

## EXPLORING THE POTENTIAL OF NEXT GENERATION SEQUENCING IN PLANT BREEDING AND GENETICS

Kashif Ali<sup>1</sup>, Hafiza Faiza Ghous<sup>1</sup>, Nawal Shehzadi<sup>1</sup>, Omaira Haroon<sup>1</sup>,  
Samia Rashid<sup>1</sup> and Mehdi Rahimi<sup>2</sup>

<sup>1</sup>Department of Plant Breeding and Genetics, University of Agriculture, Faisalabad, Pakistan

<sup>2</sup>Department of Biotechnology, Institute of Science and High Technology and Environmental Sciences,  
Graduate University of Advanced Technology, Kerman, Iran

\*Corresponding author: [2018ag6734@uaf.edu.pk](mailto:2018ag6734@uaf.edu.pk)

### ABSTRACT

Next generation sequencing (NGS) technologies have had a significant impact on plant breeding and genetics, enabling researchers to rapidly and accurately analyze large amounts of genetic data and identify and characterize important plant traits. NGS technologies, such as the study of gene expression and regulation in plants has made extensive use of RNA and genome sequencing, revealing important details about the mechanisms governing these activities. Plant-pathogen interactions have been studied using NGS, which has also been utilized to find pertinent genes. By identifying the genes and genetic variants that are associated with plant resistance to diseases, researchers can cross different plant varieties that are more disease- and pest-resistant. This can help to reduce crop losses and improve crop yields. NGS has been used to study the genetics of plant populations, enabling researchers to identify genetic variations that are associated with important plant traits and to develop new plant varieties with improved performance or desirable characteristics. Despite some challenges, such as complexity of NGS technologies and the limited availability of reference genomes for many plant species, NGS has had a major impact on plant breeding and genetics and is likely to continue to play a significant part in the future evolution of novel plant varieties.

**Keywords:** NGS, Plant Breeding and Genetics, NGS Technologies, Gene Expression

---

Article History (2022-11135) || Received: 21 Dec 2022 || Revised: 30 Dec 2022 || Accepted: 04 Jan 2023 || Published Online: 06 Jan 2023

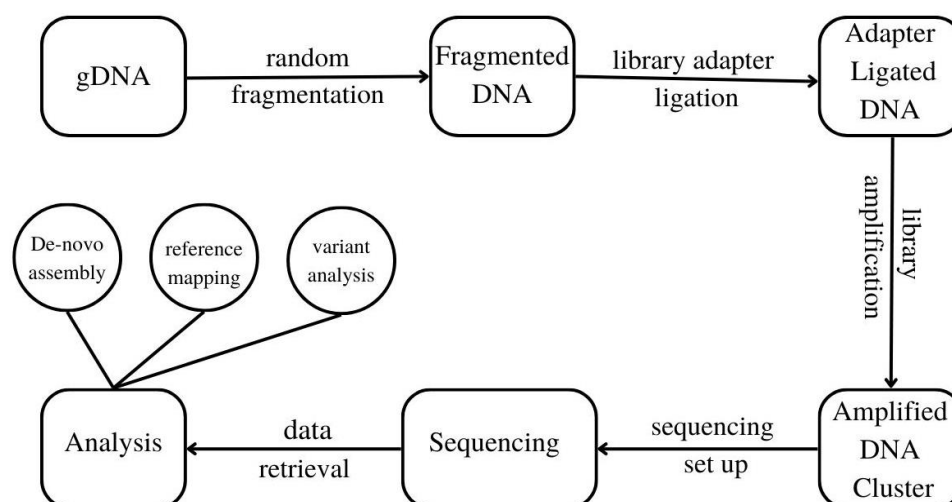
---

## 1. INTRODUCTION

NGS is a powerful tool that has revolutionized many aspects of plant breeding and genetics. It allows researchers to rapidly and accurately analyze large amounts of genetic data (Liu et al. 2012), enabling them to identify and characterize important traits (Salgotra et al. 2014), study gene expression and regulation, and understand the genetic foundation of interactions between plants and pathogens (Campos et al. 2021), and between other plants and microorganisms. NGS allows researchers to generate large amounts of genetic data quickly, cost-effectively, and efficiently, making it possible to study numerous samples in a short period of time. Multiple genomic areas can be analyzed simultaneously using NGS, making it possible to acquire a comprehensive view of plants genetic makeup (Song et al. 2017). It is sensitive enough to detect rare genetic variants that may be missed by traditional techniques. So, it provides high-resolution data, making it possible to detect even small genetic changes or variations (Allard et al. 2012).

### 1.1. Plant Traits Identification and Characterization

NGS technologies have had a significant impact on the identification and characterization of important plant traits, such as those related to growth, development, and response to environmental stresses (Unamba et al. 2015). These technologies have made it much easier and faster to study how plants' gene expression is regulated, as well as the identification and description of plant genes and their roles. This has guided to a better understanding of the genetic basis of important plant traits and has enabled the development of new strategies for improving plant traits. This can be done by comparing the gene expression patterns in plants with different phenotypes (characteristics) and identifying genes that are differentially expressed (expressed at different levels) in two different phenotypes (Vlk and Repkova 2017).



**Fig. 1:** Flowchart diagram of NGS Methodology (Gupta and Gupta 2014).

## 1.2. Gene Expression and Regulation

NGS technologies, such as sequencing by synthesis, sequencing by hybridization, 454 pyrosequencing, Ion torrent, Illumina (Bi et al. 2019). Pacific Biosciences sequencing (Slatko et al. 2018), RNA sequencing (RNA-seq) and genome sequencing, have been widely used to study gene expression and regulation in plants (Camilli et al. 1994; Dmitriev et al. 2022). RNA-seq allows researchers to detect and compute all of the RNA molecules appear in a sample (Yang et al. 2012), including mRNA, which is transcribed from protein-coding genes and can be used as a proxy for gene expression (Muskovic et al. 2022). By comparing the levels of mRNA in different plant tissues or in response to different stimuli, such as drought, heat, or disease, researchers can identify genes (Mwadingeni 2016), that are differentially expressed and study the mechanisms that regulate their expression. Genome sequencing provides a complete map of the DNA sequence of a plant genome (Jiao and Schneeberger 2017; Sun et al. 2022), which can be used to identify the locations of genes and other regulators like silencers, enhancers, and promoters. By studying the sequence and structure of these regulatory elements, researchers can gain insight into the mechanisms that control gene expression in plants.

ChIP-seq is applied to locate the transcription factors' binding locations and other regulatory proteins in the genome (Furey 2012). This provides valuable information about the mechanisms that regulate gene expression and can help to identify new targets for improving plant performance. One such technology is digital gene expression (DGE) profiling, which allows quantitatively measure the expression of all genes in a sample using high-throughput sequencing (Li et al. 2021). This method has been used to pinpoint genes that are differently expressed in various plant tissues or in response to various stimuli, and it can offer a comprehensive perspective of gene expression in plants (Ben Khedher et al. 2022). These technologies can provide valuable insights into the mechanisms that regulate gene expression in plants and can help to identify new targets for improving plant performance (Sahu et al. 2020).

## 1.3. Exploring Genetics of Plant-Pathogen Interactions

Plant-pathogen interactions have a significant impact on the productivity and health of crops and natural ecosystems. By NGS technologies, scientists have been able to delve into the genomic basis of these interactions, uncovering the genes and gene products involved in plant defense responses to pathogens and how pathogens evade or suppress these defenses (Kumari et al. 2017). By comparing the gene expression patterns of plants that are resistant or susceptible to a particular pathogen, researchers can identify key players in plant defense responses and study their functions. RNA-seq is a powerful tool for exploring the genetics of plant-pathogen interactions and identifying genes involved in plant defense responses (Fass et al. 2020). It provides valuable insights into the mechanisms by which plants resist or succumb to diseases.

Furthermore, NGS has been used to study the genetics of plant-pathogen interactions at the molecular level (Mukherjee et al. 2012), by analyzing the sequences of genes and gene products involved in plant defense responses and pathogen evasion or suppression. For example, researchers have used NGS to identify the sequences and analyze the functions of receptor-like kinases (RLKs) in plant defence mechanisms. RLKs are a group of transmembrane receptors that are crucial for plant growth and disease defense (Afzal et al. 2008), and their

sequences and functions have been characterized using NGS technologies. This can provide understanding of the mechanisms that underlie these interactions and can help to identify new targets for improving plant resistance to diseases.

#### 1.4. NGS Challenges and Limitations

There are several current challenges and limitations of using NGS in plant breeding and genetics. One challenge is the complexity of analyzing and interpreting the large amounts of data produced by NGS technologies. This requires specialized bioinformatics tools and expertise (Kulski 2016), which may not be available to all researchers. Another challenge is the limited availability of reference genomes for many plant species (Michael and VanBuren 2015), which can make it difficult to accurately annotate and interpret the sequences generated by NGS technologies. Particularly with regard to plant species that are not models, for which there may be limited genomic resources available. Non-model plant species are those that are not widely used as experimental organisms in research, and therefore have not been as well-characterized at the genomic level (Armengaud et al. 2014).

Also, there are technical challenges associated with using NGS technologies in plants, such as the difficulty of obtaining high-quality DNA and RNA from plant tissues, and the complexity of designing and optimizing NGS experiments for plant samples (Inglis 2018). Despite these challenges, NGS technologies have had a major impact on plant breeding and genetics and have provided valuable insights into the genetic basis of important plant traits. As these technologies continue to evolve and improve, it is likely that many of these challenges will be overcome and that NGS will become an increasingly important tool for plant breeding and genetics (Ray and Satya 2014).

#### 1.5. Potential Future Impacts of NGS on Plant Breeding Programs

NGS technologies have the potential to have a significant impact on plant breeding programs in the future. These technologies are likely to continue to play a key role in the development of new plant varieties. One potential impact of NGS is the acceleration of plant breeding programs (Altman et al. 2021; Savadi et al. 2021). These have the possibility to significantly influence plant breeding programs in the future. These technologies allow researchers to analyze the genomes, transcriptomes, and proteomes of plants quickly and accurately, providing valuable information about the genes and gene products engaged in key agronomic aspects like yield, quality, and disease resistance (Singh et al. 2020).

The possibility to breed plants that are more adapted to changing climatic factors including rising temperatures, shifting precipitation patterns, and growing atmospheric carbon dioxide levels is another possible benefit of NGS (Leisner 2020). By enabling researchers to identify the genes and genetic variants that are correlated with plant responses to these environmental challenges, NGS technologies can help to develop new plant varieties that are better suited to these changing conditions (Bansal et al. 2013). Scientists at the International Maize and Wheat Improvement Center (CIMMYT) in Mexico used NGS technologies to identify genetic variants that are associated with drought tolerance in maize and developed new maize varieties that are more resilient to drought stress (Aryee et al. 2021). These varieties have been released in several states in Latin America, Asia, and Africa have contributed to improved crop yields in these regions.

## 2. Conclusion

Despite some challenges, such as high cost and limited reference genomes, NGS has greatly impacted plant breeding and is expected to continue to do so in the future. Overall, the potential of NGS in plant breeding and genetics is vast, and its use is likely to continue to grow in the coming years. It has already had a significant impact on the way that plant breeding is done, and it has the potential to further transform the field by enabling the development of more efficient and precise breeding strategies.

### ORCID

Kashif Ali	<a href="https://orcid.org/0000-0002-8866-0753">https://orcid.org/0000-0002-8866-0753</a>
Hafiza Faiza Ghous	<a href="https://orcid.org/0000-0002-6430-5482">https://orcid.org/0000-0002-6430-5482</a>
Nawal Shehzadi	<a href="https://orcid.org/0000-0001-5159-285X">https://orcid.org/0000-0001-5159-285X</a>
Omaima Haroon	<a href="https://orcid.org/0000-0002-3311-0172">https://orcid.org/0000-0002-3311-0172</a>
Samia Rashid	<a href="https://orcid.org/0000-0002-0310-3721">https://orcid.org/0000-0002-0310-3721</a>
Mehdi Rahimi	<a href="https://orcid.org/0000-0001-5625-3275">https://orcid.org/0000-0001-5625-3275</a>

### REFERENCES

- Afzal AJ, Wood AJ and Lightfoot DA, 2008. Plant receptor-like serine threonine kinases: roles in signaling and plant defense. *Molecular Plant-Microbe Interactions* 21(5): 507-517. <https://doi.org/10.1094/MPMI-21-5-0507>

- Allard MW, Luo Y, Strain E, Li C, Keys CE, Son I, Stones R, Musser SM and Brown EW, 2012. High resolution clustering of salmonella enterica serovar montevideo strains using a Next-Generation sequencing approach. *BMC Genomics* 13(1): 1-19. <https://doi.org/10.1186/1471-2164-13-32>
- Altman A, Fan L, Foyer C, Cowling W, Mittler R, Qaim M, Weber AP, Reynolds M, Varshney RK and Fernie A, 2021. Past and future milestones of plant breeding. *Trends in Plant Science* 26(6): 530-538. <https://doi.org/10.1016/j.tplants.2021.03.013>
- Armengaud J, Trapp J, Pible O, Geffard O, Chaumot A and Hartmann EM, 2014. Non-model organisms a species endangered by proteogenomics. *Journal of Proteomics* 105: 5-18. <https://doi.org/10.1016/j.jprot.2014.01.007>
- Aryee SND, Owusu-Adjei D, Osei-Amponsah R, Skinner B, Sowatey E, and Sargent CA, 2021. Sustainable genomic research for food security in Sub-Saharan Africa. *Agriculture and Food Security* 10(1): 1-12. <https://doi.org/10.1186/s40066-021-00287-9>
- Bansal KC, Lenka SK, and Mondal TK, 2013. Genomic resources for breeding crops with enhanced abiotic stress tolerance. *Plant Breeding* 133(1): 1-11. <https://doi.org/10.1111/pbr.12117>
- Ben Khedher M, Ghedira K, Rolain J M, Ruimy R, and Croce O, 2022. Application and challenge of 3rd generation sequencing for clinical bacterial studies. *International Journal of Molecular Sciences* 23(3): 1395. <https://doi.org/10.3390/ijms23031395>
- Camilli A, Beattie DT and Mekalanos JJ, 1994. Use of genetic recombination as a reporter of gene expression. *Proceedings of the National Academy of Sciences* 91(7): 2634-2638. <https://doi.org/10.1073/pnas.91.7.2634>
- Campos MD, Félix MDR, Patanita M, Materatski P and Varanda C, 2021. High throughput sequencing unravels tomato-pathogen interactions towards a sustainable plant breeding. *Horticulture Research* 8. <https://doi.org/10.1038/s41438-021-00607-x>
- Dmitriev AA, Pushkova EN and Melnikova NV, 2022. Plant genome sequencing: modern technologies and novel opportunities for breeding. *Molekuliarnaia Biologija* 56(4): 531-545. <https://doi.org/10.31857/S0026898422040048>
- Fass MI, Rivarola M, Ehrenbolger GF, Maringolo CA, Montecchia JF, Quiroz F and Lia VV, 2020. Exploring sunflower responses to Sclerotinia head rot at early stages of infection using RNA-seq analysis. *Scientific Reports* 10(1): 1-14. <https://doi.org/10.1038/s41598-020-70315-4>
- Furey TS, 2012. Chip-Seq and beyond new and improved methodologies to detect and characterize Protein-DNA interactions. *Nature Reviews Genetics* 13(12): 840-852. <https://doi.org/10.1038/nrg3306>
- Gupta AK and Gupta UD, 2014. Next Generation Sequencing and Its Applications. In *Animal Biotechnology* (pp. 345-367). Elsevier. <https://doi.org/10.1016/b978-0-12-416002-6.00019-5>
- Inglis PVW, Pappas MDCR, Resende LV and Grattapaglia D, 2018. Fast and inexpensive protocols for consistent extraction of high quality DNA and RNA from challenging plant and fungal samples for high-throughput SNP genotyping and sequencing applications. *PLoS One* 13(10): e0206085. <https://doi.org/10.1371/journal.pone.0206085>
- Jiao WB and Schneeberger K, 2017. The impact of third generation genomic technologies on plant genome assembly. *Current Opinion in Plant Biology* 36: 64-70. <https://doi.org/10.1016/j.pbi.2017.02.002>
- Kulski J, 2016. Next-Generation sequencing — An overview of the history tools and Omic applications. In: *Next Generation Sequencing-Advances Applications and Challenges*, Kulski J (ed). IntechOpen Limited London, UK. <https://doi.org/10.5772/61964>
- Kumar J and Sen Gupta D, 2020. Prospects of next generation sequencing in lentil breeding. *Molecular Biology Reports* 47(11): 9043-9053. <https://doi.org/10.1007/s11033-020-05891-9>
- Kumari A, Sumer S, Jalan B, Nongbri PL and Laskar MA, 2017. Impact of next generation sequencing technology in Plant-Microbe interaction study. *Microbial Applications* 1: 269-294. <https://doi.org/10.1007/978-3-319-52666-913>
- Leisner CP, 2020. Climate change impacts on food Security-Focus on perennial cropping systems and nutritional value. *Plant Science* 293: 110412. <https://doi.org/10.1016/j.plantsci.2020.110412>
- Li Y, Wang L, Sun G, Li X, Chen Z, Feng J and Yang Y, 2021. Digital gene expression analysis of the response to ralstonia solanacearum between resistant and susceptible tobacco varieties. *Scientific Reports* 11(1): 1-16. <https://doi.org/10.1038/s41598-021-82576-8>
- Liu L, Li Y, Li S, Hu N, He Y, Pong R and Law M, 2012. Comparison of Next-Generation sequencing systems. In: *The Role of Bioinformatics in Agriculture*, Kumar S (ed), 1st Ed, Apple Academic Press, New York, USA. <https://doi.org/10.1155/2012/251364>
- Michael TP and VanBuren R, 2015. Progress challenges and the future of crop genomes. *Current Opinion in Plant Biology* 24: 71-81. <https://doi.org/10.1016/j.pbi.2015.02.002>
- Mukherjee M, Mukherjee PK, Horwitz BA, Zachow C, Berg G and Zeilinger S, 2012. Trichoderma-Plant-Pathogen interactions: advances in genetics of biological control. *Indian Journal of Microbiology* 52(4): 522-529. <https://doi.org/10.1007/s12088-012-0308-5>
- Muskovic W, Slavich E, Maslen B, Kaczorowski DC, Cursons J, Crampin E and Kavallaris M, 2022. High temporal resolution RNA-Seq time course data reveals widespread synchronous activation between mammalian lncRNAs and neighboring Protein-Coding genes. *Genome Research* 32(8): 1463-1473. <https://doi.org/10.1101/2021.08.25.457323>
- Mwadingeni L, 2016. Genetic analysis of drought tolerance in selected bread wheat (*Triticum aestivum* L.) genotypes (Doctoral dissertation).
- Ray S and Satya P, 2014. Next generation sequencing technologies for next generation plant breeding. *Frontiers in Plant Science* 5: 367. <https://doi.org/10.3389/fpls.2014.00367>
- Sahu PK, Sao R, Mondal S, Vishwakarma G, Gupta SK, Kumar V, Singh S, Sharma D and Das BK, 2020. Next generation sequencing based forward genetic approaches for identification and mapping of causal mutations in crop plants: A comprehensive review. *Plants* 9(10): 1355. <https://doi.org/10.3390/plants9101355>

- Salgotra RK, Gupta BB and Stewart CN, 2014. From genomics to functional markers in the era of next-generation sequencing. *Biotechnology Letters* 36(3): 417-426. <https://doi.org/10.1007/s10529-013-1377-1>
- Savadi S, Mangalassery S and Sandesh MS, 2021. Advances in genomics and genome editing for breeding next generation of fruit and nut crops. *Genomics* 113(6): 3718–3734. <https://doi.org/10.1016/j.ygeno.2021.09.001>
- Singh RK, Prasad A, Muthamilarasan M, Parida SK and Prasad M, 2020. Breeding and biotechnological interventions for trait improvement: status and prospects. *Planta* 252(4): 54. <https://doi.org/10.1007/s00425-020-03465-4>
- Slatko BE, Gardner AF and Ausubel FM, 2018. Overview of Next-Generation Sequencing Technologies. In: *Current Protocols in Molecular Biology* (Vol. 122, Issue 1). Wiley. <https://doi.org/10.1002/cpmb.59>
- Song J, Li Z, Liu Z, Guo Y and Qiu LJ, 2017. Next-Generation sequencing from Bulk-Segregant analysis accelerates the simultaneous identification of two qualitative genes in soybean. *Frontiers in Plant Science* 8: 919. <https://doi.org/10.3389/fpls.2017.00919>
- Sun Y, Shang L, Zhu QH, Fan L and Guo L, 2022. Twenty years of plant genome sequencing: achievements and challenges. *Trends in Plant Science* 27(4): 391–401. <https://doi.org/10.1016/j.tplants.2021.10.006>
- Unamba CI, Nag A and Sharma RK, 2015. Next generation sequencing technologies: the doorway to the unexplored genomics of non-model plants. *Frontiers in Plant Science* 6: 1074. <https://doi.org/10.3389/fpls.2015.01074>
- Vlk D and Řepková J, 2017. Application of next-generation sequencing in plant breeding. *Czech Journal of Genetics and Plant Breeding* 53(3): 89-96. <https://doi.org/10.17221/192/2016-CJGPB>
- Yang GX, Li X and Snyder M, 2012. Investigating Metabolite–Protein interactions: an overview of available techniques. *Methods* 57(4): 459-466. <https://doi.org/10.1016/j.ymeth.2012.06.013>