*Biological Forum – An International Journal* **15(5a): 538-548(2023)**

**ISSN No. (Print): 0975-1130**
**ISSN No. (Online): 2249-3239**

# Advances in Sequencing Technologies in Plant Pathology

***Prashanth Kumar A.***
*Scientist (Plant Pathology), JVR Horticultural Research Station, Malyal, SKLTSHU (Telangana), India.*

*(Corresponding author: Prashanth Kumar A.)*
*(Received: 20 March 2023; Revised: 30 April 2023; Accepted: 05 May 2023; Published: 15 May 2023)*
*(Published by Research Trend)*

**ABSTRACT: Molecular techniques offer enhanced precision in species identification, such as Sanger sequencing for fungi and environmental DNA samples. High-throughput DNA sequencing (HTS) methodologies have significantly transformed research in plant and soil biology, enabling a more accurate evaluation of biodiversity within terrestrial ecosystems. However, certain biases persist and require resolution. Metabarcoding, which involves examining the identification of fungi and oomycetes, is another method used to evaluate fungal biodiversity across various settings. The effectiveness of DNA metabarcoding depends on the careful selection of a suitable DNA marker gene. Researchers commonly use DNA barcoding and metabarcoding techniques to analyze fungal communities, but certain groups may not exhibit complete resolution at the species level using the ITS region. Amplicon sequence variant (ASV) methodologies enhance taxonomic identification, while databases like FUN Guild and Fungal Traits can identify ITS sequences from fungal and other eukaryotic organisms. Metabarcoding is a molecular technique used to identify and quantify species within environmental samples, providing cost-efficient approaches for characterizing microbial communities. It has been used to examine plant illnesses such as *Fusarium* Head blight; grapevine trunk diseases and apple replant disease. High-throughput sequencing has enhanced our capacity to evaluate biodiversity in fungal communities across ecosystems. High-throughput sequencing (HTS) enables the sequencing of the entire transcriptome, facilitating the identification of isoforms, unique transcripts, alternative splice variants and genomic variants. However, the accurate taxonomic classification of fungal transcripts at the species level heavily relies on the presence of full genomes. Metabarcoding sequencing is the most widely employed method for plant pest detection and identification due to its favorable cost-efficiency ratio and low risk of false-positive results. However, there is a lack of research focused on the validation of high-throughput sequencing (HTS) approaches for diagnosing phytopathogenic fungi. The analytical sensitivity of high-throughput sequencing can be influenced by factors such as the number of reads produced per sample, the DNA extraction technique employed, and the competition for primers in the PCR reaction. The importance of HTS technologies for the diagnosis of filamentous plant diseases is now recognized, but the cost of sequencing per sample remains unaffordable for several facilities. Long-read sequencing techniques are proposed to address the presence of soil sample sequences lacking homologies in many databases. Next-generation sequencing (NGS) technology can be used for routine detection of fungal pathogens, but the volume of NGS data requires enhancements in management. A specialized pipeline has been created using machine learning classifiers to benefit metabarcoding studies.**

**Keywords:** Epidemic, Metabarcoding, High-throughput sequencing, PCR, transcriptome, DNA extraction and next generation sequencing.

## INTRODUCTION

Plant pathologists are cognizant of the existence of a disease epidemic commonly referred to as "square pyramid disease. The plant itself, the illness-causing pathogen, the environment (specifically the microbiome), the passage of time, and human social systems are just a few of the factors influencing this epidemic. These factors are regarded as the key determinants in the development of the disease (Slippers *et al.*, 2020; Bernardo *et al.*, 2020). This assertion holds significant validity in the context of soilborne illnesses, given the substantial quantity of microorganisms present per gram of soil (Cavicchioli *et al.*, 2019), as well as the extensive and yet unexplored proportion of unidentified fungal species (Blackwell *et al.*, 2011; Hawksworth *et al.*, 2017; Tedersoo *et al.*, 2020). The accurate identification and classification of the species implicated in a certain disease is a crucial aspect to be taken into account during the investigation of the disease's epidemiology and to ensure the efficacy of any disease-control measures, including those of a genetic, chemical, or biological nature. The prompt emphasizes the pressing need for expedient and precise identification of fungal and oomycete diseases since these organisms are responsible for significant agricultural and forestry damages. This urgency stems

from the desire to minimize reliance on chemical interventions and mitigate losses in crop output (Comtet *et al*., 2015). Diagnostic techniques that rely on the cultivation of pathogens and direct observation of their morphology are frequently time-consuming and prone to unreliability, particularly in the case of non-culturable infections. Molecular techniques employing targeted oligonucleotide primers or probes, such as polymerase chain reaction (PCR)-based protocols, DNA hybridization-based methods, and DNA sequence analysis, offer enhanced precision in species identification. However, these approaches may occasionally produce false-positive outcomes and may not have the capability to identify unknown species or all species within a diverse community (Grosdidier *et al*., 2017). In order to analyze the DNA of fungi, the ITS and TEF marker genes must be looked at using Sanger sequencing of PCR products (Aragona *et al*., 2022). This method, although time-consuming and requiring several days, offers a high level of resolution in taxonomic classification. However, it should be noted that the amplification of polymorphic alleles, which is a common occurrence in noncoding areas of eukaryotes, can affect the accuracy of the results. The emergence of next-generation sequencing technologies (NGS) has facilitated the ability to rapidly sequence numerous microbes in a highly parallel manner while also reducing costs and enhancing accuracy. The ability to monitor diseases with complicated causes or track shifts in microbial populations across various settings (such as marine, soil, and air) has been greatly enhanced through the quick collection and analysis of environmental DNA samples (referred to as eDNA). This approach eliminates the requirement for taxonomic expertise and has become increasingly cost-effective. The traditional methods used to isolate the pathogen from symptomatic tissue or soil, followed by morphological identification, have been supplemented and frequently surpassed by the utilization of DNA-based techniques. These tools have enhanced our capacity to promptly and cost-effectively diagnose the emergence and dissemination of a specific pathogen, as well as quantify its biomass within plants or soil (Aragona *et al*., 2022). However, these methodologies require a significant investment of time. The application of techniques for molecular pathogen diagnostics is constrained by the existing understanding of the disease-causing agent. Consequently, the identification of novel or emerging pathogens of concern, particularly in the context of climate change scenarios, becomes challenging (Burdon *et al*., 2020). Nevertheless, these technologies are undergoing constant and rapid advancements, making it challenging to predict which one will emerge as the dominant force in the industry in the foreseeable future (Tedersoo *et al*., 2019; Jamy *et al*., 2020; Piombo *et al*., 2021). The application of high-throughput DNA sequencing (HTS) methodologies has significantly transformed research in the fields of plant and soil biology, enabling a more accurate evaluation of biodiversity within terrestrial ecosystems (Nilsson *et al*., 2019). There are various high-throughput sequencing (HTS) platforms available, categorized as second and third generations. The recent emergence of new sequencer models, including Illumina's MiSeq, Life Technologies' Ion Personal Genome Machine (PGM), PacBio's Sequel IIe, and Oxford Nanopore Technologies' Min-ION system, is causing a significant transformation in this industry. Furthermore, there is a growing number of data analysis pipelines being developed, which has led to the emergence of numerous recent studies focusing on the utilization of high-throughput sequencing (HTS) technologies for diagnosing filamentous fungal plant diseases with species-level precision. This review presents a comprehensive analysis of omics techniques and aims to substantiate the claim that the application of high-throughput sequencing (HTS) technologies in the taxonomic classification of fungal pathogens is becoming increasingly prevalent and tangible rather than remaining a mere aspiration. However, it is important to acknowledge that certain biases persist and require resolution, which are thoroughly examined in this study.

In this review, we examine the current utilization of high-throughput sequencing (HTS) technologies in the diagnostic process of filamentous fungal diseases. The comparison between the identification of viruses and bacteria as plant pathogens is discussed, with references to additional works. The present study also addresses the basic principles, advantages and disadvantages of HTS technologies. Additionally, it provides instances of their application in the diagnosis of some fungal infections and the investigation of genetic diversity. This paper presents a comprehensive examination of the fundamental experimental procedures involved in metabarcoding as well as its use in the identification of fungi and oomycetes, encompassing plant and forest pathogens. Additionally, it explores the utilization of metabarcoding in the evaluation of fungal biodiversity across various settings, such as soil and water. The applicability of metagenomic and meta-transcriptomic methods for diagnostic purposes is also discussed. In this study, we present a case study on the utilization of multilocus analysis (MLST) as an illustrative example of a methodology that has been significantly expedited in terms of experimental investigations and data analysis while also experiencing cost reduction due to the advent of next-generation sequencing (NGS) technology. Phylogenomics and genome-wide association studies (GWAS) are also considered due to their potential to elucidate the genomic underpinnings of fungal speciation and identify novel effectors and/or pathogenic genetic determinants that could serve as potential diagnostic markers at a genome-wide level. This paper primarily examines the impact of high-throughput sequencing (HTS) technology on fungal taxonomy, which has resulted in significant revisions in certain instances. A specific case study with a fungal pathogen affecting tomatoes is presented to illustrate these effects. Next, we address the potential uses in the diagnosis of fungal diseases for phytosanitary objectives, encompassing activities such as surveillance, certification of plant propagation material, monitoring of imported plant material, and other related

areas. The application of high-throughput sequencing (HTS) techniques in the field of fungal pathogen detection is a significant challenge and remains an aspiration. However, the successful implementation of HTS methods for the identification of viruses and bacteria provides promising precedents.

## HIGH-THROUGHPUT DNA SEQUENCING (HTS) TECHNOLOGIES

The second and third generations of high-throughput sequencing (HTS) technology are used in the high-throughput platforms that are currently used to study fungal communities. High-throughput sequencing (HTS) methodologies facilitate comprehensive analysis of the community structure of prokaryotic microorganisms (bacteria and archaea) as well as eukaryotic microorganisms (protists, fungi, and microfauna), allowing for detailed characterization. In the context of soil, a significant proportion of microorganisms are non-cultivable and can only be identified by the utilization of molecular techniques (Francioli *et al*., 2021). High-throughput sequencing (HTS) has facilitated the estimation of species identity and relative abundance in complicated mixtures. This can be achieved in two primary ways (Aragona *et al*., 2022).

a) Metabarcoding uses polymerase chain reaction (PCR) to increase the number of taxonomic marker genes, such as DNA barcodes. High-throughput sequencing (HTS) and subsequent comparison of the obtained sequences to a DNA barcoding database come next.

b) Metagenomics is the utilization of shotgun sequencing to evaluate the combined genomes of a population by extracting DNA from a bulk sample.

In comparison to Sanger sequencing after PCR amplification of specific target genes (Valentini *et al*., 2016), both methods make it easier to find species, require less work, don't harm the ecosystem, and allow identification without prior knowledge of the species.

However, it is imperative that the operations be improved and standardized. The absence of a standardized framework poses challenges in the comparative analysis of diverse studies. Furthermore, the challenge of accurately measuring rates of DNA degradation may lead to a potential misrepresentation of the presence of some species. These methodologies encompass multiple stages of processing, each of which has the potential to introduce substantial biases that can greatly undermine the credibility of the metabarcode or metagenome results.

### *A. Metabarcoding*

The efficacy of DNA metabarcoding is mostly contingent upon the careful selection of a suitable DNA marker gene. The selection of primers should possess suitable coverage of the target group, effective exclusion of outgroups, and the capability to differentiate taxa based on the nucleotide variability of the amplified marker (Tedersoo *et al*., 2022). The utilization of customer-designed primer pairs, as exemplified in the genome-enhanced detection and identification (GEDI) approach outlined below, is a viable option. However, it is equally crucial to incorporate universal primers on barcodes to facilitate cross-study comparisons. In order to analyze fungal communities, researchers commonly employ DNA barcoding and metabarcoding techniques. Among the various markers available, the internal transcribed sequence (ITS) region of the ribosomal RNA (rRNA) is widely utilized due to its numerous copies and ability to provide accurate species-level identification in most fungal groups (Tedersoo *et al*., 2022). However, it is important to note that certain groups, such as Trichoderma, *Fusarium*, or Oomycetes, may not exhibit complete resolution at the species level using the ITS region.

The fundamental process employed in the metabarcoding approach is high-throughput sequencing (HTS), which enables the simultaneous sequencing of all amplicons generated during the PCR step. This PCR step is indicative of the presence of all organisms within the sample. Various sequencing platforms are currently accessible and are undergoing substantial advancements. The sequencing data undergoes a series of processing stages, which include (i) demultiplexing the barcoded samples, (ii) performing pair-end assembly using Illumina technology, and (iii) eliminating chimeric reads. The three main steps involved in this process are: (iv) application of quality filtering; (v) implementation of sequence clustering; and (vi) comparison of the representative sequences with a reference database.

The fundamental aspects of analyzing metabarcoding sequencing data encompass (i) the process of grouping sequences and (ii) the attribution of taxonomic or functional information through comparison with existing databases. Sequence readings are grouped together based on their similarity. In most metabarcoding studies, operational taxonomic units (OTUs) are made by grouping readings that are similar to each other by a certain amount, usually between 95% and 99%. Typically, a threshold of 97% homology is employed. In order to enhance the accuracy and consistency of taxonomic identification, researchers have devised amplicon sequence variant (ASV) methodologies. For future assessments of the community, these methods only look at unique, identical sequences. They focus on OTUs (operational taxonomic units) that are 100% alike. After the clustering process, the reads need to be allocated to a taxon or a function, depending on the reference databases, in order to make taxonomic or functional assignments. The quality of the information obtained improves as the databases become more carefully selected and comprehensive. The identification of ITS sequences from fungal and other eukaryotic organisms is commonly conducted using the UNITE reference data set, which is accessed through the website https://unite.ut.ee (accessed on March 16, 2022). This database is widely utilized due to its extensive curation and inclusion of diverse non-fungal sequences, which aid in the differentiation of fungi from other eukaryotes (Anslan *et al*., 2018; Nilsson *et al*., 2019). Functional

assignments can be made using either the FUNGuild database (Nguyen et al., 2016) or the Fungal Traits database (Polme *et al*., 2020).

It is noteworthy to emphasize that next-generation sequencing (NGS) microbiome-based diagnostics provide a substantial volume of data, necessitating the utilization of machine learning or other resources for the evaluation of human, plant, and soil health (Oh *et al*., 2020; Topcuoglu *et al*., 2020; Krause *et al*., 2021; Wilhelm *et al*., 2022). The efficacy of bioinformatic methodologies in the retrieval of fungal strains and the corresponding proportions of the retrieved strains exhibited significant heterogeneity. The sequence analysis tools, namely USEARCH and VSEARCH, were successful in detecting nearly all strains present in the mock community. However, both methods tended to exaggerate the richness of the community. On the other hand, the DADA2 tool demonstrated more accuracy in retrieving both the true richness and composition of the mock community. The first two methods are better suited for identifying specific species, whereas the third method is better suited for doing studies on community ecology (Pauvert *et al*., 2019).

Metabarcoding refers to a molecular technique used to identify and quantify many species within a given environmental sample. Metabarcoding has demonstrated its efficacy as a cost-efficient approach for characterizing microbial communities. This method enables the assessment of biodiversity within samples, offering a high level of taxonomic precision. Additionally, it facilitates the comparison of sample communities that have been treated with various treatments. From a bioinformatic perspective, shotgun metagenomics is somewhat more challenging to handle because of its higher storage requirements and processing demands. Currently, this method is widely employed as the predominant molecular technique for characterizing microbiota in environmental samples (Francioli *et al*., 2021). This work provides novel insights into the examination of plant illnesses characterized by intricate causes occurring in both the aboveground and belowground components of plants. There are numerous instances of disease complexes that pose significant concerns for agricultural crops, impacting both.

*Fusarium* Head Blight (FHB) is a major wheat disease caused by many *Fusarium* species, many of which can make mycotoxins (Karlsson *et al*., 2021). FHB can affect both above-ground and below-ground tissues. The wheat ear fungus community in a topographically varied environment was characterized using Illumina MiSeq with V3 Chemistry (Schiro *et al*., 2019). In their study, Walder *et al*. (2017) employed PacBio CCS long read sequencing to investigate the alterations occurring in *Fusarium* spp. by targeting a combination of the highly variable internal transcribed spacer (ITS) and the D1-D2-D3 portions of the large subunit (LSU) region. The investigation examines the impact of various cover crops on crop leftovers. In a similar manner, the amplification of bacterial 16S rRNA, fungal ITS, and *Fusarium* spp. is performed using the Illumina MiSeq platform in maize. The utilization of TEF1 areas provides valuable insights into the intricate epidemiology of *Fusarium* head blight (FHB) through the identification and concurrent appearance of various phytopathogenic and beneficial bacteria in maize stalks cultivated in conjunction with wheat (Cove-Diaz *et al*., 2019). The investigation involved the analysis of fungal and bacterial community profiles in wheat straws that were intentionally inoculated with *Zymoseptoria tritici*. This allowed for comprehensive knowledge of the interactions and dynamics between the pathogen and the entire microbial community for a specified period of time (Kerdraon *et al*., 2019). The utilization of next-generation sequencing (NGS) techniques in grapevine cultivation presents numerous advantageous applications in determining the microbial species composition that is pertinent to the process of winemaking (Singh *et al*., 2019; Griggs *et al*., 2021). The utilization of next-generation sequencing (NGS) techniques has been employed to characterize clusters of grapevine trunk diseases, specifically *Eutypa*, *Esca*, *Botryosphaeria*, *Phomopsis* dieback, and black foot. The accurate characterization of these diseases is essential in order to determine and implement the most suitable control strategies. The utilization of Illumina short read technology, together with optimized and universal primers designed to target both the ITS1 and ITS2 rDNA regions, has been employed to validate the existence of the most prominent species associated with each condition. Furthermore, this approach has facilitated the identification of species that have not yet been classified within this particular complex (Morales *et al*., 2018; Del-Frari *et al*., 2019). In contrast to the relatively high level of attention given to wood diseases, the detection of Vitis phylloplane diseases using next-generation sequencing (NGS) has garnered comparatively less interest. However, it is anticipated that the utilization of NGS for this purpose will increase in the near future (Cureau *et al.,* 2021). Next-generation sequencing (NGS) was employed to investigate the impact of elicitors or biocontrol agents on the populations of microorganisms residing on the surface of leaves (Gobbi *et al.,* 2020; Nerva *et al*., 2019). Apple Replant Disease (ARD) is a significant ailment characterized by a multifaceted origin that mostly impacts fruit trees, specifically apples and other members of the Rosaceae family that are replanted in a location previously used for agriculture (Mazzola *et al*., 1998; Winkelmann *et al*., 2019). The primary species involved in many apple locales globally are oomycetes, including *Pythium* spp. and *Phytophthora* spp., as well as fungi, particularly *Cylindrocarpon* spp. and *Rhizoctonia solani*. The management of acid rock drainage (ARD) poses challenges attributed to the limited availability of approved chemical treatments, which are further compounded by the intricate nature of the disease's causative factors. The application of next-generation sequencing (NGS) methods has revealed substantial disparities in microbial composition between newly established sites and replanted sites, particularly in populations of beneficial bacteria such as *Burkholderia* spp., Microcoleus, Nocardioides, sulfur-

oxidizing bacteria, and those involved in nitrogen cycling (Sun et al., 2014). Additionally, these differences have been observed following the use of green manure with *Brassica* spp. The user's text does not contain any information. Next-Generation Sequencing (NGS) has been employed in other studies to investigate and describe the pathobiome of many crops, including oaks (Ruiz-Gomez *et al*., 2019), ginseng (Miao *et al*., 2016), tomato (Poli *et al*., 2016; Johnston-Monje *et al*., 2017), strawberry (Xu *et al*., 2015; Abdelfattah *et al*., 2016), potato (Sugiyama *et al*., 2010), banana (Shen *et al*., 2018) and ramie (*Boehmeria nivea*) (Zhu *et al*., 2018). The metabarcoding technique is commonly employed for the investigation of oomycetes, with a particular focus on *Phytophthora* spp. The second box The utilization of high-throughput sequencing (HTS) has significantly enhanced our capacity to evaluate biodiversity in fungal communities across various ecosystems, including soil, phylloplane, air, and water. This technological advancement has particularly facilitated the monitoring of a specific pathogen's dynamics within its dynamic environment. For instance, it enables the examination of the pathogen's behavior following chemical or biological interventions, as well as the investigation of the impact of climate change or agricultural practices on disease development. Understanding the structure and function of microbiota linked to various settings, such as roots, leaves, suppressive soils, and degraded soils, requires comprehensive knowledge of soil microbial communities and their compositions and diversity (Berg *et al.*, 2015; Crecchio *et al*., 2018; Jiao *et al*., 2019; Topalovic *et al*., 2020; Mazzola *et al*., 2020; Tagele *et al*., 2021).

Several fungi have the ability to adopt several lifestyles, including harmful, saprophytic, or symbiotic. The boundaries of idioms are frequently ambiguous, resulting in a lack of clarity. Individuals have the ability to modify their way of existence, for instance, through the process of endophytes transitioning into parasites and vice versa. This can be achieved through the utilization of omics, which refers to a set of instruments and novel methodologies employed to investigate plant-microbe interactions and gain a better understanding of the various behaviors involved (Bahram *et al*., 2022). Microbiome studies have gained significant traction in the field of environmental research. These studies have been employed to assess biodiversity levels and facilitate conservation efforts in protected regions (Pascher *et al*., 2022; Alem *et al*., 2022). Additionally, they have been utilized to investigate the influence of various factors such as host taxon, tissues, and seasonality on the composition of fungi and bacteria in tropical forests (Li *et al.*, 2022). Furthermore, microbiome studies have been instrumental in comparing the microbiomes of trees and associated herbaceous plants used for phytoremediation purposes (Yung *et al*., 2021).

*B. Metagenomics and Meta transcriptomics*

Shotgun metagenomic research has been of significant importance in the characterization of the taxonomic and functional profiles of microbial communities during the past two decades. Next-generation sequencing (NGS)-based metagenomics technologies were initially employed in the clinical domain for the purpose of pathogen detection (Gu *et al*., 2019). Subsequently, these technologies have gained increasing prominence in the field of plant disease diagnostics (Piombo *et al*., 2021). Shotgun metagenomics is a technique that enables the comprehensive sequencing of the complete genomes of microorganisms found in a given sample. This sample might include many sources, such as symptomatic or asymptomatic host plants, as well as soil and other environmental matrices. This approach is considered reliable and effective for the detection and identification of pathogens. Consequently, a meticulous diagnosis has been established (Mechan *et al*., 2020). This technology offers a significant benefit by eliminating the need for prior isolation of pathogens in culture. This is particularly advantageous for obligate pathogens, which cannot be cultured. Additionally, this technology does not rely on specific probes or primers for individual pathogens, thereby avoiding the biases commonly associated with PCR and metabarcoding amplification. The complexity of the analysis involved in sequencing data necessitates a thorough examination of potential flaws. Following the quality control of the reads and the assembly of metagenomic contigs, the data undergoes a process known as "binning" to generate entire genomes. By enabling the construction of comprehensive genomes through the use of specialized software tools like BUSCO (Simao *et al*., 2015); CheckM (Parks *et al*., 2015), the process of binning makes it easier to identify novel pathogens. But this method has problems when used on samples with fragmented genomes. This is likely because of things like not enough coverage or the presence of very different communities with closely related species (Knight *et al*., 2018). Shotgun metagenomics, when implemented with a high sequencing depth, enables accurate taxonomic identification at the species level. This method is particularly effective for identifying the most prevalent species, especially those with whole genomes stored in databases. In recent literature, notable instances of achieving species-level resolution through the utilization of high-throughput sequencing (HTS) technologies have been documented. For instance, Yang *et al*. (2022) employed the Oxford Nanopore Technologies MinION sequencing platform to differentiate between the boxwood blight fungal pathogens *Calonectria pseudonaviculata* and *Calonectria henricotiae*. Loit *et al*. (2019) conducted a comparative analysis of the capabilities of two third-generation sequencing devices, namely MinION by Oxford Nanopore Technologies and Sequel by Pacific Biosciences. The objective of their investigation was to evaluate the efficacy of these instruments in the identification and diagnosis of fungal and oomycete infections found in Pinaceae and Solanum tissues. This assessment was carried out using a metagenomic method.

The integration of metagenomics and meta-transcriptomics is of significant importance in order to

comprehensively understand the genetic capabilities and metabolically active species within the entire microbiome (Zifcakova *et al*., 2017; Schneider *et al*., 2021; Djemiel *et al*., 2022). High-throughput sequencing (HTS) plays a crucial role in the field of meta-transcriptomics since it enables the sequencing of the entire transcriptome. This comprehensive approach facilitates the identification of isoforms, unique transcripts, alternative splice variants, and subsequently, genomic variants. In their study, Garalde *et al*. (2018) employed Oxford Nanopore Technologies to perform direct sequencing of natural RNA. This approach circumvented the need for reverse transcription and amplification, enabling the acquisition of whole RNA sequences. The annotation of expressed genes is facilitated by the absence of introns, which allows for rapid identification. However, in the context of metagenomics, the accurate taxonomic classification of fungal transcripts at the species level heavily relies on the presence of full genomes. In a recent study, Chialva *et al.* (2019) employed an RNA-seq dataset that had been previously created for tomato plants in order to identify and analyze the taxonomic and functional diversity of the root microbiota.

## CONVENTIONAL TO NEXT-GENERATION SEQUENCING TECHNIQUES

Multilocus sequence typing (MLST), which involves the targeted sequencing of many genomic loci, is widely regarded as a highly reliable and informative approach for molecular genotyping (Chen *et al*., 2015). The method has been extensively employed for evaluating the genetic and pathogenic diversity among several populations within the same species, such as *Colletotrichum* spp. as well as for taxonomic classification purposes. Besides the use of ITS barcoding, MLST has emerged as the prevailing technique for genotyping several fungi in investigations pertaining to molecular epidemiology, pathogenicity and phylogenetics. The utilization of this technique has significant importance in the process of characterizing and classifying novel species (Taylor *et al*., 2003) across several domains. This is particularly relevant in the context of fungal species that are of particular interest to the human population, such as *Candida* (Jackson *et al*., 2009). Ergillus (Bain *et al*., 2007) and Pseudallescheria (Bernhardt *et al*., 2013), as well as Colletotrichum (Veir *et al*., 2012), in the context of plant disease, are examples of relevant organisms. Two fungal genera that are commonly referred to as *Ilyonectria* (Cabrale *et al*., 2012) and *Diaporthe* (Gao *et al*., 2017).

The conventional method of multilocus sequence typing (MLST) entails the isolation of genomic DNA and the amplification of distinct marker gene sequences using polymerase chain reaction (PCR). In many instances, as few as three loci are adequate for the identification of species and strains. In addition, there exist online databases for multiple bacterial and fungal species that serve the purpose of facilitating molecular epidemiological investigations and surveillance. Nevertheless, the existing traditional MLST methodology is burdened with drawbacks like time consumption and high costs, primarily attributed to the utilization of Sanger sequencing (Chen *et al*., 2015).

A study was done on walnut anthracnose (Da Lio *et al*., 2018), and Colletotrichum species were found using a combination strategy that included multilocus analysis. Both the "traditional" Sanger method, which involves isolating and growing pure isolates, and metabarcoding using the ITS region (Illumina MiSeq PE300) were used. Consequently, the identification of the species involved was feasible, exhibiting a notable concurrence between the two methodologies. However, certain limitations of the metabarcoding technique became apparent, primarily due to the presence of both false negative and false positive results. These problems were also caused by the low resolution of the internal transcribed spacer (ITS) region in Colletotrichum and the use of universal ITS1xITS4 primers for plant DNA amplification, which led to a shallow analysis. But in some cases, using metabarcoding may be a more reliable method. This is especially true when marker genes at the species level are known or when the method is based on a pair of perfectly matched primers, as was shown in the case of *Fusarium* spp. (Cobo-Diaz *et al*., 2019).

## PHYLOGENOMIC

Phylogenomics is a field of study that focuses on the reconstruction of evolutionary relationships among species. The significance of sequencing entire fungal genomes is widely recognized across various fields, encompassing sectors such as human and animal medicine, ecology, taxonomy, and agriculture (Aylward *et al*., 2017). In contrast to prior phylogenetic investigations that were based on a limited number. Out of a total of 143 strains, the classification of the two *P. lycopersici* "biotypes" as distinct species, namely *Pseudopyrenochaeta lycopersici* and *Pseudopyrenochaeta terrestris*, has been confirmed through genome analysis.

## NEXT-GENERATION SEQUENCING (NGS)

The utilization of Next-Generation Sequencing (NGS) technology for the identification and characterization of novel fungal pest species signifies a significant advancement in the fields of biosecurity, regulatory compliance, and commercial implications. The increasingly frequent identification of uncharacterized pathogens poses challenges in anticipating the implications of these novel species on quarantine legislation and effectively managing the lists of quarantined and certified pathogens (Martin *et al*., 2016). On the one hand, the utilization of HTS technology has expedited the identification of novel plant pests, thereby necessitating complex decision-making processes by National Plant Protection Organizations (NPPOs). Conversely, the availability of information pertaining to the biology, epidemiology, and pest risk analysis of these pests remains limited and is acquired at a comparatively slower pace. Frameworks have been proposed to address the gap in managing plant viruses, with the aim of characterizing these

recently identified pests and swiftly implementing suitable measures for the intended objectives (Massart *et al*., 2017). It is recommended that this framework be implemented and formalized for any newly discovered organism by high-throughput sequencing (HTS). Olmos *et al*. (2018) provided a comprehensive analysis of the potential and prospects of utilizing high-throughput sequencing (HTS) for regular plant pest diagnosis. The authors outlined four primary areas where HTS can be effectively applied, namely: (i) implementation of surveillance programs; (ii) certification of plant propagation material; (iii) quarantine testing at borders.

The monitoring of imported plant material for emerging potential threats is essential. In the year 2019, the International Plant Protection Convention (IPPC) issued a publication titled "Recommendation on the Utilization of High-Throughput Sequencing (HTS) as a Diagnostic Tool for Phytosanitary Purposes. This document aimed to provide a framework for achieving consistency, uniformity, verification, and quality control in the application of HTS tools (FAO, 2019). At present, metabarcoding and shotgun metagenomics are the predominant high-throughput sequencing (HTS) methodologies employed in the realm of plant pest diagnosis and surveillance. The selection of these approaches is contingent upon the specific objectives and domain of implementation, taking into account the desired outcomes of ascertaining the identities of organisms present and/or elucidating their activities. The metabarcoding study provides insights on the composition of organisms inside a given sample, specifically focusing on fungi and other plant pests. To ensure consistency and comparability, the European Plant Protection Organization (EPPO) has established a standard, PM7/129 (Standard for Diagnostics, 2021), which outlines the use of standardized barcodes for this purpose. The scope of targeted sequencing is constrained to the examination of species that possess genomic markers that are taxonomically identified and documented within the existing database (Charpton *et al*., 2014). Although targeted sequencing in routine plant pest diagnostics has certain limitations, metabarcoding sequencing is currently the most widely employed method for plant pest detection and identification. This preference can be attributed to its favorable cost-efficiency ratio, the availability of numerous well-developed tools (including bioinformatics tools), and the low risk of false-positive results. The aforementioned advantage holds significant importance, particularly in the context of officially diagnosing quarantine pests, due to its underlying economic ramifications.

The incorporation of Next-Generation Sequencing (NGS) analysis as a standard method for monitoring regulated plant pests necessitates the assessment of key factors, including sensitivity, specificity, repeatability, and reproducibility. Regrettably, the current body of peer-reviewed literature about the utilization of Next-Generation Sequencing (NGS) for phytosanitary certification remains limited in its acceptance by National Plant Protection Organizations (NPPOs), European and Mediterranean Plant Protection Organizations (EPPO), and the International Plant Protection Convention (IPPC). Although there have been numerous studies conducted on harmful viruses and viroids (Maree *et al*., 2018; Bester *et al*., 2021), there is currently a lack of research focused on the validation of high-throughput sequencing (HTS) approaches for diagnosing phytopathogenic fungi. The standardization of diagnostic techniques is a crucial aspect of plant protection policy, and this poses a significant challenge for Next-Generation Sequencing (NGS) technologies. The validation of these novel diagnostic tests necessitates the development of new experimental protocols and expertise in bioinformatics, which are not typically accounted for in traditional routine analyses. Furthermore, these conventional methods are continuously evolving and undergoing transformation. At present, numerous phytosanitary services that bear the responsibility of conducting official diagnoses have obtained accreditation in accordance with ISO 17025. This accreditation, at the European level, has been effectively implemented through the criteria established by the European and Mediterranean Plant Protection Organization (EPPO) (Standard-Diagnostics, 2021). While it is necessary for the validation of high-throughput sequencing (HTS) analysis processes to adhere to the established validation criteria outlined in the International Plant Protection Convention (IPPC) and European and Mediterranean Plant Protection Organization (EPPO) standards, additional guidelines are required to ensure quality assurance in the HTS methodology. The analytical sensitivity of high-throughput sequencing (HTS) can be influenced by various factors. These factors include the number of reads produced per sample, the DNA extraction technique employed, the competition for primers in the polymerase chain reaction (PCR) reaction in the case of amplicon sequencing, and variations in the copy number of the specific region being targeted among different organisms. The establishment of this should be conducted through the utilization of reference samples, wherein the results obtained from dilution series of target samples are compared with those derived from other reference tests (Santala *et al*., 2018). Furthermore, the validation of the analytical sensitivity value can be achieved by establishing the relevant parameters in the field of bioinformatic analysis. The degree of analytical specificity in a high-throughput sequencing (HTS) test is contingent upon several factors. These include the chosen sequencing approach, the genetic diversity of the fungal target being analyzed, the software and parameters employed in the bioinformatic studies, the reference sequence database utilized, and the desired level of taxonomic resolution. The specificity of metagenomics sequencing in relation to analytical sensitivity can be influenced by factors such as the quantity of sequenced reads and the level of target coverage. To ensure reproducibility, it is essential to establish a defined threshold level by utilizing reference target and non-target samples and subsequently validating these parameters through bioinformatic analysis.

## CONCLUSIONS

The importance of HTS technologies for the diagnosis of filamentous plant diseases caused by fungi and oomycetes, as well as their role in enhancing plant disease management, is now generally recognized. Nevertheless, it is imperative to acknowledge that there remain economic and technical factors that necessitate careful consideration prior to the realization and widespread implementation of this aspiration. Despite the potential solution of adding additional samples in each run, the cost of sequencing per sample remains unaffordable for several facilities. This suggests that more measures of DNA purification are necessary in order to prevent the occurrence of unanticipated and undesirable sequencing artifacts. These measures may include the utilization of a dummy sample and the inclusion of DNA derived from healthy plant tissues as controls. The presence of soil sample sequences lacking homologies in many databases, sometimes referred to as "dark taxa" or "dark matter fungi," has been frequently seen (Page *et al.,* 2016). The aforementioned phenomenon can be attributed to the substantial presence of fungal species that are non-cultivable and have not yet been adequately documented. Additionally, the limited taxonomic resolution achieved by the utilization of "short-reads" sequences of the rRNA barcodes has also played a role in this matter (Tedersoo *et al*., 2014). The proposed solution to address this issue is the utilization of long-read sequencing techniques to sequence the complete rRNA operon, encompassing the large subunit (LSU), internal transcribed spacer (ITS), and small subunit (SSU) (Jamy *et al*., 2020; Tedersoo *et al*., 2017; Latz *et al*., 2022). The scientific community has not fully embraced the suggestion to include intracellular DNA, sometimes known as metagenomic DNA or mgDNA, as a type (Burgaz *et al*., 2018; Lucking *et al*., 2021; Hongsanan *et al*., 2018). One problem with using Next-Generation Sequencing (NGS) technology to find fungal pathogens on a regular basis is that it needs to be made easier to manage the large amount of NGS data. This includes things like server capacity and memory power, as well as the availability of bioinformatic skills, such as algorithms and expert personnel. In order to achieve this objective, a specialized pipeline has been created, utilizing machine learning classifiers, as a viable approach for assigning error-prone sequence-long readings to certain taxa (Krause *et al*., 2021; Enjes *et al*., 2021; Yang *et al*., 2020). Metabarcoding studies, which encompass investigations into human diseases, can benefit from the application of machine learning (ML) modeling. ML modeling can aid in the prediction of disease outcomes and the analysis of environmental factors that influence microbial composition. This approach is relevant not just in the context of agricultural and natural ecosystems but also in the broader field of metabarcoding research (Chang *et al*., 2017; Zhou *et al*., 2019; Sharma *et al*., 2022; Namkung *et al*., 2020).

## REFERENCES


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

Alem, D., Dejene, T., Geml, J., Oria-de-Rueda, J.A. and Martín-Pinto, P. (2022). Metabarcoding analysis of the soil fungal community to aid the conservation of underexplored church forests in Ethiopia. *Sci. Rep., 12*, 4817.

Anslan, S., Nilsson, R.H., Wurzbacher, C., Baldrian, P., Tedersoo, L. and Bahram, M. (2018). Great differences in performance and outcome of high-throughput sequencing data analysis platforms for fungal metabarcoding. *MycoKeys, 39*, 29–40.

Aragona, M., Haegi, A. and Valenta, M. T. (2022). New-Generation Sequencing Technology in Diagnosis of Fungal Plant Pathogens: A Dream Comes True? *J. Fungi*, *8*, 737.

Bahram, M. and Netherway, T. (2022). Fungi as mediators linking organisms and ecosystems. *FEMS Microbiol. Rev*., *46*, fuab058.

Burgaz, A. R., Calatayud, V., Callac, P., Campo, E., Candusso, M., Capoen, B. and Roy, E. (2018). Considerations and consequences of allowing DNA sequence data as types of fungal taxa. *IMA Fungus*, *9*, 167–175.

Bain, J.M., Tavanti, A., Davidson, A.D., Jacobsen, M.D., Shaw, D., Gow, N.A.R. and Odds, F.C. (2007). Multilocus sequence typing of the pathogenic fungus Aspergillus fumigatus. *J. Clin. Microbiol.*, *45*, 1469–1477.

Behrens, F. H. and Fischer, M. (2022). Evaluation of different phyllosphere sample types for parallel metabarcoding of Fungi and Oomycetes in Vitis vinifera. *Phytobiomes J.*

Berg, G. (2015). Beyond borders: Investigating microbiome interactivity and diversity for advanced biocontrol technologies. *Microb. Biotechnol.*, *8*, 5–7.

Bernardo-Cravo, A. P., Schmeller, D., Chatzinotas, A., Vredenburg, V. and Loyau, A.(2020). Environmental factors and host microbiomes shape host–pathogen dynamics. *Trends Parasitol.*, *36*, 616–633.

Bernhardt, A., Sedlacek, L., Wagner, S., Schwarz, C., Würstl, B. and Tintelnot, K. (2013). Multilocus sequence typing of Scedosporium apiospermum and Pseudallescheria boydii isolates from cystic fibrosis patients. *J. Cyst. Fibros. Off. J. Eur. Cyst. Fibros. Soc.*, *12*, 592–598.

Bester, R., Cook, G., Breytenbach, J.H., Steyn, C., De Bruyn, R. and Maree, H. J. (2021). Towards the validation of high-throughput sequencing (HTS) for routine plant virus diagnostics: Measurement of variation linked to HTS detection of citrus viruses and viroids. *Virol. J.*, *18*, 61.

Blackwell, M. (2020). The fungi: 1, 2, 3 5.1 million species? *Am. J. Bot.* 2011, *98*, 426–438.

Burdon, J. J. and Zhan, J. (2020). Climate change and disease in plant communities. *PLoS Biol*., *18*, e3000949.

Cabral, A., Groenewald, J. Z., Rego, C., Oliveira, H. and Crous, P. W. (2012). Cylindrocarpon root rot: Multi-gene analysis reveals novel species within the *Ilyonectria radicicola* species complex. *Mycol. Prog.*, *11*, 655–688.

Cavicchioli, R., Ripple, W. J., Timmis, K. N., Azam, F., Bakken, L. R., Baylis, M., Behrenfeld, M. J., Boetius, A., Boyd, P. W. and Classen, A. T. (2020). Scientists' warning to humanity: Microorganisms and climate change. *Nat. Rev. Microbiol.* 2019, *17*, 569–586.

Chang, H. X., Haudenshield, J. S., Bowen, C. R. and Hartman, G. L. (2017). Metagenome-wide association study and machine learning prediction of bulk soil microbiome and crop productivity. *Front. Microbiol., 8*, 519.

Chen, Y., Frazzitta, A. E., Litvintseva, A. P., Fang, C., Mitchell, T. G., Springer, D. J., Ding, Y., Yuan, G. and Perfect, J. R. (2015). Next generation multilocus sequence typing (NGMLST) and the analytical software program MLSTEZ enable efficient, cost-effective, high- throughput, multilocus sequencing typing. *Fungal Genet. Biol., 75*, 64–71.

Chialva, M., Ghignone, S., Novero, M., Hozzein, W.N., Lanfranco, L. and Bonfante, P. (2019). Tomato RNA-seq data mining reveals the taxonomic and functional diversity of root-associated microbiota. *Microorganisms*, *8*, 38.

Cobo-Díaz, J. F., Baroncelli, R., Le Floch, G. and Picot, A. (2019). Combined metabarcoding and co-occurrence network analysis to profile the bacterial, fungal and *Fusarium* communities and their interactions in maize stalks. *Front. Microbiol., 10*, 261.

Comtet, T., Sandionigi, A., Viard, F. and 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.

Crecchio, C., Mimmo, T., Bulgarelli, D., Pertot, I., Pii, Y., Perazzolli, M., Scagliola, M. and Cesco, S. (2018). Beneficial soil microbiome or sustainable agriculture production. In *Sustainable Agriculture Reviews 31: Biocontrol*, Lichftouse, E., Ed., Springer Nature: Cham, Switzerland, pp. 482–483.

Cureau, N., Threlfall, R., Savin, M., Marasini, D., Lavefve, L. and Carbonero, F. (2021). Year, location, and variety impact on grape-, soil-, and leaf-associated fungal microbiota of arkansas-grown table grapes. *Microb. Ecol., 82*, 73–86.

Da Lio, D., Cobo-Díaz, J. F., Masson, C., Chalopin, M., Kebe, D., Giraud, M., Verhaeghe, A., Nodet, P., Sarrocco, S. and Le Floch, G. (2018). Combined metabarcoding and multi-locus approach for genetic characterization of Colletotrichum species associated with common walnut (*Juglans regia*) anthracnose in France. *Sci. Rep., 8*, 10765.

Damm, U., Cannon, P. F., Woudenberg, J. H. C. and Crous, P. W. (2012). The *Colletotrichum acutatum* species complex. *Stud. Mycol., 73*, 37–113.

Del Frari, G., Gobbi, A., Aggerbeck, M. R., Oliveira, H., Hansen, L. H. and Ferreira, R. B. (2019). Characterization of the wood mycobiome of Vitis vinifera in a vineyard affected by esca. Spatial distribution of fungal communities and their putative relation with leafsymptoms. *Front. Plant Sci., 10*, 910.

Djemiel, C., Maron, P. A., Terrat, S., Dequiedt, S., Cottin, A. and Ranjard, L. (2022). Inferring microbiota functions from taxonomic genes: A review. *Giga Science*, *11*, giab090.

Eenjes, T., Hu, Y., Irinyi, L., Hoang, M. T. V., Smith, L. M., Linde, C. C., Milgate, A. W., Meyer, W., Stone, E. A. and Rathjen, J. P. (2021). Linked machine learning classifiers improve species classification of fungi when using error-prone long-reads on extended metabarcodes. *bioRxiv*.

FAO (2019). Preparing to use high-throughput sequencing (HTS) technologies as a diagnostic tool for phytosanitary purposes. *Comm. Phytosanitary Meas. Recomm., 8*, 1–6.

Francioli, D., Lentendu, G., Lewin, S. and Kolb, S. (2021). DNA Metabarcoding for the Characterization of Terrestrial Microbiota-Pitfalls and Solutions. *Microorganisms*, *9*, 361.

Gao, Y., Liu, F., Duan, W., Crous, P.W. and Cai, L. (2017). Diaporthe is paraphyletic. *IMA Fungus*, *8*, 153–187.

Garalde, D. R., Snell, E. A., Jachimowicz, D., Sipos, B., Lloyd, J. H., Bruce, M., Pantic, N., Admassu, T., James, P. and Warland, A. (2018). Highly parallel direct RNA sequencing on an array of nanopores. *Nat. Methods*, *15*, 201–206.

Gobbi, A., Kyrkou, I., Filippi, E., Ellegaard-Jensen, L. and Hansen, L. H. (2020). Seasonal epiphytic microbial dynamics on grapevine leaves under biocontrol and copper fungicide treatments. *Sci. Rep., 10*, 681.

Griggs, R. G., Steenwerth, K. L., Mills, D. A., Cantu, D. and Bokulich, N. A. (2021). Sources and assembly of microbial communities in vineyards as a functional component of winegrowing. *Front. Microbiol., 12*, 673810.

Grosdidier, M., Aguayo, J., Marçais, B. and Ioos, R. (2017). Detection of plant pathogens using real-time PCR: How reliable are late Ct values? *Plant. Pathol., 66*, 359–367.

Gu, W., Miller, S. and Chiu, C. Y. (2019). Clinical metagenomic next-generation sequencing for pathogen detection. *Annu. Rev. Pathol., 14*, 319–338.

Hawksworth, D. L. and Lücking, R. (2017). Fungal diversity revisited: 2.2 to 3.8 million species. *Microbiol. Spectr., 5*, 4.

Hongsanan, S., Jeewon, R., Purahong, W., Xie, N., Liu, J. K., Jayawardena, R. S., ... & Peršoh, D. (2018). Can we use environmental DNA as holotypes?. *Fungal Diversity*, *92*, 1-30.

Jackson, A. P., Gamble, J. A., Yeomans, T., Moran, G. P., Saunders, D., Harris, D. and Berriman, M. (2009). Comparative genomics of the fungal pathogens *Candida dubliniensis* and *Candida albicans*. *Genome Res., 19*, 2231–2244.

Jamy, M., Foster, R., Barbera, P., Czech, L., Kozlov, A., Stamatakis, A., Bending, G., Hilton, S., Bass, D. and Burki, F. (2020). Long-read metabarcoding of the eukaryotic rDNA operon to phylogenetically and taxonomically resolve environmental diversity. *Mol. Ecol. Resour., 20*, 429–443.

Jiao, S., Chen, W. and Wei, G. (2019). Resilience and assemblage of soil microbiome in response to chemical contamination combined with plant growth. *Appl. Environ. Microbiol., 85*, e02523-18.

Johnston-Monje, D., Loewen, S. and Lazarovits, G. (2017). Mycobiomes of tomato plants with vine decline. *Can. J. Plant Pathol., 39*, 184-200.

Karlsson, I., Persson, P. and Friberg, H. (2021). *Fusarium* Head Blight from a microbiome perspective. *Front. Microbiol., 12*, 628373.

Kerdraon, L., Barret, M., Laval, V. and Suffert, F. (2019). Differential dynamics of microbial community networks help identify microorganisms interacting with residue-borne pathogens: The case of *Zymoseptoria tritici* in wheat. *Microbiome*, *7*, 125.

Knight, R., Vrbanac, A., Taylor, B. C., Aksenov, A., Callewaert, C., Debelius, J., Gonzalez, A., Kosciolek, T., McCall, L. I. and McDonald, D. (2018). Best practices for analysing microbiomes. *Nat. Rev. Microbiol., 16*, 410–422.

Krause, T., Wassan, J.T., McKevitt, P., Wang, H., Zheng, H. and Hemmje, M. (2021). Analyzing large microbiome datasets using machine learning and big data. *Bio. Med. Informatics*, *1*, 138–165.

Latz, M. A. C., Grujcic, V., Brugel, S., Lycken, J., John, U., Karlson, B., Andersson, A. and Andersson, A. F. (2022). Short- and long-read metabarcoding of the eukaryotic rRNA operon: Evaluation of primers and comparison to shotgun metagenomics sequencing. *Mol. Ecol. Resour.*, 1–15.

Li, M., Hong, L., Ye, W., Wang, Z. and Shen, H. (2022). Phyllosphere bacterial and fungal communities vary with host species identity, plant traits and seasonality in a subtropical forest. *Environ. Microbiome*, *17*, 29.

Lücking, R., Aime, M. C., Robbertse, B., Miller, A. N., Aoki, T., Ariyawansa, H. A. and Schoch, C. L. (2021). Fungal taxonomy and sequence- based nomenclature. *Nat. Microbiol.*, *6*, 540–548.

Maree, H. J., Fox, A., Al Rwahnih, M., Boonham, N. and Candresse, T. (2018). Application of HTS for routine plant virus diagnostics: State of the art and challenges. *Front. Plant Sci*., *9*, 1082.

Martin, R. R., Constable, F. and Tzanetakis, I. E. (2016). Quarantine regulations and the impact of modern detection methods. *Annu. Rev. Phytopathol.*, *54*, 189–205.

Massart, S., Candresse, T., Gil, J., Lacomme, C., Predajna, L., Ravnikar, M. and Wetzel, T. (2017). A framework for the evaluation of biosecurity, commercial, regulatory, and scientific impacts of plant viruses and viroids identified by NGS technologies. *Front. Microbiol*., *8*, 45.

Mazzola, M. (1998). Elucidation of the microbial complex having a causal role in the development of apple replant disease in Washington. *Phytopathology*, *88*, 930–938.

Mazzola, M., Graham, D., Wang, L., Leisso, R. and Hewavitharana, S. S. (2020). Application sequence modulates microbiome composition, plant growth and apple replant disease control efficiency upon integration of anaerobic soil disinfestation and mustard seed meal amendment. *Crop Prot.*, *132*, 105125.

Mazzola, M., Hewavitharana, S. S. and Strauss, S. L. (2015). Brassica seed meal soil amendments transform the rhizosphere microbiome and improve apple production through resistance to pathogen reinfestation. *Phytopathology*, *105*, 460–469.

Mechan Llontop, M. E., Sharma, P., Aguilera Flores, M., Yang, S., Pollok, J., Tian, L., Huang, C., Rideout, S., Heath, L. S. and Li, S. (2020). Strain-level identification of bacterial tomato pathogens directly from metagenomic sequences. *Phytopathology*, *110*, 768–779.

Miao, C. P., Mi, Q. L., Qiao, X. G., Zheng, Y. K., Chen, Y. W., Xu, L. H., Guan, H. L. and Zhao, L. X. (2016). Rhizospheric fungi of *Panax notoginseng*: Diversity and antagonism to host phytopathogens. *J. Ginseng Res.*, *40*, 127–134.

Morales-Cruz, A., Figueroa-Balderas, R., García, J. F., Tran, E., Rolshausen, P. E., Baumgartner, K. and Cantu, D. (2018). Profiling grapevine trunk pathogens in planta: A case for community-targeted DNA metabarcoding. *BMC Microbiol.*, *18*, 214.

Namkung, J. (2020). Machine learning methods for microbiome studies. *J. Microbiol.*, *58*, 206–216.

Nerva, L., Pagliarani, C., Pugliese, M., Monchiero, M., Gonthier, S., Gullino, M.L., Gambino, G., Chitarra, W. (2019). Grapevine phyllosphere community analysis in response to elicitor application against powdery mildew. *Microorganisms*, *7*, 662.

Nguyen, N. H., Song, Z., Bates, S. T., Branco, S., Tedersoo, L., Menke, J., Schilling, J. S. and Kennedy, P. G. (2016). FUNGuild: An open annotation tool for parsing fungal community datasets by ecological guild. *Fungal Ecol.*, *20*, 241–248.

Nilsson, R.H., Anslan, S., Bahram, M., Wurzbacher, C., Baldrian, P. and Tedersoo, L. (2019). Mycobiome diversity: High-throughput sequencing and identification of fungi. *Nat. Rev. Microbiol.*, *17*, 95–109.

Nilsson, R. H., Larsson, K. H., Taylor, A. F., Bengtsson-Palme, J., Jeppesen, T. S., Schigel, D., Kennedy, P., Picard, K., Glöckner, F. O. and Tedersoo, L. (2019). The UNITE database for molecular identification of fungi: Handling dark taxa and parallel taxonomic classifications. *Nucleic Acids Res.*, *47*, D259–D264.

Oh, M. and Zhang, L. (2020). DeepMicro: Deep representation learning for disease prediction based on microbiome data. *Sci. Rep.*, *10*, 6026.

Olmos, A., Boonham, N., Candresse, T., Gentit, P., Giovani, B., Kutnjak, D., Liefting, L., Maree, H. J., Minafra, A. and Moreira, A. (2018). High-throughput sequencing technologies for plant pest diagnosis: Challenges and opportunities. *EPPO Bull.*, *48*, 219–224.

Page, R. D. (2016). DNA barcoding and taxonomy: Dark taxa and dark texts. *Philos. Trans. R. Soc. B Biol. Sci.*, *371*, 20150334.

Parks, D. H., Imelfort, M., Skennerton, C. T., Hugenholtz, P. and Tyson, G. W. (2015). CheckM: Assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. *Genome Res.*, *25*, 1043–1055.

Pascher, K., Švara, V. and Jungmeier, M. (2022). Environmental DNA-based methods in biodiversity monitoring of protected areas: Application range, limitations, and needs. *Diversity*, *14*, 463.

Pauvert, C., Buee, M., Laval, V., Edel-Hermann, V., Fauchery, L., Gautier, A., Lesura, I., Vallance, J. and Vacher, C. (2019). Bioinformatics matters: The accuracy of plant and soil fungal community data is highly dependent on the metabarcoding pipeline. *Fungal Ecol.*, *41*, 23–33.

Piombo, E., Abdelfattah, A., Droby, S., Wisniewski, M., Spadaro, D. and Schena, L. (2021). Metagenomics approaches for the detection and surveillance of emerging and recurrent plant pathogens. *Microorganisms*, *9*, 188.

Poli, A., Lazzari, A., Prigione, V., Voyron, S., Spadaro, D. and Varese, G. C. (2016). Influence of plant genotype on the cultivable fungi associated to tomato rhizosphere and roots in different soils. *Fungal Biol.*, *120*, 862–872.

Polme, S., Abarenkov, K., Nilsson, R. H., Lindahl, B.D., Clemmensen, K. E., Kauserud, H., Nguyen, N., Kjøller, R., Bates, S. T. and Baldrian, P. (2020). Fungal Traits: A user-friendly traits database of fungi and fungus-like stramenopiles. *Fungal Divers.*, *105*, 1–16.

Ruiz Gómez, F.J., Navarro-Cerrillo, R.M., Pérez-de-Luque, A., Oβwald, W., Vannini, A. and Rodríguez, C. M. (2019). Assessment of functional and structural changes of soil fungal and oomycete communities in holm oak declined dehesas through metabarcoding analysis. *Sci. Rep.*, *9*, 5315.

Santala, J. and Valkonen, J. P. T. (2018). Sensitivity of small RNA-based detection of plant viruses. *Front. Microbiol.*, *9*, 939.

Schiro, G., Colangeli, P. and Müller, M. E. H. (2019). A metabarcoding analysis of the mycobiome of wheat ears across a topographically heterogeneous field. *Front. Microbiol.*, *10*, 2095.

Schneider, A. N., Sundh, J., Sundström, G., Richau, K., Delhomme, N., Grabherr, M., Hurry, V. and Street, N. R. (2021). Comparative fungal community analyses using metatranscriptomics and internal transcribed spacer amplicon sequencing from Norway spruce. *Msystems*, *6*, e00884-20.

Sharma, M., Kumar, C. J. and Dekac, A. (2022). Early diagnosis of rice plant disease using machine learning techniques. *Arch. Phytopathol. Plant Prot.*, *55*, 259–283.

Sharpton, T. J. (2014). An introduction to the analysis of shotgun metagenomic data. *Front. Plant Sci.*, *16*, 209.

Shen, Z., Penton, C. R., Lv, N., Xue, C., Yuan, X., Ruan, Y., Li, R. and Shen, Q. (2020). Banana *fusarium* wilt disease incidence is influenced by shifts of soil microbial communities under different monoculture spans. *Microb. Ecol.*, *75*, 739–750.

Simao, F. A., Waterhouse, R. M., Ioannidis, P., Kriventseva,

E. V. and Zdobnov, E. M. (2015). BUSCO: Assessing genome assembly and annotation completeness with single-copy orthologs. *Bioinformatics*, *31*, 3210–3212.

Singh, P., Santoni, S., Weber, A., This, P. and Péros, J. P. (2019). Understanding the phyllosphere microbiome assemblage in grape species (Vitaceae) with amplicon sequence data structures. *Sci. Rep.*, *9*, 14294.

Slippers, B. (2020). The plant disease pyramid: The relevance of the original vision of plant pathology in 2020. *S. Afr. J. Sci.*, *116*, 1–3.

Standard-Diagnostics, E. P. P. O. (2021). PM 7/129 (2) DNA Barcoding as an Identification Tool for a Number of Regulated Pests. *EPPOBull.*, *51*, 100–143.

Standard-Diagnostics, E. P. P. O. (2021). PM 7/98 (5). Specific requirements for laboratories preparing accreditation for a plant pest diagnostic activity. *EPPO Bull.*, *51*, 468–498.

Sugiyama, A., Vivanco, J. M., Jayanty, S. S. and Manter, D. K. (2010). Pyrosequencing assessment of soil microbial communities in organic and conventional potato farms. *Plant Dis.*, *94*, 1329–1335.

Sun, J., Zhang, Q., Zhou, J. and Wei, Q. (2014). Illumina amplicon sequencing of 16s rRNA tag reveals bacterial community development in the rhizosphere of apple nurseries at a replant disease site and a new planting site. *PLoS ONE*, *9*, e111744.

Tagele, S. B., Kim, R. H. and Shin, J. H. (2021). Interactions between Brassica biofumigants and soil microbiota: Causes and impacts. *J. Agric. Food Chem.*, *69*, 11538–11553.

Taylor, J. W. and Fisher, M. C. (2003). Fungal multilocus sequence typing: It's not just for bacteria. *Curr. Opin. Microbiol.*, *6*, 351–356.

Tedersoo, L., Anslan, S., Bahram, M., Kõljalg, U. and Abarenkov, K. (2020). Identifying the 'unidentified' fungi: A global-scale long-read third-generation sequencing approach. *Fungal Divers.*, *103*, 273–293.

Tedersoo, L., Bahram, M., Põlme, S., Kõljalg, U., Yorou, N. S., Wijesundera, R., Ruiz, L. V., Vasco-Palacios, A. M., Thu, P. Q. and Suija, A. (2014). Global diversity and geography of soil fungi. *Science*, *346*, 1078.

Tedersoo, L., Bahram, M., Puusepp, R., Nilsson, R. H. and James, T. Y. (2017). Novel soil-inhabiting clades fill gaps in the fungal tree of life. *Microbiome*, *5*, 42.

Tedersoo, L., Bahram, M., Zinger, L., Nilsson, R.H., Kennedy, P.G., Yang, T., Anslan, S. and Mikryukov, V. (2022). Best practices in metabarcoding of fungi: From experimental design to results. *Mol. Ecol.*, *31*, 2769–2795.

Tedersoo, L., Drenkhan, R., Anslan, S., Morales-Rodriguez, C. and Cleary, M. (2019). High-throughput identification and diagnostics of pathogens and pests: Overview and practical recommendations. *Mol. Ecol. Resour.*, *19*, 47–76.

Topalovic´, O., Hussain, M. and Heuer, H. (2020). Plants and associated soil microbiota cooperatively suppress plant-parasitic nematodes. *Front. Microbiol.*, *11*, 313.

Topçuoğlu, B. D., Lesniak, N. A., Ruffin, M. T., Wiens, J. and Schloss, P. D. (2020). A framework for effective application of machine learning to microbiome-based classification problems. *mBio*, *11*, e00434-20.

Valentini, A., Taberlet, P., Miaud, C., Civade, R., Herder, J., Thomsen, P. F., Bellemain, E., Besnard, A., Coissac, E. and Boyer, F. (2016). Next-generation monitoring of aquatic biodiversity using environmental DNA metabarcoding. *Mol. Ecol.*, *25*, 929–942.

Weir, B. S., Johnston, P. R. and Damm, U. (2012). The *Colletotrichum gloeosporioides* species complex. *Stud. Mycol.*, *73*, 115–180.

Wilhelm, R. C., van Es, H. M. and Buckley, D. H. (2022). Predicting measures of soil health using the microbiome and supervised machine learning. *Soil Biol. Biochem.*, *164*, 108472.

Winkelmann, T., Smalla, K., Amelung, W., Baab, G., Grunewaldt Stöcker, G., Kanfra, X., Meyhöfer, R., Reim, S., Schmitz, M. and Vetterlein, D. (2019). Apple replant disease: Causes and mitigation strategies. *Curr. Issues Mol. Biol.*, *30*, 89–106.

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

Yang, A., Zhang, W., Wang, J., Yang, K., Han, Y. and Zhang, L. (2020). Review on the application of machine learning algorithms in the sequence data mining of DNA. *Front. Bioeng. Biotechnol.*, *8*, 1032.

Yang, S., Johnson, M. A., Hansen, M. A., Bush, E., Li, S. and Vinatzer, B. A. (2022). Metagenomic sequencing for detection and identification of the boxwood blight pathogen *Calonectria pseudonaviculata*. *Sci. Rep.*, *12*, 1399.

Yung, L., Bertheau, C., Tafforeau, F., Zappelini, C., Valot, B., Maillard, F., Selosse, M. A., Viotti, C., Binet, P. and Chiapusio, G. (2021). Partial overlap of fungal communities associated with nettle and poplar roots when co-occurring at a trace metal contaminated site. *Sci. Total Environ.*, *782*, 146692.

Zhou, Y.-H. and Gallins, P. (2019). A review and tutorial of machine learning methods for microbiome host trait prediction. *Front. Genet.*, *10*, 579.

Zhu, S., Wang, Y., Xu, X., Liu, T., Wu, D., Zheng, X., Tang, S. and Dai, Q. (2018). Potential use of high-throughput sequencing of soil microbial communities for estimating the adverse effects of continuous cropping on ramie (*Boehmeria nivea* L. Gaud). *PLoS ONE*, *13*, e0197095.

Žifčáková, L., Větrovský, T., Lombard, V., Henrissat, B., Howe, A. and Baldrian, P. (2017). Feed in summer, rest in winter: Microbial carbonutilization in forest topsoil. *Microbiome*, *5*, 122.

**How to cite this article:** Prashanth Kumar A. (2023). Advances in Sequencing Technologies in Plant Pathology. *Biological Forum – An International Journal*, *15*(5a): 538-548.