The Role of Molecular Markers in Crop Improvement and Plant Breeding Programs: A Review

Getu Hailu and Yirgashewa Asfere
College of Natural and Computational Sciences, Department of Biotechnology, Wolkite University, Wolkite, Ethiopia

Key words: RFLP, AFLP, RAPD, SNP, SSR, ISSR, Polymorphism

Corresponding Author:
Getu Hailu
College of Natural and Computational Sciences, Department of Biotechnology, Wolkite University, Wolkite, Ethiopia

Page No.: 171-175
Volume: 15, Issue 6, 2020
ISSN: 1816-9155
Agricultural Journal
Copy Right: Medwell Publications

Abstract: Using the latest achievements in fields of genetics, molecular biology and biotechnology, breeders can create new varieties with improved useful traits. Introduction of DNA markers, especially those based on the Polymerase Chain Reaction (PCR) has led to breakthrough in the plants genetic research, molecular markers are tags that can be used to identify specific genes and locate them in relation to other genes. Now days, molecular markers are the markers of attention for advanced research. These can be classified into two categories including Hybridizations based techniques and PCR-based techniques. Restriction Fragment Length Polymorphism (RFLP) represents the hybridization based marker while PCR dependent includes more reliable and advanced polymorphic markers like Amplified Fragment Length Polymorphism (AFLP), Random Amplified Polymorphic DNA (RAPD), Inter-Simple Sequence Repeats (ISSR), Single Nucleotide Polymorphism (SNP) techniques have been developed.

INTRODUCTION

Agricultural productivity faces the problem of increasing demand from an expanding population, coupled to threats of climate change. Molecular markers are a promising technology to increase the selection efficiency of superior individuals in breeding populations which will accelerate the rate of yield improvement.

Unlike conventional breeding, molecular breeding using very detailed genetic maps and better understanding of the molecular structure of a plant's DNA, breeders can analyze only a tiny bit of plant tissue, even from a newly germinated seedling. Once the tissue is analyzed, researchers know whether that seedling contains the appropriate gene. If it doesn't, they can quickly move on and concentrate on analysis of another seedling, eventually working only with the plants which contain a specific trait. Researchers don't need to wait until the plant reaches sexual maturity. DNA isolation can be carried out at juvenile stage of plant development.

Molecular Markers are identifiable sequences of DNA that are found at specific locations of the genome and are transmitted by the standard laws of inheritance from one generation to the next. Genetic markers are closely related with the gene of interest and they act as sign or flags[1]. Molecular markers are tools used to study diversity at DNA level (polymorphism) and help breeders to identify specific chromosome segments that contain genes of interest. The present review aims to describe different marker systems which are employed in plant identification, crop improvement, genome analysis, phylogenetic and population diversity studies.

Cytological markers: In cytology, the structural features of chromosomes can be shown by Chromosome karyo type and bands. Markers that are related with variations present in the numbers, banding patterns, size, shape, order and position of chromosomes are known as cytological markers. These variations reveal differences in the distributions of euchromatin and heterochromatin. For example bands are produced by Giemsastain bands are produced by quinacrine hydrochloride and R bands are the reversed G bands. These chromosome landmarks can be used in the differentiation of normal and mutated chromosomes. Such markers can also be used in the identification of linkage groups and in physical mapping[2].

Protein (biochemical) markers: Protein markers (seed storage proteins and isozymes) are generated through electrophoresis, taking advantage of the migrational properties of proteins and enzymes and revealed by histochemical stains specific to the enzymes being assayed. Isozymes, variants of specific enzymes that retained their activity but had a slightly different amino acid composition that modified their shape or electric charge.

In such analysis, a tissue extract is prepared and proteins of the extract are separated according to their net charge and size by electrophoresis using a polyacrylamide or starch gel. The gel is stained for a particular enzyme by adding a substrate and a dye under appropriate reaction conditions, resulting in band(s) at position where the enzyme polypeptide has migrated showing relative enzyme activity[3].

Types of molecular maker: Molecular markers are broadly divided into three groups:

- Hybridization based, e.g., RFLPs
- PCR techniques based, e.g., RAPD, AFLP, microsatellites or SSR
- Single Nucleotide Polymorphisms (SNPs)

Molecular markers can also be classified based on mode of inheritance (genomic inheritance and cytoplasmic inheritance) and based on mode of gene action (dominant or codominant markers). A large variety of molecular markers are available presently but choice of suitable markers to attain objectives is necessary. Ideal markers must show high level of polymorphism, co-dominant and ease of allele detection.

The molecular marker technique has been widely used for characterization of cotton, maize and *Brassica napus* and several other species[4-7].

Properties of ideal DNA markers:

- Level of polymorphism. Ideally, the marker should be highly polymorphic in breeding material (i.e., it should discriminate between different genotypes)
- Not subject to environmental influences. The inference of a marker's genotype should be independent of the environment in which the individual lives or its developmental stage
- Occurrence and uniform distribution throughout the genome (ubiquitous distribution in the genome)
- Required small amounts of leaves and DNA sample
- Require no prior information about the genome
- It should have repeatability/reproducibility of results
- Cheap, simple and quick assay
- Co-dominant (able to differentiate homozygous from heterozygous individual)

A breeder has to make a decision for choice of a good molecular marker that meets most of the requirements according to the conditions and resources available for the breeding programme[8].

Description and types of genetic markers

Restriction fragment length polymorphism: The RFLPs are simply inherited naturally occurring Mendelian characters. Genetic information is stored in the DNA sequence on a chromosome and variation in this sequence is the basis for the genetic diversity within species. RFLP makes use of restriction endonucleases. These are enzymes which recognize and cut specific nucleotide sequences in DNA. Plants are able to replicate their DNA with high accuracy and rapidity but many mechanisms causing changes (mutation) in the DNA are operative. This leads to simple or large-base pair changes as a result of inversion, translocation, transpositions or deletion which may occur, resulting in a loss or gain of a recognition sites and in turn lead to restriction fragment of different lengths. This marker was first reported by Botstein etc. in the detection of DNA polymorphism[9]. The separated DNA fragments are transferred to nitrocellulose membrane by Southern blot technique[10].

The restriction enzymes will not cut the fragment if a single base-pair variation occurs in the recognition site. However, if this point mutation occurs in one chromosome but not the other, it is called heterozygous for the marker as both bands are present[11].

RFLP marker is very efficient in rice and can detect a high degree of polymorphism is very suitable for the estimation of genetic diversity among highly similar cultivars of rice[12]. Molecular map of the cotton genome was first constructed using 705 RFLP loci and partitioned into 41 linkage groups[13] the utility of RFLP markers in Marker Assisted Selection (MAS) of cotton is reported and RFLP linked to resistance allele for pathogen of bacterial blight was validated[14].

Random Amplified Polymorphic DNA (RAPD): Random Amplified Polymorphic DNA (RAPD) is considered the most widely used molecular marker type in molecular studies.

RAPDs are based on the PCR amplification of random DNA segments with primers of random nucleotide sequencesthat were inexpensive and easy to use. In this technique short primers (10 bp) of random sequence are used. These oli gonucleotides are used both forward and reverse-primers and amplified fragments length are basically within the range of 0.5-5 kb in size are separated by agarose gel electrophoresis and polymorphisms can be estimated after stained with ethidium bromide dyes as the presence or absence of bands. The Primers bind to complementary DNA Sequences and where two primers bind to the DNA sample in close enough for successful PCR reaction. The amplified DNA products can then be visualized by gel electrophoresis[15]. The DNA fragments generated are then separated and detected by gel electrophoresis. RAPDs can be detected by running PCR products through electrophoresis on an agarose or acrylamide gel. In both cases, the gel is stained with ethidium bromide. The difference obtained by running RAPD products in acrylamide versus agarose lies only in the degree of resolution of bands. In most cases, agarose gel electrophoresis gives sufficient resolution.

Due to the commercial availability of random primers, no sequence data for primer construction is needed. Moreover, RAPDs have a very high genomic abundance and are randomly distributed throughout the genome. RAPD primers are readily available being universal.

Amplified Fragment Length Polymorphism (AFLP): AFLP is a technique used to detect polymorphisms in DNA when no information about the genome is known. The technique employs both RFLP and PCR by ligating primer recognition sequences to the DNA fragments produced through restriction digestion[16]. The technique involves three steps: Initially, oligonucleotide adapters are ligated to both ends of the resulting restriction fragments and genomic DNA is digested. Subsequently, the fragments are selectively amplified using the adapter and restriction site sequences as primer binding sites for following PCR reactions. As the 3’ ends of the primers extend into the restriction Fragments by 1-4 bp, only those fragments are amplified, whose ends are absolutely complementary to the 3’ends of the selective primers. Therefore, only a certain amount of the restriction fragments is amplified. Finally, the amplified fragments are resolved by gel electrophoresis and visualized by either silver staining, autoradiography or fluorescence, resulting in a unique reproducible fingerprint for Each individual[17].

The advantage of this method lies in the visualization of polymorphism by PCR without the knowledge of nucleotide sequence. Rapid generation, high reproducibility and high frequency polymorphism makes DNA analysis using AFL Pan attractive tool for identifying genetic polymorphism, constructing genetic linkage maps and detecting linkage by analyzing individuals in an appropriate segregating population.

Simple Sequence Repeats (SSRs): SSRs are randomly tandem repeats of short nucleotide motifs (2-6 bp)[2]. SSRs are frequently highly polymorphic sequences normally present in animal and plant species[18] and can be used to study the relationship between inherited traits within a species[19]. Microsatellite markers are often derived from noncoding/anonymous genomic regions, such as Bacterial Artificial Chromosomes (BACs) and Genomic Survey Sequences (GSSs). Therefore, development of SSR markers used to be expensive and laborious[20].

This assay is easily detectable by gel electrophoresis for few to hundreds of samples which could be inexpensive by researchers with limited resources. Polymorphism is based on the variation in the number of repeats in different genotypes[21]. Since, polymorphisms in longer penta-nucleotide and tetra repeats are easier to make a distinction. In a variety of detection systems and longer repeats may be more robust[22].

It requires much time and cost to isolate and characterize each SSR locus when the DNA sequence of a plant species is not available. Another drawback is the occurrence of null alleles. This may be due to the poor primer annealing because of nucleotide sequence divergence, inconsistent DNA quality or low DNA quantity[23] or it might be due to mutations in the primer binding site. This can cause difficulty in the determination of allelic and genotypic frequencies and an underestimation of heterozygosity[24].

Inter-Simple Sequence Repeats (ISSR): Inter Simple Sequence Repeats (ISSR) is one of dominant markers involved in amplification of DNA segment present at an amplifiable distance in between two identical microsatellite repeat regions oriented in opposite direction. This technique relies on the principle of using microsatellite as primers which target multiple genomic loci to amplify inter simple sequence repeats of the genomic DNA of different sizes.

Inter-Simple Sequence Repeat (ISSR) amplification is a novel technique which can rapidly differentiate between closely related individuals[25]. The ISSR technique involves anchoring of designed primers to a subset of SSRs and amplify the region between two closely spaced, oppositely oriented SSRs. The advantage

of this technique is its multiplexed banding profiles, high frequency of polymorphism, high throughput and relatively low cost.

Single Nucleotide Polymorphisms (SNPs): Single Nucleotide Polymorphisms (SNPs) involve the substitution of one nucleotide for another or the addition or deletion of one or a few nucleotides. There are four major reasons for the increasing interest in the use of SNPs as markers for genetic analysis. Firstly, they are prevalent and provide more potential markers near or in anylocus of interest than other types of polymorphism such as microsatellites[26]. Secondly, some SNPs are located in coding regions and directly affect protein function. These SNPs may be directly responsible for some of the variations among individuals in important traits[27]. Thirdly; SNPs are more stably inherited than microsatellites, making them more suited as long-term selection markers. Finally, SNPs are more suitable than microsatellites for high through putgenetic analysis using DNA microarray technology[28].

Expressed Sequence Tag Markers (EST): Each gene, transcribed into Messenger RNA (mRNA) serves as a template for protein synthesis. As mRNA is very unstable outside of cells; scientists use an enzyme called reverse transcriptase to convert mRNAs to complementary DNA (cDNA). cDNA production is the reverse of the usual process of transcription in cells because here mRNA acts as a template rather than DNA. cDNA is a stable compound only representing the expressed DNA sequence, generated from mRNA which represents exons by excising (splicing) in trons. Once cDNA representing an expressed gene is isolated, scientists can sequence nucleotides from either the 5' or 3' end to create 5' expressed sequence tags (5'ESTs) and 3' ESTs, respectively[29]. Often EST-based RFLP markers allow comparative mapping across different species because sequence conservation is high in the coding regions. A DNA library needs to be established for the production of specific STS primer sets[30].

SCAR (Sequence Characterized Amplified Region): A SCAR marker is a genomic DNA fragment that is identified by PCR amplification using a pair of specific oligo-nucleotide primers ,DNA markers like Simple Sequence Repeats (SSRs), Single-Nucleotide Polymorphisms (SNPs), Amplified Fragment Length Polymorphisms (AFLPs), Random Amplified Polymorphic DNAs (RAPDs), Cleaved Amplified Polymorphic Sequences (CAPS) and Restriction Fragment Length Polymorphisms (RFLPs) have been identified to be linked with blast resistance genes in rice[31].

CONCLUSION

To fulfill the need of food for the growing world population there is need to develop the high Yielding hybrids, varieties and superior population of food crops. This can be achieved through the use of molecular marker system. Molecular markers are independent on the environment, so it offers important tool for breeders to directly select for the genotype.

REFERENCES

01. Collard, B.C.Y., M.Z.Z. Jahufer, J.B. Brouwer and E.C.K. Pang, 2005. An introduction to markers, Quantitative Trait Loci (QTL) mapping and marker-assisted selection for crop improvement: The basic concepts. Euphytica, 142: 169-196.
02. Jiang, G.L., 2013. Molecular Markers and Marker-Assisted Breeding in Plants. In: Plant Breeding from Laboratories to Fields, Andersen, S.B. (Ed.). IntechOpen, Rijeka, Croatia, pp: 45-83.
03. Kumar, M., V. Chaudhary, U. Sirohi, M.K. Singh, S. Malik and R.K. Naresh, 2018. Biochemical and molecular markers for characterization of chrysanthemum germplasm: A review. J. Pharmacogn. Phytochem., 7: 2641-2652.
04. Bhatt, K.D., S.K. Girnari, V.B. Mandaliya, L.D. Chariya and V.S. Thaker, 2011. Use of RAPD marker to confirm mutation in morphological variants on Neem tree. Electron. J. Plant Breed., 2: 473-478.
05. Girnari, S.K., K.D. Bhatt, V.B. Mandaliya, L.D. Chariya and V.S. Thaker, 2011. RAPD markers studies in some selected medicinally important plants. Arhiv Poljoprivredne Nauke, 72: 5-14.
06. Ma, X., C. Xing, L. Guo, Y. Gong, H. Wang, Y. Zhao and J. Wu, 2007. Analysis of differentially expressed genes in genic male sterility cotton (*Gossypium hirsutum* L.) using cDNA-AFLP. J. Genet. Genomics, 34: 536-543.
07. Ramavat, J.M., R.K. Ramanuj, V.B. Mandaliya, L.D. Chariya, Y.M. Bapodariya and V.S. Thaker, 2010. A rapid and realible protocol for DNA extraction from *Catharanthus roseus*. J. Sci. Agric. Res., 71: 47-52.
08. Semagn, K., A. Bjornstad and M.N. Ndjiondjop, 2006. An overview of molecular marker methods for plants. Afr. J. Biotechnol., 5: 2540-2568.
09. Agarwal, M., N. Shrivastava and H. Padh, 2008. Advances in molecular marker techniques and their applications in plant sciences. Plant Cell Rep., 27: 617-631.
10. Southern, E.M., 1975. Detection of specific sequences among DNA fragments separated by gel electrophoresis. J. Mol. Biol., 98: 503-517.
11. Mishra, K.K., R.S. Fougat, A. Ballani, V. Thakur, Y. Jha and M. Bora, 2014. Potential and application of molecular markers techniques for plant genome analysis. Int. J. Pure Applied Biosci., 2: 169-189.

12. Paux, E., P. Sourdille, I. Mackay and C. Feuillet, 2012. Sequence-based marker development in wheat: Advances and applications to breeding. Biotechnol. Adv., 30: 1071-1088.
13. Reinisch, A.J., J.M. Dong, C.L. Brubaker, D.M. Stelly, J.F. Wendel and A.H. Paterson, 1994. A detailed RFLP map of cotton, *Gossypium hirsutum* × *Gossypium barbadense*: Chromosome organization and evolution in a disomic polyploid genome. Genetics, 138: 829-847.
14. Wright, R.J., P.M. Thaxton, K.M. El-Zik and A.H. Paterson, 1998. D-subgenome bias of Xcm resistance genes in tetraploid Gossypium (cotton) suggests that polyploid formation has created novel avenues for evolution. Genetics, 149: 1987-1996.
15. Williams, J.G.K., A.R. Kubelik, K.J. Livak, J.A. Rafalski and S.V. Tingey, 1990. DNA polymorphisms amplified by arbitrary primers are useful as genetic markers. Nucleic Acids Res., 18: 6531-6535.
16. Lynch, M. and B. Walsh, 1998. Genetics and Analysis of Quantitative Traits. Sinauer Associates Inc., Sunderland, USA., ISBN-13: 978-0878934812, Pages: 985.
17. Nicod, J.C. and C.R. Largiader, 2003. SNPs by AFLP (SBA): A rapid SNP isolation strategy for non-model organisms. Nucleic Acids Res., 31: 19e-19.
18. Kalia, R.K., M.K. Rai, S. Kalia, R. Singh and A.K. Dhawan, 2011. Microsatellite markers: An overview of the recent progress in plants. Euphytica, 177: 309-334.
19. Dunn, G., A.L. Hinrichs, S. Bertelsen, C.H. Jin, J.S. Kauwe, B.K. Suarez and L.J. Bierut, 2005. Microsatellites versus single-nucleotide polymorphisms in linkage analysis for quantitative and qualitative measures. BMC. Genetics Vol. 6, 10.1186/1471-2156-6-S1-S122.
20. Mir, R.R., P.J. Hiremath, O. Riera-Lizarazu and R.K. Varshney, 2013. Evolving Molecular Marker Technologies in Plants: From RFLPs to GBS. In: Diagnostics in Plant Breeding, Lubberstedt, T. and R.K. Varshney (Eds.). Springer Science-Business, New York, pp: 229-247.
21. Ellegren, H., 2000. Microsatellite mutations in the germline: Implications for evolutionary inference. Trends Genet., 16: 551-558.
22. Koelling, J., M.C. Coles, P.D. Matthews and A. Schwekendiek, 2012. Development of new microsatellite markers (SSRs) for *Humulus lupulus*. Mol. Breed., 30: 479-484.
23. Ellegren, H., 2004. Microsatellites: Simple sequences with complex evolution. Nat. Rev. Genet., 5: 435-445.
24. Kumar, P., V.K. Gupta, A.K. Misra, D.R. Modi and B.K. Pandey, 2009. Potential of molecular markers in plant biotechnology. Plant Omics J., 2: 141-162.
25. Zietkiewicz, E., A. Rafalski and D. Labuda, 1994. Genome fingerprinting by Simple Sequence Repeat (SSR)-anchored polymerase chain reaction amplification. Genomics, 20: 176-183.
26. Primmer, C.R., T. Borge, J. Lindell and G.P. Saetre, 2002. Single-nucleotide polymorphism characterization in species with limited available sequence information: high nucleotide diversity revealed in the avian genome. Mol. Ecol., 11: 603-612.
27. Vignal, A., D. Milan, M. SanCristobal and A. Eggen, 2002. A review on SNP and other types of molecular markers and their use in animal genetics. Genet. Select. Evol., 34: 275-305.
28. Werner, M., M. Sych, N. Herbon, T. Illig, I.R. Konig and M. Wjst, 2002. Large-scale determination of SNP allele frequencies in DNA pools using MALDI-TOF mass spectrometry. Hum. Mutat., 20: 57-64.
29. Jongeneel, C.V., 2000. Searching the Expressed Sequence Tag (EST) databases: Panning for genes. Briefings Bioinformatics, 1: 76-92.
30. Bang, K.H., J.W. Lee, Y.C. Kim, D.H. Kim, E.H. Lee and J.U. Jeung, 2010. Construction of genomic DNA library of Korean ginseng (Panax ginseng CA M EYER) and development of sequence-tagged sites. Biol. Pharm. Bull., 33: 1579-1588.
31. Ashkani, S., M.Y. Rafii, M. Shabanimofrad, A. Ghasemzadeh, S. Ravanfar and M.A. Latif, 2014. Molecular progress on the mapping and cloning of functional genes for blast disease in rice (*Oryza sativa* L.): Current status and future considerations. Crit. Rev. Biotechnol., 14: 1-15.