

## FUTURISTIC TRENDS IN THE AUTHENTICATION AND EVALUATION OF PLANTS

### Abstract

Alec Jeffreys' papers on the application of minisatellite probes for human DNA fingerprinting were published almost three decades ago. Methodologies have rapidly evolved and diversified since those early days. Many different techniques have been created, improved upon, and then finally discarded as newer, more effective, and/or more reliable solutions emerged. While Southern blot hybridization and restriction fragment analysis dominated the initial stages of DNA fingerprinting, the invention of the PCR in the late 1980s paved the way for the creation of single- or multi-locus profiling techniques based on PCR. Plant DNA fingerprinting routines still frequently use PCR-based markers. Now, high-speed DNA sequencing was used for DNA fingerprinting in plants, either indirectly through the facilitation of marker development or directly through "genotyping-by-sequencing". Proper identification of plant species belonging to related taxa and finding their relatedness is necessary for the botanical world. DNA fingerprinting plays a crucial role in the authentication of plant specimens as it overcomes almost all of the limitations of traditional genotyping methods. In this chapter we aim to explain various molecular marker tools along with their advantages and disadvantages.

**Keywords:** DNA fingerprinting, Plants, Genetic mapping, Population genetics, Single nucleotide polymorphisms, Systematics

### Authors

**Padmanabhan Jayanthikumari Vivek**  
Post-Graduate & Research Department of Botany  
Sree Neelakanta Government Sanskrit College  
Pattambi, Palakkad, Kerala, India.

**Madathil Nishma**  
Post-Graduate & Research Department of Botany  
Sree Neelakanta Government Sanskrit College Pattambi  
Palakkad, Kerala, India.

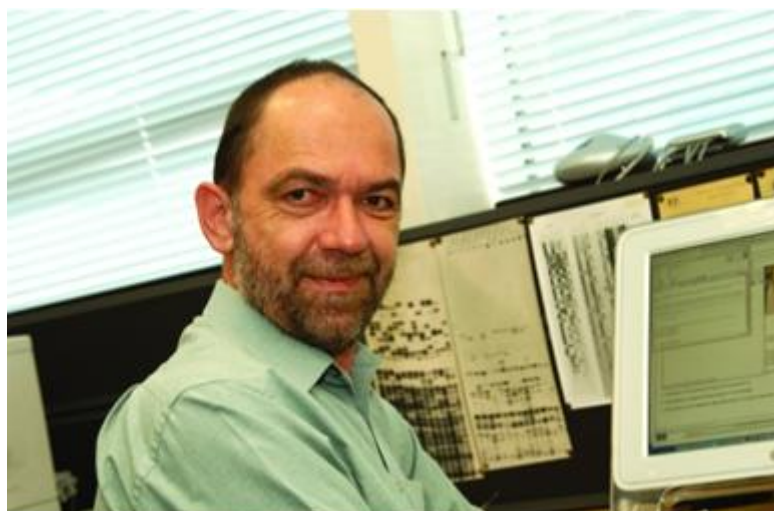
**Kuzhiparambil Amrutha**  
Post-Graduate & Research Department of Botany  
Sree Neelakanta Government Sanskrit College Pattambi  
Palakkad, Kerala, India.

**Mohankumar Saraladevi Resmi**  
Post-Graduate & Research Department of Botany  
Sree Neelakanta Government Sanskrit College Pattambi  
Palakkad, Kerala, India.

## I. INTRODUCTION OF DNA FINGERPRINTING IN PLANTS

DNA fingerprinting, also referred to as DNA typing or profiling, is the process of amplifying an individual's chromosomal DNA to create patterns that resemble bar codes and can be used to identify one person from another. The DNA molecules were composed of nucleotides like adenine, guanine, thymine, cytosine, and pentose sugar joined by phosphate bonds. The uniqueness of an individual at the molecular level, or its genetic characteristics, is the foundation of DNA fingerprinting.

It is unusual for one person to dominate an entire scientific field, but this is unquestionably the case with molecular marker-based investigative genetics; Head of the pack is Alec Jeffreys (figure 1). His primary concern was whether there was any variation in the restriction fragments of different individuals, and his initial method of addressing this was through restriction fragment length polymorphisms (RFLPs). Jeffreys was somewhat dissatisfied with RFLPs because they did not provide much genetic information, despite the fact that they would advance several areas of genetics research. Detecting and using it at the time was also difficult. From there, he opens the door for later researchers to discover DNA fragments that are a great deal more variable than typical RFLPs (26). After successfully isolating the myoglobin gene from the seal, Jeffreys discovered a minisatellite, a brief segment of DNA with tandem repeats, within the human myoglobin gene, from where his studies leads to human DNA finger printing (25).



**Figure 1:** Alec J. Jeffreys

For the plant taxon that is closely related, precise designation is required. Differentiating various plant genotypes and finding relatedness among plant species is inevitable in various branches of plant science and Ayurveda. Botanical materials are differentiated using morphological, and anatomical traits along with chemical constituents present in them. Various scientific methods were employed to analyse and identify various botanicals from each other. Morphological evaluation, anatomical inspection and chemical profiling were the initial methods for the identification of botanicals. Comparative microscopic examination of the unprocessed drugs is part of the microscopic evaluation

process. In order to look for impurities in crude drugs, different chromatographic tools were used to profile the chemical patterns of the plant materials. However, all of these methods have limitations because plants can change their morphology and anatomical characteristics as well as their chemical composition simply by changing their habitat (63).

Limited variations in plant species, prejudiced data collection and treatment, and environmentally induced flexibility were the major limitations in traditional analytical methods. A more advantageous and neutral method was offered by the molecular marker-based identification approach. DNA markers are used in the disciplines of molecular biology to distinguish a specific DNA sequence from a group of unknowns and to preserve biodiversity. Genetic typing is mostly employed in plants for a variety of purposes, including the detection of genetic diversity and variability, paternity test, genetic mapping, identification of hereditary disease related mutant genes, crop cultivars selection, marker-assisted breeding of wild cultivars, population studies and history, pharmacology, food safety, and the discovery of trait markers along with biodiversity protection (19).

The 1980s saw a significant rise in the use of molecular markers in agriculture, which increased both commercial and scientific exploitation of knowledge regarding the abundance of agronomic and disease-resistant genes in significant crop species (17). Traditional breeding programmes require whole genome crossing and superior recombinant selection with a lengthy procedure involving numerous crosses and several generations of the candidate plants, making them arduous, time-consuming, and laborious to complete. It might be even more challenging to accomplish the desired goal if there is a close linkage between the undesirable loci and the desired loci. Molecular Markers should be easily available, reproducible and highly polymorphic showing co-dominant inheritance and recurrent occurrence in genome. It should be selectively neutral to environmental conditions. Getting the molecular marker for the aforementioned criteria is very difficult (34). Hybridization-based markers and polymerase chain reaction (PCR)-based markers were used to categorise DNA typing. In a previous method, a DNA molecule was visualised after being digested by a restriction endonuclease and hybridised with a probe of known sequence. In a later technique, a specific DNA sequence was amplified *in vitro* using PCR-based markers and thermostable DNA polymerase enzyme. The amplified product was separated and examined using the gel electrophoretic method, staining it with ethidium bromide or using autoradiography. Comparison of some widely used DNA marker systems in plant is shown in Table 1.

Molecular markers based on isoenzyme protein molecules were initially used for genotype recognition (18). Protein extraction from plants with high amounts of polysaccharides and polyphenols was a major problem the scientists faced at that time. DNA molecules are easy to work with and will be more potent than proteins in producing an infinite amount of polymorphic data (51).

DNA-based restriction fragment length polymorphism (RFLP) technique eliminated challenges related to the marker molecule extraction procedure as well as to the analysis of plants growing in remote areas. Samples were usually transported to the laboratory in dried form on silica gel, and could be kept frozen until genomic DNA extraction. DNA isolation, restriction digestion of the obtained DNA, southern blotting filtering of the fragments, hybridization with locus-specific probes for detecting the fragments were the procedures of RFLP (11). Time consumption and developing species-specific hybridization probes were the

major constraints of RFLP. Even though it constitutes a well-known approach for genetic map construction as well as for finding relatedness among major crop plant species (66) (38), RFLP analysis was applied to chloroplast DNA, but not to mitochondrial DNA due to its highly conserved nature. Chloroplast DNA-based RFLP studies have been carried out mainly on an inter-specific level, as there is not much intra-specific variation in the cpDNA molecule (53).

RFLP analysis based on probes synthesised from tandemly repeated DNA sequences in humans, other mammals, and birds inspired botanists to investigate the possibilities of applying this tool to plants as well (68). In 1988, restriction-digested DNA of *Oryza sativa* was hybridised with the human 33.6 minisatellite probe and using this, Dallas and coworkers (10) distinguished different rice cultivars from each other. Identical fingerprinting was shown by the offspring of an individual crop plant which was also expected as the plant is self-pollinating. Reproductive mode determines the degree of variation between the parental and offspring lanes. Immense variation was present in the fingerprints of sexually derived cottonwood plants were as identical in inbred Tomato (58). Crop plants propagating vegetatively generally said to have monomorphic genotype. Appearance of certain variations in plant genotype due to somatic mutation can be investigated and manipulated using DNA fingerprinting (49).

**Table 1: Comparison of some widely used DNA Marker Systems in Plants**

Marker description	RFLP	RAPD	AFLP	ISSR	SSR	SNP
Genomic abundance	High	High	High	Moderate	Moderate to high	Very high
Genomic coverage	Low–moderate	Whole genome	Whole genome	Whole genome	Whole genome	Whole genome
Inheritance	Codominant	Dominant	Codominant/ Dominant	Codominant	Codominant	Codominant
Codominant Level of polymorphism	Moderate	High	High	Moderate	High	High
Reproducibility	High	Low	High	Moderate	High	High
Effective multiplex ratio	Low	Moderate	High	Moderate	Medium/ High	Moderate to high
Marker index	Low	Moderate	Moderate to high	Moderate	Moderate to high	Moderate
Quality of DNA required	High	Low	Moderate	Low	Moderate	High
Technical expertise required	High	Low	Moderate	Low	Low/ moderate	High
Automation	Low	Moderate	Moderate/ high	Moderate	Moderate/ high	High
Cost per assay	Moderate to high	Low	Moderate	Low	Moderate to high	High

(GACA)<sub>4</sub> and (GATA)<sub>4</sub> were synthetic oligonucleotides used as RFLP hybridization probes in the initial stages of DNA fingerprinting, which were hybridized to short, tandem-repeated sequences like microsatellites and simple sequence repeats (SSRs) in the genome. It will produce polymorphic fragment patterns (47). SSRs and SNPs were widely accepted as marker tools for genotype identification of crop plants and their wild relatives. As SSRs are multi-allelic, they have more polymorphic content than SNPs. SNPs were biallelic in nature and were distributed uniformly throughout the genome (79). Their nomenclature was much simpler than that of SSRs so it was more approachable in analyzing and handling the data (69). The International Union for the Protection of New Varieties of Plants (UPOV) approved these two markers as additional measures along with traditional morphological analysis for plant variety identification (64).

The first report of the invention of a panel of SNP markers was done on pepper, through which they distinguished 100% of 17 sweet peppers and 97.5% of 81 commercial pepper cultivars (27). In 2011, Cabezas and his colleagues (6) developed a stable standard set of markers for grapevine genotyping by choosing 48 SNP markers that were evenly distributed in the genome. In 2018, Li et al (42), aligned the resequencing data of an inbred cabbage line, and selected 50 core markers from 2.54 million SNPs, and introduced DNA fingerprint data for 59 cabbage varieties. Zhang et al, in the year 2020 (76), aligned 182 resequencing data sets of 261 cucumber cultivars by target SNP-seq, selected 163 core SNP markers from 4,612,350 SNPs, and introduced DNA fingerprint data of that many cucumber cultivars. In 2018, the 3000 Rice Genome Project (3 K RGP) opened up the sequence data of 3024 rice germplasm library from the 780,000 rice samples in the global rice germplasm library with an average sequencing depth of 14X (69).

In the field of population genetics, now it is possible to evaluate the genetic variation pattern throughout the whole genome due to our capability to sample the genome at much higher densities (40). Ecology and conservation will be significantly influenced by the outcomes and understandings gained from this "population genomic" approaches. Plant phylogenies are currently constructed using numerous sets of nuclear genes and will be advantageous to plant systematists (15). Sequence characterization and comparison within large taxonomic distances will help to introduce novel nuclear markers. Orthology issues in polyploid phylogenetics may be addressed by direct sequence analysis of barcoded individuals (13). Ultimately, more precise phylogenies will make it easier to understand how species within recent radiations relate to one other.

**Table 2: Summary of studies using DNA fingerprinting for Varietal Identification (54).**

<b>Author</b>	Kosmowski et al.	Maredia et al. Rabbi et al.	Maredia et al. Rabbi et al. Floro et al.	Floro et al.	lukor et al. n.d.	Kilic (n.d.)	Wossen et al.	Yirga et al.
<b>Year</b>	2016	2015–2016	2015–2016	2017	2017		2018	2016
<b>Country</b>	Ethiopia	Ghana	Zambia	Colombia	Malawi	Uganda	Nigeria	Ethiopia
<b>Organism</b>	Sweet potato ( <i>Ipomoea batatas</i> )	Cassava ( <i>Manihot esculenta</i> Cranz)	Bean ( <i>Phaseolus vulgaris</i> )	Cassava ( <i>Manihot esculenta</i> Cranz)	Cassava ( <i>Manihot esculenta</i> Cranz)	Maize ( <i>Zea mays</i> )	Cassava ( <i>Manihot esculenta</i> Cranz)	Maize ( <i>Zea mays</i> )
<b>Ploidy number</b>	Hexaploid ( $2N=6x=90$ )	Diploid ( $2N=2x=36$ )	Diploid ( $2N=2x=22$ )	Diploid ( $2N=2x=36$ )	Diploid ( $2N=2x=36$ )	Diploid ( $2N=10x=20$ )	Diploid ( $2N=2x=36$ )	Diploid ( $2N=2x=36$ )
<b>Genomesize</b>	2.37 Gbp	772 Mbp	587 Mbp	772 Mbp	772 Mbp	2.5 Gbp	772 Mbp	2.5 Gbp
<b>Tissue collected</b>	Leaf (individual)	Apical leaf (individual)	Seed (individual)	Stems (individual)	Leaf (individual)	Seed (bulk)	Leaf (individual)	Seed (bulk)
<b>Tissue for DNA extraction</b>	Leaf	Leaf	Young leaf	Stems	Leaf	Bulked seed/ crop cut	Leaf	Bulked seed/ crop cut
<b>Genotyping technology</b>	DArTseq	GBS	KASP (SNP)	SNPY-chip (yuca chip)	DArTseq	DArTseq	GBS	DArTseq
<b>No. markers per sample</b>	Not reported	56,849	66	93	Not reported	Not reported	52,899	Not reported

**Source:** Poets A, Silverstein K, Pardey P, Hearne S, Stevenson J. DNA Fingerprinting for crop varietal identification: Fit-for-protocols, their costs and analytical implications. 2020 Apr

## II. METHODS OF DNA FINGERPRINTING

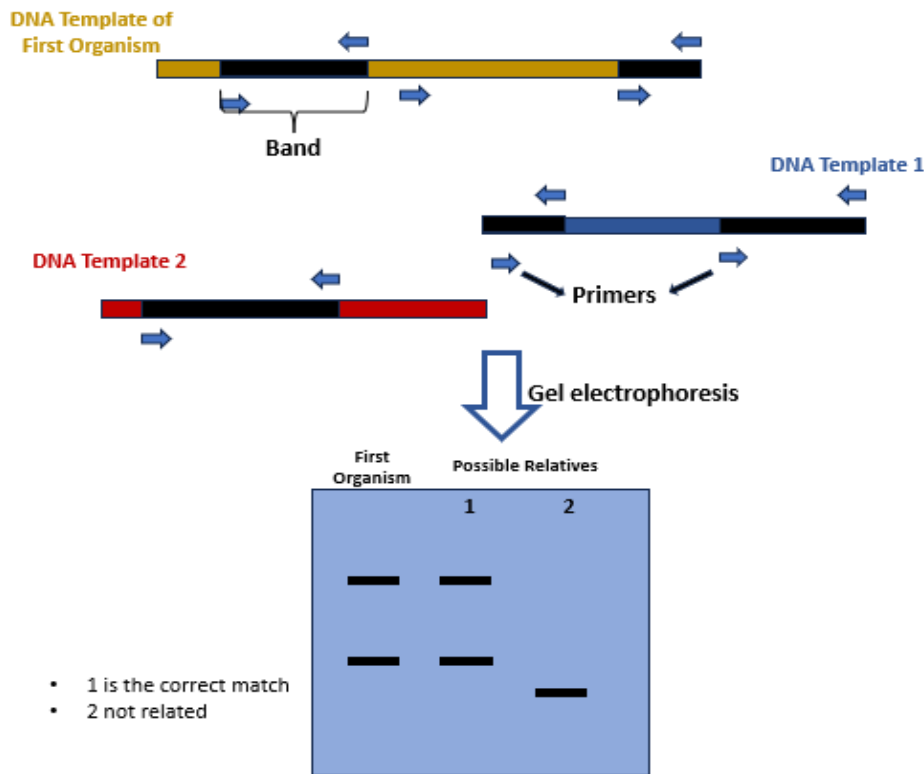
Numerous marker-based genotyping systems were developed after the establishment of RFLP in genotyping. The selection of an efficient marker method from these systems should be based on our experiment setup, cost, lab facility, expertise, and marker availability. Locus-specific microsatellite analysis (SSR) was the most widely used method for differentiating between plant cultivars in the mid-2010s, subsequent to RAPD, ISSR, AFLP, other nuclear DNA-based methods such as CAPS, DAMD, IRAP, REMAP, SNPs, SCAR, and SRAP, and organellar DNA-based methods (50).

The most widely used marker-assisted technique was RFLP, which was very dependable and reproducible between labs. As the markers are inherited codominantly, genetic characterisation was readily apparent. Using this method, which combines various probes and enzymes, it was possible to identify both individual species and entire populations. The variable number of tandem repeats (VNTRs) is the source of variation here. The frequency of RFLP usage today is decreased by the requirement for high-quality DNA samples, high levels of automation, and laborious procedures.

PCR-Based marker techniques use single oligonucleotide primers with arbitrary sequences for the PCR amplification of genomic DNA fragments, which produce multi-locus bands during electrophoretic separation. It gives information about multiple loci simultaneously. But these methods are not exactly reliable, but they provide robust identification.

Random Amplified Polymorphic DNA (RAPD) (Figure 2), Arbitrarily Primed PCR (AP-PCR), and DNA Amplification Fingerprinting (DAF) are all referred to collectively as Multiple Arbitrary Amplicon Profiling (MAAP). RAPD is easier to understand than DAF and AP-PCR. The variation between them is brought on by changes to the amplification mechanism. According to Idrees and Irshad (24), changes can be made to the thermostable DNA polymerase type, the number of PCR cycles, the primer characteristics, and the annealing temperature. RAPD approach, also utilizes arbitrary primers for the amplification of gene fragments in a specific range and has become the most popular molecular marker technique (72). The 10 bp dominant primer will amplify a DNA fragment that is complementary to its sequence and does not require knowledge of the genomic sequence of the target prior to the experiment. It is a reliable and cost-effective technique for DNA bacterial strain typing (2). Mutation can cause variation in the banding pattern of the target gene fragment in RAPD diagram (46). Simple and cost-effective methodology and the requirement of a small amount of DNA samples were the resolutions of this technique. A single primer used in the RAPD technique primes multiple loci randomly distributed throughout the genome. So, it will produce a wide spectrum of amplicons, and hence the entire genome is available for amplification and analysis (7).

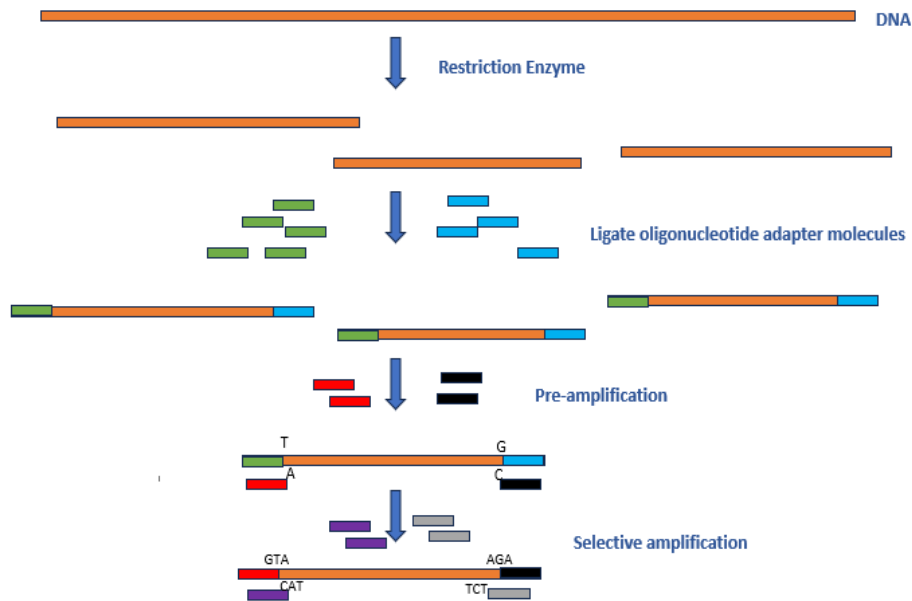




**Figure 2:** Random Amplified Polymorphic DNA

Amplified Fragment Length Polymorphism (AFLP) technique (Figure 3), introduced by Zabeau and Vos (75), has become very popular due to its multi-locus accessibility in a single experiment. It combines RFLP with PCR analysis by ligating the adaptors to the restricted DNA to obtain a large number of polymorphic bands. The representative DNA regions of the larger species distributed throughout the genome can be simultaneously screened and characterised by AFLP. It is not necessary for this technique to have prior sequence knowledge or to have made an initial investment in probe development. With the exception of the absence of restriction enzymes or any other inhibitors, the quality of the DNA sample can still be compromised in this situation. Here, a rare cutter (EcoRI or PstI) and a frequent cutter (MseI or TaqI) of two different types of restriction enzymes were used.

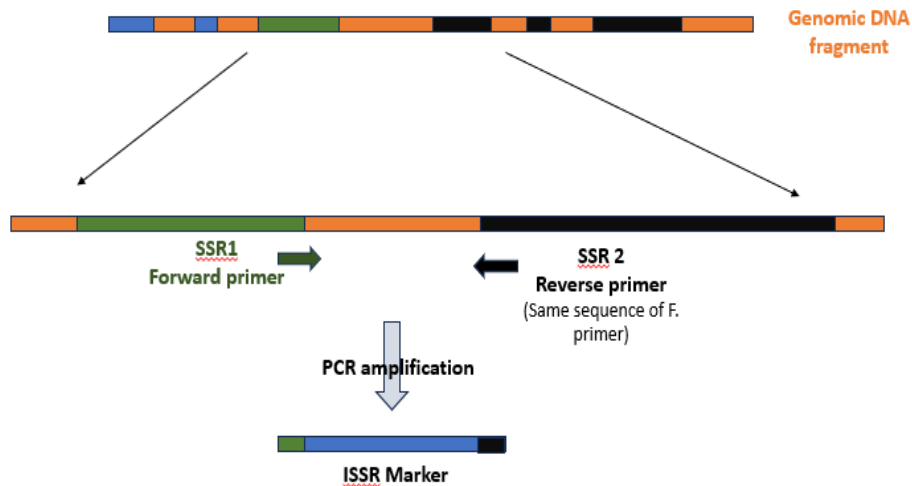
Since the adaptors were made to ligate to both ends of the fragments to provide known sequences for PCR amplification, the original restriction site will not be restored after ligation. Only fragments with adaptor sequences and complementary base pairs to the additional selective nucleotides will result in amplification. Pre-amplification will be done with a single base pair of primers, and selective amplification will be done with a primer pair with a three base pair extension, both of which have complementary regions to the adaptor sequences. It can amplify a variety of DNA fragment subsets because the primers only differ by one base pair and the PCR conditions are strict. Visualisation methods included autoradiography, silver nitrate staining, and automated genomic sequencer (24).



**Figure 3:** Amplified Fragment Length Polymorphism (AFLP)

Compared to other multi-locus molecular marker techniques, AFLP markers are highly versatile and productive and hence can be used for analysing a huge number of breeding accessions of crop plants together. The benefit of AFLP-PCR is the ability to characterise the entire plant genotype immediately by producing hundreds of precise and reliable markers developed from both polymorphisms of restriction and primer annealing sites. Biodiversity estimation can also be precisely carried out by this marker technique (12).

In 1994, Gupta et al. (16) developed INTER SIMPLE SEQUENCE REPEAT (ISSR), a molecular marker technique based on a microsatellite-complimentary primer sequence. It amplifies the DNA fragment located between two adjacent but oppositely oriented microsatellite repeat regions in the target DNA that are present within an amplifiable distance (Figure 4). Simple Sequence Repeats (SSRs) typically range in size from 15 to 25 bp and target multiple DNA loci to produce amplification products with inter-SSR sequences of various sizes. They are polymorphic in nature and can be di or tri-nucleotides. Due to its high polymorphism, ISSR is chosen as a marker in genetic diversity and mapping, evolutionary biology, and phylogenetic studies (56). Despite mostly following simple Mendelian inheritance, these markers' ability to distinguish between homozygote and heterozygote was demonstrated by the segregation of some of them as co-dominant markers (59). Similar to RAPD and AFLP, no prior knowledge of the candidate's genomic data is necessary here. Novel gene loci in maize cultivars were found using one ISSR and one RAPD primer (R-ISSR) (73).



**Figure 4:** INTER SIMPLE SEQUENCE REPEAT (ISSR) marker made up of two closely spaced SSR sequences with the same repeat unit but different orientation.

ISSR has advantages over RAPD and RFLP with strong polymorphism, stability, reliability, reproducibility, cost-effectiveness, and smooth operation (78). In spite of this, when all three methods were employed on the same plant material, they typically yielded almost the same estimates of genetic diversity and genetic distances (71).

Microsatellites are highly variable tandem repeat motifs, also referred to as short tandem repeats (STR), simple sequence repeats (SSR), and variable number tandem repeats (VNTR). In the nuclear DNA of various organisms, mono-, di-, tri-, tetra-, penta-, and hexanucleotide repeats can be found. Trinucleotide repeats are more typical in plants than in vertebrates, which are distinguished by their high frequency of occurrence (9). The primers used in the PCR amplification of microsatellite DNA are two distinct sequences that are complementary to the flanking regions. The formulation of correctly working primers is frequently a time-consuming and expensive process. Microsatellites are strong for a variety of applications though once they are developed and characterised in an organism.

Sequence-related amplified polymorphism (SRAP), one of the most frequently used methods in genetic analysis, uses PCR to identify variations in intron, promoter, and spacer lengths between individuals and species. In SCAR marker technology, lengthy primers with an average size of 23 bp were created based on the sequencing results of Multiple Arbitrary Amplicon Profiling (MAAP) fragments or RAPD marker terminals after their cloning. Short random repeats from different target organisms show a high degree of similarity. The presence or absence of a band on the autoradiogram or the stained agarose gel was used to analyse sequence variation. These molecular markers exhibit both dominant and co-dominant inheritance and are extremely reliable and reproducible. SRAP has been widely used in the analysis of genetic diversity, creation of genetic maps, mapping of essential traits, and cloning of related genes in several plant species, including grape, plum, and mangoes, due to its simplicity and effectiveness (77).

Targeting ESTs, the TRAP fingerprinting technique identified and analysed polymorphism of candidate genes. A random primer with an AT- or GC-rich core that can anneal with an intron or exon will be used as one of the two primers in this case, while the other will be complementary to the EST sequence, Sequence-Related Amplified Polymorphism (SRAP) and Target Region Amplification Polymorphism (TRAP) methods have two major common characteristics, including the use of two primers of about 18 nucleotides in length and flexible PCR conditions during the initial five amplification cycles.

Cleaved Amplified Polymorphic Sequence (CAPS) markers combine RFLP and PCR based marker assisted techniques as they work based on the restriction enzyme digestion of PCR amplified products. PCR-RF-LP was the initial acronym used to describe CAPS. The number or distribution of polymorphic sites, DNA extraction, and PCR conditions are all crucial factors in the CAPS marker approach. It completely avoids the challenges associated with probe development and blotting. The polymorphism in CAPS markers only depends on the variability in the restriction fragmentation site and exhibits co-dominant inheritance. CAPS primers were created using the EST sequences of potential organisms rather than non-functional regions, such as microsatellite markers (77).

The PCR-based marker known as random amplified microsatellite polymorphism (RAMP) combines markers from the simple sequence repeat (SSR) and random amplified DNA polymorphism (RAPD) classes. The repeat sequences in the primers are labelled, have a 5' anchor, and may or may not have a RAPD primer. Two temperatures are used for PCR in order to allow for proper melting and annealing. To ensure that only bands obtained by anchored labels are detected, PAGE with label detection is used (44).

SSCP stands for Single-Strand Conformation Polymorphism. Due to variations in the quantity and type of bases, SSCP can identify differences in the electrophoretic mobility of single-stranded DNA. A quick and accurate method for genotyping mutations is single-strand conformational polymorphism (SSCP) analysis. The SSCP analysis method is based on the idea that single-stranded DNA has a specific conformation (55). Target sequences are amplified by PCR, and then the double-stranded DNA is denatured to produce single strands. These latter sequences develop three-dimensional shapes that vary in electrophoretic mobility between sequences with even minor base differences.

Molecular markers can utilize direct amplification of minisatellite DNA as well as microsatellite DNA. Direct Amplification of Minisatellite DNA (DAMD) and Selective Amplification of Polymorphic Microsatellite Loci (SAMPL) use primers that are specific for minisatellites and microsatellites respectively (8). Conserved domains in the plant-resistant genes of DNA samples can be used as PCR primers for marker techniques, and that approach was termed Resistance Gene-Analog Polymorphism (RGAP). The conserved regions of resistance genes, such as the protein kinase domain, leucine-rich repeats, and nucleotide-binding site, are used to create RGAP markers (3). The Diversity Arrays Technology (DArT) involves the priming of fluorescent DNA probes and a set of target DNAs spotted onto a microarray, and sequence information is not needed for this (41).

There are many retrotransposon elements in the eukaryotic genome that show specificity and uniqueness and help in inventing relatedness among species. The greater the divergence between species, the more variation there will be in a particular sequence of

retrotransposon elements. So, mobile elements can also be used as molecular primers for finding the relatedness of related species (30). Inter-Retrotransposon Amplified Polymorphism (IRAP) includes the identification and analysis of mobile element mutation site polymorphisms combined with the PCR technique. IRAP is based on PCR amplification of the DNA fragments between two adjacent LTR retrotransposons using a single LTR primer (28).

The Inter-Primer Binding Site (iPBS) amplification technique does not require knowledge of the sequence databases of retrotransposons. Reverse transcriptase site (PBS) in LTR retrotransposons can normally bind to the universal complement tRNA. This binding mechanism combined with PCR methods was used in the IPBS amplification technique (29). Another PCR-based marker technique, the Retrotransposon Microsatellite Amplified Polymorphism (REMAP) like IRAP, requires retrotransposon-specific primer pairs. One of the primer pairs binds to the LTR retrotransposon sequence, and the other binds to the microsatellite sequence and acts as an anchored primer (28).

RBIP, or RETROTRANSPOSON insert polymorphism produces a PCR fragment containing the insertion by using a transposon-specific c primer in conjunction with a primer complementary to the flanking region. Primer sets specific to both of the flanking regions will produce a fragment if the insertion is missing (20).

In microarrays, also called chips, a significant amount of probe-hybridized DNA samples are immobilised on a nylon or glass slide. The main applications of the microarray technique are the detection of adulteration in food and medicine, identification of genomic constituents, and SNPs. Some of the microarray-based marker technologies include DArT, SSHA, SDA, oligonucleotide and gene-based microarrays. Single base differences in the gene pool were found using PCR amplification, cloning, and microarray spotting of chosen gene fragments after restriction digestion of the gene pool. The Diversity Array Technology (DArT) used for these molecular markers, although the method is reliable and instructive, it is very expensive and necessitates high levels of automation (52).

Molecular genotyping using Subtracted Diversity Array (SDA) have two gene pools in which the latter one is the target sequence and used as microarray probe. Here pooled genomic DNA sample was extracted from the second Genomic pool from which DNA fragments unique to the second set is obtained. In Oligonucleotide and Gene-based Microarrays, Species-specific oligonucleotide sequences are used as probes. Target genomic DNA is sheared labelled and studied. Suppressive Subtractive Hybridisation Array (SSHA) greatly increases the probability of capturing polymorphic DNA fragments, even in small occurrence, by eliminating common sequences from the target samples (35).

PCR-amplified single-locus microsatellite markers pave the way to becoming the major choice of scientists because of their high reproducibility and excellent polymorphism in the number of tandem repeats (62). Here, a pair of microsatellite-flanking primers were used for the PCR amplification of the targeted gene fragment. Sequence-Characterised Amplified Regions (SCARs) are locus-specific markers in which specific primers were designed for amplifying a particular band of RAPD results (4). Polymorphic Sequences (CAPS) approach (37).

SNPs became the prominent genotyping technique as it simultaneously screened a large number of SNP loci. Numerous technologies have been established for SNP discovery and genotyping (1). The occurrence of SNPs in plant genomic DNA depends mainly on the specimen selected and the genomic region investigated. Commonly, there will be one SNP within 100 to 500 bp of genomic DNA, which also varies. Highly established High-throughput sequencing systems increase the efficiency of SNP identification, which is done by analysing multiple copies of similar DNA fragments. All significant crop species already have SNP markers in effect, especially those for which complete genome sequences are available (14). The development of precise, high-throughput, and low-cost fingerprinting techniques provides SNPs with a major role in genetic relatedness checks, plant breeding, and molecular-assisted selection (5).

The DNA sequencing method developed by Sanger and colleagues involved *in vitro* replication of the target DNA through the selective addition of chain termination nucleotides. Different DNA typing techniques use sequencing as a key step. One of the main technologies that uses sequencing is DNA barcoding to identify and characterise plant genotypes from a mixed gene pool. Some of the universal primers used singly or in combination for the barcoding process include *trnH - psbA*, *trnL (P6)*, *nrITS*, *nrITS2*, *accD*, *matK*, *ndhJ*, *rpoB*, *rpoC1*, *ycf5*, *atpF - H*, *psbK - 1*, *rbcL*, and *rbcLa* (67). Next-generation sequencing (NGS), a new generation sequencing technique, first appeared in the middle of 2012. The three primary sequencing techniques used in NGS are single-molecule sequencing, sequencing by ligation, and sequencing by synthesis. Exploration of whole genome and transcriptome data of model and non-model plants occurs abruptly as a result of the exponential growth of the bioinformatics domain and next-generation sequencing. By enriching specific regions prior to sequencing, transcriptome sequencing and restriction site associated DNA sequencing (RADseq) reduce the complexity of enormous datasets (23).

Chloroplast DNA is most commonly used in organellar DNA-based technique. Lack of recombination in the chloroplast genome leads to the haplotyping of that particular individual. Investigation of phylogenetic and phylogeographic relatedness requires a low mutation rate in the specimen genome, which is found in plastid DNA. These markers may also be used to trace uniparental lineages over considerable gaps in time and space.

## **1. Advantages of DNA fingerprinting**

Proper identification of plant species belonging to related taxa and finding their relatedness is necessary for the botanical world. Such identification will allow the examiners to determine the adulterant materials in the original specimens. In the early decades, morphological characters such as shape, size, colour, texture, fracture characteristics, odour, and taste were used for discriminating plant samples. Anatomical, chemical, and comparative microscopic inspections were also done during such investigations. All these protocols worked well together and were usually employed. The main issues that researchers faced when using these conventional methodologies were less variation between the plant materials, sampling errors, treatment errors, and environmental plasticity.

Chemical fingerprinting establishes a characteristic chemical pattern for each plant sample. Qualitative and quantitative determinations of impurities in plant precursors of drugs were done using various chromatographic tools. These techniques have limitations, though,

because the composition and relative amounts of chemicals in a given plant species vary depending on the growing environment, the time of harvest, and the storage conditions. Each herb contains a large number of chemicals, and therefore, it is a tedious process to check the presence of all the compounds of interest in plants. DNA fingerprinting plays a crucial role in the authentication of plant specimens as it overcomes almost all of the limitations of traditional genotyping methods (63).

Molecular markers have a significant role in the detection of adulteration because they can distinguish one genotype of plant from other mixed genotypes. By utilising its primer amplicon length, the SCAR genetic marker was used to distinguish between the main adulterants of the *Angelica* species (commonly referred to as Jeonho in Korea). *A. decursiva* has an amplicon length of 363 bp, whereas *A. sylvestris* has a specific amplicon length of 273 bp. The amplicon primers for *A. decursiva* and *P. praeruptorum* are 145 bp and 305 bp, respectively. Using an AFLP genetic marker, adulterant samples of *Zanthoxylum acanthopodium* and *Z. oxyphyllum* were distinguished from genuine samples.

In the late nineties, genotypes of both wild and crop plants were identified using DNA fingerprinting. In some instances, estimates of genotype similarity based on DNA show a relatively close correlation with earlier estimates based on morphology, but in other instances, there are also significant discrepancies. In some case studies, variations within a taxonomic level may be a positive factor, but in others, they will be negative. In both cases, various questions can be answered by inventing plant species by molecular marker techniques.

DNA fingerprints derived from plant fragments should be able to provide significant evidence in criminal investigations, but progress has been slow thus far, most likely as a result of difficulties in isolating DNA of sufficient quality from poorly preserved plant material. SRR markers can overcome this problem as they can use even highly degraded DNA samples. However, one well-known early case involved RAPD examination of Palo Verde tree seed pods found both at the crime scene and in a suspect's truck (74). A variety of DNA marker techniques have been employed to prove violations of plant breeders' rights. A related area of study involves the identification of plants whose possession is prohibited. As a result, numerous studies on the identification of *Cannabis sativa* specimens as part of drug enforcement have been published (36). 15 SSR loci were combined into a single multiplex to discriminate between *Cannabis* genotypes perfectly (22).

Evolutionary influences of natural selection, mutation, gene flow, and genetic drift were common in wild plant populations and were studied using molecular markers. The partitioning of genetic variation between and within populations, as well as gene flow, are both significantly influenced by the breeding system. With DNA markers, interbreeding between populations has been shown to occur in a variety of outcrossing plant species. Artificial seed disposal methods, like anthropogenic disposal, also cause variation in the gene flow of a species. Spatial autocorrelation analysis for studying spatial scale-dependent changes in DNA marker polymorphism within a group of closely occurring populations is an important advantage of DNA fingerprinting.

Polyploidy and the fusion of gametes from two different entities are common in plants. Several types of markers can be used to identify the putative progenitor species of

polyploids, such as the internal transcribed spacer (ITS) region in the nuclear and ribosomal RNA gene clusters, which was sequenced and examined in, for instance, polyploid rose cultivars and species (57). In plant systematics, multi-locus DNA profiling techniques like AFLPs are now the most popular DNA fingerprinting tool, especially when DNA sequencing results in insufficient phylogenetic resolutions.

The degree of relatedness between the candidate species is demonstrated by the uniqueness and variation of retrotransposon elements found in the genomes of various plant species. The sequences of a particular retrotransposon show a high degree of similarity when two botanical species are closely related, whereas if they are not, the sequences diverge. It is a fact that if these plant species are distantly related, the conserved regions of the sequence may also diverge. For instance, the sequences of particular mobile elements were the same in the grass family. Nearly all of the sequences of particular retrotransposon elements were the same in wheat and *agelops*. It will decrease by almost 10% when the relatedness decreases, as in the case of *Hordeum*. Retrotransposon components are therefore useful for genetic typing. The most conserved regions in these mobile elements can be used to design PCR primers (30).

Advancement in high-throughput sequencing approaches leads to the large-scale identification of SNPs and SSRs. Multiplexed SNP-based genotyping by sequencing methods has been used in many crop plants and model plants. Huge analytical procedures were taking place for the identification and characterization of wild varieties of crop plants that lack a reference genome (61). The limitations of conventional marker methodology are now overcome by the capacity to sample the genome at much higher densities than ever before, enabling for the first time the analysis of genetic variation patterns throughout the entire genome (65).

After splicing the intron sequences, unstable mRNA samples were reverse transcribed to cDNA that only contains protein-coding exons. These cDNA sequences are known as expressed sequence tags (ESTs), which are created by sequencing the 5' or 3' ends of the sequence to create 5' ESTs or 3' ESTs, respectively. Exons, which are conserved throughout the gene family, are used to create 5' ESTs. Untranscribed regions of genomic DNA are the source of 3' ESTs. ESTs are created from cDNA libraries and will offer a vast amount of information about each organism's genes, including information on their structure and function (79).

EST-based SSR primers are designed using pattern-finding computer programmes from the target organism's available sequence data. The best sources for SSRs marker preparation are EST sequence databases because the twenty-mer markers can be best screened from them. These molecular markers typically have more conserved transcribed regions than untranscribed regions. Compared to SSRs derived from untranscribed regions, SSR markers of transcribed genomic origin are more closely associated with differential gene expression.

Geneticists and plant breeders use molecular marker assisted technologies for molecular breeding, MAS, genetic mapping, genome profiling, and characterization. A number of new molecular marker technologies can enter the sector through high-throughput genotyping platforms and develop and advance at a rapid rate. GoldenGate assay,



Genotyping-in-Thousands by Sequencing (GTseq), Diversity Arrays Technology, and NGS-based high-throughput hybridization platform systems (digitalMLPA [Multiplex Ligation-dependent Probe Amplification] and Molecular Inversion Probes [MIP]) are a few of the techniques that were developed through this door and improved the cost-effectiveness and sturdiness of these marker technologies. Along with high throughput sequencing, low-throughput molecular markers, such as Kompetitive allele-specific PCR (KASP) (45), were created to track specific genomic regions and assisted plant breeders in analysing a large sample of mixed species (31) (33).

Single nucleotide polymorphism (SNP) markers will be the best option for the genome-wide marker scanning in high-throughput genotypic platforms. However, the discovery of SNP-based markers quickens the pace of inquiries into genetic diversity. Insertion/deletion polymorphism (InDel), which involves the addition or deletion of a sequence from the genome, has significant consequences. In order to obtain high throughput knowledge about genomic diversity, variation, relatedness, and evolution, mobile element sequence-based molecular markers are crucial. Genome profiling based on interspersed repeats is an effective tool for genetic plant breeding programmes. They exhibit a high level of polymorphism and are used in population genetics studies. Along with whole genome sequencing, next-generation sequencing platforms now concentrate on specific regions of interest like TE with high levels of polymorphism. Almost all current PCR methods that involve locating and analysing repetitive elements can be adapted for use with contemporary NGS platforms.

Sequence tagged site (STS) is a unique small sequence that can only be found once throughout the entire genome. This specific sequence at that locus is regarded as the primer. After its discovery and examination in the human genome, STS was adopted to establish plant genomic studies. As long as the locus can be cloned and sequenced, it is possible to develop the primers for such markers from any sequence. High reproducibility and robustness of STS DNA markers demonstrate co-dominant inheritance. After cloning, the terminal ends of RAPD bands were sequenced to create STS in the form of SCARs. The repeatable sequences of the amplified loci were used to create RAMS primers, which are an adaptation of RAPD. It offers a better molecular marker for the scientific community because it is easier to transfer and duplicate. Numerous crop species were successfully genotyped using RAMS markers and AFLP and STS combinations (77).

By omitting the time-consuming and expensive protocols, several NGS platforms, along with genomics, proteomics, transcriptomics, epigenomics, and metabolomics, revolutionised the genotyping and genetic approaches to plants. The fields of genomics, proteomics, transcriptomics, epigenomics, and metabolomics collectively known as "omics" advanced high-throughput genotyping platforms. As model and non-model plants can be genetically profiled using these new technologies, NGS and omics have become the method of choice for scientists (32). High-quality reference genome sequences are typically present in model plants, but not in non-model crop plants. Even the actual whole genome data is lacking. But now we can genotype non-model plants using this high throughput sequencing platform. These marker-assisted NGS platforms can be used for quantitative genetic studies, genotyping progeny for MAS, or resolving phylogeographies of wild populations. Identification of genomic regions with the recombinant breakpoint that are differentially expressed in botanicals is aided by linkage mapping or (Quantitative Trait Locus) QTL.

Omics investigates a vast amount of metabolites, genes, and proteins for the continuous and thorough analysis of tissue-specific variation found in plants, which are typically extremely sensitive to even the smallest habitat changes. It will provide the characteristics of botanicals available on digital platforms as database repositories. By utilising the data repositories in the digital platform, the integration of omics with agricultural breeding programmes enhances the production of crop varieties with a value-added trait. It will assist the commercial wing of seed production in plant breeding programmes (43).

## **2. Limitations of DNA fingerprinting**

Excessive mutation rates in plant species cause methodological artefact such as bad reproducibility of molecular markers as well as biological artefacts such as insufficient germline stability of DNA fragment sequences. This problem is very common in SSRs, which are the highly sensitive markers (60). It was proved that S-SAP markers can distinguish clones developed after somatic mutation or after recombination within two similar species (21). Typically, species-specific sequence data is used to design S-SAP primers, but positive results have also been obtained using universal retrotransposon-based databases (70).

DNA isolation from processed plant samples is a challenge for scientists. Molecular marker system is costly also. This marker technique cannot identify the active chemical constituents in the sample along with genotyping. Hence molecular and chemical fingerprinting techniques has to be done together. When the two samples used for fingerprinting get mixed, markers then do not have any role in it due to meiotic rearrangements (39). High throughput sequencing in plant species provided a huge set of sequence databases to the scientific community. So today it is necessity to employ more than one locus to attain species-level discrimination across the angiosperm species. It is now compulsory to analyze algorithms for combining barcoding sequences from more than two regions of DNA to yield species-level unique markers.

## **III. CONCLUSION**

Proper identification of plant species belonging to related taxa and finding their relatedness is necessary for the botanical world. Traditionally such identification procedures were conducted through morphological, anatomical, and chemical investigations. Chemical fingerprinting markers are complex in nature and has limitation in standardisation procedure of plant specimen. Efficient sequencing and genetics studies innovated the researchers to explore various molecular markers for genotyping various crop plants, medicinal plants and economically important plants. Various DNA fingerprinting methods and their advantages and limitations were briefly reviewed in this chapter. Molecular markers are the most effective fingerprinting technique than all other ones due to their precise analytical nature. Sudden flourishing of various markers after the discovery of RFLP, increased the choices for selection of markers in the investigations. And according to our data set and sample material we can choose any of this marker. Molecular and chemical fingerprinting techniques has to be done together for genotyping along with chemical profiling. Even though molecular markers open up a lot of possibilities for genotyping and related fields, they cannot completely replace other markers. It can serve as an illustrative complement to other genotyping techniques. Sustainable development is improved through increased food

production and nutrient availability through a better understanding of diversity and genetic makeup.

## REFERENCES

- [1] Appleby N, Edwards D, Batley J. New technologies for ultra-high throughput genotyping in plants. *Plant genomics: methods and protocols*. 2009:19-39.
- [2] Aravena P, Pulgar R, Ortiz-Severín J, Maza F, Gaete A, Martínez S, Serón E, González M, Cambiazo V. PCR-RFLP detection and genogroup identification of *Piscirickettsia salmonis* in field samples. *Pathogens*. 2020 May 8;9(5):358.
- [3] Babar M, Ijaz S, Haq IU, Khan MS. Development of Molecular Marker Linked with *Cercospora* Leaf Spot (CLS) Disease Resistance in *Vigna radiata*, Cloning, and Expression for Evaluating Antifungal Activity against *Cercospora canescens*. *Phyton* (0031-9457). 2023 Apr 1;92(4).
- [4] Bautista R, Crespillo R, Cánovas FM, Claros MG: Identification of olive-tree cultivars with SCAR markers. *Euphytica* 2002, 129:33–41
- [5] Broccanello C, Chiodi C, Funk A, McGrath JM, Panella L, Stevanato P. Comparison of three PCR-based assays for SNP genotyping in plants. *Plant methods*. 2018 Dec;14:1-8.
- [6] Cabezas JA, Ibáñez J, Lijavetzky D, Vélez D, Bravo G, Rodríguez V, Carreño I, Jermakow AM, Carreño J, Ruiz-García L, Thomas MR. A 48 SNP set for grapevine cultivar identification. *BMC plant biology*. 2011 Dec;11(1):1-2.
- [7] Caetano-Anolles G, Brant J B, Peter M G. DNA amplification fingerprinting using very short arbitrary oligonucleotide primers. *Bio/technology*. 1991 Jun 1;9(6):553-7.
- [8] Chen XM, Line RF, Leung H. Genome scanning for resistance-gene analogs in rice, barley, and wheat by high-resolution electrophoresis. *Theoretical and Applied Genetics*. 1998 Aug;97:345-55.
- [9] Chen C, Zhou P, Choi YA, Huang S, Gmitter FG. Mining and characterizing microsatellites from citrus ESTs. *Theoretical and Applied Genetics*. 2006 May;112:1248-57.
- [10] Dallas JF. Detection of DNA "fingerprints" of cultivated rice by hybridization with a human minisatellite DNA probe. *Proceedings of the National Academy of Sciences*. 1988 Sep;85(18):6831-5.
- [11] Daniel H Paris, Allen L Richards, Nicholas PJ:2015 Day,Chapter 112 - Orientia,
- [12] Domblides A, Domblides E. Rapid Genetic Assessment of Carrot Varieties Based on AFLP Analysis. *Horticulturae*. 2023 Feb 22;9(3):298.
- [13] Duarte JM, Wall PK, Edger PP, Landherr LL, Ma H, Pires PK, Leebens-Mack J, Depamphilis CW. Identification of shared single copy nuclear genes in *Arabidopsis*, *Populus*, *Vitis* and *Oryza* and their phylogenetic utility across various taxonomic levels. *BMC evolutionary biology*. 2010 Dec;10:1-8.
- [14] Ganal MW, Altmann T, Röder MS. SNP identification in crop plants. *Current opinion in plant biology*. 2009 Apr 1;12(2):211-7.
- [15] Griffin PC, Robin C, Hoffmann AA. A next-generation sequencing method for overcoming the multiple gene copy problem in polyploid phylogenetics, applied to *Poa* grasses. *BMC biology*. 2011 Dec;9(1):1-8.
- [16] Gupta M, Chyi YS, Romero-Severson J, Owen JL. Amplification of DNA markers from evolutionarily diverse genomes using single primers of simple-sequence repeats. *Theoretical and applied genetics*. 1994 Dec;89:998-1006.
- [17] Gupta, P.K. and Varshney, R.K. 2004. "Cereal genomics: an overview." In *Cereal genomics*, edited by Gupta P.K. and Varshney, R.K., 639. Dordrecht, The Netherlands: Kluwer Academic Press
- [18] Hamrick JL, Godt MW. Allozyme diversity in plant species. *Plant population genetics, breeding, and genetic resources*. 1990:43-63.
- [19] Hartl DL, Jones EW. DNA structure and DNA manipulation. *Genetics: analysis of genes and genomes*. 2005;5:36-85.
- [20] Hirata C, Waki T, Shimomura K, Wada T, Tanaka S, Ikegami H, Uchimura Y, Hirashima K, Nakazawa Y, Okada K, Namai K. DNA markers based on retrotransposon insertion polymorphisms can detect short DNA fragments for strawberry cultivar identification. *Breeding Science*. 2020;70(2):231-40.
- [21] Hoffmann P, Blaich R, Forneck A. Retrotransposon-based mutations in clones of 'pinot' varieties. *Acta Horticulturae*. 2009(827):83-6.
- [22] Michael Höltsken A, Schröder H, Wischniewski N, Degen B, Magel E, Fladung M. Development of DNA-based methods to identify CITES-protected timber species: a case study in the Meliaceae family.
- [23] Hu T, Chitnis N, Monos D, Dinh A. Next-generation sequencing technologies: An overview. *Human Immunology*. 2021 Nov 1;82(11):801-11.

- [24] Idrees MU, Irshad MU. Molecular markers in plants for analysis of genetic diversity: a review. *European academic research*. 2014;2(1):1513-40.
- [25] Jeffreys AJ, Wilson V, Thein SL. Individual-specific 'fingerprints' of human DNA. *Nature*. 1985 Jul 4;316(6023):76-9.
- [26] Jeffreys AJ. DNA sequence variants in the  $\gamma$ -,  $\alpha$ -,  $\delta$ - and  $\beta$ -globin genes of man. *Cell*. 1979 Sep 1;18(1):1-0.
- [27] Jung JK, Park SW, Liu WY, Kang BC. Discovery of single nucleotide polymorphism in Capsicum and SNP markers for cultivar identification. *Euphytica*. 2010 Sep;175:91-107.
- [28] Kalendar R, Schulman AH. IRAP and REMAP for retrotransposon-based genotyping and fingerprinting. *Nature protocols*. 2006 Dec;1(5):2478-84.
- [29] Kalendar R, Antonius K, Smýkal P, Schulman AH. iPBS: a universal method for DNA fingerprinting and retrotransposon isolation. *Theoretical and Applied Genetics*. 2010 Nov;121:1419-30.
- [30] Kalendar R, Raskina O, Belyayev A, Schulman AH. Long tandem arrays of Cassandra retroelements and their role in genome dynamics in plants. *International Journal of Molecular Sciences*. 2020 Apr 22;21(8):2931.
- [31] Kalendar R, Baidyussen A, Serikbay D, Zotova L, Khassanova G, Kuzbakova M, Jatayev S, Hu YG, Schramm C, Anderson PA, Jenkins CL. Modified "Allele-specific qPCR" method for SNP genotyping based on FRET. *Frontiers in plant science*. 2022 Jan 10;12:747886.
- [32] Kalendar R, Hunter C, Orbovic V. Innovative applications of sequencing technologies in plant science. *Frontiers in plant science*. 2022 Oct 26;13:1058347.
- [33] Kalendar R, Shustov AV, Akhmetollayev I, Kairov U. Designing allele-specific competitive-extension PCR-based assays for high-throughput genotyping and gene characterization. *Frontiers in Molecular Biosciences*. 2022 Mar 1;9:773956.
- [34] Karihaloo JL. DNA fingerprinting techniques for plant identification. *Plant Biology and Biotechnology: Volume II: Plant Genomics and Biotechnology*. 2015:205-21.
- [35] Kim DH, Kang WH, Yeom SI, Kim BD. Isolation of putative pepper defense-related genes against the pathogen *Phytophthora capsici* using suppression subtractive hybridization/microarray and RNA-sequencing analyses. *Horticulture, Environment, and Biotechnology*. 2019 Oct;60:685-99.
- [36] Köhnemann S, Nedele J, Schwotzer D, Morzfeld J, Pfeiffer H. The validation of a 15 STR multiplex PCR for Cannabis species. *International journal of legal medicine*. 2012 Jul;126:601-6.
- [37] Konieczny A, Ausubel FM. A procedure for mapping Arabidopsis mutations using co-dominant ecotype-specific PCR-based markers. *The plant journal*. 1993 Aug;4(2):403-10.
- [38] Korir NK, Han J, Shangguan L, Wang C, Kayesh E, Zhang Y, Fang J. Plant variety and cultivar identification: advances and prospects. *Critical reviews in biotechnology*. 2013 Jun 1;33(2):111-25.
- [39] Kress WJ, Wurdack KJ, Zimmer EA, Weigt LA, Janzen DH. Use of DNA barcodes to identify flowering plants. *Proceedings of the National Academy of Sciences*. 2005 Jun 7;102(23):8369-74.
- [40] Lasky JR, Des Marais DL, McKAY JK, Richards JH, Juenger TE, Keitt TH. Characterizing genomic variation of Arabidopsis thaliana: the roles of geography and climate. *Molecular Ecology*. 2012 Nov;21(22):5512-29.
- [41] Lezar S, Myburg AA, Berger DK, Wingfield MJ, Wingfield BD. Development and assessment of microarray-based DNA fingerprinting in Eucalyptus grandis. *Theoretical and Applied Genetics*. 2004 Nov;109:1329-36.
- [42] Li Z, Yu H, Fang Z, Yang L, Liu Y, Zhuang M, Zhang Y. Development of SNP markers in cabbage and construction of DNA fingerprinting of main varieties. *Scientia Agricultura Sinica*. 2018;51(14):2771-87.
- [43] Li Y, Xiao J, Chen L, Huang X, Cheng Z, Han B, Zhang Q, Wu C. Rice functional genomics research: past decade and future. *Molecular plant*. 2018 Mar 5;11(3):359-80.
- [44] Liu X, Cheng J, Mei Z, Wei C, Khan MA, Peng J, Fu J. SCAR marker for identification and discrimination of specific medicinal Lycium chinense Miller from Lycium species from ramp-PCR RAPD fragments. *3 Biotech*. 2020 Aug;10:1-7.
- [45] Makhoul M, Rambla C, Voss-Fels KP, Hickey LT, Snowdon RJ, Obermeier C. Overcoming polyploidy pitfalls: a user guide for effective SNP conversion into KASP markers in wheat. *Theoretical and Applied Genetics*. 2020 Aug;133:2413-30.
- [46] Mbwana J, Bolin I, Lyamuya E, Mhalu F, Lagergard T. Molecular characterization of Haemophilus ducreyi isolates from different geographical locations. *Journal of clinical microbiology*. 2006 Jan;44(1):132-7.

- [47] Melo AT, Bartaula R, Hale I. GBS-SNP-CROP: a reference-optional pipeline for SNP discovery and plant germplasm characterization using variable length, paired-end genotyping-by-sequencing data. *BMC bioinformatics*. 2016 Dec;17(1):1-5.
- [48] Molecular Medical Microbiology (Second Edition), Academic Press, 2015, Pages 2057-2096
- [49] Nybom H, Weising K, Rotter B. DNA fingerprinting in botany: past, present, future. *Investigative genetics*. 2014 Dec;5(1):1-35.
- [50] Nybom H, Weising K. 6 DNA-Based Identification of Clonally Propagated Cultivars. *Plant breeding reviews*. 2011;34:221.
- [51] Nybom H. DNA fingerprints in sports of 'Red Delicious' apples. *HortScience*. 1990 Dec 1;25(12):1641-2.
- [52] Onley IR, Austin JJ, Mitchell KJ. Sex assignment in a non-model organism in the absence of field records using Diversity Arrays Technology (DArT) data. *Conservation Genetics Resources*. 2021 Sep;13:255-60.
- [53] Palmer JD, Herbon LA. Plant mitochondrial DNA evolved rapidly in structure, but slowly in sequence. *Journal of Molecular evolution*. 1988 Dec;28:87-97.
- [54] Poets A, Silverstein K, Pardey PG, Hearne S, Stevenson J. DNA fingerprinting for crop varietal identification: Fit-for-purpose protocols, their costs and analytical Implications. 2020 Apr
- [55] Poghosyan A, Carrillo A, Hernandez J, Barraza A, Lebsky V. Microbiomes of soil and roots of two palm species infected with 'Candidatus Phytoplasma palmae' in two different ecosystems: Single strand conformation polymorphism analysis. *Phytopathogenic Mollicutes*. 2023;13(1):27-8.
- [56] Reddy PM, Sarla N, Siddiq EA. Inter simple sequence repeat (ISSR) polymorphism and its application in plant breeding. *euphytica*. 2002 Nov;128:9-17.
- [57] Ritz CM, Schmutts H, Wissemann V. Evolution by reticulation: European dogroses originated by multiple hybridization across the genus *Rosa*. *Journal of Heredity*. 2005 Jan 1;96(1):4-14.
- [58] Rogstad SH, Patton 2nd JC, Schaal BA. M13 repeat probe detects DNA minisatellite-like sequences in gymnosperms and angiosperms. *Proceedings of the National Academy of Sciences*. 1988 Dec;85(23):9176-8.
- [59] Sankar AA, Moore GA. Evaluation of inter-simple sequence repeat analysis for mapping in Citrus and extension of the genetic linkage map. *Theoretical and Applied Genetics*. 2001 Feb;102:206-14.
- [60] Schulman AH. Molecular markers to assess genetic diversity. *Euphytica*. 2007 Dec;158(3):313-21.
- [61] Seeb JE, Carvalho G, Hauser L, Naish K, Roberts S, Seeb LW. Single- nucleotide polymorphism (SNP) discovery and applications of SNP genotyping in nonmodel organisms. *Molecular ecology resources*. 2011 Mar;11:1-8.
- [62] Selkoe KA, Toonen RJ. Microsatellites for ecologists: a practical guide to using and evaluating microsatellite markers. *Ecology letters*. 2006 May;9(5):615-29.
- [63] Selvakumari E, Jenifer J, Priyadarshini S, Vinodhini R. Application of DNA fingerprinting for plant identification. *JAIR*. 2017 Mar;5(10):149-51.
- [64] Singh R, Iqbal MA, Mishra CN, Jaiswal S, Kumar D, Raghav N, Paul S, Sheoran S, Sharma P, Gupta A, Tiwari V. Development of model web-server for crop variety identification using throughput SNP genotyping data. *Scientific Reports*. 2019 Mar 26;9(1):5122.
- [65] Stöltgen KN, Nipper R, Lindtke D, Caseys C, Waeber S, Castiglione S, Lexer C. Genomic scan for single nucleotide polymorphisms reveals patterns of divergence and gene flow between ecologically divergent species. *Molecular ecology*. 2013 Feb;22(3):842-55.
- [66] Tanksley SD, Young ND, Paterson AH, Bonierbale MW. RFLP mapping in plant breeding: new tools for an old science. *Bio/technology*. 1989 Mar 1;7(3):257-64.
- [67] Thakur VV, Tiwari S, Tripathi N, Tiwari G. Molecular identification of medicinal plants with amplicon length polymorphism using universal DNA barcodes of the *atpF-atpH*, *trnL* and *trnH-psbA* regions. *3 Biotech*. 2019 May;9(5):188.
- [68] Tian HL, Wang FG, Zhao JR, Yi HM, Wang L, Wang R, Yang Y, Song W. Development of maizeSNP3072, a high-throughput compatible SNP array, for DNA fingerprinting identification of Chinese maize varieties. *Molecular Breeding*. 2015 Jun;35:1-1.
- [69] WANG LX, Jun QI, CHANG LF, LIU LH, LI HB, PANG BS, ZHAO CP. Assessment of wheat variety distinctness using SSR markers. *Journal of Integrative Agriculture*. 2015 Oct 1;14(10):1923-35.
- [70] Wegscheider E, Benjak A, Forneck A. Clonal variation in Pinot noir revealed by S-SAP involving universal retrotransposon-based sequences. *American Journal of Enology and Viticulture*. 2009 Mar 1;60(1):104-9.
- [71] Weising K, Nybom H, Pfenninger M, Wolff K, Kahl G. DNA fingerprinting in plants: principles, methods, and applications. CRC press; 2005 Feb 28.

- [72] Williams JGK, Kubelik AR, Livak KJ, Rafalski JA, Tingey SV. DNA polymorphisms amplified by arbitrary primers are useful as genetic markers. *Nucleic Acids Res.* 1990; 18:6231–6235.
- [73] Ye C, Yu Z, Kong F, Wu S, Wang B. R-ISSR as a new tool for genomic fingerprinting, mapping, and gene tagging. *Plant molecular biology reporter.* 2005 Jun;23:167-77.
- [74] Yoon CK. Botanical witness for the prosecution. *Science.* 1993 May 14;260(5110):894-5.
- [75] Zabeau M, Vos P. Selective restriction fragment amplification: a general method for DNA fingerprinting. European Patent Office, publication 0 534 858 A1, bulletin 93/13
- [76] Zhang J, Yang J, Zhang L, Luo J, Zhao H, Zhang J, Wen C. A new SNP genotyping technology Target SNP-seq and its application in genetic analysis of cucumber varieties. *Scientific reports.* 2020 Mar 27;10(1):5623.
- [77] Zhang Z, Zhang A, Gao D, Wei Z. Genetic diversity analysis and rapid identification of Grapevine Jufeng Cultivars by MCID method. *Journal of Northeast Agricultural Sciences.* 2022; 4:38-42.
- [78] Zheng ZH, Zhou MJ, He XY, Luo JW, Cheng RB. Application status of molecular marker technology in medicinal plants and research progress of *Christina loosestrife*. *Agric. Technol. Serv* 2021; 38:43–46.
- [79] ZI S, Li W, Jin X, Zhang H. A new nuclear DNA marker revealing both microsatellite variations and single nucleotide polymorphic loci: a case study on classification of cultivars in *Lagerstroemia indica* L. *J Microb Biochem Technol.* 2016;8:266-71.