Metabarcoding: A powerful tool to investigate microbial communities and shape future plant protection strategies

Metabarcoding: A powerful tool to investigate microbial communities and shape future plant protection strategies

Ahmed Abdelfattah[a], Antonino Malacrinò[b], Michael Wisniewski[c], Santa O. Cacciola[d], Leonardo Schena[a,*]

[a] Dipartimento di Agraria, Università Mediterranea di Reggio Calabria, Località Feo di Vito, 89122 Reggio Calabria, Italy

[b] Department of Physics, Chemistry and Biology (IFM), Linköping University, Linköping, Sweden [c] U.S. Department of Agriculture-Agricultural Research Service (USDA-ARS), 2217 Wiltshire Road, Kearneysville, WV 25430, United States

[d] Dipartimento di Agricoltura, Alimentazione e Ambiente, Università degli Studi di Catania, Via S. Sofia 100, 95123 Catania, Italy

∗ Corresponding author. E-mail address: lschena@unirc.it

Keywords: High throughput sequencing, Marker gene, Biological control, Bioinformatics, Plant pathology

A B S T R A C T

Microorganisms are the main drivers shaping the functioning and equilibrium of all ecosystems, contributing to nutrient cycling, primary production, litter decomposition, and multi-trophic interactions. Knowledge about the microbial assemblies in specific ecological niches is integral to understanding the assemblages interact and function the function, and becomes essential when the microbiota intersects with human activities, such as protecting crops against pests and diseases. Metabarcoding has proven to be a valuable tool and has been widely used for characterizing the microbial diversity of different environments and has been utilized in many research endeavors. Here we summarize the current status of metabarcoding technologies, the advantages and challenges in utilizing this technique, and how this pioneer approach is being applied to studying plant diseases and pests, with a focus on plant protection and biological control. Current and future developments in this technology will foster a more comprehensive understanding of microbial ecology, and the development of new, innovative pest control strategies.

1. Introduction

Microorganisms are ubiquitous and inhabit every living and nonliving habitat that has been studied, including a wide range of extreme physical environments. A total of approximately 10^{30} prokaryotic cell types have been estimated to inhabit the Earth, and they play an essential role in the vast majority of the biological processes occur in the divers ecosystems that exist, including the human body (Whitman et al., 1998). One gram of soil, or a milliliter of either salt or fresh water has been estimated to contain 10^9 or 10^6 prokaryotic cells, respectively (Sogin et al., 2006; Turnbaugh et al., 2007; Whitman et al., 1998). Microbes are the main drivers of several fundamental physical, chemical and biological processes and the study of their ecology, has been rapidly expanding (Prosser, 2015).

Taxonomic identification and functional studies of microorganisms in an environment, are crucial to understanding how different microbiota impact each other, or in other words understanding how different ecosystems function (Wisniewski et al., 2016b). In the past, taxonomic identification of microbes has relied on their isolation and culture on different general and selective types of growth media, in concert with the characterization of their morphological appearance, and their physiological and nutritional requirements. It has been estimated that up to 70% of the organisms in the human gut identified by DNA sequencing can be cultured, and that similar statistics can be obtained for soil and water samples (Browne et al., 2016). Though it should be noted that, in relation to the gut microbiome, capturing the number of species detected by DNA sequencing would require more than eight million colonies to be cultured. Presumably, similar statistics exist for identifying microbial taxa in soil, plant, and water samples. Thus, traditional culture-dependent approaches generally capture only a very limited portion of the existing microbial diversity naturally present in a sample, unless extremely extensive and diverse culturing approaches are used that provide a variety of nutrient conditions and do not mask the presence of slow-growing microbes (Sanzani et al., 2014). Molecular techniques, such as Restriction Fragment Length Polymorphism (RFLP), Amplified Fragment Length Polymorphism (AFLP), Short Tandem Repeat (STR), or Simple Sequence Repeats (SSRs), have been used as alternatives to conventional culturing methods to characterize microbial diversity (Kirk et al., 2004). These techniques, however, possess a major drawback, as they do not allow one to obtain precise taxonomical information from complex environmental samples.

The development of “DNA barcoding” represents an important advance in the molecular identification of microorganisms. This approach relies on the sequencing of one or more short fragments of DNA from standardized regions of the genome to identify different species (Hebert et al., 2003). These barcode genes can be defined as any fragment of DNA that: (i) contains significant species-level genetic variability and divergence, and (ii) possesses conserved flanking sites that can be used to design universal PCR primers, thus providing wide taxonomic application. Additionally, the primers used to amplify the standardized region of the genome must produce amplicons short length in order for them to be compatible with current, High-Throughput DNA Sequencing (HTS) technologies (Kress and Erickson, 2008).

DNA barcoding and the evolutionary information inferred by specific marker genes, have allowed the investigation of previously hidden microbial worlds (Moreira and López-Garcıa, 2002; Pace, 1997). The ever-increasing number of metagenomic studies, fostered by metabarcoding technology and supporting analytical software, have characterized microbial communities from complex environmental samples, and provided new details about their response to environmental factors. HTS, along with high-throughput proteomic and metabolomic technologies, represent a major advancement in our ability to investigate whole microbial communities and discover the role of specific genes, proteins, and metabolites in the interactions that occur within the microbial community and the hosts they occupy (Segata et al., 2013). High-throughput technologies have evolved rapidly in the last few years and revolutionized the way we study microbial communities. Previous analyses, based on Sanger sequencing, were also possible, but limited due to the requirement of cloning the PCR fragments obtained from environmental samples, and the sequencing of only a representative number of clones, due to the expenses and time required to process samples (Malacrinò et al., 2015; Mosca et al., 2014; Prigigallo et al., 2015; Scibetta et al., 2012; Torsvik et al., 2002).

The ability to characterize microbial communities from any kind of sample and matrix has had a great impact on the study and use of plantassociated microorganisms (endophytic, epiphytic, and soil-borne). The present review provides a brief summary of metabarcoding technology and an overview of how it can be used to address questions related to crop protection and biological control.

2. Metabarcoding – An outline

Among 'meta-omic' technologies, the term metagenomics refers to the sequencing and analysis of collective microbial genomes contained in an environmental sample through shotgun sequencing (Riesenfeld et al., 2004). In Contrast, the terms amplicon metagenomics and metabarcoding refer to the amplification and high-throughput sequencing of a specific barcode region (e.g. 16S for bacteria, 18S for most eukaryotes, and ITS regions for fungi) (Dollive et al., 2012; Segata et al., 2013; Taberlet et al., 2012). The term metabarcoding was introduced by Taberlet et al. (2012), and defined as an "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). Metabarcoding analyses enable the study of virtually all the microbial taxa in a given environment, which is defined as the microbiota. Although the term microbiota has been used interchangeably with the term microbiome, distinctions in the use of the term do exist. Microbiome refers to the study of the entirety of the microbial genetic material recovered directly from the environment, also known as shotgun metagenomics, while microbiota refers to the taxonomic composition of the microbial community as determined by metabarcoding analysis (Ursell et al., 2012). While the former term (microbiome) provides information about composition and function of the microbial community, the latter more simply allows one to answer the question: "who is there?".

The current HTS devices utilized for metabarcoding sequencing enable the generation of millions of short reads, which allow for a comprehensive reconstruction of the microbial community in a sampled environment. The utilization of metabarcoding to characterize the composition of a microbial community involves five main steps: i) sampling; ii) DNA extraction; iii) targeted amplification; iv) sequencing; and v) data analysis (Fig. 1).

Sampling is a crucial step in this process, since an adequate sampling design is essential to achieve an accurate representation of the microbial community. Similarly, DNA extraction and PCR amplifications must be conducted with particular care since they may greatly affect the quality and diversity of the sequences that are obtained (Thomas et al., 2012). Since microbes are ubiquitously present, care must be taken to avoid sample contamination. Careless sampling as well as the contamination from solvents, buffers, and reagents used during DNA extraction, purification, and amplification procedures, may lead to misleading results and/or diversity inflation (Salter et al., 2014).

Sequencing can be performed on different platforms. HTS techniques were revolutionized by the availability of the benchtop sequencer Roche 454 GS-FLX/+, which has been widely used in the past decade but is now being phased out (D'Amore et al., 2016). Illumina® and Thermo Fisher Scientific® now produce sequencers that are the most commonly used in metabarcode sequencing of microbial communities. They include the MiSeq™, NextSeq™ and HiSeq™ from Illumina® (San Diego, California) and IonTorrent™ and IonProton™ from Thermo Fisher Scientific® (Waltham, Massachusetts). The recent ability to obtain longer DNA reads (> 600 bp) has received increasing interest, since longer reads enable a more reliable taxonomic inference. Devices such as the PacBio RSII™ (Pacific Biosciences®, Menlo Park, California), MinION from Oxford Nanopore Technologies® (Oxford, UK), and new library preparation methods e.g. 10x Genomics, allow extremely long reads to be obtained that are especially useful for reconstructing the genome of an organism, especially when combined with shotgun sequencing. The number of reads produced by these various platforms can vary from 700,000 (Roche 454) to 300 million (Illumina HiSeq), which impacts the number of samples that can be multiplexed together during a single sequencing run. As the number of reads obtained is relatively fixed, the number of samples sequenced in a single sequencing run is typically based on the depth of sequencing required to adequately capture the microbial diversity present in a sample vs. the number of samples required to adequately represent a specific environment. Several reports have described the specific parameters associated with various metabarcoding protocols (Bálint et al., 2014; Cristescu, 2014; D'Amore et al., 2016; Ficetola et al., 2016).

Downstream data analysis of the obtained sequence data requires various software programs, adequate computing capacity, and the periods of time needed to process large datasets. The analysis of the DNA data relies on a variety of bioinformatics tools, of which QIIME, MOTHUR, and OBITOOLS are among the most commonly used (Boyer et al., 2016; Caporaso et al., 2010; Schloss et al., 2009). The downstream analysis can be generally divided into two main parts:

1. Generation of an Operational Taxonomic Units (OTUs) table: this step involves processing of sequence data including, filtering out lowquality sequencing reads, by similarity, and assigning a taxonomy a taxonomic classification to each bin of similar sequences (OTUs).

2. Statistical analyses once the 'OTU table' is generated, several statistical analyses can be performed on the dataset.

The generation of the OTU table is a complex process that can be divided into six main steps: i) demultiplexing; ii) quality filtering; iii) OTUs creation; iv) selection of representative sequences; v) taxonomic assignment; and vi) OTUs table building. The demultiplexing step assigns reads to each biological sample. Each sample is tagged during the PCR by adding a single or a double barcode, which allows the samples to be pooled together but still remain identifiable i.e. multiplexing. After sequencing, the barcodes are used by the bioinformatic software to assign reads to specific samples (Hamady et al., 2008). The purpose of quality filtering is to discard reads with low quality base calls, chimeras, and short reads that could lead to erroneous results. This step includes a denoising procedure when 454 pyrosequencing is used, which aims to reduce the impact of sequencing errors due to long homopolymers. Once the sequences are filtered, they are clustered using a wide variety of algorithms to create OTUs (reviewed in Chen et al. (2013). Most algorithms rely on the selection of sequences from the whole dataset to generate clusters that contain similar sequences according to a pre-selected threshold of similarity (commonly 97%). After generating these clusters, singletons (clusters containing only 1 sequence) are usually discarded to avoid erroneous results due to sequencing errors, and a representative set of sequences is generated. This set includes one representative sequence for each cluster, which speeds up the taxonomic assignment. The representative OTU sequence is then queried against a set of established databases, such as Greengenes for 16S, SILVA for 18S, and UNITE for ITS2 (DeSantis et al., 2006; Kõljalg et al., 2013; Quast et al., 2013). Alternatively, custom databases can be generated and used. This procedure results in the generation of an OTU table which contains the number of reads for each OTU within each sample, and allows a wide range of statistical comparisons to be made between samples. Since not all individual samples will have the same number of reads, variations in library size must be normalized before conducting further downstream analyses (Weiss et al., 2017). The representative set of sequences within each OTU can also be aligned and phylogenetically analyzed.

Statistical analyses can be carried out using the same software packages associated with R statistical software which possess a very diverse set of sub-programs, including 'vegan' or 'Bioconductor' (Dixon, 2003; Huber et al., 2015; Team, 2013). The analysis of the OTU data commonly includes the characterization of alpha and beta diversity, and statistical comparisons between samples. Alpha diversity represents the diversity within each sample, whereas beta diversity represents the diversity between different samples by comparing samples with a wide set of multivariate statistical procedures (e.g. ANOSIM, PERMANOVA) and visualization tools e.g. Principal Component Analysis (PCA), Principal Coordinates Analysis (PCoA), Non-metric

multidimensional scaling (NMDS) (Thomas et al., 2012; Yu et al., 2012). Further insights can be achieved by computing the network of OTU occurrence, which graphically illustrates the relationship among microorganisms in the analyzed samples and visually highlights interactions.

3. Challenges and technical issues related to the use ofmetabarcoding in plant pathology applications

Meta-omics technologies and related bioinformatic tools have been developed and utilized to characterize the microbial diversity in many different environments. Most studies have focused on the ecology of microbial populations and on the analysis of complex host-microbeenvironment interactions. In these studies, microbial populations were generally analyzed as part of an ensemble using a higher taxonomic level (genus and above) and individual species were assumed to be of secondary importance. In contrast, the precise identification of specific taxa is extremely important in plant pathology, since related species, with very similar barcode genes, can have a completely different effect on plant health. Therefore, major issues which need to be addressed in the future are: i) identification of marker genes that will capture the greatest level of diversity; ii) creation of specific and reliable reference databases; and iii) specific taxonomic assignments (sequences of typespecies).

In addition, unlike other molecular techniques such as qPCR, amplicon sequencing can only provide relative abundance information i.e. compositional data, and therefore this approach cannot be used for the absolute quantification of microbial communities. This indeed, represents a major challenge towards understanding the true significance of a given taxa in the environment.

3.1. Identification of appropriate marker genes

The use of a short genetic marker to identify taxa (DNA barcoding) is a simple but powerful idea that gave an enormous boost to research in a variety of biological fields. The first DNA barcodes were proposed for animal identification, and based on the mitochondrial gene, cytochrome c oxidase subunit I (COX1) (Hebert et al., 2003). This was followed by the development of DNA barcodes for plants (Hollingsworth et al., 2009). The development of a universal, barcodebased method to identify microorganisms, such as bacteria and fungi, was a major advance since their identification by other means is often problematic. Thus, their identification has become increasingly dependent on the DNA sequences of standardized genetic markers. As previously stated, the bacterial 16S and the fungal ITS genes are the most widely used target barcodes for inferring phylogenetic relationships among microorganisms (Lane et al., 1985; Schoch et al., 2012). However, studying the full length of either barcodes is not yet feasible using the current HTS techniques. This has led several studies to focus on the selection of sub-regions such as hypervariable regions from V1V9 in bacteria and ITS1 or ITS2 for fungi. The consensus of these studies suggest that V4-V6 is the optimal sub-regions for the design of universal primers with superior phylogenetic resolution for bacterial phyla (Yang et al., 2016). In fungi instead, the

alternative use of ITS1 or ITS2 region in metabarcoding studies seems to give comparable results (Blaalid et al., 2013).

Insufficient genetic variability within the ITS regions, however, has been acknowledged. The limited variability present in this barcode gene can make the discrimination of closely related taxa very difficult or impossible and therefore limits its utility (Schena and Cooke, 2006). This shortcoming has been problematic when ITS sequences have been used in phylogenetic studies aimed at redefining the classification of fungi (Stielow et al., 2015). Nevertheless, some studies have suggested the use of supplementary barcodes in concert with ITS as the primary barcode (Stielow et al., 2015). The goal of defining a unique barcode gene for all fungal species is not realistic, and that one or more general barcode genes may need to be associated with specific targets, and properly developed and validated for specific fungal taxa (e.g. fungal genera or classes). In this context, the search for alternative barcode genes is of great interest (Chakraborty et al., 2014). New DNA barcodes need to have simple amplification conditions, low variability in the length of the obtained sequence, presence of a single copy for each cell, and low intra-species variability, in order to be suited for HTS and provide an adequate level of resolution (at species or sub-species level). In a recent study, the translation elongation factor 1-α (TEF1α) was proposed as a universal fungal DNA barcode with superior resolution compared to ITS, while other gene selections, such as the topoisomerase I (TOPI), phosphoglycerate kinase (PGK), and a hypothetical protein (LNS2) were also found to represent promising targets for Ascomycota and/or Pucciniomycotina (Stielow et al., 2015). Similarly, the RAS-related YPT1 gene has been proposed to be an effective barcode gene for Phytophthora species (Abad et al., 2016; Schena and Cooke, 2006).

Regardless of the target gene, the selection of appropriate primers for amplification is essential in metabarcoding analyses. Indeed, the selection of primers is one of the most important factors determining the DNA sequences obtained, and several studies have been conducted using both in silico and experimental approaches to adequately design appropriate primers for use in metabarcoding analyses (Bokulich and Mills, 2013; Stielow et al., 2015; Toju et al., 2012). Selected primers need to be generic to amplify target DNA from most, if not all taxa, including those still unknown to the scientific community (e.g. all fungi or bacteria), but also need to be specific enough to avoid the amplification of DNA from non-target organisms (e.g. plant DNA). Furthermore, primers need to amplify fragments whose length is compatible with current HTS platforms, and guarantee high levels of PCR efficiency (Op De Beeck et al., 2014). Another important aspect in the selection of appropriate primers is the objective of the analysis to be conducted. Broad, community-level studies require more “universality” in primer selection, while studies on specific taxa, such as plant pathogens or biocontrol agents, may require more specificity since the microorganisms of interest may represent a small minority of the microbial diversity present. In this regard, the use of more selective primers can greatly facilitate metabarcoding studies, as has been demonstrated for Ascomycota, Basidiomycota, Chytridiomycota, Zygomycota, and Oomycota (Nikolcheva and Bärlocher, 2004; Vettraino et al., 2012), as well as studies of specific genera, including

Colletotrichum spp. (Mosca et al., 2014), Phytophthora spp. (Prigigallo et al., 2015), and Trichoderma spp. (Friedl and Druzhinina, 2012).

Software programs such as ecoPrimers (Riaz et al., 2011), ecoPCR (Ficetola et al., 2010), and PrimerProspector (Walters et al., 2011) can greatly assist in the design of barcode primers. These software programs possess specific features to test the capability of primers to amplify target organisms and the ability to discriminate between taxa. To help standardize this process, and to ensure repeatability and consistency among studies, standards for reporting information about a new marker gene have been developed. For example, the Minimum information about a MARKer Gene Sequence (MIMARKS), has been developed by the Genomic Standards Consortium (Yilmaz et al., 2011).

3.2. Creation of specific and reliable reference databases

One of the objectives of metabarcoding is to characterize microbial communities by providing a taxonomic assignment to each OTU. Therefore, it is essential to have a comprehensive taxonomic reference database for each barcode gene, and preferably, one that has been generated using reliably identified voucher specimens or cultures (Cristescu, 2014). One example of the global efforts to establish such a database is the creation of unified DNA barcode databases for the identification of fungi and bacteria, such as UNITE, Greengenes, Silva, and RDP. Standardized analytical procedures and lab protocols are now used and required by many large-scale metagenomic projects investigating 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). The UNITE database (http://unite.ut.ee) is an ITS sequence database for the identification of fungi that includes only high-quality sequences of well identified fungi. Other databases covering the bacterial 16S rRNA are: Silva (Pruesse et al., 2007); Greengenes (DeSantis et al., 2006); and the International Nucleotide Sequence Database (INSD) (http://insdc.org). The latter database, however, has a greater chance of misidentification due to the lower accuracy and reliability of the deposited sequences (Nilsson et al., 2006). Validated databases, such as UNITE, have gained increasing popularity in the last few years for studying the fungal diversity and ecology of environmental samples. A major challenge in the near future will be the development of a greater number of highly reliable databases generated through the results obtained from individual and community projects. The comprehensiveness and accuracy of these databases will be essential to support analyses that utilize barcode genes with greater variability and selected for a specific purpose or for a specific group of target organisms.

3.3. Taxonomy assignment

Current standard analyses using bioinformatic software enable an accurate identification of taxa to the genus level. The same algorithms, however, may fail to reach a higher level of discrimination, required for the identification of fungal species. This drawback is primarily due to the insufficient

level of genetic variation present within currently utilized target regions, which is required in order to be able to identify closely-related species. The similarity threshold (commonly 97%) applied to cluster sequences is also generally insufficient to differentiate closely-related taxa (Nilsson et al., 2008). Increasing the similarity threshold to 99% can improve the identification of specific taxa, but may cause other problems without improving taxonomic resolution (Català et al., 2015). In fact, the intragenomic variation in multicopy genetic regions such as ITS, can lead to ambiguous clustering and classification. Furthermore, sequencing errors and the unreliable annotations of sequences in public DNA databases remain a serious obstacle to sequence-based species identification and these errors may be magnified using a very high similarity threshold (Nilsson et al., 2006). Moreover, a significant number of the deposited ITS sequences are not updated and therefore may not reflect recent updates in fungal taxonomy (Stielow et al., 2015).

Attempts to overcome the challenges related to the identification of microbial species include the extraction of sequences associated with specific taxa in order to conduct classical phylogenetic analyses using validated panels of reference sequences for each group of related taxa (Abdelfattah et al., 2015; Prigigallo et al., 2016). This approach, has enabled the exploitation of all the available genetic variations within the ITS region and, thus, the identification of taxa with the highest possible level of accuracy. This approach, however, is very time-consuming and requires the existence of well-described and validated reference sequences, which in many cases are not available.

Another possible approach is the identification of alternative barcodes with higher genetic variability. However, this is not an easy task, since it requires the creation of comprehensive and reliable databases for each new barcode gene. Furthermore, appropriate similarity thresholds will need to be defined for each barcode gene in relation to the microorganisms under investigation and in relation to the specific objective of the analysis.

4. Metabarcoding and integrated control strategies

4.1. Analysis of microbial diversity

Meta-omic analyses have been increasingly utilized in recent years to investigate microbial diversity in several different agro-environments (Table 1). In most cases, these studies have revealed a level of microbial biodiversity that was entirely missed by conventional culture-based methods. Although most of these studies were not strictly related to the development and/or optimization of control methods, they have increased our knowledge of microbial communities present on agronomic plant species and, in some cases, have had important practical implications (Massart et al., 2015). For example, the investigation of the microbiota associated with olives in southern Italy enabled the identification of several Colletotrichum species including C. godetiae, C. boninense, C. acutatum s.s, C. gleosporioides, and C. karstii (Abdelfattah et al., 2015; Schena et al., 2017). Since these species are characterized by different levels of virulence, the information about their presence and population level are valuable to the development of management practices. The same study revealed an abundant presence of Pseudocercospora cladosporioides, the causal agent of olive cercosporiosis, and a low abundance of Spilocaea oleagina, commonly considered as the main causal

agent of olive defoliation. These data suggested that the relative importance of these pathogens should be reconsidered (Abdelfattah et al., 2015). Several other studies have investigated the microbial diversity associated with the aerial parts of cultivated plants, such as tomatoes, grapes, strawberries, citrus, apple, and cereals (Abdelfattah et al., 2015; Abdelfattah et al., 2016a,b,c; Ottesen et al., 2013; Pinto et al., 2014). In all these studies, a high level of genetic diversity was documented, and many new organisms were reported for the first time. For example, a study on the spatial distribution of fungi on apple fruit detected more than 8000 OTUs, among which were fungal genera never reported to be associated with harvested apples, including Malassezia, Candida, and Trichosporon, and entire phyla, such as Glomeromycota and Chytridiomycota (Abdelfattah et al., 2016b). The OTUs detected in these studies included the presence of both beneficial and phytopathogenic microorganisms, with important practical implications on the management of plant health (Pinto et al., 2014). The genus Devriesia was found to be one of the most abundant fungal colonizers of the canopy of olive trees, although this genus has never been reported as an olive fungal colonizer, and its role in the olive phyllosphere remains unknown (Abdelfattah et al., 2015). Similar unexpected discoveries have been reported in studies on the phyllosphere and carposphere of other host plants (Table 1). In these studies, plant species always had a primary role in defining the composition of the microbial populations. It is however recommended to complement metabarcoding with other methods to understand the role of these newly discovered microorganism in the environment.

Rhizosphere microbiota also play a fundamental role in plant health (Bulgarelli et al., 2013). Plants have been shown to influence the microorganisms associated with their roots, while specific microbes or assemblages of microbes also influence the immunity response to plant pathogenic organisms (Berendsen et al., 2012). A recent study on poplar trees, reported that rhizosphere microbiota even influence the composition of the endophytic communities present within plant hosts (Beckers et al., 2016) Although sufficient evidence is currently unavailable, there is a general concurrence that microbial communities associated with the rhizosphere are more diverse and complex than those occupying the phyllosphere (Sapp et al., 2016). This is partially based on the fact that plant roots are in direct contact with one of the most diverse environments found in nature, the soil. Soil by its nature, has a strong effect on microbial community composition. For example, while species of dark-septate, endophytic fungi were found to be more diverse in acidic soils, and species of ectomycorrhizal fungi were more highly represented by OTUs in calcareous sites, the opposite was observed with arbuscular mycorrhizal genera (Geml et al., 2014).

Metabarcoding can also be used to investigate how plants and their microbiota respond to different environmental conditions. Bálint et al. (2015) analyzed the fungal diversity associated with different genotypes of balsam poplar, and demonstrated that the host-microbiome interaction strongly depends on the plant genotype and that this also affects plant response to environmental changes, such as global warming. Similarly, regional differences in climatic conditions, and varietal differences were reported to have an influence on the microbiota associated with grapes (Bokulich et al., 2014; Taylor et al., 2014). These data indicate that the selection and exchange of microbes

from different regions and or varieties should be considered when developing biological control strategies and that the inherent microbiota needs to be considered when examining plant response to the environment (Pinto et al., 2014).

4.2. Impact of the microbiota on disease management strategies

The application of chemical pesticides and/or biological control agents is a common practice in the production of food and ornamental crops. Our current understanding of the impact of these management practices on native microorganisms, however, is poor. Although a few studies have been conducted, it is evident that metabarcoding protocols represent an opportunity for conducting the in-depth investigations needed to provide the foundation for a more holistic use of control strategies, based on the potential of increasing the population of indigenous microbial antagonists. In a recent study, a comparison between organic and conventionally grown apple fruit, revealed that management practices had a significant impact on the composition of fruit-associated fungal communities (Abdelfattah et al., 2016b). Several unique taxa were detected exclusively on organic apples, suggesting that management practices may have been a contributing factor in determining the taxa present. In another study on strawberry, an interruption in the application of conventional pesticides resulted in a rapid modification of the structure of the fungal community of leaves and flowers (Abdelfattah et al., 2016c).

Some metabarcoding studies have focused on the effect of the exogenous application of biocontrol agents on native microbiota communities. Sylla et al. (2013) reported that the application of a biocontrol strain of Aureobasidium pullulans, had both short-term and longterm effects on the composition and diversity of fungal communities as a direct result of the successful establishment of the biocontrol agent. Different results, however, were obtained in two consecutive years, indicating the need for additional and more detailed studies. In another study, the application of Trichoderma harzianum on strawberries to control Botrytis cinerea was demonstrated to displace indigenous phyllosphere fungi, even though the biocontrol agent had no effect on the target pathogen (Sylla et al., 2013).

The composition of the soil microbiota, as well as the population of specific species, can be influenced by other agricultural practices not strictly related to crop protection. A study of the microbiome of soils subjected to different agricultural management practices for two decades, revealed that systems not receiving manure harbored a dispersed and functionally versatile community, characterized by oligotrophic organisms adapted to nutrient-limited environments (Hartmann et al., 2015). On the other hand, systems receiving organic fertilizer were characterized by specific microbial guilds known to be involved in the degradation of complex organic compounds, such as manure and compost.

4.3. Disease aetiology

Since metabarcoding provides a comprehensive picture of the genetic diversity present in a sample, it can be a valuable tool for identifying putative causal agents responsible for complex diseases or disease syndromes. In this regard, the emerging concept of a “pathobiome”, which takes into consideration the pathogen and its interactions within the whole microbial community, is receiving greater discussion (Vayssier-Taussat et al., 2014). This concept has been used to study the microbial

community associated with Acute Oak Decline Syndrome and enabled the detection of a set of putative causal agents (Sapp et al., 2016). Another study focused on the analysis of citrus leaves exhibiting typical symptoms of greasy spot, a disease commonly associated with Mycosphaerella citri and other related fungal species (Abdelfattah et al., 2016a). Surprisingly, none of the species previously reported as causal agents of the disease were detected, which led to the conclusion that observed symptoms associated with greasy spot in citrus in southern Italy may be caused by a different fungal species. Another study on the microbial community associated with strawberry plants demonstrated that the observed reductions in yield were not caused by Verticillium dahliae (Xu et al., 2015). A similar analysis was utilized to analyze the presence of phytopathogenic Enterobacteriaceae and to examine their biology within the context of the wider microbial community (Vayssier-Taussat et al., 2014).

4.4. Detection of plant pathogens

PCR-based molecular detection methods have greatly contributed to the accurate, sensitive, and reliable detection and quantification of many different species of plant pathogens (Sanzani et al., 2014; Schena et al., 2013). The majority of currently available diagnostic assays, however, are specifically designed to detect only one or just a few species. Therefore, their use is inappropriate for broader surveys of plant pathogens in natural and agricultural ecosystems, in which a method capable of detecting multiple species or even undescribed species is required. Such broad detection methods can play an important role in monitoring the introduction and spread of “foreign” or “alien” species of microbes, which has become an ever greater global problem due to international trade and travel (Brasier, 2008). Several recently described invasive Phytophthora species, were previously unknown to science and were only identified when they caused severe diseases in non-native environments (Kroon et al., 2012). A strain of Xylella fastidiosa, currently threatening olive orchards in southern Italy, was also unknown to science until a few years ago (Martelli et al., 2016). Having co-evolved with their natural hosts, many of these pathogens do not cause noticeable damage in their native ecosystems, and so are less likely to be detected. In this regard, the detection of all putative plant pathogens (including those currently unknown to the scientific community) remains a challenge and metabarcoding approaches may represent a powerful culture-independent technique to monitor and prevent new epidemic outbreaks caused by exotic pathogens (Comtet et al., 2015).

Metabarcoding analyses with genus-specific primers (Scibetta et al., 2012) revealed many new host–pathogen associations, the presence of microbial species previously unreported in the environment under investigation, and new phylotypes that remain to be taxonomically defined (Català et al., 2017, 2015). More specifically, the primary role of plant nurseries in fostering the introduction and dissemination of Phytophthora species has been confirmed (Prigigallo et al., 2016, 2015). Metabarcoding has also been proposed as a tool to track the spread of economically important diseases of cashew plants (Monteiro et al., 2015).

4.5. Biological control

Although most recent metabarcoding studies were not strictly focused on biological control strategies, it is clear that metabarcoding technologies represent a powerful way to identify the microflora associated with plants, and can be used to better understand complex plantmicrobiota interactions, including those involving pathogenic and beneficial species (Massart et al., 2015; Wisniewski et al., 2016a). Such studies may play a pivotal role in developing "biological" approaches for managing plant pathogens that exploit the complete microbial diversity associated with plants.

A major objective of metabarcoding studies focused on biocontrol is the identification of novel biocontrol communities with high adaptability or synergisms within the biotic community (Mazzola and Freilich, 2016). Friedl and Druzhinina (2012) used genus-specific primers to analyze the presence of Trichoderma, a genus comprising several species widely used for biological control (Howell, 2003). The information generated may be valuable in formulating new Trichodermabased products for protecting crops against soil-borne plant pathogens. Studies on the presence and identification of mycoviruses has also attracted a great deal of attention due to their potential use as biocontrol agents (Marzano et al., 2016; Xie and Jiang, 2014).

Meta-omics can also be useful in characterizing the microflora of disease-suppressive soils, which are able to inhibit the development of plant-pathogenic microorganisms (Kinkel et al., 2011; Mazzola, 2004). The suppression of plant disease is achieved by the active competition of the general microbial community with pathogens, the ability of some species of microbes to induce systemic resistance in plants, and by the production of antibiotics (van Elsas et al., 2008). The ability of the soil microflora to suppress specific diseases has been the subject of many studies whose objectives have been to identify the mechanism by which pathogenic microflora are suppressed and host immunity-responses are induced, and determining how plants are able to recruit beneficial microorganisms as a response to stress (Berendsen et al., 2012). Metaomics approaches have been used to reveal how the microbial community of suppressive soils is structured (metabarcoding), and how it functions (metatranscriptomics). Sun et al. (2015), used metabarcoding to characterize the microbiome of disease-suppressive soils in relation to the occurrence and severity of potato common scab. They found that soil properties have a significant influence on the spatial variation of Streptomyces population densities, the proportion of inhibitory isolates, and the pathogen suppressive capacity between different locations in a field planting and between field plantings.

A variety of soil amendments are being evaluated for use in integrated management strategies designed to manage soil-borne diseases. The mechanism underlying the increased suppression in amended soils is frequently poorly understood (Meghvansi and Varma, 2015). De Tender et al. (2016) used a metabarcoding approach to study how charcoal used as a soil amendment helps strawberry plants to resist Botrytis cinerea infections and found an increased below-ground recruitment of beneficial microbes. Metabarcoding was used to document a similar mechanism underlying the use of Brassica seed meal soil amendments to suppress apple replant disease (Mazzola et al., 2014), the use of biofertilizers to suppress Fusarium wilt disease in banana (Shen et al., 2015), the use of almond shells to promote the suppression of Rosellinia necatrix (Vida et al.,

2016), and use of several other composts (Mehta et al., 2014). Lastly, the analysis of the metagenome from environmental samples has been proposed as a useful method for the screening of microbial communities and genes useful for weed management (Kao-Kniffin et al., 2013).

4.6. Insect microbiota

The characterization of insect microbial communities could provide information that would be very useful in designing microbiota-based insect pest management strategies. In particular, knowledge of the microorganisms associated with specific niches or with particular conditions (e.g. health status of insects) may facilitate the selection of effective biocontrol agents. For example, a metabarcoding analysis of the red palm weevil (Rhynchophorus ferrugineus) was used to gain information on the microflora associated with this pest (Liu et al., 2013). Further studies could focus on the practical use of these microbes in controlling the weevil. In another study, metatranscriptomics was used to identify a viral pathogen of the invasive Caribbean Crazy Ant, Nylanderia pubens (Valles et al., 2012). Interestingly, metabarcoding has also been used to characterize the antagonist complex associated with specific insect pests (Pompanon et al., 2012; Rougerie et al., 2011; Symondson, 2002).

Insects also have an important role as vectors in the spread of plant pathogens, including viruses, viroids, phytoplasmas, bacteria, and fungi. The relationship between insects and fungi has been widely studied for bark and ambrosia beetles, which are known vectors of fungal agents of plant diseases (Hulcr and Dunn, 2011; Miller et al., 2016). Since these beetles can share fungal associates through horizontal transfer, and their spread through wood trades worldwide is widely acknowledged, their potential role as 'in-out' travelling populations can result in the spread of fungi and other microbes by their acquisition of new fungal species from one area and their spread to a new location (Malacrinò et al., 2017; Wingfield et al., 2016). The characterization of the fungal microbiota of the olive fruit fly (Bactrocera oleae) revealed a very diverse community in which sooty mold fungi and plant pathogens were highly represented (Malacrinò et al., 2016, 2015). Although the nature of the interaction between the olive fly and the highly-represented fungi remains to be determined, a comparative analysis of the insect fungal microbiota and the olive phylloplane in the same geographic location indicated the existence of specific interactions between the insect and some abundant fungal taxa, including Cladosporium spp. (Abdelfattah et al., 2015; Malacrinò et al., 2016).

5. Conclusions and future perspectives

The availability and accessibility of HTS technologies, together with the development of metabarcoding protocols and bioinformatics software, as well as other meta-omics technologies, has clearly provided a powerful new way to characterize microbial communities and to address specific questions related to microbial ecology and crop protection. Future studies will benefit greatly from the ability to conduct extensive and reliable taxonomic identifications. Combining metabarcoding data with metatranscriptomic data will greatly increase our knowledge about the

structure and functioning of microbial communities in diverse environments. Bioinformatic strategies, based on curated comprehensive and validated databases, could be used to predict the role of each detected taxa, giving further insights into the ecology of these microorganisms (Nguyen et al., 2016).

Recent studies have provided data documenting that a significant portion of the microbial biodiversity that exists in nature is still unknown or has an unclear ecological role. Researchers are still only just beginning to exploit these new diagnostic techniques, especially in their application to the study of biocontrol systems. Future research efforts to standardize techniques and procedures, together with the development of improved and accurate databases, will play a fundamental role in increasing our knowledge of the microbial ecology of naturally-occurring microbial communities in diverse environments and in shaping the future, biologically-based pest management strategies.

Acknowledgments

This research was funded by The Italian Ministry of Education, University and Research (MIUR) with "PON Ricerca e competitività 2007–2013": i) Innovazione di prodotto e di processo nelle filiera dei prodotti da forno e dolciari (PON03PE_00090_01); ii) "Modelli sostenibili e nuove tecnologie per la valorizzazione delle olive e dell'olio extravergine di oliva prodotto in Calabria (PON03PE_00090_02); and iii) Modelli sostenibili e nuove tecnologie per la valorizzazione delle filiere vegetali mediterranee. The authors wish to thank Mrs. Ann Davies for the revision of the English style.

References

Abad, Z.G., Bienapfl, J., Knight, L., Jennings, K., Galvez, M., Schena, L., 2016. Developing an innovative molecular toolbox for identification of Phytophthora species with emphasis on species of regulatory concern., In: 9th APS Annual Meeting, Tampa, Florida, USA.

Abdelfattah, A., Nicosia, M.G.L.D., Cacciola, S.O., Droby, S., Schena, L., 2015. Metabarcoding analysis of fungal diversity in the phyllosphere and carposphere of olive (Olea europaea). PLoS One 10, e0131069.

Abdelfattah, A., Cacciola, S.O., Mosca, S., Zappia, R., Schena, L., 2016a. Analysis of the fungal diversity in citrus leaves with greasy spot disease symptoms. Microb. Ecol. 1–11.

Abdelfattah, A., Wisniewski, M., Droby, S., Schena, L., 2016b. Spatial and compositional variation in the fungal communities of organic and conventionally grown apple fruit at the consumer point-of-purchase. Horticult. Res. 3, 16042.

Abdelfattah, A., Wisniewski, M., Li Destri Nicosia, M.G., Cacciola, S.O., Schena, L., 2016c. Metagenomic analysis of fungal diversity on strawberry plants and the effect of management practices on the fungal community structure of aerial organs. PLoS One 11, e0160470.

Bálint, M., Bartha, L., O'Hara, R.B., Olson, M.S., Otte, J., Pfenninger, M., Robertson, A.L., Tiffin, P., Schmitt, I., 2015. Relocation, high-latitude warming and host genetic identity shape the foliar fungal microbiome of poplars. Mol. Ecol. 24, 235–248.

Bálint, M., Schmidt, P.A., Sharma, R., Thines, M., Schmitt, I., 2014. An Illumina metabarcoding pipeline for fungi. Ecol. Evol. 4, 2642–2653.

Bálint, M., et al., 2013. Host genotype shapes the foliar fungal microbiome of balsam poplar (Populus balsamifera). PLoS ONE 8 (1), e53987.

Beckers, B., Op De Beeck, M., Thijs, S., Truyens, S., Weyens, N., Boerjan, W., Vangronsveld, J., 2016. Performance of 16s rDNA primer pairs in the study of rhizosphere and endosphere bacterial microbiomes in metabarcoding studies. Front. Microbiol. 7, 650.

Berendsen, R.L., Pieterse, C.M.J., Bakker, P.A.H.M., 2012. The rhizosphere microbiome and plant health. Trends Plant Sci. 17, 478–486.

Blaalid, R., Kumar, S., Nilsson, R.H., Abarenkov, K., Kirk, P.M., Kauserud, H., 2013. ITS1 versus ITS2 as DNA metabarcodes for fungi. Mol. Ecol. Resour. 13, 218–224.

Bokulich, N.A., Mills, D.A., 2013. Improved selection of internal transcribed spacer-specific primers enables quantitative, ultra-high-throughput profiling of fungal communities. Appl. Environ. Microbiol. 79, 2519–2526.

Bokulich, N.A., Thorngate, J.H., Richardson, P.M., Mills, D.A., 2014. Microbial biogeography of wine grapes is conditioned by cultivar, vintage, and climate. Proc. Natl.

Acad. Sci. 111, E139–E148.

Boyer, F., Mercier, C., Bonin, A., Le Bras, Y., Taberlet, P., Coissac, E., 2016. Obitools: a unix-inspired software package for DNA metabarcoding. Mol. Ecol. Resour. 16, 176–182.

Brasier, C.M., 2008. The biosecurity threat to the UK and global environment from international trade in plants. Plant. Pathol. 57, 792–808.

Browne, H.P., Forster, S.C., Anonye, B.O., Kumar, N., Neville, B.A., Stares, M.D., Goulding, D., Lawley, T.D., 2016. Culturing of 'unculturable' human microbiota reveals novel taxa and extensive sporulation. Nature 533, 543–546.

Bulgarelli, D., Schlaeppi, K., Spaepen, S., van Themaat, E.V.L., Schulze-Lefert, P., 2013.

Structure and functions of the bacterial microbiota of plants. Annu. Rev. Plant Biol. 64, 807–838.

Caporaso, J.G., Kuczynski, J., Stombaugh, J., Bittinger, K., Bushman, F.D., Costello, E.K., Fierer, N., Pena, A.G., Goodrich, J.K., Gordon, J.I., 2010. QIIME allows analysis of high-throughput community sequencing data. Nat. Methods 7, 335–336.

Català, S., Berbegal, M., Pérez-Sierra, A., Abad-Campos, P., 2017. Metabarcoding and development of new real-time specific assays reveal Phytophthora species diversity in holm oak forests in eastern Spain. Plant. Pathol. 66, 115–123.

Català, S., Pérez-Sierra, A., Abad-Campos, P., 2015. The use of genus-specific Amplicon pyrosequencing to assess phytophthora species diversity using eDNA from soil and water in Northern Spain. PLoS One 10, e0119311.

Chakraborty, C., Doss, C.G.P., Patra, B.C., Bandyopadhyay, S., 2014. DNA barcoding to map the microbial communities: current advances and future directions. Appl.

Microbiol. Biotechnol. 98, 3425–3436.

Chen, W., Zhang, C.K., Cheng, Y., Zhang, S., Zhao, H., 2013. A comparison of methods for clustering 16S rRNA sequences into OTUs. PLoS One 8, e70837.

Comtet, T., Sandionigi, A., Viard, F., Casiraghi, M., 2015. DNA (meta)barcoding of biological invasions: a powerful tool to elucidate invasion processes and help managing aliens. Biol. Invasions 17, 905–922.

Cordier, T., et al., 2012. The composition of phyllosphere fungal assemblages of European beech (Fagus sylvatica) varies significantly along an elevation gradient. New Phytol.

196 (2), 510–519.

Cristescu, M.E., 2014. From barcoding single individuals to metabarcoding biological communities: towards an integrative approach to the study of global biodiversity.
Trends Ecol. Evol. 29, 566–571.
D'Amore, R., Ijaz, U.Z., Schirmer, M., Kenny, J.G., Gregory, R., Darby, A.C., Shakya, M., Podar, M., Quince, C., Hall, N., 2016. A comprehensive benchmarking study of protocols and sequencing platforms for 16S rRNA community profiling. BMC Genom. 17, 1–20.
De Tender, C., Haegeman, A., Vandecasteele, B., Clement, L., Cremelie, P., Dawyndt, P., Maes, M., Debode, J., 2016. Dynamics in the strawberry Rhizosphere microbiome in response to biochar and botrytis cinerea leaf infection. Front. Microbiol. 7.
DeSantis, T.Z., Hugenholtz, P., Larsen, N., Rojas, M., Brodie, E.L., Keller, K., Huber, T., Dalevi, D., Hu, P., Andersen, G.L., 2006. Greengenes, a Chimera-Checked 16S rRNA gene database and workbench compatible with ARB. Appl. Environ. Microbiol. 72, 5069–5072.
Dixon, P., 2003. VEGAN, a package of R functions for community ecology. J. Veg. Sci. 14, 927–930.
Dollive, S., Peterfreund, G.L., Sherrill-Mix, S., Bittinger, K., Sinha, R., Hoffmann, C.,
Nabel, C.S., Hill, D.A., Artis, D., Bachman, M.A., Custers-Allen, R., Grunberg, S., Wu, G.D., Lewis, J.D., Bushman, F.D., 2012. A tool kit for quantifying eukaryotic rRNA gene sequences from human microbiome samples. Genome Biol. 13, 1–13.
Ficetola, G.F., Coissac, E., Zundel, S., Riaz, T., Shehzad, W., Bessière, J., Taberlet, P., Pompanon, F., 2010. An in silico approach for the evaluation of DNA barcodes. BMC Genom. 11, 1.
Ficetola, G.F., Taberlet, P., Coissac, E., 2016. How to limit false positives in environmental DNA and metabarcoding? Mol. Ecol. Resour. 16, 604–607.
Foulon, J., et al., 2016. Environmental metabarcoding reveals contrasting microbial communities at two poplar phytomanagement sites. Sci. Total Environ. 571, 1230–1240.
Friedl, M.A., Druzhinina, I.S., 2012. Taxon-specific metagenomics of Trichoderma reveals a narrow community of opportunistic species that regulate each other's development. Microbiology 158, 69–83.
Geml, J., Gravendeel, B., van der Gaag, K.J., Neilen, M., Lammers, Y., Raes, N., Semenova, T.A., de Knijff, P., Noordeloos, M.E., 2014. The contribution of DNA metabarcoding to fungal conservation: diversity assessment, habitat partitioning and mapping redlisted fungi in protected coastal salix repens communities in the Netherlands. PLoS One 9, e99852.
Hamady, M., Walker, J.J., Harris, J.K., Gold, N.J., Knight, R., 2008. Error-correcting barcoded primers for pyrosequencing hundreds of samples in multiplex. Nat Methods 5, 235–237.
Hartmann, M., Frey, B., Mayer, J., Mäder, P., Widmer, F., 2015. Distinct soil microbial diversity under long-term organic and conventional farming. ISME J. 9, 1177–1194.
Hebert, P.D.N., Cywinska, A., Ball, S.L., deWaard, J.R., 2003. Biological identifications through DNA barcodes. Proc. R. Soc. Lond. B Biol. Sci. 270, 313–321.
Hollingsworth, P.M., Forrest, L.L., Spouge, J.L., Hajibabaei, M., Ratnasingham, S., van der Bank, M., Chase, M.W., Cowan, R.S., Erickson, D.L., Fazekas, A.J., 2009. A DNA barcode for land plants. Proc. Natl. Acad. Sci. 106, 12794–12797.
Howell, C.R., 2003. Mechanisms employed by trichoderma species in the biological control of plant diseases: the history and evolution of current concepts. Plant Dis. 87, 4–10.
Huber, W., Carey, V.J., Gentleman, R., Anders, S., Carlson, M., Carvalho, B.S., Bravo,
H.C., Davis, S., Gatto, L., Girke, T., Gottardo, R., Hahne, F., Hansen, K.D., Irizarry,

R.A., Lawrence, M., Love, M.I., MacDonald, J., Obenchain, V., Oles, A.K., Pages, H., Reyes, A., Shannon, P., Smyth, G.K., Tenenbaum, D., Waldron, L., Morgan, M., 2015. Orchestrating high-throughput genomic analysis with bioconductor. Nat. Methods 12, 115–121.

Hulcr, J., Dunn, R.R., 2011. The sudden emergence of pathogenicity in insect–fungus symbioses threatens naive forest ecosystems. Proc. R. Soc. B: Biol. Sci. 278, 2866–2873.

Izuno, A., et al., 2016. Structure of phyllosphere fungal communities in a tropical dipterocarp plantation: A massively parallel next-generation sequencing analysis.

Mycoscience 57 (3), 171–180.

Junker, R.R., Keller, A., 2015. Microhabitat heterogeneity across leaves and flower organs promotes bacterial diversity. FEMS Microbiol. Ecol. 91 (9) fiv097–fiv097.

Kao-Kniffin, J., Carver, S.M., DiTommaso, A., 2013. Advancing weed management strategies using metagenomic techniques. Weed Sci. 61, 171–184.

Kecskeméti, E., et al., 2016. Are epiphytic microbial communities in the carposphere of ripening grape clusters (Vitis vinifera L.) different between conventional, organic, and biodynamic grapes? PLoS ONE 11 (8), e0160852.

Kinkel, L.L., Bakker, M.G., Schlatter, D.C., 2011. A coevolutionary framework for managing disease-suppressive soils. Annu. Rev. Phytopathol. 49, 47–67.

Kirk, J.L., Beaudette, L.A., Hart, M., Moutoglis, P., Klironomos, J.N., Lee, H., Trevors, J.T., 2004. Methods of studying soil microbial diversity. J. Microbiol. Methods 58, 169–188.

Kõljalg, U., Nilsson, R.H., Abarenkov, K., Tedersoo, L., Taylor, A.F., Bahram, M., Bates, S.T., Bruns, T.D., Bengtsson-Palme, J., Callaghan, T.M., 2013. Towards a unified paradigm for sequence-based identification of fungi. Mol. Ecol. 22, 5271–5277.

Kress, W.J., Erickson, D.L., 2008. DNA barcodes: genes, genomics, and bioinformatics.

Proc. Natl. Acad. Sci. 105, 2761–2762.

Kroon, L.P.N.M., Brouwer, H., Cock, A.W.A.M.d., Govers, F., 2012. The phytophthora Genus Anno 2012. Phytopathology 102, 348–364.

Lane, D.J., Pace, B., Olsen, G.J., Stahl, D.A., Sogin, M.L., Pace, N.R., 1985. Rapid determination of 16S ribosomal RNA sequences for phylogenetic analyses. Proc. Natl.

Acad. Sci. 82, 6955–6959.

Liu, J., Sui, Y., Wisniewski, M., Droby, S., Liu, Y., 2013. Review: utilization of antagonistic yeasts to manage postharvest fungal diseases of fruit. Int. J. Food Microbiol. 167, 153–160.

Malacrinò, A., Rassati, D., Schena, L., Mehzabin, R., Battisti, A., Palmeri, V., 2017. Fungal communities associated with bark and ambrosia beetles trapped at international harbours. Fungal Ecol. 28, 44–52.

Malacrinò, A., Schena, L., Campolo, O., Laudani, F., Mosca, S., Giunti, G., Strano, C.P., Palmeri, V., 2016. A metabarcoding survey on the fungal microbiota associated to the olive fruit fly. Microb. Ecol. 1–8.

Malacrinò, A., Schena, L., Campolo, O., Laudani, F., Palmeri, V., 2015. Molecular analysis of the fungal microbiome associated with the olive fruit fly Bactrocera oleae. Fungal Ecol. 18, 67–74.

Martelli, G.P., Boscia, D., Porcelli, F., Saponari, M., 2016. The olive quick decline syndrome in south-east Italy: a threatening phytosanitary emergency. Eur. J. Plant Pathol. 144, 235–243.

Marzano, S.-Y.L., Nelson, B.D., Ajayi-Oyetunde, O., Bradley, C.A., Hughes, T.J., Hartman, G.L., Eastburn, D.M., Domier, L.L., 2016. Identification of diverse mycoviruses through metatranscriptomics characterization of the viromes of five major fungal plant pathogens. J. Virol. 90, 6846–6863.

Massart, S., Martinez-Medina, M., Jijakli, M.H., 2015. Biological control in the microbiome era: challenges and opportunities. Biol. Control 89, 98–108.
Mazzola, M., 2004. Assessment and management of soil microbial community structure for disease suppression 1. Annu. Rev. Phytopathol. 42, 35–59.
Mazzola, M., Freilich, S., 2016. Prospects for biological soilborne disease control: application of indigenous versus synthetic microbiomes. Phytopathology 107, 256–263.
Mazzola, M., Hewavitharana, S.S., Strauss, S.L., 2014. Brassica seed meal soil amendments transform the rhizosphere microbiome and improve apple production through resistance to pathogen reinfestation. Phytopathology 105, 460–469.
Meghvansi, M.K., Varma, A., 2015. Organic Amendments and Soil Suppressiveness in Plant Disease Management. Springer.
Mehta, C.M., Palni, U., Franke-Whittle, I.H., Sharma, A.K., 2014. Compost: its role, mechanism and impact on reducing soil-borne plant diseases. Waste Manage. 34, 607–622.
Miao, C.-P., et al., 2016. Rhizospheric fungi of Panax notoginseng: diversity and antagonism to host phytopathogens. J. Ginseng Res. 40 (2), 127–134.
Miller, K.E., Hopkins, K., Inward, D.J.G., Vogler, A.P., 2016. Metabarcoding of fungal communities associated with bark beetles. Ecol. Evol. 6, 1590–1600.
Monteiro, F., Romeiras, M.M., Figueiredo, A., Sebastiana, M., Baldé, A., Catarino, L., Batista, D., 2015. Tracking cashew economically important diseases in the West African region using metagenomics. Front. Plant Sci. 6.
Moreira, D., López-Garcıa, P., 2002. The molecular ecology of microbial eukaryotes uń veils a hidden world. Trends Microbiol. 10, 31–38.
Mosca, S., Li Destri Nicosia, M.G., Cacciola, S.O., Schena, L., 2014. Molecular analysis of colletotrichum species in the carposphere and phyllosphere of olive. PLoS One 9, e114031.
Nguyen, N.H., Song, Z., Bates, S.T., Branco, S., Tedersoo, L., Menke, J., Schilling, J.S., Kennedy, P.G., 2016. FUNGuild: An open annotation tool for parsing fungal community datasets by ecological guild. Fungal Ecol. 20, 241–248.
Nikolcheva, L.G., Bärlocher, F., 2004. Taxon-specific fungal primers reveal unexpectedly high diversity during leaf decomposition in a stream. Mycological Progr. 3, 41–49.
Nilsson, R.H., Kristiansson, E., Ryberg, M., Hallenberg, N., Larsson, K.-H., 2008. Intraspecific ITS variability in the kingdom fungi as expressed in the international sequence databases and its implications for molecular species identification. Evol.
Bioinf. Online 4, 193–201.
Nilsson, R.H., Ryberg, M., Kristiansson, E., Abarenkov, K., Larsson, K.-H., Kõljalg, U., 2006. Taxonomic reliability of DNA sequences in public sequence databases: a fungal perspective. PLoS One 1, e59.
Op De Beeck, M., Lievens, B., Busschaert, P., Declerck, S., Vangronsveld, J., Colpaert, J.V., 2014. Comparison and validation of some its primer pairs useful for fungal metabarcoding studies. PLoS One 9, e97629.
Ottesen, A.R., Peña, A.G., White, J.R., Pettengill, J.B., Li, C., Allard, S., Rideout, S., Allard, M., Hill, T., Evans, P., 2013. Baseline survey of the anatomical microbial ecology of an important food plant: Solanum lycopersicum (tomato). BMC Microbiol. 13, 114.
Pace, N.R., 1997. A molecular view of microbial diversity and the biosphere. Science 276, 734–740.
Pinto, C., Pinho, D., Sousa, S., Pinheiro, M., Egas, C., Gomes, A.C., 2014. Unravelling the diversity of grapevine microbiome. PLoS One 9, e85622.

Pompanon, F., Deagle, B.E., Symondson, W.O.C., Brown, D.S., Jarman, S.N., Taberlet, P., 2012. Who is eating what: diet assessment using next generation sequencing. Mol.
Ecol. 21, 1931–1950.
Prigigallo, M.I., Abdelfattah, A., Cacciola, S.O., Faedda, R., Sanzani, S.M., Cooke, D.E.L., Schena, L., 2016. Metabarcoding Analysis of Phytophthora Diversity Using GenusSpecific Primers and 454 Pyrosequencing. Phytopathology 106, 305–313.
Prigigallo, M.I., Mosca, S., Cacciola, S.O., Cooke, D.E.L., Schena, L., 2015. Molecular analysis of Phytophthora diversity in nursery-grown ornamental and fruit plants.
Plant. Pathol. 64, 1308–1319.
Prosser, J.I., 2015. Dispersing misconceptions and identifying opportunities for the use of'omics' in soil microbial ecology. Nat. Rev. Microbiol. 13, 439–446.
Pruesse, E., Quast, C., Knittel, K., Fuchs, B.M., Ludwig, W., Peplies, J., Glöckner, F.O., 2007. SILVA: a comprehensive online resource for quality checked and aligned ribosomal RNA sequence data compatible with ARB. Nucleic Acids Res. 35, 7188–7196.
Quast, C., Pruesse, E., Yilmaz, P., Gerken, J., Schweer, T., Yarza, P., Peplies, J., Glöckner, F.O., 2013. The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. Nucleic Acids Res. 41, D590–D596.
Riaz, T., Shehzad, W., Viari, A., Pompanon, F., Taberlet, P., Coissac, E., 2011. EcoPrimers: inference of new DNA barcode markers from whole genome sequence analysis.
Nucleic Acids Res. 39 e145-e145.
Riesenfeld, C.S., Schloss, P.D., Handelsman, J., 2004. Metagenomics: genomic analysis of microbial communities. Annu. Rev. Genet. 38, 525–552.
Rougerie, R., Smith, M.A., Fernandez-Triana, J., Lopez-Vaamonde, C., Ratnasingham, S., Hebert, P.D.N., 2011. Molecular analysis of parasitoid linkages (MAPL): gut contents of adult parasitoid wasps reveal larval host. Mol. Ecol. 20, 179–186.
Salter, S.J., Cox, M.J., Turek, E.M., Calus, S.T., Cookson, W.O., Moffatt, M.F., Turner, P., Parkhill, J., Loman, N.J., Walker, A.W., 2014. Reagent and laboratory contamination can critically impact sequence-based microbiome analyses. BMC Biol. 12, 87.
Sanzani, S.M., Li Destri Nicosia, M.G., Faedda, R., Cacciola, S.O., Schena, L., 2014. Use of quantitative PCR detection methods to study biocontrol agents and phytopathogenic fungi and oomycetes in environmental samples. J. Phytopathol. 162, 1–13.
Sapp, M., Lewis, E., Moss, S., Barrett, B., Kirk, S., Elphinstone, J., Denman, S., 2016. Metabarcoding of bacteria associated with the Acute Oak decline syndrome in England. Forests 7, 95.
Schena, L., Abdelfattah, A., Mosca, S., Li Destri Nicosia, M.G., Agosteo, G.E., Cacciola, S.O., 2017. Quantitative detection of Colletotrichum godetiae and C. acutatum sensu stricto in the phyllosphere and carposphere of olive during four phenological phases. Eur. J. Plant Pathol.
Schena, L., Cooke, D.E., 2006. Assessing the potential of regions of the nuclear and mitochondrial genome to develop a “molecular tool box” for the detection and characterization of Phytophthora species. J. Microbiol. Methods 67, 70–85.
Schena, L., Li Destri Nicosia, M., Sanzani, S., Faedda, R., Ippolito, A., Cacciola, S., 2013. Development of quantitative PCR detection methods for phytopathogenic fungi and oomycetes. J. Plant Pathol. 7–24.
Schloss, P.D., Westcott, S.L., Ryabin, T., Hall, J.R., Hartmann, M., Hollister, E.B., Lesniewski, R.A., Oakley, B.B., Parks, D.H., Robinson, C.J., 2009. Introducing mothur: open-source, platform-

independent, community-supported software for describing and comparing microbial communities. Appl. Environ. Microbiol. 75, 7537–7541.

Schoch, C.L., Seifert, K.A., Huhndorf, S., Robert, V., Spouge, J.L., Levesque, C.A., Chen, W., Fungal Barcoding, C., Fungal Barcoding Consortium Author, L., 2012. Nuclear ribosomal internal transcribed spacer (ITS) region as a universal DNA barcode marker for Fungi. Proc. Natl. Acad. Sci. 109, 6241–6246.

Scibetta, S., Schena, L., Chimento, A., Cacciola, S.O., Cooke, D.E.L., 2012. A molecular method to assess Phytophthora diversity in environmental samples. J. Microbiol. Methods 88, 356–368.

Segata, N., Boernigen, D., Tickle, T.L., Morgan, X.C., Garrett, W.S., Huttenhower, C., 2013. Computational meta'omics for microbial community studies. Mol. Syst. Biol. 9.

Shen, Z., Ruan, Y., Wang, B., Zhong, S., Su, L., Li, R., Shen, Q., 2015. Effect of biofertilizer for suppressing Fusarium wilt disease of banana as well as enhancing microbial and chemical properties of soil under greenhouse trial. Appl. Soil Ecol. 93, 111–119.

Sogin, M.L., Morrison, H.G., Huber, J.A., Welch, D.M., Huse, S.M., Neal, P.R., Arrieta, J.M., Herndl, G.J., 2006. Microbial diversity in the deep sea and the underexplored “rare biosphere”. Proc. Natl. Acad. Sci. 103, 12115–12120.

Stielow, J.B., Lévesque, C.A., Seifert, K.A., Meyer, W., Iriny, L., Smits, D., Renfurm, R.,
Verkley, G.J.M., Groenewald, M., Chaduli, D., Lomascolo, A., Welti, S., Lesage-
Meessen, L., Favel, A., Al-Hatmi, A.M.S., Damm, U., Yilmaz, N., Houbraken, J.,
Lombard, L., Quaedvlieg, W., Binder, M., Vaas, L.A.I., Vu, D., Yurkov, A., Begerow,
D., Roehl, O., Guerreiro, M., Fonseca, A., Samerpitak, K., van Diepeningen, A.D.,
Dolatabadi, S., Moreno, L.F., Casaregola, S., Mallet, S., Jacques, N., Roscini, L., Egidi, E., Bizet, C., Garcia-Hermoso, D., Martín, M.P., Deng, S., Groenewald, J.Z., Boekhout, T., de Beer, Z.W., Barnes, I., Duong, T.A., Wingfield, M.J., de Hoog, G.S., Crous, P.W.,
Lewis, C.T., Hambleton, S., Moussa, T.A.A., Al-Zahrani, H.S., Almaghrabi, O.A.,
Louis-Seize, G., Assabgui, R., McCormick, W., Omer, G., Dukik, K., Cardinali, G., Eberhardt, U., de Vries, M., Robert, V., 2015. One fungus, which genes? Development and assessment of universal primers for potential secondary fungal DNA barcodes.
Persoonia: Molecular Phylogeny and Evolution of Fungi 35, 242–263.

Sun, P., Otto-Hanson, L.K., Arenz, B.E., Ma, Q., Kinkel, L.L., 2015. Molecular and functional characteristics of streptomycete communities in relation to soil factors and potato common scab. Eur. J. Soil Biol. 70, 58–66.

Sylla, J., Alsanius, B.W., Krüger, E., Reineke, A., Strohmeier, S., Wohanka, W., 2013. Leaf microbiota of strawberries as affected by biological control agents. Phytopathology 103, 1001–1011.

Symondson, W.O.C., 2002. Molecular identification of prey in predator diets. Mol. Ecol. 11, 627–641.

Taberlet, P., Coissac, E., Pompanon, F., Brochmann, C., Willerslev, E., 2012. Towards next-generation biodiversity assessment using DNA metabarcoding. Mol. Ecol. 21, 2045–2050.

Taylor, M.W., Tsai, P., Anfang, N., Ross, H.A., Goddard, M.R., 2014. Pyrosequencing reveals regional differences in fruit-associated fungal communities. Environ.
Microbiol. 16, 2848–2858.

Team, R.C., 2013. R: A language and environment for statistical computing.

Thomas, T., Gilbert, J., Meyer, F., 2012. Metagenomics - a guide from sampling to data analysis. Microb. Inf. Exp. 2, 3.

Toju, H., Tanabe, A.S., Yamamoto, S., Sato, H., 2012. High-coverage ITS primers for the DNA-based identification of ascomycetes and basidiomycetes in environmental samples. PLoS One 7, e40863.

Torsvik, V., Øvreås, L., Thingstad, T.F., 2002. Prokaryotic diversity-magnitude, dynamics, and controlling factors. Science 296, 1064–1066.

Turnbaugh, P.J., Ley, R.E., Hamady, M., Fraser-Liggett, C., Knight, R., Gordon, J.I., 2007. The human microbiome project: exploring the microbial part of ourselves in a changing world. Nature 449, 804.

Ursell, L.K., Metcalf, J.L., Parfrey, L.W., Knight, R., 2012. Defining the human microbiome. Nutr. Rev. 70, S38–S44.

Valles, S.M., Oi, D.H., Yu, F., Tan, X.-X., Buss, E.A., 2012. Metatranscriptomics and pyrosequencing facilitate discovery of potential viral natural enemies of the invasive caribbean crazy ant, Nylanderia pubens. PLoS One 7, e31828.

van Elsas, J.D., Costa, R., Jansson, J., Sjöling, S., Bailey, M., Nalin, R., Vogel, T.M., van Overbeek, L., 2008. The metagenomics of disease-suppressive soils – experiences from the METACONTROL project. Trends Biotechnol. 26, 591–601.

Vayssier-Taussat, M., Albina, E., Citti, C., Cosson, J.F., Jacques, M.-A., LEBRUN, M.-H., Le Loir, Y., Ogliastro, M., Petit, M.-A., Roumagnac, P., Candresse, T., 2014. Shifting the paradigm from pathogens to pathobiome: new concepts in the light of meta-omics.

Front. Cell. Infect. Microbiol. 4.

Vettraino, A.M., Bonants, P., Tomassini, A., Bruni, N., Vannini, A., 2012. Pyrosequencing as a tool for the detection of Phytophthora species: error rate and risk of false Molecular Operational Taxonomic Units. Lett. Appl. Microbiol. 55, 390–396.

Vida, C., Bonilla, N., de Vicente, A., Cazorla, F.M., 2016. Microbial profiling of a suppressiveness-induced agricultural soil amended with composted almond shells. Front. Microbiol. 7.

Walters, W.A., Caporaso, J.G., Lauber, C.L., Berg-Lyons, D., Fierer, N., Knight, R., 2011. PrimerProspector: de novo design and taxonomic analysis of barcoded polymerase chain reaction primers. Bioinformatics 27, 1159–1161.

Weiss, S., Xu, Z.Z., Peddada, S., Amir, A., Bittinger, K., Gonzalez, A., Lozupone, C., Zaneveld, J.R., Vázquez-Baeza, Y., Birmingham, A., Hyde, E.R., Knight, R., 2017. Normalization and microbial differential abundance strategies depend upon data characteristics. Microbiome 5, 27.

Whitman, W.B., Coleman, D.C., Wiebe, W.J., 1998. Prokaryotes: the unseen majority.

Proc. Natl. Acad. Sci. 95, 6578–6583.

Wingfield, M.J., Garnas, J.R., Hajek, A., Hurley, B.P., de Beer, Z.W., Taerum, S.J., 2016.

Novel and co-evolved associations between insects and microorganisms as drivers of forest pestilence. Biol. Invasions 18, 1045–1056.

Wisniewski, M., Droby, S., Norelli, J., Liu, J., Schena, L., 2016a. Alternative management technologies for postharvest disease control: the journey from simplicity to complexity. Postharvest Biol. Technol. 122, 3–10.

Wisniewski, M., Norelli, J., Droby, S., Ballester, A.R., Abdelfattah, A., Levin, E., 2016b. Genomic tools for developing markers for postharvest disease resistance in Rosaceae fruit crops. International Society for Horticultural Science (ISHS), Leuven, Belgium, pp. 7–16.

Xie, J., Jiang, D., 2014. New insights into mycoviruses and exploration for the biological control of crop fungal diseases. Annu. Rev. Phytopathol. 52, 45–68.

Xu, X., Passey, T., Wei, F., Saville, R., Harrison, R.J., 2015. Amplicon-based metagenomics identified candidate organisms in soils that caused yield decline in strawberry. Horticult. Res. 2, 15022.

Yang, B., Wang, Y., Qian, P.-Y., 2016. Sensitivity and correlation of hypervariable regions in 16S rRNA genes in phylogenetic analysis. BMC Bioinformatics 17, 135.

Yilmaz, P., Kottmann, R., Field, D., Knight, R., Cole, J.R., Amaral-Zettler, L., Gilbert, J.A., Karsch-Mizrachi, I., Johnston, A., Cochrane, G., Vaughan, R., Hunter, C., Park, J.,

Morrison, N., Rocca-Serra, P., Sterk, P., Arumugam, M., Bailey, M., Baumgartner, L., Birren, B.W., Blaser, M.J., Bonazzi, V., Booth, T., Bork, P., Bushman, F.D., Buttigieg, P.L., Chain, P.S., Charlson, E., Costello, E.K., Huot-Creasy, H., Dawyndt, P., DeSantis, T., Fierer, N., Fuhrman, J.A., Gallery, R.E., Gevers, D., Gibbs, R.A., San Gil, I.,

Gonzalez, A., Gordon, J.I., Guralnick, R., Hankeln, W., Highlander, S., Hugenholtz, P., Jansson, J., Kau, A.L., Kelley, S.T., Kennedy, J., Knights, D., Koren, O., Kuczynski, J., Kyrpides, N., Larsen, R., Lauber, C.L., Legg, T., Ley, R.E., Lozupone, C.A., Ludwig, W.,

Lyons, D., Maguire, E., Methe, B.A., Meyer, F., Muegge, B., Nakielny, S., Nelson, K.E., Nemergut, D., Neufeld, J.D., Newbold, L.K., Oliver, A.E., Pace, N.R., Palanisamy, G., Peplies, J., Petrosino, J., Proctor, L., Pruesse, E., Quast, C., Raes, J., Ratnasingham, S.,

Ravel, J., Relman, D.A., Assunta-Sansone, S., Schloss, P.D., Schriml, L., Sinha, R.,

Smith, M.I., Sodergren, E., Spo, A., Stombaugh, J., Tiedje, J.M., Ward, D.V.,

Weinstock, G.M., Wendel, D., White, O., Whiteley, A., Wilke, A., Wortman, J.R., Yatsunenko, T., Glockner, F.O., 2011. Minimum information about a marker gene sequence (MIMARKS) and minimum information about any (x) sequence (MIxS) specifications. Nat. Biotechnol. 29, 415–420.

Yu, D.W., Ji, Y., Emerson, B.C., Wang, X., Ye, C., Yang, C., Ding, Z., 2012. Biodiversity soup: metabarcoding of arthropods for rapid biodiversity assessment and biomonitoring. Methods Ecol. Evol. 3, 613–623.

Table 1. Examples of recent studies based on the use of metabarcoding approach to investigate the microbiota associated with different plant organs.[a]

Year	Target plant	Phyllosphere	Flower	Carposphere	Rhizosphere	Reference
2012	Fagus sylvatica	*				Cordier et al. (2012)
2013	Solanum lycopersicum	*	*	*	*	Ottesen et al. (2013)
2013	Populus balsamifera	*				Bálint et al. (2013)
2014	Vitis vinifera	*				Pinto et al. (2014)
2014	Vitis vinifera			*		Taylor et al. (2014)
2014	Salix repens				*	Geml et al. (2014)
2015	Olea europaea	*	*	*		Abdelfattah et al. (2015)
2015	Metrosideros polymorpha	*	*			Junker and Keller (2015)
2016	Fragaria x ananassa	*	*	*		Abdelfattah et al. (2016c)
2016	Malus domestica			*		Abdelfattah et al. (2016b)
2016	Citrus sinensis	*				Abdelfattah et al. (2016a)
2016	Shorea leprosula	*				Izuno et al. (2016)
2016	Quercus spp.				*	Sapp et al. (2016)
2016	Vitis vinifera			*		Kecskeméti et al. (2016)
2016	Panax notoginseng				*	Miao et al. (2016)
2016	Quercus ilex				*	Català et al. (2017)
2016	Populus spp.				*	Foulon et al. (2016)

[a] There are numerous published reports. An attempt was made to list representative references, and we apologize to investigators whose specific results could not be cited due to space limitations.

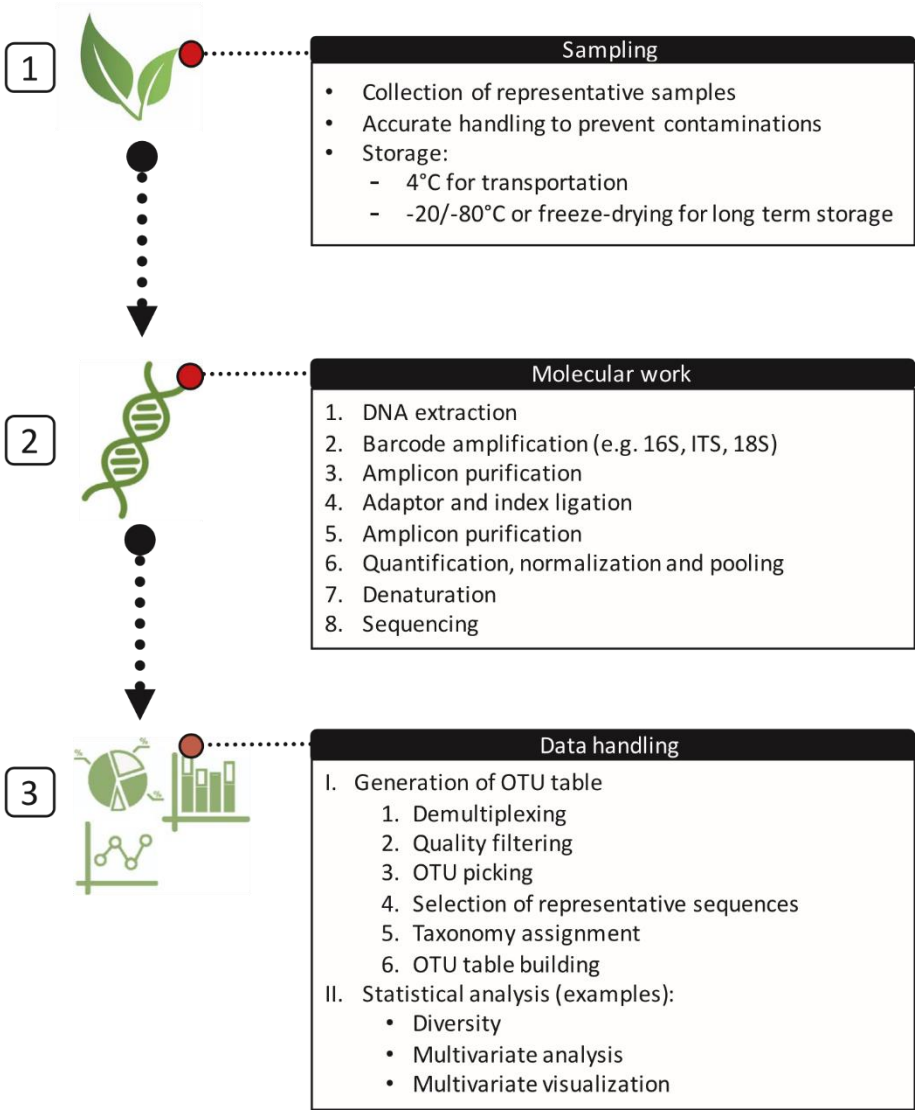

Fig. 1. An outline of the metabarcoding approach. The workflow groups these steps into main categories including sampling (specimen collection), molecular work (wet lab), and data handling (bioinformatics, and biostatistics steps).