ISBN: 978-81-974088-6-1

10. Role of Molecular Marker in Plant Biotechnology

Yashowardhan Singh

Junior Research Fellow, Department of Plant Pathology,
Rani Lakshmi Bai,Central Agriculture University,
Jhansi, Uttar Pradesh.

Likhita J.

Ph.D. Scholer Department of Plant Pathology,
RLBCAU, Rani Lakshmi Bai,
Central Agriculture University, Jhansi, Uttar Pradesh.

Meenakshi Arya

Scientist AICRP Chickpea,
Department of Plant Pathology, Rani Lakshmi Bai,
Central Agriculture University, Jhansi, Uttar Pradesh.

Swarna Kurmi

Assistant Professor,
Department of Plant Pathology,
JNKVV, Jabalpur (M.P.)

Abstract:

Molecular markers are indispensable tools in plant biotechnology, revolutionizing the field by offering insights into genetic diversity, trait inheritance, and marker-assisted selection. These markers, comprised of DNA sequences with distinct variations, enable researchers to precisely locate genes, identify genetic variations, and facilitate the development of improved plant varieties. In the realm of plant biotechnology, molecular markers serve multifaceted roles, playing pivotal roles in marker-assisted breeding programs, genetic mapping endeavors, genome characterization studies, and phylogenetic analyses. Marker-assisted breeding, a cornerstone of modern plant breeding programs, harnesses the power of molecular markers to expedite the selection of desired traits in plant populations. By pinpointing genomic regions associated with favorable characteristics, such as disease resistance, abiotic stress tolerance, and enhanced yield potential, breeders can efficiently incorporate these traits into elite cultivars, accelerating the breeding process and enhancing agricultural productivity. Genetic mapping efforts leverage molecular markers to construct detailed genetic maps of plant genomes, facilitating the elucidation of gene functions, trait associations, and chromosomal architecture. Through linkage analysis and quantitative trait loci (QTL) mapping, researchers can unravel the genetic basis of complex traits and decipher the underlying mechanisms governing trait expression and inheritance patterns.

Furthermore, molecular markers are instrumental in characterizing plant genomes, aiding in the assembly, annotation, and comparative analysis of genomic sequences. High-throughput genotyping technologies, coupled with bioinformatics tools, enable researchers to unravel the intricacies of plant genomes, unraveling gene regulatory networks, and deciphering the evolutionary history of plant species.

Keywords: *Molecular markers, Plant biotechnology, Marker-assisted breeding, Genetic diversity, Genetic mapping.*

10.1 Introduction:

DNA sequences known as molecular markers can be used to analyze individual variations or polymorphisms. Numerous genetic alterations, including insertions, deletions, point mutations, duplications, and translocations, are the source of these polymorphisms.

Crucially, these modifications might not necessarily have a direct impact on gene function. According to Nadeem *et al.* (2018), the ideal DNA marker is co-dominant, which indicates that both alleles are expressed, evenly distributed across the genome, extremely consistent in replication, and able to detect a high level of polymorphism.

A DNA segment that is easily recognized and traceable using inheritance patterns is called a molecular marker. These markers are useful tools that take advantage of naturally existing DNA polymorphisms for a variety of applications (Collard *et al.*, 2005).

To distinguish between chromosomes with mutant genes and those with normal genes through marker association, a marker has to be polymorphic, or to exist in several forms. The presence of two different genotypes or phenotypes within the same population is referred to as genetic polymorphism (Nadeem *et al.*, 2018).

In particular, DNA markers are useful for effectively assessing and choosing plant material. DNA markers segregate as single genes and are not impacted by environmental conditions, in contrast to protein markers. Plant samples may easily have their DNA extracted, and DNA markers can be analyzed in a labour- and money-efficient manner (Poczai *et al.*, 2013). Restriction fragment length polymorphisms (RFLPs), which are DNA markers produced by restriction digestion, were among the first to be used. Since then, a number of marker systems have been created to improve studies based on molecular markers (Kordrostami and Rahimi, 2015).

10.2 History of Evaluation of Molecular Marker:

Botsten *et al.* (1980) employed restriction fragment length polymorphism (RFLP) markers, which were the first DNA-based genetic markers. These were employed to find the first QTL (Stuber *et al.*, 1987) and create the first genome-wide linkage maps (Helentjans *et al.*, 1986). A novel class of DNA markers, the random amplified polymorphic DNA (RAPD) marker, was introduced in the 1990s as the focus shifted to assays based on the polymerase chain reaction (PCR) (Williams *et al.*, 1990; Welsh and McClelland *et al.*, 1990).

The advent of single nucleotide polymorphisms (SNPs) (Jordan and Humphries, 1994), amplified fragment length polymorphisms (AFLPs) (Vos *et al.,* 1995), and simple sequence repeats (SSRs) (Akkaya *et al.,* 1992;) made it possible to use marker technology extensively in the genome and progeny.

10.3 Types of Molecular Markers:

A. Restriction Fragment Length Polymorphisms (RFLPs):

DNA polymorphism known as RFLP is based on differences in DNA sequence lengths caused by restriction enzymes cleaving genomic DNA at certain recognition sites. To find these differences, this technique uses Southern blotting to hybridize DNA probes. Variation in DNA sequences can be determined by examining the existence or lack of fragments of varying sizes that result from the digestion of DNA sample. Bacterial enzymes called restriction endonucleases are essential for RFLP analysis because they cut DNA at certain recognition sites to produce shorter pieces. For this reason, scientists usually use restriction enzymes that have been industrially isolated. These enzymes can digest double-stranded DNA at particular places that contain sequences that the enzyme recognizes. Different recognition sites are the targets of each restriction endonuclease, usually

B. Random Amplified Polymorphic DNA (RAPD):

With this process, genomic DNA is amplified by PCR (Polymerase Chain Reaction) using a single, randomly selected, short primer (10 nucleotides). When two hybridization sites on the DNA template match and are orientated in opposing directions, amplification takes place during PCR (Jiang 2013). The lengths of the primer that was used and the target genome both affect the amplified fragments that are produced. The primer design is essential to the PCR amplification process's success. It is best to use a primer with at least 40% GC content since primers with lower GC contents might not be able to resist the 72°C annealing temperature needed for DNA polymerase to elongate the DNA (Wolff *et al.,* 1993). The PCR result is separated on an agarose gel stained with DNA dye after amplification.

C. Simple Sequence Repeats (SSRs) or Microsatellites:

SSRs, or simple sequence repeats, are another name for microsatellites.They are short tandem repeats that show basic variations in sequence length. These repetitions are extensively dispersed across the genomes of different species and are made up of motifs with one to six nucleotides. Mononucleotide repeats (like A), dinucleotide repeats (like GT), trinucleotide repeats (like ATT), tetranucleotide repeats (like ATCG), pentanucleotide repeats (like TAATC), and hexanucleotide repeats (like TGTGCA) are a few examples of microsatellites. Although the nuclear genome contains the majority of microsatellites, they are also present in mitochondria, chloroplasts, and protein-coding genes (Tautz 1989) Microsatellites are highly polymorphic, mostly because of differences in the number of repeats in their regions, which makes PCR a simple method of detecting them.

D. Inter-Simple Sequence Repeat (ISSR) Marker:

Semi-arbitrary markers known as inter-simple sequence repeats (ISSRs) are produced by PCR using a single primer that is complementary to a target microsatellite. Another name for amplification in the presence of non-anchored primers is MP-PCR, or microsatellite-primed PCR. According to Zietkiewicz *et al.* (1994), this type of amplification produces multi-locus, highly polymorphic patterns without requiring knowledge of the genome sequence. A DNA sequence that is separated by two inverted microsatellites is represented by each band. Similar to RAPDs, ISSR markers are fast and simple to employ, however due to their longer primer length, they appear to have a repeatability issue (Gupta *et al.*, 2004).

E. Single-Nucleotide Polymorphism (SNP):

SNPs are single base-pair variations that can be found in an individual's genetic sequence. Depending on which nucleotide is substituted, SNPs can be transversions (C/G, A/T, C/A, or T/G) or transitions (C/T or G/A). Single base alterations, such as SNPs that are insertions or deletions (InDel) in a single base, are typically seen in mRNA. The lowest unit of heredity is a single nucleotide base, and SNP may provide the greatest number of markers in an easy-to-understand manner. Both plants and animals have large numbers of SNPs; in plants, there is one SNP for every 100–300 bp. SNPs are extensively dispersed across the genome and can be found in either the coding or non-coding portions of genes, as well as in the intergenic region—the area between two genes—with varying rates (Sobrino *et al.,* 2005). Many platforms for allelic discrimination and detection have formed the foundation for the development of numerous SNP genotyping technologies. In 2005, Sobrino *et al.* SNPs are the most desirable markers for genotyping due to some new high-throughput genotyping techniques, including allele-specific PCR, NGS, GBS, and chip-based NGS (Agarwal *et al.,* 2008).

F. Retrotransposon-Based Insertion Polymorphism (RBIP):

Flavell *et al.* 1998 developed a method to detect retrotransposon sequences as molecular markers. This method amplifies DNA using a primer, requires sequence information, and uses agarose gel electrophoresis for polymorphism detection. High-throughput retrotransposon-based markers are employed using fluorescent microarray markers.

G. SCAR (Sequence-Characterized Amplified Regions):

In 1993, Paran and Michelmore created the first sequence-characterized amplified region (SCAR) markers in lettuce for downy mildew resistance genes (Paran and Michelmore 1993). Compared to RAPD, SCAR markers are more repeatable and specific (Yang *et al.,* 1993). Co-dominant and mono-locus markers, or SCAR markers, are mostly used in physical mapping. Purification of PCR fragments and SCAR primer design are the first steps in the process of creating SCAR markers (Kiran *et al.,* 2010). Agarose gel is used to identify polymorphic bands, after which the nucleotide sequence of the chosen DNA fragment is examined. By comparing the polymorphic DNA sequence to the database of known DNA sequences maintained by the NCBI.

10.4 Techniques for Molecular Marker Analysis:

Molecular marker analysis is a key technique in molecular biology used to detect variations in DNA sequences among individuals or populations. These markers are essentially genetic variations that can be used to identify individuals, track genetic traits, or analyze evolutionary relationships. Here are some common techniques used in molecular marker analysis:

10.4.1 Polymerase Chain Reaction (PCR):

Polymerase Chain Reaction (PCR) is a powerful molecular biology technique used to amplify a specific segment of DNA, making millions of copies of that segment in a relatively short amount of time. Developed in the 1980s by Kary Mullis, PCR has become an indispensable tool in various fields of biology, including genetics, forensics, medicine, and biotechnology (Mullis *et al.*, 1986).

A. Principle and applications in marker analysis:

PCR is a widely used molecular biology technique that amplifies a specific segment of DNA copies of a target DNA sequence, making it easier to study and analyze through repeated cycles of heating and cooling. It allows researchers to make millions of DNA. The working principle of PCR is as follows:

- Denaturation: First, a high temperature (usually about 95°C) is applied to the DNA sample, which splits the double-stranded DNA into two single strands. Denaturation is the term for this process (Mullis and Faloona, 1987).
- Annealing: Short DNA primers are then able to attach to complementary sequences on each single-stranded DNA template by lowering the temperature. The synthesis of DNA begins with these primers (Mullis and Faloona, 1987).
- Extension: DNA polymerase, an enzyme that synthesizes new DNA strands, extends the primers by adding complementary nucleotides along the template strands. The temperature is raised again to optimize the activity of the DNA polymerase, typically around 72°C (Mullis and Faloona, 1987).

B. Application in Marker Analysis:

- **Genetic Marker Analysis:** PCR is used to amplify genetic markers such as microsatellites, SNPs, and indels, allowing for the study of genetic variation within and between populations (Bostein *et al.*, 1980).
- **Genotyping:** PCR-based genotyping assays determine the presence or absence of specific alleles at genetic loci, aiding in disease association studies and forensic analysis (Hacia, 1999).
- **Molecular Evolutionary Studies:** PCR amplification of gene regions enables the study of genetic diversity, population structure, and evolutionary relationships among species (Avise *et al.*, 1987).

- **Forensic DNA Analysis:** PCR amplifies DNA markers from crime scene samples, allowing for the identification of suspects or victims based on their DNA profiles (Gill *et al.*, 1985).
- **Disease Diagnosis:** PCR-based assays detect pathogens, mutations, or genetic biomarkers associated with diseases such as cancer, aiding in early diagnosis and treatment (kwok *et al.*, 1990).
- **Environmental DNA (eDNA) Analysis:** PCR detects DNA traces shed by organisms into the environment, allowing for species presence monitoring and biodiversity assessment (Dejean *et al.*, 2011).
- **Gene Expression Analysis:** RT-PCR amplifies cDNA synthesized from RNA, allowing for the quantification of gene expression levels in various tissues or cell types (Bustin *et al.*, 2009). They offer insights into the development, optimization, and utilization of PCR techniques in genetic research, diagnostics, and forensic science.

10.4.2 Gel Electrophoresis:

Gel electrophoresis is a fundamental technique used in molecular biology to separate and analyze macromolecules such as DNA, RNA, and proteins based on their size and charge.

In this method, molecules are forced to migrate through a gel matrix under the influence of an electric field.

The rate of migration is primarily determined by the size and charge of the molecules, allowing for their separation into distinct bands. (Laemmli 1070).

10.4.3 Sequencing Technologies:

Sequencing technologies, also known as DNA sequencing methods, are instrumental in determining the precise order of nucleotides in a DNA molecule. Since the development of the first DNA sequencing method in the 1970s, sequencing technologies have evolved significantly, becoming faster, more accurate, and more cost-effective.

A. Next-Generation Sequencing (NGS):

NGS technologies, developed in the 2000s, revolutionized DNA sequencing by massively parallelizing the sequencing process, allowing for high-throughput sequencing of millions of DNA fragments simultaneously.

Various NGS platforms, such as Illumina, Ion Torrent, and Pacific Biosciences, employ different sequencing chemistries and technologies, including sequencing-by-synthesis, ion semiconductor sequencing, and single-molecule real-time (SMRT) sequencing. NGS has significantly reduced the cost and time required for sequencing, making large-scale genomic projects and personalized medicine more accessible (Metzker, 2010).

Principle: NGS platforms use massively parallel sequencing to simultaneously sequence millions of DNA fragments in a single reaction.

Applications:

- Whole Genome Sequencing (WGS)
- Exome Sequencing
- Transcriptome Sequencing (RNA-Seq)
- ChIP-Seq and DNA Methylation Sequencing
- Metagenomic Sequencing
- NGS Platforms
- Illumina Sequencing
- Ion Torrent Sequencing
- Pacific Biosciences (PacBio) Sequencing
- Oxford Nanopore Sequencing

B. Sangar Sequencing:

Sanger sequencing, which Frederick Sanger invented in 1977, is still regarded as one of the best methods for sequencing DNA. DNA strands are synthesized for Sanger sequencing in the presence of modified nucleotides called deoxynucleotides, which stop DNA synthesis when they are absorbed into the developing strand. Using gel electrophoresis to segregate the terminated fragments by size, the sequence is ascertained by identifying the fluorescent signals that correspond to each nucleotide (Sanger *et al.*, 1977).

Principle: The selective insertion of chain-terminating dideoxynucleotides (ddNTPs) during DNA synthesis is the foundation of Sanger sequencing. When single-stranded DNA templates are present together with modest quantities of fluorescently labeled deoxynucleotides (ddNTPs), DNA polymerase stretches a primer that has annealed to them. This process is known as Sanger sequencing. Since a ddNTP lacks the 3'-OH group required for further extension, its insertion stops DNA synthesis, resulting in DNA fragments of different lengths terminating at each point where a ddNTP is inserted. Gel electrophoresis is used to separate the terminated DNA fragments based on size, and each nucleotide's fluorescent signal is detected to identify the sequence (Sanger *et al.*, 1977).

Applications:

- DNA Sequencing
- Mutation Analysis
- Genetic Engineering

Advantages:

Accuracy: Sanger sequencing provides high accuracy, with error rates typically less than 1 in 1,000 base pairs.

Read Length: Sanger sequencing can generate long read lengths, up to 1,000 base pairs or more, making it suitable for sequencing individual genes and characterizing genetic variants.

Limitations:

Throughput: Sanger sequencing has limited throughput compared to next-generation sequencing (NGS) technologies, making it less suitable for large-scale sequencing projects.

Cost: Sanger sequencing can be more expensive per base compared to NGS, particularly for projects requiring high coverage or large numbers of samples.

10.5 Application of Molecular Marker in Plant Biotechnology:

10.5.1 Genetic Diversity, DNA Fingerprinting, Germplasm Conservation:

To achieve long-term genetic improvement and quickly adjust to changing breeding objectives, it is essential to evaluate and use genetic variety in breeding endeavors. The commercialization of early and extra-early maturing maize varieties has greatly increased the amount of maize grown in sub-Saharan Africa (SSA) over the past 20 years, enhancing food and nutritional security and making maize a vital part of the region's agricultural economy across a variety of agro-ecologies (Kim *et al.*, 2021). In this work, 439 early and extra-early maize inbred lines produced by the International Institute of Tropical Agriculture Maize Improvement Programme (IITA-MIP) were examined for population structure and genetic diversity. Thirteen narrow-based and twenty-seven broad-based populations gave rise to these inbred lineages. 9642 DArTseq-based single nucleotide polymorphism (SNP) markers, which were uniformly dispersed throughout the maize genome, were used for genotyping (Aparaku Bada *et al.*, 2021).

Using ISSR and RAPD markers, the study evaluated the genetic diversity of 54 castor samples—3 wild and 51 cultivated—from throughout the world. Whereas RAPD analysis revealed 90 bands with 60.98% polymorphism, ISSR analysis provided 83 bands, 74.53% of which were polymorphic. The genotypes were sorted into three groups by the UPGMA dendrogram, which was in line with morphological clustering. The tight relationship between the majority of Chinese types suggests a limited genetic foundation. It is advised to introduce wild castor resources for reproduction. The accessions from Nigeria, Malaysia, Indonesia, and Vietnam were grouped together. The results of PCoA and UPGMA analysis agreed. Future research on castor breeding will benefit greatly from these findings (Kim *et al.*, 2021).

RAPD analysis revealed 466 fragments, with 96.43% polymorphism, averaging 46.6 bands per primer. ISSR primers produced 328 fragments, 91.83% of which were polymorphic, with an average of 49.7 bands per primer. (Mazumdar *et al.*, 2020).

10.5.2 Assessment of Heterosis in Breeding:

Heterosis was observed in all hybrids, with traits like plant height, boll number, and lint percentage showing higher levels than fiber quality traits. Significant correlations were found between parental and F1 performances, with F1 performances varying among hybrid sets.

Genetic distance (GD) between parents was found to predict heterosis for specific traits, suggesting molecular marker analysis can aid in hybrid cotton breeding and guide parental selection (Geng *et al.*, 2021).

The study at Bowen University, Nigeria, assessed hybrid efficiency and genetic variability between white and yellow parental maize varieties. The experiment involved two parental lines and two hybrids, with a randomized block design. Results showed significant variations in traits, with hybrid 1 having the highest number of rows per cob and longest cobs. Both hybrids outperformed their parents, with hybrid 1 showing the highest rates of mid-parent and better-parent heterosis (Esan *et al.*, 2021; Tomkowiak *et al.*, 2020).

This study investigated the genetic distance (GD) between parental genotypes using single nucleotide polymorphism (SNP) DNA markers and its correlation with heterosis in common wheat. A field experiment was conducted with parents and hybrids, assessing various traits like plant height, ear number, grain weight, and yield. Using a wheat 90K SNP chip, GD between parents was determined and its relationship with hybrid performance analyzed. Cluster analysis based on SNP-derived GD categorized elite parents into five groups, aligning closely with parental pedigrees. Significant correlation was found between GD and mid-parent heterosis (MPH) of 1000-grain weight, while no significant correlation was observed between GD and high-parent heterosis (HPH) of 1000-grain weight. Weak correlations were noted between GD and spikelet number, harvested spikes, and yield at MPH or HPH. This suggests SNP analysis can aid in grouping wheat parents for heterosis, though the relationship between SNP-based GD and hybrid performance warrants further investigation (Nie *et al.*, 2019).

10.5.3 Identification of Haploid/Diploid Plants:

In maize breeding, DH technology is essential. The R1-navajo gene functions as a marker, although one that is not always reliable, for choosing seeds containing haploid embryos.

To evaluate genome size at an early stage, we offer a complete technique that combines molecular marker-based genotyping from seedling DNA, flow cytometry of root tip nuclei, and phenotypic screening. Merely 59% of the seeds that were chosen based on color markers were verified to be haploid in a sample experiment. By using the UMC1152 SSR marker, we were able to identify haploid candidates, of which 83% did not exhibit inducer line alleles. The use of colchicine resulted in a 3% frequency of doubled haploids, improving the accuracy of haploid identification in hybrid populations (Radi *et al.*, 2020).

10.5.4 Bulk Segregant Analysis:

Dwarfism, a desirable trait in non-timber woody plants like castor bean, is linked to divergent cell growth in all tissues. Two QTLs associated with plant height were identified, with one QTL revealing a putative IAA transport protein. This study provides insights into the physiological and molecular mechanisms of dwarfing, potentially aiding genetic breeding efforts for castor bean and related crops. (Wang *et al.*, 2021). Identifying cold-tolerance quantitative trait loci (QTL) in rice, particularly during the bud-bursting stage, is essential for breeding varieties resilient to cold environments (Yang *et al.*, 2021).

Drought limits soybean yield in rainfed areas. Researchers developed F2 mapping populations by crossing drought-tolerant and susceptible genotypes. Segregation analysis confirmed a 3:1 ratio of tolerant to susceptible plants in both F2 populations. Satt277, a polymorphic marker, was linked to drought survival. These tolerant genotypes can be used in soybean improvement programs and aid in screening large germplasm collections. Hybrid 1 and hybrid 2 showed superior performance (Sreenivasa *et al.*, 2020).

10.5.5 Marker- Assisted Selection:

Cassava mosaic disease (CMD) poses a significant threat to cassava production in sub-Saharan Africa. Utilizing genomic-assisted selection in seedling trials can expedite breeding programs, reducing time, costs, and resources. Five cassava populations underwent CMD resistance screening at 1-, 3-, and 5-months post-planting, graded on a scale of 1–5. Additionally, genotypes were screened using six molecular markers linked to the CMD2 gene, with correlations between phenotypic and marker data estimated. CMD severity scores revealed 53–82% progeny resistance, averaging 70.5%, while marker data identified around 70% resistance across populations (62–80%). Combining market data and CMD scores, 40–60% of progenies were resistant (average 52%). Correlation coefficients between marker data and CMD scores ranged from 0.2024 to 0.3460, suggesting the involvement of novel genes not linked to markers. Resistant genotypes, potentially carrying other desirable traits, were selected for advanced trial evaluation, expediting breeding program timelines (Olasanmi *et al.*, 2021). Marker-assisted selection (MAS) is used to improve Camellia oleifera yield and fruit quality in China. A three-generational hybridization experiment used SRAP markers to authenticate hybrids and identify inheritance. Genetic diversity increased across generations, and traits showed high broad-sense heritability, indicating potential for hybrid breeding (Feng *et al.*, 2020).

10.5.6 Genetic and Physical Mapping:

Marker-assisted selection (MAS) is applied to enhance the yield and fruit characteristics of Camellia oleifera to meet the growing demand for new varieties in China. We conducted a three-generational (G0, G1, and G2) diallel mating and selection experiment, employing Sequence-Related Amplified Polymorphism (SRAP) markers within a MAS framework. These SRAP markers played a crucial role in ensuring hybrid authenticity and identifying the presence of matrilineal or patrilineal inheritance, aiding in the selection and direction of mating pairs. Over the three generations studied estimates of genetic diversity parameters consistently increased, indicating no reduction in diversity associated with selection. Genetic distance estimates and their correlation with heterosis proved valuable in selecting mating pairs to achieve desired yield and fruit attributes, such as fruit diameter, height, weight, peel thickness, and seed characteristics. Most yield and fruit attributes exhibited high broad-sense heritability, indicating the potential for increased hybrid breeding efficiency in this species over successive generations (Feng *et al.*, 2020).

The study evaluated aphid tolerance in F1 and F6-recombinant inbred lines (RILs) after Uroleucon infestation. Dominance effects were observed in F1 plants, while RILs showed quantitative variation. Genotyping revealed two quantitative trait loci (QTLs) associated with DW after aphid infestation: QUc-Ct3.1 and QUc-Ct5.1.

This marks the first identification of genes linked to safflower aphid tolerance (Jegadeeswaran *et al.*, 2021). The study developed a high-density composite genetic linkage map of A. sellowiana using two populations, H5 and H6 and identified 4,921 SNP markers using the genotyping by sequencing method. The composite map, spanning 1,314 cM, included 1,897 SNP markers. This approach captures meiosis information from individuals of two linked populations in a single estimator, contributing to future genetic research and breeding initiatives (Quezada *et al.*, 2021).

10.5.7 QTL Mapping and Characterization:

This study constructed a genetic linkage map for the Reyan5-11×PR107 population using SSR and AFLP markers. They integrated this map with a previously published one and conducted QTL mapping for rubber yield-related traits. The genetic map comprised 132 markers (88 SSR and 44 AFLP), while the joint map contained 143 markers (99 SSR and 44 AFLP), spanning 19 linkage groups and covering 1298.5 cM of the rubber tree genome. Each linkage group contained 2–40 markers, with an average distance of 9.08 cm between them. The study identified 2 QTLs for girth, 3 for dry rubber content, 3 for dry rubber yield per tree, 4 for unit secant length dry rubber yield, 5 for the number of conducting phloem laticifer, and 4 for the number of yellow phloem laticifer under LOD≥2.5 conditions (Hou *et al.*, 2021).

Researchers created four F2 populations of upland cotton, combining normal lines with introgression lines for better fiber quality. They analyzed eight agronomic and economic traits, identifying fifty quantitative trait loci (QTLs) for plant height, fiber quality, and yield. Some QTLs were consistent across populations, providing insights into improving yield and fiber quality in upland cotton breeding (Li *et al.*, 2021). The study evaluates genetic relationships among Capsicum annuum and Capsicum frutescens varieties in West Africa using 10 ISSR primers. The results show 75 loci, with 14 polymorphic. The study also reveals genetic relatedness among the varieties, with C. frutescens var. baccatum not separated from C. annuum. The findings suggest that C. frutescens var. baccatum could be considered a variety of the C. annuum species (Olantunji and Afalayan, 2019).

10.5.8 Association Mapping:

The goal of this research was to strengthen chickpea resistance to Fusarium wilt (FW), which is essential to maintaining chickpea output worldwide. Over two years, 75 FW-responsive chickpea breeding lines and control samples were examined in an area impacted by wilt. A total of 267 alleles, with an average of 3.56 alleles per marker, were found by genetic diversity analysis utilizing 75 simple sequence repeat (SSR) markers. Based on clustering and factorial analysis, the samples were grouped into two primary classes. Using STRUCTURE analysis, the samples were further separated into three unique sub-groups (Jha *et al.* (2021).

Our study utilized association mapping to analyze 200 SSR markers and 8 seed traits across 10 near-isogenic lines (NILs) of waxy maize and parental lines (HW3 and HW9) of the "Mibaek 2" variety. By comparing allele frequencies, we identified 32 SSR markers linked to seed trait variations.

Notably, five SSR markers were associated with multiple traits like elongation, kernel weight, coat color, roundness, lightness, and volume. Population structure analysis divided the lines into two groups, highlighting genetic diversity. This underscores SSR analysis's effectiveness in understanding genetic diversity and conducting association mapping for traits (Kim *et al.*, 2021).

10.5.9 Marker-Assisted Back Cross Breeding:

Molecular-assisted backcrossing (MABC) was employed to incorporate drought tolerance, Striga, and root-knot nematode resistance QTLs into a popular cowpea landrace used for intercropping in Burkina Faso. Two donor lines with specific resistances were crossed with the recurrent parent, Moussa Local. Genotyping with 184 SNP markers, including trait-associated and background markers, aided in selecting BC1F1 plants with desired traits.

This selection process was repeated, resulting in BC3F1 families for each donor. The top-performing families were selfed to produce BC3F2 seeds for further testing. This approach effectively enhanced the cowpea variety's resilience to drought and biotic stresses, showcasing the utility of SNP markers in MABC for crop improvement (Batieno *et al.*, 2016).

The study aimed to enhance rice production in Malaysia by developing high-yielding fragrant rice varieties, which are often imported due to domestic demand. Using a marker-assisted backcrossing (MABC) approach, the fragrance gene (fgr) was introduced into two Malaysian varieties, MR84 and MR219, in less than two years and with fewer than fifty molecular markers. (Cheng *et al.*, 2017).

10.5.10 Hybrid Identification:

In this work, seven genotypes of E. sibiricus with different seed-shattering properties were crossed to create five hybrid populations. We used DNA fingerprinting to identify hybrids and evaluate genetic diversity between hybrids and their parents using nine EST-SSR primers. Fifteen hybrids were found, and two primary categories of markers were noted: those that were shared by the hybrids and both parents, and those that were shared solely by the hybrids and the male parents.

The amplified bands in each hybrid population ranged from 37 to 57, and the polymorphism percentages ranged from 65.12% to 75.68%, with an average of 70.51%. In every hybrid population, we found new bands that suggested genetic rearrangements. Principal coordinate analysis (PCoA) produced results that were consistent with the structural analysis that divided all hybrid populations and parents into eight categories. Overall, our research shows that EST-SSR markers are useful for identifying hybrids, and it raises the possibility of using crossbreeding to increase the genetic diversity of hybrid populations (Zhao et al., 2017). This work used SSR molecular markers, a technique not as often used in Mexican lime breeding programs, to identify hybrids produced by controlled pollination between C. aurantifolia var. "Colimex" and C. limon var. "Rosenberg" and their reciprocal (Guzman et al., 2017).

10.6 References:

1. Agarwal, M., Shrivastava, N., & Padh, H. (2008). Advances in molecular marker techniques and their applications in plant sciences. Plant Cell Reports, 27(4), 617–631. https://doi.org/10.1007/s00299-008-0507-z
2. Akkaya, M. S., Bhagwat, A. A., & Cregan, P. B. (1992). Length polymorphisms of simple sequence repeat DNA in soybean. Genetics, 132(4), 1131-1139.
3. https://doi.org/10.1093/genetics/132.4.1131
4. Asif, A. A., Garcia, M., Tilbrook, J., Brien, C., & Dowling, K. (2020). Identification of salt tolerance QTL in wheat RIL mapping population using destructive and nondestructive phenotyping. Functional Plant Biology, 48(2), 131–140. https://doi.org/10.1071/FP20042
5. Avise, J. C., Arnold, J., Ball, R. M., Bermingham, E., Lamb, T., Neigel, J. E., & Saunders, N. C. (1987). Intraspecific phylogeography: the mitochondrial DNA bridge between population genetics and systematics. Annual Review of Ecology and Systematics, 18(1), 489–522. https://doi.org/10.1146/annurev.es.18.110187.002421
6. Badu-Apraku, B., Garcia-Oliveira, A. L., Petroli, C. D., Hearne, S., Adewale, S. A., & Gedil, M. (2021). Genetic diversity and population structure of early and extra-early maturing maize germplasm adapted to sub-Saharan Africa. BMC Plant Biology, 21(1), 1-15. https://doi.org/10.1186/s12870-020-02741-y
7. Batieno, B. J., Danquah, E., Tignegre, J., Huynh, B., Drabo, I., Close, T. J., Ofori, K., Roberts, P., & Ouedraogo, T. J. (2016). Application of marker-assisted backcrossing to improve cowpea (Vigna unguiculata L. Walp) for drought tolerance. Journal of Plant Breeding and Crop Science, 8(12), 273–286. https://doi.org/10.5897/JPBCS2016.0606
8. Beckmann, J. S., & Weber, J. L. (1992). Survey of human and rat microsatellites. Genomics, 12(4), 627–631. https://doi.org/10.1016/0888-7543(92)90437-Z
9. Bhagyawant, S. S. (2016). RAPD-SCAR markers: An interface tool for authentication of traits. Journal of Biosciences and Medicines, 4(1), 1–9. https://doi.org/10.4236/jbm.2016.41001
10. Botstein, D., White, R. L., Skolnick, M., & Davis, R. W. (1980). Construction of a genetic linkage map in man using restriction fragment length polymorphisms. American Journal of Human Genetics, 32(3), 314–331. https://doi.org/10.1016/S0002-9297(80)80132-0
11. Bustin, S. A., Benes, V., Garson, J. A., Hellemans, J., Huggett, J., Kubista, M., & Wittwer, C. T. (2009). The MIQE guidelines: Minimum information for publication of quantitative real-time PCR experiments. Clinical Chemistry, 55(4), 611–622. https://doi.org/10.1373/clinchem.2008.112797
12. Cheng, J. L., Li, J., and Qi, Y. M., (2016). Development of novel SCAR markers for genetic characterization of Lonicera japonica from high GC-RAMP-PCR and DNA cloning. Genetics and Molecular Research, 15(10), 4238. https://doi.org/10.4238/gmr.15038238
13. Collard, B. C. Y., Jahufer, M. Z. Z., and Brouwer, J. B., (2005). An introduction to markers, quantitative trait loci (QTL) mapping and marker-assisted selection for crop improvement: The basic concepts. Euphytica, 142(1-2), 169–196. https://doi.org/10.1007/s10681-005-1681-5

14. Dejean, T., Valentini, A., Duparc, A., Pellier-Cuit, S., Pompanon, F., Taberlet, P., & Miaud, C. (2011). Persistence of environmental DNA in freshwater ecosystems. PLoS ONE, 6(8), e23398. https://doi.org/10.1371/journal.pone.0023398
15. Esan, V. I., Omilani, O. O., Olajide, A. A., Adebayo, R., & Sangoyomi, T. E. (2021). Genetic diversity, advance, heritability and heterosis estimation in F1 hybrids and parental lines of maize. Agriculture & Food Security, 10(1), 1-12. https://doi.org/10.1186/s40066-021-00320-5
16. Feng, J., Jiang, Y., Yang, Z., Chen, S., El-Kassaby, Y. A., & Chen, H. (2020). Marker-assisted selection in hybrid population. Silvae Genetica, 69(1), 63–72. https://doi.org/10.2478/sg-2020-0008
17. Flavell, A. J., Knox, M. R., Pearce, S. R., & Ellis, T. H. N. (1998). Retrotransposon-based insertion polymorphisms (RBIP) for high throughput marker analysis. The Plant Journal, 16(5), 643–650. https://doi.org/10.1046/j.1365-313x.1998.00356.x
18. Geng, X., Qu, Y., Jia, Y., et al. (2021). Assessment of heterosis based on parental genetic distance estimated with SSR and SNP markers in upland cotton (Gossypium hirsutum L.). BMC Genomics, 22, 123. https://doi.org/10.1186/s12864-021-07447-9
19. Gill, P., Jeffreys, A. J., Werrett, D. J., & Budowle, B. (1985). An investigation of the rigor of interpretation rules for STRs derived from less than 100 pg of DNA. Forensic Science International, 56(2), 135–149. https://doi.org/10.1016/0379-0738(85)90046-4
20. Gupta, M., Chyi, Y. S., Romero-Severson, J., & Owen, J. L. (1994). Amplification of DNA markers from evolutionarily diverse genomes using single primers of simple-sequence repeats. Theoretical and Applied Genetics, 89(7-8), 998–1006. https://doi.org/10.1007/BF00224530
21. Guzman, M. J. L. F., Rodríguez, L. F. G. P. G., Mariscal, P. A. P. S., & Santos, M. O. (2017). Identifying hybrids of Citrus aurantifolia and Citrus limon using simple sequence repeats (SSR) markers. Revista Mexicana de Ciencias Agrícolas, 8(6), 1369-1380. https://doi.org/10.29312/remexca.v8i6.727
22. Hacia, J. G. (1999). Resequencing and mutational analysis using oligonucleotide microarrays. Nature Genetics, 21(1 Suppl), 42–47. https://doi.org/10.1038/4475
23. Hou, B., Li, W., Gao, X., Zhang, X., Feng, S., & Wu, Y. (2021). Construction and integration of genetic map and quantitative trait locus (QTL) mapping of yield-related traits. Acta Agriculturae Scandinavica, Section B - Soil & Plant Science. https://doi.org/10.1080/09064710.2021.1886554
24. Igwe, D. O., Ihearahu, O. C., Osano, A. A., et al. (2021). Genetic diversity and population assessment of Musa L. (Musaceae) employing CDDP markers. Plant Molecular Biology Reporter. https://doi.org/10.1007/s11105-021-01264-2
25. Jegadeeswaran, M., Kadirvel, P., Srinivas, P. S., Senthilvel, S., Selvaraj, V. M., et al. (2021). Genetic mapping reveals a major QTL associated with tolerance to the aphid, Uroleucon compositae (Theobald) in safflower (Carthamus tinctorius). Plant Breeding. https://doi.org/10.1111/pbr.12958
26. Jha, U., Jha, R., Bohra, A., Manjunatha, L., Saabale, P., Parida, S., & Singh, N. (2021). Association mapping of genomic loci linked with Fusarium wilt resistance (Foc2) in chickpea. Plant Genetic Resources: Characterization and Utilization, 1–8. https://doi.org/10.1017/S1479262121000115
27. Jiang, G. L. (2013). Molecular markers and marker-assisted breeding in plants. In S. B. Andersen (Ed.), Plant breeding from laboratories to fields (pp. 45–83). InTech. https://doi.org/10.5772/52583

28. Jing, R., Bolshakov, V., & Flavell, A. J. (2007). The tagged microarray marker (TAM) method for high-throughput detection of single nucleotide and indel polymorphisms. Nature Protocols, 2(1), 168–177. https://doi.org/10.1038/nprot.2007.12
29. Jones, C. J., Edwards, K. J., Castaglione, S., et al. (1997). Reproducibility testing of RAPD, AFLP, and SSR markers in plants by a network of European laboratories. Molecular Breeding, 3(5), 381–390. https://doi.org/10.1023/A:1009639529013
30. Jordan, S. A., & Humphries, P. (1994). Single nucleotide polymorphism in exon 2 of the BCP gene on 7q31-q35. Human Molecular Genetics, 3(10), 1915.
31. https://doi.org/10.1093/hmg/3.10.1915
32. Kim, H., Lei, P., Wang, A., Liu, S., Zhao, Y., Huang, F., Yu, Z., Zhu, G., He, Z., Tan, D., et al. (2021). Genetic diversity of Castor bean (Ricinus communis L.) revealed by ISSR and RAPD markers. Genetics and Molecular Research, 11, 457. https://doi.org/10.4238/gmr.457
33. Kim, H. R., Sa, K. J., Nam-Gung, M., Park, K. J., Ryu, S.-H., et al. (2021). Genetic characterization and association mapping in near-isogenic lines of waxy maize using seed characteristics and SSR markers. Genes & Genomics, 43(1), 79–90. https://doi.org/10.1007/s13258-020-00992-2
34. Kiran, U., Khan, S., Mirza, K. J., & Ram, M. (2010). SCAR markers: A potential tool for authentication of herbal drugs. Fitoterapia, 81(8), 969–976. https://doi.org/10.1016/j.fitote.2010.07.016
35. Kordrostami, M., & Rahimi, M. (2015). Molecular markers in plants: Concepts and applications. In Conference on Genetics in the Third Millennium (Vol. 13, pp. 4024–4031).
36. Kwok, S., Kellogg, D. E., McKinney, N., Spasic, D., Goda, L., Levenson, C., & Sninsky, J. J. (1990). Effects of primer-template mismatches on the polymerase chain reaction: Human immunodeficiency virus type 1 model studies. Nucleic Acids Research, 18(4), 999–1005. https://doi.org/10.1093/nar/18.4.999
37. Laemmli, U. K. (1970). Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature, 227(5259), 680–685. https://doi.org/10.1038/227680a0
38. Lopez-Malvar, A., Butron, A., Malvar, R. A., et al. (2021). Association mapping for maize stover yield and saccharification efficiency using a multiparent advanced generation intercross (MAGIC) population. Scientific Reports, 11, 3425. https://doi.org/10.1038/s41598-021-83000-7
39. Madhumati, B. (2014). Potential and application of molecular markers techniques for plant genome analysis. International Journal of Pure & Applied Bioscience, 2(1), 169–188. https://doi.org/10.18782/2320-7051.6243
40. Mazumdar, S. R., Hoque, H., Sinha, B., Chowdhury, W. R., Hasan, M. N., et al. (2020). Genetic variability analysis of partially salt tolerant local and inbred rice (Oryza sativa L.) through molecular markers. Heliyon, 6, e04333. https://doi.org/10.1016/j.heliyon.2020.e04333
41. Metzker, M. L. (2010). Sequencing technologies - the next generation. Nature Reviews Genetics, 11(1), 31–46. https://doi.org/10.1038/nrg2626
42. Mullis, K., Faloona, F., Scharf, S., Saiki, R., Horn, G., & Erlich, H. (1986). The polymerase chain reaction: An enzymatic approach to DNA replication. Cold Spring Harbor Symposia on Quantitative Biology, 51(Pt 1), 263–273. https://doi.org/10.1101/SQB.1986.051.01.032

43. Nadeem, M. A., Nawaz, M. A., Shahid, M. Q., et al. (2018). DNA molecular markers in plant breeding: Current status and recent advancements in genomic selection and genome editing. Biotechnology & Biotechnological Equipment, 32(2), 261–285. https://doi.org/10.1080/13102818.2018.1426447
44. Nie, Y., Ji, W., & Ma, S. (2019). Assessment of heterosis based on genetic distance estimated using SNP in common wheat. Agronomy, 9(2), 66. https://doi.org/10.3390/agronomy9020066.
45. Olatunji, T. L., & Afalayan, A. J. (2019). Evaluation of genetic relationship among varieties of Capsicum annuum L. and Capsicum frutescens L. in West Africa using ISSR markers. Heliyon, 5, e01700. https://doi.org/10.1016/j.heliyon.2019.e01700
46. Paran, I., & Michelmore, R. W. (1993). Development of reliable PCR-based markers linked to downy mildew resistance genes in lettuce. Theoretical and Applied Genetics, 85(8), 985–993. https://doi.org/10.1007/BF00215038
47. Poczai, P., Varga, I., Laos, M., et al. (2013). Advances in plant gene-targeted and functional markers: A review. Plant Methods, 9, 6. https://doi.org/10.1186/1746-4811-9-6
48. Provan, J., Powell, W., & Hollingsworth, P. M. (2001). Chloroplast microsatellites: New tools for studies in plant ecology and evolution. Trends in Ecology & Evolution, 16(3), 142–147. https://doi.org/10.1016/S0169-5347(00)02097-8
49. Quezada, M., Amadeu, R. R., Vignale, B., Cabrera, D., Pritsch, C., & Garcia, A. A. F. (2021). Construction of a high-density genetic map of Acca sellowiana (Berg.) Burret, an outcrossing species, based on two connected mapping populations. Frontiers in Plant Science.
50. Radi, F., Torok, K., Nagymihaly, M., et al. (2020). Improved reliability in production of maize inbred lines by the combination of the R1-Navajo marker with flow cytometry or microsatellite genotyping. Cereal Research Communications, 48, 423–430.
51. Sanger, F., Air, G. M., Barrell, B. G., Brown, N. L., Coulson, A. R., & Fiddes, J. C. (1977). Nucleotide sequence of bacteriophage ΦX174 DNA. Nature, 265(5596), 687–695. https://doi.org/10.1038/265687a0
52. Sanger, F., Nicklen, S., & Coulson, A. R. (1977). DNA sequencing with chain-terminating inhibitors. Proceedings of the National Academy of Sciences, 74(12), 5463–5467. https://doi.org/10.1073/pnas.74.12.5463
53. Shendure, J., & Ji, H. (2008). Next-generation DNA sequencing. Nature Biotechnology, 26(10), 1135–1145. https://doi.org/10.1038/nbt1486
54. Sobrino, B., Brión, M., & Carracedo, A. (2005). SNPs in forensic genetics: A review on SNP typing methodologies. Forensic Science International, 154(2-3), 181-194.
55. https://doi.org/10.1016/j.forsciint.2004.09.084
56. Sreenivasa, V., Lal, S., Kiran Babu, P., Mahadeva Swamy, H., Yadav, R., Talukdar, A., & Rathod, D. (2020). Inheritance and mapping of drought tolerance in soybean at seedling stage using bulked segregant analysis. Plant Genetic Resources: Characterization and Utilization, 18(2), 63–70. https://doi.org/10.1017/S1479262119000388.
57. Tautz, D. (1989). Hypervariability of simple sequences as a general for polymorphic DNA markers. Nucleic Acids Research, 17(16), 6463–6471. https://doi.org/10.1093/nar/17.16.6463

58. Tilwari, A., & Sharma, R. (2021). Random amplified polymorphic DNA and inter simple sequence repeat markers reveal genetic diversity between micropropagated, wild and field-cultivated genotypes of Gloriosa superba: An endangered medicinal plant. Molecular Biology Reports, 48, 2437–2452. https://doi.org/10.1007/s11033-021-06257-4
59. Tomkowiak, A., Bocianowski, J., Kwiatek, M., & Kowalczewski, P. L. (2020). Dependence of using SNP in common wheat. Agronomy, 9(2), 66. https://doi.org/10.3390/agronomy9020066
60. Vos, P., Hogers, R., Bleeker, M., Reijans, M., Lee, T. V. D., Hornes, M., Friters, A., Pot, J., Paleman, J., Kuiper, M., & Zabeau, M. (1995). AFLP: A new technique for DNA fingerprinting. Nucleic Acids Research, 23(21), 4407–4414. https://doi.org/10.1093/nar/23.21.4407
61. Wang, Z., Yu, A., Li, F., et al. (2021). Bulked segregant analysis reveals candidate genes responsible for dwarf formation in woody oilseed crop castor bean. Scientific Reports, 11, 6277. https://doi.org/10.1038/s41598-021-85462-5
62. Welsh, J., & McClelland, M. (1990). Fingerprinting genomes using PCR with arbitrary primers. Nucleic Acids Research, 18(24), 7213-7218. https://doi.org/10.1093/nar/18.24.7213
63. Williams, J. G., Kubelik, A. R., Livak, K. J., Rafalski, J. A., & Tingey, S. V. (1990). DNA polymorphisms amplified by arbitrary primers are useful as genetic markers. Nucleic Acids Research, 18(22), 6531-6535. https://doi.org/10.1093/nar/18.22.6531
64. Wolff, K., Schoen, E. D., & Peters-Van Rijn, J. (1993). Optimizing the generation of random amplified polymorphic DNAs in chrysanthemum. Theoretical and Applied Genetics, 86(8), 1033-1037. https://doi.org/10.1007/BF00223889
65. Yang, L., Fu, S., Khan, A., et al. (2013). Molecular cloning and development of RAPD-SCAR markers for Dimocarpus longan variety authentication. SpringerPlus, 2, 501.
66. https://doi.org/10.1186/2193-1801-2-501
67. Yang, L., Lei, L., Li, P., Wang, J., Wang, C., et al. (2021). Identification of candidate genes conferring cold tolerance to rice (Oryza sativa L.) at the bud-bursting stage using bulk segregant analysis sequencing and linkage mapping. Frontiers in Plant Science.
68. Zhao, X., Zhang, J., Zhang, Z., et al. (2017). Hybrid identification and genetic variation of Elymus sibiricus hybrid populations using EST-SSR markers. Hereditas, 154, 15. https://doi.org/10.1186/s41065-017-0038-5
69. Zietkiewicz, E., Rafalski, A., & Labuda, D. (1994). Genome fingerprinting by simple sequence repeat (SSR)-anchored polymerase chain reaction amplification. Genomics, 20(2), 176–183. https://doi.org/10.1006/geno.1994.1151