Review

Metagenomics Next Generation Sequencing (mNGS): An Exciting Tool for Early and Accurate Diagnosing Plant Fungal Pathogens

F. Şeyma Gökdemir[1,2], Özlem D. İşeri[2,3], Abhishek Sharma[4*], Premila Achar[5], Füsun Eyidoğan[2,3]

1Michigan State University, Plant Soil and Microbial Science, East Lansing, Michigan, USA; fsgokdemir@gmail.com

2 Başkent University, Department of Molecular Biology and Genetics, Faculty of Arts and Sciences, Ankara, Turkey; oiseri@gmail.com

3 Başkent University, Institute of Food, Agriculture and Livestock Development, Ankara, Turkey; fusunie@gmail.com

4 Amity Food and Agriculture Foundation, Amity University, Noida-201313, Uttar Pradesh, India; asharma5@amity.edu

5 Department of Molecular and Cellular Biology, Kennesaw State University, Kennesaw, GA 30144, USA; pachar@kennesaw.edu

*Correspondence: asharma5@amity.edu (AS)

Abstract: Crop output is directly impacted by plant infections (with fungi as the major pathogen), making accurate diagnosis of these threats crucial. Developing technology and multidisciplinary approaches are turning to genomic analyses in addition to traditional culture methods in diagnosing fungal plant diseases. The metagenomic next-generation sequencing (mNGS) method is preferred for genotyping identification of organisms, identification at the species level, illumination of metabolic pathways, and determination of microbiota. Moreover, the data obtained so far show that this new approach shows promise as an emerging new trend in fungal disease detection. Another approach covered by mNGS technologies, known as metabarcoding, enables use of specific markers specific to a genetic region and allows for genotypic identification by facilitating the sequencing of certain regions. Although the core concept of mNGS remains constant across applications, the specific sequencing methods and bioinformatics tools used to analyze the data do differ. In this review, we focus on how mNGS technology, including metabarcoding, is applied in fungal pathogenesis and its promising developments for the future.

Keywords: mNGS; Metabarcoding; ITS; Biyoinformatic Analysis; Fungal Pathogenesis

1. Introduction

Farmers worldwide have struggled with crop losses caused by pathogens, including bacteria, viruses, and fungi. The main biotic stress that causes the most economic damage and losses are fungal pathogens. Although the course of the disease and the loss of crops vary according to the host plant, sometimes up to 100% crop losses are experienced. These losses will pave the way for food shortages and ecological degradation in the future. Difficulties in culturing and diagnosing organisms are at the forefront of the unavoidable reasons for yield losses. Therefore, It is crucial to have state-of-the-art methods for detecting and preventing crop diseases if we reduce the loss due to disease at all stages of crop production (from growth through harvest and postharvest processing) and ensure agricultural sustainability. The discovery of ribosomal RNA (rRNA) genes and the realization that not all microorganisms can be cultured in the laboratory led to the creation of the Sanger sequencing technique [1,2,47]. Subsequently, the development of gene expression techniques that enable the discovery of new genes and metabolic products inspired the "metagenomic" science, which provides all genomic information that can be obtained without culturing under in vitro conditions.

DNA sequencing approaches provide basic information about the diversity of living things of biological importance. Despite their high cost, Sanger sequencing technologies are one of the most preferred methods in sequencing technologies. However, as an alternative to this; Many sequencing technologies are widely used, including third or next generation sequencing technologies (NGS) such as Illumina, Ion Torrent, HeliScope, Pacific Biosciences (PacBio),

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 26 August 2022 doi:10.20944/preprints202208.0457.v1

454/Roche, Se-quencing by Oligo Ligation Detection (SOLiD), and Oxford Nanopore. It is preferred and reduces the high sequencing cost [3]. Next-generation sequencing technologies enable the sequencing of part or all of an organism's genome. However, mNGS, which includes 3rd generation technologies, also allows us to learn about living variance and population genetics. Moreover, metagenomic next-generation sequencing (mNGS) can be used to provide information on the diversity of biologically important resources, analyze DNA sequences, uncover details of metabolic pathways, identify homology-based genes, discover industrially important enzymes, etc. Solves important problems such as detection of viral and fungal pathogens.

mNGS technologies are now regularly employed to assess the phylogeny and functionality of non-cultivable microbes, though human pathogens take precedence over plant pathogens. Although metagenomic sequencing technologies have just begun to be used in plant sciences, promising results are obtained for the future and are beginning to gain importance in agronomic sciences. This study explains how mNGS technology is used in fungal pathogen detection in Agronomic sciences.

2. Multiple Real-World Applications for mNGS

mNGS technologies can be optimized for use in many areas today (Figure 1). Even if, each usage area seems different, mNGS is a common point thanks to the similarity of the specific barcodes used and the method (Figure 2). One of the primary purposes of mNGS is detecting all culturable and non-culturable organisms in the medium or a host. For this reason, varieties can be scanned by using barcodes specific to the species to be determined. 16 S rRNA-based universal barcodes are used most for bacteria, while barcodes from the ITS region are preferred for fungal pathogenicity. Evolutionary and ecological studies have a vital role in the development of metagenomic science. The first discovery of proteorhodopsin proteins occurred in environmental DNA. Complete genome data of microbial communities found in samples from environmental samples can be obtained today, with scientists aiming to reveal whole genomes. Environmental genomes obtained in this way allow us to decipher the details of the existence of metabolic pathways for organisms and create a gene inventory. Environmental DNA or mixed DNA samples help us understand the genetic microheterogeneity of bell groups [4].

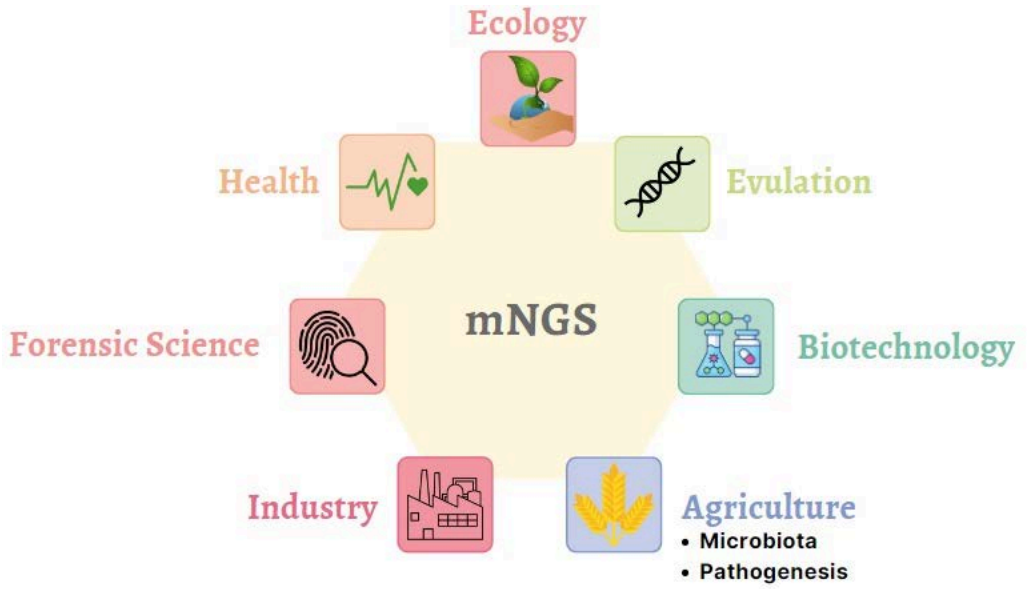


Figure 1. Applications of mNGS technology in different fields

According to the literature review, it is seen that metabarcoding or metagenomic sciences have been widely used in health sciences until now [5]. Especially by analyzing body fluids [6], detection of pulmonary infection in lung tissues [7,8], microbial organisms underlying chronic meningitis, determination of organisms causing tuberculous meningitis in cerebrospinal fluid [9] and even identifying pathogens responsible for uncultured prosthetic joint infection [10] have become a trend of choice. Most of the studies used viral, bacterial, and fungal kits. In addition to all infection detections, scientists aimed to map human-associated microbial communities, such as the gut, mouth, skin, and vagina, as part of the Human Microbiome project [11,12]. mNGS technology is also used in forensic sciences.

 doi:10.20944/preprints202208.0457.v1

mNGS technologies in the agricultural and industrial fields have led to important discoveries. New generation sequencing studies, primarily available in plant roots, are increasingly preferred as they enable the discovery of important secondary metabolites, enzymes, and metabolites [13]. With the influence of industrial applications of the metagenomic approach, the discovery of stress-sensitive bioactive compounds reveals the genetic information of organisms living in extreme conditions. This discovery is used for efficient crop production and elucidation of plant stress mechanisms.

Agronomically, the scope of mNGS technologies is expanding day by day. Researchers think microbial diversity data is essential for sustainable black pepper production [14]. The organisms that make up the plant microbiota provide the necessary nutrients for the growth and development of the plant. Therefore, it is vital for the sustainability of agriculture. Moreover, using metagenomic data to detect and control biotic stress factors affecting crop yield offers optimistic promises for the future. For example, the metagenomic method with 16 S rRNA barcodes was applied to samples obtained from black pepper roots grown in Vietnam [14].

3. **mNGS Methodology for Detecting Fungal Pathogens in Plants**

3.1. Wet Lab Applications

Obtaining a suitable sample is very important for mNGS technology to be applicable. Especially in detecting vegetative pathogens, it is necessary to use plants that are still alive but well infected. When taking the sample, the plant should be preferred where the infection symptom is most evident. The conditions of infection of vegetative pathogens may differ according to the experimental design. Suppose it is aimed to determine the infection on plants in an uncontrolled area, which is expected for metagenomic sampling. In that case, it should be well differentiated to which microorganism group the symptoms belong to. Although fungal stress is the main biotic stress that cause a considerable decrease in yield, virus and bacterial infections are also common in plants under natural conditions [15]. Trying to identify a fungal pathogen with bacterial barcodes is pointless. The distinction between which abiotic stress causes infection in the plant should be made with observational techniques [16]. Collected samples should be kept in a cold environment, transported to the laboratory environment, and stabilized. Because storing the samples under standard ambient settings for an extended period poses a danger of DNA contamination from other organisms. This may compromise the sensitivity of metagenomic analysis and lead to misinterpretation of the data [17].

Nucleic acid extraction is the initial step of mNGS analysis. Extraction can be done using either commercial kits or standard manual procedures, but the former is recommended to rule out the possibility of environmental contamination. Extraction experiments should be performed in an aseptic environment. Since the extracted nucleic acids will comprise DNA from multiple species, they are referred to as mix-DNA, and if they are collected from environmental samples, they are known as environmental DNA or eDNA [18]. Sometimes traditional culturing method can be used to confirm vegetative infection. This will ensure that the dominant pathogen in the plant is reproduced in vitro, and it will be possible to determine whether the plant is indeed an organism-borne infection. It can be considered as a control mechanism for metagenome sequencing. However, this is optional. DNA isolation can also be performed directly from the infected leaf using appropriate kits**.**

3.2. Preparation of Library

The purpose of not preparing a library for mNGS is to make the resulting nucleic acid mixture compatible with sequence analysis. While preserving the diversity of DNA sequences in microbiota analysis, it is necessary to protect or enrich the sequences in pathogenicity studies. Therefore, the library preparation is a complex process. In some metagenomic analyses, the entire nucleic acid obtained can be sequenced, or strategic barcodes of a particular microorganism population can be used. That is because even the most efficient DNA sequencing technologies can sequence only a small fraction of DNA and RNA [19]. Therefore, the prepared library should be representative of the original sample.

 doi:10.20944/preprints202208.0457.v1

In investigations designed to detect pathogenicity or microbiota on the plant, it is anticipated that most nucleic acid extracted will be from the plant. However, using the necessary purification kits, the DNA of the pathogen or microbiota can be separated from the plant's DNA. mNGS libraries can be constructed using minimal amounts of obtained microbial nucleic acid. Microbial enrichment techniques can be used for both DNA and RNA. For the determination of pathogenicity a comprehensive DNA library is created [20]. Pathogenic fungal, bacterial, and viral fragments can be amplified by PCR amplification to increase the nucleic acid content of existing pathogenicity.

When nucleic acid samples are ready for sequencing, sample barcodes and sequencing adapters are added. Barcoding technology involves using short strings of specific markers (Barcodes) added to the end of the sample booklet [19]. This allows multiple samples to be used together for sequencing and to generate sample ID for each sequence read. It is determined by bioinformatic analysis. Library preparation kits such as the high-tech Nextera XT (Illumina, San Diego) are sensitive enough to work with one ng of DNA.

3.3. Sequencing

Various high-throughput platforms are used for the sequencing of mNGS samples. The most used methods in metagenomic studies are Illumina sequencing, Nanopore sequencing, and Roche/454 pyrosequencing. Ilumina sequencing can provide more sensitive and unique results than others, with a read depth of 1 to 5 million at 75 to 100 base pair alignments. Specific 16S rRNA barcodes are used to detect bacterial infections, while barcodes used for the ITS-23S rRNA region are used to detect viral organisms [21]. Some studies may require the use of both barcode types together. In studies where the plant species is unknown, barcodes explicitly defined for the plant can be included in the study by using a method called metabarcoding. The most preferred universal plant barcodes are rbcL, trnL-trnF, rpc36-8, trnT2-rps4, and 2 mitochondrial genes nad7 and atpA [22]. 16S rRNA for detection of bacterial organisms, barcodes of ITS, and 18 S rRNA genes for fungi and archaea are preferred.

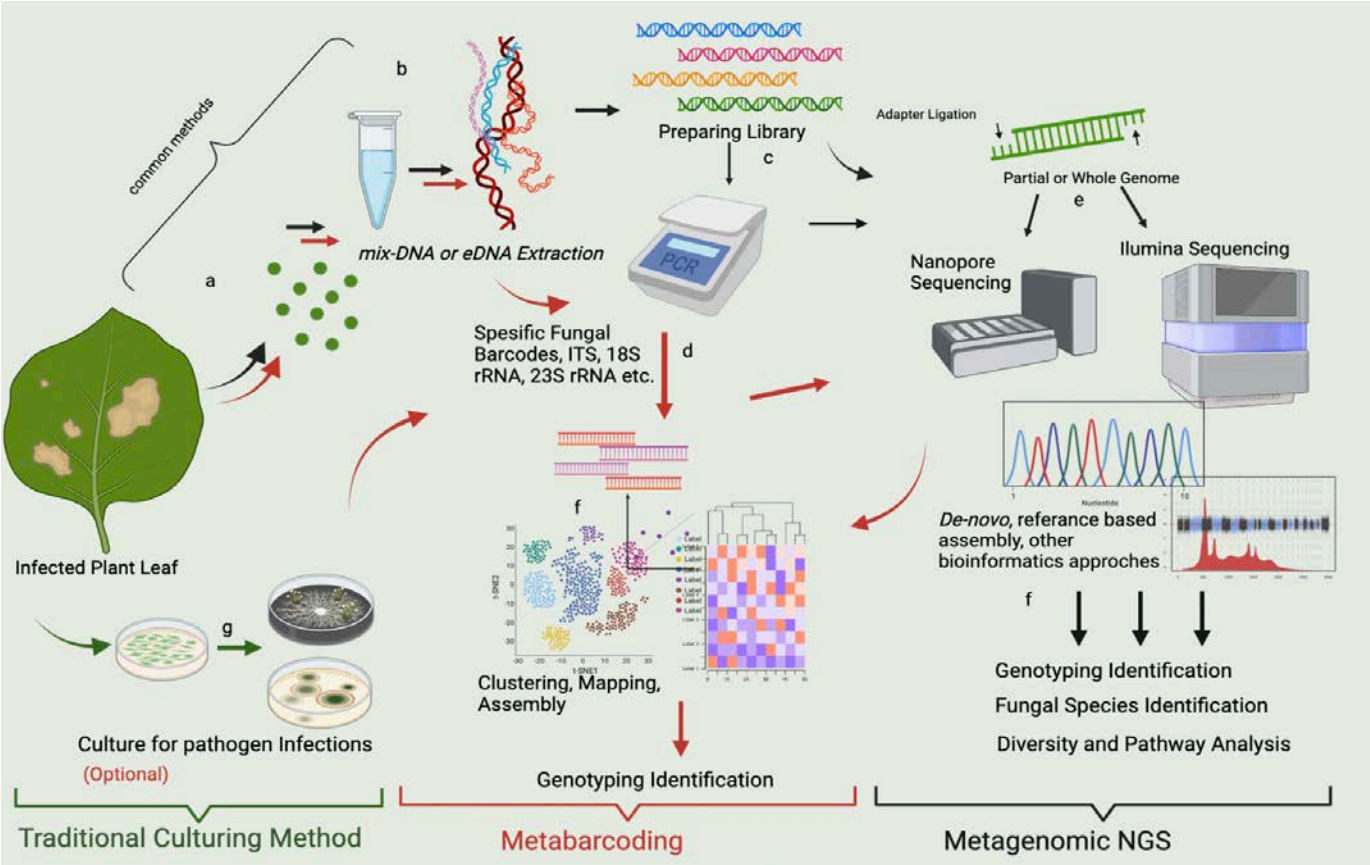
3.4. Bioinformatics Data Analysis

After the metagenomic next-generation sequencing process, a series of bioinformatic analyses is required to analyze the data. The hundreds of short reads obtained in the sequencing must first be filtered. The aim is to extract poor-quality sequences and host genome data. To extract short sequence reads, including the plant genome, a comparison with a reference genome is used to extract matched reads [23]. After filtering, the remaining sequences are compared with reference microbial sequence databases. NCBI is the most preferred database. Because it is possible to reach genomic data of many organisms to be detected. Large sequence reads are combined de novo in clusters each called a contig, which is derived from the word “contiguous”. A contig, in geneomic sequencing, is defined a set of DNA sequences that overlap and provides a contiguous representation of a genomic region enabling links to physical maps. The aim is to assign as many groups as possible to every possible taxonomic group (species, genus phylum). Reads that do not match any sequence are combined de novo with unique algorithms developed for metagenomics (Table 1). *De novo* joins can be done with Meta velvet and Meta-IDBA software [24,25].

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 26 August 2022 doi:10.20944/preprints202208.0457.v1

Table1. Algorithms tools (Bioinformatics Analysis) employed post mNGS process

Purpose	Algorithm tools	References
OTU Clustering	MOTHUR, SUMACLUST, SWARM, METACLUSTER, UCLUST, CD-HIT-OUT, TBC	[26, 27]
Phylogenetic Classifications	Phymm, BLAST, CARMA	[28]
Denoising	Pyronoise, Denoiser, DADA, Acacia	[29, 30]
Chimera Detection	UCHİME, ChimeraSlayer, Persus, DECIPHER	[30, 31]
ITS Database for Fungal Detection	UNITE	[32]
All in one	MOTHUR, QIIME, MEGAN	[30, 33, 34]



 doi:10.20944/preprints202208.0457.v1

Figure 2. mNGS and metabarcoding workflow chart for The sample obtained from the infected leaf. The workflow designated in "red" shows metabarcoding pathways, which use specific metabarcodes for fungal detection, and the "black" one shows that mNGS pathways. In the workflows, the PCR stage is optional. After sequencing and bioinformatic analysis, metabarcoding gives genotyping identification, whereas mNGS gives fungal species identification, microbial diversity, and pathway detection besides genotyping identification. Though both techniques seem to include the same methods, the bioinformatics analysis algorithms differ. In metabarcoding, certain parts of the genome are sequenced using target-specific barcodes. In mNGS, either partial or whole genome is sequenced by reference-based comparison with the prepared library. Both approaches provide a fundamental approach and solution for metagenomics. The workflow designated in "green" represents the traditional culturing method at the researcher's discretion. It may allow culturing of some of the possible microorganisms prior to mNGS and metabarcoding. However, this gives an assignment far below sufficient for mNGS and metabarcoding. The stages represented in the figure can be summarized as follows; a) Sampling from the infected parts of the plant, leaf discs are preferred, b) mixed DNA extraction from leaf samples, c) library preparation, d) PCR for amplification of gene regions of microbial pathogens with specific gene barcodes, e) sequencing with Illumina, Nanopore, etc. f) bioinformatic analysis for mNGS contain *de novo* approaches and referenced based assembly, bioinformatic analysis for metabarcoding assembly, clustering, prediction and g) control culture of infected leaves.

4. Successful Applications of mNGS to Detect Fungal Plant Pathogens

mNGS technology holds promise for pathogen detection in plants. It is used for definitive diagnosis, primarily since it provides sequencing of all nucleic acids found in a sample taken from infected tissue, regardless of traditional culture methods. Since the barcodes specific to the disease agent are not used, there is no need for pre-sequencing information of the infecting organism. However, it is recommended to prefer a database with genome information of fungal pathogens while analyzing the results. Because databases containing genomic data of plant pathogenic fungi are rare.

One of the primary studies on determining fungal pathogenesis in plants is the use of metagenomic analyses by Yang (2022) and colleagues for the detection and identification of *Calonectria pseudonaviculata*, which causes boxwood blight in plants [35]. Boxwood blight is a fungal disease that causes significant economic losses for ornamental plants. For this reason, readings are done with a combination of barcodes used for bacterial and fungal infection detection in existing studies. According to data obtained using different DNA isolation protocols and different bioinformatics algorithms, more than 9% of the reads performed in high-infection plant tissue were identified as *C. pseudonaviculata* [35]. This study, which shows how metagenomics can be applied to plant pathogens, is promising for future fungal pathogen studies. Because the fungal infection is the leading biotic stress affecting yield and product quality in agricultural areas. Unfortunately, approaches to revealing the plant-fungus interaction at the molecular level are progressing slowly. This is because, as mentioned before, plant pathogenic fungi are less well defined than bacterial and viral infections in databases containing genomic data.

One of the pioneering studies in fungal pathogenicity determination studies is the study aiming to determine the microbiome of plants infected by *Zymoseptoria tritici* fungus, one of the common diseases of wheat plants [36]. The result obtained in this study using 450 leaf samples shows significant differences between healthy tissue and infected tissue microbiota. However, the microbiomes of infected leaves collected from different cultivars show very high similarities. These data indicate that it may help prevent infection by *Zymosepttoria tritici* and improve wheat health.

Although sequencing analyses using mNGS technology have become widespread, many approaches using metagenomic techniques have determined fungal microorganisms. Some of these studies are given in Table 2.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 26 August 2022 doi:10.20944/preprints202208.0457.v1

Table 2 The Successful Applications Of Metagenomic Techniques To Diagnose Fungal Pathogens.

Plant	The aim of study	Metagenomics Tecniques	References
Grape	Determination of fungi and oomycetes in different phyllosphere samples	Metabarcoding	[37]
Grape	Determination of soil and leaf-associated fungal microbiota	mNGS-Ilumina	[38]
Wheat	Detection of fungal microorganisms in the wheat phyllosphere	Microbiome Metabarcoding using ITS barcodes	[39]
Grape	Identification of fungal diseases on the vine trunk	mNGS-Ilumina	[40]
Maize	Determination of fungal microbiota after harvest	Metabarcoding	[41]
Wheat	Determination of fungal communities in wheat residues	Metabarcoding	[42]
Grapevine	Determination of fungal disease agents associated with grapevine	Metabarcoding	[43]
Banana	Investigation of the effect of variable soil microbiota on fusarium disease	Metabarcoding	[44]
Wheat, maize	To determine fusarium species in various plants	PaCBio SMRT Sequencing	[45]
Strawberry	Determination of microbial communities in strawberry growing soils with different yields	Amplicon Based Metagenomic	[46]

5. Future Approaches

The use of mNGS technology to determine fungal pathogenicity and microbial relationships of plants is promising for the future. Uncovering the Plant-Microbiota interaction will, in turn, enable the discovery of new genomic data and new industrially important biological materials. Moreover, its dissemination in agronomic sciences will enable the development of methods to combat biotic stress in food-related problems that threaten the future. Detection and identification of infectious agents in the plant's phyllosphere region are essential in increasing agricultural and crop yields.

Although the metagenomic technology applied today requires a high cost, it has the potential to reduce the cost thanks to the increasing demand and developing technology. As the cost decreases, mNGS technology will become more widespread.

Knowing the fungal pathogen causing plant infection beforehand creates a limiting effect for metagenomics. However, some infections are necessary for diagnosing and identifying organisms that cannot be cultured by conventional methods. In addition, the obtained mixed DNA samples also enable the identification of new genes. In particular, the discovery of stress-tolerant genes can be used in crop yield and agricultural improvement processes.

Funding: This research received no external funding.

Acknowledgments: FSG and AS acknowledge Professor Brad Day, Michigan State University Plant, Soil and Microbial Science, for his support. AS acknowledge Founder President, Amity University for his kind support. And also, all authors thank Molecular Biologist Begüm Coşar for her support in drawing.

Conflicts of Interest: There is no conflict of interest between the authors.

References

1. Woese, C. R. The problem of evolving a genetic code. 1970 *Bioscience*, *20*(8), 471-485.
2. Walters, W. A., & Knight, R. Technology and techniques for microbial ecology via DNA sequencing. 2014. Annals of the American Thoracic Society, 11(Supplement 1), S16-S20.Author 1, A.; Author 2, B. *Book Title*, 3rd ed.; Publisher: Publisher Location, Country, 2008; pp. 154–196.
3. ller, R. R., Montoya, V., Gardy, J. L., Patrick, D. M., & Tang, P. Metagenomics for pathogen detection in public health. 2013. *Genome medicine*, *5*(9), 1-14.
4. Hugenholtz, P., & Tyson, G. W. Metagenomics. 2008. *Nature*, *455*(7212), 481-483.
5. Gu, W., Miller, S., & Chiu, C. Y. Clinical metagenomic next-generation sequencing for pathogen detection. 2019. *Annual review of pathology*,
6. Gu, W., Deng, X., Lee, M., Sucu, Y. D., Arevalo, S., Stryke, D., ... & Chiu, C. Y. Rapid pathogen detection by metagenomic next-generation sequencing of infected body fluids. 2021. *Nature Medicine*, *27*(1), 115-124.
7. Li, H., Gao, H., Meng, H., Wang, Q., Li, S., Chen, H., ... & Wang, H. Detection of pulmonary infectious pathogens from lung biopsy tissues by metagenomic next-generation sequencing. 2018. *Frontiers in cellular and infection microbiology*, *8*, 205.
8. 8. Qian, Y. Y., Wang, H. Y., Zhou, Y., Zhang, H. C., Zhu, Y. M., Zhou, X., ... & Ai, J. W. Improving pulmonary infection diagnosis with metagenomic next generation sequencing. 2021. *Frontiers in cellular and infection microbiology*, *10*, 567615.
9. 9. Wang, S., Chen, Y., Wang, D., Wu, Y., Zhao, D., Zhang, J., ... & Hu, Y. The feasibility of metagenomic next-generation sequencing to identify pathogens causing tuberculous meningitis in cerebrospinal fluid. 2019. *Frontiers in microbiology*, *10*, 1993.
10. Wang, C., Huang, Z., Li, W., Fang, X., & Zhang, W. Can metagenomic next-generation sequencing identify the pathogens responsible for culture-negative prosthetic joint infection? 2020. *BMC infectious diseases*, *20*(1), 1-7.
11. Turnbaugh, P. J., Ley, R. E., Hamady, M., Fraser-Liggett, C. M., Knight, R., & Gordon, J. I. The human microbiome project. (2007). *Nature*, *449*(7164), 804-810.
12. Gevers, D., Knight, R., Petrosino, J. F., Huang, K., McGuire, A. L., Birren, B. W., ... & Huttenhower, C. The Human Microbiome Project: a community resource for the healthy human microbiome. 2012.
13. Lorenz, P., & Eck, J. Metagenomics and industrial applications. 2005. *Nature Reviews Microbiology*, *3*(6), 510-516.
14. Tran, D. M., Nguyen, T. H., Huynh, T. U., Do, T. O., Nguyen, Q. V., & Nguyen, A. D. Analysis of endophytic microbiome dataset from roots of black pepper (Piper nigrum L.) cultivated in the Central Highlands region, Vietnam using 16S rRNA gene metagenomic next-generation sequencing. 2022. *Data in Brief*, *42*, 108108.
15. Hussain, F., & Usman, F. Fungal biotic stresses in plants and its control strategy 2019. *Abiotic and biotic stress in plants*.
16. Burdon, J. J., & Zhan, J. Climate change and disease in plant communities. 2020. *PLoS biology*, *18*(11), e3000949.
17. Mzava, O., Cheng, A. P., Chang, A., Smalling, S., Djomnang, L. A. K., Lenz, J. S., ... & De Vlaminck, I. A metagenomic DNA sequencing assay that is robust against environmental DNA contamination. 2022. *Nature Communications*, *13*(1), 1-10.
18. Ruppert, K. M., Kline, R. J., & Rahman, M. S. Past, present, and future perspectives of environmental DNA (eDNA) metabarcoding: A systematic review in methods, monitoring, and applications of global eDNA. 2019. *Global Ecology and Conservation*, *17*, e00547.
19. Chopra, R. S., Chopra, C., & Sharma, N. R. (Eds.). *Metagenomics: techniques, applications, challenges and opportunities*. 2020. Springer Singapore.
20. Daniel, R. The metagenomics of soil. 2005. *Nature Reviews Microbiology*, *3*(6), 470-478.
21. Cuscó, A., Catozzi, C., Viñes, J., Sanchez, A., & Francino, O. Microbiota profiling with long amplicons using Nanopore sequencing: full-length 16S rRNA gene and the 16S-ITS-23S of the rrn operon. 2018). *F1000Research*,
22. Ward, J., Gilmore, S. R., Robertson, J., & Peakall, R. A grass molecular identification system for forensic botany: a critical evaluation of the strengths and limitations. 2009. *Journal of Forensic Sciences*, *54*(6), 1254-1260.

23. Miller, S., & Chiu, C. The role of metagenomics and next-generation sequencing in infectious disease diagnosis. 2022. *Clinical Chemistry*, *68*(1), 115-124.
24. Peng, Y., Leung, H. C., Yiu, S. M., & Chin, F. Y. Meta-IDBA: a de Novo assembler for metagenomic data. 2011. *Bioinformatics*, *27*(13), i94-i101.
25. Bakshi, A., Moin, M., & Madhav, M. S. Metagenomics in Agriculture: State-of-the-Art. In *Metagenomics: Techniques, Applications, Challenges and Opportunities*, 2020. (pp. 167-187). Springer, Singapore.
26. Fu, L., Niu, B., Zhu, Z., Wu, S., & Li, W. CD-HIT: accelerated for clustering the next-generation sequencing data. 2012. *Bioinformatics*, *28*(23), 3150-3152.
27. Bhat, A. H., Prabhu, P., & Balakrishnan, K. A critical analysis of state-of-the-art metagenomics OTU clustering algorithms. 2019 *Journal of Biosciences*, 44(6), 1-9.
28. Brady, A., & Salzberg, S. L. Phymm and PhymmBL: metagenomic phylogenetic classification with interpolated Markov models. 2009. *Nature methods*, *6*(9), 673-676.
29. Rosen, M. J., Callahan, B. J., Fisher, D. S., & Holmes, S. P. Denoising PCR-amplified metagenome data. 2012). *BMC Bioinformatics*, *13*(1), 1-16
30. Kim, M., Lee, K. H., Yoon, S. W., Kim, B. S., Chun, J., & Yi, H. Analytical tools and databases for metagenomics in the next-generation sequencing era. 2013. *Genomics & informatics*, *11*(3), 102
31. Edgar, R. C., Haas, B. J., Clemente, J. C., Quince, C., & Knight, R. UCHIME improves sensitivity and speed of chimera detection. 2011. *Bioinformatics*, *27*(16), 2194-2200.
32. Abarenkov, K., Nilsson, R. H., Larsson, K. H., Alexander, I. J., Eberhardt, U., Erland, S., ... & Kõljalg, U. The UNITE database for molecular identification of fungi–recent updates and future perspectives. 2010. *The New Phytologist*, *186*(2), 281-285.
33. Schloss, P. D., Westcott, S. L., Ryabin, T., Hall, J. R., Hartmann, M., Hollister, E. B., ... & Weber, C. F. Introducing mothur: open-source, platform-independent, community-supported software for describing and comparing microbial communities. 2009. *Applied and environmental microbiology*, *75*(23), 7537-7541.
34. Huse, S. M., Welch, D. M., Morrison, H. G., & Sogin, M. L. Ironing out the wrinkles in the rare biosphere through improved OTU clustering. 2010. *Environmental Microbiology*, *12*(7), 1889-1898.
35. Yang, S., Johnson, M. A., Hansen, M. A., Bush, E., Li, S., & Vinatzer, B. A. Metagenomic sequencing for detection and identification of the boxwood blight pathogen *Calonectria pseudonaviculata*. 2022. *Scientific reports*, *12*(1), 1-14.
36. Barroso-Bergadà, D., Massot, M., Vignolles, N., Faivre d'Arcier, J., Chancerel, E., Guichoux, E., ... & Suffert, F. Metagenomic next- generation sequencing (mNGS) data reveals the phyllosphere microbiome of wheat plants infected by the fungal pathogen Zymoseptoria tritici. 2022. *Phytobiomes Journal*, (ja).
37. Behrens, F. H., & Fischer, M. Evaluation of different phyllosphere sample types for parallel metabarcoding of Fungi and Oomycetes in Vitis vinifera. 2022. Phytobiomes Journal, (ja).
38. Cureau, N., Threlfall, R., Savin, M., Marasini, D., Lavefve, L., & Carbonero, F. Year, location, and variety impact on grape-, soil-, and leaf-associated fungal microbiota of Arkansas-grown table grapes. 2021. Microbial ecology, 82(1), 73-86.
39. Schiro, G., Colangeli, P., & Müller, M. E. A metabarcoding analysis of the mycobiome of wheat ears across a topographically heterogeneous field. 2019. *Frontiers in microbiology*, *10*, 2095.
40. Del Frari, G., Gobbi, A., Aggerbeck, M. R., Oliveira, H., Hansen, L. H., & Ferreira, R. B. Characterization of the wood mycobiome of Vitis vinifera in a vineyard affected by esca. Spatial distribution of fungal communities and their putative relation with leaf symptoms. 2019. *Frontiers in plant science*, *10*, 910.
41. Cobo-Díaz, J. F., Baroncelli, R., Le Floch, G., & Picot, A. Combined metabarcoding and co-occurrence network analysis to profile the bacterial, fungal and Fusarium communities and their interactions in maize stalks. 2019. *Frontiers in Microbiology*, *10*, 261
42. Kerdraon, L., Barret, M., Laval, V., & Suffert, F. Differential dynamics of microbial community networks help identify microorganisms interacting with residue-borne pathogens: the case of *Zymoseptoria tritici* in wheat. 2019. Microbiome, 7(1), 1-17.
43. Morales-Cruz, A., Figueroa-Balderas, R., García, J. F., Tran, E., Rolshausen, P. E., Baumgartner, K., & Cantu, D. Profiling grapevine trunk pathogens in planta: a case for community-targeted DNA metabarcoding. 2018. BMC microbiology, 18(1), 1-14.
44. Shen, Z., Penton, C. R., Lv, N., Xue, C., Yuan, X., Ruan, Y., ... & Shen, Q. Banana Fusarium wilt disease incidence is influenced by shifts of soil microbial communities under different monoculture spans. 2018. Microbial ecology, 75(3), 739-750.
45. Walder, F., Schlaeppi, K., Wittwer, R., Held, A. Y., Vogelgsang, S., & van der Heijden, M. G. Community profiling of Fusarium in combination with other plant-associated fungi in different crop species using SMRT sequencing. 2017 *Frontiers in Plant Science*, *8*, 2019.
46. Xu, X., Passey, T., Wei, F., Saville, R., & Harrison, R. J. Amplicon-based metagenomics identified candidate organisms in soils that caused yield decline in strawberry. 2015 . *Horticulture research*, *2*.
47. Yadav, S., Gettu, N., Swain, B., Kumari, K., Ojha, N., & Gunthe, S. S. (2020). Bioaerosol impact on crop health over India due to emerging fungal diseases (EFDs): an important missing link. *Environmental Science and Pollution Research*, *27*(11), 12802-12829.