



Next Generation Sequencing Pathogen Detection

(*Pulkit Mittal¹ and Dr. Bapulal Roat²)

¹M.Sc. Research Scholar, Department of Plant Pathology, MPUAT, Udaipur

²Professor, KVK, faloj (Dungarpur), MPUAT, Udaipur

*Corresponding Author's email: pulkitmittal824@gmail.com

Next-generation sequencing (NGS) is a method for detecting plant pathogens that uses DNA or RNA sequencing to identify plant viruses, viroids and small RNAs. NGS is a fast and efficient way to sequence millions of DNA sequences at once, which can help identify pathogens that cause plant disease. High-throughput sequencing (HTS) is developing quickly and steadily, and its costs have drastically decreased. This has led to new and unforeseen advances in the field of plant pathology. Whole-genome sequencing, which was too expensive for many projects until recently, is now so inexpensive that phylogenomics, a new field is being established. Genome comparison is also having a significant impact on fungal taxonomy. Finding new genes that could be targets for a precise diagnosis of novel or emerging pathogens particularly those that are quarantine-related has become simpler. Similar to the advancement of metagenomics and metabarcoding techniques has made it feasible to roughly identify fungus, bacteria, nematodes and other microorganisms in complex diseases or to provide answers to important queries like "What's in my soil?" The new technologies make it possible to rethink disease control techniques by taking into account the pathogens in their surroundings and figuring out the intricate relationships between microorganisms and the crops that are grown.

High-Throughput DNA Sequencing (HTS) Technologies

Multiple infections can be detected in a single sample thanks to NGS's ability to produce enormous volumes of data in a single run. When plants are infected with several pathogens at once or when pathogens exist in low abundance, this is particularly valuable. High-throughput platforms for analysis of fungal communities currently include the second and third generation of HTS technologies.

Second-generation HTS methods—Short reads sequencing

The development of second-generation high-throughput sequencing (HTS) techniques in the 2000s revolutionized molecular biology science. DNA is sequenced by ligation or synthesis in these sequencing platforms, which can only handle short reads (less than 550 bases). A polymerase is employed in the synthesis method, which is employed by the majority of platforms. A signal, such as a fluorophore or a shift in ionic concentration, indicates that a nucleotide has been incorporated into an elongating strand. Clonal template populations are created when DNA is blocked on a solid surface (solid-state, bead-based, and DNA nanoball production).

Third-generation HTS methods—Long-reads sequencing

In the 2010s, long-read platforms, such as PacBio single-molecule real-time (SMRT) sequencing (Pacific BioSciences Inc., California, USA, com) and nanopore sequencing (ONT, Oxford Nanopore Technologies Inc., Oxford, UK.) Were introduced. PacBio's SMRT sequencing method is now the most popular long-read platform. The device makes use of zero-mode waveguides (ZMW), a specialized flow cell with hundreds of distinct picoliter

wells. With polymerase attached to the bottom of the ZMW, which synthesizes a single DNA molecule, the sequencing process is dependent on DNA synthesis. The SMRT platform created a circular consensus sequence (CCS), which is a distinct circular template that is continually sequenced by the polymerase, because this approach had poor sequencing depth (only hundreds rather than thousands of reads per sample).

Metabarcoding

A cost-effective technique for characterizing microbial communities is metabarcoding, which enables comparison of sample communities under various treatments, deep taxonomic resolution, and biodiversity assessment. It is simpler to handle than shotgun metagenomics from a bioinformatic perspective (less computing power and storage space). It has grown into the most popular molecular method for describing the microbiota in environmental samples. It provides fresh insights into the investigation of complicated etiology plant diseases, both in the soil and in the air. There are numerous instances of disease complexes that damage both the aboveground and subsurface tissues of agricultural crops and are of concern: Numerous *Fusarium* species, the majority of which produce mycotoxin, are implicated in the dangerous wheat disease known as *Fusarium* Head Blight (FHB). The wheat ear fungus community was defined using Illumina MiSeq with V3 Chemistry in a topographically diverse setting.

Metagenomics and Metatranscriptomics

Shotgun metagenomic research has been crucial in analyzing the taxonomic and functional characteristics of microbial communities during the last 20 years. Plant disease diagnosis is increasingly using NGS-based metagenomics technologies, which were initially employed for pathogen identification in clinical settings. Shotgun metagenomics is a legitimate method for identifying pathogens and, consequently, for making an accurate diagnosis since it enables the sequencing of the complete genome of the microorganisms present in a sample, including soil and other environmental matrices, as well as the symptomatic or asymptomatic host plant.

Phylogenomics

Phylogenomics is an interdisciplinary field that combines phylogenetics (the study of evolutionary relationships among species) with genomics (the study of the complete set of genes and their functions). This approach uses whole genome data to infer evolutionary histories, trace the origins of species or pathogens, and understand the genetic basis of adaptation and diversity across different organisms. The main focus of phylogenomics is to build more accurate, robust, and comprehensive evolutionary trees (phylogenies) using whole-genome data instead of just a few genetic markers. This method provides more detailed insights into evolutionary processes and helps resolve phylogenetic relationships that are difficult to determine using traditional, marker-based approaches.

Multilocus Analysis in Plant pathology: From Traditional to New-Generation Sequencing Technology

One of the most accurate and enlightening techniques for molecular genotyping is multilocus sequence typing, or MLST, which involves target sequencing of several genomic loci. It has been frequently employed for taxonomic designation and to evaluate the degree of genetic and pathogenic heterogeneity among populations of the same species (e.g., *Colletotrichum* spp.). Apart from ITS barcoding, MLST has emerged as the common technique for genotyping a large number of fungus in molecular epidemiology, pathogenicity, and phylogenetic research. It is a useful technique for characterizing and classifying new species in a variety of domains, including plant pathology (e.g., *Colletotrichum*, *Ilyonectria*, or *Diaporthe*) and fungal species of human significance (e.g., *Candida*, *Aspergillus* and *Pseudallescheria*).

Conclusions

It is now commonly accepted that HTS technologies are crucial for diagnosing filamentous plant pathogens, such as fungi and oomycetes, and for enhancing the control of plant diseases. But before the goal is realized and widely applied, there are still technological and economic factors to take into account. Even though this challenge can be solved by loading additional samples in each cycle, many laboratories are still unable to afford the sequencing cost per sample. This suggests that extra DNA purification procedures are required to prevent unforeseen and undesired sequencing artifacts, such as using a dummy sample and DNA from healthy plant tissues as controls.