Next Generation Sequencing in Plant Breeding

Chapter · September 2023

DOI: 10.5281/zenodo.8406007

4 authors, including:

Shahiba A. M.
Kerala Agricultural University
49 PUBLICATIONS **35** CITATIONS

SEE PROFILE

Modern Approaches in Plant Breeding

Enhancing Crop Genetics

Dr. Shahiba A.M.

G. Pranay

Pallavi Sonaniya

Jeetendra Singh Kushwaha

Dr. Sonali Ambadas Aware

Elite Publishing House

First Edition 2023

ISBN : 978-93-58990-60-7

Published by:

Elite Publishing House
A-10/28, Sector - 18, Rohini, New Delhi - 110089
Tele Info: 9289051518, 9289051519
Email: ephinternational@gmail.com, ephpublishers@gmail.com
Website: www.elitepublishing.in

12

Next Generation Sequencing in Plant Breeding

Arya Balachandran[1], Dr. Lovely B.[2], Shazna A.[3] and Shahiba A. M.[4]

[1] Msc. Student, Department of Plant Breeding and Genetics, Kerala Agricultural University, (Kerala)
[2]Assistant Professor, Department of Plant Breeding and Genetics, Kerala Agricultural University, (Kerala)
[3]Msc. Student, Department of Plant Breeding and Genetics, Kerala Agricultural University, (Kerala)
[4]Ph.D. Scholar, Department of Plant Breeding and Genetics, Kerala Agricultural University, (Kerala)

Abstract

The investigation of diverse biological topics is made possible by the development of Next-Generation Sequencing (NGS) technologies, which are fast advancing and becoming more affordable. NGS has transformed the field, enabling plant breeders to unravel the genetic diversity, identify key traits, and expedite the development of improved plant varieties. NGS enhances crop productivity, resilience, and sustainability. NGS has immense potential in revolutionizing plant breeding and contributing to global food security.

Keywords: Crop productivity, genetic diversity, resilience, sustainability, unravel

Introduction

The investigation of diverse biological topics is made possible by the development

of next-generation sequencing (NGS) technologies, including as synthesis, ligation, and single-molecule sequencing, which are fast advancing and becoming more affordable. When selecting superior genotypes for plants, techniques based on genome decoding improve genetic gain during selection. Genetic diversity for plant breeding is shown through NGS techniques including genome-wide association mapping and genomic selection. Along with mapping and cloning quantitative trait loci, NGS is also capable of decoding the transcriptome and epigenome. Instead of offering an exhaustive list of sequencing projects or methods, this study explains the logic behind various plant genome sequencing methods and gives examples of some potential uses for them.

Kumar et al. (2012) evaluated the evolution of sequencing technology. The identification of agronomically significant genes controlling yield and tolerance to biotic and abiotic stressors is made possible by next-generation sequencing (NGS), which permits comprehension of the genetic composition and behaviour of plant genomes. NGS technologies allow for high-resolution study of plant genomes, reasonably priced nucleotide variation profiling, and extensive functional marker discoveries. Utilising traditional methods and methodologies, this data helps in the selection of economically significant features in plant breeding. A gene revolution in plant breeding is being sparked by the discovery of molecular markers for marker-assisted selection (MAS) and the analysis of genome connections made possible by NGS techniques and internet resources. Different applications of next-generation sequencing technique in genetics are shown below Fig.1.

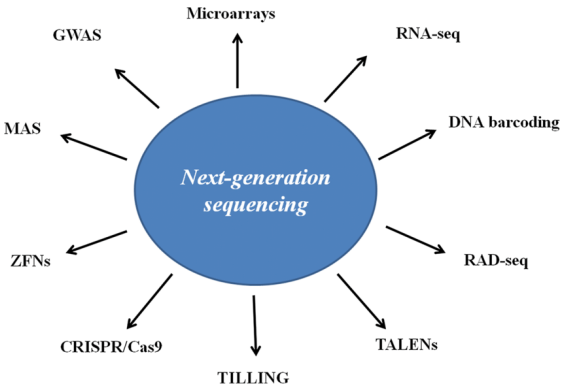

Fig.1. Different applications of next-generation sequencing technique in genetics (Niazian,2019)

Plant Breeding: Challenges and Opportunities

The issues brought on by a burgeoning world population are greatly helped by plant breeding. Food, feed, fibre, and other agricultural goods are in greater demand as the world's population continues to grow. Arable land cannot be expanded much longer, and the problems facing agriculture are only becoming worse due to climate change. Plant breeding becomes a crucial tactic in creating superior crop varieties that are more productive, robust, and nutritionally dense in order to fulfil the rising demand for food and resources responsibly. Plant breeding may increase agricultural production and productivity by creating cultivars with desired features, lowering yield losses, boosting disease and insect resistance, increasing abiotic stress tolerance, raising nutritional value, and utilising sustainable farming methods. Breeders can create crops that are more productive by choosing and breeding plants with desired features that are better suited to particular conditions and agronomic techniques. The development of types that are more resistant to harsh weather conditions can also be aided through plant breeding, guaranteeing reliable yields even in challenging circumstances.

Next-Generation Sequencing (NGS) has revolutionized plant breeding by enabling faster genetic analysis, precision breeding, genomic selection, identification of genetic variation, and understanding plant biology. NGS enables rapid sequencing of the entire genome of a plant species or individual varieties, enabling breeders to select and introgress desirable traits into new varieties. This approach reduces time and resources needed in breeding, enables genomic selection, and enhances adaptability and resilience in cultivated crops. Additionally, NGS deepens our understanding of plant biology, enabling more informed and effective breeding strategies.

In conclusion, plant breeding, empowered by advancements in technology like Next-Generation Sequencing, plays a pivotal role in addressing the challenges posed by a growing population. By developing improved crop varieties that are more productive, resilient, and nutritious, plant breeding contributes significantly to global food security and sustainable agriculture.

The Emergence of Next-Generation Sequencing

NGS, or next-generation sequencing, transformed genomic research in the middle of the 2000s. High-throughput sequencing is achieved using a variety of techniques by developed platforms as Illumina/Solexa, 454 Pyrosequencing, Ion Torrent, PacBio, and Oxford Nanopore. The finding of genetic variants, functional components, and disease-causing mutations has been hastened by the use of NGS, which has also permitted large-scale investigations. The development of personalised treatment and the ability to diagnose uncommon genetic illnesses are two further effects of

NGS on medicine and diagnostics. As long as efforts are made to increase accuracy, lower prices, and broaden applications, NGS has a bright future. NGS is expected to become a pillar of contemporary biology as the technology develops and has an increasing influence on our understanding of human health and illness.

Objectives

This chapter's goals are to clarify the crucial role of plant breeding in addressing the problems brought on by a growing world population, such as rising food demand and climate change, and to investigate how Next-Generation Sequencing (NGS) developments help hasten the creation of improved crop varieties. In addition to demonstrating how NGS technologies enable accurate genetic analysis, trait discovery, and informed breeding decisions, the chapter aims to highlight the importance of plant breeding in enhancing agricultural productivity, disease resistance, and stress resistance. This will ultimately lead to sustainable agriculture and help ensure long-term food security.

Genome Sequencing and Assembly

The fundamental procedures of genome sequencing and assembly, which entail figuring out the whole DNA sequence of an organism's genome and rebuilding it into a useable form, are used in genomics research. These procedures are essential for figuring out how genes work, researching genetic variants, and identifying functional components in the genome.

Whole Genome Sequencing Techniques: Whole genome sequencing (WGS) is a method that identifies the whole DNA sequence of an organism, including both coding and non-coding sections. Oxford Nanopore Technologies (ONT), PacBio Sequencing, and Illumina Sequencing are examples of common technologies. Long reads are produced by PacBio using single-molecule, real-time sequencing technology, whereas short reads are produced by Illumina using reversible dye-terminator chemistry. Long reads are generated using ONT sequencing, which employs nanopores to monitor electrical current changes in order to sequence contiguous genomic areas and identify structural alterations. These sequencing methods have revolutionised genomes research and are a contribution to it.

De novo Assembly and Reference-Based Assembly: A whole genome sequence is reconstructed using sequencing reads in a process known as genome assembly. De novo assembly and reference-based assembly are the two basic strategies. De novo assembly assembles sequencing reads into contigs by locating overlaps and combining them into larger sequences. Reference-based assembly aligns sequencing reads to a known reference genome. The decision between these methods is influenced by

the study goals, the complexity of the genome being sequenced, the availability and quality of a reference genome, and other factors. There are tools that provide user-friendly interfaces for gene discovery in *de novo* transcriptomes, such as Trapid (Van Bel *et al.* 2013).

Genomic Resources for Plant Breeding: Genetic information for crop development is provided via genomic resources, which are crucial in plant breeding programmes. A few of the most important tools used in plant breeding include reference genomes, genetic variation databases, transcriptome databases, QTL databases, and germplasm collections. Reference genomes offer a thorough representation of the genetic code of a plant species, allowing for the mapping of sequencing reads, the detection of genetic variants, and the annotation of genes and functional components. When compared to QTL databases, transcriptomic databases include gene expression patterns across tissues, developmental stages, and environmental circumstances. Plant species' genetic variety is preserved in germplasm collections, which also make it possible to conduct genetic research, identify novel traits, and create new breeding lines. By enabling plant breeders to make educated decisions, speed up breeding cycles, and create improved crop varieties with desirable traits, these genomic resources, in conjunction with cutting-edge sequencing technologies and bioinformatics tools, help ensure the world's food security and agricultural sustainability.

Platform	Pros	Cons
Illumina (Short read)	✓ Used widely by the NGS industry ✓ **Lowest per-Gb sequencing cost range** ✓ Highest confirmed output ✓ Low error rate (<1%)	× Short read length × **Limited *de novo* genome assembly**
PacBio (Long read)	✓ Real long reads ✓ **Extremely high accuracy with CCS** (>99.999% at 20 passes) ✓ Direct detection of epigenetic modifications ✓ Real-time data output ✓ No problem with repeats, low/high % GC	× Relatively **high cost** per Gb × **Large amounts of starting material required for library preparation** × High error rate of CLR mode (single pass) × Shorter reads in CCS mode × Maximum read length limited by polymerase processivity
ONT (Long read)	✓ Real (ultra-) long reads (**Longest confirmed reads**) ✓ Direct detection of epigenetic modifications ✓ Real-time data output ✓ Direct sequencing of RNA and detection of RNA modifications	× **High overall error rate** and systematic errors with homopolymers × **Large amounts of starting material required for library preparation** × Sensitivity of biological nanopores to changes in experiment environment × **Frequent changes of software versions, kits**

Fig.2. Pros and Cons of illumine, PacBio and ONT

Genetic Variation and Diversity Analysis

An essential component of genomics research is genetic variation and diversity analysis, which is especially important for comprehending the genetic basis of characteristics, population dynamics, and germplasm characterisation for crop development. To investigate genetic diversity within populations and species, a

variety of methods and analysis are used. Here are the key topics related to Genetic Variation and Diversity Analysis:

Genotyping-by-Sequencing (GBS): A low-cost, high-throughput approach for genotyping many Single Nucleotide Polymorphisms (SNPs) throughout the genome is genotyping-by-sequencing (GBS). Reduced-representation libraries are sequenced using a combination of restriction enzyme digestion and next-generation sequencing. SNP calling, library preparation, and sequencing are all steps in the procedure Fig.3. GBS is often employed in plant and animal species for genomic selection, linkage mapping, association research, and diversity analyses. Utilising the genotype will allow for improved cultivar selection and trait prediction in plant breeding. Genomic-assisted breeding will be feasible at the scale necessary to increase the world's food supplies in the face of declining arable land and climate change by enhancing our understanding of the link between heritable genetic determinants and the phenotypic outputs (Poland and Rife, 2012).

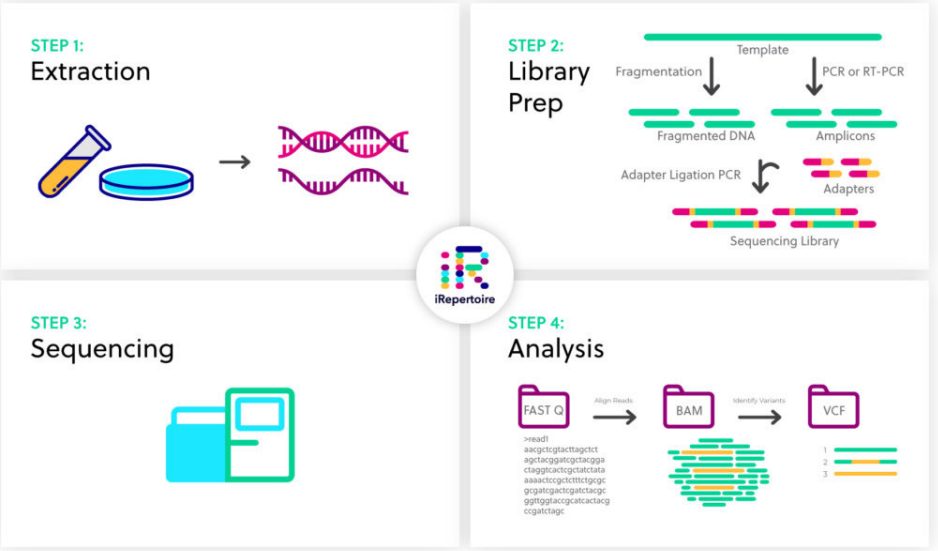

Fig.3. Extraction, library preparation, sequencing and analysis.

Single Nucleotide Polymorphism (SNP) Analysis: SNPs, or single nucleotide polymorphisms, are genetic changes that occur at specific locations in the genome. Through the use of techniques like GBS, whole genome sequencing, or targeted genotyping arrays, SNP analysis includes the identification and characterisation of SNPs within a population or species. Understanding SNP's possible functional ramifications and effects on protein function is made easier by SNP annotation. SNP information is used to investigate population structure, evolutionary connections

within and across populations, and genetic diversity. Population genetic techniques that characterise population structure and gene flow include computing genetic diversity indices and principal component analysis (PCA).

Structural Variants and Indels: Larger structural variants seen in genomes, like as insertions, deletions, duplications, inversions, and translocations, can have a major influence on phenotypic variation and disease risk. The identification and characterization of these alterations depend on structural variant analysis and Indel analysis. In order to comprehend evolutionary processes and conservation efforts, population genetics and germplasm characterisation are essential. In germplasm collections, accessions' genetic diversity and relatedness are evaluated using genotyping techniques including SNP genotyping arrays and GBS. While phylogenetics reconstructs evolutionary connections between individuals or populations using genetic data, population structure analysis uses genetic similarity to identify subgroups within populations. By offering insightful knowledge about the genetic basis of features, the evolutionary past, and conservation measures, these investigations improve science in areas including agriculture, ecology, and medicine.

Marker-Assisted Selection (MAS)

A breeding technique called marker- assisted selection (MAS) makes better use of molecular markers to identify individuals with desired features. Breeders can choose the best candidates by analysing certain DNA areas linked to the desired features. Traditional breeding techniques may be considerably accelerated using MAS, and trait selection accuracy can be increased. Here's an overview of Marker-Assisted Selection:

Molecular Markers and MAS Principles: Molecular markers are genetic differences present in people's DNA that act as indications of particular features of interest. MAS focuses on locating molecular markers connected to specific qualities and employing them to choose individuals that possess such features throughout the breeding process. Trait characterisation, marker discovery, marker validation, marker-assisted selection, and recurrent selection are crucial MAS stages. Identification of the target characteristic is necessary for trait characterisation. The target trait may be disease resistant, yield-related, or a quality feature. Finding closely related markers includes using genetic mapping, association research, or gene analysis. Markers with strong relationships are chosen for future usage after validation studies identify those that consistently display these correlations.

SNP-Based Markers and High-Throughput Genotyping: Single Nucleotide Polymorphisms (SNPs) are abundant genetic variation in the genome and widely used as molecular markers in Marker-Assisted Selection. High-throughput genotyping

technologies, like SNP genotyping arrays and Genotyping-by-Sequencing (GBS), enable cost-effective and efficient genotyping of thousands to millions of SNPs, enabling simultaneous trait analysis in large populations.

Case Studies of MAS in Plant Breeding: In plant breeding programmes to enhance crop types, marker- assisted selection (MAS) has been extensively employed. Abiotic stress tolerance, quality attributes, improved yield, disease resistance, and hybrid breeding are a few examples. In areas prone to dryness, MAS has been utilised to generate drought-resistant cultivars, add resistance genes against particular diseases, increase grain quality features, and boost yield components.

In these case studies, MAS has allowed breeders to make more precise and efficient selections, leading to the development of improved crop varieties with enhanced agronomic traits and stress resilience. It has significantly accelerated the breeding process and contributed to global food security by delivering crops with improved productivity and adaptability.

Quantitative Trait Locus (Qtl) Mapping

QTL mapping is a genetics and genomics technique that identifies genomic regions associated with quantitative traits, influenced by multiple factors. It offers insights into complex traits, genetic architecture, enabling breeders to develop improved varieties.

QTL Mapping Strategies: The process of QTL mapping entails locating genomic areas (QTLs) that contribute to the observed variation in a quantitative characteristic. There are two main approaches: association mapping and linkage mapping, which examine the relationships between markers and phenotypic characteristics in various populations. Association mapping examines the co-segregation of markers and the trait. Mendelian-inherited characteristics and biparental populations can both benefit from linkage mapping; however, populations with substantial genetic variety in wild or breeding populations benefit more from association mapping.

Linkage Mapping and Association Mapping: Both linkage mapping and association mapping are effective methods for locating QTLs, although they have different populations and genetic make-ups. Due to a minimal number of recombination events in bi-parental populations, QTLs may be mapped with a high degree of accuracy. It is however restricted to Mendelian-inherited features and can be diminished in a small population. The use of varied individuals from breeding or wild populations in association mapping enables the mapping of QTLs in various genetic backgrounds and complicated characteristics with numerous contributors. The considerable linkage disequilibrium in several populations, however, may make it difficult to control for population structure and find causal QTLs.

Genomic Selection and Breeding Value Estimation: Genomic Selection is a breeding strategy that uses dense markers across the genome to predict individuals' breeding value for target traits. This strategy is particularly effective for complex traits influenced by multiple genes, where traditional phenotypic selection may be less efficient. It involves genotyping a large number of markers, phenotyping individuals for target traits, building a statistical model relating marker genotypes to phenotypic values, and using the model to predict individuals' breeding value based on their genotypic data.

Integrating Genomic Data with Phenotypic Information: Integrating genomic data with phenotypic information is crucial for effective QTL mapping and genomic selection. This enables breeders and researchers to identify QTLs associated with target traits, estimate breeding values, and perform genome-wide association studies (GWAS) to predict genetic potential. This comprehensive understanding of complex traits enables informed breeding programs, resulting in improved and resilient crop varieties.

Transcriptome Analysis

Transcriptome analysis is a crucial genomics method for studying gene expression patterns, understanding functional elements, and examining biological processes. Here are the key aspects of transcriptome analysis:

RNA-Seq and Gene Expression Profiling: RNA-Seq is a high-throughput sequencing technique for quantifying and profiling the transcriptome. It involves library preparation, converting RNA into cDNA, sequencing using NGS platforms, and data analysis to identify genes and expression levels in the sample. This process helps detect and quantify all RNA molecules, including coding and non-coding RNAs, in a sample. Data acquired by RNA-seq are universal. Furthermore, they can be used in gene characterization (Dassanayake *et al.* 2009) and molecular marker development (Trick *et al.* 2009).

Differential Gene Expression Analysis: Differential gene expression analysis identifies genes that are differentially expressed between conditions or experimental groups, allowing researchers to understand upregulation or downregulation under specific treatments or tissues. The process involves data preprocessing, statistical analysis using methods like the negative binomial model, and visualization using plots, heatmaps, or other graphical representations.

Functional Annotation and Pathway Analysis: Identifying differentially expressed genes involves functional annotation and pathway analysis to understand their biological significance and processes. Functional annotation associates genes with known functional information, while pathway analysis enriches them in specific

biological pathways, providing insights into underlying mechanisms and potential molecular pathways affected by experimental conditions.

Transcriptome-Based Marker Discovery: Transcriptome analysis enables marker discovery of SNPs and SSRs, enabling researchers to develop molecular markers for genotyping and breeding. These markers are valuable for non-model organisms and species without a reference genome, enabling genetic mapping, marker-assisted selection, and population genetics studies. In summary, transcriptome analysis plays a crucial role in understanding gene expression regulation, identifying differentially expressed genes, annotating functional elements, and discovering markers for genomics and breeding applications. It has broad applications in various fields, including agriculture, medicine, and ecology. Roche technology has been used effectively to sequence a variety of non-model plants, such as the comparative sequencing of transcripts from two different olive trees during fruit development (*Olea europaea* L.; Alagna *et al.* 2008) and in Yunong 201, a cultivar of bread wheat, transcriptome study was conducted (*Triticum aestivum* L.; Zhang *et al.* 2016).

In addition to the Sanger sequencing method used to identify expressed sequence tags, Illumina technology (EST; Swarbreck *et al.* 2011) is typically beneficial due to its better plant covering transcriptomes. Microarray and Roche technologies were used by Nigam et al. (2014) to determine the genes and by-products linked to cotton fibre quality.

Epigenetic Studies

The focus of epigenetic research is on heritable modifications of gene expression that do not entail changes to the DNA sequence. Epigenetic changes, including as DNA methylation and histone modifications, are crucial for the control of genes and can have an impact on a range of biological functions and plant properties. Studies on epigenetics have substantial effects on plant breeding and our understanding of the molecular underpinnings of complex characteristics. Let's explore the key aspects of epigenetic studies:

DNA Methylation and Histone Modifications: DNA methylation, an epigenetic modification, adds a methyl group to the cytosine base of DNA, affecting gene expression by regulating transcriptional activity. Hypermethylation leads to gene silencing, while hypomethylation increases expression. Histone modifications, like methylation, acetylation, phosphorylation, and ubiquitination, alter DNA's accessibility to transcriptional machinery, activating or repressing gene expression depending on the modification's location on the histone tail.

Epigenomics Approaches in Plant Breeding: Epigenomics involves the genome-

wide analysis of epigenetic modifications. In plant breeding, epigenomics approaches are used to study the epigenetic landscape and its association with important traits. Epigenomics techniques include Whole Genome Bisulfite Sequencing (WGBS), Chromatin Immunoprecipitation Sequencing (ChIP-Seq), and Reduced Representation Bisulfite Sequencing (RRBS). WGBS determines DNA methylation status at single-base resolution, revealing genome-wide patterns and differentially methylated regions. ChIP-Seq identifies histone modifications related to gene regulation and active or repressive chromatin regions linked to specific traits. RRBS is a cost-effective method for DNA methylation analysis in CpG-rich regions.

Epigenetic Regulation of Plant Traits: Epigenetic modifications impact plant traits such as developmental processes, abiotic stress responses, disease resistance, and nutritional quality. They affect seed germination, flowering time, organ differentiation, and response to environmental stresses. Epigenetic changes affect defense-related genes, affecting plant resistance to pathogens. Additionally, epigenetic regulation affects the accumulation of secondary metabolites, affecting crop nutritional quality.

Epigenetics-Assisted Breeding Strategies: Epigenetics-assisted breeding strategies improve plant breeding outcomes by utilizing epigenetic information. These strategies include epigenetic selection, recombination, variation analysis, and genome editing tools like CRISPR-Cas9. These methods aim to identify individuals with desirable traits, promote recombination, identify novel epialleles, and control gene expression and phenotypic traits. Epigenetic studies are an emerging field with immense potential for improving plant breeding strategies, enhancing our understanding of complex traits, and developing more resilient and productive crop varieties. Integrating epigenetic knowledge with traditional genetics can contribute to the sustainable development of agriculture and food security.

NGS Applications in Crop Improvement: NGS technologies have revolutionized crop improvement by providing a deeper understanding of genetic basis, enabling more efficient breeding strategies, and accelerating the development of improved crop varieties. Applications include genome sequencing, transcriptome analysis, genotyping and marker discovery, association mapping and genomic selection, and genome editing. Genome sequencing allows for high-quality reference genomes for functional annotation, comparative genomics, and marker discovery. Transcriptome analysis helps identify genes associated with specific traits and understand gene regulatory networks. NGS-based genotyping techniques, such as GBS and SNP genotyping arrays, facilitate high-throughput marker discovery and genotyping, enhancing breeding efficiency. Radish (Raphanus sativus L.) roots have been used to identify possible cadmium-responsive miRNAs and the genes they target using Illumina technology (Xu *et al*. 2013b).

SNPs are being converted into genetic markers for useful crop species through extensive genetic variation discovery initiatives and sequencing technology (Deschamps and Campbell, 2010).

Long non-coding RNAs (lncRNAs) have been shown to constitute yet another significant regulating mechanism. The RNA molecules in question are longer than 200 bp and do not encode any protein. They are related to reactions to abiotic stress, blooming time control, and gene silencing, according to current studies (Wang *et al*. 2014; Zhang *et al*. 2014).

Examples of Improved Traits through NGS-Assisted Breeding: NGS-assisted breeding has improved crop traits, such as disease resistance, abiotic stress tolerance, yield improvement, and nutritional enhancement. By identifying genetic variants linked to disease resistance, abiotic stress tolerance, and yield-related traits, NGS technologies have reduced chemical inputs and improved crop yield and quality. Additionally, NGS data has identified genes associated with yield-related traits, leading to increased crop productivity. Overall, NGS-assisted breeding has significantly contributed to the development of healthier crop varieties.

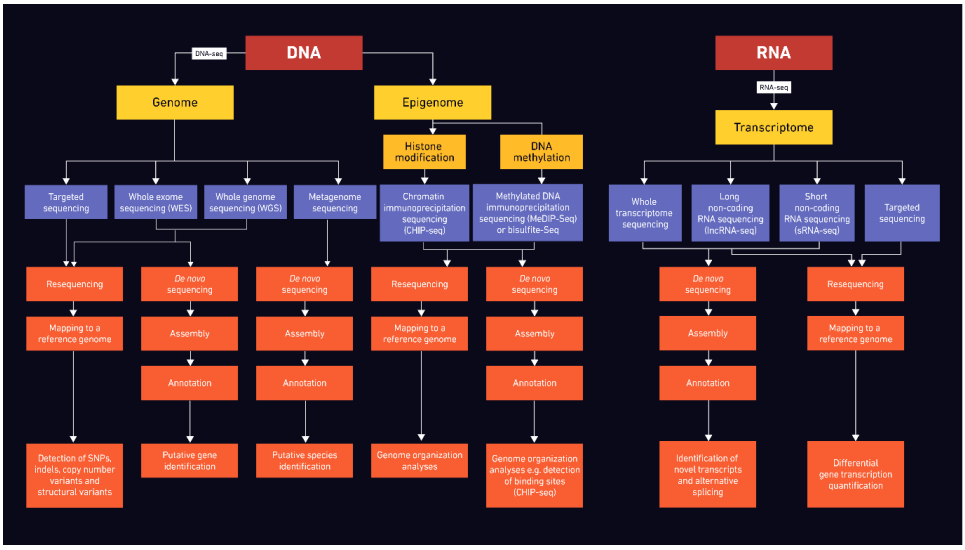

Fig.4. Flow diagram indicating possible sequencing strategies for different sample types (Salko et. al.,2018)

Prospects and Implications for Plant Breeding: Genomics in plant breeding offers promising prospects for understanding complex traits, developing resilient, high-yielding, and nutritionally enhanced crop varieties. Advancements in NGS

and bioinformatics provide access to genetic information, enhancing precision breeding strategies. Epigenomics offers insights into transgenerational effects and environmental adaptation, and integrating multi-omics data revolutionizes breeding programs. Flow diagram indicating possible sequencing strategies for different sample types is shown below **fig.4.**

Constraints and the Future

Next-generation sequencing technology has significantly benefited crop genome sequencing, gene regulation, and functional genomics. Third-generation sequencing systems like Oxford Nanopore and PacBio offer long-read sequencing, enabling the identification of epigenetic markers like DNA methylation and faster gene identification. NGS methods also enable genome-wide SNP discovery, allele mining, and molecular marker creation.

NGS technology has the potential to be applied in plant breeding by finding novel allelic variations in crops' genomes. Genome sequences for several crops have been made available, enabling genome editing techniques for site-directed mutagenesis. Nucleases targeting specific sequences, such as CRISPR/Cas9, TALEN, and ZFN, can be used for this. The CRISPR/Cas9 technology has been used to plants, such as rice, to control the mutation of genes linked to morphological and qualitative features (Shan *et al.* 2013) and in cucumber to develop resistance to *Cucumber vein yellowing virus, Zucchini yellow mosaic virus,* and *Papaya ringspot mosaic virus-W* (Chandrasekaran *et al.* 2016). Herbicide resistance was produced in modified tobacco plants by acetolactate synthase gene mutations caused by Zinc finger nucleases (ZNF). (Townsend *et al*. 2009), and Tomato genes associated with gibberellin signalling were engineered using TALEN-mediated mutagenization (Lor *et al*. 2014). Future research will increasingly use irradiation techniques and advanced NGS pipelines to produce mutants and characterize them. Combining DNA sequencing technology with irradiation tools like cosmic rays and heavy ion beams will increase mutagenesis efficiency and improve the utilization of genetic materials in plant breeding and functional genomics research. The newest sequencing machines of the family are shown below fig.5.

Fig.5. The latest sequencing machines

Conclusion

Genomics in plant breeding, focusing on whole genome sequencing techniques, genome assembly, genetic resources, genetic variation and diversity analysis, marker-assisted selection, transcriptome analysis, and epigenetic studies. WGS techniques enable complete genome sequencing, while assembly and reference-based assembly help understand genome structure and organization. Genetic resources facilitate trait discovery and marker development, while marker-assisted selection and genomic selection improve breeding processes. Transcriptome analysis and epigenetic studies expand our understanding of gene regulation and offer new avenues for crop improvement.

Mutant breeding is a crucial aspect of modern plant breeding, playing a vital role in global nutritional and food security. Advancements in molecular marker techniques, mapping and cloning, next-generation sequencing, and functional genomics have made induced mutagenesis more practical for crop improvement and identifying new candidate genes and their biological functions. Next-generation sequencing (NGS) methods are essential for locating and characterizing genomic variants linked to economically significant characteristics. NGS applications include WGR and transcriptome profiling, providing complete information on genetic diversity and regulatory mechanisms. These strategies are applicable to classic mutations with qualitative features and are usable even without a reference genome or genetic instruments. Bioinformatics and statistical techniques are necessary for precise assembly, alignment, and variant identification. A contemporary plant breeding team must include experts from various disciplinary backgrounds, including plant biology, genetics, physiology, molecular biology, bioinformatics, statistics, and mathematics. The development of third-generation sequencing technology will further enhance the accuracy of variant and mutation detection.

Genomics revolutionizes plant breeding by unlocking genetic potential and addressing global challenges like food security, climate change, and sustainable agriculture. Integrating genomics with traditional approaches offers precision breeding, enabling efficient and sustainable development of desired traits.

References

Alagna, F., D'Agostino, N., Torchia, L., Servili, M., Rao, R., Pietrella, M., ... & Perrotta, G. (2009). Comparative 454 pyrosequencing of transcripts from two olive genotypes during fruit development. *BMC genomics*, *10*(1), 1-15.

Chandrasekaran, J., Brumin, M., Wolf, D., Leibman, D., Klap, C., Pearlsman, M., ... & Gal-On, A. (2016). Development of broad virus resistance in non-transgenic cucumber using CRISPR/Cas9 technology. *Molecular plant pathology*, *17*(7), 1140-1153.

Dassanayake, M., Oh, D. H., Haas, J. S., Hernandez, A., Hong, H., Ali, S., ... & Cheeseman, J. M. (2011). The genome of the extremophile crucifer Thellungiella parvula. *Nature genetics*, *43*(9), 913-918.

Deschamps, S., & Campbell, M. A. (2010). Utilization of next-generation sequencing platforms in plant genomics and genetic variant discovery. *Molecular breeding*, *25*, 553-570.

Kumar, S., Banks, T. W., & Cloutier, S. (2012). SNP discovery through next-generation sequencing and its applications. *International journal of plant genomics*, *2012*.

Lor, V. S., Starker, C. G., Voytas, D. F., Weiss, D., & Olszewski, N. E. (2014). Targeted mutagenesis of the tomato PROCERA gene using transcription activator-like effector nucleases. *Plant physiology*, *166*(3), 1288-1291.

Myles, S., Peiffer, J., Brown, P. J., Ersoz, E. S., Zhang, Z., Costich, D. E., & Buckler, E. S. (2009). Association mapping: critical considerations shift from genotyping to experimental design. *The Plant Cell*, *21*(8), 2194-2202.

Nigam, D., Kavita, P., Tripathi, R. K., Ranjan, A., Goel, R., Asif, M., ... & Sawant, S. V. (2014). Transcriptome dynamics during fibre development in contrasting genotypes of G ossypium hirsutum L. *Plant biotechnology journal*, *12*(2), 204-218.

Niazian, M. (2019). Application of genetics and biotechnology for improving medicinal plants. *Planta*, *249*, 953-973.

Poland, J. A., & Rife, T. W. (2012). Genotyping-by-sequencing for plant breeding and genetics. *The plant genome*, *5*(3).

Slatko, B. E., Gardner, A. F., & Ausubel, F. M. (2018). Overview of next-generation

sequencing technologies. *Current protocols in molecular biology*, *122*(1), e59.

Shan, Q., Wang, Y., Li, J., Zhang, Y., Chen, K., Liang, Z., ... & Gao, C. (2013). Targeted genome modification of crop plants using a CRISPR-Cas system. *Nature biotechnology*, *31*(8), 686-688.

Swarbreck, S. M., Lindquist, E. A., Ackerly, D. D., & Andersen, G. L. (2011). Analysis of leaf and root transcriptomes of soil-grown Avena barbata plants. *Plant and cell physiology*, *52*(2), 317-332.

Townsend, J. A., Wright, D. A., Winfrey, R. J., Fu, F., Maeder, M. L., Joung, J. K., & Voytas, D. F. (2009). High-frequency modification of plant genes using engineered zinc-finger nucleases. *Nature*, *459*(7245), 442-445.

Trick, M., Long, Y., Meng, J., & Bancroft, I. (2009). Single nucleotide polymorphism (SNP) discovery in the polyploid Brassica napus using Solexa transcriptome sequencing. *Plant biotechnology journal*, *7*(4), 334-346.

Van Bel, M., Proost, S., Van Neste, C., Deforce, D., Van de Peer, Y., & Vandepoele, K. (2013). TRAPID: an efficient online tool for the functional and comparative analysis of de novoRNA-Seq transcriptomes. *Genome biology*, *14*(12), 1-10.

Wang, H., Chung, P. J., Liu, J., Jang, I. C., Kean, M. J., Xu, J., & Chua, N. H. (2014). Genome-wide identification of long noncoding natural antisense transcripts and their responses to light in Arabidopsis. *Genome research*, *24*(3), 444-453.

Xu, L., Wang, Y., Zhai, L., Xu, Y., Wang, L., Zhu, X., ... & Liu, L. (2013). Genome-wide identification and characterization of cadmium-responsive microRNAs and their target genes in radish (Raphanus sativus L.) roots. *Journal of experimental botany*, *64*(14), 4271-4287.

Zhang, W., Han, Z., Guo, Q., Liu, Y., Zheng, Y., Wu, F., & Jin, W. (2014). Identification of maize long non-coding RNAs responsive to drought stress. *PloS one*, *9*(6), e98958.

Zhang, N., Wang, S., Zhang, X., Dong, Z., Chen, F., & Cui, D. (2016). Transcriptome analysis of the Chinese bread wheat cultivar Yunong 201 and its ethyl methanesulfonate mutant line. *Gene*, *575*(2), 285-293.