

Review Article

<https://doi.org/10.20546/ijcmas.2021.1004.078>

## The Enormous World of Diverse Molecular Markers

Ankita Chauhan\*, Megha Bhatt and Pushpa Lohani

*Department Of Molecular Biology and Genetic Engineering, College of Basic Science and Humanities, G.B Pant University of Agriculture and Technology, Pantnagar-263145 (Udham Singh Nagar) Uttarakhand, India*

*\*Corresponding author*

### ABSTRACT

The creation and use of molecular markers for the recognition and exploitation of DNA polymorphism is one of the most important developments in the field of molecular biology. Molecular markers, due to their consistency, low cost and effortless use have become a popular tool for a variety of applications including gene tagging, phylogenetic and genetic diversity analysis and genome mapping. Many molecular marker techniques were utilized for over a decades. However, only a few of these techniques, like RFLPs, RAPDs, AFLPs, ISSRs, SSRs and SNP shave received universal acceptance. A current revolution in DNA sequencing techniques has taken the encounter and usefulness of molecular markers to high-throughput and ultrahigh-throughput levels. Although the choice of the marker will noticeably depend on the targeted usage, microsatellites and SNPs largely fulfill many of the user's necessities. This review is intended to be a synopsis of recent developments in molecular markers and their applications in research fields.

#### Keywords

Molecular markers, RAPDs, SSRs, SNPs.

#### Article Info

*Accepted:*

18 March 2021

*Available Online:*

10 April 2021

### Introduction

DNA sequences that are eagerly identified or readily detected and whose inheritance can effortlessly observed are referred as molecular markers. Due to their capability to distinguish between genotypes, they are in high demand in the genetics-based research fields. On other way a genetic marker can also be defined in

one of the following ways: (i) a specific section of DNA with a known location on the genome;(ii)a chromosomal landmark that permits the tracing of a particular section of DNA or (iii) representative whose phenotypic expression is generally easily distinguished, used to recognize a species that carries it, or as a probe to spot chromosomes or nucleus (King and Stansfield, 1990). The reason for the

enormous use of these markers is that they are highly polymorphic and are not influenced by the environment. They also provide the necessary information for biodiversity studies (Elshibli and Korpelainen, 2008). The genes they mark are close together or distinct from each other on the same chromosome, they tend to stay together as each generation of plants. So that one can learn where markers occur on a chromosome, and how close they are to specific genes, they can be used to create a genetic linkage map. A marker may have useful significances, such as altering the expression or function of a gene that directly involved for the development of diseases.

Apart from locating genes, they can prove very useful for a variety of purposes relevant to crop improvement. For instance, these markers have been utilized broadly for the preparation of saturated molecular maps (genetical and physical). Markers association with genes/QTLs controlling the traits of economic importance has also been employed in some cases for indirect marker-assisted selection (MAS) (e.g. Koebner 2004, Korzun 2002). And can even be used for cistron introgression by backcrossing, germplasm characterization, genetic diagnostics, genotype diversity analysis, characterization of transformants, the study of genome organization and phylogenetic analysis (Jain *et al.*, 2002). Genetic markers are further divided into 3 broad categories: those supported visually assessable traits (morphological and agronomic traits), those supported on gene product (biochemical markers), and those depend on a DNA assay (molecular markers). Genetic markers based on visual traits involve crossing whole genomes, followed by the selection of the superior recombinants from among the several segregation products. The entire process is laborious and time-consuming, including numerous crosses, many generations, and attentive phenotypic investigation, and the linkage drag make it

more difficult to attain the wanted results. Though genetic markers appeared to be used early in works of literature (Sax, 1932; Wexelsen, 1933) but the development of molecular markers has greatly enhanced our understanding of life sciences. Molecular markers are constant landmarks in the genome, they are not genes because they do not have any biological effect. They are recognizable DNA sequences, found at precise locations of the genome, and transmitted by the standard laws of inheritance from one generation to the next.

Here, we present a detailed review on the basic principles, necessities, pros and cons of the most widely used molecular markers for genetic diversity studies, genetic mapping, marker-trait association studies, and marker-assisted selection programs. The molecular markers can be categorized into different types (Fig 1) based on the process of detection.

### **PCR-based markers for random genome profiling**

#### **Random amplified polymorphic DNA (RAPD)**

RAPD is simplest and first PCR based molecular marker technique. This technique is based on the amplification of genomic DNA with single primer of an arbitrary nucleotide sequence with no previous information about the genome. It is impossible to distinguish either the amplified DNA segment is heterozygous (two different copies) or homozygous (two identical copies) at a definite locus as most of the RAPD markers are dominant. Although codominant RAPD markers are also present in some cases, in which different-sized DNA segments are amplified from the same locus. PCR amplification from RAPD primers would be expected only when the priming site occurs twice in opposite orientation within

approximately 2000 bases. Sequence variation at priming sites and length variation in the target sequence lying between primer binding sites results in RAPD polymorphism.

RAPD markers gives numerous advantages like requirement of only a minute amount of DNA, technical simplicity, no requirement of prior information of the DNA sequence and viability of automation and higher frequency of polymorphism. RAPDs have extensively needed for identification of polymorphisms at a number of loci concurrently (Grattapaglia & Sederoff, 1994; Hu & Quiros, 1991). Further benefit of RAPDs is that huge numbers of markers are obtained which are often distributed over the genome. A number of RAPD variants are present which includes: Arbitrarily primed PCR (AP-PCR), DNA amplification fingerprinting (DAF), Mini-hairpin primer-driven DNA amplification fingerprinting (mhpDAF), Nucleic acid scanning by hybridization (NASBH), Arbitrary signatures from amplification profiles (ASAP) and many more. These modified together called multiple arbitrary amplicon profiling (MAAP) or single primer amplified regions (SPAR).

The major drawback of RAPD markers is that they are not always reproducible. Moreover, they are dominant markers and cannot differentiate between homologous and heterologous loci.

### **Inter-simple sequence repeat (ISSR)**

ISSR first described by Thomas *et al.*, in 1993. According to them these are amplification reactions with a single long oligonucleotide (microsatellite) primer anchored at either of the ends. ISSR markers basically involves amplification of DNA segments present at a significant (amplified) distance in between two identical microsatellite repeat regions oriented

in opposite direction in the genome of organism. ISSR technique targeting multiple genomic loci to amplify mainly inter simple sequence repeats of different sizes by using microsatellite as primers in a single primer PCR reaction. ISSRs contain reproducibility and specificity with the advantages of RAPD. Reproducibility is produced because longer primers are used for PCR amplification compared to RAPDs. Specificity is enhanced as higher annealing temperature is employed in the PCR. As in the case of RAPDs, practically no prior target sequence knowledge is required for ISSRs, and can thus be applied to any species with ease. Richness of microsatellites in genomes and their hyper-variability among different individuals of inter or intra species confirms the effectiveness of ISSR markers. Fastening of primer with non-repeat sequences ensure that the amplification commences at the same position in each cycle.

ISSRs display the specificity of microsatellite markers, but do not requires equence information for primer synthesis. Apart from these the technique is simple, quick, and the use of radioactivity or fluorescence is not required. These markers generally exhibit high polymorphism (Kojima *et al.*, 1998). However, the level of polymorphism totally depend upon the detection method used. Numerous other variants of ISSR are present like Random amplified microsatellite polymorphism (RAMP), Double stringency PCR (DS-PCR) and Random amplified hybridization microsatellites (RAHM) etc. The main disadvantages of ISSRs are homology of co-migrating amplification products, dominant inheritance and reproducibility.

### **Amplified fragment length polymorphism (AFLP)**

AFLP technique depends on the PCR amplification of restriction fragments obtained as a result of digestion of genomic DNA. The

PCR products obtained are further ligated to oligonucleotide adapters of a few nucleotide bases. PCR amplification is carried out under highly rigorous conditions. These primers, despite being adapter specific, also carry random extensions of 1–3 nucleotides at 3'-end to selectively amplify genomic fragments from random sites. AFLP is the newest form of marker assisted selection and is a highly sensitive method based on the mutual concepts of RFLP and RAPD. This technique is relevant to most species giving very reproducible outcomes. Polymorphism may be because of alterations in the restriction recognition sites. High frequency of identifiable polymorphisms, more reproducibility and rapid generation make AFLP analysis an attractive technique for determining linkages by analysing individuals from segregating populations of the different organism (Alexander *et al.*, 2012). AFLPs are far more reproducible than RAPDs (Savelkoul *et al.*, 1999), thus are more reliable for population and genetic studies. Like RAPDs, AFLPs too are distributed all over the genome. Therefore AFLP study is suitable for construction of genetic linkage maps (Jorg *et al.*, 1995). However, polymorphic information content of AFLPs is usually smaller than that of SSR markers (Pejic *et al.*, 1998).

AFLPs normally show dominant inheritance and are expensive, labour intensive and require technical skill as the fragments are identified by either silver staining or fluorescence or radioactivity. The bands are resolved on large sequencing gels or automated DNA sequencers. The alternatives used to mention AFLPs are Selective Fragment Length Amplification (SFLA) and Selective Restriction Fragment Amplification (SRFA).

AFLPs has different variants namely, Selective Amplification of Microsatellite Polymorphic Loci (SAMPL), Sequence-

specific amplified polymorphism (S-SAP), Amplification of insertion mutagenized sites (AIMS), Resistance gene analogue anchored amplified fragment length polymorphism (AFLP-RGA) and Three endonuclease AFLP (TE-AFLP).

The main limitations of AFLPs are i) it is laborious and expensive ii) requires more number of steps iii) requires template DNA free of inhibitor compounds that interfere with the restriction enzyme iv) AFLP loci are dominant, which do not separate dominant homozygotes from heterozygotes. This lessens the precision of AFLP markers in the marker-assisted selection, genetic mapping, and population genetic analysis.

### **Sequence-characterized amplified region (SCAR)**

SCAR marker, a genomic DNA fragment is identified by PCR amplification employing a pair of specific oligonucleotide primers (Paran and Michelmore 1993). Sequence-characterized amplified regions (SCARs) is also derived from RAPDs (Paran & Michelmore, 1993) or AFLPs (Sharma *et al.*, 2007) or any other multi-locus marker technique, overcoming the constraints of arbitrary profiling. It comprises cloning and sequencing of amplified DNA fragment of interest and designing long primers complementary to their ends. Unlike RAPDs, SCARs do not have poor reproducibility because of the use of longer PCR primers. Compared to arbitrary primers, SCARs exhibit several benefits in mapping studies. Co-dominant SCARs are more informative for genetic mapping than dominant RAPDs, also upright in map-based cloning as they can be expended to screen pooled genomic libraries by PCR, locus specificity, physical mapping etc. Only small quantity of template DNA is required for analysis. The technique is rapid and easy to practice. The major drawback

about SCARs is need of sequence data for designing of the PCR primers.

### **Resistance gene analogue polymorphism (RGAP)**

Resistance gene analogue polymorphism (RGAP) is another PCR-based technique useful in plants, in which degenerate primers are required to intensify plant resistance genes and their equivalents. RGAPs can be produced by prior sequence knowledge, or through data mining of genomic or EST databases.

### **Cleaved Amplified Polymorphic Sequence (CAPS)**

CAPS are DNA fragments amplified by PCR using precise 20-25 bp primers. The products of PCR are digested with a specific restriction enzyme. Then, variation in the occurrence of restriction sites in the digested products resulted in length polymorphism, which is identified by gel electrophoresis of the digested products (Akopyanz *et al.*, 1992; Konieczny and Ausubel 1993).

The main benefits of CAPS includes: the high reproducibility, the codominance of allele and it require low quantities of template DNA (50-80 ng per reaction).

On Comparing with RFLPs, CAPS examination doesn't include the technically demanding and laborious stages of Southern blot hybridization and radioactive recognition procedures.

The polymorphism in CAPS is more challenging to discovery because of restricted size of the amplified fragments and sequence requirement for PCR primers designing. CAPS is sometimes stated as PCR-Restriction Fragment Length Polymorphism (PCR-RFLP) and predominantly applied in gene mapping studies.

### **Repetitive sequence based PCR (rep-PCR)**

The rep-PCR method is grounded on the use of DNA primers comprising of naturally occurring interspersed repetitive elements such as, ERIC, BOX etc. Rep-PCR reaction was first established in bacteria (Versalovic *et al.*, 1994). As they are long, defined and conserved, LTRs serve as perfect priming sites in this technique. These markers are particularly useful in establishing genetic relationships and studying phylogeny among closely related groups of individuals (Alavi-Kia *et al.*, 2008).

The research carried out on different marker technology as well as their popularity has been summarised in Fig 2. Molecular markers have gained vast acceptance in past years with markers designed for targeting specific sites in genome being used widely

### **PCR-based markers targeting specific genome sites**

#### **Micro satellites as genetic markers (SSR)**

Microsatellites, also called as simple sequence repeats (SSR), variable number tandem repeats (VNTR) and short tandem repeats (STR) and also sub-categories of short tandem repeat (STR). SSRs are polymorphic DNA loci comprising repeated nucleotide sequences, generally from 2 to 7 nucleotides per unit. The length of the repeater unit is that the same for the bulk of the repeats within an individual microsatellite locus; the number of repeats for a particular locus may differ, leading to alleles of varying lengths. These repeated motifs are denoted (GACC)<sub>n</sub>, where n is that the number of tandem repeats. The sequences flanking these microsatellites are often conserved and may be accustomed to design primers. These primers will be designed by constructing a unique genomic library and sequencing segments of the subject

genome. Previously discovered sequence (i.e.: GENEBANK online database) can also be used for SSRs and primers designing. Polymorphism relies on the number of tandem repeats and so the length of the PCR products.

Like RFLP, SSR are also co dominant marker and is usually metaphor agarose or polyacrylamide gels are used for visualization. Existence of short tandem repeats of variable length is the main characteristic of microsatellite loci (Akkaya *et al.*, 1992). The simple sequence repeats (SSR) or microsatellites (sometimes remarked to as a variable number of tandem repeats or VNTRs) that are similar in nature are shown to be abundant and highly polymorphic in eukaryotic genomes. They can be di-nucleotide repeats (AG)<sub>n</sub>, (AT)<sub>n</sub> and (AC)<sub>n</sub> or tri-nucleotide or tetra-nucleotide repeats. Microsatellite markers are examined by PCR-amplification of a small genomic region comprising of the repeated sequence and use gel separation for size estimation of the repeated sequence. SSRs have been the principal used markers for genotyping plants over the ancient years because they're highly co-dominant, multi-allele and informative genetic markers. They are experimentally reproducible and transferable among closely related organisms (Mason, 2015). SSRs are specifically useful for wild species for carrying out studies on diversity measured on the premise of genetic distance. It may also be useful to estimate gene flow and crossing over rates. SSR find a great importance in evolutionary studies as well as inference of intra-specific genetic relations. On the opposite hand, for cultivated species; SSRs are commonly used for constructing linkage maps, mapping loci involved in quantitative traits (QTL), estimating the degree of association between genotypes using marker-assisted selection and defining cultivar DNA fingerprints (Jonah *et al.*, 2011; Kalia *et al.*, 2011). Around many algorithms are present

for SSR search which includes TRF (Tandem, 1999), MISA (Thiel *et al.*, 2003), TROLL (Castelo *et al.*, 2002), SSRIT (Temnykh, 2001), SciRoKo (Kofler *et al.*, 2007) and SSRFinder (Gao *et al.*, 2003). SSR are powerful genetic markers because of their high genetic variability. SSRs have been established as a valuable tool for genome mapping in many individuals (Knapik *et al.*, 1998). Lately, a source of SSRs development from expressed sequence tag (EST) databases has been used (Kota *et al.*, 2001; Kantety *et al.*, 2002; Michalek *et al.*, 2002). With the provision of huge numbers of ESTs and other DNA sequence data, development of EST-based SSR markers through data processing has become an efficient, comparatively inexpensive and quick compared with the progress of genomic SSRs (Gupta *et al.*, 2003). In this method the expensive and time-consuming methods of producing genomic libraries and sequencing of large numbers of replicates for finding the SSR regions are not needed (Eujayl *et al.*, 2004). Though, the expansion of EST-SSRs is limited to organism for which this type of database is present. Furthermore, the EST-SSR markers are reported to possess lower rate of polymorphism compared to the SSR markers derived from genomic libraries (Cho *et al.*, 2000; Scott *et al.*, 2000; Eujayl *et al.*, 2002; Chabane *et al.*, 2005).

The main disadvantages of SSRs are i) the high development expenses ii) suitable primer sequences may not be available for unstudied collections iii) mutations within the primer annealing sites may end in the occurrence of null alleles which may result in errors in genotype scoring.

### **Single-Strand Conformation Polymorphism (SSCP)**

These are short DNA fragments (200-800 bp) amplified by PCR with precise primers of 20-

25 bp. To identify nucleotides variation between the amplified fragments, gel electrophoresis of this single-strand DNA is used. The marker basically depend on the fact that the electrophoretic movement of single-strand DNA depends on the conformation (secondary structure) of the molecule, which changes significantly due of the mutation in the DNA. Thus, SSCP provides a method to detect nucleotide variation among DNA samples without sequence reactions (Orita *et al.*, 1989).

SSCP is a potential tool for studying high throughput DNA polymorphism. It is useful in the detection of heritable human diseases and mutations in genes using gene sequence information. This technique is not extensively used in plant studies. However, its application in differentiating progenies can be used when precise primers are designed for agronomically useful characters (traits).

Benefits of SSCP comprise of the need of low quantities of template DNA (10-100 ng per reaction). Other advantages include the codominance of alleles and the ease of carrying out the technique. The drawbacks of this technique are i) requirement of sequence information for primer designing ii) the requirement of extremely optimized electrophoretic conditions in order to get reproducible out comes iii) occasionally mutations can remain unidentified and henceforth absence of mutation maynot be verified.

### **Sequence Tagged Site (STS)**

Olsen *et al.*, (1989) developed a short and unique sequence called STS, as DNA landmarks in the physical mapping of the animal genome, and later accepted in plants. Clones (two or more) comprising the same STS must overlap and the overlap must contain STS. Although, any clone can be

sequenced and can be used as STS only if it comprises a unique sequence. STS is characterized by a couple of PCR primers designed by sequencing either an RFLP probe representing a mapped low copy number sequence or AFLP fragments, mainly in case of plants. Though the major technical challenge is the conversion of AFLP markers into STS markers and which repeatedly frustrating in many polyploids such as wheat, rice but despite of these it has been successful in several other cereals (Guo *et al.*, 2003).

The primers designed on the basis of RAPD are also referred to as STS, sometime also called as SCARs. STS markers are highly reproducible, codominant, technically simple to use and suitable for high throughput analysis.

### **DNA Amplification Fingerprinting (DAF)**

DAF are single arbitrary primers up to 5-8 bases long used to amplify DNA using PCR. They were first employed by Caetano-Anolles *et al.*, (1991). In a range of products found, simple patterns are useful as genetic markers for mapping, while more complex patterns are useful for DNA fingerprinting. Reproducible band patterns can be studied using silver staining and polyacrylamide gel electrophoresis. DAF needs careful standardization of different parameters.

However, it is extremely agreeable to automation and tagging of primers fluorescently for early and easy identification of amplified fragments. This approach, is simple, rapid and sensitive.

It does not use cloned probes and is independent of prior DNA sequence information. DAF contours can be tailored by using various alterations such as predigesting of templates etc. It is very useful in mapping and genetic typing.

### **Randomly Amplified Microsatellite Polymorphisms(RAMPO)**

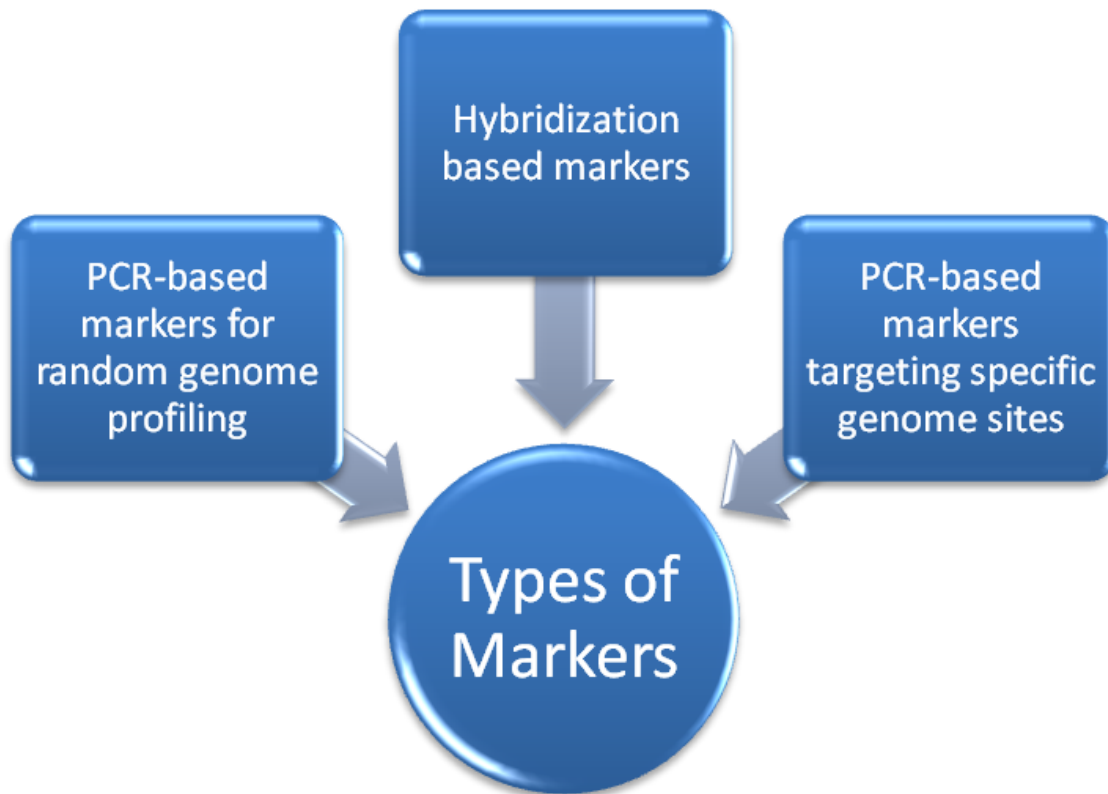
RAMPO is a PCR-based method that amplifies genomic DNA using random (RAPD) primers. The amplified fragments are separated by gel electrophoresis and with the help of microsatellite oligonucleotide probes, the dried gel is hybridized. The collective benefits of microsatellite-primed PCR (Gupta *et al.*, 1994; Weising *et al.*, 1995), oligonucleotide fingerprinting and RAPD (Williams *et al.*, 1990) are: high sensitivity, quick assay, the non-requirement of prior DNA sequence information and high level of variability detection. These markers have been established to be a potentially precious marker for the study of genetic relationships in cultivated plant individuals. This technique

has also used in the genetic fingerprinting of closely-related genotypes of *D. bulbifera*, kiwi fruit and tomato(Richardson *et al.*, 1995).

### **Single nucleotide polymorphism (SNP)**

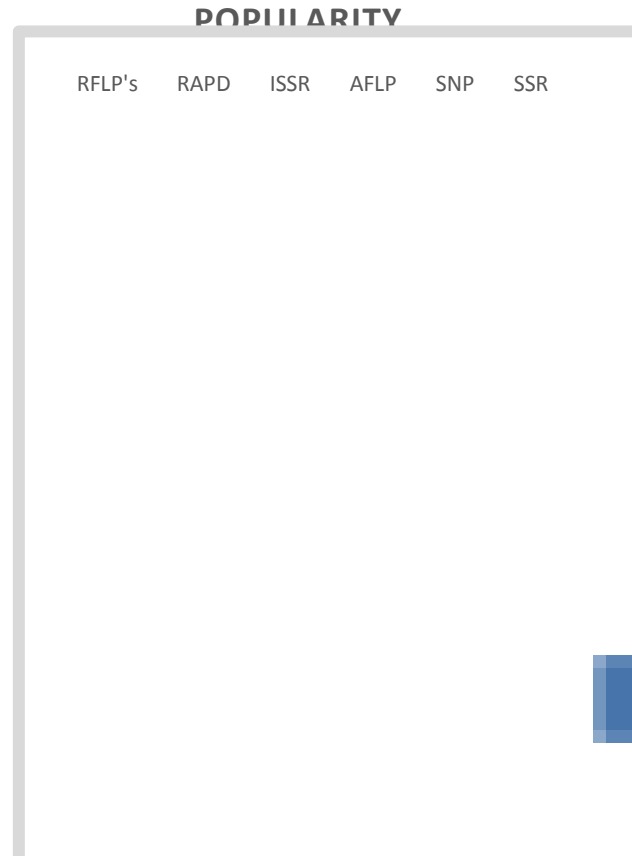
The Single nucleotide polymorphisms, frequently called SNPs, are the most common type of molecular markers. Every Single nucleotide polymorphism characterizes a variation in a single nucleotide in different sequences. On an average SNP arise nearly once in every 1,000 nucleotides. These variations may be unique or occur in same pattern in many organisms. They act as biological markers, helps to locate genes that are associated with disease or agronomically important traits.

**Fig.1** Classification of different types of markers





**Fig.2** Popularity of different molecular markers based on total number of citations in year 2003-2019 on google scholar.



**Table.1** Comparative description of major classes of molecular marker

	Development cost	Polymorphism	Quality of DNA required	Quantity of DNA required	Running cost	Prior sequence knowledge	Throughput	Technical expertise
<b>RFLP and variants</b>	High	High	High	High	High	High	Low	High
<b>AFLP and variants</b>	High	High	High	High	High	Low	Low	High
<b>SSR</b>	Low	High	Low	Low	Low	High	High	Low
<b>SNP</b>	Medium	High	High	Low	Medium	High	High	High
<b>DArT and variants</b>	High	High	High	Low	High	High	Very High	High
<b>SSCP</b>	Low	High	High	Low	Medium	Low	Low	High
<b>SPARs</b>	Low	Medium	Low	Low	Low	Low	Low	Low
<b>GAAP and Other functional markers</b>	Low	Medium	High	Low	Low	High	Low	Low

In other words, SNP is a DNA sequence in which variation occurs when a single nucleotide (A, T, G or C) differs among rest of the individuals. Although, most of the SNPs exists within non-coding genomic regions, but they may exist in coding region too. SNP is the most abundant marker both in plant and animal genomes and has recently been developed as a new generation molecular marker for carrying out different studies. These markers are co-dominant and therefore are able to efficiently differentiate between homozygous and heterozygous alleles. Besides, unlike microsatellites their power not only comes from the number of alleles but also from the large number of loci that can be evaluated (Foster *et al.*, 2010).

The SNP can be discovered in low diversity organisms where the genetic population discrimination power can be equivalent to the same number of loci in genetically varied individuals. The problem of homoplasy is less subject due to the evolutionary preserved nature of SNPs (Brumfield *et al.*, 2003). Most significantly, SNPs are open to rapid and efficient genotyping of large numbers of samples and high throughput automation (Tsuchihashi and Dracopoli 2002). In plants, SNP can be designed from single-stranded pyrosequence (Miller *et al.*, 2003) and ESTs (Coles *et al.*, 2005). The technique in SNP where thousands of oligonucleotide probes attached to a solid surface (e.g., glass, silicon wafer) allowing for a large number of SNPs to be examined instantaneously called SNP array. (Rapley and Harbron 2001).

SNP is able to identify genetic diversity in plants, mainly in organisms with partial genetic diversity. Identification of population genetic structure of Castor bean (*Ricinus communis*) using SNPs from genome wide comparisons presented low levels of genetic diversity and mixing of genotypes, leading to lower geographic organizing of worldwide castor bean populations (Foster *et al.*, 2010).

Many assays have been established for the development and detection of SNPs for the purpose of genotyping. NGS based SNP discovery methods such RAD and crops have become extremely widespread (Appleby *et al.*, 2009; Davey *et al.*, 2011; Martin *et al.*, 2013).

Automated methods like Golden Gate Technology of Illumina have amplified applications in genotyping (Mammadov *et al.*, 2010; Taniguchi *et al.*, 2013; Thakur *et al.*, 2014; Trebbi *et al.*, 2011). Single Nucleotide Polymorphism calling is based on a reference, which in most cases is a high quality genome sequence of the species under examination (Ramos *et al.*, 2009).

### **Inter-retrotransposon amplified polymorphism (IRAP)**

IRAP is a PCR dependent technique which identify a high level of polymorphism. This technique does not require DNA digestion, ligations and hybridization with probe to produce a marker. It therefore increases the reliability and sturdiness of the method. It is a retrotransposon-type marker developed by Kalendar *et al.*, in 1999. They are alternative retrotransposon-based markers. In IRAP method, PCR oligonucleotide primers face away from terminal retrotransposon sections. Therefore, they amplify the fragment between two retrotransposon insertions.

The integration patterns of retrotransposon differ between different genotypes, the size and number of the resulting amplicons can be used, to reconstruct phylogenies or to measure genetic diversity or for differentiation of genotypes or cultivars.

Recently, in plants species IRAP has been used to study phylogenetic and genetic diversity (Dalong *et al.*, 2006; Alavi-kia *et al.*, 2008) and fungal pathogens (Shull and Hamer, 1996; Pasquali *et al.*, 2007).

## **Hybridization based markers**

### **Allele-Specific Associated Primers (ASAPs)**

AS-PCR also known as PCR amplification of specific alleles (PASA) or amplification refractory mutation system (ARMS) is a PCR-dependent method which can be used to identify the known SNPs. Newton *et al.*, (1989) gave the concept of AS-PCR. In this method, the precise primers are designed for amplification by DNA polymerase when the nucleotide at the 3'-end of the primer complements the base at the variant sequences. The patterns of precise PCR products permit the variation of the SNP, after gel electrophoresis. Many advanced approaches have been used to identify the presence of exact PCR product.

Out of which some are based on probe hybridization which involves precise tagged probes and some requires nucleic acid stains for meting curve analysis. In many areas of study such as genetic disorder, microbiology, pharmaco genetics etc, Here Allele-Specific PCR been used widely.

Obtaining an allele-specific marker requires the sequence of a specific allele (either in homozygous or heterozygous state). The precise primers are designed which are important for amplification of DNA templates to produce a single fragment at rigorous annealing temperatures in PCR. If talked about the similarity with SCARs, these markers (tag specific alleles) are more or less similar to SCARs (Gu *et al.*, 1995).

### **Diversity arrays technology (DArT)**

It is a hybridization-based technique for typing numerous thousand loci in a single assay (Jaccoud *et al.*, 2001), and creating whole-genome fingerprints by scoring the presence or absence of DNA fragments in the genome.

Currently, DArT has assumed the status of highly robust, reliable and suitable markers for analysis of genetic mapping, as well as genetic diversity using linkage or association studies both in a variety of model and non-model crops such as Sugarcane (Aitken *et al.*, 2014), eucalyptus (Sansaloni *et al.*, 2010), pearl millet (Supriya *et al.*, 2011), Brassica (Raman *et al.*, 2012), wheat (Orabi *et al.*, 2014; Terracciano *et al.*, 2013), barley (Lex *et al.*, 2014), carrot (Grzebellus *et al.*, 2014). Table1 gives a descriptive analysis of different markers. Subtracted diversity array (SDA; Li *et al.*, 2006) is advancement over DArT incorporating the benefits of subtractive suppression hybridization (SSH) based microarray, in order to enrich the resulting data with polymorphic genomic DNA sequences.

Therefore, making DArT more appropriate genotyping method for molecular mapping and gene tagging.

The diverse molecular markers utilized in various research have been discussed along with their applications in genetic variability evaluation, genome fingerprinting and population genetic studies. Due to the fast developments in the field of molecular genetics, a variety of techniques have been emerged for gene bank management and identifying genetic variation in germplasm. Efficient molecular markers are codominant, inherit high polymorphism, and exhibit even distribution and frequent occurrence throughout the genome. Other qualities of markers include open access, selectively neutral behavior, low cost, easy and fast assay, transferability between populations and high reproducibility.

In the current scenario as such no molecular markers are available yet that fulfill all the above requirements needed by molecular markers. Therefore careful selection of

molecular marker is needed which combines at least some of these desirable properties. During the last few decades numerous research works has been carried out in field of molecular marker technology. Associated genetics, comparative genomics and allele mining are promising new approaches to obtain insight into the organization and variation of genes in the genome that affects relevant phenotypic traits of the particular individual. The developments and advancement can further be harnessed by combining expertise from several other disciplines like molecular genetics, statistics, bioinformatics etc. Hybridization based DNA markers which were the pioneer markers for genetic diversity studies have been outmoded by the development of more reliable and reproducible PCR based markers.

For the particular study, selection of the most suitable marker will eventually depend on the specific research approach used and degree of polymorphism and resolution needed. In near future with rapid progress in molecular biology, more superior and effective markers may appear, which can significantly accelerate research on plants as well as animal science.

### Abbreviations

RFLPs- Restriction Fragment Length Polymorphism, RAPDs- Random Amplification of Polymorphic DNA, AFLPs- Amplified Fragment Length Polymorphism, ISSRs- Inter Simple Sequence Repeat, SSRs- Simple Sequence Repeat and SNPs- Single Nucleotide Polymorphism

### References

Aitken, K. S., McNeil, M. D., Hermann, S., Bundock, P. C., Kilian, A., Heller-Uszynska, K., Henry, R.J., Li, J., 2014. A comprehensive genetic map of sugarcane that provides enhanced map coverage and integrates high-throughput Diversity Array

Technology (DArT) markers. *BMC Genomics*. Vol 15(152).

Akkaya, M. S., Bhagwat, A. A., Cregan, P. B., 1992. Length polymorphisms of simple sequence repeat DNA in soybean. *Genetics*. 132:1131–1139.

Akopyanz, N., Bukanov, N. O., Westblom, T. U., Kresovic, S., Berg, D., 1992. DNA diversity among clinical isolates of *Helicobacter pylori* detected by PCR-based RAPD fingerprinting. *Nucleic Acids Research*. pp5137-5142.

Alavi-Kia, S. S., Mohammadi S. A., Aharizad S., Moghaddam M., 2008. Analysis of genetic diversity and phylogenetic relationships in crocus genus of Iran using inter-retrotransposon amplified polymorphism. *Biotechnol Biotechnol Eq*. 22: 795-800.

Alavi-Kia, S. S., Mohammadi, S. A., Aharizad, S., Moghaddam, M., 2008. Analysis of genetic diversity and phylogenetic relationships in *Crocus* genus of Iran using inter-retrotransposon amplified polymorphism. *Biotechnol Biotechnol Equip*. 22:795–800

Alexander, E. K., Kennedy, G. C., Baloch, Z. W., Cibas, E. S., Chudova, D., Diggans, J., Friedman, L., Kloos, R. T., LiVolsi, V. A., Mandel, S. J., 2012. Preoperative diagnosis of benign thyroid nodules with indeterminate cytology. *N Engl J Med*. 367:705–715.

Appleby, N., Edwards, D., Batley, J., 2009. New technologies for ultra-high throughput genotyping in plants. *Plant Genomics*. 513:19–39.

Brumfield, R. T., Beerli, P., Nickerson, D. A., Edwards, S.V., 2003. The utility of single nucleotide polymorphisms in inferences of population history. *Trends Ecol. Evol* 18: 249–256.

Caetano-Anollés, G., Bassam, B. J., Gresshoff, P.M., 1991. DNA amplification finger printing using very short arbitrary oligonucleotide primers. *Bio/Technology*. 9: 553–557.

Castelo, A.T., Martins, W., Gao, G.R., 2002. TROLL—Tandem Repeat Occurrence Locator. *Bioinformatics*. 18:634–636.

Chabane, K., Ablett, G. A., Cordeiro, G. M., Valkoun, J., Henry, R.J., 2005. EST versus

- genomic derived microsatellite markers for genotyping wild and cultivated barley. *Genetic Resources and Crop Evolution*. 52, 903-909.
- Cho, Y. G., Ishii, T., Temnykh, S., Chen, X., Lipovich, L., Park, W. D., Ayres, N., Cartinhour, S., McCouch, S. R., 2000. Diversity of microsatellites derived from genomic libraries and GenBank sequences in rice (*Oryza sativa* L.) *Theor Appl Genet*. 100:713–722.
- Coles, B. F., Kadlubar, F. F., 2005. Human Alpha Class Glutathione S- Transferases: Genetic Polymorphism, Expression, and Susceptibility to Disease, *Journal Methods in Enzymology*. Vol 401, Pages 9-42.
- Dalong, G. U. O., Haiqing, Z., Zhengrong, L. U. O., 2006. Genetic relationships of *Diospyros kaki* Thunb. and related species revealed by IRAP and REMAP analysis. *Plant Sci* 170: 528-533.
- Davey, J. W., Hohenlohe, P. A., Etter, P. D., 2011. Genome-wide genetic marker discovery and genotyping using next-generation sequencing. *Nat Rev. Genet*, *Nature Reviews Genetics*. 12(7).
- Elshibli, S., Korpelainen, H., 2008. Microsatellite markers reveal high genetic diversity in date palm (*Phoenix dactylifera* L.) germplasm from Sudan. *J. Genetica*. 134(2):251-60.
- Eujayl, I., Sledge, M., Wang, L., 2004. *Medicago truncatula* EST-SSRs reveal cross-species genetic markers for *Medicago* spp. *Theoretical and Applied Genetics*. 108:414–422.
- Eujayl, I., Sorrells, M. E., Baum, M., Wolters, P., Powell, W., 2002. Isolation of EST-derived microsatellite markers for genotyping the A and B genomes of wheat. *Theoretical and Applied Genetics*. 104, 399-407.
- Foster, J. T., Allan, G. J., Chan, A. P., Rabinowicz, P. D., Ravel, J., Jackson, P. J., Keim, P., 2010. Single nucleotide polymorphisms for assessing genetic diversity in castor bean (*Ricinus communis*). *BMC Plant Biology*. Vol.10, No.13, pp. 1-11, ISSN 1471-2229.
- Gao, L., Tang, J., Li, H., Jia, J., 2003. Analysis of microsatellites in major crops assessed by computational and experimental approaches. *Mol Breed*. 12:245–261.
- Grattapaglia, D., Sederoff, R., 1994. Genetic linkage maps of *Eucalyptus grandis* and *Eucalyptus urophylla* using a pseudo-testcross: mapping strategy and RAPD markers. *Genetics*. 137(4):1121-37.
- Grzebelus, D., Iorizzo, M., Senalik, D., Ellison, S., Cavagnaro, P., Macko-Podgorni, A., Heller-Uszynska, K., Kilian, A., Nothnagel, T., Allender, C., Simon, P.W., Baranski, R., 2014. Diversity, genetic mapping, and signatures of domestication in the carrot (*Daucus carota* L.) genome, as revealed by Diversity Arrays Technology (DART) markers. *Mol Breed*. 33(3): 625–637.
- Gu, W. K., Weeden, N. F., Yu, J., Wallace, D. H., 1995. Large-scale, cost-effective screening of PCR products in marker-assisted selection applications, *Theoretical and Applied Genetics*. vol 91 pp465–470.
- Guo, W., Zhang, T., Shen, X., Yu, J. Z., Kohel, R. J., 2003. Development of SCAR Marker Linked to a Major QTL for High Fiber Strength and Its Usage in Molecular-Marker Assisted Selection in Upland Cotton. *CROP SCIENCE*. VOL. 43.
- Gupta, M., Chyi, Y. S., Romero-Severson, J., Owen, J. L., 1994. Amplification of DNA markers from evolutionary diverse genomes using single primers of simple-sequence repeats. *Theor Appl Genet*. 89: 998-1006.
- Gupta, P. K., Rustgi, S., Sharma, S., Singh, R., Kumar, N., Balyan, H S., 2003. Transferable EST-SSR markers for the study of polymorphism and genetic diversity in bread wheat. *Mol Genet Genom*. 270: 315-323.
- Hu, J., Quiros, C. E., 1991. Identification of broccoli and cauliflower cultivars with RAPD markers. *Plant Cell Reports* 10:505-511.
- Jaccoud, D., Peng, K., Feinsein, D., Kilian, A., 2001. Diversity Arrays: a solid state technology for sequence information independent genotyping. *Nucleic Acids Res*. 29(4): e25.
- Jain, S. M., Brar, D.S., 2009. *Molecular Techniques in Crop Improvement*,

- springer.
- Jonah, P., Bello, L., Lucky, O., Midau, A., Moruppa, S., 2011. Review: the importance of molecular markers in plant breeding programmes. *Glob J Sci Front Res.* 11 eV-vers1.
- Jörg, B., Pieter, V., Martin, K., Francesco, S., Manfred, H., 1995. Combined mapping of AFLP and RFLP markers in barley, *Molecular and General Genetics MGG* volume 249. pp65–73.
- Kalendar, R., Grob, T., Regina, M., Suoniemi, A., Schulman, A., 1998. IRAP and REMAP: two new retrotransposon-based DNA fingerprinting techniques. *Theor Appl Genet.* 98 : 704-711.
- Kalia, R. K., Rai, M. K., Kalia, S., Singh, R., Dhawan, A.K., 2011. Microsatellite markers: an overview of the recent progress in plants. *Euphytica.* 177:309–334.
- Kantety, R. V., Rota, M. L., Matthews, D. E., Sorrells, M.E., 2002. Data mining for simple sequence repeats in expressed sequence tags from barley, maize, rice, sorghum and wheat. *Plant Mol. Biol.* 48: 501–510
- King, R. C., Stansfield W. D., 1990. A dictionary of genetics. 4th ed., Oxford University Press, New York-Oxford, pp 188.
- Knapik, E. W., Goodman, A., Ekker, M., Chevrette, M., Delgado, J., Neuhauss, S., Shimoda, N., Driever, W., Fishman, M.C., and Jacob H. J., 1998. A microsatellite genetic linkage map for zebrafish (*Danio rerio*). *Nature Genetics.* 18: 338-343.
- Koebner, R. M. D., 2004. Marker assisted selection in the cereals: the dream and the reality, in: Gupta P. K., Varshney R. K.(Eds.). *Cereal genomics*, kluwer academic publishers. pp 317-329.
- Kofler, R., Schlotterer, C., Lelley, T., SciRoKo, 2007. a new tool for whole genome microsatellite search and investigation. *Bioinformatics.* 23:1683–1685.
- Kojima, T., Nagaoka, T., Noda, K., Ogihara, Y., 1998. Genetic linkage map of ISSR and RAPD markers in Einkorn wheat in relation to that of RFLP markers. *Theor Appl Genet.* 96: 37–45.
- Konieczny, A. I., Ausubel, F. M., 1993. A procedure for mapping Arabidopsis mutations using co-dominant ecotype-specific PCR-based markers. *Plant J.* 4(2):403-10.
- Korzun V., 2002. Use of molecular markers in cereal breeding. *Cell Mol Biol Lett.* 7(2B) :811-20.
- Kota, R., Varshney, R. K., Thiel, T., Dehmer, K. J., Graner, A., 2001. Generation and comparison of EST-derived SSRs and SNPs in barley (*Hordeum vulgare* L.). *Hereditas.* 135: 145-151.
- Lex, J., Ahlemeyer, J., Friedt, W., Ordon, F., 2014. Genome-wide association studies of agronomic and quality traits in a set of German winter barley (*Hordeum vulgare* L.) cultivars using Diversity Arrays Technology (DArT). *Journal of Applied Genetics.* vol55 pp295–305.
- Li, T. X., Wang, J. K., Bai, Y. F., Lu, Z. H., 2007. Diversity Suppression- Subtractive Hybridization Array for Profiling Genomic DNA Polymorphisms. *journal of integrative plant biology.* vol 48(4) pp 460-467.
- Mammadov, J., Chen, W., Mingus, J., Thompson, S., Kumpatla, S., 2012. Development of versatile gene-based SNP assays in maize (*Zea mays* L.). *Mol Breeding.* 29:779–790 DOI 10.1007/s11032-011-9589-3.
- Martín, J. A., Witzell, J., Blumenstein, K., Rozpedowska, E., Helander, M., Sieber, T. N., 2013. Resistance to Dutch Elm disease reduces presence of xylem endophytic fungi in Elms (*Ulmus* spp.) *PLoS ONE.* 8:e56987 10.1371/journal.pone.0056987.
- Mason, A. S., 2015. SSR Genotyping. In: *Barley. J. Plant Genotyping.* Springer, New York, pp77–89.
- Michalek, W., Weschke, W., Pleissner, K P., Graner, A., 2002. EST analysis in barley defines a unigene set comprising 4,000genes. *Theor Appl Genet.* 104:97–103.
- Newton, C. R., Graham, A., Heptinstall, L. E., Powell, S. J., Summers, C., Kalsheker, N., Smith, J. C., Markham, A. F., 1989. Analysis of any point mutation in DNA. The amplification refractory mutation system (ARMS). *Nucleic Acids Res* 17: 2503–2516.

- Olsen, M., Hood, L., Cantor, C. H., Botstein, D., 1989. A common language for physical mapping of the human genome. *Science* 424 : 1434—1435.
- Orabi, J., JahoorA, J., Backes, G., 2014. Changes in allelic frequency over time in European bread wheat (*Triticum aestivum* L.) varieties revealed using DArT and SSR markers. *Euphytica*. 197(3).
- Orita, M., Suzuki, Y., Sekiya, T., Hayashi, K., 1989. Rapid and sensitive detection of point mutation and DNA polymorphisms using polymerase chain reaction. *Genomics* 5:874-879.
- Pacey-Miller, T., Henry, R. J., 2003. Single-nucleotide polymorphism detection in plants using a single-stranded pyrosequencing protocol with a universal biotinylated primer. *Analytical Biochemistry* 317(2).
- Paran, I., Michelmore, R. W., 1993. Development of reliable PCR-based markers linked to downy mildew resistance genes in lettuce. *Theor Appl Genet.* 85, pp. 985-993.
- Pasquali, M., Dematheis, Lodovica, F., Gullino, M., Garibaldi, A., 2007. Identification of race1 of *Fusarium oxysporum*f. sp. *lactucaeon* lettuce by inter-retrotransposon sequence-characterized amplified region technique. *Phytopathol.* 97: 987-996.
- Pejic, I., Ajmone-Marsan, P., Ajmone-Marsan, P., Morgante, M., 1998. Comparative analysis of genetic similarity among maize inbred lines detected by RFLPs, RAPDs, SSRs, and AFLPs, *December Theoretical and Applied Genetics* 97(8):1248-1255.
- Raman, H., Raman, R., Kilian, A., Detering, F., Long, Y., Edwards, D., Parkin, I.A.P, Sharpe, A. G., Nelson, M. N., Larkan, N., Zou, J., Meng, J., Aslam, M.N., Batley, J., Cowling, W.A., Lydiate, D., 2013. A consensus map of rapeseed (*Brassica napus* L.) based on diversity array technology markers: applications in genetic dissection of qualitative and quantitative traits, *BMC Genomics.* vol 14(277).
- Ramos, A. M., Crooijmans, R. P. M. A., Affara, N. A., Amaral, A. J., Archibald, A. L., Beever, J.E., Bendixen, C., Churcher, C., Clark, R., Dehais, P., Hansen, M. S., Hedegaard, J., Hu, Z. L., 2009. Design of a High Density SNP Genotyping Assay in the Pig Using SNPs Identified and Characterized by Next Generation Sequencing Technology, *PLoS ONE.* 4(8): e6524. <https://doi.org/10.1371/journal.pone.0006524>.
- Rapley, R., Harbron, S., 2001. *Molecular Analysis and Genome Discovery.* Chichester, UK: John Wiley & Sons Ltd.
- Richardson, G. D., Robson, C. N., Lang, S. H., Neal, D. E., Maitland, N. J., Anne, Collins, T., 2004. CD133, a novel marker for human prostatic epithelial stem cells. *Journal of Cell Science* 117, 3539-3545.
- Sansaloni, C. P., Petroli, C. D., Carling, J., Hudson, C. J., Steane, D. A., Myburg, A. A., Grattapaglia, D., Vaillancourt, R. E., Kilian, A., 2010. A high-density Diversity Arrays Technology (DArT) microarray for genome-wide genotyping in Eucalyptus. *Plant Methods.* vol 6(16).
- Savelkoul, P. H., Aarts, H. J., de Haas, J., Dijkshoorn, L., Duim, B., Otsen, M., Rademaker, J. L., Schouls, L., Lenstra, J. A., 1999. Amplified-fragment length polymorphism analysis: the state of an art. *J Clin Microbiol.*37(10):3083-91.
- Sax K., 1932. The association of size differences with seed-coat pattern and pigmentation in *Phaseolus vulgaris*. *Genetics*, 8:552-560.
- Scott, K. D., Eggler, P., Seaton, G., Rossetto, M., Ablett, E. M., Lee, L.S., Henry, R. J., 2000. Analysis of SSRs derived from grape ESTs. *Theor Appl Genet.* 100:723-726.
- Sharma, A., Sharma, P. C., Prakash, S., Bhat, S., Bhat, S., 2007. Identification of AFLP markers linked to the male fertility restorer gene of CMS (*Moricandia arvensis*) *Brassica juncea* and conversion to SCAR marker. *Theoretical and Applied Genetics.* 114(2):385-92.
- Shull, V., Hamer, J. E., 1996. Genetic differentiation in the rice blast fungus revealed by the distribution of the fosbury retrotransposon. *Fungal Genet Biol.* 20: 59–69.
- Supriya, A. I., Senthilvel, S., Nepolean, T., Eshwar, K., Rajaram, V., Shaw, R., Hash,

- C T., Kilian, A., Yadav, R C., Narasu, M L., 2011. Development of a molecular linkage map of pearl millet integrating DArT and SSR markers. *Theor Appl Genet.*123(2):239-50. doi: 10.1007/s00122-011-1580-1.
- Tandem, B. G., 1999. repeats finder: a program to analyze DNA sequences. *Nucleic Acids Res.* 27:573–580.
- Taniguchi, Y., Matsuda, H., Yamada, T., Sugiyama, T., Homma, K., Kaneko, Y., Yamagishi, S., Iwaisaki, H., 2013. Genome-Wide SNP and STR Discovery in the Japanese Crested Ibis and Genetic Diversity among Founders of the Japanese Population. *PLoS One.* 8(8): e72781.
- Temnykh, S., 2001. Computational and experimental analysis of microsatellites in rice (*Oryza sativa* L.): Frequency, length variation, transposon associations, and genetic marker potential. *Genome Res.* 11:1441–1452.
- Terracciano, I., Maccaferri, M., Bassi, F. M., Roberto, T., 2013. Development of COS-SNP and HRM markers for high-throughput and reliable haplotype-based detection of Lr14a in durum wheat (*Triticum durum* Desf.). *Theoretical and Applied Genetics.* 126(4).
- Thakur, S., Singh, P. K., Rathour, R., Variar, M., Prashanthi, S. K., Gopalakrishnan, S., 2014. Genotyping and development of single-nucleotide polymorphism (SNP) markers associated with blast resistance genes in rice using GoldenGate assay. *Mol. Breeding.* 34, 1449–1463. 10.1007/s11032-014-0129-9.
- Thiel, T., Michalek, W., Varshney, R., Graner, A., 2003. Exploiting EST databases for the development and characterization of gene-derived SSR-markers in barley (*Hordeum vulgare* L.). *Theor Appl Genet.*106:411–422.
- Thomas, W. M., Thomas, G. M., Mitchell, E. G., Freedman, J. R., Vilgalys, J., 1993. Hybridization probes for conventional DNA fingerprinting can be used as single primers in the PCR to distinguish strains of *Cryptococcus neoformans*. *Journal of Clinical Microbiology.* 31(9):2274-80.
- Trebbi, D., Maccaferri, M., Heer, P. D., Sørensen, A., Giuliani, S., Salvi, S., Sanguineti, M. C., Massi, A., Gerard van der Vossen, E. A., Tuberosa, R., 2011. High-throughput SNP discovery and genotyping in durum wheat (*Triticum durum* Desf.). *Theoretical and Applied Genetics* volume 123. Pp 555–569.
- Tsuchihashi, Z., Dracopoli, N. C., 2002. Progress in high throughput SNP genotyping methods. *Pharmacogenomics Journal.* 2, 103–110.
- Versalovic, J., Koeuth, T., Lupski, J. R., 1991. Distribution of repetitive DNA sequences in eubacteria and application to fingerprinting of bacterial genomes. *Nucleic Acids Research.* pp. 6823-6831.
- Weising, K., Winter, P., Hüttel, B., Kahl, G., 1997. Microsatellite Markers for Molecular Breeding, *Journal of Crop Production.* Vol 1 (1) pp 113-143.
- Wexelsen, H., 1933. Linkage between quantitative and qualitative characters in barley. *Hereditas (Lund),* 17:323-341.
- Williams, J. G. K., Kubelik, A. R., Livak, K. J., Rafalski, J. A., Tingey, S. V., 1990. DNA polymorphisms amplified by arbitrary primers are useful as genetic markers. *Nucleic Acids Res.* 18: 6531–6535.

### How to cite this article:

Ankita Chauhan, Megha Bhatt and Pushpa Lohani. 2021. The Enormous World of Diverse Molecular Markers. *Int.J.Curr.Microbiol.App.Sci.* 10(04): 757-772.  
doi: <https://doi.org/10.20546/ijcmas.2021.1004.078>