Review paper

Assessment of the genetic diversity in plants using molecular markers: a review and perspective

Arshad Naji Alhasnawi[*1], Yassir Dakheel Kremsh Alasadiy[1], and Febri Doni[2]

[1]Department of Biology, College of Education for Pure Sciences, Al-Muthanna University, Samawah, 66001, Iraq
[2]Department of Biology, Faculty of Mathematics and Natural Sciences, Universitas Padjadjaran, Jatinangor 45363, West Java, Indonesia

[*]Corresponding author email: arshad@mu.edu.iq

The main reasons for food insecurity in the developing world are the rising global population, climate change, and decreasing amount of arable land. Hence, the significance of plant genetic diversity as a subject of research is universally considered important. A significant improvement in the field of molecular genetics includes the apposition of molecular markers for exploring and identifying plant genetic diversity. Molecular markers are now often used in plant breeding research, from identifying the genes that are responsible for the desirable traits to managing different backcrossing breeding strategies. They offer practical solutions to solve the challenges noted in agricultural genomics. In this review article, we have comprehensively reviewed the source of genetic diversity in plants. This review also describes and highlights the application of DNA markers AFLP, ASAP, ASO, CAPS, CAS, Cytoplasmic genome- derived-DNA markers, DGGE, ESTs, ISSR, SSR, RAPD, RBIP, and SNP, which can be used as reliable and efficient tools for identifying genetic diversity of plants.

Keywords: Genetic diversity, plants, DNA, molecular markers

Biological diversity refers to the diversity found in all plant species, their genetic makeup, and the habitats where they exist (Waldman and Shevah 2000). Three types of diversities are noted in biological samples, i.e., species diversity (species richness), genetic diversity (differences in genes and genotypes), and ecosystem diversity (community of species and their environments) (Rao and Hodgkin 2002).

Plant evolution, whether natural or induced by humans, is essentially determined by the population's genetic diversity. The level within or between species is known as diversity. All crop improvement programmes are based on intra- and inter-specific variations (Bhanu 2017). If the members within the species were the same, there would be little opportunity for improving the different traits in the plants as well as their performance (Glaszmann et al. 2010).

Natural variability and divergence among crops have been identified and utilised in crop improvement strategies since the inception of systematic plant breeding techniques. However, there has been a decrease in natural diversity over time due to: (i) Unbalanced breeding strategies which focus on improving only a few features (such as yield and other traits), (ii) The frequent use of a few chosen genotypes as the parents in variety development programmes, and (iii) The transfer of a few excellent lines to several nations, which increases the genetic similarity across modern crop cultivars (Bhanu 2017).

Genetic diversity refers to the degree of genetic variation found within a population or variety of individuals in a species (Govindaraj et al. 2015). It results from the recombination of genetic material i.e., deoxyribonucleic acid (DNA), throughout the inheritance procedure, gene flow, mutations, and genetic drift, and it promotes variations in epigenetic profiles, DNA sequence, protein structures or isoenzymes, and physiology-morphological characteristics traits (Bouchard et al. 2011). Consequently, the genetic makeup of the population determines the

0041–3216/2024/010120–134

diversity of the plant populations (Salgotra and Chauhan 2023).

One approach for evaluating genetic diversity and species relationship is the use of DNA fingerprints, which help in identifying several closely-related plant species (Feng et al. 2020). Because of hydrogen bonding between base pairs, DNA is a very stable material that is not influenced by the changing seasons or plant age (Every and Russu 2007).

Plant extracts are more easily authenticated using two molecular techniques that use genetic markers, such as DNA-based amplified fragment length polymorphism (AFLP) (Choi and Lee 2021; Sorkheh et al. 2019) and arbitrary primed PCR (AP-PCR) (Babu et al. 2021). The other techniques that employ genetic markers are allele-specific associated primers (ASAP) (Mayer et al. 1997), cleaved amplified polymorphic sequences (CAPS) (Matuszczak et al. 2020), cytoplasmic genome-derived-DNA markers (Nam et al. 2021), DNA amplification fingerprinting (DAF) (Caetano-Anollés et al. 1995), expressed sequence tags (ESTs) (Kantety et al. 2002), inter simple sequence repeat (ISSR) (Alhasnawi et al. 2019), microsatellites (Nybom et al. 2014), (Garrido-Cardenas et al. 2018), random amplified polymorphic DNA (RAPD) (Al-Janabi and Alhasnawi 2021), restriction fragment length polymorphism (RFLP) (Amiteye 2021), sequence characterised amplified region (SCAR) (Ravi et al. 2021), single nucleotide polymorphisms (SNPs) (Xu et al. 2021), single strand conformation polymorphism (SSCP) (Hančević et al. 2020), simple sequence repeats (SSR) (Guan et al. 2017), sequence tagged microsatellite loci (STMS) (Caballo et al. 2018). This review examines the potential impact of genetic diversity and the type of DNA markers that are currently being utilised to elucidate the genetic diversity in plants.

Genetic diversity

Genetic diversity is the total number of genetic traits in the genetic composition of a specific species. It differs from genetic variability, which is the propensity for hereditary traits to change (Priyadarshan 2016). Some of the important roles played by genetic diversity in a species include the crop's ability to adapt to population fluctuations, disease threats, and environmental changes. Genetic diversity analysis can be employed for understanding how individual populations and species are affected by mating patterns, population size, migration, mutation, the spatial distribution of individuals, and natural selection (Vinson et al. 2018).

Genetic variation is the diversity in DNA/RNA sequences or gene alleles found in the gene pool of an individual species or population. This can be expressed in several phenotypic forms. However, genetic variety is a broad phrase that refers to all the genetic diversity that exists within and across genotypes that are linked to a single species or between species. The number of unique genes existing in the gene pool can be counted to determine the genetic variety, but genetic variation cannot be counted and is only predicted. Hence, genetic diversity can be viewed as the foundation of genetic diversity (Bhanu 2017).

Several laboratory-based approaches, for example DNA or allozyme analysis, which directly quantify the levels of variation, are frequently used to assess genetic variability within and between inhabitants at the molecular level. It is also possible to characterise and assess genetic diversity using biochemical and morphological methods.

Although morphological characterisation does not entail expensive technology, it may nonetheless be more expensive compared to molecular evaluation because it frequently requires large land areas. These characteristics are frequently subject to phenotypic plasticity; however, it helps in the evaluation of diversity in the case of environmental variation.

Proteins are categorised into different banding patterns for biochemical examination. It is a quick process that uses very little biological material. The resolution of diversity is constrained since there are a limited number of available enzymes.

Molecular analyses include a wide range of DNA molecular markers that can be used for variation analysis. Various genetic traits are

expressed via different markers; they could be co-dominant or dominant, or could amplify anonymous or characterised loci, or include non-expressed or expressed sequences, etc. (Mondini et al. 2009).

Significance of genetic diversity in plant genetic resources

Crop improvement and the existence of cultivated plants in nature are both based on genetic diversity. Genetic diversity provides opportunities for improving the cultivars with desirable features, including both farmer- and breeder-preferred traits. In the early days of agriculture, genetic variability was employed to provide enough food for subsistence. Owing to the changing climate factors that are negatively affecting plant abilities to grow and develop normally, plant breeders have started focusing on developing climate-adapted cultivars. The prevalence of desired alleles is closely correlated with the presence of genetic diversity, which aids in the development of climate-resilient varieties. The growing unpredictability and intensity of drought- related stress brought on by global climate change endangers the sustainability of agricultural production and poses a threat to global food security. Breeding programmes that incorporate modified natural genetic differences can broaden the genetic diversity of stress-tolerant plants and increase the yield of these crops under stress conditions (Begna 2021).

Modern cultivars are designed for a variety of qualities, including increased nutritional value, pest and disease resistances, high yield, and stress tolerance. Plant breeders choose a variety of parents from plant genetic resources in their breeding programmes to create novel crop varieties (Salgotra and Chauhan 2023).

Molecular assessment of genetic diversity

Molecular markers function by drawing attention to variations (polymorphisms) between the nucleic acid sequences in different individuals. A few of these variations include deletions, insertions, duplications, translocations, and point mutations. However, they exclude the activity of particular genes.

Molecular markers are relatively resistant to different environmental factors and display the following advantages: (i) They can be applied to any area in the genome (such as exons, introns, and regulation regions); (ii) They do not display any epistatic or pleiotropic effects; (iii) They can differentiate between the polymorphisms that do not exhibit any phenotypic variation, and, (iv) A few of them are co-dominant (Mondini et al. 2009).

When using molecular approaches to generate data for diversity research, it is crucial to understand the various ways that data might be examined. There are two primary forms of data analysis: analysis of the genetic correlations between the samples, and population genetic parameters (especially, diversity and its division at various levels). The initial process of analysing the genetic links between samples is initiated with the development of a sample × sample pair-wise genetic distance or (similarities) matrix (Govindaraj et al. 2015).

Designing a molecular marker that fulfills the requirements is quite difficult. A marker system that would meet the molecular biology requirements can be identified depending on the sort of study being conducted. DNA polymorphism is measured using a variety of molecular markers, which can be divided into two groups: PCR-based markers and hybridisation-dependent markers. The PCR-based markers include the *in vitro* amplification of specific DNA sequences using the specifically- or arbitrarily-selected oligonucleotide sequences (primers) and a thermostable DNA polymerase enzyme. On the other hand, the hybridisation-based markers visualise the DNA profiles by hybridising the restriction endonuclease-digested DNA fragment to a labelled probe, which is a DNA fragment of a known sequence (Idrees and Isrhad 2014). Electrophoresis is used to separate the amplified DNA fragments, which are visualised using SYBR® safety stains (Noyce et al. 2018).

Due to advancements made in molecular biology, it is possible to conserve both

uncommon and endangered species. The cost-effective method of conserving PGRs is through genetic resource preservation. Several species in biodiversity are in danger of extinction and are difficult to preserve. One of the strongest options for preserving the genetic variability in these resources - which may contain novel genes or alleles that help in ensuring future food security - is DNA storage. Genomic fragments from individual genes or complete genotypes are preserved in the gene library or the library containing DNA samples in a gene bank. Various forms of nuclear matter, such as DNA, RNA, and cDNA, are the different forms of genetic data storage. Such libraries are regarded as the main source of crucial germplasm for upcoming scientific studies around the world (Salgotra and Chauhan 2023).

Various genetic traits are linked to various markers (they could be codominant or dominant, could amplify anonymous or characterised loci, or contain expressed or non-expressed sequences, etc.). A genomic locus that is identified using a probe or specialised starter (primer) is referred to be a molecular marker if its presence identifies the chromosomal trait that it describes, as well as its flanking areas at the 3' and 5' ends (Govindaraj et al. 2015).

Application of the molecular markers in the plants

Molecular markers are being used by plant conservation geneticists to evaluate the long- and short- term survival of plant species with small population sizes, found in changed landscapes, and/or experiencing genetic exchange from different gene pools. The anticipated effects of these risks must be carefully considered by conservation managers (Henry 2012). DNA markers are also popularly used for determining the genetic diversity among wild plants (Nam et al. 2021). In the past, many molecular markers were used for determining genetic diversity during plant breeding experiments (Kumar et al. 2022).

Molecular markers are regarded as tools for several applications that range from gene localisation to improving plant varieties using a marker-assisted selection process. They are also regarded as popular markers that can be used for phylogenetic analysis, which could add novel dimensions to different evolutionary theories. While analysing the history of the development of different molecular markers, it is clear that molecular markers have been enhanced over the past 20 years to offer simple, quick, and automatic support for scientists and researchers. Molecular marker-based genome analysis generates a lot of data and many databases have been constructed for preserving and popularising these markers (Dhawan 2016).

Desirable properties of molecular markers

DNA markers

The variation in DNA nucleotide sequences that are close to or intricately linked to the target gene expressing a trait in different organisms or species refers to a molecular or DNA marker (Hasan et al. 2021). Usually, the tightly-linked molecular marker, target gene-expressing phenotypes, and biological functions are transmitted simultaneously. The locus or loci, which are the specific genomic locations of molecular markers in the chromosomes, could be identified or unidentified (Amiteye 2021).

Several molecular markers focus on various genomic regions, which offer specific and varied information regarding the actual diversity present in the genotypes (Rao et al. 2020). Using particular or randomly-constructed oligonucleotide primers, the PCR method enables the practical amplification of specific DNA sequences from genomic DNA segments. Molecular markers are incredibly helpful tools that are available for plant improvement research. Due to point mutations in the oligonucleotide priming sites, genotypes genetically display different pools of fragments. In some instances, polymorphism may result from the insertion or deletion mutation events that change the distances between the terminal sequences.

Genomic analysis of plants now frequently uses DNA marker techniques that are mediated by polymerase chain reaction (PCR) applications. Progress in many areas of plant genetics has tremendously benefited a greater understanding of molecular markers, the genetic diversity of plant species, and the effectiveness of plant molecular breeding programmes. The current rapid advancements in molecular marker technology have led to unique methodologies that substantially facilitate research in every field of crop development and improvement (Amiteye 2021).

DNA fingerprinting

The process of creating a set of different DNA fragments using a single DNA sample is known as DNA fingerprinting. The synthesised DNA fragments are utilised to generate genotypic data. Many techniques are used for generating DNA fingerprinting patterns. DNA fingerprinting is thought to be objective and, apparently, a less error-prone means of determining the varieties compared to the traditional approaches, especially when certain varieties exhibit little morphological differences. This method extracts genetic material from a field sample and compares it to a reference library, which is a collection of genetic sequences from improved and unimproved varieties. The samples are categorised as a specific variety depending on the genetic similarity and dissimilarity to the sequences in the reference libraries, within a certain tolerance (Poets et al. 2020).

The selection of an appropriate technique is based on the organism that is studied and the research subject under investigation. All DNA fingerprinting methods investigate the sequence patterns that are linked to the genetic markers; however, each method examines a different set of genetic markers, both with regards to quantity and variety. For instance, some methods (referred to as single-locus markers) permit the evaluation of a molecular marker at one locus, while others permit the simultaneous examination of many loci (known as multi-locus markers). Some strategies concentrate on the co-dominant markers, which reveal details about the two alleles at a specific locus. Other strategies focus on dominant markers, which can only reveal whether a specific allele is present or absent and cannot reveal whether an individual is homozygous for that specific allele (Rana and Bhat 2017).

PCR-based markers

Amplified fragment length polymorphism (AFLP)

The ability of AFLP to simultaneously screen representative DNA areas that are scattered randomly over the genome is known as "genome representation" capability, which is its key strength. Each organism's DNA can produce AFLP markers without investing in the development of primer/probes and sequence analysis. Nucleic acid digestion can use high-quality and partially- degraded DNA samples, however, the DNA has to be free of PCR and restriction enzyme inhibitors (Idrees and Isrhad 2014). The application of a single primer results in the amplification of multiple DNA fragments that are dispersed randomly throughout the genome. Before amplifying the DNA for differentiation, AFLP uses restriction enzyme digestion and insertion of adapters for primer annealing. All these procedures generate several bands, and the nucleic acid fragments are separated using gel electrophoresis, which helps in examining the polymorphism in different isolates. RAPD primers are typically short, i.e., 8 – 12 mer long with a random sequence composition. The amplification process is sensitive to the reaction conditions, specifically the annealing temperature (Sachse and Frey 2002).

Allele-specific associated primers (ASAP)

SNP are also identified by constructing allele-specific primers for specific SNPs locations. The ASAP method employs distinct fluorochromes connected at the 5`-end and polymorphic nucleotides at the 3`-end of both primers.

Amplification of the pre-amplified DNA using the allele- specific primers results in an allele-specific produce that is easily visible on polyacrylamide gel or an automatic sequencer (Adhikari et al. 2017).

Allele-specific oligonucleotide (ASO)

ASO probes are typically short oligonucleotides (15 – 17 mers) having a low Guanine-Cytosine content (GC content) (≈ 30 - 50%) and are constructed in a manner that the distinguishing nucleotide lies in the centre of the probe. The proper selection of ASO probes necessitates a substantial investment of time and testing of multiple candidate probes, each of which must be examined separately using both the positive and the negative control samples, under fixed hybridisation and wash conditions. Additionally, polymorphic sequences need to be avoided, which may result in misleading negative results if the polymorphism affects annealing and/or destabilises the ASO/target hybrid (Bertorini 2008).

Cleaved amplified polymorphic sequences (CAPS)

CAPS, also known as PCR-RFLPs, refers to a technique that combines PCR and RFLP. It uses amplified DNA fragments that have been digested using a restriction endonuclease to identify restriction site polymorphisms. The PCR technique is used for amplifying the target DNA, followed by the restriction enzyme digestion step. As a result, CAPS markers depend on the variations in the PCR fragment-digested patterns using restriction enzymes that result from nucleotide polymorphism between the samples (Nam et al. 2020). The PCR method uses primers (generally 20 – 25 bps in size) to amplify the DNA fragments from target genotypes. The DNA sequences that were used for designing the PCR primers were acquired from gene banks like NCBI and cDNA sequences, or could be acquired from the isolated and cloned polymorphic RAPD bands. It must be noted that the primers utilised in CAPS-PCR are tailored to target introns or 30 untranslated regions, increasing the likelihood of finding polymorphisms. The PCR- amplified products of the target gene or genomic region typically range between 300 – 2000 bps. Following PCR amplification, each of the resulting PCR fragments is then individually restricted using the right combination of restriction enzymes. A few restriction enzymes could be more polymorphic compared to others. Hence, using a variety of restriction enzymes to screen polymorphism is effective. By separating the restriction fragments on gel electrophoresis, the digested PCR fragments are then tested for polymorphism. CAPS can be described as polymorphic co-dominant markers that provide important information and CAPS genotypes can be easily determined, scored, and analysed. The CAPS molecular technique is a highly popular technique that can be used for identifying single nucleotide changes or polymorphisms (Amiteye 2021).

Coupled amplification and sequencing (CAS)

Coupled amplification and sequencing helps in directly sequencing the DNA segment from the genomic DNA. The target is chosen and amplified during the first stage of PCR amplification. Thereafter, the target segment's two strands are concurrently sequenced and amplified. Researchers have shown how CAS can quickly recognise the different base pairs. By limiting CAS to the variant bases, it is possible to genotype a population quickly and directly for the pre- determined sequence variation, from the genomic DNA for every organism. The method replaces the development and modification of alternative typing assays that depend on oligonucleotide ligation or hybridisation. Additionally, it is easy to combine the second stage of CAS with competitive oligonucleotide priming using allelic primers. Combining different techniques enables the direct haplotyping of polymorphism at the priming site with others included in the amplified segment, along with the sequencing of a single

chromosome from a heterozygous genomic sample (Ruano and Kidd 1991).

Cytoplasmic genome-derived-DNA markers

Cytoplasmic genome-derived DNA markers, also known as organelle DNA markers, are genetic markers that are derived from the DNA found in the cytoplasmic organelles of a cell, specifically the mitochondria and chloroplasts. These markers are often used in molecular biology and genetics research to study various aspects of evolution, population genetics, and phylogenetics (Camus et al. 2022).

The cytoplasmic genome, such as the DNA of chloroplast genome, and mitochondrial genome can also yield DNA markers (Nam et al. 2020). Because the plant cytoplasmic genome, which contains the mitochondrial and chloroplast genomes, is only transmitted through the maternal parent, cytoplasmic markers can be employed in breeding programmes to distinguish cytoplasmic genome types and trace maternal ancestors.

The cytoplasmic genome is inherited from the mother in the majority of angiosperms and is made up of mitochondrial and chloroplast DNA (Sanetomo and Gebhardt 2015). Mitochondrial DNA (mtDNA) markers such as Cytochrome c oxidase subunit I (COI) are frequently used for DNA barcoding in animal species identification (Hebert et al. 2003). The D-loop (control region) is a non-coding region in mitochondrial DNA used for studying population genetics and phylogenetics in various organisms (Searle 2000).

Denaturing gradient gel electrophoresis (DGGE)

Denaturing gradient gel electrophoresis is a common method for analysing the composition of microbial populations (Lv et al. 2017), analysing DNA fragments based on their nucleotide sequence-dependent melting behaviour. It is an effective technique for identifying mutations and variations in DNA samples. (Gafan and Spratt 2005). Analysis of 16S rRNA genes using DGGE to profile complex microbial populations laid the foundation for the application of DGGE in microbial ecology and diversity studies (Muyzer et al. 1993). The sensitivity of DGGE for detecting single base substitutions and mutations in DNA fragments when a GC-clamp is added to the target DNA, highlights the diagnostic potential of DGGE in mutation analysis (Myers et al. 1985). Use of DGGE, in combination with terminal restriction fragment length polymorphism (T- RFLP) analysis, to study microbial diversity by analysing 16S rRNA genes, demonstrates the utility of DGGE as a tool for fingerprinting microbial communities (Liu et al. 1997).

Expressed sequence tags (ESTs)

Expressed sequence tag (EST) and database for expressed sequence tags (dbEST) markers are important resources in genomics and molecular biology. ESTs are short DNA sequences representing expressed genes, while DBEST is a database containing these EST sequences. These markers are valuable for gene discovery, expression analysis, and functional genomics. The EST division of GenBank, dbEST, is a large repository of data (Wolfsberg and Landsman 1997).

Expressed sequence tags (ESTs) are short DNA sequences, 200 – 500 nucleotides, that are produced by sequencing the 5′ and/or 3′ ends of the complementary DNA (cDNA), which are then grouped and counted. Analysis of ESTs is seen to be a novel technique that helps in screening the genes that could be linked to variations in prognoses (Tsui et al. 2014). ESTs are used for novel gene discovery, particularly for organisms whose economies of scale prevent comprehensive genome sequencing (Parkinson and Blaxter 2009). Since full genome sequences for diverse plant species are unavailable, EST analysis is regarded as an effective method for identifying conserved miRNAs and studying the evolution and conservation of miRNAs in various species. The identification of novel

miRNAs from a variety of plant species has been accomplished using EST analysis techniques (Gupta et al. 2015).

Inter-simple sequence repeat (ISSR)

Inter-simple sequence repeat (ISSR) markers have a wide range of uses and are polymorphic, simple to design, and affordable in comparison to other approaches (Mandal et al. 2019). Furthermore, despite the diverse plant phylogenies, it is possible to utilise the same ISSR primers, universally. The fundamental ISSR methodology is adaptable and is modified with alternatives for implementing a wide range of projects and budgets. Generally, manual agarose-based ISSRs can be regarded as a good, safe, simple, and inexpensive way to infer plant genetic diversity (Gemmill and Grierson 2021). This technique makes use of primers called microsatellites, which included the di-, tri-, and tetra- or pentanucleotide repeats. Generally, primers with a length of 15 – 30 bases are employed in this method. These primers used in the ISSR could be unanchored or anchored at the 3' or 5'-end with 1 - 4 degenerate bases that can be extended into flanking sequences (Nadeem et al. 2018). The ISSR markers are useful and polymorphic and help in evaluating phylogeny, genome mapping, gene tagging, evolutionary biology, and genetic diversity. Furthermore, ISSR can detect the variation without requiring any sequence data (Mohamad et al. 2017).

Microsatellites (SSR simple sequence repeats)

DNA fragments containing a tandem repeat motif of 1 - 6 nucleotides are called simple sequence repeats (SSR), also referred to as microsatellites. Some appealing characteristics of SSR markers include robust and high repeatability, genome-wide coverage, transferability between species, high polymorphism with many alleles in every locus, co-dominant inheritance, and low instrumentation and knowledge requirements (Bhattarai et al. 2021).

Polymerase chain reaction (PCR) is frequently employed for analysing the SSRs or microsatellites, and it presents many advantages. In general, SSR markers are robust, repeatable, simple to score (or record), and can be easily automated for high-throughput detection. They are also co-dominantly expressed. Moreover, numerous SSRs are multiplexed and identified using gel electrophoresis, non-radioactive DNA sequences, or capillary electrophoresis. SSRs have been mapped in a variety of crops because of recent developments in genomics. It is crucial to choose high-quality SSR that are polymorphic, unlinked, dependable, reproducible, and free of stuttering for successful microsatellite analysis. A reliable source should synthesise the primers for any type of marker system. Certain SSRs are accessible to the public. High-quality markers must be developed at a significant cost if they are not already protected by intellectual property rights (Rana and Bhat 2017).

Random amplified polymorphic DNA (RAPD)

The RAPD application methodology uses arbitrarily short oligonucleotide primers, ranging between 8 - 15 nucleotides, to randomly amplify the DNA areas in the large genomic DNA (Amiteye 2021).

A single set of primers can be used to amplify numerous loci using a small amount of DNA, which makes RAPD a quick, easy, and affordable method. The RAPD application can be used to identify important features (Nam et al. 2021). Genetic diversity studies continue to make extensive use of several markers that display a good performance during these analyses, like RAPDs. Since no method can be used for all applications, scientists should weigh the advantages and disadvantages of each method before initiating a new study (Idrees and Isrhad 2014).

Retrotransposon-based insertion polymorphism (RBIP)

The RBIP technique focuses on a single retrotransposon insertion site and determines the polymorphism based on whether the insertion is present or not. When an allele has a retrotransposon insertion, the flanking region, and long terminal repeat, LTR-specific primer sets can be used to amplify the genomic region. Hence, RBIP markers provide a precise DNA profile and help in detecting polymorphism caused by the integration of an element at a specific locus (Nashima et al. 2017). Information on the sequence of sections near the retrotransposon insertion sites is required for RBIP characterisation, which is a significant problem (Nashima et al. 2017).

The benefits of RBIP include the ease with which it can be automated and the sample throughput is increased by using techniques such as TaqManTM or DNA chip technology, which do not require gel analysis (Amiteye 2021).

Single nucleotide polymorphism (SNP)

Haplotype-driven breeding, a promising and contemporary approach to developing customised crop varieties, focuses on the recognition of superior haplotypes and their integration into breeding programmes (Sinha et al. 2020). Haplotype blocks have been proposed as independent variables in genomic prediction due to their potential to offer additional information when compared to single SNPs and utilising haplotype blocks enhances the precision of genomic prediction for both quantitative and qualitative traits in plants (Difabachew et al. 2023). When there is a difference of a single nucleotide (A, T, G, or C) across members of a species, it is known as a SNP. SNP is the most prevalent marker system in plant and animal genomes, and they have lately been used for the generation of molecular technique, for numerous applications. They have co-dominant or binary status and can effectively distinguish between the homozygous and heterozygous alleles (Idrees and Isrhad 2014).

An SNP can be described as the difference between a single nucleotide base between two individuals or DNA sequences. Since a single nucleotide base is the smallest unit of heredity, SNPs offer the most basic and comprehensive type of molecular markers. Thus, they can also offer a high marker density. The high density of SNP markers enhances the likelihood of discovering polymorphisms in the target gene, giving them a significant advantage over earlier markers that were only marginally associated with the locus of interest (Kumar et al. 2018).

Recent advances in sequencing technology have made it easier to identify SNPs and insertions/deletions, which account for most changes across alleles. Given many gene and EST sequences, as well as the high level of intraspecific nucleotide variability found in different crop species, such as maize and soybean, SNP identification is often not difficult. SNP haplotypes can be directly read out for these crop species. Analysis based on haplotypes is more insightful than individual SNP-based analysis, and it shows a higher analytical power when assessing the associations with phenotypes. A few of the elite germplasms in the crops have experienced bottlenecks, which increased the amount of existing linkage disequilibrium (LD) and made it easier to associate the SNP haplotypes at the candidate gene loci to phenotypes. Furthermore, the whole-genome scans might be effective in identifying the specific regions in the genome that are linked to intriguing traits, if sufficient LD was noted. The application of the SNPs and indel markers is more appealing due to technological advancements for high-throughput marker-assisted breeding, EST mapping, and the fusion of physical and genetic maps (Rafalski 2002).

Genetic diversity analysis in plants

Compared to techniques for transcriptome, whole genome, or resequencing, molecular markers are an effective and inexpensive tool for examining genetic variation and constructing core germplasm libraries. Microsatellite markers are frequently used in genetic breeding, diversity

identification, evaluation of germplasm diversity, and conservation (Liu et al. 2022).

The survival and growth of humans depend heavily on plants. Plants have been domesticated for a very long time using desirable features. DNA markers have garnered a lot of research interest in genetic diversity. DNA markers have been employed in plant breeding to promote cross-breeding and evaluate propagation materials (Nam et al. 2021). Genetic variation refers to the variety in DNA/RNA sequences or gene alleles found in the gene pool of a population or species. It is expressed as alternate phenotypic forms. While genetic diversity can be quantified by counting the number of distinct genes in a gene pool, genetic variation is an inherent phenomenon that is expected to exist but cannot be directly measured. Therefore, genetic variability can be regarded as the foundational element contributing to genetic diversity (Bhanu 2017).

Conclusions

The genetic diversity in crop plants forms the basis for the enhancement of the production of novel varieties which leads to better food and nutritional security. Thus, it is essential to determine the various genetic resources using various statistical approaches and use them in a breeding strategy. Molecular markers are a class of molecular tools that are constantly evolving because they are particularly responsive to new genome-based discoveries and technological developments. Most molecular marker approaches are used to construct genetic and physical maps and assess genetic diversity. It is simpler to connect the genetic distances and physical distances when the linked markers are physically mapped. Using molecular markers to measure and preserve genetic diversity is crucial because it acts as a reservoir that helps in adapting to a variety of environmental and other changes.

Competing interests

The authors declare that they have no competing interests.

References

Adhikari, S., S. Saha, A. Biswas, T.S. Rana, T.K. Bandyopadhyay, and P. Ghosh. 2017. "Application of Molecular Markers in Plant Genome Analysis: A Review." *The Nucleus* **60 (3):** 283–297. https://doi.org/10.1007/s13237-017-0214-7.

Al-Janabi, A.S.A., and A.N. Alhasnawi. 2021. "Determination of Genetic Diversity Based on RAPD Molecular Marker and ParARF3 Gene Expressions in Some Apricot Genotypes in Iraq." *Basrah Journal of Agricultural Sciences* **34 (2):** 60–74. https://doi.org/10.37077/25200860.2021.34.2.05

Alhasnawi, A.N., A.M. Mandal, and H.M. Jasim. 2019. "Using DNA Fingerprinting to Detect the Genetic Relationships in Acacia by Inter-Simple Sequence Repeat Markers." *Journal of Pure and Applied Microbiology* **13 (1):** 281–288. https://dx.doi.org/10.22207/JPAM.13.1.30.

Amiteye, S. 2021. "Basic Concepts and Methodologies of DNA Marker Systems in Plant Molecular Breeding." *Heliyon* **7 (10):** e08093. https://doi.org/10.1016/j.heliyon.2021.e08093

Babu, K.N., T.E. Sheeja, D. Minoo, M.K. Rajesh, K. Samsudeen, E.J. Suraby, and I.P.V. Kumar. 2021. "Random Amplified Polymorphic DNA (RAPD) and Derived Techniques." *Methods Mol Biol.* **2222:**219–247. doi: 10.1007/978-1-0716-0997-2_13

Begna, T., and T. Begna. 2021. "Role and Economic Importance of Crop Genetic Diversity in Food Security." *International Journal of Agricultural Science and Food Technology* **7 (1):** 164–169. https://doi.org/10.17352/2455-815x.000104

Bertorini, T.E. 2008. "Neurologic Evaluation and Ancillary Tests." In *Neuromuscular Case Studies*, 27–76. Elsevier. https://doi.org/10.1016/b978-0-7506-7332-7.50006-4.

Bhanu, A.N. 2017. "Assessment of Genetic Diversity in Crop Plants - An Overview." *Adv Plants Agric Res.* **7 (3):** 279–286. https://doi.org/10.15406/apar.2017.07.00255.

Bhattarai, G., A. Shi, D.R. Kandel, N. Solís-Gracia, J.A. da Silva, and C.A. Avila. 2021. "Genome-Wide Simple Sequence Repeats (SSR) Markers Discovered from Whole-Genome Sequence Comparisons of Multiple Spinach Accessions." *Scientific Reports* **11 (1):** 9999. https://doi.org/10.1038/s41598-021-89473-0.

Bouchard, C., T. Rankinen, and J.A. Timmons. 2011. "Genomics and Genetics in the Biology of Adaptation to Exercise." *Compr Physiol*. **1 (3):** 1603–48. doi: 10.1002/cphy.c100059.

Caballo, C., P. Castro, J. Gil, I. Izquierdo, T. Millan, and J. Rubio. 2018. "STMS (Sequence Tagged Microsatellite Site) Molecular Markers as a Valuable Tool to Confirm Controlled Crosses in Chickpea (*Cicer arietinum* L.) Breeding Programs." *Euphytica* **214 (12):** 1–9. https://doi.org/10.1007/S10681-018-2314-0/METRICS

Caetano-Anollés, G., L.M. Callahan, P.E. Williams, K.R. Weaver, and P.M. Gresshoff. 1995. "DNA Amplification Fingerprinting Analysis of Bermudagrass (Cynodon): Genetic Relationships Between Species and Interspecific Crosses." *Theoretical and Applied Genetics* **91 (2):** 228–235. https://doi.org/10.1007/BF00220882/METRICS

Camus, M.F., B. Alexander-Lawrie, J. Sharbrough, and G.D.D. Hurst. 2022. "Inheritance Through the Cytoplasm." *Heredity* **129 (1):** 31–43. https://doi.org/10.1038/s41437-022-00540-2

Choi, T.Y., and S.R. Lee. 2021. "A Review of Intraspecific Genetic Diversity on Wild Plants in Korea Estimated from Varying Nuclear DNA Markers." *Journal of Asia-Pacific Biodiversity* **14 (4):** 622–627. https://doi.org/10.1016/j.japb.2021.08.004.

Dhawan, S.S. 2016. "Role of Molecular Markers in Assessing Genetic Diversity in Mentha: A Review." *Scientific J Genet Gene Ther*. **2 (1):** 022–026.

Difabachew, Y.F., M. Frisch, A.L. Langstroff, A. Stahl, B. Wittkop, R.J. Snowdon, M. Koch, M. Kirchhoff, L. Cselényi, M. Wolf, J. Förster, S. Weber, U.J. Okoye, and C. Zenke-Philippi. 2023. "Genomic Prediction with Haplotype Blocks in Wheat. *Frontiers in Plant Science* **14:** 1168547. https://doi.org/10.3389/fpls.2023.1168547

Every, A.E., and I.M. Russu. 2007. "Probing the Role of Hydrogen Bonds in the Stability of Base Pairs in Double-Helical DNA." *Biopolymers* **87 (2–3):** 165–173. https://doi.org/10.1002/bip.20811

Feng, T., Q. Jia, X. Meng, X. Chen, F. Wang, W. Chai, and Z. Liang. 2020. "Evaluation of Genetic Diversity and Construction of DNA Fingerprinting in *Polygonatum* Mill. based on EST-SSR and SRAP Molecular Markers." *3 Biotech.* **10 (7):** 322. https://doi.org/10.1007/s13205-020-02316-z

Gafan, G.P., and D.A. Spratt. 2005. "Denaturing Gradient Gel Electrophoresis Gel Expansion (DGGEGE) – An Attempt to Resolve the Limitations of Co-Migration in the DGGE of Complex Polymicrobial Communities." *FEMS Microbiology Letters* **253 (2):** 303–307. https://doi.org/10.1016/j.femsle.2005.09.048.

Garrido-Cardenas, J.A., C. Mesa-Valle, and F. Manzano-Agugliaro. 2018. "Trends in Plant Research Using Molecular Markers." *Planta* **247 (3):** 543–557. https://doi.org/10.1007/S00425-017-2829-Y/METRICS

Gemmill, C.E.C., and E.R.P. Grierson. 2021. "Inter-Simple Sequence Repeats (ISSR), Microsatellite-Primed Genomic Profiling Using Universal Primers." *Methods in Molecular Biology* **2222:** 249–262. https://doi.org/10.1007/978-1-0716-0997-2_14/COVER

Glaszmann, J.C., B. Kilian, H.D. Upadhyaya, and R.K. Varshney. 2010. "Accessing Genetic Diversity for Crop Improvement." *Current Opinion in Plant Biology* **13 (2):** 167–173. https://doi.org/10.1016/J.PBI.2010.01.004

Govindaraj, M., M. Vetriventhan, and M. Srinivasan. 2015. "Importance of Genetic Diversity Assessment in Crop Plants and its Recent Advances: An Overview of its Analytical Perspectives." *Genetics Research International* **2015:** 1–14. https://doi.org/10.1155/2015/431487

Guan, X., N. Yuyama, A. Stewart, C. Ding, N. Xu, T. Kiyoshi, and H. Cai. 2017. "Genetic Diversity and Structure of *Lolium* Species Surveyed on Nuclear Simple Sequence Repeat and Cytoplasmic Markers." *Frontiers in Plant Science* **8:**1–9. https://doi.org/10.3389/fpls.2017.00584

Gupta, H., T. Tiwari, M. Patel, A. Mehta, and A. Ghosh, A. 2015. "An Approach to Identify the Novel miRNA Encoded from H. Annuus EST Sequences." *Genomics Data* **6:**139–144. https://doi.org/10.1016/J.GDATA.2015.09.005

Hančević, K., P. Saldarelli, M. Čarija, S. Černi, G. Zdunić, A. Mucalo, and T. Radić. 2020. "Predominance and Diversity of GLRaV-3 in Native Vines of Mediterranean Croatia." *Plants* **1 (1):** 17. https://doi.org/10.3390/PLANTS10010017

Hasan, N., S. Choudhary, N. Naaz, N. Sharma, and R.A. Laskar. 2021. "Recent Advancements in Molecular Marker-Assisted Selection and Applications in Plant Breeding Programmes. *Journal of Genetic Engineering and Biotechnology* **19 (1):** 128. https://doi.org/10.1186/s43141-021-00231-1

Hebert, P.D.N., A. Cywinska, S.L. Ball, and J.R. DeWaard. 2003. "Biological Identifications Through DNA Barcodes." *Proceedings of the Royal Society of London. Series B: Biological Sciences* **270 (1512):** 313–321. https://doi.org/10.1098/rspb.2002.2218

Henry, R.J. 2012. "Molecular Markers in Plants. In Molecular Markers in Plants." https://doi.org/10.1002/9781118473023

Idrees, M., and M. Isrhad. 2014. "Molecular Markers in Plants for Analysis of Genetic Diversity: A Review." *European academic Research* **2 (1):** 1513–1540.

Kantety, R.V., M. La Rota, D.E. Matthews, and M.E. Sorrells. 2002. "Data Mining for Simple Sequence Repeats in Expressed Sequence Tags from Barley, Maize, Rice, Sorghum and Wheat." *Plant Molecular Biology* **48 (5–6):** 501–510. https://doi.org/10.1023/A:1014875206165/METRICS

Kumar, A., N. Longmei, P. Kumar, and P. Kaushik. 2022. "Molecular Marker Analysis of Genetic Diversity in Maize: A Review." *OBM Genetics* **6 (1):** 1–21. https://doi.org/10.21926/obm.genet.2201150

Kumar, M., V. Haudhary, U. Siroh, M.K. Singh, S. Malik, and R. Naresh. 2018. "Biochemical and Molecular Markers for Characterization of Chrysanthemum Germplasm: A Review." *Journal of Pharmacognosy and Phytochemistry* **7 (5):** 2641–2652.

Liu, J., S. Gao, Y. Xu, M. Wang, J.J. Ngiam, N.C. Rui Wen, J.J.J. Yi, X. Weng, L. Jia, and J. Salojärvi. 2022. "Genetic Diversity Analysis of Sapindus in China and Extraction of a Core Germplasm Collection using EST-SSR Markers." *Frontiers in Plant Science* **13:** 1698. https://doi.org/10.3389/FPLS.2022.857993/BIBTEX

Liu, W.T., T.L. Marsh, H. Cheng, and L.J. Forney. 1997. "Characterization of Microbial Diversity by Determining Terminal Restriction Fragment Length Polymorphisms of Genes Encoding 16S rRNA." *Applied and Environmental Microbiology* **63 (11):** 4516–4522. https://doi.org/10.1128/aem.63.11.4516-4522.1997

Lv, X.C., Y.J. Jiang, J. Liu, W.L. Guo, Z.B. Liu, W. Zhang, P.F. Rao, and L. Ni. 2017. "Evaluation of Different PCR Primers for Denaturing Gradient Gel Electrophoresis (DGGE) Analysis of Fungal Community Structure in Traditional Fermentation Starters used for Hong Qu Glutinous Rice Wine." *International Journal of Food Microbiology* **255:**58–65. https://doi.org/10.1016/j.ijfoodmicro.2017.05.010

Mandal, A., A. Alhasnawi, H. Jasim, and A. Mohamad. 2019. "Evaluation of Salt Stress and Molecular Analysis of Genetic Variation of Iraqi Rice Cultivars." *Biodiversitas Journal of Biological Diversity* **20 (11):** 3309–3314. https://doi.org/ 10.13057/biodiv/d201125

Matuszczak, M., S. Spasibionek, K. Gacek, and I. Bartkowiak-Broda. 2020. "Cleaved Amplified Polymorphic Sequences (CAPS) Marker for Identification of Two Mutant Alleles of the Rapeseed BnaA.FAD2 Gene." *Molecular Biology Reports* **47 (10):** 7607–7621. https://doi.org/ 10.1007/S11033-020-05828-2/FIGURES/4

Mayer, M.S., A. Tullu, C.J. Simon, J. Kumar, W.J. Kaiser, J.M. Kraft, and F.J. Muehlbauer. 1997. "Development of a DNA Marker for Fusarium Wilt Resistance in Chickpea." *Crop Science* **37 (5):** 1625–1629. https://doi.org/10.2135/CROPSCI1997.0011183X003700050036X

Mohamad, A., A.N. Alhasnawi, A.A. Kadhimi, A. Isahak, W.M.W. Yusoff, and C.C. Radziah. 2017. "DNA Isolation and Optimization of ISSR-PCR Reaction System in *Oryza sativa* L." *International Journal on Advanced Science, Engineering and Information Technology* **7 (6):** 2088–5334.

Mondini, L., A. Noorani, and M.A. Pagnotta. 2009. "Assessing Plant Genetic Diversity by Molecular Tools." *Diversity* **1 (1):** 19–35. https://doi.org/10.3390/d1010019

Muyzer, G., E.C. de Waal, and A.G. Uitterlinden. 1993. "Profiling of Complex Microbial Populations by Denaturing Gradient Gel Electrophoresis Analysis of Polymerase Chain Reaction-Amplified Genes Coding for 16S rRNA." *Applied and Environmental Microbiology* **59 (3):** 695–700. https://doi.org/10.1128/aem.59.3.695-700.1993

Myers, R.M., S.G. Fischer, L.S. Lerman, and T. Maniatis. 1985. "Nearly All Single Base Substitutions in DNA Fragments Joined to a GC-Clamp can be Detected by Denaturing Gradient Gel Electrophoresis." *Nucleic Acids Research* **13 (9):** 3131–3145. https://doi.org/10.1093/nar/13.9.3131

Nadeem, M.A., M.A. Nawaz, M.Q. Shahid, Y. Doğan, G. Comertpay, M. Yıldız, R. Hatipoğlu, F. Ahmad, A. Alsaleh, N. Labhane, H. Özkan, G. Chung, and F.S. Baloch. 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/10.1080/13102818.2017.1400401

Nam, V.T., P.L.B. Hang, N.N. Linh, L.H. Ly, H.T.T. Hue, N.H. Ha, H.H. Hanh, and L.T.T. Hien. 2021. "Molecular Markers for Analysis of Plant Genetic Diversity." *Vietnam Journal of Biotechnology* **18 (4):** 589–608. https://doi.org/10.15625/1811-4989/18/4/15326

Nashima, K., Terakami, S., Kunihisa, M., Nishitani, C., Shoda, M., Matsumura, M., Onoue- Makishi, Y., Urasaki, N., Tarora, K., Ogata, T., and Yamamoto, T. 2017. "Retrotransposon- based Insertion Polymorphism Markers in Mango." *Tree Genetics and Genomes* **13 (5):** 1–10. https://doi.org/10.1007/s11295-017-1192-2.

Noyce, R.S., S. Lederman, and D.H. Evans. 2018. "Construction of an Infectious Horsepox Virus Vaccine from Chemically Synthesized DNA Fragments." *PLOS ONE* **13 (1):** e0188453. https://doi.org/10.1371/journal.pone.0188453.

Nybom, H., K. Weising, and B. Rotter. 2014. "DNA Fingerprinting in Botany: Past, Present, Future." *Investigative Genetics* **5 (1):** 1–35. https://doi.org/10.1186/2041-2223-5-1

Parkinson, J., and M. Blaxter. 2009. "Expressed Sequence Tags: An Overview." In *Methods in Molecular Biology*. **533**: 1–12. https://doi.org/10.1007/978-1-60327-136-3_1

Poets, A., K. Silverstein, P. Pardey, S. Hearne, and J. Stevenson. 2020. "DNA Fingerprinting for Crop Varietal Identification: Fit-for-Purpose Protocols

and Their Cost and Analytical Implications." In *Rome: Standing Panel on Impact Assessment (SPIA)*. **April**: 1– 65.

Priyadarshan, P.M. 2016. *Genetic Diversity and Erosion in Hevea Rubber*, edited by M.R. Ahuja and S.M. Jain. **8**:233–267. Springer International Publishing. https://doi.org/10. 1007/978-3-319-25954-3_6.

Rafalski, A. 2002. "Applications of Single Nucleotide Polymorphisms in Crop Genetics." *Current Opinion in Plant Biology* **5 (2):** 94–100. https://doi.org/ 10.1016/S1369-5266(02)00240-6

Rana, M., and K. Bhat. 2017. "DNA Fingerprinting in Plants: Standard Operating Methods and Protocols." ICAR-National Bureau of Plant Genetic Resources.

Rao, G.K., C. Kapadia, N.B. Patel, K.D. Desai, and P.N. Narasimha Murthy. 2020. "Genetic Diversity Analysis of Greater Yam (*Dioscorea alata* L.) Genotypes through RAPD and ISSR Markers." *Biocatalysis and Agricultural Biotechnology* **23:** 101495. https://doi.org/ 10.1016/j.bcab.2020.101495.

Rao, V.R., and T. Hodgkin. 2002. "Marian Feldman." *Plant Cell, Tissue and Organ Culture* **68:** 1–19. https://www.cbd. int/doc/articles/2002-/A-00109.pdf

Ravi, D., E.A. Siril, and B.R. Nair. 2021. "SCAR Marker Development for the Identification of Elite Germplasm of *Moringa Oleifera* Lam.-A Never Die Plant." *Plant Molecular Biology Reporter* **39 (4):** 850–861. https://doi.org/10. 1007/s11105-021-01300-y.

Ruano, G., and K.K. Kidd. 1991. "Genotyping and Haplotyping of Polymorphisms Directly from Genomic DNA Via Coupled Amplification and Sequencing (CAS)." *Nucleic Acids Research* **19 (24):** 6877–6882. https://doi.org/10.1093/nar/19.24. 6877.

Sachse, K., and J. Frey. 2002. "PCR Detection of Microbial Pathogens." In *PCR Detection of Microbial Pathogens*. https://doi.org/10.1385/1592593445

Salgotra, R.K., and B.S. Chauhan. 2023. "Genetic Diversity, Conservation, and Utilization of Plant Genetic Resources." *Genes* **14 (1):** 174. https://doi.org/10.3390 /genes14010174.

Sanetomo, R., and C. Gebhardt. 2015. "Cytoplasmic Genome Types of European Potatoes and Their Effects on Complex Agronomic Traits." *BMC Plant Biology* **15 (1):** 1–16. https://doi.org/10.1186/S12870-015-0545-Y/TABLES/7.

Searle, J.B. 2000. *Phylogeography - The History and Formation of Species*, edited M. Pessarakl, 2nd ed. Harvard University Press, Cambridge Marcel Dekker, Inc, New York.

Sinha, P., V.K. Singh, R.K. Saxena, A.W. Khan, R. Abbai, A. Chitikineni, A. Desai, J. Molla, J. H.D. Upadhyaya, A. Kumar, and R.K. Varshney. 2020. "Superior Haplotypes for Haplotype-Based Breeding for Drought Tolerance in Pigeonpea (*Cajanus cajan* L.)." *Plant Biotechnology Journal* **18 (12):** 2482–2490. https://doi.org/10.1111/pbi.13422.

Sorkheh, K., M. Masaeli, M.H. Chaleshtori, A. Adugna, and S. Ercisli. 2019. "Expression of Concern: AFLP-Based Analysis of Genetic Diversity, Population Structure, and Relationships with Agronomic Traits in Rice Germplasm from North Region of Iran and World Core Germplasm Set." *Biochemical Genetics* **57 (745):** https://doi.org/10. 1007/s10528-019-09933-1.

Tsui, K.-H., P. Yang, Y.-C. Hsu, C.-P. Hou, Y.-H. Lin, P.-L. Chang, and H.-H. Juang. 2014. "The Expression Sequence Tag is an Effective Method for Screening DNA Segments that Predict Urinary Bladder Transitional Cell Carcinoma Prognosis." *Onco Targets and Therapy* **7:** 1777. https://doi.org/10.2147/OTT.S69239.

Vinson, C.C., E. Mangaravite, A.M. Sebbenn, and T.A. Lander. 2018. "Using Molecular Markers to Investigate Genetic Diversity, Mating System and Gene Flow of Neotropical Trees." *Revista Brasileira de Botanica* **41 (2):** 481–496. https://doi.org/ 10. 1007/s40415-018-0472-x.

Waldman, M., and Y. Shevah. 2000. "Biological Diversity - An Overview." *Water, Air, and Soil Pollution* **123 (1–4):**

299–310. https://doi.org/10.1007/978-94-011-4369-1_24/COVER.

Wolfsberg, T.G., and D. Landsman. 1997. "A Comparison of Expressed Sequence Tags (ESTs) to Human Genomic Sequences." *Nucleic Acids Research* **25 (8):** 1626–1632. https://doi.org/10.1093/nar/25.8.1626

Xu, Y., S. Tian, R. Li, X. Huang, F. Li, F. Ge, W. Huang, and Y. Zhou. 2021. "Transcriptome Characterization and Identification of Molecular Markers (SNP, SSR, and Indels) in the Medicinal Plant *Sarcandra glabra* spp." *BioMed Research International* **2021:** https://doi.org/10.1155/2021/9990910.