J. Appl. Res Plant Sci. Vol. 4(1), 512-517, 2023, https://doi.org/10.38211/joarps.2023.04.01.61

Review Article

OPEN ACCESS

Available on https://www.joarps.org
Journal of Applied Research in Plant Sciences (JOARPS)
ISSN: 2708-3004 (Online), 2708-2997 (Print)

Role of Next Generation Sequencing (NGS) in Plant Disease Management: A Review

Muhammad Saeed[1,2], Zainab Jamil[1], Tayyab Shehzad[3], Syed Zia ul Hasan[4], Riffat Bibi[5], Safia Naureen Malik[5], Hafiz Matee-ur-Rehman[6], Raees Ahmed[6]*

[1]Department of Plant Pathology, PMAS-Arid Agriculture University, Rawalpindi, P akistan.
[2]Wheat Research Sub-Station, Sunny Bank Murree, Pakistan.
[3]Department of Plant Breeding and Genetics, PMAS-Arid Agriculture University, Rawalpindi, Pakistan.
[4]Hill Fruit Research Station Sunny Bank, Murree, Pakistan.
[5]Soil and Water Conservation Research Institute Chakwal, Pakistan.
[6]Department of Plant Pathology, University of Poonch, Rawalakot, AJK, Pakistan.
*Corresponding author's email; raees@upr.edu.pk

Article Received 09-01-2023, Article Revised 14-02-2023, Article Accepted 16-02-2023

Abstract:

A high throughput technique used to determine a part of the nucleotide sequence of an organism's genome is called next generation sequencing (NGS). NGS has been Proven revolutionary in genomics. Clinical diagnostics, Plant diseases diagnostic and other aspects of medical are now made possible by sequencing. Techniques of NGS: there are different techniques of NGS which are being used in real life sciences i.e., Illumina sequencing, Pyrosequencing, Roche 454 sequencing and Ion torrent sequencing. All vintage methods like culturing in bacterial, fungal, and viral samples are being suppressed by next generation sequencing. The potential for random metagenomic sequencing of sick samples to find potential pathogens has surfaced with the development of next-generation high-throughput parallel sequencing technology. NGS enables highly efficient, rapid, low-cost DNA or RNA high-throughput sequencing of plant virus and viroids genomes, as well as specific small RNAs generated during infection. Although this technique is not so much familiar in the field of plant diseases. However, its widespread application in agronomic sciences will make it possible to create solutions to future food-related challenges that involve biotic stress.

Keywords: *NGS, Plant, Pathogen, Diagnosis*

Introduction

A high throughput technique used to determine a part of the nucleotide sequence of an organism's genome is called next generation sequencing (NGS). This methodology is also known as second generation sequencing. This Method uses the sequencing technologies of DNA which are good capable of processing numbers of DNA sequences in parallel. Next generation sequencing also called massively parallel sequencing. Simply an advanced DNA, RNA sequencing technique to sequence big or large genome sample at accurately and quickly as compared to Sanger technique.

Next-generation sequencing techniques utilizing methylation, RNA, or DNA sequencing left a great impact on life science. As compared to Sanger traditional method which is also called first generation sequencing technique, NGS cost lower and time saving techniques with gigabase range of base pairs. The utilization of molecular biology and genetic information by different species to reproduce and live with or without disease, mutations, and diversity within their population networks and changing surroundings is now more or much more studied by NGS than ever before.

The history of DNA sequencing starts from 1965 when Robert Holley was awarded the Nobel Prize as he sequenced first tRNA, in 1986. In his Nobel Prize speech, he said, "without minimizing the pleasure of receiving awards and prizes, I think it is true that the greatest satisfaction for a scientist comes from carrying a major piece of research to a successful conclusion" (Holley, 1968). From that different technologies are being used to sequence whole genome. In 1977, a scientist named Sanger sequenced the first-time whole genome of a virus (Berg, 2014). Human genome was sequenced in 20 years with 3 billion dollars expenditures. Different techniques got evolutional to reach next generation sequencing. In a many of previous reviews (Lam *et al.*, 2012; Mardis, 2008; Margulies *et al.*, 2005) history of next generation was published in detailed.

Techniques of NGS: there are different techniques of NGS which are being used in real life sciences.

Illumina sequencing: Clonal array building and its unique reversible terminator technology are used in Illumina's sequencing approach for efficient and accurate large-scale sequencing. The process identifies DNA bases while also including them in a chain of nucleic

acids. The specific fluorescent signal that each base produces as it is integrated into the growing strand establishes the DNA sequence's order (Quail *et al.*, 2008). Sample-specific barcodes must be incorporated into sequencing libraries in order to parallelize target capture and sequencing for numerous samples while maintaining the ability to identify the origin of each sequence. This technique outlines a quick and dependable process for creating barcoded ("indexed") sequencing libraries for the Genome Analyzer platform from Illumina (Meyer & Kircher, 2010).

A significant number of sequence data may now be produced quickly and for a lot less money than with capillary sequencing because to improvements in DNA sequencing technology. These new technologies have unique traits and constraints that must be taken into account either during project design or throughout data analysis (Figure 1). In both the laboratory and statistical stages of project planning and analysis, specific knowledge is required to produce high quality data from these new platforms. The Illumina sequencers that allow parallel reading of several hundred million immobilized sequences using fluorescent-dye reversible-terminator chemistry include the new HI Scan, Genome Analyzers, and Hi Seq. The effectiveness of the sequencing library, sample handling, instrument settings, and sequencing chemistry all have a big impact on the sequencing process (Kircher, Heyn, & Kelso, 2011).

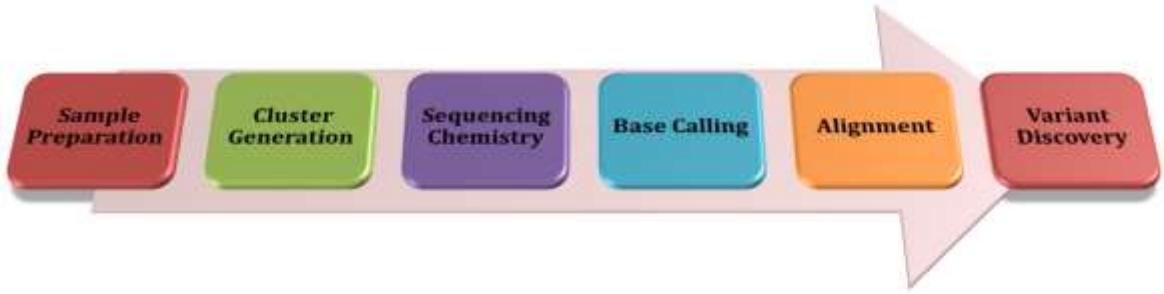

Figure 1. Factors contribution platform accuracy in Illumina sequencing.

Pyrosequencing: Pyrosequencing is a method of genetic material sequencing that is based on the synthesis principle. It works by locating nucleotides that a DNA polymerase has incorporated into the sequence. A chain reaction-based light detection system is also used in pyrosequencing. This system is also based on an enzymatic system. There are many present and future applications of this method (Ahmadian, Ehn, & Hober, 2006).

Roche 454 sequencing: The scientific community now conducts far more sequencing thanks to the 454 Sequencer, and there is a wider range of issues that can be solved using direct readouts of DNA sequence. Higher throughput streamlined in vitro sample preparation procedures, and the miniaturization of sequencing chemistries all of which contributed to the development of the 454-sequencing platform made massively parallel sequencing reactions feasible at a scale and price that were previously unthinkable (Figure 2). The 454 platform has begun to democratize sequencing along with other recently announced next-generation technologies, giving independent laboratories access to capacities that were previously only available at a small number of major sequencing centers (Rothberg & Leamon, 2008).

Ion torrent sequencing: The detection of hydrogen ions generated during DNA polymerization is the foundation of the DNA sequencing technique known as ion semiconductor sequencing or Ion torrent sequencing. A complementary strand is constructed using this "sequencing by synthesis" technique, which bases it on the sequence of a template strand (Merriman, D Team, & Rothberg, 2012).

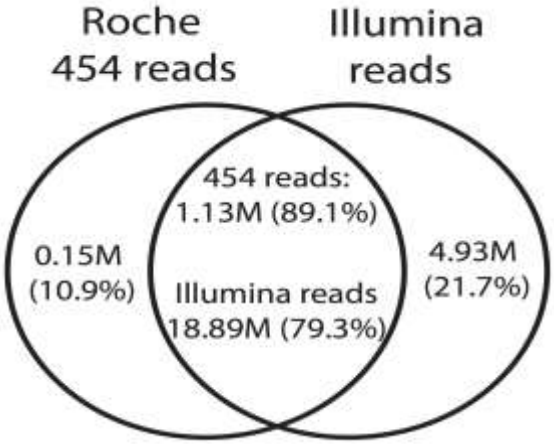

Figure 2. Genetic diversity and abundance of genes in Roche 454 and Illumina (Luo, Tsementzi, Kyrpides, Read, & Konstantinidis, 2012).

NGS as disease diagnostic tool: The new molecular diagnostic tools in the medical field have changed the scenario in detection of different diseases. All vintage methods like culturing in bacterial, fungal, and viral samples are being suppressed by next generation sequencing. Culturing technique is not perfect reliable

method to diagnostic, because we can't study the whole genome but only the specific targeted area. Similarly, clinically useful method quantitative bacterial load analysis is not often used because of heavy expenditures and time lengthen (Lecuit & Eloit, 2015).

Since next generation sequencing is non-targeted identification technique which is being used in whole genome sequencing, data mining and sorting out of microbes without a priori. In this technique we can save hundreds of different primers and time consuming protocols and can easily combat different diseases and hindered infections. In this technique can also identify the continuous altering of sequences in different variant and species. De Vlaminck *et al.* (2013) reported that WG-NGS has a positive correlation between the number of NGS reads and CT values and can also give diagnostic PCRs additional in-depth taxonomic information, such as viral subtypes or serotypes (Figure 3).

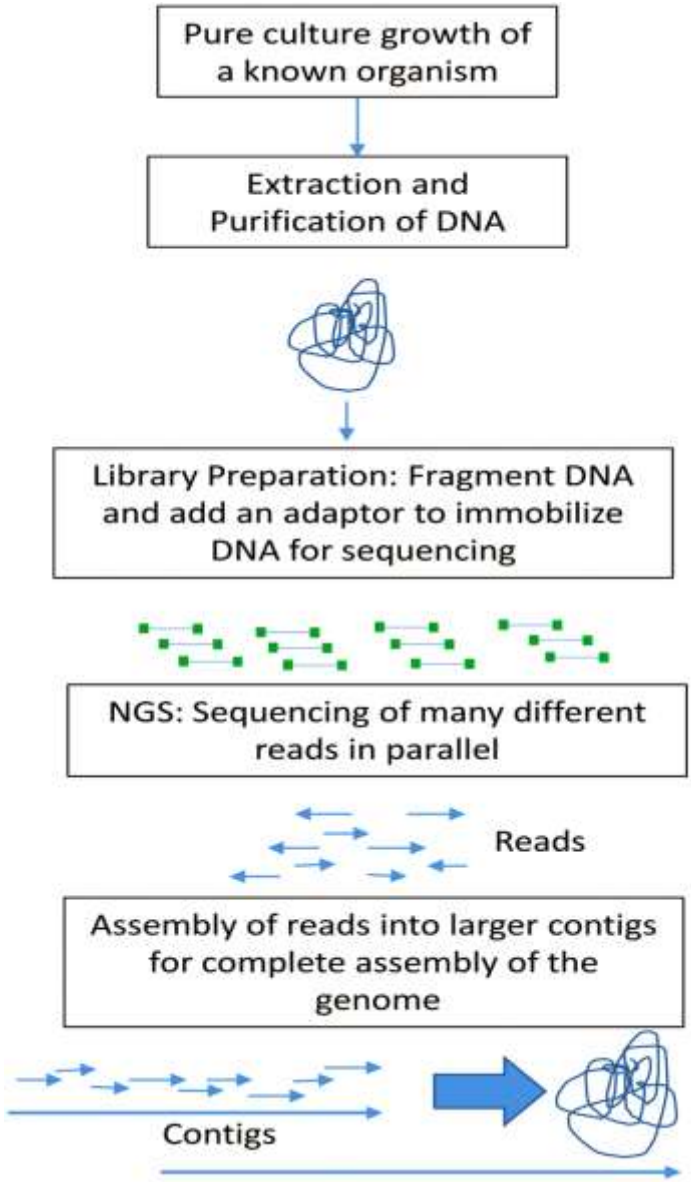

Figure 3. An overview of whole-genome sequencing of a pure organism from cultured growth (Simner, Miller, & Carroll, 2018).

Ion torrent sequencing: The detection of hydrogen ions generated during DNA polymerization is the foundation of the DNA sequencing technique known as ion semiconductor sequencing or Ion torrent sequencing. A complementary strand is constructed using this "sequencing by synthesis" technique, which bases it on the sequence of a template strand (Merriman *et al.*, 2012).

NGS as disease diagnostic tool: The new molecular diagnostic tools in the medical field have changed the scenario in detection of different diseases. All vintage methods like culturing in bacterial, fungal, and viral samples are being suppressed by next generation sequencing. Culturing technique is not perfect reliable method to diagnostic, because we can't study the whole genome but only the specific targeted area. Similarly, clinically useful method quantitative bacterial load analysis is not often used because of heavy expenditures and time lengthen (Lecuit & Eloit, 2015).

Since next generation sequencing is non-targeted identification technique which is being used in whole genome sequencing, data mining and sorting out of microbes without a priori. In this technique we can save hundreds of different primers and time consuming protocols and can easily combat different diseases and hindered infections. In this technique can also identify the continuous altering of sequences in different variant and species. De Vlaminck *et al.* (2013) reported that WG-NGS has a positive correlation between the number of NGS reads and CT values and can also give diagnostic PCRs additional in-depth taxonomic information, such as viral subtypes or serotypes (Figure 3).

A study conducted by Lecuit and Eloit (2015), proved that next generation sequencing can easily sequence the infectious illness whole genome. This uncommon disease has a huge impact on 30 million Europeans 25 to 50 million americans10 million Australians. Rare diseases provide a significant challenge to clinicians because they can be difficult to precisely diagnose and because they significantly raise the cost of healthcare globally.

This article will understand us the way that how next generation sequencing methodologies left a great impact on discoveries of rare infectious diseases and also the impact of NGS on research and technology. Next generation sequencing facilitates the development of potential novel medications that are easy to use and reasonably priced. Genetic abnormalities can now be detected with this technique, the importance of mosaic and de novo mutations have been highlighted, cases of digenic inheritance have been discovered, the vast phenotypic range of the majority of genes has been exposed, and patients with numerous uncommon diseases have been identified. NGS has made all these things feasible. Although there are some back draws of this technique too. But this is really an amazing technique in modern medicine.

NGS role in plant diseases: Crop disease resistance has been improved by using next generation sequencing because these genomic based approaches have the ability to sequence the whole genome at once. Gratitude must be paid to NGS technology, due to which it become simple to sequence the whole genome of plants and their pathogens. By using this technology, metagenomic studies of different microbes and their hosts are now conceivable. Thanks to NGS techniques, as it is now routinely and affordably possible to generate plant and pathogen genome or their transcriptome marker sequences, that are related to virulence phenotypes in the pathogen or resistance phenotypes in the plant. This might lead to better management of plant diseases.

Now a days next generation sequencing tools are being used to detect plant parasitic pathogens i.e., fungi, bacteria, nematodes, and viruses. The potential for random metagenomic sequencing of sick samples to find potential pathogens has surfaced with the development of next-generation high-throughput parallel sequencing technology (Cox-Foster *et al.*, 2007). From tissue samples acquired from three transplant recipients who had passed away after getting transplants from the same donor, Palacios *et al.* (2008) sequenced the cDNA.

NGS technology has a good potential for detecting diseases in plants. It is used for certain diagnoses because it provides sequencing of every nucleic acid found in a sample taken from diseased tissue or cell, regardless of standard culture protocols. Because the disease agent-specific barcodes are not used, the pre-sequencing information of the illness-causing organism is not necessary in this technique. However, it is advised to favor a database having information on the fungal pathogens' genomes when examining the outcomes. Because plant pathogenic fungus genetic data databases are uncommon. One of the primary studies in this regard is that Yang 2022 and colleagues used metagenomic analyses for the detection and identification of *Calonectria pseudonaviculata*, which causes boxwood blight in plants, which is one of the key works on defining fungal pathogenesis in plants (Yang *et al.*, 2022). A fungus called boxwood blight results in severe financial losses for attractive plants. Due to this, readings are performed using a mixture of barcodes that are currently utilized in research to detect bacterial and fungal infections.

NGS has created new opportunities for the study of the microbial diversity associated with either specific plants or the environment in which they thrive, in addition to enabling us to comprehend the range of bacterial plant illnesses. Plant microbiome or phytobiome research is a young area that has great potential for unravelling the complex connections between plants and their phytobiomes. The growing knowledge might potentially be used to construct phytobiomes to lessen the prevalence and severity of disease or by employing components of the phytobiome as biopesticides to control particular infections. A systems framework is suggested by Poudel *et al.* (2016) in their research to clarify microbiome networks and, more significantly, discover potential microorganisms.

Next generation sequencing is also getting itself beef up in the field of metagenomics. The metagenomics is a procedure to evaluate microbial populations in a material

at nucleotide sequence. Different metagenomes like, viral metagenomes (present in human guts) and bacterial metagenomes, (which are present in underground mines, in oceans and in sweet water) are of few examples (Angly *et al.*, 2006; Breitbart *et al.*, 2003; Edwards *et al.*, 2006; Fierer *et al.*, 2007; Sogin *et al.*, 2006; Venter *et al.*, 2004; Williamson *et al.*, 2008; Zhang *et al.*, 2006). This large scale molecular genomic research has become more accurate and time lapsing to next generation sequencing.

Experiments are also being conducted to detect plant viruses using NGS. For the detection of Potato Virus Discovery, Surveillance, and Hunting, Alinda *et al.* (2020) used next generation sequencing. Similarly, another scientist Mwaipopo *et al.* (2021) used the same molecular genomic technique to find out the viruses in wild plants. He used this technique with reverse-transcription PCRs (a traditional technique). Another work was carried out in Tanzania, where leaf samples from more than 1400 wild plants were collected to separate RNA. From these samples they found more than 25 genera of plant viruses, which is only possible by NGS techniques.

Soil microbiome is a recently developed method that uses next-generation sequencing (NGS) technology to give soil microbial community structure in order to provide insights about soil microbial activity and diversity. Prokaryotic and eukaryotic soil bacteria can now be more easily characterized thanks to this sequencing method (Ramirez *et al.*, 2018). Ribosomal gene snippets from shotgun metagenomics are used in this sequencing technology for microbial community characterization (Guo, 2016).

Different scientists have used the next generation sequencing technology in plant nematology too. Different species of root-knot nematodes are being separated easily by this technique. According to Besnard *et al.* (2014) different genomic rearrangements were detected profound as when compared with other species of the Hoplolaimoidea subfamily. In this study repeat elements of different base pairs were discovered. This study illustrate the capability of next-generation sequencing to produce whole mitochondrial genomes, even in the absence of a reference sequence, and suggests new directions for nematode species/race identification, phylogenetics, and population genetics.

Conclusion

Next generation sequencing (NGS) is fast, up to date, time lapsing and precise data recording technique, which is being used in all the fields of life sciences. Although many Phyto doctors are using NGS techniques but, not much work was found as to use NGS techniques in detecting plant pathogens and to address them. There are several reasons behind this, i.e., it is very costly technique still now. However, its widespread application in agronomic sciences will make it possible to create solutions to future food-related challenges that involve biotic stress. To improve agricultural and crop yields, infectious agents must be found and identified in the phyllo-sphere of plants.

References

Ahmadian, A., Ehn, M., & Hober, S. (2006). Pyrosequencing: History, biochemistry and future. *Clinica Chimica Acta*, ***363***(1-2), 83-94.

Alinda, A. K., Okoth, P. K., Onamu, R., David, R., Genevieve, T., & Muoma, J. O. (2020). Next generation sequencing platforms for potato virus hunting, surveillance and discovery. *African Journal of Bacteriology Research*, ***12***(1), 1-11.

Angly, F. E., Felts, B., Breitbart, M., Salamon, P., Edwards, R. A., Carlson, C., . . . Rohwer, F. (2006). The marine viromes of four oceanic regions. *PLoS Biology*, ***4***, e368.

Berg, P. (2014). *Fred Sanger: A memorial tribute*. Paper presented at the Proceedings of the National Academy of Sciences.

Besnard, G., Jühling, F., Chapuis, E., Zedane, L., Lhuillier, E., Mateille, T., & Bellafiore, S. (2014). Fast assembly of the mitochondrial genome of a plant parasitic nematode (*Meloidogyne graminicola*) using next generation sequencing. *Comptes Rendus Biologies*, ***337***(5), 295-301.

Breitbart, M., Hewson, I., Felts, B., Mahaffy, J. M., Nulton, J., Salamon, P., & Rohwer, F. (2003). Metagenomic analyses of an uncultured viral community from human feces. *Journal of Bacteriology*, ***185***, 6220–6223.

Cox-Foster, D. L., Conlan, S., Holmes, E. C., Palacios, G., Evans, J. D., Moran, N. A., . . . Geiser, D. M. (2007). A metagenomic survey of microbes in honey bee colony collapse disorder. *Science*, ***318***(5848), 283-287.

De Vlaminck, I., Khush, K. K., Strehl, C., Kohli, B., Luikart, H., Neff, N. F., . . . Nicolls, M. R. (2013). Temporal response of the human virome to immunosuppression and antiviral therapy. *Cell*, ***155***(5), 1178-1187.

Edwards, R. A., Rodriguez-Brito, B., Wegley, L., Haynes, M., Breitbart, M., Peterson, D. M., . . . Rohwer, F. (2006). Using pyrosequencing to shed light on deep mine microbial ecology. *BMC genomics*, ***7***(1), 1-13.

Fierer, N., Breitbart, M., Nulton, J., Salamon, P., Lozupone, C., Jones, R., . . . Rayhawk, S. (2007). Metagenomic and small-subunit rRNA analyses reveal the genetic diversity of bacteria, archaea, fungi, and viruses in soil. *Applied and Environmental Microbiology*, ***73***(21), 7059-7066.

Guo, J. (2016). *Rhizosphere metagenomics of three biofuel crops*. USA: Michigan State University.

Holley, R. (1968). Alanine transfer RNA, Nobel lecture. In.

Kircher, M., Heyn, P., & Kelso, J. (2011). Addressing challenges in the production and analysis of illumina sequencing data. *BMC genomics*, ***12***, 1-14.

Lam, H. Y., Clark, M. J., Chen, R., Chen, R., Natsoulis, G.,

O'huallachain, M., . . . Gerstein, M. B. (2012). Performance comparison of whole-genome sequencing platforms. *Nature Biotechnology*, ***30***(1), 78-82.

Lecuit, M., & Eloit, M. (2015). The potential of whole genome NGS for infectious disease diagnosis. In (Vol. **15**, pp. 1517-1519): Taylor and Francis.

Luo, C., Tsementzi, D., Kyrpides, N., Read, T., & Konstantinidis, K. T. (2012). Direct comparisons of Illumina vs. Roche 454 sequencing technologies on the same microbial community DNA sample. *PloS one*, ***7***(2), e30087.

Mardis, E. R. (2008). Next-generation DNA sequencing methods. *Annual Review of Genomics and Human Genetics*, ***9***, 387-402. doi:10.1146/annurev.genom.9.081307.164359

Margulies, M., Egholm, M., Altman, W. E., Attiya, S., Bader, J. S., Bemben, L. A., . . . Chen, Z. (2005). Genome sequencing in microfabricated high-density picolitre reactors. *Nature*, ***437***(7057), 376-380.

Merriman, B., D Team, I. T., & Rothberg, J. M. (2012). Progress in ion torrent semiconductor chip based sequencing. *Electrophoresis*, ***33***(23), 3397-3417.

Meyer, M., & Kircher, M. (2010). Illumina sequencing library preparation for highly multiplexed target capture and sequencing. *Cold Spring Harbor Protocols*, ***2010***(6), 5448.

Mwaipopo, B., Rajamäki, M.-L., Ngowi, N., Nchimbi-Msolla, S., Njau, P. J., Valkonen, J. P., & Mbanzibwa, D. R. (2021). Next-generation sequencing-based detection of common bean viruses in wild plants from Tanzania and their mechanical transmission to common bean plants. *Plant Disease*, ***105***(9), 2541-2550.

Palacios, G., Druce, J., Du, L., Tran, T., Birch, C., Briese, T., . . . Marshall, J. (2008). A new arenavirus in a cluster of fatal transplant-associated diseases. *New England Journal of Medicine*, ***358***(10), 991-998.

Poudel, R., Jumpponen, A., Schlatter, D. C., Paulitz, T., Gardener, B. M., Kinkel, L. L., & Garrett, K. (2016). Microbiome networks: a systems framework for identifying candidate microbial assemblages for disease management. *Phytopathology*, ***106***(10), 1083-1096.

Quail, M. A., Kozarewa, I., Smith, F., Scally, A., Stephens, P. J., Durbin, R., . . . Turner, D. J. (2008). A large genome center's improvements to the Illumina sequencing system. *Nature Methods*, ***5***(12), 1005-1010.

Ramirez, K. S., Knight, C. G., de Hollander, M., Brearley, F. Q., Constantinides, B., Cotton, A., . . . de Vries, F. T. (2018). Detecting macroecological patterns in bacterial communities across independent studies of global soils. *Nature microbiology*, ***3***(2), 189-196. doi:10.1038/s41564-017-0062-x

Rothberg, J. M., & Leamon, J. H. (2008). The development and impact of 454 sequencing. *Nature Biotechnology*, ***26***(10), 1117-1124.

Simner, P. J., Miller, S., & Carroll, K. C. (2018). Understanding the promises and hurdles of metagenomic next-generation sequencing as a diagnostic tool for infectious diseases. *Clinical Infectious Diseases*, ***66***(5), 778-788.

Sogin, M. L., Morrison, H. G., Huber, J. A., Welch, D. M., Huse, S. M., Neal, P. R., . . . Herndl, G. J. (2006). Microbial diversity in the deep sea and the underexplored "rare biosphere". *Proceedings of the National Academy of Sciences*, ***103***(32), 12115-12120.

Venter, J. C., Remington, K., Heidelberg, J. F., Halpern, A. L., Rusch, D., Eisen, J. A., . . . Smith, H. O. (2004). Environmental genome shotgun sequencing of the *Sargasso sea*. *Science*, ***304***, 66–74.

Williamson, S. J., Rusch, D. B., Yooseph, S., Halpern, A. L., Heidelberg, K. B., Glass, J. I., . . . Sutton, G. (2008). The Sorcerer II Global Ocean Sampling Expedition: Metagenomic characterization of viruses within aquatic microbial samples. *PloS one*, ***3***(1), e1456.

Yang, S., Johnson, M. A., Hansen, M. A., Bush, E., Li, S., & Vinatzer, B. A. (2022). Metagenomic sequencing for detection and identification of the boxwood blight pathogen Calonectria pseudonaviculata. *Scientific reports*, ***12***(1), 1-14.

Zhang, T., Breitbart, M., Lee, W. H., Run, J.-Q., Wei, C. L., Soh, S. W. L., . . . Ruan, Y. (2006). RNA viral community in human feces: Prevalence of plant pathogenic viruses. *PLoS Biology*, ***4***(1), e3.