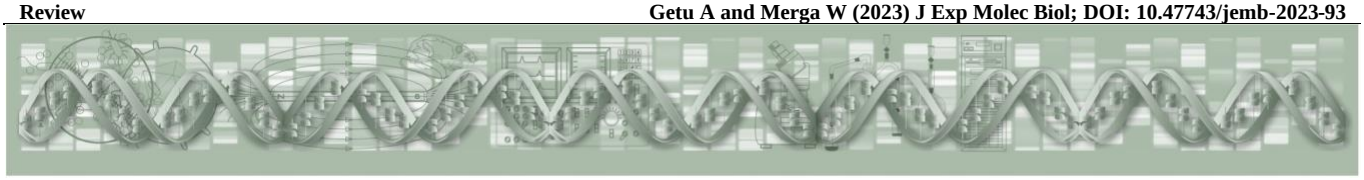

Application of Genetic Marker for Diversity Assessment and Conserving of Plant Genetic Resource

Wakuma Merga and Abukiya Getu

Ethiopian Institute of Agricultural Research Teppi Agricultural Research Centre
Corresponding Author: Wakuma Merga, Ethiopian Institute of Agricultural Research Teppi Agricultural Research Centre, P.O.Box 34, Teppi, Ethiopia,
Email: wakumerga@gmail.com

Abstract

Several molecular markers have been used for various purposes since the beginning of contemporary molecular technology: genetic resource characterization, core genetic resource collection, mapping, marker-assisted selection and marker-assisted backcrossing etc. The variety of all plant species, their genetic make-up and the environments in which they dwell are considered to be components of biological diversity. Plant classification and identification can be done most quickly and simply using morphological examination. It has been established that molecular markers are effective instruments for evaluating genetic diversity in groups and individuals as well as germplasm resources. To preserve biodiversity more than just genetic testing and DNA polymorphism detection are needed. It examines significant issues regarding managing plant germplasm both ex-situ and in-situ in order to support decision-making. Significant progress has been made in recent years in mapping, tagging and isolating many important genes for agriculture using molecular markers like (Restriction Fragment Length Polymorphisms, Random Amplified Polymorphic DNAs, Amplified Fragment Length Polymorphisms, Simple Sequence Repeats and Single Nucleotide Polymorphism). Each marking technique has unique benefits, drawbacks and applications. If one was aware of the presence of relevant traits, genes and alleles one might make judgments on the extension of accessions and the preservation of seed stocks to meet an anticipated rise in demand for materials. By including genotypes with well-known and useful genes and alleles in the core collections and then breeders can use them. Utilizing crop plant genetic resources will be simpler as a result increasing their potential.

Keywords: DNA, Germplasm Conservation, Genome editing, Marker assisted selection, Quantitative Trait Loci, Polymerase Chain Reaction,

Introduction

The diversity of all plant and animal species, their genetic makeup and the ecosystems in which they cohabit are all seen to be instances of biological diversity. The three basic types of diversity are genetic, theoretical and ecological. The assortment of genotype and gene variations is referred to as genetic variety (Species communities and the ecosystems they inhabit). Many would argue that diversity is vital for the feasibility and sustainability of many human efforts,and in recent years the worth of biodiversity to humans has received significant acknowledgement. (Shiva, 1994). Morphological analysis is the quickest and least technical way for classifying and identifying plants. The strategy involves writing down and keeping track of features that are simple to see such form and structure. Prior to the development of biotechnology, distinct cultivars were characterized by their morphological and physical characteristics. Numerous studies have evaluated the genetic diversity of various crops based on differences in morphological and agronomic features or ancestry data. (Sneller et al. 1997).

It has been demonstrated that molecular markers are useful tools for assessing germplasm resources and assessing genetic variation both among and within populations. The area of plant genomics that has really made the most strides in the use of DNA marker technology is population genetics. Nevertheless, markers produced from the polymerase chain reaction (PCR) and restriction length polymorphism (RFLP) have also been extensively used in the mapping of Mendelian genes and QTLs in plants. It is crucial to use molecular markers for genetic resource management and research so that breeders can add critical genetic and behavioral data to core collections. Recent advances in the techniques have led to the mapping, tagging and isolation of a number of agriculturally significant genes. Microsatellites, restriction fragment length polymorphism (RFLP), amplified fragment length polymorphism (AFLP), random amplified polymorphic DNA (RAPD) and PCR-based DNA markers like sequence-characterized amplifiable regions (SCARs) or sequence-tagged sites are among the methods that are particularly promising (STS). These methods make use of F2 and backcross populations near isogenic lines, doubling haploids and recombinant inbred lines to aid in the direct selection of numerous desired traits simultaneously (FAO, 2002).

Managing biodiversity entails more than just genetic characterization via DNA polymorphism detection; it also calls for data that may be used to address important challenges in the management of both ex situ and in situ plant germplasm and to aid in decision-making. Molecular technologies may help with sampling, management, the creation of "core" collections and the use of genetic diversity for in situ agricultural germplasm maintenance. The best management and utilization strategies for the in situ and "on farm" preservation strategies of genetic resources may be determined using molecular markers. These markers can

also help identify the most representative populations within a landrace's "gene pool" as well as the best management and usage tactics (Lanteri and Barcaccia, 2005).Therefore,, the objective of these review article is to describe molecular marker approaches and discuss how to use molecular markers to assess and protect plant genetic resources.

Methods of Molecular Markers and Their Application

A DNA sequence on an area in an organism's genome where the DNA genomic sequence changes between members of a population are referred to as a molecular marker. The way that molecular markers function is by exposing variations in DNA sequences (polymorphism) among various members of the population. Insertions, deletions, point mutations and translocations are a few of these alterations. A desirable biomarker has the following qualities: high levels of polymorphism that are evenly distributed throughout the genome, the absence of a requirement for prior knowledge of an organism's genome, the ability to generate multiple, distinct and reliable markers that provide an adequate level of resolution, the need for little startup material, simplicity, speed and low cost and the lack of pleiotropic or epistatic links to distant phenotypes.

Restriction Fragment Length Polymorphism (RFLP)

Refers to Variations in the size of DNA fragments generated from the genomic DNA of two or more individuals of a species by a particular restriction endonuclease (Kahl, 2001). An infinite number of RFLPs can be produced by digesting entire DNA with particular restriction enzymes. RFLPs are co-dominant in nature and are comparatively tiny in size. The restriction enzyme will only cut the DNA of one person if there is even a little nucleotide difference between the two in the restriction site. Thus, restriction fragments of various lengths are produced. It is bands that identify DNA fragments that have been produced as a result of the digestion of genomic DNA by restriction enzymes. These pieces of DNA are typically between 2 and 10 kb in size.

Southern blot hybridization to a tagged DNA probe is used to detect DNA fragments after they have been separated using agarose gel electrophoresis. Digoxigenin or fluorescein are examples of non-radioactive stains that can be used to mark the probe instead of a radioactive isotope. The homologous sequence of a certain chromosomal region makes up the locus-specific RFLP probes. The creation of genomic or complementary DNA (cDNA) libraries yields the creation of probes, which can be made of either a specific sequence of unknown identity (genomic DNA) or a segment of the sequence of a functional gene. (Exons only, cDNA). In suitable bacterial vectors, RFLP probes are kept as clones, making it simple to isolate the DNA fragments they contain.

Techniques applied: - To separate the fragments of genomic DNA based on size, electrophoresis is performed through an agarose gel after one or more restriction enzymes have broken down the DNA. The DNA is subsequently moved from the gel to a nylon membrane following in situ denaturation. The positioning of the DNA fragments relative to one another is not altered during passage to the filter. Following the hybridization of the DNA with radioactive or (in this case) non-radioactively labelled DNA probes, autoradiography or alternative enzyme-linked detection methods can be used to determine the locations of the probe's complementary bands.

Random Amplified Polymorphic DNA (RAPD)

Any DNA fragment amplified using the polymerase chain reaction with short oligodeoxy nucleotide primers of any nucleotide sequence (amplifiers) (Kahl, 2001). RAPDs are DNA fragments generated by PCR using short, synthetic primers with random sequences (often 10 bp). These oligonucleotides often have the capacity to amplify fragments from 1–10 genomic locations simultaneously and function as forward and reverse primers. Agarose gel electrophoresis is used to separate amplified fragments, which are typically in the 0.5–5 kb size range. Ethidium bromide staining and the presence or absence of bands of particular diameters are used to identify polymorphisms. These polymorphisms can occur for a variety of reasons, including variations in the primer annealing sites and length differences in the amplified sequence between the primer annealing sites.

Amplified Fragment Length Polymorphism (AFLP)

AFLPs are DNA fragments (80–500 bp) produced by restriction enzyme digestion, oligonucleotide adapter ligation to the digestion products, and targeted PCR amplification. Therefore, RFLP and PCR are both used in AFLPs. The restriction enzyme-specific sequence, 1–5 selected nucleotides and a core sequence (a component of the adaptor) make up the PCR primers. Variations in the restriction sites or the surrounding area lead to changes in the AFLP banding profiles. The AFLP method produces fragments from numerous genomic locations concurrently (about 50–100 fragments per reaction), which are then sorted by polyacrylamide gel electrophoresis.. Selective Fragment Length Amplification (SFLA) and Selective Restriction Fragment Amplification (SRFA) are synonyms sometimes used to refer to AFLPs. A variation of the AFLP technique is known as Selectively Amplified Microsatellite Polymorphic Locus (SAMPL). This technology amplifies microsatellite loci by using a single AFLP primer in combination with a primer complementary to compound microsatellite sequences, which do not require prior cloning and characterization.

Simple Sequence Repeats SSR (Microsatellite)

Any of a number of DNA sequences that are scattered throughout the genomes of fungi, plants, animals and humans that are very short (2-10 bp), largely repetitive, tandemly organized and highly variable (hyper variable) (Kahl, 2001). A kind of repeated DNA components are called microsatellites or simple sequence repeats (SSR). (Tautz and Rentz, 1984; Tautz, 1989). The 5- to 50-copy tandem arrays of di-, tri-, or tetra-nucleotide repeats such as (AT)29, (CAC)16, or (GACA)32, are used to organize the repeats. Plants have many SSRs, with one on average per 6-7 kb. In order to PCR-amplify the DNA portion

containing the SSR, forward and reverse primers can be created using the conserved nucleotide sequences that flank these repeat motifs. By using gel electrophoresis, SSR alleles, which are amplified products with varied lengths may be distinguished and seen by silver-staining, autoradiography (if primers are radioactively labelled) orvia automation (if primers are fluorescently labelled). SSR analysis is amenable to automation and multiplexing), and allows genotyping to be performed on large numbers of lines and multiple loci to be analyzed simultaneously. SSRs can be identified by searching among DNA databases (e.g. EMBL and Gene bank) or alternatively small insert (200-600bp) genomic DNA libraries can be produced and enriched for particular repeats (Powell *et al*., 1996). From the sequence data primer pairs.

Single nucleotide polymorphism (SNP)

Any variation between two genomes that is caused by a minor deletion, insertion or exchange of a single nucleotide (Kahl, 2001). The novel marker technology known as small nucleotide polymorphism (SNP) was first created in nonhuman primates. SNPs, which have two to three severe polymorphic sites per site, are the most prevalent polymorphism markers (Cooper *et al*., 1985). SNPs were initially found in humans but they are currently used to genotype plants. Sequence information is a key component of SNP technology.

Table I Widely used molecular markers in conservation of plant genetic resource

Feature and description	RFLP	RAPD	AFLP	SSR	SNP
Genomic abundance	High	High	High	Moderate to high	Very high
Expression/inheritance Co-dominant	Co-dominant	Dominant	Dominant/co-dominant	Co-dominant	Co-dominant
Number of loci	Small (< 1000)	Small (< 1000)	Moderate (1000s)	High (1000s–10,000s)	Very high (> 100,000)
Level of polymorphism	Moderate	High	High	High	High
Type of polymorphism	Single base change, indel	Single base change, Indel	Single base change, indel	Changes in length repeat	Single base change, indel
Cloning and/or sequencing	Yes	No	No	Yes	Yes
Type of probes/primer	Low-copy DNA or cDNA clones	10 bs random	Nucleotides Specific	sequence Specific	sequence Allele-specific PCR primer
PCR-based	Usually no	Yes	Yes	Yes	Yes
Radioactive detection	Usually yes	No	Yes or NO	Usually no	No
Reproducibility/reliability	High	Low	High	High	High
Amount of DNA required	Large (5–50 µg)	Small (0.01–0.1 µg)	Moderate (0.5–1.0 µg)	Small (0.05–0.12 µg)	Small (> 0.05 µg)
Genotyping throughput	Low	Low	High	High	High
Cost	Moderate to high	Low	Moderate	Moderate to high	High
Marker index Moderate	Low	Moderate	Moderate	Moderate to high	Moderate
Time demanding	High	Low	Moderate	Low	Low
Number of polymorphic per loci	1-3	1.5-5	20-100	1-3	1
Primary application	Genetic	Diversity	Diversity and genetic	All purpose	All purpose

Due to the timeline of technological advancement, molecular markers are divided into three main categories or generations: (i) hybridization-based markers, such as restriction fragment length polymorphism (RFLP), (ii) PCR-based markers, such as random amplification of polymorphic DNA (RAPD), amplified fragment length polymorphism (AFLP) and microsatellites or simple sequence repeat (SSR) and (iii) sequence-based markers, such as single nucleotide (SNP). In various areas of plant science study, including genetic and phylogenetic investigations and ecological, evolutionary and taxonomic studies, DNA-based molecular markers are frequently used.

Table II Strength, weakness and application of molecular markers

Markers type	**Strength**	**Weakness**	**Application**
Restriction Fragment Length Polymorphism (RFLP)	Moderately polymorphic. Show co-dominant alleles and having high reproducibility	Requirement of laborious and technically demanding methodological procedures Not amenable to automation and collaboration among research teams requires distribution of probes.	Applied in diversity and phylogenetic studies Widely used in gene mapping studies They also have been used to investigate relationships of closely related taxa
Random Amplified Polymorphic DNA (RAPD)	They are quick and easy to assay. No sequence data for primer construction are needed. Have a very high genom ic abundance and are randomly distributed throughout the genome.	Low reproducibility highly standardized experimental procedures are needed Markers are not locus-specific, band profiles cannot be interpreted in terms of loci and alleles (dominance of markers), and similar sized fragments may not be homologous.	Studies at the individual level. Applied in gene mapping studies. Include Arbitrarily Primed Polymerase Chain Reaction (AP-PCR)
Amplified Fragment Length Polymorphism (AFLP)	High genomic abundance Considerable reproducibility No sequence data for primer construction are required.	The need for purified, high molecular weight DNA The dominance of alleles, and the possible non-homology of co-migrating fragments belonging to different loci.	Applied in studies involving genetic identity, parentage and identification of clones and cultivars, and phylogenetic studies of closely related species. Widely valued technology for gene mapping studies Considered more applicable to intraspecific than to interspecific studies
Simple Sequence Repeats SSR (Microsatellite)	Fast, highly polymorphic Required only very small DNA Does not require radioactivity.	High developmental and startup cost Species-specific Sometimes difficult interpretation	Fingerprinting, Varietal/line identification Genetic maps, Diversity studies, Marker-assisted selection
Single nucleotide polymorphism (SNP)	Robust in usage Polymorphism are identifiable Different detection methods available Suitable for high throughput Can be automated	Very high development costs Requires sequence information Can be technically challenging	Genetic maps, Diversity studies, Marker-assisted selection

Role of Molecular Marker for Plant Genetic Resource Conservation

Characterization of plant genetic resources refers to the procedure and equipment used to assess, classify and identify accessions. This identification, in broad terms, can apply to any morphological descriptor or molecular polymorphism of an accession. Characterization refers to the description of qualitative or quantitative features that are highly heritable, detectable by the eye, and equally expressed in all situations. It is a word commonly used in the management of germplasm collections and gene banks (Van and M.C, 2005). Strong genetic characterization makes it possible for decisions to be made regarding conservation measures to be backed by this data which leads to better management of the germplasm. Plant genetic resources are enhanced by experimental work related to them, as well as by morpho-phenological and molecular analyses of germplasm.

Customary efforts to directly employ plant accessions kept in germplasm banks in breeding programs have generally been focused on locating the origins of interesting genes, including resistance to plant diseases or pests, and their transfer to cultivated materials. Because better material from sophisticated breeding programs is significantly more alluring than any germplasm resource with unknown genetic origin and phenotypic adaptability or performance, linkage drag has frequently prevented breeders from taking the initiative to employ accessions from germplasm banks (Ferreira, 2006).

Molecular Markers' Function in Characterization

The use of genomic DNA-based marker assays has modernized and changed the way we can characterize genetic diversity and explain genetic selection (Lanteri and Barcaccia, 2006). For the characterization of genome architectures and the analysis of gene polymorphisms in agricultural plants, molecular markers are recognized as being particularly efficient and trustworthy methods. The plant DNA polymorphism assays are effective instruments for assessing and examining germplasm resources and genetic relatedness, in addition to linkage mapping, gene targeting and aided breeding. The most popular methods for characterizing crop plant genetic resources are SSR and AFLP markers, which both provide an almost infinite variety of molecular traits for distinct fingerprinting and genotyping of plant materials, both with and without prior knowledge of the target DNA sequences.

In order to detect polymorphism, molecular markers typically assess a portion of the total amount of DNA sequence variation in a genome. The polymorphisms that the RFLP assay finds are a reflection of the range of restriction fragment sites. The causes of PCR-based polymorphisms are changes in DNA sequence at primer binding sites as well as variations in DNA length between primer binding sites. In the SSR experiment, which employs pairs of primers flanking each simple sequence repeat, polymorphisms depend on the number of repeating di-, tri-, or tetra nucleotide units that are present at one locus. To find polymorphisms at various loci, the AFLP assay combines primers specific for two separate four-base and six-base long restriction sites that flank the target sequence unit.

The selective potency of RFLP markers has generated significant debate, despite the fact that they have assisted in quantifying the genetic diversity and relatedness of crop plants. More probe-enzyme combinations may increase the number of loci that RFLP markers can identify, but PCR-based markers can still detect a higher level of polymorphism. SSR markers often detect multiple alleles at a specific locus because of their own genetic make-up, in contrast to AFLP tests, which typically detect single alleles at various randomly distributed loci in the genome. SNP markers, more contemporary methods, directly sequence the DNA of target gene regions to detect single-nucleotide polymorphisms. In actuality, AFLP markers have so far provided the widest the largest.

Population-Level Characterization of Germplasm

Determining the genetic make-up of societies is greatly aided by measures of genetic diversity and similarity. The genetic composition of a crop plant species natural populations is significantly influenced by both an individual's reproductive system and the kinds of unions that occur within populations. Plant reproductive barriers and mating systems determine the breeding methods that can be used and the types that can be produced. Normally dominated by a small number of genotypes with advantageous adaptations for genotype separation, natural populations of animals that reproduce vegetatively or via apomixes are polyclonal, consisting of multiple genetically distinct clones (Spooner, 2005).

Land races of self-pollinated plants, such beans, lentils, wheat, and barley, are made up of pure lines that are reproductively separate from one another genetically. Genetic and phenotypic variation is most obvious within lines because natural populations frequently have fixed genotypes, primarily homozygous for different alleles. Yet, depending on the species, environmental factors and germplasm availability, some spontaneous hybridization is conceivable. The cultivated forms of selfing species are often pure lines made through the repeated self-pollination of a large number of hybrid individuals descended from two parental lines chosen for complementary morphological and commercial features. One of the most significant cross-pollinated plants for trade is maize (Spooner, 2005).

Molecular Markers for Plant Genetic Resource Conservation

Knowing the genetic make-up of gene bank accessions facilitates decision-making for conservation operations, which range from acquiring and managing genetic resources to choosing genes that provide value for breeding purposes. The adoption of educated sampling methods for germplasm material intended for ex-situ conservation and the choice of prioritized in-situ conservation sites are essential for the success of conservation initiatives. Hence, developing solutions requires knowledge about the location, distribution and degree of genetic diversity. Molecular marker-based characterization provides reliable information on a variety of factors, including the degree of genetic diversity/similarity, the structure of genetic variation in samples and populations rates of genetic differentiation among populations and the distribution of biodiversity in populations from different geographical locations.

Molecular characterization is useful for understanding the reproduction of species, the reproduction and adaptation of individuals and the existence of gene flow across individuals, i.e., the movement of alleles within and among populations of the same or related species (Papa and Gepts, 2003). Molecular data, which also provide the essential knowledge required to understand taxonomy, domestication and evolution increase or even enable the understanding of phylogeny. Moreover, molecular marker data sets provide a baseline for monitoring changes in the genetic makeup of accessions that occur naturally as well as those brought on by human intervention help prevent issues later (De Vicente *et al.*, 2006). Many molecular genetics-related methods, such as Mendelian gene tagging and QTL mapping, have been helpful in characterizing. This research, where cutting-edge techniques have found beneficial variation that may aid in varietal development, has highlighted the significance of wild relatives.

Current Advancement of Genetic Marker

The inter simple sequence repeat (ISSR)-PCR technique uses microsatellite sequences as primers in a polymerase chain reaction to produce multi-locus markers. The bulk of the advantages of microsatellites (SSRs) and amplified fragment length polymorphism (AFLP) are combined in a quick and simple technology. This technology combines the universality of random amplified polymorphic DNA with these advantages (RAPD). Given that they have a high degree of polymorphism, ISSR markers are useful in studies of genetic diversity, phylogeny, gene tagging, genome mapping and evolutionary biology (Reddy 2002). Sequences amplified by ISSR-PCR can be used to identify DNA. Sequence diversity is larger than in actual gene sequences but lower than in SSR-PCR. This method is better suited for phylogeographical research or perhaps taxonomic delimitation rather than human identification because an ISSR may be a conserved or non-conserved area. In addition, microsatellite sequencing and ISSR sequencing are advantageous to one another since each produces primers for the other (Pradeep *et al.*, 2002). (Uddin, M.S. and Cheng, Q., 2015)

Applications of ISSR

For the first time in spiders, the Inter Simple Sequence Repeat (ISSR) method was used to assess the genetic diversity of Brachypelma vagans populations in Mexico. Seven ISSR primers were tested and six populations on the Yucatan peninsula were sampled non-lethally. Four of these primers generated fragments (bands) that were sufficiently distinct and reproducible to generate a binary matrix and determine the parameters of genetic variability. There was the highest level of polymorphism (P 5 98.7%) ever seen in tarantula spiders. The results show that the ISSR-PCR strategy is promising for tarantula spider intraspecific variation (Uddin, M.S. and Cheng, Q., 2015)

Application of Marker Assisted Selection

Local damages in the genome were caused by targeting the majority of crop plants can benefit from the non-transgenic reverse genetics technique known as targeted induced local lesions in the genome (TILLING). In an effort to comprehend how two genes in Arabidopsis plants function, McCallum created TILLING in 1990. (McCallum CM *et al.*, 2000). When using TILLING procedures, the mutagenic population is initially established by treating seeds with a common chemical mutagen like methyl methanesulfonate (MMS) or ethyl methanesulfonate (EMS). Using the most crucial methods; including mass spectroscopy, liquid chromatography, array-based technologies and enzymatic mismatch cleavage. It is possible to determine the differences in target nucleotide sequences of mutant people in the population (Kurowska M.*et al*., 2011). Eventually more significant bioinformatics methods are used to analyze mutations brought on by certain mutagens, such as project aligned relevant sequences and assessed SNPs (PARSESNP). While TILLING is applicable to all plant species, ploidy level and genome size shouldn't have an impact on it. The discovery of a higher rate of gene mutations is the technique's most significant benefit. This method can be utilized in molecular genetics during plant breeding projects since it allows for the rapid and accurate identification of novel alleles at a cheaper cost and in less time (Hasan *et al.*, 2021)

Genome editing (CRISPR)

The use of the CRISPR genome editing method has enhanced many crop plants (Feng Z. *et al*., 2013). Because to its many benefits, including ease of use, the capacity to cleave the methylation loci and the versatility of genome editing, the new emerging technique of Cas9 technology is quickly replacing other methods (Hsu PD *et al.*, 2013, Lozano-Juste J, Cutler SR ., 2014). The two most crucial components of the CRISPR method are CRISPR RNAs and the Cas protein. Two short-length RNA molecules, trans-encoded CRISPR RNA (tracrRNA) and CRISPR RNA (crRNA), can cleave a specific target site with the aid of Cas9 endonuclease (the most explored Cas protein). Single guide RNA, or sgRNA, is the hybrid formed when tracrRNA and crRNA are artificially combined (Qi LS, Larson MH. *et al*., 2013). The sgRNA and Cas proteins come together to produce the RNA-guided endonuclease, which facilitates the cleave of a specific sequence in the genome (Niewoehner O., *et al* 2013). The CRISPR-Cas system is divided into three types—I, II, and III—based on this Cas protein. Two distinct proteins called Cas1 and Cas2 are frequently found in all three types. Type III is most frequently found in archaea but can also be found in some bacteria, whereas type I is present in both archaea and bacteria and type II is only present in bacteria.(Makarova KS. *et al*., 2011). Genome editing have been performed fruitfully in model plants like Nicotiana tabacum (Shan Q.*et al*., 2013), Arabidopsis (Ali Z. *et al*. 2018) and some economically important crops like maize (Svitashev S.*et al*., 2018) and wheat (Hasan et al., 2021)

Investigations of genome-wide associations in plants in order to fairly precisely identify the genomic areas causing symptoms genome-wide association studies (GWAS) take advantage of an ancient recombination event. After a decade of intense research, the GWAS approach was well established in the field of human genetics. GWAS are already developing as a potent tool for discovering natural variation underlying complex traits in cops due to the rapid advancement of sequencing technologies and computational methodologies (Zargar SM. *et al*., 2015). GWAS in crops typically utilizes a permanent resource—a population of various (and preferably homozygous) varieties that can be rephenotyped for multiple traits and only need to be genotyped once—and one can then produce specific mapping populations for a particular trait or QTLs in crops(Atwell S.*et al* 2010). GWAS have now been successfully conducted in a variety of crops, including sorghum, foxtail millet, rice and maize (Zhao J. *et al*., 2016) . Rice and maize are the two main crop GWAS models based on the volume of resources already created and published, and both include panels of tens of thousands of genotyped inbred and multiple environment experiments undertaken for numerous characteristics. Low genome coverage sequencing was done on 446 wild rice accessions (Oryza rufipogon) and 1083 cultivated O. sativa spp. indica and O. sativa spp. japonica varieties of rice(Huang X. *et al*., 2012). Data imputation was used to create a high-density haplotype map of the rice genome, and a GWAS was then carried out utilizing the extensive data set of 1.3 million SNPs to describe the allele linked to 10 grain-related variables and blooming time. A GWAS was also carried

out in 446 O.sativa for the traits leaf sheath color and tiller angle since these traits would have better mapping power because the wild species exhibit a greater level of genetic diversity Rufipogon accessions. In addition, the GWAS was carried out using a genotyping method based on microarrays. The genotyping of 413 different O. sativa accessions at 44,100 SNP variations and the phenotyping of 34 phenotypes revealed the intricate genetic architecture of rice attributes. By combining linkage mapping and GWAS in the NAM panel, the genetic architecture of maize's flowering time, leaf angle, leaf size and disease resistance features was analyzed, and numerous associated candidate genes were found (Juliana P. *et al*., 2018). The GWAS results showed that several QTLs with minor effects dominate the genetic architecture of these traits. A GWAS in maize was recently conducted to determine the composition of maize kernel oil, which is a significant food and energy source (Li H. 2013). A total of 368 maize lines were examined using a genome-wide SNP analysis and 74 loci were discovered to be connected to the quantity and makeup of maize kernel oil. These results demonstrate that the GWAS method in crops is a reliable and practical approach that complements traditional biparental cross mapping and has the capacity to simultaneously genetically map many variables. In order to better research the genetic basis of plant shape, yield, and physiology in more grasses as well as close wild family of cultivated crops, GWAS results are anticipated to be used. It is significant to emphasize that uncommon alleles, which account for a considerable fraction of natural variation, have a low power in GWAS 44% of the SNPs in rice have low frequency. The use of a large sample size or method for high-throughput virus-induced gene silencing is recommended when studying uncommon alleles (VIGS). VIGS is performed by cloning a short stretch of sequence from a candidate gene or random cDNAs into a virus genome under the control of promoter within a binary vector (Hasan *et al*., 2021).

Conclusion

Molecular markers give better conservation tactics a strong base. Some molecular methods can be used to accurately and cheaply identify the genotype profiles and gene haplotypes of accessions generated by DNA sequencing and fingerprinting. These methods can also detect contaminants, especially in the case of mixtures, contaminating genes from commercial or other accessions, as well as the existence of redundant materials or duplicated accessions. In terms of the use of molecular marker technologies, applied breeding programs are probably the area of agricultural plant genomics that will experience the most expansion. Genetic characterization serves an increasingly important function of identification. Using cutting-edge molecular markers, it is now possible to conserve desirable treat at the DNA level. The most recent developments in molecular plant breeding technology include marker aided selection, marker assisted back crossing and marker assisted plant breeding. On the other hand, genome editing or CRISPR technology is currently being used to improve several crop plants. Although each of the aforementioned markers has its own benefit and negative, they are all utilized to conserve and make use of plant genetic resources.

References

Agarwal M, Shrivastava N, Padh H., 2008. Advances in molecular marker techniques and their applications in plant sciences. Plant Cell Rep 27:617–31.

Ali Z, Eid A, Ali S, Mahfouz MM (2018) Pea early-browning virusmediated genome editing via the CRISPR/Cas9 system in Nicotiana benthamiana and Arabidopsis. Virus Res 244:333–337. https://doi.org/1 0.1016/j.virusres.2017.10.009

Atwell S, Huang YS, Vilhjálmsson BJ, Willems G, Horton M, Li Y, Nordborg M (2010) Genome-wide association study of 107 phenotypes in Arabidopsis thaliana inbred lines. Nature 465(7298):627–631. https://doi.org/10.1038/na ture08800

Cooper, D. N., B. A. Smith, H. J. Cooke, S. Niemann, and J. Schmidtke, 1985. An estimate of unique DNA sequence heterozygosity in the human genome. Hum.Genet. **69**(3): 201-205.

FAO/IAEA 2002. Division of Nuclear Techniques in Food and Agriculture. *international atomic energy agency*, Vienna.

Feng Z, Zhang B, Ding W, Liu X, Yang DL, Wei P, Cao F, Zhu S, Zhang F, Mao Y, Zhu JK (2013) Efficient genome editing in plants using a CRISPR/Cas system. Cell Res 23(10):1229–1232. https://doi.org/10.1038/cr.2013.114

Ferreira ME (2006) Molecular analysis of gene banks for sustainable conservation and increased use of crop genetic resources. In: Ruane J, Sonnino A (eds) The role of biotechnology in exploring and protecting agricultural genetic resources, FAO, Rome. pp 121–127

Hasan et al. Journal of Genetic Engineering and Biotechnology (2021) 19:128 Pp 15-26 the construction of multiple biparental cross population (e.g. NAM or MAGIC) may be helpful

Hsu PD, Scott DA, Weinstein JA, Ran FA, Konermann S, Agarwala V, Li Y, Fine EJ, Wu X, Shalem O, Cradick TJ, Marraffini LA, Bao G, Zhang F (2013) DNA targeting specificity of RNA-guided Cas9 nucleases. Nat Biotechnol 31(9):827–832. https://doi.org/10.1038/nbt.2647 .

Huang X, Zhao Y, Li C, Wang A, Zhao Q, Li W, Han B (2012) Genome-wide association study of flowering time and grain yield traits in a worldwide collection of rice germplasm. Nat Genet 44(1):32–39. https://doi.org/10.103 8/ng.1018

Joint, FAO. (2002). Mutant germplasm characterization using molecular markers. A manual (No. IAEA-TCS--19). Joint FAO/IAEA Division of Nuclear Techniques in Food and Agriculture.

Juliana P, Singh RP, Singh PK, Poland JA, Bergstrom GC, Huerta-Espino J, Bhavani S, Crossa J, Sorrells ME (2018) Genome-wide association mapping for resistance to leaf rust, stripe rust and tan spot in wheat reveals potential candidate genes. Theor Appl Genet 131(7):1405–1422. https://doi.org/10.1 007/s00122-018-3086-6

Kahl, G. (2001). *Dictionary of gene technology* (No. Ed. 2). Wiley-VCH Verlag GmbH.

Kisha, T. J., Sneller, C. H., & Diers, B. W. (1997). Relationship between genetic distance among parents and genetic variance in populations of soybean. *Crop Science*, *37*(4), 1317-1325.

Kurowska M, Daszkowska-Golec A, Gruszka D, Marzec M, Szurman M, Szarejko I, Maluszynski M (2011) TILLING – a shortcut in functional genomics. J Appl Genet 52(4):371–390. https://doi.org/10.1007/s13353-011- 0061-1
Lanteri S, Barcaccia G (2006) Molecular markers based analysis for crop germplasm preservation. In: Ruane J, Sonnino A (eds) The role of biotechnology in exploring and protecting agricultural genetic resources, FAO, Rome. pp 55–66
Lanteri Sergio and Barcaccia Gianni, 2005. Molecular markers based analysis for crop
Li H, Peng Z, Yang X, Wang W, Fu J, Wang J, Yan J (2013) Genome-wide association study dissects the genetic architecture of oil biosynthesis in maize kernels. Nat Genet 45(1):43–50. https://doi.org/10.1038/ng.2484
Lozano-Juste J, Cutler SR (2014) Plant genome engineering in full bloom. Trends Plant Sci 19(5):284–287. https://doi.org/10.1016/j.tplants.2014.02.014
Makarova KS, Haft DH, Barrangou R, Brouns SJ, Charpentier E, Horvath P, Moineau S, Mojica FJ, Wolf YI, Yakunin AF, van der Oost J, Koonin EV (2011) Evolution and classification of the CRISPR– Cas systems. Nat Rev Microbiol 9(6):467–477. https://doi.org/10.1038/nrmicro2577
McCallum CM, Comai L, Greene EA, Henikoff S (2000) Targeting induced local lesions in genomes (TILLING) for plant functional genomics. Plant Physiol 123(2):439–442. https://doi.org/10.1104/pp.123.2.439
Morgante M, OlivieriAM (1993) PCR-amplified microsatellites as markers in plant genetics. Plant J 3:175–182
Niewoehner O, Jinek M, Doudna JA (2013) Evolution of CRISPR RNA recognition and processing by Cas6 endonucleases. Nucleic Acids Res 42(2): 1341–1353. https://doi.org/10.1093/nar/gkt922
Papa R, Gepts P (2003) Asymmetry of gene flow and differential geographical structure of molecular diversity in wild and domesticated common bean (*Phaseolus vulgaris* L.) from Mesoamerica. TheorAppl Genet 106:239–250
Powell, W., G. C. Machray, and J. Provan**,** 1996. Polymorphism revealed by simple sequence repeats. Trends in Plant Sci. **1**(7): 215-222
Pradeep Reddy, M., Sarla, N. and Siddiq, E.A., 2002. Inter simple sequence repeat (ISSR) polymorphism and its application in plant breeding. *euphytica*, *128*, pp.9-17.
Primmer CR, 2009. From conservation genetics to conservation genomics. Ann N Y Acad Sci 1162:357–68.
Qi LS, Larson MH, Gilbert LA, Doudna JA, Weissman JS, Arkin AP, Lim WA (2013) Repurposing CRISPR as an RNA-guided platform for sequencespecific control of gene expression. Cell 152(5):1173–1183. https://doi.org/1 0.1016/j.cell.2013.02.022
Shan Q, Wang Y, Li J, Zhang Y, Chen K, Liang Z, Zhang K, Liu J, Xi JJ, Qiu JL, Gao C (2013) Targeted genome modification of crop plants using a CRISPRCas system. Nat Biotechnol 31(8):686–688. https://doi.org/10.1038/nbt.2650
Shiva V (1994) Agriculture and food production. UNESCO/ Environmental Education Dossiers No. 9 (May): 2–3
Spooner, D. (2005). *Molecular markers for genebank management* (No. 10). Bioversity International.
Svitashev S, Young JK, Schwartz C, Gao H, Falco SC, Cigan AM (2015) Targeted mutagenesis, precise gene editing, and site-specific gene insertion in maize using Cas9 and guide RNA. Plant Physiol 169(2):931–945. https:// doi.org/10.1104/pp.15.00793
Tautz, D. and M. Rentz**,** 1984. Simple sequences are ubiquitous repetitive components of eucariotic genomes. Nature. **322**: 652-656
Tautz, D.**,** 1989. Hypervariability of simple sequences as a general source for polymorphic DNA markers. Nucleic Acids Res. **17**: 6463-6471
Uddin, M.S. and Cheng, Q., 2015. Recent application of biotechniques for the improvement of mango research. In *Applied plant genomics and biotechnology* (pp. 195-212). Woodhead Publishing.
Uddin, M.S. and Cheng, Q., 2015. Recent application of biotechniques for the improvement of mango research. In *Applied plant genomics and biotechnology* (pp. 195-212). Woodhead Publishing.
Van Treuren and M.C. de Vicente. 2005. Molecular markers for genebank management. IPGRI Technical Bulletin No. 10. International Plant Genetic Resources Institute, Rome, Italy.
Vendramin G, Morgante M (2006) Genetic diversity in forest tree populations and conservation:analysis of neutral and adaptive variation. In: Ruane J, Sonnino A (eds) The role of biotechnology in exploring and protecting agricultural genetic resources, FAO, Rome, pp 145–147
Zargar SM, Raatz B, Sonah H, Bhat JA, Dar ZA, Agrawal GK, Rakwal R (2015) Recent advances in molecular marker techniques: insight into QTL mapping, GWAS and genomic selection in plants. J Crop Sci Biotechnol 18(5):293–308. https://doi.org/10.1007/s12892-015-0037-5
Zhao J, Perez M, Hu J, Salas FMG (2016) Genome-wide association study for nine plant architecture traits in sorghum. Plant Genome 9(2):1–14