensen et al. 2013, Sylla et al. 2013b), these previous studies were based on isolation/culture techniques and/or did not study the full complement of strawberry aerial organs, e.g. leaves, fruit and flowers, at the same time. Furthermore, these studies did not investigate the impact of conventional chemical treatments of fungal populations. Next generation sequencing (NGS), together with the evolution of bioinformatics tools, have made it easier to comprehensively study microbial communities on any type of matrix, including plants. To date, however, relatively few researchers have utilized these new technologies in relation to plant pathology or plant microbial ecology studies in general (Unterseher et al. 2013, Pinto et al. 2014).

The ITS regions of the ribosomal DNA (rDNA) (ITS) are the most utilized DNA barcodes in fungal metabarcoding. These regions can be easily amplified and sequenced with universal primers and the corresponding ITS sequence data is highly represented in GenBank and other databases (Kõljalg et al. 2005, Pruesse et al. 2007). The choice of utilizing either ITS1 or ITS2 is optional since these regions share many properties, and enable similar levels of discrimination (Bazzicalupo et al. 2013). Since ITS2 is less variable in length and lacks the problem of co-amplification of a 5' SSU intron, ITS2 sequence data is generally easier to work with. Additionally, ITS2 sequence data are better represented than ITS1 in databases (Nilsson et al. 2009).

The objective of the present study was to identify the composition of the fungal community on different strawberry organs and determine if these fungal communities were affected by the application of standard chemical treatments used in the management program of conventional strawberry production systems. The identification and quantification of endophytic and epiphytic microflora present in and on plants will provide information that can facilitate an understanding of the complex interactions that take place between plants and resident fungal microflora, including pathogenic and beneficial

species. Such data may also have potential practical implications for strawberry disease management, particularly of economically important pathogens, like *B. cinerea*, which can impact fruit yield and quality at both pre- and post-harvest stages of production. Documenting the relative abundance (RA) of fungal pathogens on different plant organs will greatly assist in understanding the etiology and control of complex plant diseases. Moreover, the presence of less serious secondary pathogens or rare taxa is not well documented due to either their low economic impact or the difficulty associated with their isolation and identification using conventional culturing methods. Since chemical treatments are intensively applied in strawberry production it was of interest to better understand the potential effect of these treatments on fungal communities of different strawberry organs.

Materials and Methods

Sampling and DNA extractions

Samples were collected on April 30, 2014 from a strawberry farm in Lamezia Terme, southern Italy (GPS coordinates: 38°49'37.6"N 16°14'23.0"E) as summarized in Table III-1. The strawberry cultivar Rania (Consorzio Italiano Vivaisti, Italy) was grown on raised beds under high tunnels. Samples were collected from 6 different rows, 50 m in length, based on a complete randomized block design. Three of the strawberry rows were subjected to regular schedules of chemical treatments, including the use of ciprodinil and fludioxonil (Switch, Syngenta, Milan, Italy) on April 1, 2014, pyraclostrobin and Boscalid (Signum, BASF, Cesano Maderno, Italy) on April 10, 2014, and fenexamid (Teldor, Bayer Bayer CropScience, Milan, Italy) on April 20, 2014. The other three rows did not receive any chemical treatment during the month of April. To avoid the potential effect of fungicide drift, the untreated rows of strawberry were separated by two rows from the treated strawberry rows. Three biological replicate samples were collected from four different plant parts (leaves, flowers, immature fruit and mature fruit) in each row, resulting in a total of 24 samples. Each biological replicate consisted of a pool of ten leaves, flowers, or immature or mature fruit. Collected samples were kept in sterile plastic bags in a thermally insulated container for approximately 4-5 hours, lyophilized using a freeze dryer (Labconco Corp., Kansas City, MO) and stored at -80°C until the samples were homogenized by grinding under liquid nitrogen. For each sample, three technical replicate DNA extractions were performed from 0.02 g of homogenized tissue using the

DNeasy Plant Mini kit (QIAGEN, Dusseldorf, Germany). The quantity and quality of extracted DNA were determined using a Nanodrop 2000 spectrophotometer (Nano-drop Technologies, Wilmington, DE).

Table III-1. Summary of analyses and results of metagenomic surveys conducted with treated and untreated flowers (TFI and UFI), leaves (TL and UL), immature fruit (TImFr and UImFr), and mature fruit (TMFr and UMFr).

Samples	MIDs	Sequences	OTUs (Total)*	OTUs (200)**	Shannon diversity	Chao1 estimate
TFI	MID1,2,3	6268	71	12.1	1.53	23.17
UFI	MID26,28,40	8207	129	15.8	2.10	46.48
TImFr	MID4,6,13	53268	129	10.6	1.84	21.09
UImFr	MID30,31,32	42145	109	7.6	1.25	10.41
TL	MID21,22,24	2958	83	20.2	1.94	47.26
UL	MID36,37,38	3886	118	29.5	2.83	64.35
TMFr	MID7,10,16	85879	112	7.9	1.29	10.83
UMFr	MID33,34,35	15553	75	7.9	1.30	11.10

*Total number of detected OTUs;

**Number of OTUS detected with an even sequencing depth of 200 sequences.

Fungal DNA amplification and amplicon library preparation

DNA extracts from all technical replicates were amplified in triplicate using the universal fungal primers ITS3-ITS4 targeting the ITS2 region of the ribosomal DNA (Toju et al. 2012). Both primers were modified to construct fusion primers appropriate for 454 sequencing with adapters sequences A and B, key sequences and multiplex identifiers (MIDs) (<http://www.454.com/>). Twenty-four different MIDs were utilized to label different samples and biological replicates (Table III-1).

PCR reactions were conducted in a total volume of 25 μ l containing 2.5 μ l of 10X reaction buffer, 0.25 μ l of each primer (10 μ M), 0.1 μ l of AccuPrime *Taq* DNA Polymerase High Fidelity (Invitrogen, CA, USA) and 1 μ l of DNA template (10 ng/ μ l). Reactions were incubated in an Eppendorf Mastercycler gradient, (Hamburg, Germany) for 1 min at 94°C

followed by 30 cycles of 30s at 94°C, 30 s at 55°C and 30 s at 68°C. All reactions ended with a final extension of 1 min at 72°C. Amplicons from each biological replicate (3 amplifications per each of the three DNA extractions) were pooled and purified with the Agencourt AMPure XP system (Beckman Coulter, Inc.). The concentration and quality of the purified amplicons was evaluated by agarose gel electrophoresis (Sambrook et al. 1989). Amplicons were sequenced by Macrogen Inc. (Seoul, Korea) using the 454 GS FLX+System (Roche Diagnostics Corporation). Data were deposited in the BioProject database (NCBI) as PRJNA289287 (<http://www.ncbi.nlm.nih.gov/bioproject/289287>).

Data analysis and statistics

The bioinformatics pipeline, QIIME v. 1.8 (Caporaso et al. 2010) was used to process and analyze the obtained sequence data. Using default parameters, preliminary processing of the data included de-multiplexing and quality filtering with a minimum quality score of 25, a minimum/maximum length ratio of 200/1000, and a maximum number of homopolymer bases of 6. Sequences were denoised using the denoise wrapper (Reeder and Knight 2010) and the ITS2 region was extracted using ITSx application (Bengtsson-Palme et al. 2013). Chimeric sequences were identified and filtered using USEARCH 6.1 (Edgar 2010). The most abundant sequences were selected as representative sequences and used in determining the Operational Taxonomical Units (OTUs) and taxonomy assignments. OTUs were picked using the UNITE dynamic database released on 10.09.2014 (<http://unite.ut.ee/>) as a reference database. The same database was also used for taxonomy assignments using BLAST (Altschul et al. 1990) utilizing a similarity threshold of 0.97.

For downstream analysis, the OUT table was rarefied at even depth to reduce biases in sequencing depth. Alpha diversity were calculated using observed species, Shannon and Chao1 estimates and the results were compared based on a two-sample t-test using non-parametric (Monte Carlo) methods and a default number (999) of Monte Carlo permutations. The results were visualized in boxplots figures.

Weighted and unweighted UniFrac metrics were utilized to evaluate β -diversity (Lozupone and Knight 2005). A distance-based redundancy analysis (db-RDA) as implemented in Qiime 1.8 was utilized to relate the fungal community composition to sample types and to evaluate the effect of sample type (plant part) and treatment (fungicide applications vs. no fungicide applications), using weighted and unweighted UniFrac

distance matrices. Additionally, a Monte Carlo permutation test was used to determine experimental variables significantly contributing to the observed variance in fungal communities.

Identification of fungal taxa

To confirm the accuracy of QIIME taxonomic assignments, sequences associated with each OTU within each identified fungal genus with a minimum RA of 0.1% were extracted and compared with validated barcode sequences. Extracted sequences were introduced in [ElimDupes](http://hcv.lanl.gov/content/sequence/ELIMDUPES/elimdupes.html) (<http://hcv.lanl.gov/content/sequence/ELIMDUPES/elimdupes.html>) to detect identical sequences and determine their frequency. Unique representative sequences, defined as sequence types (STs) (Cf. Chapter II, Prigigallo et al. 2015), were analyzed along with genetically closely related reference sequences of the same genus to determine their phylogenetic collocation and enable their identification with the highest possible level of accuracy. To this aim, local databases of validated reference sequences were created with priority given to sequences from specific recent taxonomic studies. In cases where no matches were found in the reference sequence from selected published articles, more closely related sequences were selected using BLAST searches of GenBank. The reliability of the latter sequences was evaluated based on different parameters, including the consistency of sequences from different sources, available details on sequenced isolates, and year of publication, giving priority to more recent items. Identical reference sequences were included in the local database when representative of different accepted species.

For each genus, STs identified in the present study and reference sequences were aligned using MUSCLE and introduced into MEGA for phylogenetic analysis utilizing the Maximum Likelihood method (Tamura et al. 2013). Analyses were performed with 1000 bootstrap replications.

Results

General sequencing analysis

After quality evaluations (length trimming, denoising, ITS2 extraction, and chimeric sequence exclusion), a total of 218,164 high quality sequences were recovered and assigned to 316 OTUs (Table III-1). The rarefaction analysis indicated that the

sequencing depth had been saturated for most of the samples indicating that the great majority of OTUs had been detected (Fig. III-1).

Alpha diversity and richness

The total number of OTUs detected in individual samples varied from 71 to 129. Considering OTUs detected with an even sequencing depth of 200 sequences a higher number of OTUs were detected in leaves, followed by flowers, whereas lower values were obtained for immature and mature fruit (Table III-1). In agreement with the OTU data, the Shannon's Diversity Index and the Chao1 estimate, both of which are based on alpha diversity indexes, achieved their highest values in leaf samples, followed by flowers and fruit, with similar values for immature and mature fruit (Table III-1; Fig. III-2).

These results were confirmed by a two-sample t-test with significant differences observed

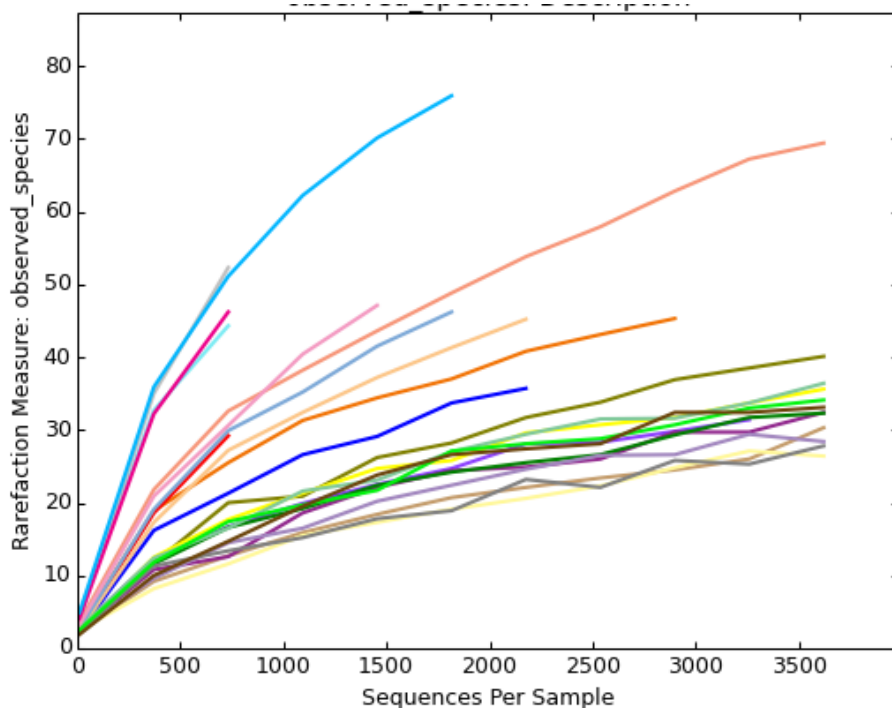


Figure III-1 Rarefaction curves determined for all strawberry samples investigated in the present study.

strawberry plants resulted in a reduction in the alpha diversity of treated leaves and, to a lesser extent, of flowers, compared to untreated samples (Table III-1, Fig. III-2 and Fig. III-3). Fungicidal treatments, however, did not have a significant effect on the fungal between strawberry organs (Fig. III-3). The application of fungicidal sprays to the composition associated with immature and mature fruit (Fig. III-2 and Fig. III-3).

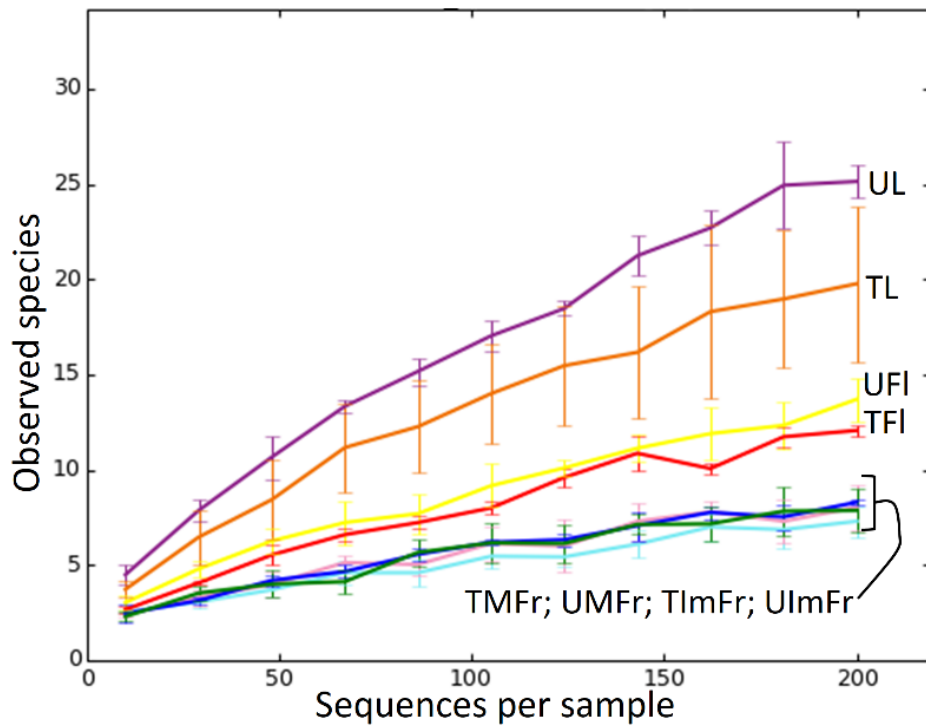


Figure III-2 Rarefaction curves of treated and untreated samples at an even depth of 200 sequences. Analyzed samples comprised treated and untreated flowers (TFI and UFI), immature fruits (TImFr and UImFr), leaves (TL and UL), and mature fruits (TMFr and UMFr). Rarefaction curves were grouped to collapse the 3 biological replicates of each analyzed sample type.

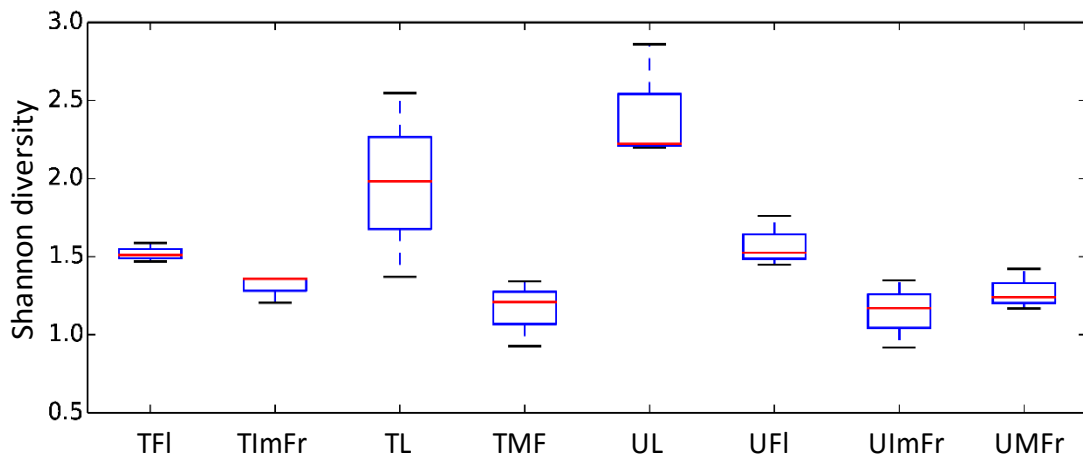


Figure III-3 Boxplots visualizing results of the nonparametric two sample t-test to compare the alpha diversity of treated and untreated samples. Analyzed samples comprised treated and untreated flowers (TFI and UFI), immature fruits (TImFr and UImFr), leaves (TL and UL), and mature fruits (TMFr and UMFr).

Beta diversity analysis

The distance-based redundancy analysis (db-RDA) based on unweighted UniFrac metrics revealed significant differences ($P < 0.01$) in the fungal community of all the investigated strawberry organs (Table III-2). Furthermore, significant differences between treated and untreated samples were revealed for leaves ($P=0.002$) and flowers ($P=0.016$), while no significant differences were observed between treated and untreated samples for immature and mature fruit (Table III-2). Using the weighted UniFrac metrics, however, a significant difference ($P=0.039$) was observed between treated and untreated mature fruit (data not shown).

Table III-2 Statistical significance of the Monte Carlo permutation test (p-value) based on unweighted UniFrac distance matrices, utilized to compare fungal communities associated to different strawberry organs and treated and untreated samples (β -diversity).

Compared categories	Pseudo-F value*	p-value
All categories	1.8353	0.001
All organs	2.2950	0.001
Treated vs. untreated leaves	1.7915	0.002
Treated vs. untreated flowers	1.6419	0.016
Treated vs. untreated mature fruit	1.2647	0.136
Treated vs. untreated immature fruit	1.1818	0.217
Flower vs. mature fruit	1.8448	0.012
Flower vs. leaves	2.0681	0.001
Flower vs. immature fruit	2, 2100	0.012
Mature fruit vs. leaves	1.6489	0.011
Mature fruit vs. immature fruit	1.7326	0.003
Immature fruit vs. leaves	1.3996	0.049

*Pseudo-F value determined according to Lattin et al. 2003

Strawberry fungal community structure and identified taxa

Based on the analysis of the complete ITS2 data set, members of the phylum Ascomycota were dominant in all samples, accounting for 95.6% of the total number of detected sequences, followed by Basidiomycota (3.9%) and unidentified fungi (0.3%).

The incidence of Ascomycota varied in different samples ranging from 88.2% in untreated leaves to 99% in untreated fruit (Fig. III-4A).

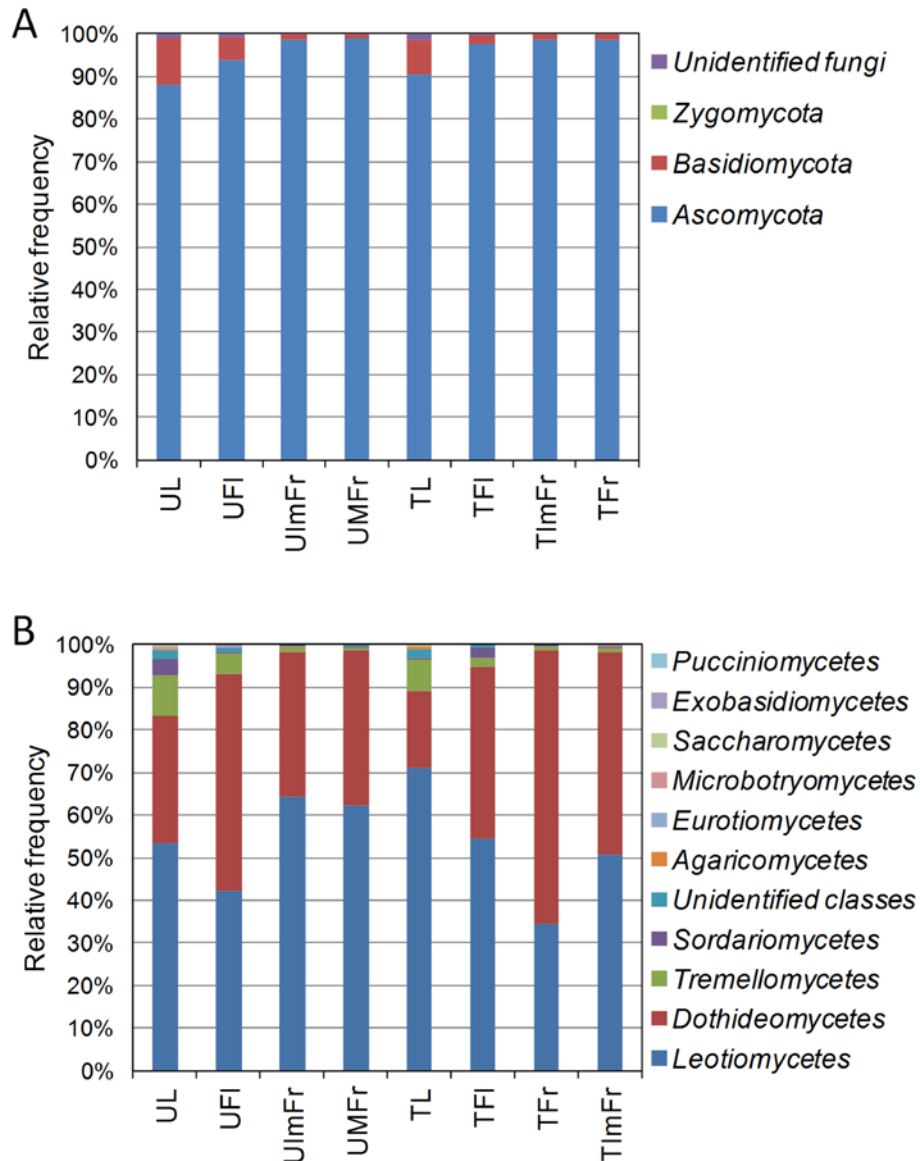


Figure III-4 Relative abundance of different fungal phyla (A) and classes (B). Analyzed samples comprised treated and untreated flowers (TFI and UFI), immature fruits (TIMFr and UIMFr), leaves (TL and UL), and mature fruits (TMFr and UMFr).

Similar values were obtained for treated samples (Fig. III-4A). Basidiomycota accounted for 8.2% and 10.6% of the detected sequences in untreated and treated leaves, respectively, and were much less abundant in other organs (immature and mature fruit). Sequences ascribed to non-identified fungi had similar levels of abundance in all samples.

Ascomycota sequences were largely identified as members of the classes Leotiomyces and Dothideomyces (54.0% and 40.2%, respectively), followed by Tremellomyces (3.40%) and Sordariomyces (1%) (Fig. III-4B). Leotiomyces were more abundant on leaves (61%) than on fruit (38.5%). In contrast, members of the class Dothideomyces were more abundant on fruit (60%) than on leaves (38.5%). Tremellomyces was the predominant class of Basidiomycota in all analyzed samples. Members of this class were more highly represented on leaves and flowers (8.5 and 3.5%, respectively) than on immature and mature fruit (1.2 and 1%, respectively).

The most abundantly detected genera were *Botrytis* spp. and *Cladosporium* spp. with the percentage of reads ranging from 70% in leaves to 99% in fruit. The genus *Botrytis* was the most abundant with an average RA of 53.8% (Fig. III-5). The phylogenetic analysis of sequences of this genus revealed high genetic diversity with 12 different sequence types (STs) (Fig. III-6). All STs clustered with reference sequences of *B. cinerea* and other closely-related species of the same genus with identical or very similar ITS2 sequences (Fig. III-6). This clustering of sequences did not enable the identification of STs at the level of species but confirmed the identification of the genus since sequences were clearly separated from those of the related genera *Monilinia* and *Sclerotinia*.

The genus *Cladosporium* had a RA of 37.80% and was most abundant on treated fruit (Fig. III-5). The phylogenetic analyses of its sequences enabled the identification of 29 different STs which were related to the *Cladosporium cladosporoides* complex (19 STs) and to the species *Cladosporium sphaerospermum* (10 STs). STs related to the *C. cladosporoides* complex were much more abundant representing 97.2% of the total *Cladosporium* spp. reads. The application of fungicides caused a significant reduction in the RA of *Botrytis* spp. and an increase in the RA of *Cladosporium* spp. on mature and immature fruit (Fig. III-5). This trend, however, was not evident in flowers or leaves (Fig. III-5).

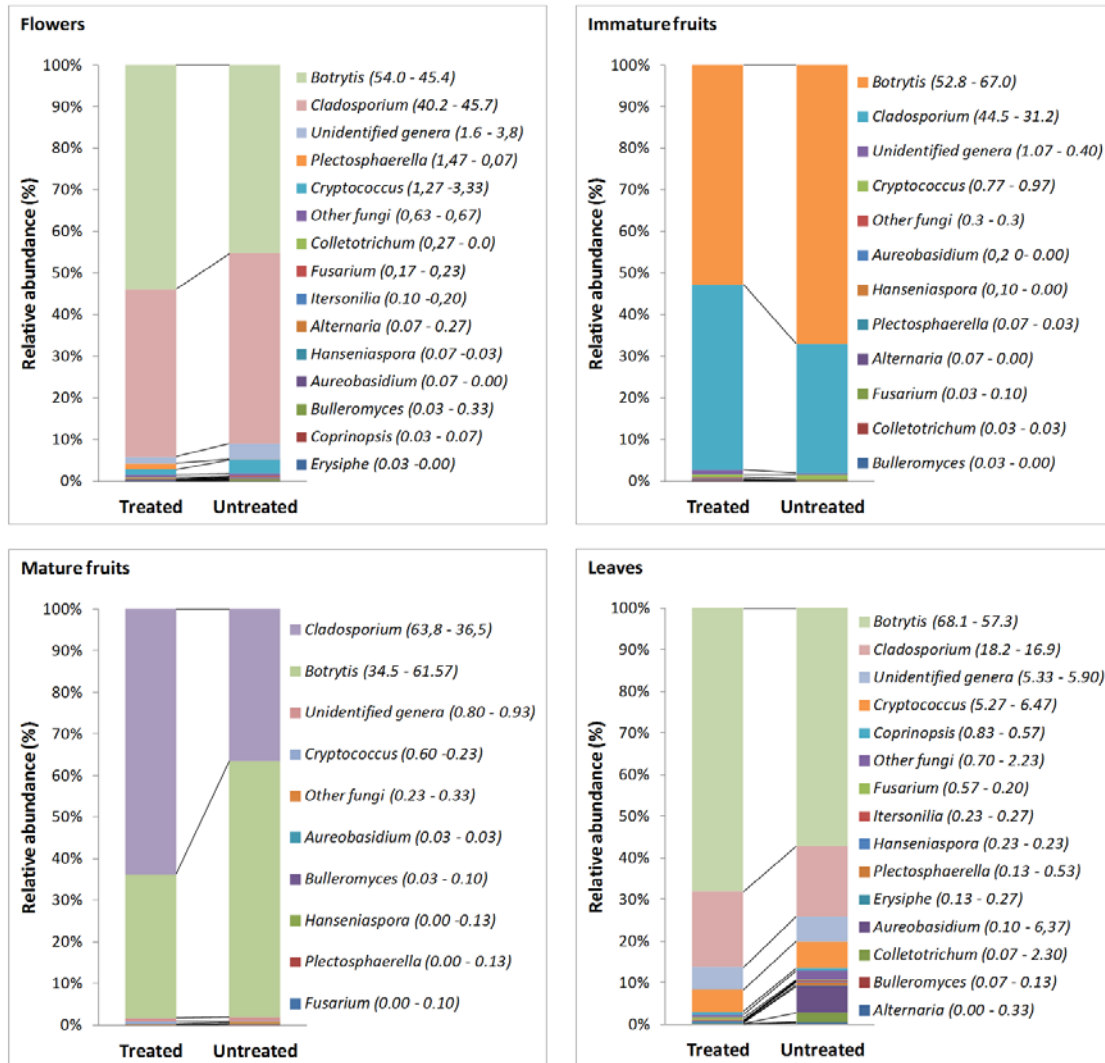


Figure III-5 Relative abundance of the fungal genera in treated and untreated strawberry flowers, leaves and immature and mature fruits. The exact relative abundance (percentage) of each genus in treated (left) and untreated (right) samples is reported between parentheses

Cryptococcus spp. was the third most abundant genus detected. It was least abundant on mature and immature fruit and more abundant on flowers and leaves (Fig. III-5). Fungicide treatments reduced the RA of *Cryptococcus* spp. in leaves from 6.47% to 5.27% and from 3.33% to 1.27% in flowers. A phylogenetic analysis of STs of this genus revealed a high genetic diversity with the identification of six different species (*C. victoriae*, *C. tephrensis*, *C. heimaeyensis*, *C. oeirensis*, *C. stepposus*, and *C. flavescens*) and two phylotypes clustering within the genus but not related to currently known species (Fig. III-6).

Other genera represented a limited portion of the detected fungal populations. The genus *Aureobasidium*, represented by 3 STs, all identified as *A. pullulans*, was most abundant on leaves and less abundant on other organs. The leaf population of *A. pullulans* was significantly influenced by the fungicide treatments as evidenced by its RA on untreated leaves (6.37%) vs. treated leaves (0.1%). Other taxa detected in the investigated organs included *Plectosphaerella cucumerina*, two different phlotypes of *Fusarium* spp. closely related to *F. avenaceum* and *F. equiseti*, *Hanseniaspora uvarum* and *Bulleromyces albus* (Fig. III-5; Fig. III-6). Sequences related to Alternaria Sect. Alternata (a single ST) and sequences related to the genus Colletotrichum identified as *C. acutatum s.str.* (a single ST) and *C. lini* or *C. americae-borealis* (two STs) were detected in flowers, immature fruit, and leaves. Finally, *Coprinopsis cinerea* (a single ST), *Itersonia perplexans* (two STs) and a ST clustering within the *Erysiphe aquilegiae* clade, were detected in flower and leaf samples (Fig. 5; Fig. III-6).

Fungal genera with a very low RA cumulatively represented a RA ranging from 0.30% (immature fruit) and 2.23% (untreated leaves). The incidence of these fungi decreased to 0.7% in treated leaves. Finally, unidentified fungal genera had a cumulative RA ranging from 0.4% (untreated immature fruit) and 5.9% (untreated leaves).

Discussion

Results of the present study indicate that a wide variety of fungal taxa inhabit strawberry plants. Some of them are known to be resident while others have not been previously reported on this host plant. The study also demonstrated that the composition and diversity of fungal communities, as evidenced by their alpha and beta diversity, vary significantly on different plant organs (leaf, flower and mature and immature fruit). The alpha diversity analysis indicated that leaves supported the most diverse fungal community, followed by flowers. This ranking of diversity is in agreement with a previous study conducted in olive trees, in which the fungal community of the same organs was analyzed (Cf. Chapter II). In that study, it was suggested that the greater diversity present on leaves may be due to the higher surface/volume ratio of leaves relative to fruit. Additionally, leaves do not change as much as fruit in shape, color, and nutrient composition as they develop. Leaves also have a longer life span relative to flowers that quickly develop into fruit.

Interestingly, the fungal population associated with strawberries was characterized by the prevalence of two genera, *Botrytis* and *Cladosporium*, while other fungi were exhibited at a very low RA. The prevalence of *Botrytis* spp. over other fungal genera was somewhat unexpected considering its low frequency of isolation in previous studies which have investigated strawberry fungal populations using conventional cultural methods based on the plating of the fruit washing water on a nutritive media (Lima et al. 1997). Sylla et al. (2013), using 454 pyrosequencing indicated that while *Botrytis* was not detected initially, its RA increased to as high as 37% in subsequent samplings and its presence was unaffected by several applications of a variety of biocontrol agents (Sylla et al. 2013b). The high RA of *Botrytis* spp. detected in the present study in asymptomatic leaves, flowers, and mature and immature fruit may be partially explained by the high incidence of gray mold infections occurring in the investigated field during the sampling period (Schena L., personal communication). These infections may have served as a source for the contamination of healthy tissues with *B. cinerea* conidia. Furthermore *B. cinerea* readily colonizes floral tissues and establishes latent infections which would be difficult to detect by plating the wash water from fruit because the infections are internal and the fungus is dormant (Sanzani et al. 2012, Jensen et al. 2013). On the other hand, the heterokaryotic nature of *B. cinerea* that can possess an average of 3–6 nuclei per cell (Buckley et al. 1966) may contribute to overestimating the fungal population using molecular detection

methods. Although, *Botrytis* was the main target organisms of the chemical treatments that were applied, the RA of *Botrytis* on leaves was unaffected by the chemical applications, though its presence was reduced on both immature and mature fruit. These results indicate that there is a large potential reservoir of *Botrytis* present that at any point in time could serve as source of inoculum for infecting fruit.

The sharp prevalence of two genera, *Botrytis* and *Cladosporium*, over all other detected fungi may be related to the intensive cultivation system utilized in the investigated field and the use of fungicides. Chemical applications have been reported to have a significant impact on non-target organisms and reduce the overall genetic diversity (Singh et al. 2015). Indeed, the phyllosphere population determined in the present study is quite different compared to that determined in a previous investigation focusing on conventionally-grown strawberries, where several fungal genera had a very high RA (Sylla et al. 2013b). Supporting this previous study, beta diversity analyses in the current study revealed that exposure of strawberries to chemical treatments had a significant effect on the fungal composition of leaves and flowers. The same result was not evident for either immature or mature fruit. The fact that a significant difference was observed between treated and control fruit when weighted metrics were used, however, may indicate that fungicide applications did have an effect on the RA of specific fungal taxa that were present but not on the overall composition of the microbial community present on fruit. It is important to note that the present study was conducted in a strawberry field in a farm using conventional practices and that the untreated plants did not receive the conventionally-used fungicidal treatments for only a month. This fact underscores the plasticity of fungal communities and their ability to establish a new equilibrium as soon as chemical applications are interrupted.

A more in-depth analysis of the DNA sequences of the identified OTUs provided additional information at the genus and species level (Cf. Chapter II). The analysis of sequences related to *Botrytis* spp. was problematic and made their identification at the level of species difficult. This was due to the existence of several accepted species with very similar or identical ITS sequences (Staats et al. 2005). Considering the relevance of grey mold for strawberries, however, it can be likely assumed that the majority of the detected sequences belong to the species *B. cinerea*. The detection of 12 distinct STs suggests that there is a great deal of strain diversity present within *B. cinerea* which perhaps is a reflection of its wide distribution and host range (Elad et al. 2004). Most of

this genetic diversity is likely to remain undetected using traditional culturing methods since, in the present study, a single ST was predominant above the sequences of the other STs of *Botrytis* and represented 83.8% of the total sequences obtained for this genus. The heterokaryotic nature of *B. cinerea*, however, should also be considered as potentially contributing to the detected level of molecular diversity (Buckley et al. 1966).

The genus, *Cryptococcus*, was the most abundant yeast identified on strawberry plant organs in the present study, being represented by six species and two non-identified taxa. The abundance of *Cryptococcus* was affected by the application of the chemical treatments and was significantly reduced in both leaves and flowers. This finding is in agreement with the high RA of this genus found in the phyllosphere of organic strawberries where it was the most abundant fungus (Sylla et al. 2013b). Interestingly, *Cryptococcus* has often been isolated from fruit washings and has been identified as a good biocontrol agent for the management of postharvest diseases (Liu et al. 2013). Due to its potential role as an effective biocontrol agent, more studies should be conducted on its role in inhibiting plant pathogens in natural plant microbial communities. Perhaps methods could be found to support and/or enhance its presence. The yeast-like fungus, *A. pullulans*, has also been isolated from plant washings and used to control postharvest diseases on a number of different commodities (Lima et al. 1997, Schena et al. 2003, Ippolito et al. 2005). The overall detected population of *A. pullulans* in the present study was lower than expected considering previous studies that focused on the phylloplane of strawberry and other plant species (Sylla et al. 2013a, Cf. Chapter II). In the current study, a high population of *A. pullulans* was only detected in untreated leaves (6.37%) but was significantly impacted by the fungicide application, being reduced in abundance down to 0.1% in treated leaves. As suggested for *Cryptococcus*, the use of chemicals in the strawberry production system examined in the present study may be responsible for the overall low abundance of *A. pullulans* and indicate the need for further studies to better understand the role of *A. pullulans* in strawberry fungal communities, as well as its potential use to prevent the establishment of infections by plant pathogens.

The detection of *Colletotrichum* species was also expected considering their major role as causal agents of strawberry anthracnose. However, among the several *Colletotrichum* species responsible for strawberry anthracnose (Damm et al. 2012, Baroncelli et al. 2015), only *C. acutatum* s. str. was detected and at a very low percentage. On the other hand, molecular analyses revealed the presence of sequences associated to

two different *Colletotrichum* species having identical ITS sequences (Damm et al. 2014). Both species have been reported on flax and several other hosts but not on strawberry (Damm et al. 2014). A single ST related to *Alternaria Sect. Alternaria* was detected with a low RA in flowers, immature fruit, and leaves. The genus *Alternaria* was found to be quite abundant in the phyllosphere of organic strawberries (Sylla et al. 2013a) and in the carposphere of conventional and organic strawberries (Jensen et al. 2013). Jensen (2013) hypothesized that this fungus may have a role as a mycotoxin producer in strawberries. Two different phylotypes of *Fusarium* spp. were detected in the present study and were found to be genetically related to *F. equiseti* and *F. avenaceum*, although the available data on genetic diversity within the ITS2 region did not enable the unequivocal identification of these two species. *Fusarium equiseti* has reported to cause damping-off and root rot in cucurbits (Chehri et al. 2011) while *F. avenaceum* is a cosmopolitan plant pathogen with a wide host range (Crous et al. 1995). The role of these microorganism in the strawberry phylloplane is unknown, but the potential production of mycotoxins cannot be completely excluded. Interestingly, our study revealed the presence of the biotrophic pathogen, *Erysiphe aquilegiae*, on leaves and flowers. This fungus is the causal agent of powdery mildew in species of the family Ranunculaceae (Cunnington et al. 2004) and in *Catharanthus roseus* (Liberato and Cunnington 2006) but has never been associated with strawberries. Surprisingly, *Sphaerotheca macularis* f. sp. *fragariae*, the causal agent of strawberry powdery mildew (Amsalem et al. 2006) was not detected. Among the other detected fungal species, *P. cucumerina* is a known pathogen of melon and the causal agent of fruit and collar rot in several plant species but not strawberry (Carlucci et al. 2012). *Itersonilia perplexans* is the causal agent of foliar and petal blight on several cut-flowers (McGovern et al. 2006). *Hanseniospora uvarum* has been isolated from the surface of fruits such as grapes and is known for its role in natural fermentation and for the secretion of a killer toxin against *Saccharomyces cerevisiae* (Radler et al. 1990).

Conclusion

The results of the present study indicate that strawberry plants support a high diversity of fungal organisms, although two genera, *Botrytis* and *Cladosporium*, were more highly abundant than all the other identified genera, representing 70 – 99% of the RA. Several plant pathogens of other plant species and that would not be intuitively expected to be present on strawberry plants were among the detected genera and species while some common strawberry pathogens were less evident or not detected. The study also indicated

a high degree of diversity in the ITS sequence within certain taxonomic groups. Both these findings raise significant questions about the natural fungal community present on plants and about our classification of plant pathogens, and fungal species in general. One could raise the question whether or not some taxonomic units are correctly assigned or alternatively if our understanding of the genetic identity of an infectious agent is comprehensive enough to account for some level of genetic diversity within a species. These questions are highlighted in the present study by the genetic diversity observed within *B. cinerea*, and the genetic (sequence-based) presence of species that are pathogenic on hosts other than strawberry, while there was a failure to identify the strawberry-equivalents of the same disease-causing organisms. These questions will only be resolved as more metagenomic data about natural microbial populations on plants accumulates and our understanding of genetic diversity within microbial species increases. Furthermore, a major finding of the present study indicates, and is supported by a previously (Cf. Chapter II), that the diversity and composition of the fungal community differs on different organs within a species. Lastly, our data suggest that applications of fungicides may have an important impact on the fungal community by affecting the RA of the individual members of the community. Although additional investigations are needed, the overwhelming abundance of just two fungal genera may be a consequence of the intensive chemical pesticide schedule commonly utilized in the investigated farm. The differences observed between treated samples and samples untreated for a month, indicate that natural fungal communities exhibit a great plasticity that can quickly change in response to modified environmental conditions.

References

- Agrios, G. 2005. Plant Pathology. 5th eds. New York: Academic Press.
- Altschul, S. F., W. Gish, W. Miller, E. W. Myers, and D. J. Lipman. 1990. Basic local alignment search tool. *Journal of Molecular Biology* **215**:403-410.
- Amsalem, L., S. Freeman, D. Rav-David, Y. Nitzani, A. Sztejnberg, I. Pertot, and Y. Elad. 2006. Effect of climatic factors on powdery mildew caused by *Sphaerotheca macularis* f. sp. *fragariae* on strawberry. *European Journal of Plant Pathology* **114**:283-292.
- Baroncelli, R., A. Zapparata, S. Sarrocco, S. A. Sukno, C. R. Lane, M. R. Thon, G. Vannacci, E. Holub, and S. Sreenivasaprasad. 2015. Molecular diversity of anthracnose pathogen populations associated with UK strawberry production suggests multiple introductions of three different *colletotrichum* species. *PLoS one* **10**:e0129140.

- Bazzicalupo, A. L., M. Bálint, and I. Schmitt. 2013. Comparison of ITS1 and ITS2 rDNA in 454 sequencing of hyperdiverse fungal communities. *Fungal Ecology* **6**:102-109.
- Bengtsson-Palme, J., M. Ryberg, M. Hartmann, S. Branco, Z. Wang, A. Godhe, P. De Wit, M. Sánchez-García, I. Ebersberger, F. de Sousa, A. Amend, A. Jumpponen, M. Unterseher, E. Kristiansson, K. Abarenkov, Y. J. K. Bertrand, K. Sanli, K. M. Eriksson, U. Vik, V. Veldre, and R. H. Nilsson. 2013. Improved software detection and extraction of ITS1 and ITS2 from ribosomal ITS sequences of fungi and other eukaryotes for analysis of environmental sequencing data. *Methods in Ecology and Evolution* **4**:914-919.
- Buckley, P. M., V. E. Sjaholm, and N. Sommer. 1966. Electron microscopy of *Botrytis cinerea* conidia. *Journal of Bacteriology* **91**:2037-2044.
- Caporaso, J. G., J. Kuczynski, J. Stombaugh, K. Bittinger, F. D. Bushman, E. K. Costello, N. Fierer, A. G. Pena, J. K. Goodrich, and J. I. Gordon. 2010. QIIME allows analysis of high-throughput community sequencing data. *Nature methods* **7**:335-336.
- Carlucci, A., M. L. Raimondo, J. Santos, and A. J. L. Phillips. 2012. Plectosphaerella species associated with root and collar rots of horticultural crops in southern Italy. *Persoonia : Molecular Phylogeny and Evolution of Fungi* **28**:34-48.
- Chehri, K., B. Salleh, T. Yli-Mattila, K. Reddy, and S. Abbasi. 2011. Molecular characterization of pathogenic *Fusarium* species in cucurbit plants from Kermanshah province, Iran. *Saudi Journal of Biological Sciences* **18**:341-351.
- Crous, P. W., O. Petrini, G. F. Marais, Z. A. Pretorius, and F. Rehder. 1995. Occurrence of fungal endophytes in cultivars of *Triticum aestivum* in South Africa. *Mycoscience* **36**:105-111.
- Cunnington, J. H., A. C. Lawrie, and I. G. Pascoe. 2004. Unexpected ribosomal DNA internal transcribed spacer sequence variation within *Erysiphe aquilegiae sensu lato*. *Fungal Diversity* **16**:1-10.
- Damm, U., P. Cannon, J. Woudenberg, and P. Crous. 2012. The *Colletotrichum acutatum* species complex. *Studies in Mycology* **73**:37-113.
- Damm, U., R. O'Connell, J. Groenewald, and P. Crous. 2014. The *Colletotrichum destructivum* species complex—hemibiotrophic pathogens of forage and field crops. *Studies in Mycology* **79**:49-84.
- Edgar, R. C. 2010. Search and clustering orders of magnitude faster than BLAST. *Bioinformatics* **26**:2460-2461.
- Elad, Y., B. Williamson, P. Tudzynski, and N. Delen. 2004. *Botrytis*: biology, pathology and control. Springer.
- Ippolito, A., L. Schena, I. Pentimone, and F. Nigro. 2005. Control of postharvest rots of sweet cherries by pre-and postharvest applications of *Aureobasidium pullulans* in combination with calcium chloride or sodium bicarbonate. *Postharvest Biology and Technology* **36**:245-252.
- Jensen, B., I. M. B. Knudsen, B. Andersen, K. F. Nielsen, U. Thrane, D. F. Jensen, and J. Larsen. 2013. Characterization of microbial communities and fungal metabolites

- on field grown strawberries from organic and conventional production. *International Journal of Food Microbiology* **160**:313-322.
- Köljalg, U., K. H. Larsson, K. Abarenkov, R. H. Nilsson, I. J. Alexander, U. Eberhardt, S. Erland, K. Høiland, R. Kjølner, and E. Larsson. 2005. UNITE: a database providing web-based methods for the molecular identification of ectomycorrhizal fungi. *New Phytologist* **166**:1063-1068.
- Lattin, J. M., J. D. Carroll, and P. E. Green. 2003. *Analyzing multivariate data*. Thomson Brooks/Cole Pacific Grove, CA.
- Liberato, J., and J. Cunnington. 2006. First record of *Erysiphe aquilegiae* on a host outside the Ranunculaceae. *Australasian Plant Pathology* **35**:291-292.
- Lima, G., A. Ippolito, F. Nigro, and M. Salerno. 1997. Effectiveness of *Aureobasidium pullulans* and *Candida oleophila* against postharvest strawberry rots. *Postharvest Biology and Technology* **10**:169-178.
- Liu, J., Y. Sui, M. Wisniewski, S. Droby, and Y. Liu. 2013. Review: utilization of antagonistic yeasts to manage postharvest fungal diseases of fruit. *International Journal of Food Microbiology* **167**:153-160.
- Lozupone, C., and R. Knight. 2005. UniFrac: a new phylogenetic method for comparing microbial communities. *Applied and Environmental Microbiology* **71**:8228-8235.
- Maas, J. L. 1998. *Compendium of strawberry diseases*. American Phytopathological Society (APS Press).
- McGovern, R., H. Horita, C. Stiles, and T. Seijo. 2006. Host range of *Iterosonilia perplexans* and management of *Iterosonilia* petal blight of China aster. *Plant Health Progress*.
- Nilsson, R. H., M. Ryberg, K. Abarenkov, E. Sjökvist, and E. Kristiansson. 2009. The ITS region as a target for characterization of fungal communities using emerging sequencing technologies. *FEMS Microbiology Letters* **296**:97-101.
- Pinto, C., D. Pinho, S. Sousa, M. Pinheiro, C. Egas, and A. C. Gomes. 2014. Unravelling the diversity of grapevine microbiome. *PloS one* **9**:e85622.
- Prigigallo, M., S. Mosca, S. O. Cacciola, D. Cooke, and L. Schena. 2015. Molecular analysis of *Phytophthora* diversity in nursery-grown ornamental and fruit plants. *Plant pathology*.
- Pruesse, E., C. Quast, K. Knittel, B. M. Fuchs, W. Ludwig, J. Peplies, and F. O. Glöckner. 2007. SILVA: a comprehensive online resource for quality checked and aligned ribosomal RNA sequence data compatible with ARB. *Nucleic Acids Research* **35**:7188-7196.
- Radler, F., M. Schmitt, and B. Meyer. 1990. Killer toxin of *Hanseniaspora uvarum*. *Archives of microbiology* **154**:175-178.
- Reeder, J., and R. Knight. 2010. Rapidly denoising pyrosequencing amplicon reads by exploiting rank-abundance distributions. *Nature methods* **7**:668-669.
- Sambrook, J., E. F. Fritsch, and T. Maniatis. 1989. *Molecular cloning*. Cold spring harbor laboratory press New York.

- Sanzani, S. M., L. Schena, V. De Cicco, and A. Ippolito. 2012. Early detection of *Botrytis cinerea* latent infections as a tool to improve postharvest quality of table grapes. *Postharvest Biology and Technology* **68**:64-71.
- Schena, L., F. Nigro, I. Pentimone, A. Ligorio, and A. Ippolito. 2003. Control of postharvest rots of sweet cherries and table grapes with endophytic isolates of *Aureobasidium pullulans*. *Postharvest Biology and Technology* **30**:209-220.
- Singh, S., R. Gupta, M. Kumari, and S. Sharma. 2015. Nontarget effects of chemical pesticides and biological pesticide on rhizospheric microbial community structure and function in *Vigna radiata*. *Environmental Science and Pollution Research*:1-11.
- Staats, M., P. van Baarlen, and J. A. van Kan. 2005. Molecular phylogeny of the plant pathogenic genus *Botrytis* and the evolution of host specificity. *Molecular Biology and Evolution* **22**:333-346.
- Sylla, J., B. Alsanius, E. Krüger, A. Reineke, M. Bischoff-Schaefer, and W. Wohanka. 2013a. Introduction of *Aureobasidium pullulans* to the Phyllosphere of Organically Grown Strawberries with Focus on Its Establishment and Interactions with the Resident Microbiome. *Agronomy* **3**:704.
- Sylla, J., B. W. Alsanius, E. Krüger, A. Reineke, S. Strohmeier, and W. Wohanka. 2013b. Leaf microbiota of strawberries as affected by biological control agents. *Phytopathology* **103**:1001-1011.
- Tamura, K., G. Stecher, D. Peterson, A. Filipski, and S. Kumar. 2013. MEGA6: Molecular Evolutionary Genetics Analysis version 6.0. *Molecular Biology and Evolution* **30**:2725-2729.
- Toju, H., A. S. Tanabe, S. Yamamoto, and H. Sato. 2012. High-coverage ITS primers for the DNA-based identification of ascomycetes and basidiomycetes in environmental samples. *PloS one* **7**:e40863.
- Unterseher, M., D. Peršoh, and M. Schnittler. 2013. Leaf-inhabiting endophytic fungi of European Beech (*Fagus sylvatica* L.) co-occur in leaf litter but are rare on decaying wood of the same host. *Fungal Diversity* **60**:43-54.

Chapter IV. Analysis of fungal diversity in orange (*Citrus sinensis*) leaves with and without greasy spot disease symptoms

Abstract

Citrus greasy spot (CGS) is a disease of citrus with worldwide distribution and recent surveys have revealed a high level of incidence and severity of the disease in Sicily, southern Italy. The etiology of the disease is still unclear although *Mycosphaerella citri* and other related species are generally considered the most important causal agents. Here we report the use of an amplicon metagenomic approach to investigate the fungal communities associated with citrus leaves with and without CGS symptoms, in order to determine whether or not there is an association between symptoms and the presence of specific fungal taxa. A total of 212 OTUs clustered at 97% from 35,537 high quality chimeric free reads were obtained. Data revealed a dominating presence of the phylum Ascomycota (92.6%) over other fungal phyla. No significant difference was observed between symptomatic and asymptomatic leaves ($P=0.8$) based on an analysis of alpha and beta diversity. Sequences corresponding to the genus *Mycosphaerella* were the most abundant sequences detected in all samples. Surprisingly, none of the species currently reported as causal agents of CGS was detected in the present study. The most abundant sequence type (ST) corresponded to the species *M. fragariae*, which is the causal agent of a common and widespread disease of strawberry. Overall, results indicated that a considerable part of the fungal diversity in citrus leaves is still unknown. Aspects of the interaction among different citrus fungal pathogens and potential beneficial fungi, including rare taxa, are discussed.

Introduction

Citrus greasy spot (CGS) is a disease of citrus with worldwide distribution that affects leaves and less frequently fruits (Mondal and Timmer 2006). Leaf symptoms appear as yellow, dark brown, or black lesions occurring first on the underside of mature citrus leaves. As the lesions develop on the underside of the leaves, they gradually become darker and a corresponding chlorotic spot will appear on the upper leaf surface. These yellow (chlorotic) spots occur when the chlorophyll in infected cells is degraded and not

replaced. Lesions are more yellowish and diffuse on lemons and grapefruit and more raised and darker on tangerines. Effected leaves fall prematurely from the tree in the fall and winter resulting in reduced tree vigor and yield. On grapefruit, small, black, necrotic spots are produced on the fruit surface with the surrounding area retaining a green color, causing a symptom referred to as greasy spot rind blotch. Rind blotch is a significant problem on grapefruits produced for fresh market (Whiteside 1970b, 1972, 1983, Gonghi 1987, Whiteside 1988, Timmer and Gottwald 2000, Mondal and Timmer 2006).

CGS symptoms are limited to the Rutaceae, and all *Citrus* spp. appear to be susceptible to some degree but the disease symptoms are most severe in lemons (*Citrus limon*), grapefruit, and their hybrids (Timmer and Gottwald 2000, Timmer et al. 2000, Mondal and Timmer 2003, 2006). Among sweet oranges, early ripening cultivars are the most susceptible, whereas Valencia-like cultivars show more intense symptoms (Whiteside 1988, Timmer and Gottwald 2000). Whiteside reported symptoms on other genera of the Rutaceae including *Poncirus*, *Fortunella*, *Murraya*, and *Aeglopsis* (Whiteside 1972). The disease has been reported in many citrus growing areas including the United States, Mexico, Cuba, Brazil, Chile, Peru, Venezuela, Argentina, Costa Rica, Egypt, China, Japan, Australia, and Italy but its etiology is not always clear as several fungi can be isolated from CGS lesions and even reproduce the same symptoms (Yamada 1956, Whiteside 1970a, Grasso and Catara 1982, Díaz et al. 1985, Gonghi 1987, Hidalgo et al. 1997, Timmer and Gottwald 2000, Haggag 2012).

CGS symptoms were first reported on citrus in Florida (Fawcett 1915). The cause of this disease was unknown for long time and it was thought to be a consequence of nutritional problems or rust mites *Phyllocoptruta oleivora* (Thompson 1948). The involvement of a fungal species in the genus *Cercospora* was first hypothesized to be a potential causal agent in Japan but then the fungus *Mycosphaerella horii* was isolated from symptomatic leaves (Yamada 1956). *Mycosphaerella citri* was first proposed as the causal agent of CGS by Whiteside (Whiteside 1970b). Ascospores of *M. citri* are produced in pseudothecia in decomposing leaf litter on the ground (Whiteside 1972). When the pseudothecia matures, ascospores are ejected following a wetting period and are dispersed by air (Mondal and Timmer 2006). Since infection occur through stomata, only ascospores deposited on the underside of the leaf germinate and form epiphytic mycelia. The pathogen penetrates into the mesophyll after the formation of an appressoria. Colonization of the leaf occurs very slowly, and symptoms appear only after 45 to 60 days, even on highly

susceptible species under optimal conditions (Mondal and Timmer 2006). Peak periods of ascospore production are closely associated with wet periods (Hidalgo et al. 1997). Leaves on the orchard floor decompose rapidly as rainfall increases and serve as a substrate for pseudothecial development and subsequent ascospore release. Sexual fruiting structures are only produced intermittently on decomposing dead leaves on the orchard floor. Both disease symptoms and the production of ascospore peak in spring and summer (Whiteside 1970b).

Whiteside has stated that “*Although it has been confirmed that the CGS fungus can cause minute black specks on orange rind, all necroses of the guard cells and underlying tissues are not necessarily the result of such infection. Nevertheless, fungal structures as seen in inoculated fruit have been observed in many samples of fruit showing stomatal necrosis; however, it was evident that stomatal penetration by this fungus did not always produce disease symptoms*” (Whiteside 1970c). While CGS symptoms have been observed in several countries, its specific association with *M. citri* has not been fully demonstrated (Baker 2008). Several other fungal species have also been associated with CGS-like symptoms on Citrus and other hosts, such as leaves infected by *Septoria citri*, although Koch's postulates were not yet fully satisfied since the fungus was not re-isolated from artificially induced lesions (Wellings 1981). *Colletotrichum gloeosporioides* is also frequently isolated from diseased leaves (Wellings 1981). More recently, four *Zasmidium* species, including *Z. indonesianum* in Indonesia, *Z. fructicola* and *Z. fructigenum* in China, and *Z. citri-griseum* in many different countries have been isolated from symptomatic leaves of several citrus species (Huang 2015). In Italy, the disease was first reported by Grasso and Catara in 1982 and more recently a high level of incidence and severity has been observed in Sicily. Detailed analyses to determine the etiology, however, have not been performed (Grasso and Catara 1982, Biasi et al. 2012). The aim of the present study was to use a metagenomic approach to investigate fungal communities present in and on citrus leaves with and without CGS symptoms, in order to determine whether or not there was an association between symptoms and specific fungal taxa. The present investigation also provided new information about the fungal communities associated with citrus leaves and contributed to understanding the complex interactions that occur among citrus fungal pathogens and potential beneficial fungi, including rare taxa and uncultivable fungi that are usually missed by traditional isolation and molecular detection methods.

Materials and Methods

Samples were collected from symptomatic and asymptomatic leaves of two sweet orange (*Citrus × sinensis*) cultivars (Tarocco Scirè and Lane Late) and from Tangelo Nova (*Citrus × tangelo*). Symptomatic leaves exhibited symptoms typical of CGSs (Fig. IV-1), and identical to those reported in the literature (Mondal and Timmer 2006). The sampled citrus orchards were located the southwest of Catania (Mineo), Italy (GPS coordinates 37°19'38.5"N 14°41'11.5"E). Leaves were individually collected according to a complete

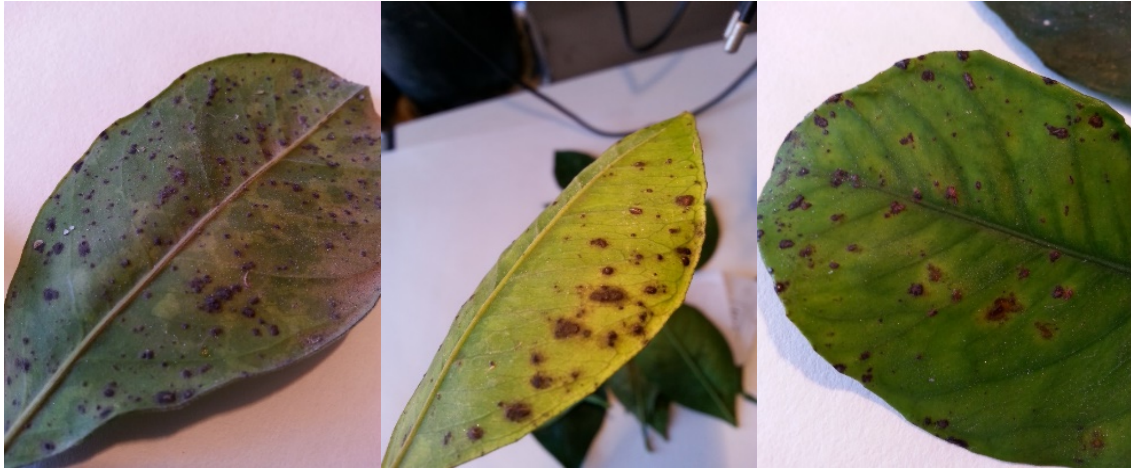


Figure IV-1 Leaf samples of the showing greasy spot like symptoms collected from Sicily, southern Italy. Citrus varieties: Sweet orange (Tarocco Scirè), Navel (Lane Late) and Tangelo (GPS coordinates 37°19'38.5"N 14°41'11.5"E) randomized block design. Collected samples were kept in sterile plastic bags in a thermally insulated container for approximately 4-5 hours until their lyophilization (Labconco Corp., Kansas City, MO). Freeze dried samples were stored at -20°C and homogenized by grinding under liquid nitrogen. DNA was extracted from three technical replicates of each sample. The extraction procedure utilized 0.02 g of homogenized tissue and DNA was extracted using the DNeasy Plant Mini kit (QIAGEN, Dusseldorf, Germany). The quantity and quality of extracted DNA was determined using a Nanodrop 2000 spectrophotometer (Nano-drop Technologies, Wilmington, DE).

DNA extracts from all technical replicates were amplified in triplicate using the universal fungal primers ITS3-ITS4 targeting the ITS2 region of ribosomal DNA (White et al. 1990). Both primers were modified to construct fusion primers appropriate for 454 sequencing with adapters sequences A and B, key sequences and multiplex identifiers (MIDs) (<http://www.454.com/>). Five different MIDs were utilized to label different samples (Table IV-1).

Table IV-1 Summary of analyses and results of metabarcoding surveys conducted with citrus leaves showing symptoms of citrus greasy spot disease (CGS) (Sy), and asymptomatic leaves (As).

Collected leaf samples*	MIDs	Observed species	Shannon	chao1
Tangelo Nova (Sy)	MID7	107.5	4.534482	135.7667
Tangelo Nova (As)	MID10	101	4.031091	128.1841
Lane late orange (Sy)	MID16	80.7	2.796404	122.3217
Lane late orange (As)	MID19	92.2	2.812339	120.2184
Tarocco scirè orange (Sy)	MID28	86.4	2.813436	119.6574

*Collected leaf samples comprised leaves with (Sy) and without (As) typical greasy post symptoms

PCR reactions were conducted in a total volume of 25 µl containing 2.5 µl of 10X reaction buffer, 0.25 µl of each primer ITS3-ITS4 (10µM), 0.1 µl of AccuPrime *Taq* DNA Polymerase High Fidelity (Invitrogen, CA, USA) and 1µl of DNA template (10 ng/µl). Reactions were incubated in an Eppendorf Mastercycler gradient (Hamburg, Germany) for 1 min at 94°C followed by 30 cycles of 30s at 94°C, 30 s at 55°C and 30 s at 68°C. All reactions ended with a final extension of 1 min at 72°C. Amplicons from the technical replicates of each sample were pooled and purified using the Agencourt AMPure XP system (Beckman Coulter, Inc.). The concentration and quality of the purified amplicons was evaluated by agarose gel electrophoresis (Sambrook et al. 1989). Amplicons were sequenced by MacroGen Inc. (Seoul, Korea) using the 454 GS FLX+System (Roche Diagnostics Corporation).

Data analysis and statistics

The bioinformatics pipeline, QIIME v. 1.8 (Caporaso et al. 2010) was used to process and analyze the obtained sequence data. Using default parameters, preliminary processing of the data included de-multiplexing and quality filtering with a minimum quality score of 25, a minimum/maximum length of 150/1000, and a maximum number of homopolymer bases of 6. Sequences were denoised using the denoise wrapper (Reeder and Knight 2010) and the ITS2 region was extracted using ITSx application (Bengtsson-Palme et al. 2013). Chimeric sequences were identified and filtered using USEARCH 6.1 software (Edgar 2010). Sequences were clustered at 97% similarity threshold using USEARCH 6.1 software and the most abundant sequences in each OTU were selected as representative sequences. These sequences were then used for the taxonomy assignment. OTUs were picked using the UNITE dynamic database released on 10.09.2014 (<http://unite.ut.ee/>) as a reference database. The same database was also used for taxonomy

assignments using the BLAST algorithm (Altschul et al. 1990) at a similarity threshold of 0.97.

For downstream analysis, the OUT table was rarefied at an even depth to reduce biases in sequencing depth. Alpha diversity was calculated using both Shannon and Chao1 estimates and results were compared using a two-sample *t*-test based on non-parametric (Monte Carlo) methods with 999 permutations.

Weighted and unweighted UniFrac metrics and the Bray Curtis method (Bray and Curtis 1957) were utilized to evaluate β -diversity (Lozupone and Knight 2005). A distance-based redundancy analysis (db-RDA) and Permanova as implemented in QIIME v. 1.8 was utilized to relate the fungal community composition to sample types and to evaluate differences between symptomatic and asymptomatic leaves and differences among varieties. Additionally, a Monte Carlo permutation test was used to determine experimental variables significantly contributing to the observed variance in fungal communities.

Identification of fungal taxa

In order to confirm the accuracy of QIIME taxonomic assignments, sequences associated with the most abundant OTUs were extracted and introduced into ElimDupes ([http:// hcv.lanl.gov/content/sequence/ELIMDUPES/elimdupes.html](http://hcv.lanl.gov/content/sequence/ELIMDUPES/elimdupes.html)) to detect identical sequences and determine the frequency within each OTU. Unique representative sequences, defined as sequence types (STs) (Cf. Chapter II; Prigigallo et al. 2015), were analyzed along with genetically closely related reference sequences of the same genus to determine their phylogenetic collocation and enable their identification with the highest possible level of accuracy. To this aim, local databases of validated reference sequences were created with priority given to sequences from specific recent taxonomic studies. In cases where no matches were found in the reference sequence from selected published articles, more closely related sequences were selected using BLAST searches of GenBank. The reliability of the latter sequences was evaluated based on different parameters, including the consistency of sequences from different sources, available details on sequenced isolates, and year of publication, giving priority to more recent items. For each selected genus, STs identified in the present study and reference sequences were aligned using MUSCLE and introduced into MEGA for phylogenetic analysis utilizing the

Maximum Likelihood method (Tamura et al. 2013). Analyses were performed with 1000 bootstrap replications.

Results and Discussion

A total of 35,537 reads were recovered after quality evaluations (length trimming, denoising, ITS2 extraction, and chimeric sequence exclusion), and assigned to 212 OTUs clustered at a 97% similarity threshold. The number of OTUs ranged between 86.4 in ‘Tarocco Scirè’ and 107.5 in ‘Tangelo Nova’ (Table IV-1)

The rarefaction analysis indicated that the sequencing depth had been saturated for all of the analyzed samples indicating that the great majority of OTUs had been detected (Fig. IV-2).

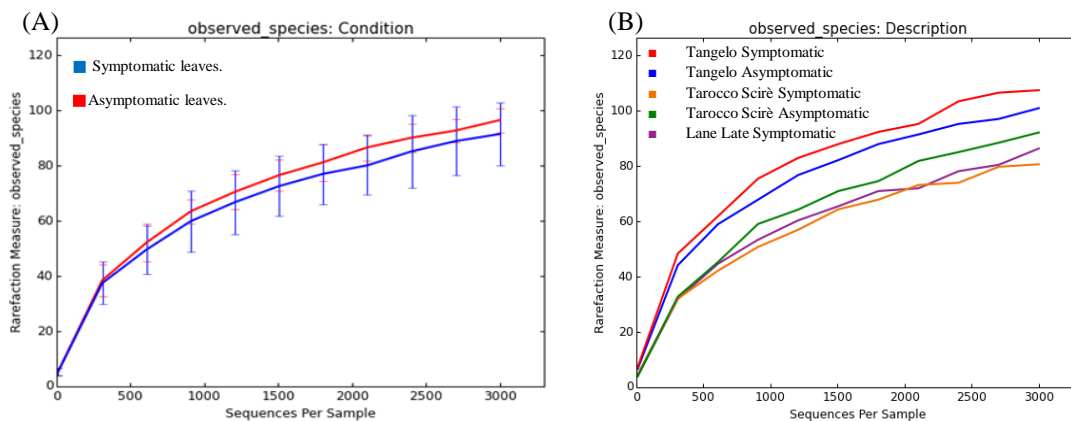


Figure IV-2 Rarefaction plots of the observed OTUs metrics based on β -diversity of the studied citrus samples. A) Samples collapsed based on the symptoms observation; blue curves: symptomatic leaves, red curves: asymptomatic leaves. B) Curves represent individual samples.

Alpha diversity, calculated based on Shannon’s Diversity Index and the Chao1 estimate, indicated that there were no significant differences between symptomatic and asymptomatic leaves (Table IV-2). ‘Tangelo Nova’, however, had both a higher number of OTUs and a greater diversity level, compared to both orange varieties (Tarocco Scirè and Lane Late) (Fig. IV-2B). In agreement with the alpha diversity, a distance-based redundancy analysis (db-RDA) and Permanova based on unweighted UniFrac metrics and Bray Curtis dissimilarity, did not reveal any significant differences between symptomatic and asymptomatic leaves ($P= 0.8$) (Fig. IV-3A).

Table IV-2 Results of nonparametric two sample t-tests to compare the alpha diversities based on Shannon index using Monte Carlo permutations (999) in symptomatic (Sy) and asymptomatic (As) citrus leaves . The table shows the mean and standard deviation of each group's alpha diversity. The t-statistic and p-value are reported in the last columns.

Group1	Group2	Group1 mean	Group1 std	Group2 mean	Group2 std	t stat	p- value
Asymptomatic	Symptomatic	3.421715	0.609376	3.381441	0.815353	0.046188	0.891

On the other hand, samples from the two orange varieties and from ‘Tangelo Nova’ were clearly differentiated in the PCoA plot, regardless of being symptomatic or not (Fig. IV-3B). These results indicate that different fungal communities were resident in or on the leaves of the different citrus species examined, although these differences did not reach the threshold of statistical significance ($P=0.052$).

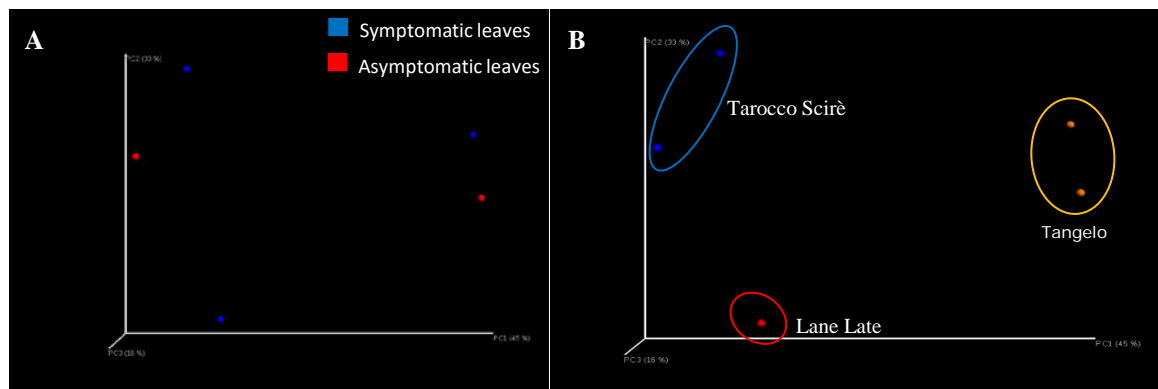


Figure IV-3 PCoA plots using unweighted UniFrac metrics based on β -diversity of the studied citrus samples. On the left side graph, colors are utilized to differentiate symptomatic and asymptomatic leaves ($P=0.80$). On the right side graph, colors differentiate citrus varieties ($P=0.05$).

Regardless of the sample condition or variety, the phylum Ascomycota dominated, representing 92.6% of the total number of detected sequences. This was followed by the phylum Basidiomycota (4.3%), and then unidentified fungi (1.2%). Within the phylum Ascomycota, the class Dothideomycetes (63%) was the most abundant followed by Eurotiomycetes (24.4%). The high incidence of the former class was primarily due to the abundance of the order Capnodiales, and more precisely the family Mycosphaerellaceae. Whereas Chaetothyriales was the sole family present in Eurotiomycetes (Fig. IV-4A, B). Within the family Mycosphaerellaceae, representing the most abundant taxa, we found unidentified Mycosphaerellaceae (44.80%), *Mycosphaerella* spp. (2.40%) and *Septoria* spp. (1.70%). Importantly, this family contains pathogens believed to be responsible for CGS disease (Timmer and Gottwald 2000, Timmer et al. 2000, Mondal and Timmer 2003, 2006). Qiime analyses did not enable the identification of most sequences of this family.

Therefore, a total of 9 STs were identified and phylogenetically analyzed using validated reference sequences containing representative sequences of the order Capnodiales (Crous et al. 2009), the family Mycosphaerellaceae, and the genus *Zasmidium* spp. which has been recently associated with CGS-like symptoms (Huang 2015).

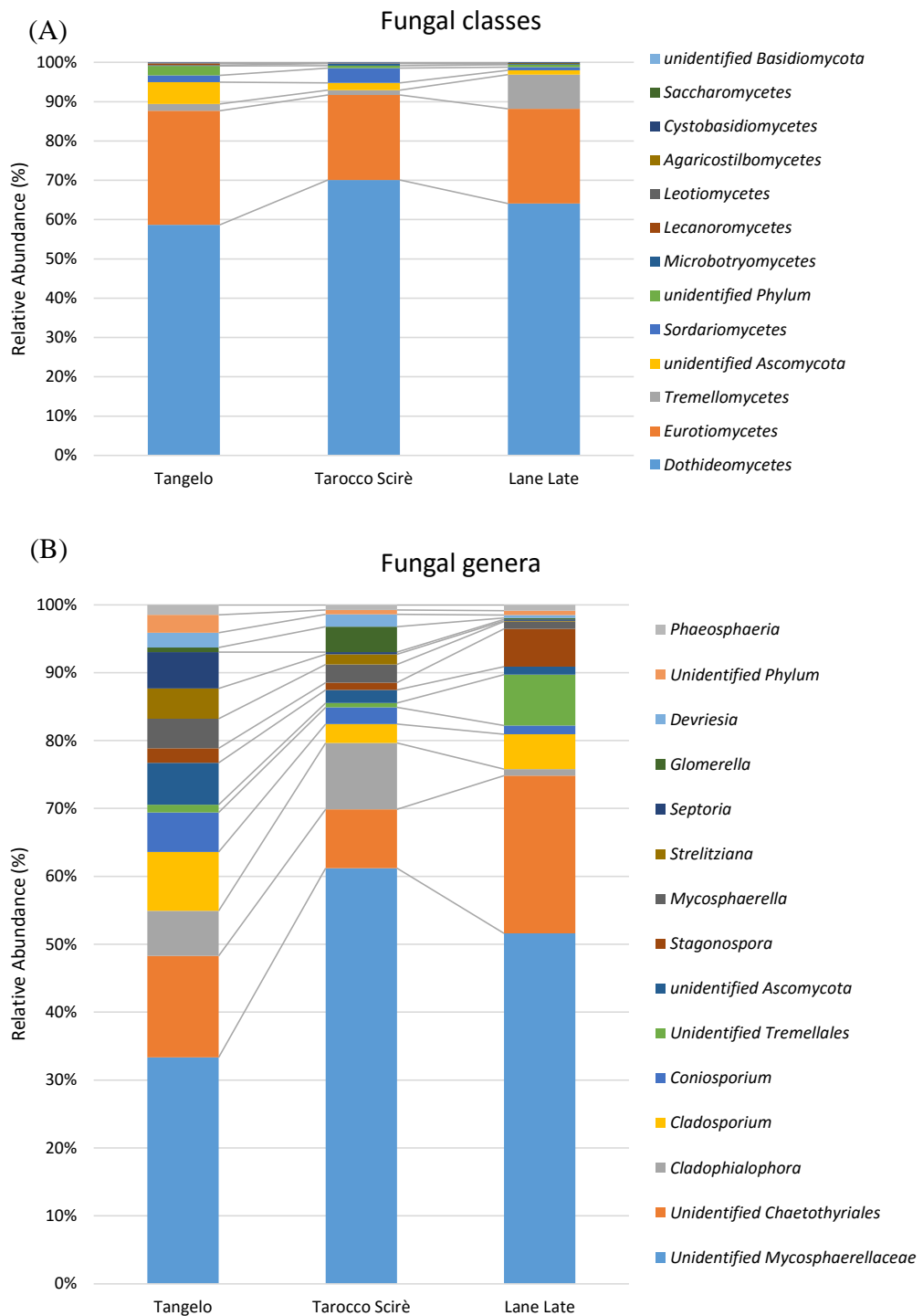


Figure IV-4 Relative abundance of different fungal classes (A) and genera (B) detected in the studied with at least 1% relative abundance. Citrus varieties: Sweet orange (Tarocco Scirè), Navel Orange (Lane Late) and Tangelo Nova.

According to this analysis, the most abundant ST within the family Mycosphaerellaceae (representing 92.9% of the sequences) corresponded to the species *Mycosphaerella fragariae*, although other sequences, mainly from *Ramularia* spp., had very similar ITS2 sequences (Fig. IV-5). One more ST, representing around 1% of Mycosphaerellaceae sequences, was identified as *M. graminicola*. All other STs clustered within the genus *Mycosphaerella* but their identification at the species level was not possible because of the presence of identical or very similar sequences in more reference species, or because of the absence of very closely related sequences. In the latter case, sequences were represented by two STs and accounted for 4.2% of the sequences in the family Mycosphaerellaceae. These STs may represent new *Mycosphaerella* species that are still unknown to the scientific community.

Surprisingly, none of the detected STs clustered with *M. citri*, indicating that the fungal species currently considered as the causal agent of CGS (Mondal and Timmer 2006) was not present on citrus leaves collected in Sicily, regardless of the presence of the disease symptoms. These results in combination with the abundant presence of STs clustering within the genus *Mycosphaerella* may suggest the involvement of other species of this genus as the causal agent of CGS disease in Sicily. Of particular interest was the prevalence of *M. fragariae* over other detected fungal sequences. This species seems to be specifically associated with strawberry based on currently available data (Zheng and Sutton 1994), however, the involvement of the same or closely related species in other pathosystems cannot be completely excluded.

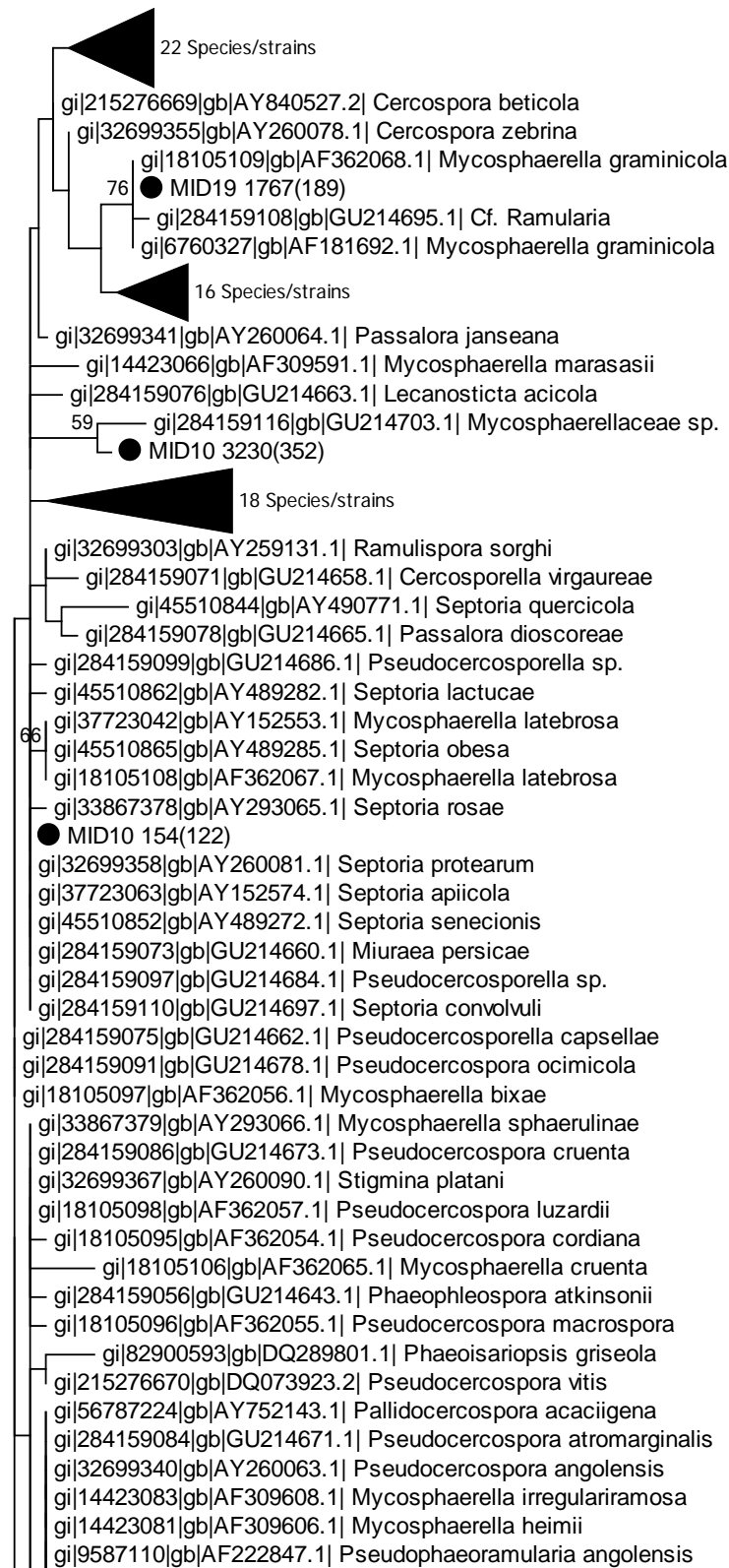


Figure IV-5 Molecular Phylogenetic analysis of STs associated to the family Mycosphaerellaceae. Numbers in parentheses along with STs (MIDs) indicate the number of sequences represented by each ST. Numbers on nodes represent the posterior probabilities for the maximum likelihood method.

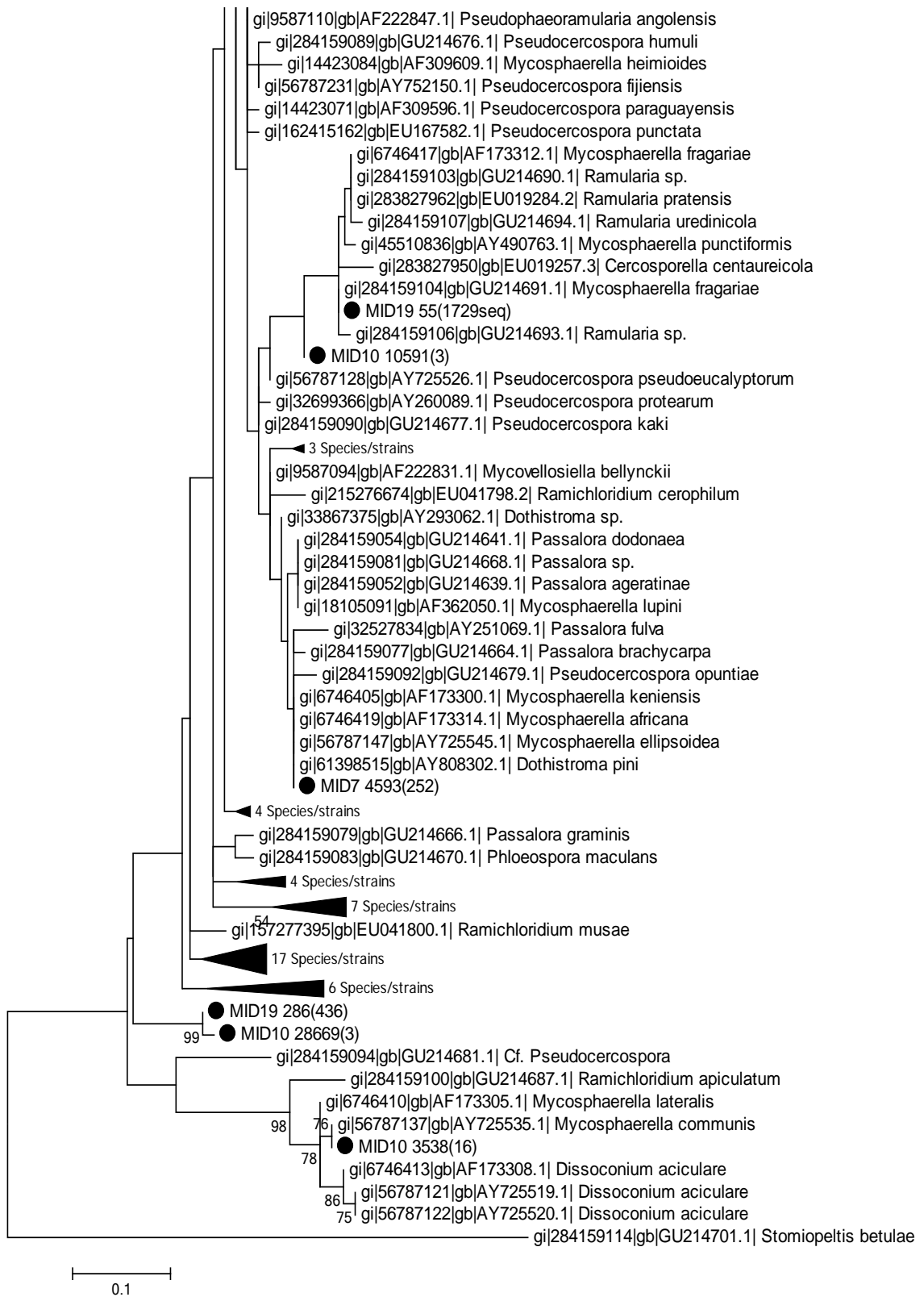


Figure IV-5 Continued

Indeed, it must be stressed that our identification at the species level is very rough, since it is only based on the phylogenetic analysis of the ITS2 sequence. Therefore, considering the complex taxonomy of Mycosphaerellaceae (Crous et al. 2009, Huang 2015), the species identification designated in the current study should be viewed with caution. Although our data did not show significant differences between symptomatic and asymptomatic leaves, the putative pathogen responsible for CGS may be present in asymptomatic leaves as quiescent or latent infections, especially when the long incubation period typical of most *Mycosphaerella* pathogens is taken into account (Mondal and Timmer 2006). In this context, it must be emphasized that symptomatic and asymptomatic leaves analyzed in the present study were collected from the same trees. Unfortunately, leaves from completely healthy trees, devoid of any CGS symptoms, were not included in the present analysis, and data from the currently available literature is insufficient to conclude anything about the presence and abundance of *Mycosphaerella* species in healthy citrus leaves.

The second most abundant group of sequences (14.3%) represented by 8 STs was p to the order Chaetothyriales. According to the phylogenetic analysis of the STs, using reference sequences from the family Herpotrichiellaceae (Davey and Currah 2007) and Chaetothyriaceae (Chomnunti et al. 2012), 4 STs exhibited high similarity to fungi of the genus *Knufia* (Fig. IV-6), while the other 4 STs did not show any close relationship to currently known fungal genera (Fig. IV-7). The fact that some of the sequences were not identified, even at the family level, suggests that there are still unknown taxa present on citrus leaves. The order Chaetothyriales is divided into two different families, the Chaetothyriaceae and the Herpotrichiellaceae (Geiser et al. 2006). The Chaetothyriaceae are known as epiphytes, colonizing the leaves and the bark of trees in the tropics (Batista et al. 1962).

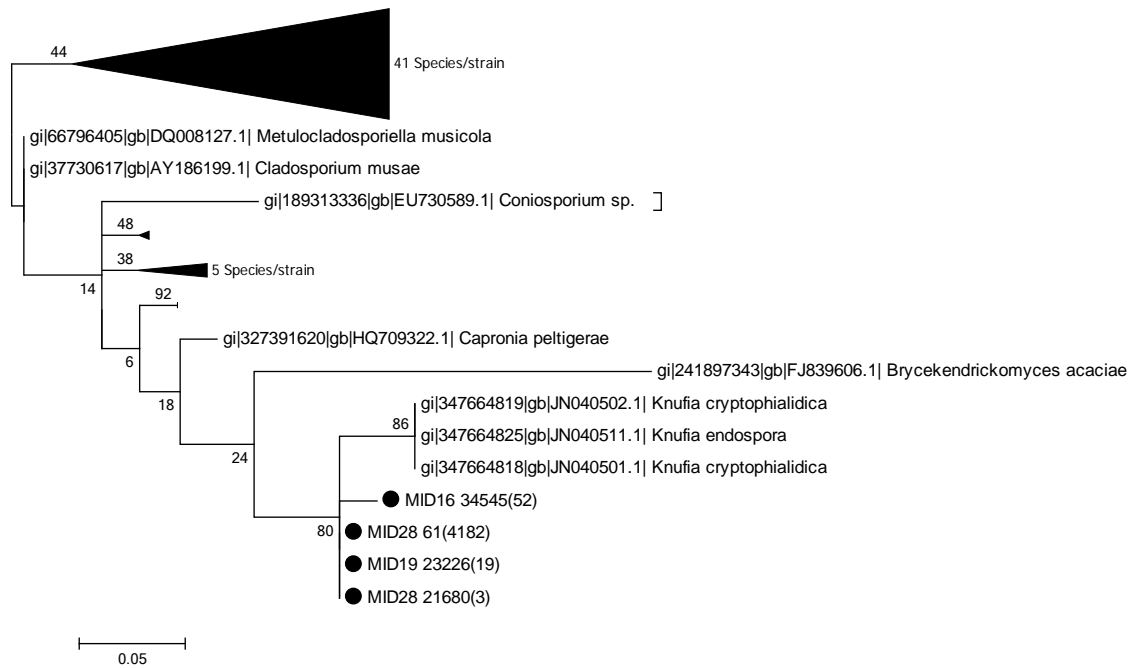


Figure IV-6 Molecular Phylogenetic analysis of STs associated to the order Chaetothyriales. Numbers in parentheses along with STs (MIDs) indicate the number of sequences represented by each ST. Numbers on nodes represent the posterior probabilities for the maximum likelihood method.

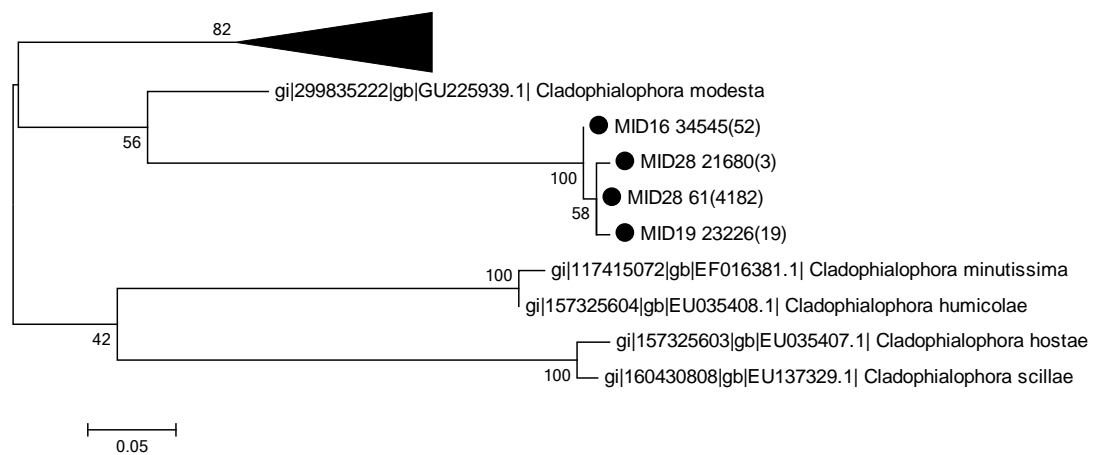


Figure IV-7 Molecular Phylogenetic analysis of STs associated to the order Chaetothyriales. Numbers in parentheses along with STs (MIDs) indicate the number of sequences represented by each ST. Numbers on nodes represent the posterior probabilities for the maximum likelihood method.

Another group of sequences (5.30%) was associated with the genus *Cladophialophora* and based on a phylogenetic analysis utilizing reference sequences (Liu et al. 2015), they were identified as *Cladophialophora protea* (Fig. IV-8). This species was originally isolated from the woody shrub, *Protea cynaroides*, on which it was assumed to be pathogenic, though no inoculation tests have ever been conducted to confirm this hypothesis (Swart et al. 1998, Crous et al. 2008). *C. Protea* also occurs on dead leaf tissues of the cycad, *Encephalartos altensteinii* (Swart et al. 1998, Crous et al. 2007). This fungal species is worthy of further investigation since its RA was higher in symptomatic (14.6%) leaf samples than in asymptomatic (9.3%) leaf samples.

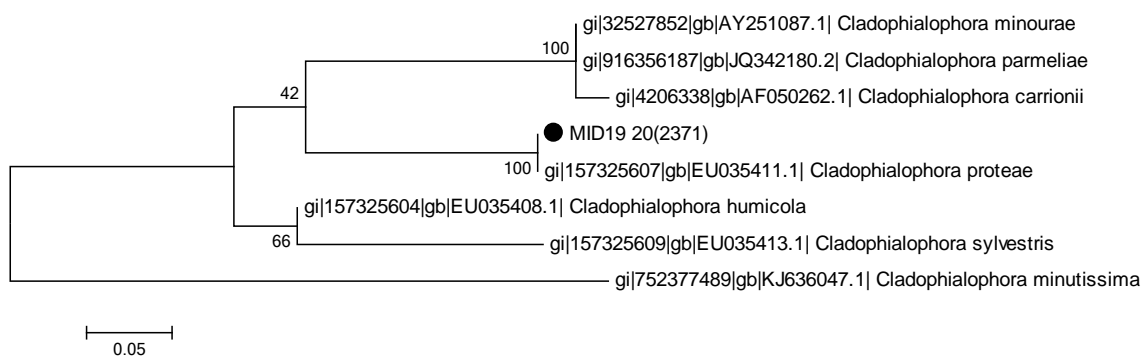


Figure IV-8 Molecular Phylogenetic analysis of STs associated to the order *Cladophialophora*. Numbers in parentheses along with STs (MIDs) indicate the number of sequences represented by each ST. Numbers on nodes represent the posterior probabilities for the maximum likelihood method.

Based on our analysis, the genus *Cladosporium* had a RA of 5% containing 1508 sequences, represented by a single ST. The identification of the species was not possible due to the complexity of the genus and the existence of many taxa with identical ITS sequences (Bensch et al. 2015). The genus *Cladosporium* represents one of the most common fungi of the dematiaceous hyphomycetes (Bensch et al. 2015), comprising human and plant pathogens, as well as beneficial fungi (Bensch et al. 2012). Therefore, any speculation regarding the role of this genus in the citrus phyllosphere is problematic.

Another sequence with an RA of 2.9% corresponded to the genus *Coniosporium* and according to BLAST search had a 99% similarity to an uncultured fungus clone (KP891446.1) isolated from the elm bark beetle, *Scolytus multistriatus*. *Coniosporium* species are known as Black Yeasts and have been found on rotten wood of *Pinus maritima*, rock, and marble in Greece and Italy (Sterflinger et al. 1998, Sert et al. 2007).

Several sequences, with a total RA of 2.9%, were associated with the order Tremellales. According to a BLAST search, the majority of these sequences showed high

identity with uncultured fungi (*i.e.* EU486124.1) and to *Cryptococcus* species such as *C. carnescens* (KP131895) and other uncultured *Cryptococcus* species.

The genus *Colletotrichum* (1.4%) was represented by 785 sequences grouped in 3 STs. According to the phylogenetic analysis of these STs, along with reference sequences of *Colletotrichum* (Damm et al. 2012), the most abundant ST (759 sequences) was identified as *C. gloeosporioides*. Another ST clustered within the *C. boninense* species complex but was not identified at the species level, while the third ST was identified as *C. karstii*. Although *C. gloeosporioides* is frequently isolated from citrus and is believed to be an endophyte, the presence of an ST associated with *C. boninense s.l.* is interesting since never reported before (Fig. IV-9). These fungi have been recently detected in the phyllosphere of olive using an amplicon metabarcoding approach based on *Colletotrichum* spp. specific primers (Mosca et al. 2014).

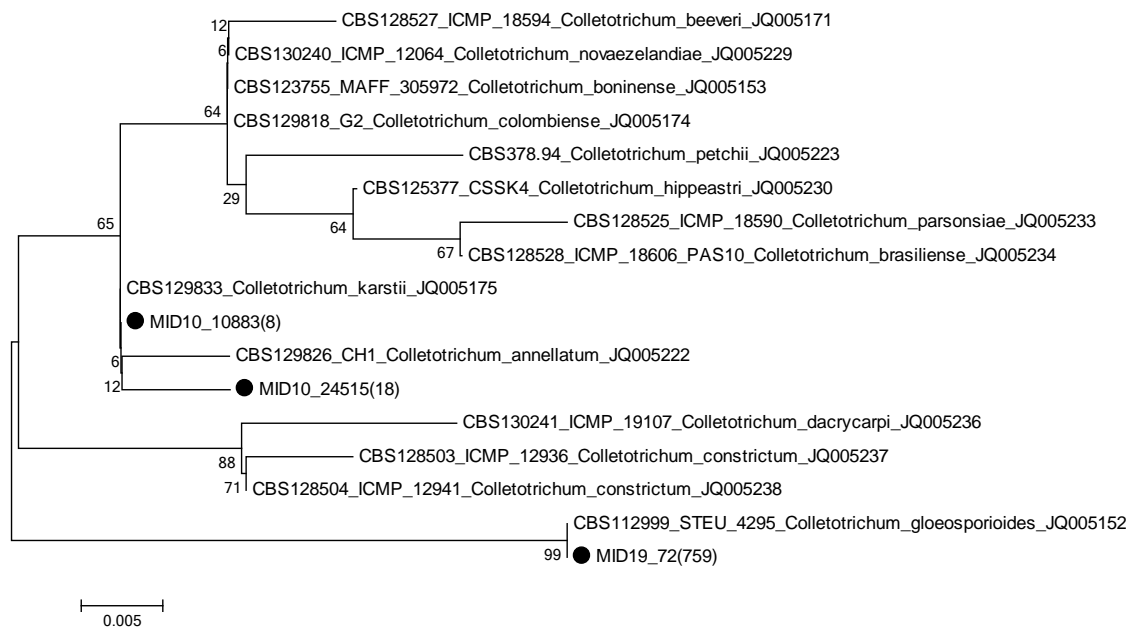


Figure IV-9. Molecular Phylogenetic analysis of STs associated to the genus *Colletotrichum*. Numbers in parentheses along with STs (MIDs) indicate the number of sequences represented by each ST. Numbers on nodes represent the posterior probabilities for the maximum likelihood method.

Lastly, 4 STs having collective RA of 1.3% were associated with the genus *Devriesia*. Most of them were identified as *D. fraseriae* and secondly as *D. hilliana* (Fig. IV-10). Although, there are no records of this genus on citrus, *D. fraseriae*, which was isolated originally in 2010 in Australia, was found with high abundance in olive leaves during October and December in Calabria, Italy (Cf. Chapter II).

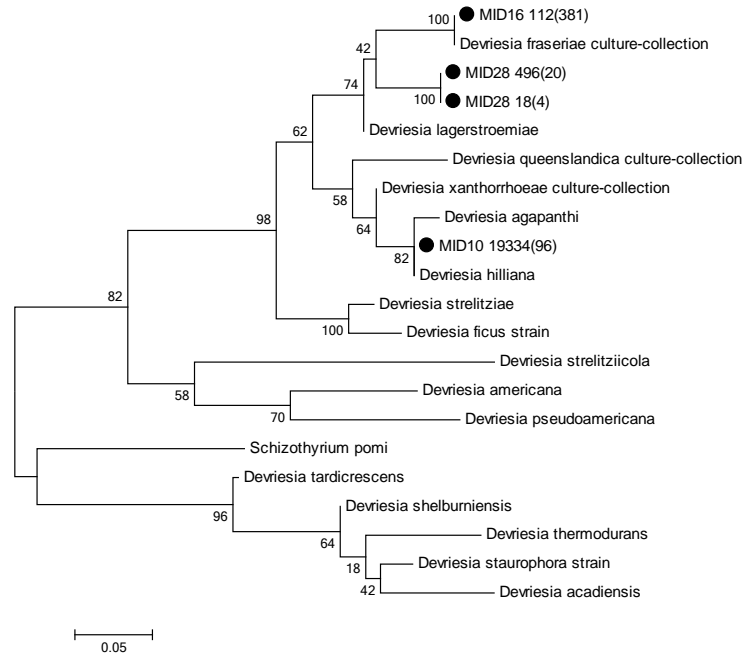


Figure IV-10 Molecular Phylogenetic analysis of STs associated to the genus *Devriesia*. Numbers in parentheses along with STs (MIDs) indicate the number of sequences represented by each ST. Numbers on nodes represent the posterior probabilities for the maximum likelihood method.

Conclusion

In conclusion, the results of the present study indicate that citrus leaves support a high diversity of fungal organisms. In general, *Mycosphaerella* species were more abundant than any of the other identified genera, followed by unidentified taxa in the Chaetothyriales. In addition, a large portion of the detected sequences were not associated with any of any currently known fungal species, indicating that a considerable portion of the fungal diversity of citrus leaves has yet to be characterized. Although, the present study did not provide definitive results about the causal agent of CGS in the investigated area, it clearly showed that *M. citri*, the most accredited causal agent of this disease, was not present in Sicily, southern Italy. The abundant detection of *M. fragariae* and other *Mycosphaerella* species may suggest the involvement of other *Mycosphaerella* species in this disease, but additional studies will be needed to confirm this possibility. Considering the difficulties in confirming Koch's postulates for *Mycosphaerella*-like organisms (Wellings 1981), additional metabarcoding analyses seem to be a logical approach to for studying the complex etiology of CGS and its causal agents. Further metagenomic studies should help to address many of the unresolved questions about this disease.

References

- Altschul, S. F., W. Gish, W. Miller, E. W. Myers, and D. J. Lipman. 1990. Basic local alignment search tool. *Journal of Molecular Biology* **215**:403-410.
- Baker, R. 2008. Pest risk assessment made by France on *Mycosphaerella citri* considered by France as harmful in French overseas department of Réunion.
- Batista, A. C., R. Ciferri, R. Ciferri, and R. Ciferri. 1962. The chaetothyriales. Berger.
- Bengtsson-Palme, J., M. Ryberg, M. Hartmann, S. Branco, Z. Wang, A. Godhe, P. De Wit, M. Sánchez-García, I. Ebersberger, F. de Sousa, A. Amend, A. Jumpponen, M. Unterseher, E. Kristiansson, K. Abarenkov, Y. J. K. Bertrand, K. Sanli, K. M. Eriksson, U. Vik, V. Veldre, and R. H. Nilsson. 2013. Improved software detection and extraction of ITS1 and ITS2 from ribosomal ITS sequences of fungi and other eukaryotes for analysis of environmental sequencing data. *Methods in Ecology and Evolution* **4**:914-919.
- Bensch, K., U. Braun, J. Z. Groenewald, and P. W. Crous. 2012. The genus *Cladosporium*. *Studies in Mycology* **72**:1-401.
- Bensch, K., J. Groenewald, U. Braun, J. Dijksterhuis, M. de Jesús Yáñez-Morales, and P. Crous. 2015. Common but different: The expanding realm of *Cladosporium*. *Studies in Mycology* **82**:23-74.
- Biasi, A., A. De Patrizio, R. Faedda, L. Schena, A. Pane, S. O. Cacciola, and G. M. di San Lio. 2012. *Septoria citri* is a common pathogen of citrus in southern Italy. *Journal of Plant Pathology* **94**.
- Bray, J. R., and J. T. Curtis. 1957. An ordination of the upland forest communities of southern Wisconsin. *Ecological Monographs* **27**:325-349.
- Caporaso, J. G., J. Kuczynski, J. Stombaugh, K. Bittinger, F. D. Bushman, E. K. Costello, N. Fierer, A. G. Pena, J. K. Goodrich, and J. I. Gordon. 2010. QIIME allows analysis of high-throughput community sequencing data. *Nature methods* **7**:335-336.
- Chomnunti, P., T. W. K. Ko, E. Chukeatirote, K. D. Hyde, L. Cai, E. B. G. Jones, R. Kodsueb, B. A. Hassan, and H. Chen. 2012. Phylogeny of Chaetothyriaceae in northern Thailand including three new species. *Mycologia* **104**:382-395.
- Crous, P. W., C. L. Schoch, K. D. Hyde, A. R. Wood, C. Gueidan, G. S. de Hoog, and J. Z. Groenewald. 2009. Phylogenetic lineages in the Capnodiales. *Studies in Mycology* **64**:17-47S17.
- Crous, P. W., K. Schubert, U. Braun, G. S. de Hoog, A. D. Hocking, H. D. Shin, and J. Z. Groenewald. 2007. Opportunistic, human-pathogenic species in the Herpotrichiellaceae are phenotypically similar to saprobic or phytopathogenic species in the Venturiaceae. *Studies in Mycology* **58**:185-217.
- Crous, P. W., A. R. Wood, G. Okada, and J. Z. Groenewald. 2008. Follicolous microfungi occurring on Encephalartos. *Persoonia : Molecular Phylogeny and Evolution of Fungi* **21**:135-146.
- Damm, U., P. Cannon, J. Woudenberg, P. Johnston, B. Weir, Y. Tan, R. Shivas, and P. Crous. 2012. The *Colletotrichum boninense* species complex. *Studies in Mycology* **73**:1-36.

- Davey, M. L., and R. S. Currah. 2007. A new species of *Cladophialophora* (hyphomycetes) from boreal and montane bryophytes. *Mycological Research* **111**:106-116.
- Díaz, J., O. Brown, L. Herrera, and V. C. C. Marta Abreu. 1985. Influencia de la mancha grasienta (*Mycosphaerella citri* Whiteside) sobre el rendimiento del naranjo Valencia. *Centro Agrícola*.(Ene-Abr **12**:127-137.
- Edgar, R. C. 2010. Search and clustering orders of magnitude faster than BLAST. *Bioinformatics* **26**:2460-2461.
- Fawcett, H. S. 1915. Citrus diseases of Florida and Cuba compared with those of California.
- Geiser, D. M., C. Gueidan, J. Miadlikowska, F. Lutzoni, F. Kauff, V. Hofstetter, E. Fraker, C. L. Schoch, L. Tibell, and W. A. Untereiner. 2006. *Eurotiomycetes: eurotiomycetidae* and *chaetothyriomycetidae*. *Mycologia* **98**:1053-1064.
- Gonghi, Z. F. Y. 1987. The fungus causing greasy yellow spot of citrus in china. *Acta Phytopathologica Sinica* **3**:012.
- Grasso, S., and A. Catara. 1982. Osservazioni su intumescenze gommose delle foglie di agrumi [dovute a funghi patogeni]. *Informatore fitopatologico*.
- Haggag, W. 2012. First report of *Mycosphaerella citri*, the agent of greasy spot of orange in Egypt. *Journal of Plant Pathology* **94**.
- Hidalgo, H., T. B. Sutton, and F. Arauz. 1997. Epidemiology and control of citrus greasy spot on Valencia orange in the humid tropics of Costa Rica. *Plant Disease* **81**:1015-1022.
- Huang, F. 2015. *Cercosporoid* diseases of Citrus. *Mycologia*:15-059.
- Liu, X.-Y., D. Udayanga, Z.-L. Luo, L.-J. Chen, D.-Q. Zhou, H. Y. Su, and K. D. Hyde. 2015. Backbone tree for Chaetothyriales with four new species of *Minimelanolocus* from aquatic habitats. *Fungal biology* **119**:1046-1062.
- Lozupone, C., and R. Knight. 2005. UniFrac: a new phylogenetic method for comparing microbial communities. *Applied and Environmental Microbiology* **71**:8228-8235.
- Mondal, S., and L. Timmer. 2003. Effect of urea, CaCO₃, and dolomite on pseudothecial development and ascospore production of *Mycosphaerella citri*. *Plant disease* **87**:478-483.
- Mondal, S., and L. Timmer. 2006. Greasy spot, a serious endemic problem for citrus production in the Caribbean Basin. *Plant disease* **90**:532-538.
- Mosca, S., M. G. Li Destri Nicosia, S. O. Cacciola, and L. Schena. 2014. Molecular analysis of *Colletotrichum* species in the carposphere and phyllosphere of olive. *PloS one* **9**:e114031.
- Prigigallo, M., S. Mosca, S. O. Cacciola, D. Cooke, and L. Schena. 2015. Molecular analysis of *Phytophthora* diversity in nursery-grown ornamental and fruit plants. *Plant pathology*.
- Reeder, J., and R. Knight. 2010. Rapidly denoising pyrosequencing amplicon reads by exploiting rank-abundance distributions. *Nature methods* **7**:668-669.
- Sambrook, J., E. F. Fritsch, and T. Maniatis. 1989. *Molecular cloning*. Cold spring harbor laboratory press New York.

- Sert, H. B., H. Sümbül, and K. Sterflinger. 2007. Microcolonial fungi from antique marbles in Perge/Side/Termessos (Antalya/Turkey). *Antonie van Leeuwenhoek* **91**:217-227.
- Sterflinger, K., W. Krumbein, T. Lellau, and J. Rullkötter. 1998. Two cases of biogenic patina formation on rock. *Ancient Biomolecules* **3**:51-65.
- Swart, L., P. W. Crous, S. Denman, and M. E. Palm. 1998. Fungi occurring on Proteaceae. I. *South African Journal of Botany* **64**:137-145.
- Tamura, K., G. Stecher, D. Peterson, A. Filipski, and S. Kumar. 2013. MEGA6: Molecular Evolutionary Genetics Analysis version 6.0. *Molecular biology and evolution* **30**:2725-2729.
- Thompson, W. 1948. Greasy spot on citrus leaves. *Citrus Industry* **29**:20-22.
- Timmer, L., and T. Gottwald. 2000. Greasy spot and similar diseases. *Compendium of citrus diseases.* Second edition. (Eds LW Timmer, SM Garnsey, JH Graham) pp:25-28.
- Timmer, L., P. Roberts, H. Darhower, P. Bushong, E. Stover, T. Peever, and A. Ibáñez. 2000. Epidemiology and control of citrus greasy spot in different citrus-growing areas in Florida. *Plant disease* **84**:1294-1298.
- Wellings, C. 1981. Pathogenicity of fungi associated with citrus greasy spot in New South Wales. *Transactions of the British Mycological Society* **76**:495-499.
- White, T. J., T. Bruns, S. Lee, and J. Taylor. 1990. Amplification and direct sequencing of fungal ribosomal RNA genes for phylogenetics. *PCR protocols: a guide to methods and applications* **18**:315-322.
- Whiteside, J. 1970a. Effect of fungicides applied to citrus trees on peritheclal development by the greasy spot fungus in detached leaves. *Plant Disease Reporter* **54**:865-869.
- Whiteside, J. 1970b. Etiology and epidemiology of Citrus greasy spot. *Phytopathology* **60**:1409-1414.
- Whiteside, J. 1970c. Symptomatology of Orange fruit infected by the Citrus greasy spot fungus. *Phytopathology* **60**:1859-1860.
- Whiteside, J. 1972. Histopathology of Citrus greasy spot and identification of the causal fungus. *Phytopathology* **62**:260-263.
- Whiteside, J. 1983. Fungicidal effects of some acaricides on *Mycosphaerella citri*. *Plant Disease* **67**:864-866.
- Whiteside, J. 1988. Greasy spot and greasy spot rind blotch. *Compendium of Citrus Diseases*. JO Whiteside, SM Garnsey, and LW Timmer, eds. APS Press, St. Paul, MN **55121**:15-17.
- Yamada, S. 1956. Studies on the greasy spot (black melanose) of citrus. II. Morphological characteristics of the causal fungus (*Mycosphaerella horii* Hara). *Horticulture Division, National Tokai-kinki Agricultural Experiment Station, Okitsu Japan. Bulletin* **3**:49-62.
- Zheng, J., and J. Sutton. 1994. Inoculum concentration, leaf age, wetness duration, and temperature in relation to infection of strawberry leaves by *Diplocarpon earlianu*. *Canadian Journal of Plant Pathology* **16**:177-186.

Chapter V. Metagenomic analysis of fungal populations in harvested organic and conventional apples and the impact of hot water treatments on fungal diversity

Abstract

In the present study, a metagenomics approach based on Illumina MiSeq v. 3 technology was utilized to evaluate the fungal diversity in apple fruit using the ITS2 region as a barcode gene. Experiments were performed to characterize the fungal communities present on fruit in relation to several factors, including management practice (organic vs conventional), the part of the fruit (stem and calyx end), the application of a hot water treatments (45°C for 10 min) used to manage postharvest pathogens. After quality manipulation, a total of 5,760,162 high quality fungal sequences were recovered and assigned to 8,504 OTUs. Members of the phylum Ascomycota were dominant in all samples and accounted for 91.6% of the total number of detected sequences. This was followed by Basidiomycota (8%), Chytridiomycota (0.1%) and unidentified fungi (0.3%). The incidence of the Phyla varied significantly between the examined fruit parts. For example, Ascomycota had a relative abundance (RA) of 88.2% in stem end (SE) samples and a RA of 95.3% in calyx end (CE) samples, whereas Basidiomycota had a RA of 11.4 and 4.6% in the SE and CE, respectively. No significant difference was observed in the alpha diversity of fungal communities present in apples grown using organic vs. conventional management practices. In contrast, significantly different alpha and beta diversity indexes were evident for CE and SE communities ($P = 0.001$). The heat treatment had a minor effect on fungal populations as indicated by the presence of significant differences in beta diversity indexes for the CE of organic apples. These results were consistent over time of sampling. The presence of specific species and their relative abundance are discussed. Results indicate that fungal communities present in the CE and SE of apples are very diverse and resilient and do not appear to be impacted by management practices or by heat treatments.

Introduction

The management of postharvest diseases plays an important role in reducing the economic impact of fruit decay and spoilage, as well as in limiting health risks due to

mycotoxin contamination of processed fruit products such as juice and baby food. Finding safe and effective alternatives to synthetic, chemical fungicides for reducing postharvest losses of harvested commodities has been a focus of much research over the past three decades (Wilson and Wisniewski 1989, Wisniewski and Wilson 1992, Spadaro and Gullino 2004, Droby et al. 2009, Sui et al. 2015). Looking for alternatives that are widely accepted and commercially viable has, however, been a challenge. Nevertheless, there is a real imperative to continue this line of research, since regulations on the use of both new and existing fungicides are becoming more and more stringent. For example, the postharvest use of fungicides is completely prohibited in some European countries or limited to just a few registered chemicals (EU Directive 2009/128/EC). Safety concerns about mycotoxins and foodborne pathogens also increase the need to find viable alternatives. Alternative control measures include the use of natural compounds such as chitosan and oligochitosan, plant extracts, essential oils, modified atmosphere, physical treatments (hot water, UV-C etc.), and microbial antagonists.

Apples are one of the most popular fruit in the world, with a total production of 80,822,521 tons in 2013 (FAOSTAT 2015). Apples are grown using both organic and non-organic (sustainable or conventional, respectively) management practices. According to the United States Department of Agriculture (USDA), organic production is defined as an ecological production system that integrates cultural, biological, and mechanical practices that foster resource cycling, ecological balance, and biodiversity. Public awareness of agricultural chemicals used to produce food crops has been increasing worldwide, especially in the United States, after the media reported in 1989 on the use of the growth regulator daminozide (Alar) on apples and health risks associated with the use of this chemical compound (Egan 1991). Since then, organic apple production as a portion of American apple production has been steadily increasing (Slattery et al. 2011). Despite the extensive literature currently available about alternative control means and the increasing spread of organic farming systems in the United States, as well as in many other geographic regions, such as Europe, little is currently known about the complex interactions that occur between microbial antagonists, host tissues, pathogens, and how they are affected by biotic and abiotic elicitors of defense mechanisms, and the environment. In this context, recent advances in -omics technologies (genomics, transcriptomics, metabolomics, proteomics, and metagenomics) may play a crucial role in increasing our understanding of fungal biology in relation to food crops, and the effects of

alternative control methods, such as biological control. In particular, massive sequencing of polymerase chain reaction (PCR) amplicons of specific barcode genes (amplicon metagenomics or metabarcoding) has proved to be a powerful culture-independent technique for investigating microbial diversity and for determining the relative quantity of community members in environmental samples (Lindahl et al. 2013). The identification and quantification of endophytic and epiphytic microflora present in and on plants may provide information that can facilitate our understanding of the complex interactions that take place between plants and resident fungal microflora, including pathogenic and beneficial species (Schena et al. 2003, Pinto et al. 2014). Furthermore, these techniques enable an in-depth investigation of the effect of conventional and alternative control methods, including physical treatments such as hot water and biological treatments such as the application of biocontrol agents, on the composition and diversity of microbial communities (Sylla et al. 2013a, Sylla et al. 2013b).

In the present study, a metagenomics approach based on the fungal ITS2 region was utilized to assess the fungal diversity of organic and conventional red delicious apple fruit, as well as to investigate the effect of hot water as postharvest treatment on the fungal communities present on apple fruit. The objective of the study was to investigate the following aspects: i) impact of management practices (organic vs conventional) on apple fungal communities (“practices” effect); ii) differences in the fungal communities present in the stem end (SE) vs. the calyx end (CE) of apples (“location” effect); iii) impact of hot water treatments on the fungal communities, taking into consideration different “locations” and “practices”(“treatment” effect); and iv) effect of time following hot water treatments on the fungal community in different “locations” and “practices” (“time” effect).

Materials and Methods

Experimental design

The conducted experiment included the use of 90 ‘Golden Delicious’ apples (45 organic and 45 conventional fruit) that had been grown in the state of Washington, USA and purchased at a local supermarket in Charles Town, WV. Organic and conventional fruit were divided into three groups of 15 fruit, which received one of the following treatments: i) dipped in hot water (Treated); ii) dipped in room temperature water (Washed); or iii) untreated (Untouched). Five apples from each treatment, representing 5

biological replicates, were analyzed at three different time points: i) T1 = directly after the treatment; ii) T4 = one week after the treatment; and iii) T5 = two weeks after the treatment.

Hot water treatment

Ultrapure Water filtered through Barnstead E-Pure Purification filters (Thermo Fisher Scientific) and autoclaved for 30 minutes at 120°C, was utilized in the experiment. Organic and conventional apples were treated separately to avoid cross contamination. Apples were dipped for 10 min in a preheated water bath at 45°C (Treated) or in water at room temperature (approximately 22°C) (washed). Untouched apples were used as a control.

Sampling of apple tissues

A sterile cork borer (11mm diameter) was used to extract the core of apple fruit and a sterile razor was used to cut it into 3 parts: i) the stem end (SE); ii) the core middle (CM) which included seeds and iii) the calyx end (CE). Each part was immediately frozen in liquid nitrogen and stored in -80 °C before lyophilization. Only the SE and CE samples were subjected to metagenomics analysis.

DNA extraction, amplification and sequencing

Collected samples were homogenized in a 2010 Geno/Grinder SPEX SamplePrep using autoclaved metal beads, and total DNA was extracted using the Wizard Genomic DNA Purification Kit (Promega, Madison, WI, USA) according to the manufacturer's protocol. Samples were amplified using the universal fungal primers ITS3_KYO2-ITS4 to amplify the ITS2 region of the ribosomal DNA (Toju et al. 2012). Both primers were modified to include Illumina adaptors (www.illumina.com). PCR reactions were conducted using in a total volume of 25 µl containing: 12.5µl of KAPA HiFi HotStart ReadyMix (2X) (Kapa Biosystems, Wilmington, MA, USA), 1.5 µl of each primer ITS3_KYO2-ITS4 (10µM), and 1µl of DNA template. Reactions were incubated in a 100 thermal cycler (Bio-Rad) for 3 min at 98°C followed by 30 cycles of 30s at 95°C, 30 s at 50°C and 30 s at 72°C. All reactions ended with a final extension of 1 min at 72°C. Nuclease-Free Water (QIAGEN, Valencia, CA, USA) subjected to PCR was processed as a negative control. All amplicons, as well as amplification mixtures from the negative

controls were sent to the DNA Services Facility (University of Illinois, Chicago, USA) for sequencing using Illumina MiSeq V3 (2×300 bp) chemistry.

Data Analysis

Paired-end reads were merged using PEAR 0.9.6 Paired-End reAd merger (Zhang et al. 2014) using default parameters. The CLC genomics workbench V 8 (Qiagen) was used for primer and quality trimming with minimum of Q20. Sequences without either of the primers were discarded. The ITS2 region was extracted using ITSx application (Bengtsson-Palme et al. 2013) and chimeric sequences were identified and filtered using VSEARCH 1.4.0 (Rognes et al. 2015). Sequences were clustered at similarity threshold of 97% using USEARCH 6.1 software and the most abundant sequences in each OTU were selected as representative sequences. These sequences were then used for the taxonomic assignment. OTUs were picked using the UNITE dynamic database released on 10.09.2014 (<http://unite.ut.ee/>) as a reference database. The same database was also used for taxonomy assignments using a BLAST algorithm.

The OTU table was rarefied to an even sequencing depth (5000) in order to remove sample heterogeneity. Rarefaction analysis was used to calculate alpha diversity. MetagenomeSeq's cumulative sum scaling (CSS) was used for other downstream analyses, including beta diversity and group significance (Paulson et al. 2013). Alpha diversity was calculated using Observed Species, Chao1, and Shannon metrics. Alpha diversity results were compared based on a two-sample t-test using non-parametric (Monte Carlo) methods and the default number of Monte Carlo permutations (999). Results were visualized in boxplots figures. Bray Curtis metrics (Bray and Curtis 1957) were utilized to evaluate β -diversity which was then used to construct UPGMA trees and PCoA plots (Vázquez-Baeza et al. 2013). The uncertainty in the UPGMA clusters and PCoA plots was estimated by performing Jackknifing (repeatedly resampling a subset of the available data from each sample). A bootstrapped tree was then generated based on the jackknife data. The distance-based redundancy analysis (db-RDA), as implemented in Qiime 1.9, was utilized to evaluate the effect of "practice", "treatments", "Time" and "Location" variables based on Bray Curtis distance matrices. OTU frequencies between sample groups were determined using ANOVA and the Kruskal-Wallis test (Kruskal and Wallis 1952). In both tests, significance was determined using 999 Monte Carlo permutations, and the adjusted $FDR-P \leq 0.05$ was considered significant.

To confirm the accuracy of QIIME taxonomic assignments, representative sequences of each OTU that were either most abundant or relevant to the objectives of the study were extracted and introduced in ElimDupes (<http://hcv.lanl.gov/content/sequence/ELIMDUPES/elimdupes.html>) to detect identical sequences and determine their frequency. Unique representative sequences, defined as sequence types (STs) (Cf. Chapter II, Prigigallo et al. 2015), were analyzed along with genetically closely-related reference sequences of the same taxa to determine their phylogenetic association and enable their identification with the highest possible level of accuracy. To this aim, local databases of validated reference sequences were created with priority given to sequences from specific recent taxonomic studies. In cases where no matches were found in the reference sequence from selected published articles, more closely-related sequences were selected using a BLAST search of GenBank. The reliability of the latter sequences was evaluated based on different parameters, including the consistency of sequences from different sources, available details on sequenced isolates, and year of publication, giving priority to more recent items. For each selected taxa, representative sequences from the present study and reference sequences were aligned using MUSCLE (Edgar 2004) and introduced into MEGA6 for phylogenetic analysis utilizing the Maximum Likelihood method (Tamura et al. 2013). Analyses were performed using 1000 bootstrap replications.

Results and Discussion

After paired-end alignments, quality filtering, deletion of chimeric and plant sequences a total of 6,053,798 reads were recovered and assigned to 302,078 fungal OTUs. The number of OTUs per sample ranged from 59 to 9,748. After the exclusion of singletons, 5,760,162 reads were recovered and assigned to 8,504 fungal OTUs ranging from 24 to 906 OTUs per sample (Table V-1).

The rarefaction analysis indicated that the sequencing depth had been saturated in all of the analyzed samples indicating that the great majority of OTUs had been detected (Fig. V-1). According to the alpha diversity metrics (observed OTUs, Chao1 and Shannon), using a rarefied OTU table to an even depth of 5000 reads per sample, no significant differences were observed between the selected groupings: “practice”, “treatment” and “time”. In contrast, significant differences were revealed between the two

sampling locations: stem end (SE) and calyx end (CE). Based on a two-sample t-test, the difference between the sample locations was highly significant ($P = 0.001$) (Fig. V-2).

Table V-1 Summary of analyses and results of metabarcoding surveys conducted with harvested apples. Alpha diversity metrics were determined at an even depth of 5000 read.

Sample detail*	Counts	OTUs	Observed OTUs 5000	Shannon 5000	Chao1 5000
Conv.CE.Wshd.T1	163329	978	104.3	2.760231	403.985
Conv.CE.Wshd.T4	73656	535	85.8	2.085854	251.6303
Conv.CE.Wshd.T5	64253	486	105	2.812244	249.3905
Conv.CE.T.T1	219593	1015	100.3	2.842886	233.101
Conv.CE.T.T4	97462	826	102.9	1.603017	328.5642
Conv.CE.T.T5	63411	537	94.3	1.924802	321.5637
Conv.CE.unt.T1	175394	499	71.4	2.056759	147.4678
Conv.CE.unt.T4	75976	623	116.6	2.976273	290.297
Conv.CE.unt.T5	120719	515	71.4	1.799183	180.3253
Conv.SE.Wshd.T1	211972	833	115.4	3.01254	242.96
Conv.SE.Wshd.T4	35754	396	138.7	3.577853	287.2899
Conv.SE.Wshd.T5	174472	971	141.8	3.527069	307.1812
Conv.SE.T.T1	151359	912	135.2	3.396818	339.4237
Conv.SE.T.T4	44714	367	107.1	2.193123	255.6382
Conv.SE.T.T5	152067	699	110.8	2.762638	253.8533
Conv.SE.unt.T1	280573	1141	125.9	3.318782	268.2154
Conv.SE.unt.T4	107593	701	126.7	3.294148	257.7244
Conv.SE.unt.T5	97339	700	127.5	3.063759	321.7001
Org.CE.Wshd.T1	103229	513	80.2	2.781637	183.5699
Org.CE.Wshd.T4	99425	436	93.3	2.261844	181.533

Org.CE.Wshd.T5	194610	1354	119	2.851743	347.5681
Org.CE.T.T1	225745	1330	141.3	3.638272	326.5776
Org.CE.T.T4	183475	1102	134.6	3.557913	297.9193
Org.CE.T.T5	55033	378	98.5	2.689433	232.5971
Org.CE.unt.T1	90136	776	111.1	2.085916	357.5585
Org.CE.unt.T4	362013	1927	124.8	2.884427	327.3815
Org.CE.unt.T5	435348	2080	121.7	3.234535	305.0563
Org.SE.Wshd.T1	145043	854	132.6	3.348298	309.4131
Org.SE.Wshd.T4	248429	1330	111.8	3.135098	307.8748
Org.SE.Wshd.T5	90515	563	130.4	2.798456	267.6749
Org.SE.T.T1	224921	1296	142.7	3.337671	324.9656
Org.SE.T.T4	144180	798	132.4	3.089098	321.222
Org.SE.T.T5	121899	869	174.3	3.063326	384.0982
Org.SE.unt.T1	188957	852	112	2.753823	252.6056
Org.SE.unt.T4	402992	1621	136.5	3.462392	304.7641
Org.SE.unt.T5	134576	693	133.9	3.58281	284.228

* Abbreviations: Conv. = Conventional; Org. = Organic; T= treated; Wshd= washed; unt = untouched; SE = Stem End; CE = Calyx End.

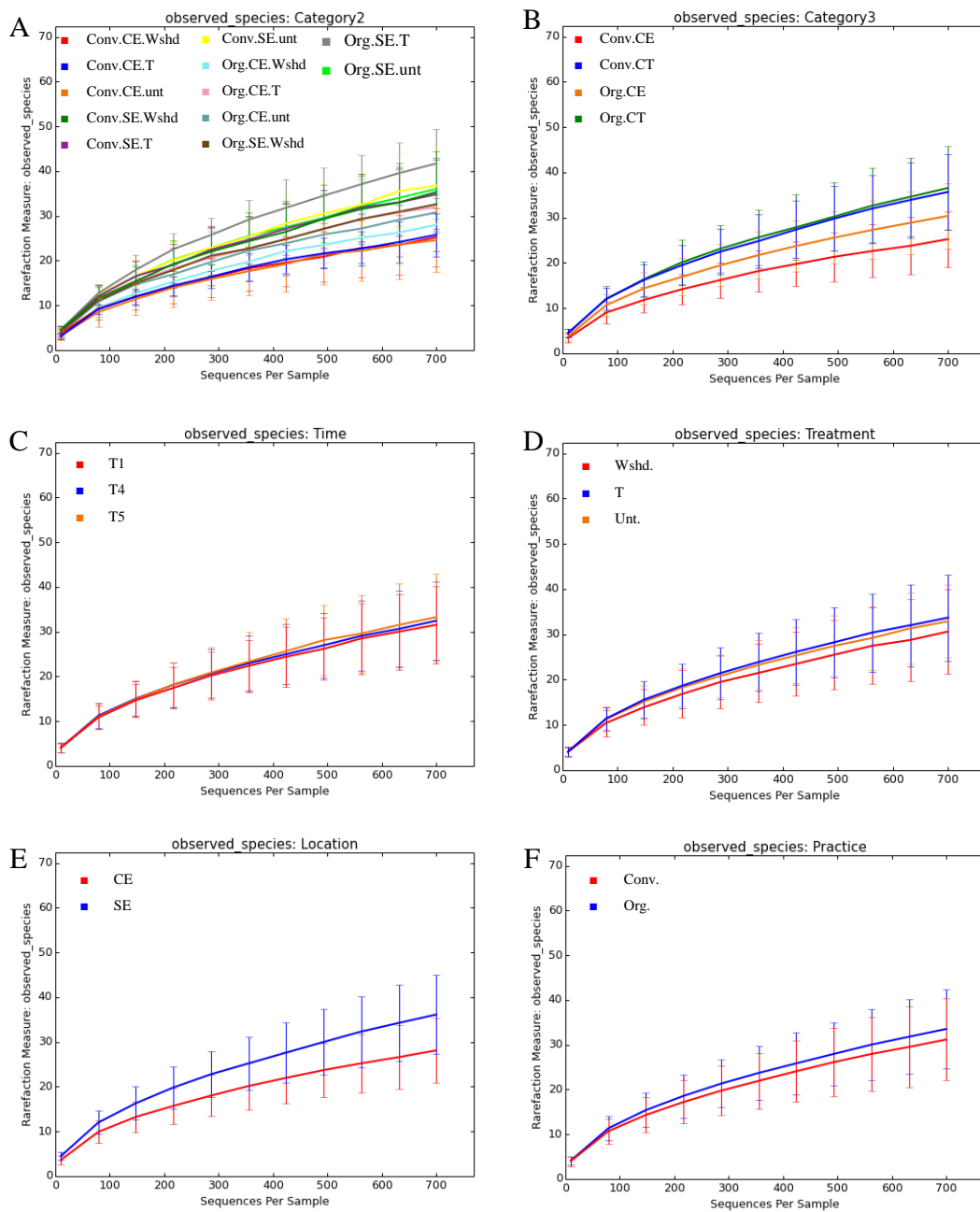


Figure V-1 Rarefaction curves determined for all samples investigated in the present study. A) Samples collapsed to show all categories regardless of Time B) Samples collapsed regardless of time and treatments C) samples collapsed regardless of practice, sample location, and treatments. D) Samples collapsed based on treatments E) Samples collapsed based on sampling location F) samples collapsed based on practice. Abbreviations: Conv. = Conventional; Org. = Organic; T= treated; Wshd= washed; unt = untouched; SE = Stem End; CE = Calyx End.

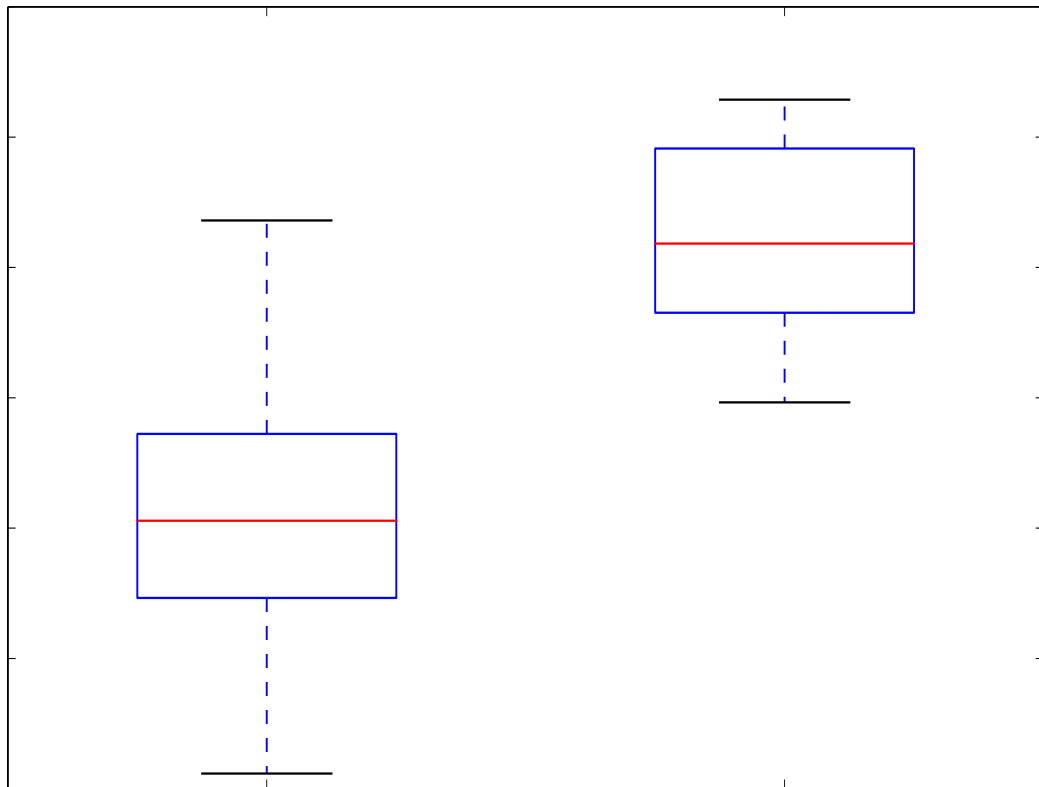


Figure V-2 Boxplot visualizing results of the nonparametric two sample t-test to compare Shannon alpha diversity of in Calyx End (CE) on the right and Stem End (SE) on the left. Results of nonparametric two sample t-tests to compare the alpha diversities based on Shannon index using Monte Carlo permutations (999). ($P=0.001$)

Impact of management practices (organic vs. conventional) on the fungal communities found on apple (“practices” effect)

Untouched samples of conventional vs. organic fruit sampled at T1 exhibited significant differences ($P=0.006$) in beta diversity based on the db-RDA analysis. Differences were significant in both SE and CE samples with values of $P=0.007$ and 0.001 , respectively (Table V-2).

Differences in fungal communities between apple stem end (SE) and calyx end (CE) samples (“location” effect);

The comparison of the fungal community in SE vs. CE samples revealed significant differences in untouched samples at T1 ($P=0.001$) (Table V-2). A high level of significance was confirmed for both organic ($P=0.016$) and conventional ($P=0.01$) apples. Moreover, SE and CE clustered in separate groups in the PCoA ($P=0.001$), regardless of “practice”, “treatment” or “time” (Fig. V-3). The different fungal populations detected in SE vs. CE samples most likely reflects different microclimatic and niche conditions present in these two locations of apple fruit. Furthermore, the presence of senesced stamens in the CE may have favored the colonization of necrotrophic fungi, as previously demonstrated for *Botrytis cinerea* on grapes (Sanzani et al. 2012).

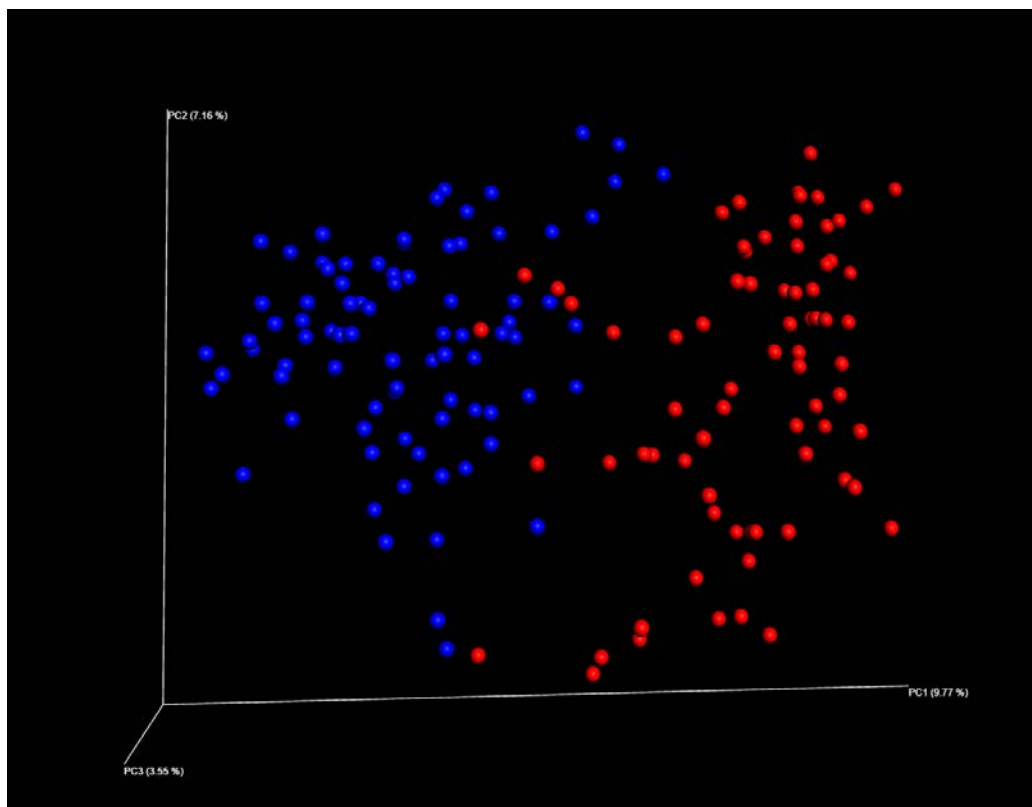


Figure V-3 3 dimensional PCoA plots based on Bray Curtis distance metrics at an even depth of 5000 reads. Red and blue spheres represent fruit CE and SE respectively. According to Distance-based redundancy analysis (db-RDA) the tested groups significantly differed ($P=0.001$).

Table V-2 Comparison of fungal communities according to Bray Curtis beta diversity metrics. The significance was determined through 999 Monte Carlo permutations using the distance-based redundancy analysis (db-RDA). Sample groups were considered significantly different if the P value was ≤ 0.05 . • Conv. = Conventional; Org. = Organic; T= treated; Wsh= washed; unt = untouched; SE = Stem End; CE = Calyx End.

i) Impact of management practices (organic vs conventional) “practices” effect	Significance
Compare Unt. T1 (Conv. vs. org.)	$P= 0.006^*$
Compare Unt. T1 CE (Conv. vs. org.)	$P= 0.001^*$
Compare Unt. T1 SE (Conv. vs. org.)	$P= 0.007^*$
ii) Differences between apple’s stem end (SE) and calyx end (CE) “location” effect	
Unt T1 (CE vs. SE)	$P= 0.001^*$
Org. Unt T1 (CE vs. SE)	$P= 0.016^*$
Conv. Unt T1 (CE vs. SE)	$P=0.01^*$
iii) Impact of hot water treatments on fungal communities “treatment” effect	
Comparison between Conv. CE (Unt. T1 vs. Wsh. T1 vs. T T1)	$P= 0.05^*$
Comparison between Conv. CE (Wsh. T1 vs. T T1)	$P= 0.05^*$
Comparison between Conv. CE (Unt. T1 vs. T T1)	$P= 0.21$
Comparison between Conv. CE (Unt. T1 vs. Wsh. T1)	$P= 0.077$
Comparison between Conv. SE (Unt. T1 vs. Wsh. T1 vs. T T1)	$P= 0.126$
Comparison between Conv. SE (Wsh. T1 vs. T T1)	$P= 0.203$
Comparison between Conv. SE (Unt. T1 vs. T T1)	$P= 0.044^*$
Comparison between Conv. SE (Unt. T1 vs. Wsh. T1)	$P= 0.539$
Comparison between Org. CE (Unt. T1 vs. Wsh. T1 vs. T T1)	$P= 0.009^*$
Comparison between Org. CE (Wsh. T1 vs. T T1)	$P= 0.171$
Comparison between Org. CE (Unt. T1 vs. T T1)	$P= 0.042^*$
Comparison between Org. CE (Unt. T1 vs. Wsh. T1)	$P= 0.013^*$
Comparison between Org. SE (Unt. T1 vs. Wsh. T1 vs. T T1)	$P= 0.441$
Comparison between Org. SE (Wsh. T1 vs. T T1)	$P= 0.087$
Comparison between Org. SE (Unt. T1 vs. T T1)	$P= 0.68$
Comparison between Org. SE (Unt. T1 vs. Wsh. T1)	$P= 0.596$
Comparison between Org. CE (Unt. T5 vs. Wsh. T5 vs. T T5)	$P= 0.02^*$
Comparison between Org. SE (Unt. T5 vs. Wsh. T5 vs. T T5)	$P= 0.294$
Comparison between Conv. CE (Unt. T5 vs. Wsh. T5 vs. T T5)	$P= 0.43$
Comparison between Conv. SE (Unt. T5 vs. Wsh. T5 vs. T T5)	$P= 0.109$
iv) Effect of time after hot water treatments “time” effect	
Comparison between Org. CE (T T1 vs. T T4 vs. T T5)	$P= 0.209$
Comparison between Org. SE (T T1 vs. T T4 vs. T T5)	$P= 0.035^*$
Comparison between Conv. CE (T T1 vs. T T4 vs. T T5)	$P= 0.369$
Comparison between Conv. SE (T T1 vs. T T4 vs. T T5)	$P= 0.029^*$
Comparison between Org. CE (Wsh. T1 vs. Wsh. T4 vs. Wsh. T5)	$P= 0.098$
Comparison between Org. SE (Wsh. T1 vs. Wsh. T4 vs. Wsh. T5)	$P= 0.136$
Comparison between Conv. CE (Wsh. T1 vs. Wsh. T4 vs. Wsh. T5)	$P= 0.404$
Comparison between Conv. SE (Wsh. T1 vs. Wsh. T4 vs. Wsh. T5)	$P= 0.64$
Comparison between Org. CE (Unt. T1 vs. Unt. T4 vs. Unt. T5)	$P= 0.046^*$
Comparison between Org. SE (Unt. T1 vs. Unt. T4 vs. Unt. T5)	$P= 0.344$
Comparison between Conv. CE (Unt. T1 vs. Unt. T4 vs. Unt. T5)	$P= 0.452$
Comparison between Conv. SE (Unt. T1 vs. Unt. T4 vs. Unt. T5)	$P= 0.05^*$

Effect of hot water treatment on fungal communities ("treatment" effect)

A comparison of treated (hot water dip) and washed (room temperature water) conventional apples indicated a significant effect of the hot water dip only at T1 in CE samples ($P=0.05$) (Table 2). In organic apples, sampled at T1, significant differences were observed in between untouched and treated CE samples ($P=0.04$), and between untouched and washed samples ($P=0.01$). However, no significant differences were observed between washed and treated apples in either CE or SE samples ($P=0.17$). After 2 weeks (T5), these results were only consistent in the organic CE samples ($P=0.02$). In contrast, no significant differences between treatments were observed in the SE of either organic or conventional apples (Table V-2).

Effect of time following hot water treatment on the fungal community ("time" effect)

A comparison of fungal communities in treated (hot water dip) apples over time (T1, T4, T5) revealed significant differences in SE samples of both organic and conventional fruit ($P=0.035$ and $P=0.029$, respectively) (Table 2). In contrast, no significant differences were revealed in the CE of either organic or conventional fruit. Considering all the treatments (untouched, washed, hot water dip), significant differences over the time were observed only in CE samples of untouched organic apples ($P=0.046$) and in SE samples of conventional untouched apples ($P=0.05$).

Apple fungal communities

Generally, members of the phylum Ascomycota were dominant in all samples and accounted for 91.6% of the total number of detected sequences. This was followed by Basidiomycota (8%), Chytridiomycota (0.1%), and unidentified fungi (0.3%). Rozellomycota and Zygomycota were also detected but at a very low frequency (Fig. V-4A).

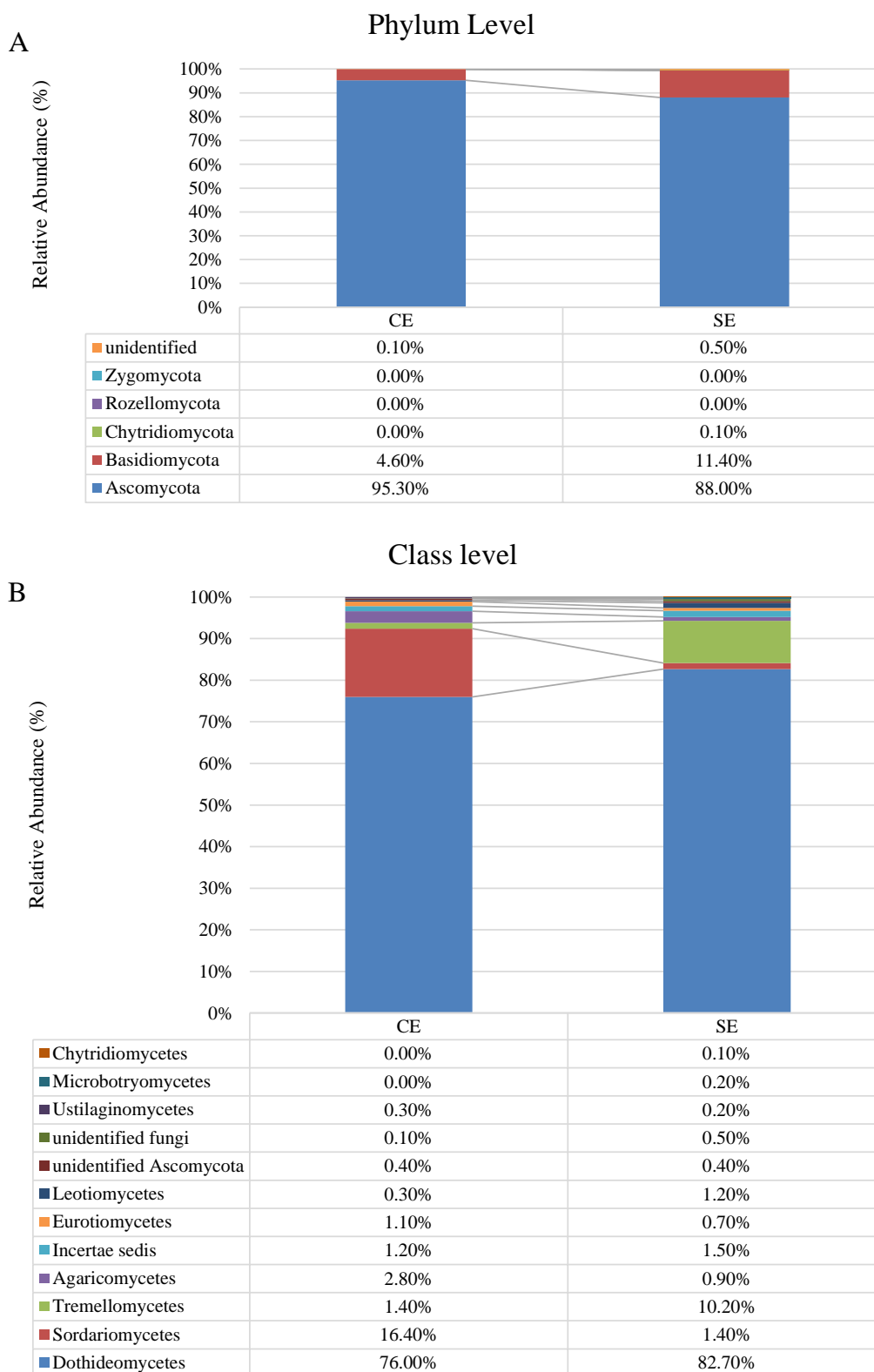


Figure V-4 Relative abundance of the fungal phyla (A); Class (B) of the taxa detected in apple fruit at the two sampling locations: Calyx End (CE) on the left and Stem End (SE) on the right.

The incidence of different phyla varied significantly between the locations (SE and CE). For example, Ascomycota represented 88.2% in SE samples and 95.3% in CE samples (FDR $P=1.12E-15$), whereas Basidiomycota had an RA of 11.4% in SE samples and 4.6% in CE samples (FDR $P=1.12E-15$) (Table V-3).

Table V-3 OTUs comparison at the phylum level between samples location using the nonparametric Kruskal-Wallis test. The table reports the P value, the P value corrected by the Benjamini-Hochberg FDR and the Bonferroni procedures. The groups means represent the mean of the genus relative abundance over all samples regardless of practice, treatment or time. Genera reported here were detected with a relative abundance (RA) of at least 1%.

OTU	Test-Statistic	P	FDR P	Bonferroni P	CE mean	SE mean
Ascomycota	64.44264	9.94E-16	1.12E-15	1.99E-15	0.979652	0.903509
Basidiomycota	64.2027	1.12E-15	1.12E-15	2.25E-15	0.020348	0.096491

Ascomycota were largely identified as members of the classes Dothideomycetes and Sordariomycetes (79.3% and 8.9%, respectively), followed by the classes, Incertae sedis 1.40%, Eurotiomycetes 0.90%, Leotiomycetes 0.70%, and unidentified Ascomycota 0.40%. Members of the Dothideomycetes were more abundant in SE (82.7%) samples than in CE (76.0%) samples. In contrast, members of the class Sordariomycetes were more abundant in CE (16.4%) samples than in SE (1.4%) samples. Additionally, Tremellomycetes (5.8%) was the predominant class of Basidiomycota. Members of this class were more represented in SE samples than in CE (10.2% and 1.4%, respectively) samples (Fig. V-4B).

Across all samples, *Davidiella* was the most abundant genus detected with an overall relative abundance (RA) of 33.7%. It had a RA of 46.90% and 18.50% in apple CE and SE samples, respectively (FDR $P=6.37E-14$). The genus *Didymella*, across all samples, had an RA of 17.9% which was 4.60% and 31.20% in CE and SE samples, respectively (FDR $P=7.34E-16$). A high RA (10.40%) was also revealed for the genus *Alternaria* (6.70% and 14.10% in CE and SE samples, respectively), followed by a group of unidentified Dothideomycetes (7.50% and 5.50% in CE and SE samples, respectively), *Cryptococcus* (1.30% and 10.00% in CE and SE samples, respectively), *Acremonium* (9.90% and 0.50% in CE and SE samples, respectively) and *Aureobasidium* (0.20% and 6.30% in CE and SE samples, respectively). A RA higher than 1% was also revealed for *Cladosporium*, *Lewia*, *Stilbella*, unidentified Agaricales and *Phaeosphaeria* (Table V-4, Fig. V-5 Fig. V-6). Moreover, other fungal taxa with a lower RA including *Lophiostoma*, *Phoma*, *Coprinellus*, a group of unidentified Ustilaginales, a group of unidentified

Ustilaginaceae, *Ustilago*, and *Tumularia* were also characterized by significant differences in abundance in CE vs. SE samples (Table V-5).

Table V-4 OTUs frequencies comparison at the genus level between samples location using the nonparametric Kruskal-Wallis test. The table reports the *P* value, the *P* value corrected by the Benjamini-Hochberg FDR and the Bonferroni procedures. The groups means represent the mean of the genus relative abundance across all samples regardless of practice, treatment or time. Genera reported here were detected with a relative abundance (RA) of at least 1%.

OTU	Test-Statistic	<i>P</i>	FDR_ <i>P</i>	Bonferroni <i>P</i>	CE mean*	SE mean*
<i>Aureobasidium</i>	85.70793	2.09E-20	2.50E-19	2.50E-19	0.003839	0.069498
<i>Cryptococcus</i>	79.58010	4.63E-19	2.78E-18	5.56E-18	0.003730	0.077647
<i>Didymella</i>	67.77201	1.84E-16	7.34E-16	2.20E-15	0.069236	0.380248
<i>Davidiella</i>	58.41512	2.12E-14	6.37E-14	2.55E-13	0.592163	0.190610
<i>Cladosporium</i>	29.69494	5.06E-08	1.21E-07	6.07E-07	0.040846	0.014982
Unidentified Dothideomycetes	5.06394	5.55E-07	1.11E-06	6.66E-06	0.056906	0.058955
<i>Alternaria</i>	18.05899	2.14E-05	3.67E-05	0.000257	0.082124	0.168716
<i>Stilbella</i>	11.795699	0.000594	0.000891	0.007124	0.009624	0.001855
<i>Phaeosphaeria</i>	9.122543	0.002525	0.003366	0.030297	0.011423	0.007992
<i>Lewia</i>	8.922716	0.002816	0.003380	0.033797	0.031750	0.020088
unidentified Agaricales	7.237326	0.007140	0.007789	0.085684	0.027905	0.009150
<i>Acremonium</i>	5.153917	0.023194	0.023194	0.278328	0.070453	0.000258

Abbreviations: SE = Stem End; CE = Calyx End.

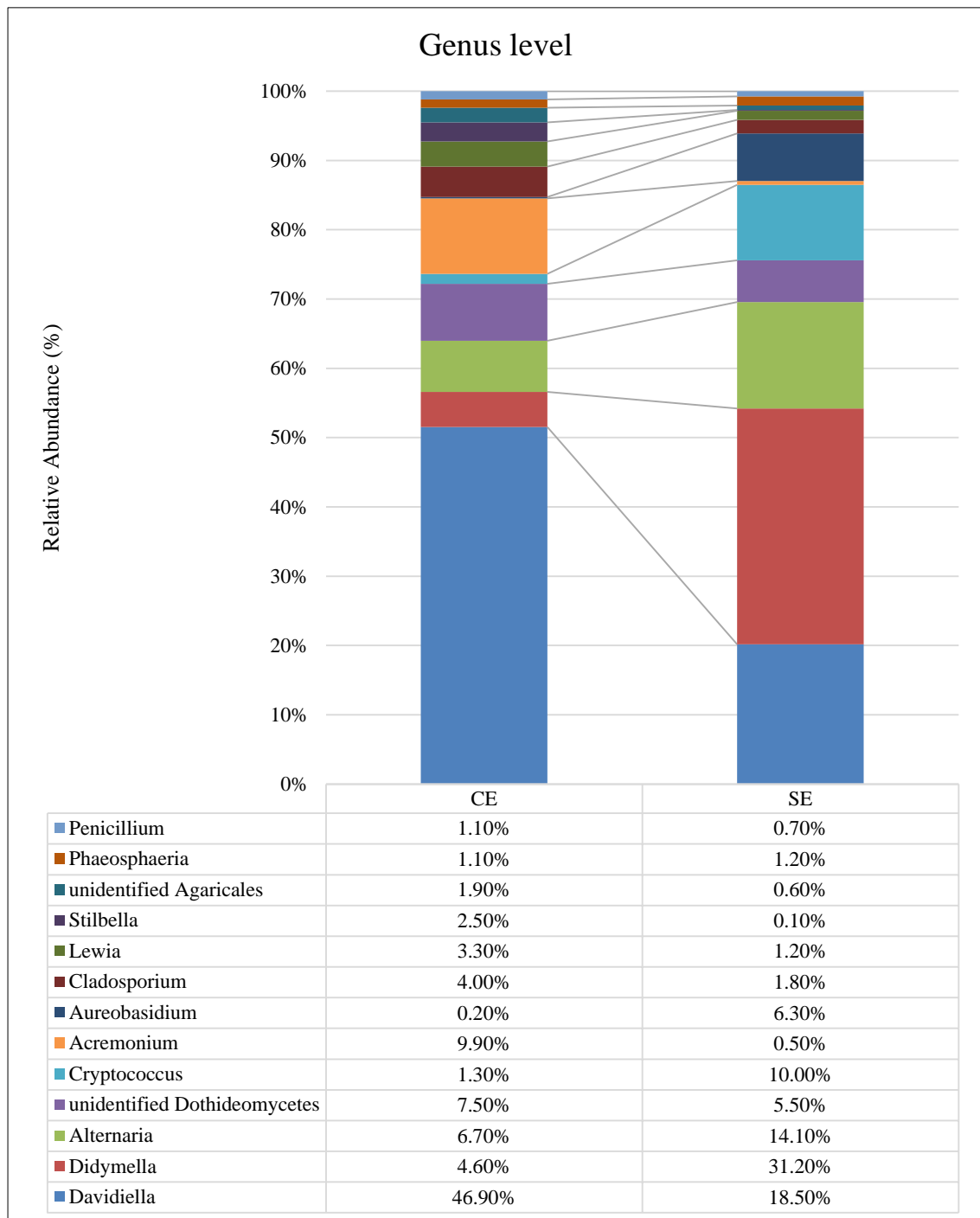


Figure V-5 Relative abundance of the fungal genera of the taxa detected in apple fruit at the two sampling locations: Calyx End (CE) on the left and Stem End (SE) on the right.

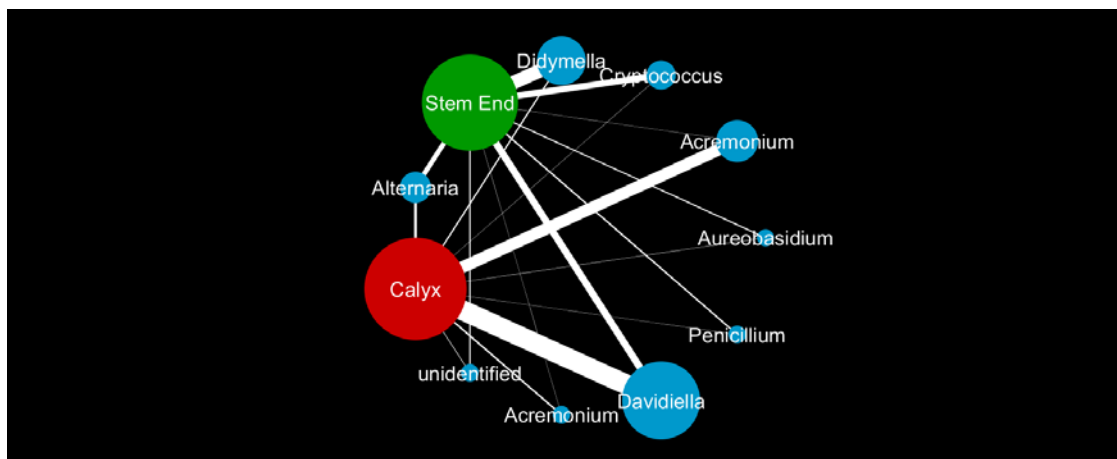


Figure V-6 A network displaying the shared phylotypes between stem end (Green) and calyx (Red) samples. Nods in blue represent OTUs, the size of the node reflects the abundance of the OTU across all samples. The width of the edges connecting between samples and OTUs, reflects the abundance of the OTUs in a specific sample.

Table V-5 OTUs frequencies comparison at the genus level between samples location using the nonparametric Kruskal-Wallis test. The table reports the *P* value, the *P* value corrected by the Benjamini-Hochberg FDR and the Bonferroni procedures. The groups means represent the mean of the genus relative abundance at Time 0 untouched samples regardless of practice. Genera reported here were detected with a relative abundance (RA) of at least 1%.

Genera	Test-Statistic	<i>P</i>	FDR <i>P</i>	Bonferroni <i>P</i>	CE mean	SE mean
<i>Davidiella</i>	13.5	0.000239	0.005356	0.010974	0.286464	0.123947
<i>Cryptococcus</i>	12.90667	0.000327	0.005356	0.015061	0.042298	0.126357
<i>Lophiostoma</i>	12.78554	0.000349	0.005356	0.016068	0	0.009007
unidentifed Pleosporales	10.76106	0.001037	0.008376	0.047683	0.010346	0.033968
<i>Didymella</i>	10.66667	0.001091	0.008376	0.050178	0.080777	0.125167
<i>Phoma</i>	10.66384	0.001093	0.008376	0.050255	0.002455	0.015981
<i>Cladosporium</i>	10.14	0.001451	0.009534	0.06674	0.068074	0.038945
<i>Coprinellus</i>	9.548718	0.002001	0.011505	0.092041	0.000741	0.010608
unidentifed Ustilaginales	9.13468	0.002508	0.012819	0.115372	0.012981	0.032722
unidentifed Ustilaginaceae	8.830328	0.002963	0.013628	0.136282	0	0.005872
<i>Ustilago</i>	8.613333	0.003337	0.013955	0.153507	0.001891	0.019574
<i>Lewia</i>	6.406667	0.011369	0.043582	0.522986	0.05601	0.033436
<i>Tumularia</i>	6.263163	0.012327	0.04362	0.56706	0.003693	0.015354

Identification of fungal species

The genus *Davidiella* (33.7%), was represented by 3061 OTUs but most sequences were associated with a single OTU (New.ReferenceOTU55). The number of OTUs was reduced to 26, however, when a cutoff of 100 sequences per OTU was used. The phylogenetic analysis of the most abundant sequence within these OTUs (Fig. V-7), did not enable the identification of the species due to the complex taxonomy within the genus (Bensch et al. 2010, Behr et al. 2013). However, this analysis revealed a close association of the most abundant ST to the species *Davidiella tassiana*, the teleomorph of *Cladosporium herbarum* (De Not.) Crous & U. Braun (Braun et al. 2003). Although this fungus was isolated from rooted apples in cold storage (Heald and Ruehle 1931), the pathogenicity of *C. herbarum* has not been characterized. *D. tassiana* was also isolated as an endophyte in grapevines (Pancher et al. 2012). Results of various studies suggest that the ability of *D. tassiana* to parasitize the plant hosts examined is generally limited. The genus *Davidiella* was proposed to accommodate the teleomorphs of the genus *Cladosporium s. str.* (Braun et al. 2003). Whereas the family Davidiellaceae was introduced as a new family to accommodate species of *Davidiella* with *Cladosporium* anamorphs (Schoch et al. 2006).

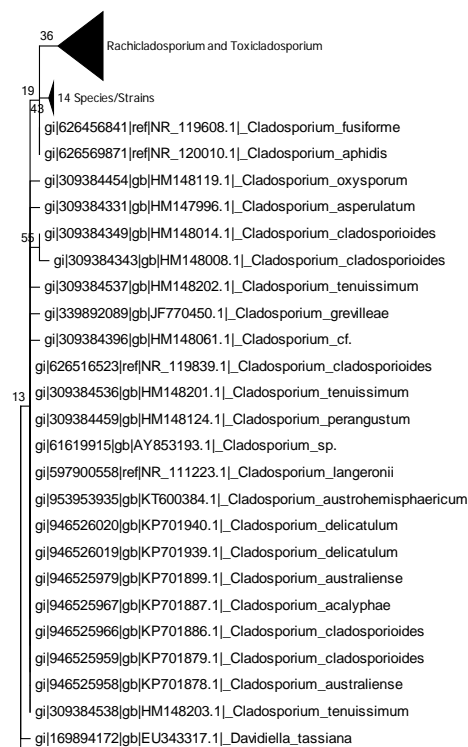


Figure V-7 Molecular Phylogenetic analysis of *Davidiella* by Maximum Likelihood method. Sequences of this study are labeled with a bullet and the number of sequences of the most abundant OTU is indicated between brackets.

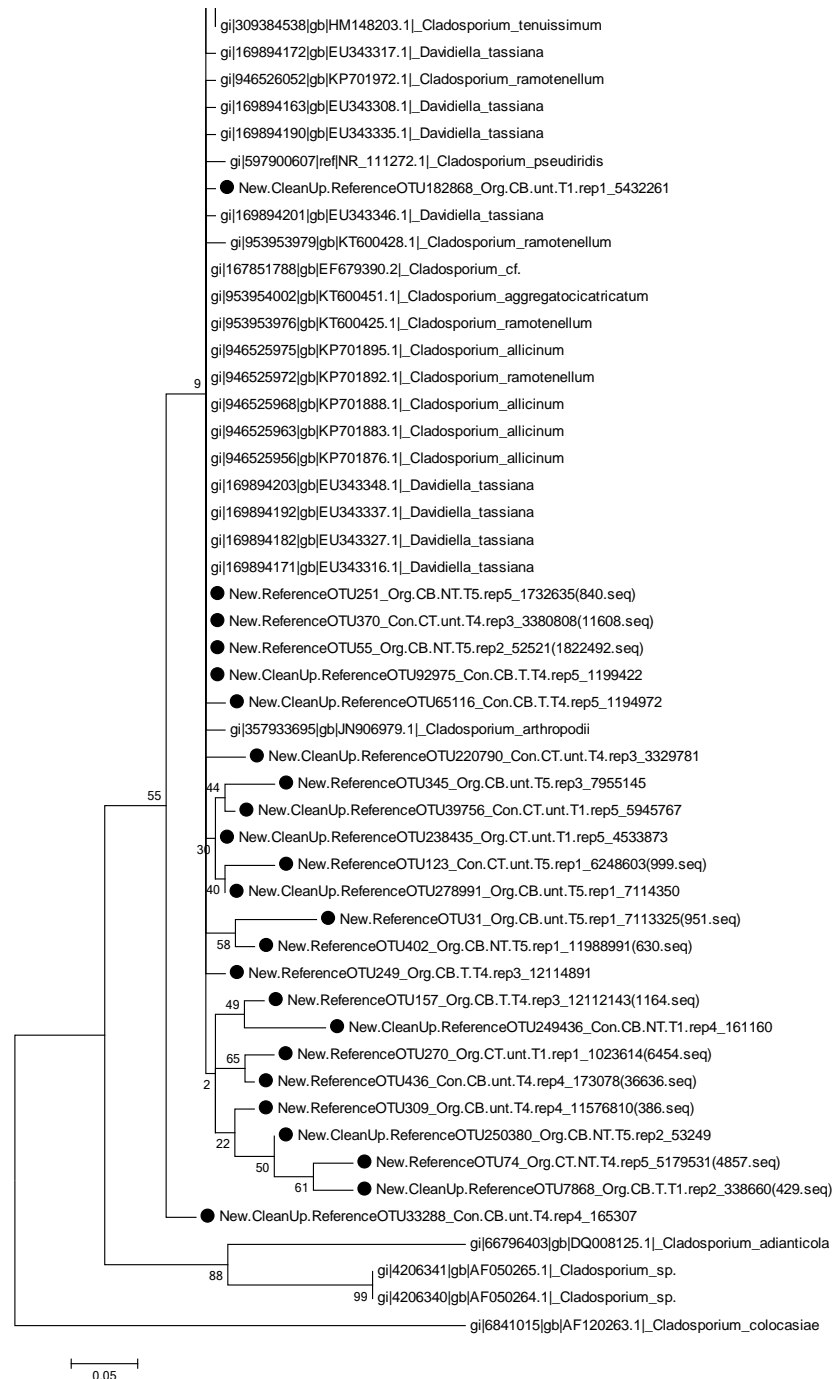


Figure V-7 continued

The genus *Didymella*, the second most abundant genus with a RA of 17.90%, was represented by 321 OTUs. Most sequences (993,808) were associated with a single OTU (New.ReferenceOTU179) with 14 OTUs containing at least 100 sequences. The phylogenetic analysis of representative sequences, along with reference sequences from the CBS *Didymella* collection in NCBI, did not enable a precise identification of the species since the most abundant OTU clustered with 6 different species (Fig. V-8). The genus *Didymella*, is considered to be the teleomorph of *Phoma s. str.* (Aveskamp et al.

2010). The majority of the taxa of *Phoma s. str.* are phytopathogenic fungi associated with land plants and mainly cause leaf and stem spots (Aveskamp et al. 2008, Aveskamp et al. 2009, Zhang et al. 2009). Some species of *Phoma* are able to switch from a saprophytic to a pathogenic lifestyle when a suitable host is encountered (Aveskamp et al. 2008, Aveskamp et al. 2009). The genus further comprises several species and varieties that are recognized as endophytic, fungicolous, and lichenicolous fungi (Aveskamp et al. 2010). *Phoma* species are known to cause stalk end rot in apple (Moore 1959). More particularly, most of the putative species detected in the present study, including *D. Glomerata*, *D. pomorum*, and *Phoma macrostoma*, are known to cause leaf spot diseases in apple (Sharma and Bhardwaj 1999). The association of this genus with leaf and stalk (stem end) diseases, in addition to being sometimes a secondary invader (Boerema et al. 1977) explain the higher abundance in SE than CE samples. Other OTUs related to this genus, according the phylogenetic tree, were distant from all of the analyzed reference sequences, suggesting the existence of unknown *Phoma* or *Didymella* species.

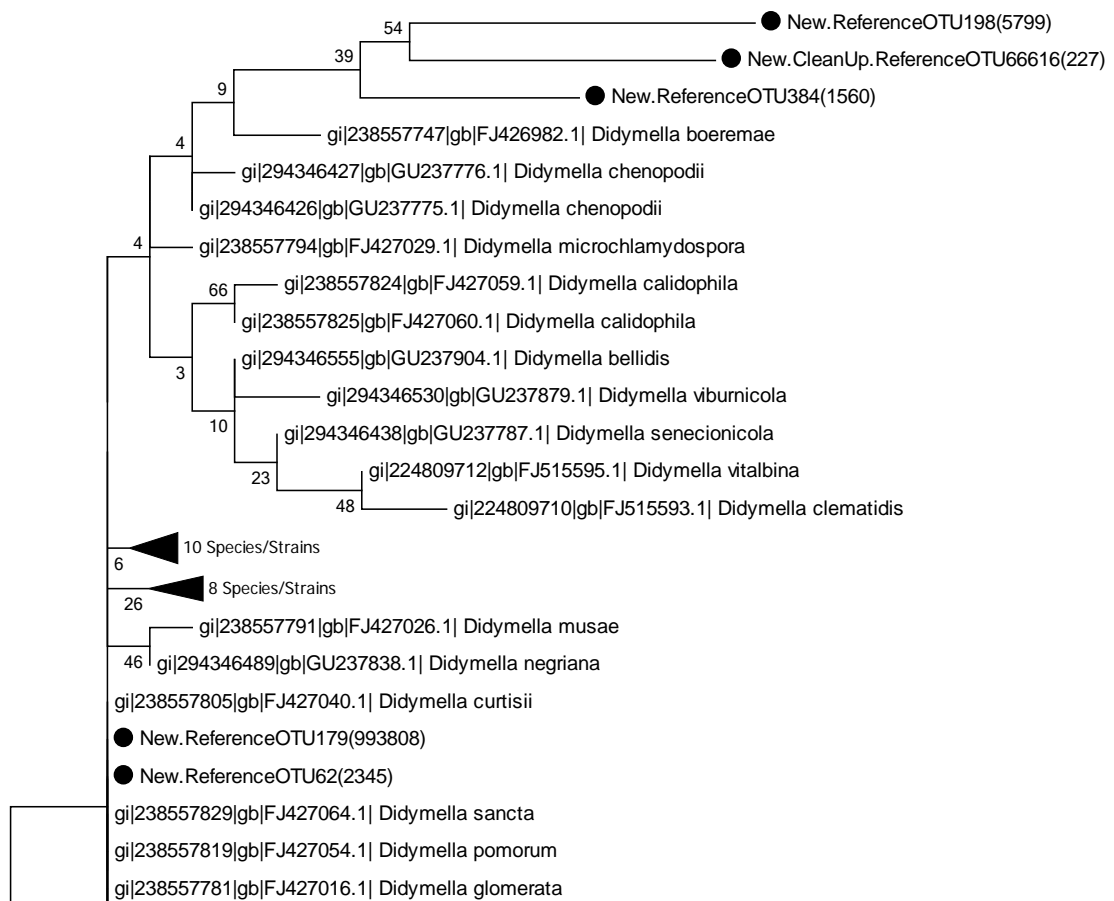


Figure V-8 Molecular Phylogenetic analysis of *Didymella* by Maximum Likelihood method Sequences of this study are labeled with a bullet and the number of sequences of the most abundant OTU is indicated between brackets.

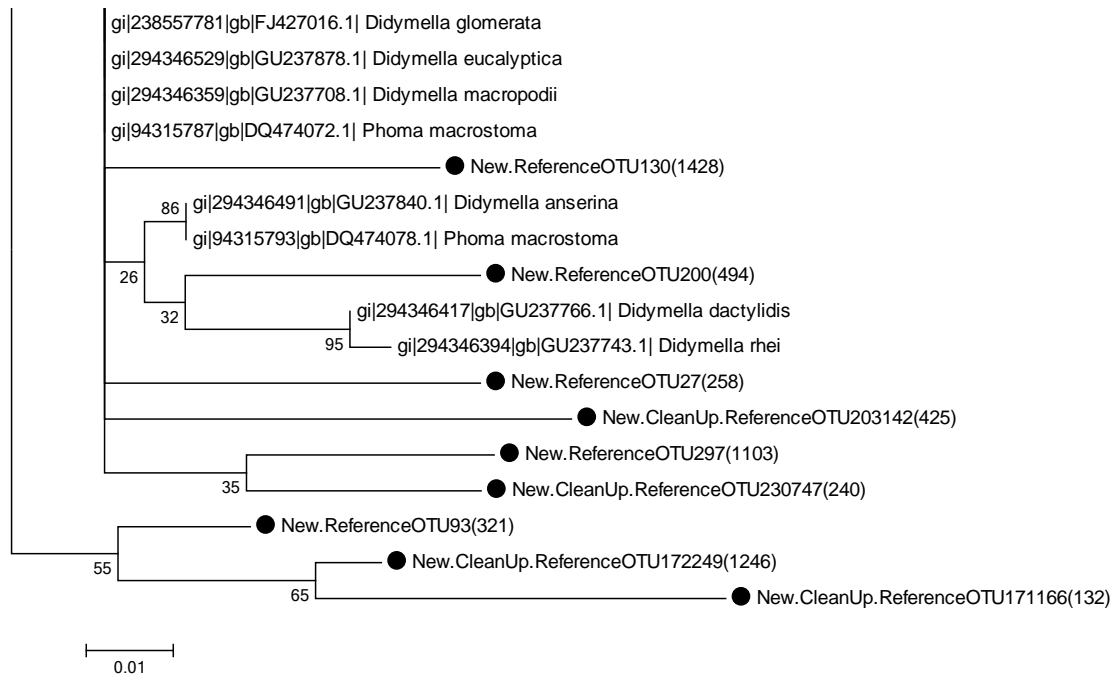


Figure V-8 continued

The genus *Alternaria* had an overall RA of 10.40% with a total of 276 OTUs and 14 OTUs containing at least 100 sequences. The phylogenetic analysis of the most abundant OTUs along with validated reference sequences of the genus *Alternaria* (Woudenberg et al. 2013) revealed a perfect similarity to *A. alternata*. However, the accurate identification of the species was not possible due to the existence of several closely related species within the genus (*A. arborescens*, *A. longipes*, *A. limoniasperae*, and *A. daucifolii*). On the other hand, the phylogenetic analysis did not confirm the identity of the other less abundant STs grouped within *Alternaria* according to Qiime analysis (Fig. V-9). According to a BLAST search, these sequences appeared to be slightly outside of this genus as they had 99% similarity to *Phoma* and *Peyronellaea*. The genus *Lewia* was also detected in this study, with an RA of 2.30%. *Lewia infectoria*, the teleomorph of *A. infectoria* (Perelló and Sisterna 2008), is a very common endophyte of apple (Ostenfeld Larsen et al. 2003) and was found to be associated with core rot of apples (Serdani et al. 1998). The genus includes species that are saprobic, endophytic, and pathogenic to human, animals, plants and can be found almost anywhere, including soil, air, plants, and plant products. In apple, *Alternaria* species cause two important diseases including *Alternaria* core rot (Ntasiou et al. 2015) and *Alternaria* blotch (Johnson et al. 2000). Species of this genus are also well known for their production of mycotoxins (Ntasiou et al. 2015), which can be a health risk especially in processed apple products.

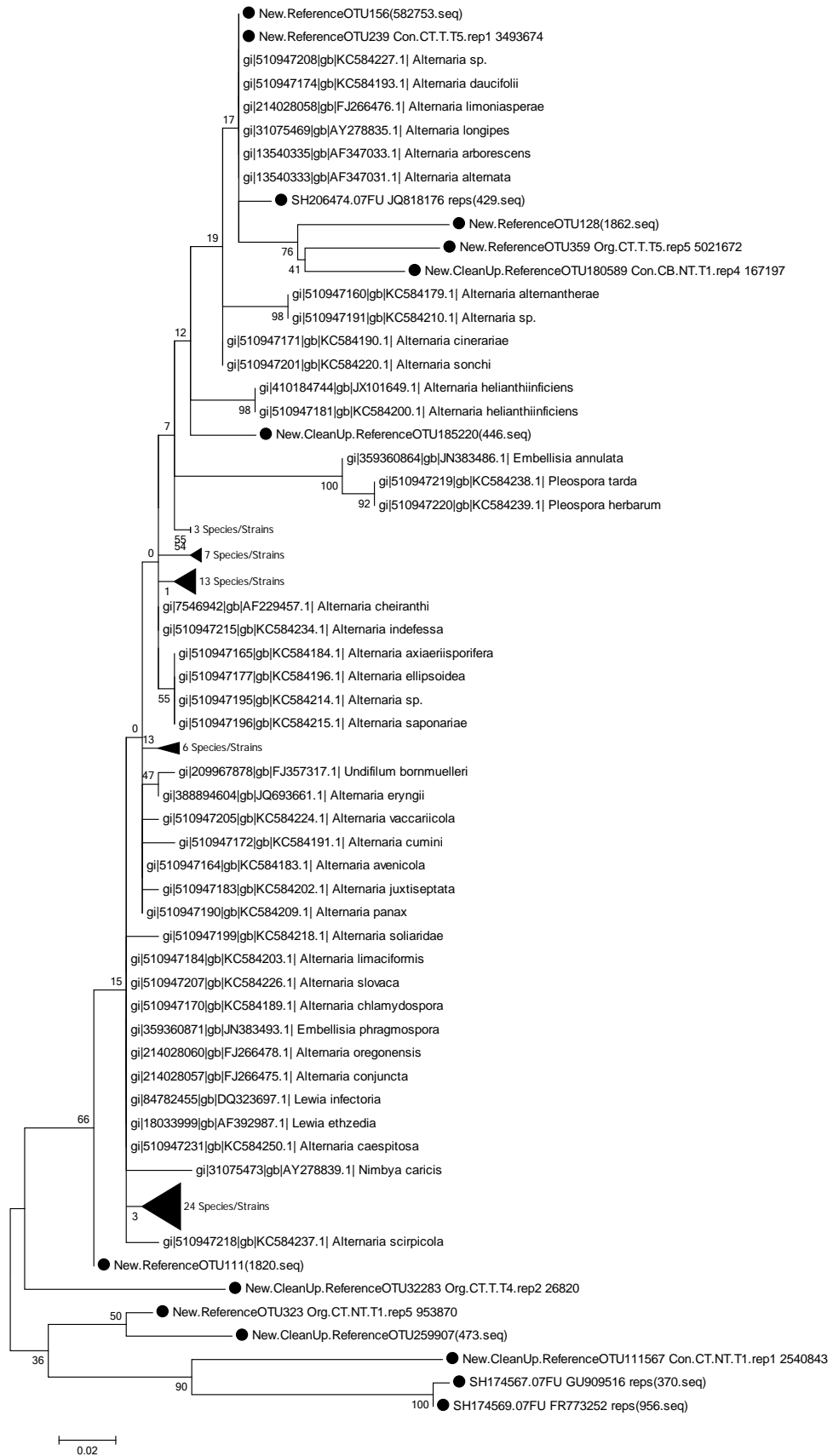


Figure V-9 Molecular Phylogenetic analysis of *Alternaria* by Maximum Likelihood method. Sequences of this study are labeled with a bullet and the number of sequences of the most abundant OTU is indicated between brackets.

The genus *Acremonium* had an RA of 5.20% and was represented by 305,517 sequences and 456 OTUs with 29 OTUs having more than 100 sequences. According to Qiime and phylogenetic analyses most of these sequences belonged to the species *A. fusidioides* (Fig. V-10). The second most abundant OTU (New.ReferenceOTU236), containing 68,469 sequences, was only identified at the level of genus, although it seemed to be mainly related to *A. alternatum*. Other detected OTUs clustered with *A. parvum*, *A. charticola*, *A. sclerotigenum* and *A. persicinum*. Finally some OTUs clustered together forming a different group not associated with any currently accepted species suggesting the existence of a still unknown new clade. *Acremonium* species are known to be endophytes, saprobes in air or soil, or even as plant, animal, and human pathogens (Hoveland 1993, Perdomo et al. 2011, Xu et al. 2015). *A. sclerotigenum* was recently reported to cause *Acremonium* brown spot on bagged apple fruit in China (Bogomolova and Kirtsideli 2009, Li et al. 2014). Interestingly, there are no records in the literature about *A. fusidioides* on apple or as a plant pathogen. The species *A. alternatum* is known for its hyperparasitic activity against powdery mildew (Malathrakis 1985) and was reported to reduce clubroot symptoms in *Arabidopsis* (Jäschke et al. 2010). *A. alternatum* is also known as an endophyte that has an effect on the development and nutrition of larvae of the diamondback moth, *Plutella xylostella* (Raps and Vidal 1998).

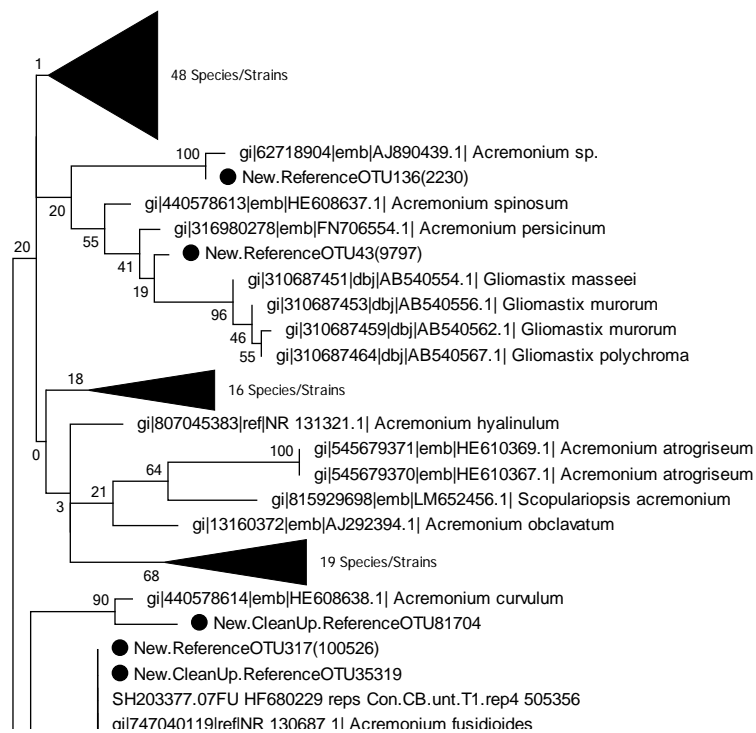


Figure V-10 Molecular Phylogenetic analysis of *Acremonium* by Maximum Likelihood method Sequences of this study are labeled with a bullet and the number of sequences of the most abundant OTU is indicated between brackets.

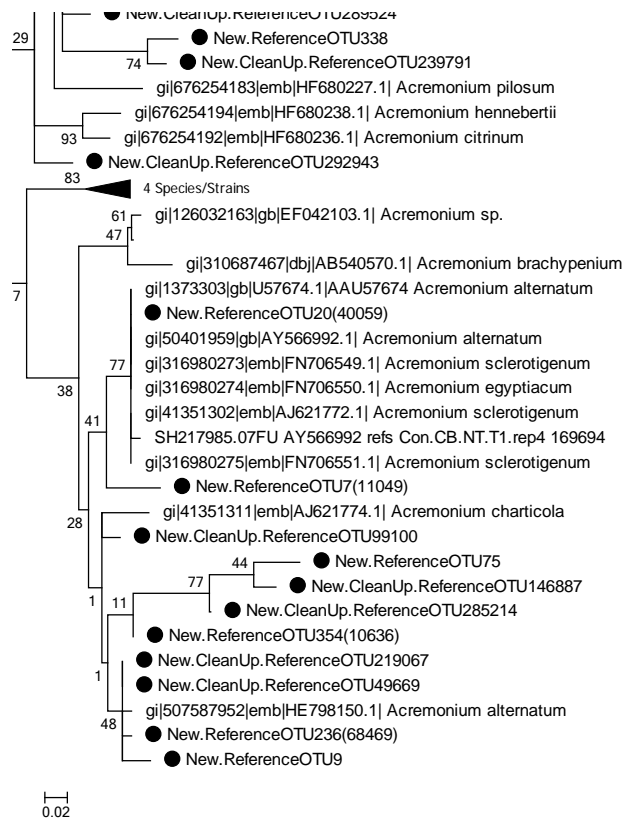


Figure V-10 continued

The genus *Cryptococcus* had a RA of 5.60% accounting for 224 OTUs containing 318,857 sequences. While 21 OTUs contained at least 100 sequences, a single OTU “New.ReferenceOTU182” contained 228,018 sequences. The identification of species within the genus *Cryptococcus* is very difficult due to the high complexity of the genus (Takashima et al. 2003). However, according to Qiime and BLAST analyses the most abundant OTUs, containing 243,077 sequences, likely to belong to the species, *C. magnus* (KR912263.1). Other sequences showed a high level of similarity (99-100%) to several other species including *C. victoriae*, *C. bhutanensis*, *C.tephrensensis*, *C. dimennae*, *C. saitoi* and *C. heimaeyensis*. Some species of *Cryptococcus* are known for their antagonistic activity against apple blue mold caused by *Penicillium expansum* (Hashem et al. 2014) and grey mold in grapes (Ligorio et al. 2007, Meng et al. 2010).

The genus *Aureobasidium* had an RA of 3.30% and was represented by 89 OTUs and 83,511 sequences. Among the 6 OTUs containing more than 100 sequences, the most representative ones (New.ReferenceOTU392) containing 181,098 sequences were identified as *Aureobasidium pullulans*, having an ITS2 sequence identical to that of reference isolates. *Aureobasidium pullulans* is a ubiquitous yeast-like fungus that can

colonize almost all environmental niches. It has been reported as one of the most abundant fungal colonizers of the phyllosphere and carposphere of a number of different plant species and was one of the most abundant fungi in the olive phyllosphere (Cf. Chapter II). Although *A. pullulans* has been reported as a causal agent of russet of apple fruit (Heidenreich et al. 1997), as well as other overripe fruit (Morgan and Michailides 2004, Kim 2014), it is generally considered as a non-pathogen and has been widely exploited as a postharvest bio-control agent on several fruit species including apple (Ippolito et al. 2000, Schena et al. 2003, Mari et al. 2012).

Among the other detected genera, *Stilbella* had an RA of 1.30% and was represented by 34 OTUs. The most abundant OTU showed 99% similar to an uncultured fungus (JQ989307.1) but only 93% similarity to the closest accepted species (*Stilbella fimetaria*; FJ939394.2). The genus *Phaeosphaeria* had an RA of 1.10% and was represented by 78 OTUs. According to Qiime and BLAST analyses the most common sequence was identified as *Phaeosphaeria triglochicola*, which has been reported as a plant endophyte (David et al. 2015). In general, *Phaeosphaeria* species are known as pathogens of cereal crops (Gladieux et al. 2008). Lastly, other detected sequences were associated with the order Agaricales (RA of 1.30%) and to the class Dothideomycetes (RA of 6.50%) but their identification, even at the level of genus, was not possible due to the absence of a high consistent homology with accepted fungal species.

Conclusion

Results of the present study provided increased insight into the fungal communities associated with apple fruit. Some of the detected taxa were already known as apple pathogens and/or endophytes. Many other taxa, however, were detected on apples for the first time or represented completely unknown taxa. The presence of unidentified fungi, even at the phylum level, indicates that the fungal diversity of apple fruit, as well as in other plant species and organs, is far from being completely characterized. The comprehensive picture of the fungal diversity on apple fruit provided in the current study may serve as the foundation for future experiments focusing on a specific group or groups of fungi, detected here for the first time. Moreover, the study has provided information about the effect of management practices, as well as the effect of hot water treatments on the composition of fungal communities present on different parts of apple fruit. The significant difference in the fungal populations of observed in CE and SE stem apples,

regardless of other investigated factors (“practices”, “location” and “time” effect) is a novel finding. Results of the study also suggest that the observed differences were also related to the way fungal communities reacted to treatments over a period of time. The results suggest the existence of a very complex equilibrium in the fungal populations present on apple fruit and a similar phenomenon is likely to occur in other fruit crops as well. Unfortunately, the limited genetic variation within the analyzed barcode gene (ITS2 region) did not enable the precise identification of species within the two most abundant genera (*Davidiella* and *Didymella*) hampering speculation about their importance and role. In this context further investigations of the fungal microbiota of apple with other more variable barcode genes are worthwhile.

References

- Aveskamp, M., d. J. Gruyter, and P. Crous. 2008. Biology and recent developments in the systematics of *Phoma*, a complex genus of major quarantine significance. *Fungal Diversity* **31**:1-18.
- Aveskamp, M. M., J. De Gruyter, J. Woudenberg, G. Verkley, and P. W. Crous. 2010. Highlights of the Didymellaceae: a polyphasic approach to characterise *Phoma* and related pleosporalean genera. *Studies in Mycology* **65**:1-60.
- Aveskamp, M. M., G. J. Verkley, J. de Gruyter, M. A. Murace, A. Perello, J. H. Woudenberg, J. Z. Groenewald, and P. W. Crous. 2009. DNA phylogeny reveals polyphyly of *Phoma* section *Peyronellaea* and multiple taxonomic novelties. *Mycologia* **101**:363-382.
- Behr, M., S. Legay, and D. Evers. 2013. Molecular identification of botrytis cinerea, *Penicillium* spp. And *Cladosporium* spp. In Luxembourg. Journal international des sciences de la vigne et du vin= International journal of vine and wine sciences **47**:239-248.
- Bengtsson-Palme, J., M. Ryberg, M. Hartmann, S. Branco, Z. Wang, A. Godhe, P. De Wit, M. Sánchez-García, I. Ebersberger, F. de Sousa, A. Amend, A. Jumpponen, M. Unterseher, E. Kristiansson, K. Abarenkov, Y. J. K. Bertrand, K. Sanli, K. M. Eriksson, U. Vik, V. Veldre, and R. H. Nilsson. 2013. Improved software detection and extraction of ITS1 and ITS2 from ribosomal ITS sequences of fungi and other eukaryotes for analysis of environmental sequencing data. *Methods in Ecology and Evolution* **4**:914-919.
- Bensch, K., J. Z. Groenewald, J. Dijksterhuis, M. Starink-Willemse, B. Andersen, B. A. Summerell, H. D. Shin, F. M. Dugan, H. J. Schroers, U. Braun, and P. W. Crous. 2010. Species and ecological diversity within the *Cladosporium cladosporioides* complex (Davidiellaceae, Capnodiales). *Studies in Mycology* **67**:1-94.
- Boerema, G., M. M. Dorenbosch, and H. Van Kesteren. 1977. Remarks on species of *Phoma* referred to *Peyronellaea*, V. *Kew Bulletin*:533-544.

- Bogomolova, E., and I. Kirtsideli. 2009. Airborne fungi in four stations of the St. Petersburg Underground railway system. *International Biodeterioration & Biodegradation* **63**:156-160.
- Braun, U., P. W. Crous, F. Dugan, J. E. Groenewald, and G. S. De Hoog. 2003. Phylogeny and taxonomy of *Cladosporium*-like hyphomycetes, including *Davidiella* gen. nov., the teleomorph of *Cladosporium* s. str. *Mycological Progress* **2**:3-18.
- Bray, J. R., and J. T. Curtis. 1957. An ordination of the upland forest communities of southern Wisconsin. *Ecological Monographs* **27**:325-349.
- David, A. S., E. W. Seabloom, and G. May. 2015. Plant Host Species and Geographic Distance Affect the Structure of Aboveground Fungal Symbiont Communities, and Environmental Filtering Affects Belowground Communities in a Coastal Dune Ecosystem. *Microbial ecology*:1-15.
- Droby, S., M. Wisniewski, D. Macarasin, and C. Wilson. 2009. Twenty years of postharvest biocontrol research: is it time for a new paradigm? *Postharvest Biology and Technology* **52**:137-145.
- Edgar, R. C. 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Research* **32**:1792-1797.
- Egan, T. 1991. Apple growers bruised and bitter after Alar scare. *New York Times* **9**:A1.
- FAOSTAT. 2015. Food and Agriculture Organisation. United Nations. <http://faostat.fao.org>.
- Gladieux, P., X.-G. Zhang, D. Afoufa-Bastien, R. V. Sanhueza, M. Sbaghi, and B. Le Cam. 2008. On the origin and spread of the scab disease of apple: out of central Asia. *PloS one* **3**:e1455.
- Hashem, M., S. A. Alamri, A. E.-L. Hesham, F. M. Al-Qahtani, and M. Kilany. 2014. Biocontrol of apple blue mould by new yeast strains: *Cryptococcus albidus* KKUY0017 and *Wickerhamomyces anomalus* KKUY0051 and their mode of action. *Biocontrol Science and Technology* **24**:1137-1152.
- Heald, F., and G. D. Ruehle. 1931. The rots of Washington Apples in cold storage. *Washington Agric. Exper. Stat.*
- Heidenreich, M. M., M. Corral-Garcia, E. Momol, and T. Burr. 1997. Russet of apple fruit caused by *Aureobasidium pullulans* and *Rhodotorula glutinis*. *Plant disease* **81**:337-342.
- Hoveland, C. S. 1993. Importance and economic significance of the *Acremonium* endophytes to performance of animals and grass plant. *Agriculture, Ecosystems & Environment* **44**:3-12.
- Ippolito, A., A. El Ghaouth, C. L. Wilson, and M. Wisniewski. 2000. Control of postharvest decay of apple fruit by *Aureobasidium pullulans* and induction of defense responses. *Postharvest Biology and Technology* **19**:265-272.
- Jäschke, D., D. Dugassa-Gobena, P. Karlovsky, S. Vidal, and J. Ludwig-Müller. 2010. Suppression of clubroot (*Plasmodiophora brassicae*) development in *Arabidopsis thaliana* by the endophytic fungus *Acremonium alternatum*. *Plant pathology* **59**:100-111.
- Johnson, R., L. Johnson, Y. Itoh, M. Kodama, H. Otani, and K. Kohmoto. 2000. Cloning and characterization of a cyclic peptide synthetase gene from *Alternaria alternata*

- apple pathotype whose product is involved in AM-toxin synthesis and pathogenicity. *Molecular Plant-Microbe Interactions* **13**:742-753.
- Kim, Y. 2014. First report of a new postharvest rot in sweet cherries caused by *Aureobasidium pullulans*. *Plant disease* **98**:424-424.
- Kruskal, W. H., and W. A. Wallis. 1952. Use of ranks in one-criterion variance analysis. *Journal of the American statistical Association* **47**:583-621.
- Li, B., C. Wang, X. Dong, Z. Zhang, and C. Wang. 2014. Acremonium Brown Spot, a New Disease Caused by *Acremonium sclerotigenum* on Bagged Apple Fruit in China. *Plant disease* **98**:1012-1012.
- Ligorio, A., G. Platania, L. Schena, V. Castiglione, I. Pentimone, F. Nigro, I. d. Silvestro, A. Ippolito, and P. Berolini. 2007. Control of table grape storage rots by combined applications of antagonistic yeasts, salts and natural substances. Pages 124-128 in *Novel approaches for the control of postharvest diseases and disorders. Proceedings of the International Congress, Bologna, Italy, 3-5 May, 2007*. CRIOF, University of Bologna.
- Lindahl, B. D., R. H. Nilsson, L. Tedersoo, K. Abarenkov, T. Carlsen, R. Kjølner, U. Kõljalg, T. Pennanen, S. Rosendahl, and J. Stenlid. 2013. Fungal community analysis by high-throughput sequencing of amplified markers—a user's guide. *New Phytologist* **199**:288-299.
- Malathrakis, N. 1985. The fungus *Acremonium alternatum* Linc: Fr., a hyperparasite of the cucurbits powdery mildew pathogen *Sphaerotheca fuliginea*. *Zeitschrift für Pflanzenkrankheiten und Pflanzenschutz* **92**:509-515.
- Mari, M., C. Martini, A. Spadoni, W. Rouissi, and P. Bertolini. 2012. Biocontrol of apple postharvest decay by *Aureobasidium pullulans*. *Postharvest Biology and Technology* **73**:56-62.
- Meng, X.-H., G.-Z. Qin, and S.-P. Tian. 2010. Influences of preharvest spraying *Cryptococcus laurentii* combined with postharvest chitosan coating on postharvest diseases and quality of table grapes in storage. *LWT-Food Science and Technology* **43**:596-601.
- Moore, W. C. 1959. British parasitic fungi. CUP Archive.
- Morgan, D., and T. Michailides. 2004. First report of melting decay of 'Red Globe' grapes in California. *Plant disease* **88**:1047-1047.
- Ntasiou, P., C. Myresiotis, S. Konstantinou, E. Papadopoulou-Mourkidou, and G. S. Karaoglanidis. 2015. Identification, characterization and mycotoxigenic ability of *Alternaria* spp. causing core rot of apple fruit in Greece. *International Journal of Food Microbiology* **197**:22-29.
- Ostenfeld Larsen, T., N. B. Perry, and B. Andersen. 2003. Infectopyrone, a potential mycotoxin from *Alternaria infectoria*. *Tetrahedron Letters* **44**:4511-4513.
- Pancher, M., M. Ceol, P. E. Corneo, C. M. O. Longa, S. Yousaf, I. Pertot, and A. Campisano. 2012. Fungal endophytic communities in grapevines (*Vitis vinifera* L.) respond to crop management. *Applied and Environmental Microbiology*:AEM. 07655-07611.
- Paulson, J. N., O. C. Stine, H. C. Bravo, and M. Pop. 2013. Differential abundance analysis for microbial marker-gene surveys. *Nature methods* **10**:1200-1202.

- Perdomo, H., D. Sutton, D. García, A. Fothergill, J. Cano, J. Gené, R. Summerbell, M. Rinaldi, and J. Guarro. 2011. Spectrum of clinically relevant *Acremonium* species in the United States. *Journal of Clinical Microbiology* **49**:243-256.
- Perelló, A., and M. Sisterna. 2008. Formation of *Lewia infectoria*, the teleomorph of *Alternaria infectoria*, on wheat in Argentina. *Australasian Plant Pathology* **37**:589-591.
- Pinto, C., D. Pinho, S. Sousa, M. Pinheiro, C. Egas, and A. C. Gomes. 2014. Unravelling the diversity of grapevine microbiome. *PloS one* **9**:e85622.
- Prigigallo, M., S. Mosca, S. Cacciola, D. Cooke, and L. Schena. 2015. Molecular analysis of *Phytophthora* diversity in nursery-grown ornamental and fruit plants. *Plant Pathology*.
- Raps, A., and S. Vidal. 1998. Indirect effects of an unspecialized endophytic fungus on specialized plant–herbivorous insect interactions. *Oecologia* **114**:541-547.
- Rognes, T., F. Mahé, and T. Flouri. 2015. vsearch
- Sanzani, S. M., L. Schena, V. De Cicco, and A. Ippolito. 2012. Early detection of *Botrytis cinerea* latent infections as a tool to improve postharvest quality of table grapes. *Postharvest Biology and Technology* **68**:64-71.
- Schena, L., F. Nigro, I. Pentimone, A. Ligorio, and A. Ippolito. 2003. Control of postharvest rots of sweet cherries and table grapes with endophytic isolates of *Aureobasidium pullulans*. *Postharvest Biology and Technology* **30**:209-220.
- Schoch, C. L., R. A. Shoemaker, K. A. Seifert, S. Hambleton, J. W. Spatafora, and P. W. Crous. 2006. A multigene phylogeny of the Dothideomycetes using four nuclear loci. *Mycologia* **98**:1041-1052.
- Serdani, M., P. Crous, G. Holz, and O. Petrini. 1998. Endophytic fungi associated with core rot of apples in South Africa, with specific reference to *Alternaria* species.
- Sharma, I., and S. Bhardwaj. 1999. Canker and foliar diseases of apple. *Diseases of Horticultural CropsC-Fruits*”(eds. Verma, LR and Sharma, RC) Indus Publishing Co., New Delhi, India:15-53.
- Slattery, E., M. Livingston, C. Greene, and K. Klonsky. 2011. Characteristics of conventional and organic apple production in the United States. US Department of Agriculture, Economic Research Service.
- Spadaro, D., and M. L. Gullino. 2004. State of the art and future prospects of the biological control of postharvest fruit diseases. *International Journal of Food Microbiology* **91**:185-194.
- Sui, Y., M. Wisniewski, S. Droby, and J. Liu. 2015. Responses of Yeast Biocontrol Agents to Environmental Stress. *Applied and Environmental Microbiology* **81**:2968-2975.
- Sylla, J., B. Alsanius, E. Krüger, A. Reineke, M. Bischoff-Schaefer, and W. Wohanka. 2013a. Introduction of *Aureobasidium pullulans* to the Phyllosphere of Organically Grown Strawberries with Focus on Its Establishment and Interactions with the Resident Microbiome. *Agronomy* **3**:704.
- Sylla, J., B. W. Alsanius, E. Krüger, A. Reineke, S. Strohmeier, and W. Wohanka. 2013b. Leaf microbiota of strawberries as affected by biological control agents. *Phytopathology* **103**:1001-1011.

- Takashima, M., T. Sugita, T. Shinoda, and T. Nakase. 2003. Three new combinations from the *Cryptococcus laurentii* complex: *Cryptococcus aureus*, *Cryptococcus carnescens* and *Cryptococcus peneaus*. *International Journal of Systematic and Evolutionary Microbiology* **53**:1187-1194.
- Tamura, K., G. Stecher, D. Peterson, A. Filipinski, and S. Kumar. 2013. MEGA6: Molecular Evolutionary Genetics Analysis version 6.0. *Molecular Biology and Evolution* **30**:2725-2729.
- Toju, H., A. S. Tanabe, S. Yamamoto, and H. Sato. 2012. High-coverage ITS primers for the DNA-based identification of ascomycetes and basidiomycetes in environmental samples. *PloS one* **7**:e40863.
- Vázquez-Baeza, Y., M. Pirrung, A. Gonzalez, and R. Knight. 2013. EMPERor: a tool for visualizing high-throughput microbial community data. *Structure* **585**:20.
- Wilson, C. L., and M. E. Wisniewski. 1989. Biological Control of Postharvest Diseases of Fruits and Vegetables: An Emerging Technology*. *Annual Review of Phytopathology* **27**:425-441.
- Wisniewski, M. E., and C. L. Wilson. 1992. Biological control of postharvest diseases of fruits and vegetables: recent advances. *American Society for Horticultural Science* **27**:94-98.
- Woudenberg, J., J. Groenewald, M. Binder, and P. Crous. 2013. *Alternaria* redefined. *Studies in Mycology* **75**:171-212.
- Xu, J., P. W. Qin, X. D. Xu, L. Hu, L. J. Wang, Y. Jiang, and Y. Y. Cao. 2015. First Report of Melon Collapse Caused by *Acremonium cucurbitacearum* in China. *Plant Disease* **99**:1184.
- Zhang, J., K. Kobert, T. Flouri, and A. Stamatakis. 2014. PEAR: a fast and accurate Illumina Paired-End reAd mergeR. *Bioinformatics* **30**:614-620.
- Zhang, Y., C. Schoch, J. Fournier, P. Crous, J. De Gruyter, J. Woudenberg, K. Hirayama, K. Tanaka, S. Pointing, and J. Spatafora. 2009. Multi-locus phylogeny of Pleosporales: a taxonomic, ecological and evolutionary re-evaluation. *Studies in Mycology* **64**:85-102.