J. Crop Sci. Biotech. 2009 (December) 12 (4) : 169 ~ 181
DOI No. 10.1007/s12892-009-0124-6

REVIEW ARTICLE

Molecular Markers: It's Application in Crop Improvement

Mahipal Singh Kesawat[1], Basanta Kumar Das[2*]

[1] National Research Centre on Plant Biotechnology, Indian Agricultural Research Institute, Pusa Campus, New Delhi-110012
[2] National Institute of Horticultural and Herbal Sciences, RDA, Suwon, Republic of Korea 440706

Received: September 9, 2009 / Accepted: December 8, 2009

Abstract

Over the past few decades, plant genomics research has been studied extensively bringing about a revolution in the field of plant biotechnology. Molecular markers, useful for plant genome analysis, have now become an important tool in crop improvement. The development and use of molecular markers for the detection and exploitation of DNA polymorphism is one of the most significant developments in the field of molecular genetics. The presence of various types of molecular markers, and differences in their principles, methodologies and applications require careful consideration in choosing one or more of such methods. No molecular markers are available yet that fulfill all requirements needed by researchers. In this article we attempt to review most of the available DNA markers that can be routinely employed in various aspects of plant genome analysis such as characterization of genetic variability, genome fingerprinting, genome mapping, gene localization, analysis of genome evolution, population genetics, taxonomy, plant breeding, and diagnostics. The emerging patterns make up a unique feature of the analyzed individual and are currently considered to be the ultimate tool for biological individualization.

Key words: AFLP, DArT, ISSR, polymerase chain reaction, RAPD, RFLP, SNP, SSR

Introduction

Molecular markers are heritable differences in nucleotide sequences of DNA at the corresponding position on homologous chromosome of two different individuals, which follow a simple Mendelian pattern of inheritance. Over the last two decades, the advent of molecular markers has revolutionized the entire scenario of biological sciences. DNA-based molecular markers are a versatile tool in the fields of taxonomy, physiology, embryology, genetic engineering, etc. (Schlotterer 2004). They are no longer looked upon in simple DNA fingerprinting markers in variability studies or in mere forensic tools. Ever since the development of molecular markers, these are constantly being modified to enhance the utility and to bring about automation in the process of genome analysis. The discovery of PCR (polymerase chain reaction) was a landmark in this effort and proved to be a unique process that brought about a new class of DNA profiling markers. This facilitated the development of marker-based gene tags, genetic mapping, map-based cloning of agronomically important genes, genetic diversity studies, phylogenetic analysis, and marker-assisted selection of desirable genotypes etc. (Joshi et al. 2000). Thus, giving new dimensions to breeding and marker-aided selection, that can reduce the time span of developing new and better varieties and the dream of super varieties come true. These DNA markers offer several advantages over traditional phenotypic markers, as they provide data that can be analyzed objectively. The existence of various molecular techniques and differences (Table 1) in their principles and methodologies require careful consideration in choosing one or more of such marker types. This review article deals on the basic principles, requirements, advantages, and disadvantages of the most widely used molecular markers developed during the last two decades of molecular biology research and utilized for various applications in the field of biological science especially in crop improvement.

Types of Molecular Markers

Different types of molecular markers described in the literature are listed in alphabetical order in Table 2. Although some of these marker types are very similar (e.g. ASAP, ASO, and AS-PCR), some synonymous (e.g., ISSR, RAMP, RAM, SPAR,

Basanta Kumar Das (✉)
E-mail: basanta.das@rediffmail.com
Tel: +82 1023348057

Table 1. Overview of the relevant characteristics of some important molecular markers

Charac teristics	RFLP	Mini sat.	RAPD	Micro sat.	ISSR	SSCP	CAPS	SCAR	AFLP
Genomic abundance	High	Medium	High	High	Medium-High	Low	Low	Low	High
Polymorphism level	Medium	High	Medium	High	Medium	Low	Low-Medium	Medium	Medium
Locus specificity	Yes	No/Yes	No	Yes	No	Yes	Yes	Yes	No
Co-dominance of alleles	Yes	No/Yes	No	Yes	No	Yes	Yes	No/Yes	No/Yes
Reproducibility	High	High	Low	High	Medium-High	Medium	High	High	Medium-High
Labor-intensity	High	High	Low	Low	Low	Low-Medium	Low-Medium	Low	Medium
Technical demands	High	High	Low	Low-Medium	Low-Medium	Medium	Low	Low	Medium
Operational costs	High	High	Low	Low	Low-Medium	Low-Medium	Low	Low	Medium
Development costs	Medium-High	Medium-High	Low-Medium	High	Low	High	Medium	Medium	Low
Required DNA Quantity	High	High	Low	Low	Low	Low	Low	Low	Medium
Amenability to automation	No	No	Yes	Yes	Yes	No	Yes	Yes	Yes

AMP-PCR, MP-PCR, and ASSR; Reddy et al. 2002), and some identical (e.g. SSLP, STMS, STR, and SSR), there is still a wide range of techniques available for researchers to choose from. One of the main challenges is, therefore, to associate the purpose(s) of a specific project with the marker types which can be classified into different groups based on the basic criteria:

I. Mode of transmission (bi-parental nuclear inheritance, maternal nuclear inheritance, maternal organelle inheritance, or paternal organelle inheritance).

II. Mode of gene action (dominant or co-dominant markers).

III. Method of analysis (hybridization-based or PCR-based markers).

Ideal properties

Desirable properties for ideal DNA markers should be a highly polymorphic nature, co-dominant inheritance (determination of homozygous and heterozygous states of diploid organisms), frequent occurrence in genome, selective neutral behavior (the DNA sequences of any organism are neutral to environmental conditions or management practices), easy and fast assay. Furthermore, it must have easy access (availability), high reproducibility, and easy exchange of data between laboratories.

It is extremely difficult to find a molecular marker which can meet all the above criteria. However, depending on the type of study to be undertaken, marker systems can be identified that would fulfill at least a few of the above characteristics (Weising et al. 1995). Various types of molecular markers are utilized to evaluate DNA polymorphism generally classified as hybridization-based markers and PCR-based markers. In the former, DNA profiles are visualized by hybridizing the restriction enzyme-digested DNA to a labelled probe, a DNA fragment of known origin or sequence. At the same time, PCR-based markers involve *in vitro* amplification of particular DNA sequences or loci with the help of specifically or arbitrarily chosen oligonucleotide sequences (primers) and a thermostable DNA polymerase enzyme. The amplified fragments are separated electrophoretically and the banding patterns are detected by different methods such as staining and autoradiography. Their application for diverse purposes has opened up a multitude of new possibilities in the field of molecular biology. For simplicity, this review is divided into two parts. The first part is about the general description of most of the available DNA marker types, while the second part includes its application in plant genomics and breeding programs in crop improvement.

Restriction Fragment Length Polymorphism (RFLP): RFLP is the most widely used hybridization-based molecular marker, initially used to identify DNA sequence polymorphisms for genetic mapping of a temperature-sensitive mutation of adenovirus serotypes. They were initially used for human genome mapping (Botstein et al. 1980) but were later adopted for plant genomes. The technique is based on restriction enzymes that reveal a pattern difference between DNA fragment sizes in individual organisms. Although two individuals of the same species have almost identical genomes, they will always differ at a few nucleotides possibly due to point mutation, insertion/deletion, translocation, inversion, and duplication. Some of the differences in DNA sequences at the restriction sites can result in the gain, loss, or relocation of a restriction site.

Table 2. A comprehensive lists of different types of molecular markers with short names

Types of molecular markers	Name
Allele specific associated primers	ASAP
Allele specific oligo	ASO
Allele specific polymerase chain reaction	AS-PCR
Amplified fragment length polymorphism	AFLP
Anchored microsatellite primed PCR	AMP-PCR
Anchored simple sequence repeats	ASSR
Arbitrarily primed PCR	AP-PCR
Cleaved amplified polymorphic sequence	CAPS
Diversity arrays technology	DArT
DNA amplification fingerprinting	DAF
Expressed sequence tags	EST
Inter-simple sequence repeat	ISSR
Microsatellite primed PCR	MP-PCR
Multiplexed allele-specific diagnostic assay	MASDA
Random amplified microsatellite polymorphisms	RAMP
Random amplified microsatellites	RAM
Random amplified polymorphic DNA	RAPD
Restriction fragment length polymorphism	RFLP
Short tandem repeats	STR
Selective amplification of microsatellite polymorphic loci	SAMPL
Sequence characterized amplified regions	SCAR
Sequence tagged microsatellite site	STMS
Sequence tagged site	STS
Simple sequence length polymorphism	SSLP
Simple sequence repeats	SSR
Single nucleotide polymorphism	SNP
Single primer amplification reactions	SPAR
Single stranded conformational polymorphism	SSCP
Variable number tandem repeat	VNTR

In RFLP analysis, restriction enzyme-digested genomic DNA is obtained by gel electrophoresis and then blotted on nitrocellulose membrane. Specific banding patterns are then visualized by hybridization with labelled probe. These probes are mostly species-specific single- or multi-locus probes of about 0.5-3.0 kb in size, obtained from a cDNA library or a genomic library. RFLP markers were used for the first time in the construction of genetic maps. Being co-dominant markers, RFLP can detect coupling phase of DNA molecules, as DNA fragments from all homologous chromosomes are detected. The major strength of RFLP markers are their high reproducibility, co-dominant inheritance, and good transferability between laboratories which provides locus-specific markers that allow synteny (conserved order of genes between related organisms) studies. For this, no sequence information is required and they are relatively easy to score due to large size differences between fragments. Still, there are several limitations for RFLP analysis: it requires the presence of high quantity and quality of DNA (Young et al. 1992). The requirement of radioactive isotope makes the analysis relatively expensive and hazardous. The assay is time-consuming and labor-intensive. RFLPs can be applied in diversity and phylogenetic studies ranging from individuals within populations or species, to closely related species. It is widely used in gene mapping studies because of their high genomic abundance due to the ample availability of different restriction enzymes and random distribution throughout the genome. Basically, it was used to investigate relationships of closely related taxa (Miller and Tanksley 1990), fingerprinting tools for diversity studies, and for studies of hybridization and introgression, as well as studies of gene flow between crops and weeds (Desplanque et al. 1999).

Random Amplified Polymorphic DNA (RAPD):
These are DNA fragments amplified by the PCR using short synthetic primers (generally 10 bp) of random sequence. These oligonucleotides serve as both forward- and reverse-primers and are usually able to amplify fragments from one to ten genomic sites simultaneously. Amplified fragments are usually within the range of 0.5-5 kb in size are separated by agarose gel electrophoresis and polymorphisms can be detected after ethidium bromide staining, as the presence or absence of bands of particular sizes. Polymorphisms are considered to be primarily due to variation in the primer annealing sites, but it can also be generated by length differences in the amplified sequence between primer annealing sites (Williams et al. 1990). The main advantages of RAPDs are: they are less time consuming, easy to assay, and low quantities of template DNA are required, usually 5-50 ng per reaction. Due to the commercial availability of random primers, no sequence data for primer construction is needed. Moreover, RAPDs have a very high genomic abundance and are randomly distributed throughout the genome. The main drawback of RAPDs is low reproducibility hence highly standardized experimental procedures are needed because of its sensitivity to the reaction conditions. RAPD analyses generally require purified, high molecular weight DNA, and precautions are needed to avoid contamination of DNA samples because the short random primers used may amplify DNA fragments in a variety of organisms. Altogether, the inherent problems of reproducibility make RAPDs unsuitable for transference or comparison of results among research teams working in similar species and subjects. As for most other multilocus techniques, RAPD markers are not locus-specific, band profiles cannot be interpreted in terms of loci and alleles (dominance of markers), and similar-sized fragments may not be homologous. The applications of RAPDs ranging from studies at the individual level (e.g. genetic identity) to studies involving closely related species and gene mapping studies to fill gaps not covered by other markers (Williams et al. 1990). Variants of the RAPD technique include Arbitrarily Primed-Polymerase Chain Reaction (AP-PCR) which uses longer arbitrary primers than RAPDs, and DNA Amplification Fingerprinting (DAF) that uses shorter, 5-8 bp primers to generate a larger number of fragments. Multiple Arbitrary Amplicon Profiling (MAAP) is the collective term for techniques using single arbitrary primers.

DNA Amplification Fingerprinting (DAF): Caetano-Anolles et al. (1991) employed single arbitrary primers as short as 5-8 bases to amplify DNA using PCR. In a spectrum of products obtained, simple patterns are useful as genetic markers for mapping, while more complex patterns are useful for DNA fingerprinting. Band patterns are reproducible and can be analyzed using polyacrylamide gel electrophoresis and silver staining. DAF requires careful optimization of parameters; however, it is extremely amenable to automation and fluorescent tagging of primers for early and easy determination of amplified products. DAF profiles can be tailored by employing various modifications such as predigesting of templates. Generally, it is useful in genetic typing and mapping.

Arbitrary Primed-Polymerase Chain Reaction (AP-PCR): This is a special case of RAPD, wherein discrete amplification patterns generated by employing single primers of 10-30 bp long in PCR of genomic DNA (Welsh and McClelland 1990). In the first two cycles, annealing is under non-stringent conditions. The final products are structurally similar to RAPD products. Compared to DAF, this variant of RAPD is not very popular as it involves autoradiography. Recently, it has been simplified by separating the fragments on agarose gels and using ethidium bromide staining for visualization.

Allele-Specific Associated Primers (ASAPs). Obtaining an allele-specific marker requires the sequence of a specific allele (either in homozygous or heterozygous state). The designed specific primers are important for amplification of DNA templates to generate a single fragment at stringent annealing temperatures. These markers tag specific alleles in the genome and are more or less similar to SCARs (Gu et al. 1995).

Randomly Amplified Microsatellite Polymorphisms (RAMPO): In this PCR-based strategy, genomic DNA is first amplified using arbitrary (RAPD) primers. The amplified products are separated electrophoretically and the dried gel is hybridized with microsatellite oligonucleotide probes. The com-

bined advantages of oligonucleotide fingerprinting, RAPD (Williams et al. 1990) and microsatellite-primed PCR (Gupta et al. 1994; Weising et al. 1995) are: quick assay, high sensitivity, high level of variability detection, and the non-requirement of prior DNA sequence information. The technique is successfully employed in the genetic fingerprinting of tomato, kiwi fruit, and closely-related genotypes of *D. bulbifera* (Richardson et al. 1995).

Amplified Fragment Length Polymorphism (AFLP):

This technique combines the power of RFLP with PCR-based technology by ligating primer recognition sequences (adaptors) to the restricted DNA. The key feature of AFLP is its capacity for "genome representation" and the simultaneous screening of representative DNA regions distributed randomly throughout the genome. AFLPs are DNA fragments (80-500 bp) obtained from digestion with restriction enzymes, followed by ligation of oligonucleotide adapters to the digestion products and selective amplification by the PCR. The PCR primers consist of a core sequence (part of the adapter) and a restriction enzyme specific sequence, and one to five selective nucleotides (the higher the number of selective nucleotides, the lower the number of bands obtained per profile). The banding profiles are the result of variations in the restriction sites or in the intervening regions. The technique simultaneously generates fragments from many genomic sites (usually 50-100 fragments per reaction) that are separated by polyacrylamide gel electrophoresis and generally scored as dominant markers (Vos et al. 1995). Selective Fragment Length Amplification (SFLA) and Selective Restriction Fragment Amplification (SRFA) are synonyms sometimes used to refer to AFLPs. A variation of the AFLP technique is known as Selectively Amplified Microsatellite Polymorphic Locus (SAMPL). This technology amplifies microsatellite loci by using a single AFLP primer in combination with a primer complementary to compound microsatellite sequences, which do not require prior cloning and characterization (Morgante and Vogel 1994). SAMPL is considered more applicable for intra-specific than inter-specific studies due to frequent null alleles. The strength of AFLPs includes its high genomic abundance, considerable reproducibility, the generation of many informative bands per reaction, wide range of applications, and the fact that no sequence data for primer construction are required (Saal and Wricke 2002). AFLPs can be analyzed on automatic sequencers, but software problems concerning the scoring of AFLPs are encountered on some systems. Disadvantages include purified high molecular weight DNA, the dominance of alleles, and the possible non-homology of co-migrating fragments belonging to different loci. In addition, due to the high number and different intensity of bands per primer combination, there is the need to adopt certain strict but subjectively determined criteria for acceptance of bands in the analysis. Special attention should be paid to the fact that AFLP bands are not always independent. For example, in the case of an insertion between two restrictions sites, the amplified DNA fragment results in increased band size. This will be interpreted as the loss of a small band and at the same time as the gain of a larger band. It is important for the analysis of genetic relations because it would enhance the weight of non-independent bands compared to the other bands. AFLPs can be applied in studies involving genetic identity, fingerprinting, identification of clones and cultivars, and phylogenetic studies of closely related species. Their high genomic abundance and generally random distribution throughout the genome make AFLPs a widely valued technology for gene mapping studies (Vos et al. 1995).

Minisatellites: Like RFLPs, it involves digestion of genomic DNA with restriction endonucleases, but minisatellites are conceptually a very different class of marker. It consists of chromosomal regions containing tandem repeat units of 10-50 base motifs, flanked by conserved DNA restriction sites. A minisatellite profile consisting of many bands, usually within 4-20 kb size range is generated by using common multilocus probes able to hybridize to minisatellite sequences in different species (Jeffreys et al. 1985). Locus-specific probes can be developed by molecular cloning of DNA restriction fragments, subsequent screening with multilocus minisatellite probes, and isolation of specific fragments. Variation in the number of repeat units, due to unequal crossing over or gene conversion, is considered to be the main cause of length polymorphisms. Due to the high mutation rate in minisatellites, the level of polymorphism is substantial, generally resulting in unique multilocus profiles for different individuals within a population. Minisatellite loci are also often referred to as Variable Number of Tandem Repeats (VNTR) loci (Nakamura et al. 1987). The main advantages of minisatellites are their high level of polymorphism and high reproducibility. Disadvantages of minisatellites are similar to RFLPs due to the high similarity in methodological procedures. If multilocus probes are used, highly informative profiles are generally observed due to the generation of many informative bands per reaction. In that case, band profiles cannot be interpreted in terms of loci and alleles, and similar-sized fragments may be non-homologous. In addition, the random distribution of minisatellites across the genome has been questioned (Schlotterer 2004). The term DNA fingerprinting was introduced for minisatellites, though DNA fingerprinting is now used in a more general way to refer to a DNA-based assay to uniquely identify individuals. Minisatellites are particularly useful in studies of genetic identity, clonal growth and structure, and identification of varieties and cultivars (Jeffreys et al. 1985).

Microsatellites: The genomes of higher organisms contain three types of multiple copies of simple repetitive DNA sequences (satellite DNAs, minisatellites, and microsatellites) arranged in arrays of vastly differing size (Litt and Luty 1989). They are also known as simple sequence repeats (SSRs), short tandem repeats (STRs) or simple sequence length polymorphisms (SSLPs) and are the smallest class of simple repetitive DNA sequences. Microsatellites represent tandem repeats but their repeat motifs are shorter (1-6 bp). If nucleotide sequences in the flanking regions of the microsatellite are known, specific primers (generally 20-25 bp) can be designed to amplify the microsatellite by PCR. Microsatellites and their flanking sequences can be identified by constructing a small-insert genomic library, screening the library with a synthetically

labelled oligonucleotide repeat and sequencing the positive clones. Alternatively, microsatellites may be identified by screening sequence databases for microsatellite sequence motifs from which adjacent primers may then be designed. In addition, primers may be used that have already been designed for closely related species. DNA Polymerase slippage during DNA replication, or slipped strand mispairing and unequal crossing over is considered to be the main cause of variation in the number of repeat units of a microsatellite, can be detected by gel electrophoresis (Matsuoka et al. 2002). The strengths of microsatellites include the codominance of alleles, high genomic abundance, and random distribution throughout the genome (Morgante et al. 2002). In general, microsatellites show a high level of polymorphism, so they are very informative markers. They can be used for population genetics studies and gene mapping, ranging from the individual level (e.g. clone and strain identification) to that of closely related species (Jarne and Lagoda 1996). This significantly decreases the analytical costs. One of the main drawbacks of microsatellites is the high development costs if adequate primer sequences for the species of interest are unavailable, making them difficult to apply to unstudied groups. Mutations in the primer annealing sites may result in the occurrence of null alleles (no amplification of the intended PCR product), which may lead to errors in genotype scoring. The potential presence of null alleles increases with the use of microsatellite primers generated from germplasm unrelated to the species used to generate the microsatellite primers (poor "cross-species amplification"). Null alleles may result in a biased estimate of the allelic and genotypic frequencies and an underestimation of heterozygosity. The underlying mutation model of microsatellites (infinite allele model or stepwise mutation model) is still under debate. Sometimes homoplasy may occur at microsatellite loci due to different forward and backward mutations, resulting in an underestimation of genetic divergence. A very common observation in microsatellite analysis is the appearance of stutter bands occur due to slipped-strand mispairing by DNA during PCR amplification (Freudenreich et al. 1997). These can complicate the interpretation of the band profiles because the size determination of the fragments is more difficult and heterozygotes may be confused with homozygotes. However, the interpretation may be clarified by including appropriate reference genotypes of known band sizes in the experiment. Conversely, its high mutation rate makes it unsuitable for studies at higher taxonomic levels.

Inter Simple Sequence Repeats (ISSR): It involves the amplification of DNA segments present at an amplifiable distance in between two identical microsatellite repeat regions oriented in opposite directions. The technique uses microsatellites as primers in a single primer PCR reaction targeting multiple genomic loci to amplify mainly ISSR of different sizes. The microsatellite repeats used as primers for ISSRs can be di-nucleotide, tri-nucleotide, tetra-nucleotide, or penta-nucleotide. The primers used can be either unanchored (Gupta et al. 1994; Meyer et al. 1993) or more usually anchored at 3` or 5` end with one to four degenerate bases extended into the flanking sequences (Zietkiewicz et al. 1994). ISSRs use longer primers (15-30 mers) as compared to RAPD primers (10 mers), which permit the subsequent use of a high annealing temperature leading to higher stringency. The annealing temperature depends on the GC content of the primer used and ranges from 45 to 65 ^{0}C. The amplified products are usually 200-2000 bp long and can be detected by both agarose and polyacrylamide gel electrophoresis. ISSRs exhibit the specificity of microsatellite markers, but do not need sequence information for primer synthesis enjoying the advantage of random markers (Joshi et al. 2000). The primers are not proprietary and can be synthesized by anyone. The technique is simple, quick, and the use of radioactivity is not essential. ISSR markers are randomly distributed throughout the genome and usually show high polymorphism although the level of polymorphism has been shown to vary with the detection method used. Disadvantages include the possibility of non-homology of similar-sized fragments. Moreover, ISSRs, like RAPDs, may have reproducibility problems. ISSR analysis can be applied in studies involving genetic identity, parentage, clone and strain identification, and taxonomic studies of closely related species as well as in gene mapping studies (Gupta et al. 1994; Zietkiewicz et al. 1994).

Expressed Sequence Tag Markers (EST): Each gene, transcribed into messenger RNA (mRNA) serves as a template for protein synthesis. As mRNA is very unstable outside of cells; scientists use an enzyme called reverse transcriptase to convert mRNA to complementary DNA (cDNA). cDNA production is the reverse of the usual process of transcription in cells because here mRNA acts as a template rather than DNA. cDNA is a stable compound only representing the expressed DNA sequence, generated from mRNA, which represents exons by excising (splicing) introns. Once cDNA representing an expressed gene is isolated, scientists can sequence nucleotides from either the 5' or 3' end to create 5' expressed sequence tags (5' ESTs) and 3' ESTs, respectively (Jongeneel 2000). A 5' EST is obtained from the portion of a transcript (exons) that usually codes for a protein. These regions tend to be conserved across species and do not change much within a gene family. The 3' ESTs are likely to fall within non-coding (introns) or untranslated regions (UTRs) and tend to exhibit less cross-species conservation than coding sequences. The identification of ESTs has proceeded rapidly, with over six million ESTs now available in computerized databases. ESTs were originally intended as a way to identify gene transcripts, instrumental in gene discovery, for obtaining data on gene expression and regulation, sequence determination, and for developing highly valuable molecular markers such as EST-based RFLPs, SSRs, SNPs, and CAPS. ESTs are also used for designing probes for DNA microarrays to determine the gene expression, construction of high-density genetic linkage maps and physical maps. Often EST-based RFLP markers allow comparative mapping across different species, because sequence conservation is high in the coding regions. Hence, marker development and map-based cloning in one species will profit directly from data, which are available in any other species. Pattern-finding programs can be employed to

identify SSRs in ESTs. A modest 1 to 5% of the ESTs in various plant species are found to have SSRs of suitable length (20 bp or more) for marker development (Kantety et al. 2002). It is possible to find a large number of these SSRs in an organism for which a great number of ESTs are generated. Kantety et al. (2002) searched 262,631 ESTs from five different grass (rice, maize, wheat, barley, and sorghum) databases for SSRs (di-, tri-, and tetranucleotide motifs with a minimum repeat length of 18 bp) and found that 3.2% of ESTs contained SSRs. EST-SSRs are generally anchored within more conserved transcribed regions across species than those from the untranscribed regions and hence expected to be more transferable to closely related genera. EST-SSRs also have a higher probability of being functionally associated with differences in gene expression than the genomic SSRs. Most of the large scale, multi species *in silico* mining efforts for developing EST-SSRs seem to have focused primarily on monocotyledonous crops (Kantety et al. 2002), although ESTs of a few dicot species have been explored for SSR mining (Varshney et al. 2005). Two strategies have been employed for SNP development based on ESTs. One strategy uses ESTs from the 3'-end of cDNA clones, which consists mainly of 3'-UTRs, to maximize the chance of finding sequence variations. Primer pairs can be derived from the EST sequences and the amplification of corresponding regions from several genotypes followed by sequence comparison may reveal SNPs. Alternatively, one can use clusters of ESTs which contain sequences from different cultivars and identify potential SNPs computationally. Basically, EST markers are useful for cloning of specific genes of interest, full genome sequencing, and mapping of functional genes in various related organisms. Thus, it is more popular in identifying new and active genes. In rice, *Arabidopsis* etc., thousands of functional cDNA clones are being converted in to EST markers.

Cleaved Amplified Polymorphic Sequence (CAPS): CAPS are DNA fragments amplified by PCR using specific 20-25 bp primers, followed by digestion of the PCR products with a restriction enzyme. Subsequently, length polymorphisms resulting from variation in the occurrence of restriction sites are identified by gel electrophoresis of the digested products (Akopyanz et al. 1992; Konieczny and Ausubel 1993). Advantages of CAPS include the involvement of PCR requiring only low quantities of template DNA (50-100 ng per reaction), the codominance of alleles and the high reproducibility. Compared to RFLPs, CAPS analysis does not require the laborious and technically demanding steps of Southern blot hybridization and radioactive detection procedures. CAPS polymorphisms are more difficult to find because of limited size of the amplified fragments (300-1800 bp) and sequence data needed to design the PCR primers. CAPS are also referred as PCR-Restriction Fragment Length Polymorphism (PCR-RFLP) and predominantly applied in gene mapping studies (Akopyanz et al. 1992; Konieczny and Ausubel 1993).

Sequence Characterized Amplified Region (SCAR): SCAR marker, a genomic DNA fragment is identified by PCR amplification using a pair of specific oligonucleotide primers (Paran and Michelmore 1993). SCARs are derived by cloning and sequencing the two ends of RAPD markers that appeared to be diagnostic for specific purposes (e.g. a RAPD band present in disease-resistant lines but absent in susceptible lines), and longer primers designed (22-24 nucleotide bases long) for specific amplification of a particular locus. Due to the use of longer PCR primers, SCARs do not have low reproducibility as generally encountered in RAPDs. Compared to arbitrary primers, SCARs exhibit several advantages in mapping studies (codominant SCARs are informative for genetic mapping than dominant RAPDs), map-based cloning as they can be used to screen pooled genomic libraries by PCR, physical mapping, locus specificity, etc. Due to the use of PCR, only low quantities of template DNA are required (10-100 ng per reaction), and they are quick and easy to use. The need of sequence data to design the PCR primers is the main limitation. Applications of SCARs are in gene mapping studies, marker assisted selection (Paran and Michelmore 1993), and comparative mapping or homology studies among related species, thus making it an extremely adaptable concept in the near future.

Single-Strand Conformation Polymorphism (SSCP): This is a powerful and rapid technique for gene analysis particularly for detection of point mutations and typing of DNA polymorphism. It can identify heterozygosity of DNA fragments of the same molecular weight and can detect changes of a few nucleotide bases as the mobility of the single-stranded DNA changes even in its GC content. Generally, SSCPs are DNA fragments of about 200-800 bp amplified by PCR using specific primers of 20-25 bp. Gel electrophoresis of single-strand DNA is used to detect nucleotide sequence variation among the amplified fragments. The method is based on the fact that the electrophoretic mobility of single-strand DNA depends on the secondary structure (conformation) of the molecule, which changed significantly due to mutation. Thus, SSCP provides a method to detect nucleotide variation among DNA samples without sequence reactions (Orita et al. 1989). In SSCP, the amplified DNA is first denatured, and then subject to non-denaturing gel electrophoresis. Related techniques to SSCP are Denaturing Gradient Gel Electrophoresis (DGGE) based on double-stranded DNA, converted to single-stranded DNA in an increasingly denaturing physical environment during gel electrophoresis and Thermal Gradient Gel Electrophoresis (TGGE) which uses temperature gradients to denature double stranded DNA during electrophoresis. Advantages of SSCP include the codominance of alleles and the low quantities of template DNA required (10-100 ng per reaction) as the technique is PCR-based. However the drawbacks are the need of sequence data to design PCR primers, the necessity of highly standardized electrophoretic conditions in order to obtain reproducible results and sometimes mutations may remain undetected and hence absence of mutation cannot be proven. SSCP, a potential tool for high throughput DNA polymorphism, useful in the detection of heritable human diseases, detect mutations in genes using gene sequence information for primer construction. In plants, it is not well developed although its application in discriminating progenies can be

exploited once suitable primers are designed for agronomically important traits.

Sequence Tagged Site (STS): STS, a short, unique sequence was first developed by Olsen et al., (1989) as DNA landmarks in the physical mapping of the human genome, and later adopted in plants. Two or more clones containing the same STS must overlap and the overlap must include STS. Any clone can be sequenced and used as STS provided that it contains a unique sequence. In plants, STS is characterized by a pair of PCR primers designed by sequencing either an RFLP probe representing a mapped low copy number sequence or AFLP fragments. Although conversion of AFLP markers into STS markers is a technical challenge and often frustrating in polyploids such as hexaploid wheat, it has been successful in several crops (Guo et al. 2003). The primers designed on the basis of RAPD are also referred to as STSs (sometimes), more appropriately called as SCARs. STS markers are codominant, highly reproducible, suitable for high throughput, automation, and technically simple to use.

Single Nucleotide Polymorphism (SNP): Single nucleotide polymorphisms (SNPs) and insertions and deletions (InDels) are highly abundant and distributed throughout the genome in various species including plants. They are an attractive tool for mapping, marker-assisted breeding, and map-based cloning. By comparing sequences from a japonica rice cultivar with an indica cultivar, for example, Yu et al. (2002) identified on average one SNP for every 170 bp and one InDel every 540 bp. As suggested by the acronym, a SNP marker is just a single-base change in the DNA sequence at which different nucleotides occur in different individuals of populations. Hence, in contrast to all previous methods, allele discrimination cannot be based on size differences on gel. Over the past years, a number of different SNP genotyping methods have been developed based on various methods of allelic discrimination and detection platforms. Recently, the majority of SNP genotyping assays can be classified into one of four groups based on molecular mechanism: allele specific hybridization, primer extension, oligonucleotide ligation, and invasive cleavage (Sobrino et al. 2005). Allele-specific hybridization, also known as allele-specific oligonucleotide hybridization (ASO), is based on a distinction between two DNA targets differing at one nucleotide position by hybridization. Two allele-specific probes are designed, usually with the polymorphic base in a central position in the probe sequence. Under optimized assay conditions, only the perfectly matched, probe-target hybrids are stable, and hybrids with one-base mismatch become unstable. Most hybridization techniques are derived from the Dot Blot, in which DNA to be tested (either genomic, cDNA or a PCR reaction) fixed on a membrane and is hybridized with a probe, usually an oligonucleotide. However, hybridization techniques are error prone and need careful probe designing and hybridization protocols. The latest improvement techniques, use of DNA chips (collection of microscopic DNA spots attached to a solid surface, such as glass, plastic or silicon chips) on which the probes are directly synthesized using a parallel procedure involving masks and photolithography (Pease et al. 1994). To take full advantage of new ASO probe formats for SNP typing, it is necessary to use detection methods which provide high accuracy, sensitivity, and results. Primer extension is based on the ability of DNA polymerase to incorporate specific deoxy-ribonucleotides complementary to the sequence of the template DNA. There are variations in the primer extension reaction, which can be divided into three main types of reactions: a) the mini-sequencing reaction or single nucleotide primer extension where the polymorphic base is determined by the addition of the dideoxynucleotide triphosphate (ddNTP) complementary to the base interrogated by a DNA polymerase, b) The allele-specific extension where the DNA polymerase amplifies only if the primers have a perfect match with the template and, c) pyrosequencing.

There are several detection methods for analyzing the products of each type of allelic discrimination reaction: gel electrophoresis, fluorescence resonance energy transfer (FRET), fluorescence polarization, arrays or chips, luminescence, mass spectrophotometry (Matrix-assisted laser desorption/ionization time-of-flight mass spectrometry or MALDI-TOF), chromatography, etc., can be followed. However, suitable technology in terms of sensitivity, reproducibility, accuracy, capability of multiplexing for high result, cost effectiveness in terms of initial investment for equipment and cost per data-point, flexibility of the technology for uses other than SNP discovery, and time-consumption for analysis should be taken into early consideration. To study large sets of samples, the use of primer extension techniques analyzed by MALDI-TOF mass spectrometry holds high promise in terms of automation, accuracy, and cost effectiveness (Tost and Gut 2002). Mass spectrometry-based methods for SNP genotyping has continuously improved and matrix-assisted laser desorption/ionization mass spectrometry (MALDIMS) is now one of the most automated and efficient detection platforms, is price competitive, and delivers results of the highest accuracy and reliability (Tost and Gut 2002).

Diversity Arrays Technology (DArT): One of the recently developed molecular techniques used especially in rice, barley, eucalyptus, Arabidopsis, cassava, wheat, and pigeon-pea. It's an open source (non-exclusive) technology with a great potential for genetic diversity and mapping studies in a number of 'orphan' crops relevant in Third World countries. DArT is a microarray hybridization-based technique, and enables the simultaneous typing of several hundred polymorphic loci spread over the genome (Jaccoud et al. 2001). For each individual DNA sample typed, genomic representations are prepared by restriction enzyme (*Pst*I and *Taq*I) digestion of genomic DNA, followed by ligation of restriction fragments to adapters. The genome complexity is reduced by PCR using primers with complementary sequences to the adapter and selective overhangs. The fragments from representations are cloned and amplified using vector-specific primers, purified, and arrayed onto a solid support (microarray) resulting in a “discovery array.” Labeled genomic representations prepared from the individual genomes included in the pool are hybridized to the discovery array

(Jaccoud et al. 2001). Polymorphic clones (DArT markers) show variable hybridization signal intensities for different individuals. These clones are subsequently assembled into a "genotyping array" for routine genotyping. It does not need prior sequence information for the species to be studied, thus are applicable to all species regardless of how much DNA sequence information exists for that species. It has high output, is highly reproducible, and cost effective. The genetic scope of analysis is easily expandable as it is not covered by exclusive patent rights but on the contrary, is open-source (designed for open use and shared improvement). However, it has some limitations as microarray-based technique require dedicated software (DArTsoft and DArTdb), laboratory facilities, high investment, and skilled manpower. DArT assays represent the presence (or amount) of a specific DNA fragment. Hence, DArT markers are primarily dominant (present or absent) or different in intensity, which limits their value in some applications.

Applications of molecular markers in crop improvement

Molecular markers are used as tools for a large number of applications ranging from localization of a gene to improvement of plant varieties by marker-assisted selection. Especially, applications in phylogenetic analysis add a new dimension to evolutionary theories. In the modern era of research on molecular biology, molecular marker-assisted section is a major achievement for scientists and breeders.

Mapping and tagging of genes: Generating tools for marker-assisted selection in plant breeding

Manipulation of large number of genes are often required for plant improvements, either by natural selection or by the efforts of breeders, always depend on creating, evaluating, and selecting the right combination of alleles (Flavell 1995). Tracing out the valuable alleles in a segregating population and markers (once mapped), enables dissection of the complex traits into component's genetic units more precisely provides new ideas for efficient breeding programs. The very first genome map in plants was reported in maize (Gardiner et al. 1993), followed by rice (McCouch et al. 1988), *Arabidopsis* (Nam et al. 1989), etc., using RFLP markers. Maps for several other crops like potato, barley, banana, and members of Brassicaceae have been constructed (Winter and Kahl 1995). Microsatellite markers, especially STMS markers, have been found to be extremely useful in genome mapping. Based on Mendelian inheritance, construction of index maps can be easily used to provide an anchor or reference point for specific regions of the genome. Initially, microsatellites maps in plants were assembled by Zhao and Kochert (1992) in rice using (GGC)*n*, followed by mapping of (GA)*n* and (GT)*n* by Tanksley et al. 1995, and (GA/AG)*n*, (ATC) 10 and (ATT) 14, in rice. Similar to microsatellites, looking to the pattern of variation, generated by retrotransposons, it is now proposed that apart from genetic variability, these markers are ideal for integrating genetic maps (Ellis et al. 1998).

Once mapped, these markers can be efficiently employed in tagging several individual traits that are extremely important for a breeding program like yield, disease resistance, stress tolerance, seed quality, etc. A large number of monogenic and polygenic loci for various traits have been identified in many plants and are currently being exploited by breeders and molecular biologists together, so as to make the dream of marker-assisted selection come true. Tagging of useful genes like the ones responsible for conferring resistance to plant pathogens, synthesis of plant hormones, drought tolerance, and a variety of other important developmental pathway genes, is a major target. Such tagged genes can also be used for detecting the presence of useful genes in the new genotypes generated in a hybrid program or by other methods like transformation. The first reports on gene tagging were from tomato (Williamson et al., 1994), availing the means for identification of markers linked to genes involved in several traits like water-use-efficiency (Martin et al. 1989), resistance to *Fusarium oxysporum* (12 genes) (Sarfatti et al. 1989), leaf rust resistance genes *LR9* and *LR24* (Schachermayr et al. 1995), and root knot nematodes (*Meliodogyne* sp. the *mi* gene). Xiao et al. (1998) have shown the utility of RFLP markers in identifying the trait improving QTL alleles from wild rice relative *O. rufipogon*. Allele-specific associated primers have also exhibited their utility in genotyping of allelic variants of loci that result from both size differences and point mutations. Some of the genuine examples of this are the waxy gene locus in maize (Shattuck-Eidens et al. 1991), the *Glu D1* complex locus associated with bread making quality in wheat (D'Ovidio and Anderson 1994), the *Lr1* leaf rust resistance locus in wheat (Feuillet et al. 1995), the *Gro1* and *H1* alleles conferring resistance to the root cyst nematode *Globodera rostochiensis* in potato (Niewohner et al. 1995), and allele-specific amplification of polymorphic sites for detection of powdery mildew resistance loci in cereals. A number of other traits have been tagged using ASAPs in tomato, lettuce, etc. (Paran et al. 1991).

STMS markers are used as potential diagnostic markers for important traits in plant breeding programs, e.g. (AT) 15 repeat located within a soybean heat shock protein gene which is about 0.5 cM from (*Rsv*) a gene conferring resistance to soybean mosaic virus (Yu et al. 1994). Several resistance genes including peanut mottle virus (*Rpv*), phytopthora (*Rps3*), and Japanese root knot nematode are clustered in this region of the soybean genome. Similar to RFLPs, STMS, and ASAPs, arbitrary markers, RAPDs have also played important role in saturation of the genetic linkage maps and gene tagging. Its use in mapping is especially important in systems, where RFLPs failed to reveal much polymorphism. One of the first uses of RAPD markers in saturation of genetic maps was reported by Williams et al. (1991). They have proven utility in construction of linkage maps among species and the inherent difficulty in producing F_2 segregating populations that have large genome size, e.g. conifers (Chaparoo et al. 1992). RAPD markers, near isogenic lines can be converted into SCARs and used as diagnostic markers. SCAR/STS markers linked to the translocated segment on chr. 4 AL of bread wheat carrying the *Lr28* gene has been tagged by Naik et al.1998. Recently, ISSRs, also an arbitrary marker have been employed as a reliable tool for gene tagging. An ISSR marker (AG) *8YC* has been found to be linked closely (3.7 ± 1.1

cM) to the rice nuclear restorer gene, *RF1* for fertility. *RF1* is essential for hybrid rice production and this marker would be useful not only for breeding both restorer and maintainer lines, but also for the purity management of hybrid rice seeds (Akagi et al. 1996). Similarly ISSR marker $(AC)_8$ YT has been found to be linked to the gene for resistance to fusarium wilt race 4 in repulsion at a distance of 5.2 cM in chickpea (Ratnaparkhe et al. 1998). Apart from mapping and tagging of genes, an important utility of RFLP markers has been observed in detecting gene introgression in a backcross breeding and synteny mapping among closely related species (Gale and Devos 1998). Similar utility of STMS markers has been observed for reliable preselection in a marker assisted selection backcross scheme. Besides specific markers, DAMD-based DNA fingerprinting in wheat has also been useful for monitoring backcross-mediated genome introgression in hexaploid wheat (Somers et al. 1996).

Phylogeny and evolution

Early theories of evolution were mostly based on morphological and geographical variations between organisms. However, the recent techniques of molecular biology becoming more and more evident in providing detailed information about the genetic structure (Slatkin 1987). RFLP, DNA sequencing and a number of PCR-based markers are being used extensively for reconstructing phylogenies of various species. The techniques are speculated to provide path-breaking information regarding the fine time-scale on which closely related species have diverged and what sort of genetic variations are associated with species formation. Furthermore, these studies hold a great promise for revealing more about the pattern of genetic variation within species (Avise 1994). Efforts are now being made to study the genetic variation in plants, so as to understand their evolution from wild progenitors and to classify them into appropriate groups. The taxonomic classification is an essential first step to determine whether any germplasm is a part of the primary, secondary, or tertiary gene pool of the system concerned. This is more important where morphological markers can prove to be inaccurate and misleading. For example, the lines Azucena and PR 304 which have been classified as indica using morphological characters behave like japonicas in crossing studies. These samples are however, clearly revealed to be japonicas, when analyzed by RAPD markers (Virk et al. 1997).

RFLPs have been used in evolutionary studies for deducing the relationship between the hexaploid genome of bread wheat and its ancestors (Gill et al. 1991). Similarly a number of transposon elements like *tos1-1*, *tos2-1*, and *tos3-1* retrotransposons have been used to detect the genetic differences between different species of rice and even between ecotypes of cultivated rice, wherein they were found to distinguish between the cultivars of Asian and African rice, *O. sativa* and *O. Glaberrima* (Fukuchi et al. 1993). Retro-element *Wis-2* has been found to detect genomic variation within individual plants of wheat variety and also within and between varieties of wheat. This element has also been found to occur in the genomes of other grasses like barley, rye, oats, *Aegilops* species, etc., and indicating common ancestral elements in grasses. Though RFLPs, microsatellites, minisatellites, and transposons are useful for carrying out genetic variability analysis, the trend is now shifting towards the use of PCR-based markers. Specific markers are preferred over arbitrary primers. However, arbitrary primers are found to be the markers of choice in the analysis of complex genomes. Like wheat, where genetic variation is extremely difficult to observe, DAF is used as a new source of molecular markers in fingerprinting of bread wheat (Sen et al. 1997).

Specific markers like STMS (sequence-tagged microsatellite markers), ALPs (Amplicon length polymorphisms), or STS markers have proved to be extremely valuable in the analysis of gene pool variation of crops during the process of cultivar development and classification of germplasm. These markers are extremely sensitive and can detect allelic variability during cultivar development (Yang et al. 1994). STS markers specific to chloroplast or mitochondrial DNA have been useful in providing seed- and pollen-specific markers. It can be utilized for the detection of length variation at multiple, physically linked sites and for haplotype data and thus genotypically unique individual plants (Wu et al. 1998). Also, a comparison of patterns of variability detected with bi-parentally (nuclear) and uniparentally (organellar) transmitted markers can provide complementary information to population and evolutionary biologists. Excellent examples are the Poly A mononucleotide repeats in maize (Powell et al. 1995), Poly (TA/AT) dinucleotide repeats found in liverworts, maize, pea and non-photosynthetic green plant *Epifagus virginiana* etc. Though all these marker types provide valuable information regarding the evolution and phylogeny of various species being studied in any given set of samples, the trend is now shifting towards the use of ESTs (expressed sequence tags) for such analysis. This may be because in such studies one actually looks at the evolution of functional genes of target or goal (Mason-Gamer et al. 1998).

Diversity analysis of exotic germplasm

Characterization of germplasm resources available in crop species is essential for estimation of genetic diversity, identification, elimination of duplicates, and protection of elite genotype. Following domestication, genetic variation in crop plants has continued to narrow due to continuous selection process for specific traits, i.e. yield, thus rendering them more vulnerable to diseases and jeopardizing the potential for sustained genetic improvement over a long term. This risk was brought sharply into focus in 1970 with the outbreak of Southern corn leaf blight which drastically reduced corn yield in the USA and was attributed to the extensive use of a single, genetic male sterility factor which was unfortunately linked to the disease susceptibility. Thus, it is extremely important to study the genetic composition of the germplasm of existing modern-day cultivars in comparison with their ancestors and related species. This will not only provide information on their phylogenetic relationship but will also indicate a chance of finding new and useful genes, as the accessions with most distinct DNA profiles are likely to contain a greater number of novel alleles. DNA profiling to make such sampling decisions is now underway in most crops. The exotic germplasm for breeding is selected on the basis of certain char-

acteristic features such as (a) the exotic germplasm must possess a significant number of unique DNA polymorphisms (throughout the genome) relative to the modern-day cultivars and (b) each exotic germplasm has to be genetically dissimilar (on the basis of DNA profiling (Brown and Kresovich 1996). This is necessary to understand the genetic variations between the existing cultivars, the cultivars in comparison with their wild progenitors, and a number of wilds that still exist in nature.

Many DNA markers both specific as well as arbitrary have been used so far, for DNA fingerprinting of various classes of germplasm (Virk et al. 1997). Further, STMS markers throw light on the domestication process and are useful criteria for enriching the gene pool of crop plants and determine how efficient plant breeders accessing pre-existing forms of variation. AFLP has gained popularity for the study of genetic polymorphism especially in species where polymorphism is extremely rare. Pakniyat et al. 1997 used AFLP for studying variation in wild barley with reference to salt tolerance and associated eco-geography and a number of reports are coming up each day for different systems. Similarly, ISSR markers are used for diversity analysis of pine, rice, and also in wheat (Blair et al. 1999). These studies helped in the classification of existing biodiversity among plants, which can be further exploited in wild gene introgression programs.

Genotyping of cultivars

The repetitive and arbitrary DNA markers are of choice in genotyping of cultivars. Microsatellites like $(CT)_{10}$, $(GAA)_5$, $(AAGG)_4$, $(AAT)_6$, $(GATA)_4$, $(CAC)_5$ (Gupta et al. 1994), and minisatellites (Ramakrishna et al. 1995) have been employed in DNA fingerprinting for the detection of genetic variation, cultivar identification, and genotyping (Sant et al. 1999; Yang et al. 1994). This information is useful for quantification of genetic diversity, characterization of accessions in plant germplasm collections, and taxonomic studies. Microsatellites have been useful for the generation of STMS markers, revealing polymorphisms within closely related cultivars (Morjane et al. 1994). In plants, the first application of microsatellites was for cultivar identification and was later used to genotype unequivocally diverse materials like rice, wheat, grapevine, soybean, etc. This is important especially for protection of proprietary germplasm. Similarly microsatellite markers have also been advantageous in pedigree analysis as they represent a single locus. The multi-allelism of these markers facilitates comparative allelic variability detection reliably across a wide range of germplasm and allows individuals to be ubiquitously genotyped, so that gene flow and paternity can be established. One of the most recent applications of these markers is in sex identification of dioeciously plants where microsatellite probe $(GATA)_4$ is found to reveal sex-specific differences in Southern analysis. It can be used as a diagnostic marker where male and female plants do not show any sex-specific morphological differences until flowering. Similarly, Di Stilio et al. (1998) have identified a randomly-amplified (RAPD) DNA marker for pseudo-autosomal plant sex chromosome in *Silene dioica* (L.).

Conclusion

Due to the rapid developments in the field of molecular genetics, a variety of techniques have been emerged to analyze genetic variation in germplasm and gene bank management especially during the last few decades,. The desirable properties of molecular markers are high polymorphism, codominant inheritance, frequent occurrence and even distribution throughout the genome, selectively neutral behavior, open access, easy and fast assay, low cost, high reproducibility and transferability between laboratories, populations, and/or species. No molecular markers are available yet that fulfill all these requirements, so it needs careful selection of molecular marker, which combines at least some of these desirable properties.

DArT is a recent technique and needs further analysis in various species. SCAR and STS markers would be developed by sequencing fragments associated with economically important traits. SNPs seem to be very exciting markers but expensive, so they are unlikely to be taken up by the national agricultural systems and universities in developing countries. The use of EST and EST-based markers, such as EST-SSR, CAPS, and EST-RFLP, are applicable only for species which have been extensively sequenced before. Therefore, RFLP, SSR, RAPD, AFLP, and ISSR are the only markers that could be used for a wide range of applications in plants. The increasing availability of sequence data for more crops via whole-genome sequencing projects, and access to EST-databases, enables the development of markers targeting coding regions of the genome or even specific genes. Technological developments continue to increase by large-scale genotyping of genetic resources. Allele mining, associated genetics and comparative genomics are promising new approaches to obtain insight the organization and variation of genes that affect relevant phenotypic traits. These developments exploited by combining expertise from several disciplines, including molecular genetics, statistics, bioinformatics, etc.

References

Akagi H, Yokozeki Y, Inagaki A, Nakamura A, Fujimura T. 1996. A codominant DNA marker closely linked to the rice nuclear restorer gene, *Rf-1*, identified with inter-SSR finger printing. Genome 39: 1205-1209

Akopyanz N, Bukanov NO, Westblom TU, Berg DE. 1992. PCR-based RFLP analysis of DNA sequence diversity in the gastric pathogen *Helicobacter pylori*. Nucleic Acid Res. 20: 6221-6225

Avise JC. 1994. Molecular Markers, Natural History and Evolution, Chapman and Hall

Blair MW, Panaud O, McCouch SR. 1999. Inter-simple sequence repeats (ISSR) amplification for analysis of microsatellite motif frequency and fingerprinting in rice (*Oryza sativa* L.). Theor. Appl. Genet. 98: 780-792

Botstein D, White RL, Skolnick M, Davis RW. 1980. Construction of a genetic linkage map in man using restriction fragment

length polymorphisms. Am. J. Hum. Genet. 32: 314 331
Brown SM, Kresovich S. 1996. In AH Paterson, ed, Genome Mapping in Plants Clandes, New York, pp. 85-93
Caetano-Anolles G, Bassam BJ, Gresshoff PM. 1991. DNA amplification fingerprinting using very short arbitrary oligonucleotide primers. Biotechnol. 9: 553-557
Chaparoo J, Werner D, O'Malley D, Sederoff R. 1992. Targeted mapping and linkage analysis in peach. Plant Genome 21: 9-11
D'Ovidio R, Anderson OD. 1994. PCR analysis to distinguish between allele of a member of multi-gene family correlated with wheat bread making quality. Theor. Appl. Genet. 88: 759-763
Desplanque B, Boudry P, Broomberg K, Saumitou-Laprade P, Cuguen J, van Dijk H. 1999. Genetic diversity and gene flow between wild, cultivated and weedy forms of *Beta vulgaris* L. (Chenopodiaceae), assessed by RFLP and microsatellite markers. Theor. Appl. Genet. 98: 1194-1201
Di Stilio VS, Kesseli RV, Mulcahy DL. 1998. A pseudoautosomal random amplified polymorphic DNA marker for the sex chromosomes of *Silene dioica*. Genetics 149: 2057-2062
Ellis THN, Poyser SJ, Knox MR, Vershinin AV, Ambrose MJ. 1998. Ty1-copia class retrotransposon insertion site polymorphism for linkage and diversity analysis in pea. Mol. Gen. Genet. 260: 9-91
Flavell RB. 1995. Plant biotechnology R & D - the next ten years. Tibtech. 13: 313-319
Freudenreich CH, Stavenhagen JB, Zakian VA. 1997. Stability of a CTG:CAG trinucleotide repeat in yeast is dependent on its orientation in the Genome. Mol. Cell Biol. 4: 2090-2098
Fukuchi A, Kikuchi F, Hirochika H. 1993. DNA fingerprinting of cultivated rice with rice retrotransposon probes. Jpn. J. Genet. 68: 195-20
Feuillet C, Messmer M, Schachermayr G, Keller B. 1995. Genetic and physical characterization of LR 1 leaf rust resistance locus in wheat (Triticum aestivum). Mol. Gen. Genet. 248: 553-562
Gale MD, Devos KM. 1998. Comparative Genetics in the Grasses. Proc. Natl. Acad. Sci. USA 95: 1971-1974
Gardiner JM, Coe EH, Melia-Hancock S, Hoisington DA, Chao S. 1993. Map in maize using an immortalized F2 population. Genetics 134: 917-930
Gill KS, Lubbers EL, Gill BS, Raupp WJ, Cox TS. 1991. A genetic linkage map of *Triticum tauschii* (DD) and its relationship to the D genome of bread wheat (AABBDD). Genome 34: 362-374
Gu WK, Weeden NF, Yu J, Wallace DH. 1995. Large-scale, cost effective screening of PCR products in marker-assisted selection applications. Theor. Appl. Genet. 91: 465-470
Guo PG, Bai GH, Shaner GE. 2003. AFLP and STS tagging of a major QTL for Fusarium head blight resistance in wheat. Theor. Appl. Genet. 106: 1011-1017
Gupta M, Chyi YS, Romero-Severson J, Owen JL. 1994. Amplification of DNA markers from evolutionarily diverse genomes using single primers of simple-sequence repeats. Theor. Appl. Genet. 89: 998-1006
Jaccoud D, Peng K, Feinstein D, Kilian A. 2001. Diversity arrays: a solid state technology for sequence information independent genotyping. Nucleic Acids Res. 29: 25-29
Jarne P, Lagoda PJL. 1996. Microsatellites, from molecules to populations and back. Trends Ecol. Evol. 11: 424-429
Jeffreys AJ, Wilson V, Thein SL. 1985. Hypervariable "minisatellite" regions in human DNA. Nature 314: 67-73
Jongeneel CV. 2000. Searching the expressed sequence tag (EST) databases: Panning for genes. Brief. Bioinformatics 1: 76-92
Joshi SP, Gupta VS, Aggarwal RK, Ranjekar PK, Brar DS. 2000. Genetic diversity and phylogenetic relationship as revealed by inter simple sequence repeat (ISSR) polymorphism in the genus *Oryza*. Theor. Appl. Genet. 100: 1311-1320
Kantety RV, Rota ML, Matthews DE, Sorrells ME. 2002. Data mining for simple-sequence repeats in expressed sequence tags from barley, maize, rice, sorghum, and wheat. Plant Mol. Biol. 48: 501-510
Konieczny A, Ausubel FM. 1993. A procedure for mapping Arabidopsis mutations using co-dominant ecotype-specific PCR-based markers. Plant J. 4: 403-410
Litt M, Luty JM. 1989. A hypervariable microsatellite revealed by in vitro amplification of a dinucleotide repeat within the cardiac muscle actin gene. Am. J. Hum. Genet. 44: 397-401
Martin B, Nienhuis J, King G, Schaefer A. 1989. Restriction fragment length polymorphisms associated with water-use efficiency in tomato. Science 243: 1725-1728
Mason-Gamer RJ, Weil CF, Kellogg EA. 1998. Granule-bound starch synthase: structure, function, and phylogenetic utility. Mol. Biol. Evol. 15: 1658-1673
Matsuoka Y, Mitchell SE, Kresovich S, Goodman M, Doebley J. 2002. Microsatellites in *Zea*-variability, patterns of mutations, and use for evolutionary studies. Theor. Appl. Genet. 104: 436-450
McCouch SR, Kochert G, Yu ZH, Wang ZY, Khush GS, Coffma, WR, Tanksley SD. 1988. Molecular mapping of rice chromosomes. Theor. Appl. Genet. 76: 815-829
Meyer W, Mitchell TG, Freedman EZ, Vilgalys R. 1993. Hybridization probes for conventional DNA fingerprinting used as single primers in the polymerase chain reaction to distinguish strains of Cryptococcus *neoformans*. J. Clin. Microbiol. 31: 2274-2280
Miller JC, Tanksley SD. 1990. RFLP analysis of phylogenetic relationships and genetic variation in the genus *Lycopersicon*. Theor. Appl. Genet. 80: 437-448
Morgante M, Vogel J. 1994. Compound microsatellite primers for the detection of genetic polymorphisms. US patent application no. 08/326456
Morgante M, Hanafey H, Powell W. 2002. Microsatellites are preferentially associated with nonrepetitive DNA in plant Genome. Nat. Genet. 30: 194-200
Morjane H, Geistlinger J, Harrabi M, Weising K, Kahl G. 1994. Oligonucleotide fingerprinting detects genetic diversity among *Ascochyta rabiei* isolates from a single chickpea field in Tunesia. Curr. Genet. 26: 191-197
Naik S, Gill KS, Prakasa Rao VS, Gupta VS, Tamhankar SA,

Pujar S, Gill BS, Ranjekar PK. 1998. Identification of an STS marker linked to the *Aegilops speltoides* derived leaf rust resistance gene Lr28 in wheat. Theor. Appl. Genet. 97: 535-54

Nakamura Y, Leppert M, O'Connell P, Wolff R, Holm T, Culver M, Martin C, Fujimoto E, Hoff M, Kumlin E,White R. 1987. Variable number tandem repeat (VNTR) markers for human gene mapping. Science 235: 1616-1622

Nam HG, Giraudat J, Den Boer B, Moonan F, Loos WDB, Hauge BM, Goodman HM. 1989. Restriction Fragment Length Polymorphism linkage map of *Arabidopsis thaliana*. Plant Cell 1: 699-705

Niewohner J, Salamini F, Gebhardt C. 1995. Development of PCR assay diagnostic for RFLP markers closely linked to allele GROI and HI conferring resistance to root cyst nematode *Globodera rostochinensis* potato. Mol. Breed. 1: 65-78

Olsen M, Hood L, Cantor C, Botstein D. 1989. A common language for physical mapping of the human genome. Science 245: 1434-1435

Orita M, Suzuki Y, Sekiya T, Hayashi K. 1989. Rapid and sensitive detection of point mutations and DNA polymorphisms using polymerase chain reaction. Genomics 5: 874-879

Pakniyat H, Powell W, Baird E, Handley LL, Robinson D, Scrimgeour EM, Nevo E, Hackett C A, Caligari PDS, Forster BP. 1997. AFLP variation in wild barley (*Hordeum spontaneum* C. Koch) with reference to salt tolerance and associated ecogeography. Genome 40: 332-341

Paran I, Michelmore RW. 1993. Development of reliable PCR-based markers linked to downy mildew resistance genes in lettuce. Theor. Appl. Genet. 85: 985-993

Paran I, Kesseli R, Michelmore R. 1991. Identification of RFLP and RAPD markers linked to downy mildew resistance gene in tomato using near isogenic lines. Genome 34: 1021-1027

Pease AC, Solas D, Sullivan EJ, Cronin MT, Holmes CP, Fodor SP. 1994. Light generated oligonucleotide arrays for rapid DNA sequence analysis. Proc. Natl. Acad. Sci. USA 91: 5022-5026

Powell W, Morgante WM, McDevitt R, Vendramin GG, Rafalski JA. 1995. Polymorphic simple sequence repeat regions in chloroplast genomes: applications to the population genetics of pines. Proc. Natl. Acad. Sci. USA 92: 7759-7763

Ramakrishna W, Chowdari KV, Lagu MD, Gupta VS, Ranjekar PK. 1995. DNA fingerprinting to detect genetic variation in rice using hypervariable DNA sequences. Theor. Appl. Genet. 90: 1000-1006

Ratnaparkhe MB, Santra DK, Tullu A, Muehlbauer FJ. 1998. Inheritance of intersimple sequence repeat polymorphisms and linkage with a fusarium wilt resistance gene in chickpea. Theor. Appl. Genet. 96: 348-353

Reddy MP, Sarla N, Siddiq EA. 2002. Inter simple sequence repeat (ISSR) polymorphism and its application in plant breeding. Euphytica 128: 9-17

Richardson T, Cato S, Ramser J, Kahl G, Weising K. 1995. Nucleic Acids Res. 23: 3798-3799

Saal B, Wricke G. 2002. Clustering of amplified fragment length polymorphism markers in a linkage map of rye. Plant Breed. 121: 117-123

Sant VJ, Patankar AG, Gupta VS, Sarode ND, Mhase LB, Sainani MN, Deshmukh RB, Ranjekar PK. 1999. Potential of DNA markers in detecting divergence and analysis in heterosis in Indian elite chickpea cultivars. Theor. Appl. Genet. 98: 1217-1225

Sarfatti M, Katan J, Fluhr R, Zamir D. 1989. An RFLP marker in tomato linked to the *Fusarium oxysporum* resistance gene *I2*. Theor. Appl. Genet. 78: 755-759

Schachermayr G, Messmer MM, Feuillet C, Winzeler H, Keller B. 1995. Identification of molecular markers linked to the *Agropyron elongatum*-derived leaf rust resistance gene *lr24* in wheat. Theor. Appl. Genet. 90: 982-990

Schlotterer C. 2004. The evolution of molecular markers-just a matter of fashion? Nat. Rev. Genet. 5: 63-69

Sen A, Balyan HS, Sharma PC, Ramesh B, Kumar A, Roy JK, Varshney RK, Gupta PK. 1997. DNA amplification fingerprinting (DAF) as a new source of molecular markers in bread wheat. Wheat Inf. Serv. 85: 35-42

Shattuck-Eidens D, Bell RN, Mitchell JT, McWorther VC. 1991. Rapid detection of maize DNA sequence variation. GATA 8: 240-245

Slatkin M. 1987. Gene flow and population structure of natural populations. Science 263: 787-792

Sobrino B, Briona M, Carracedoa A. 2005. SNPs in forensic genetics: a review on SNP typing methodologies. Forensic Sci. Int. 154: 181-194

Somers DJ, Zhou Z, Bebeli P, Gustafson JP. 1996. Repetitive, genome-specific probes in wheat (*Triticum aestivum* L. em Thell) amplified with minisatellite core sequences. Theor. Appl. Genet. 93: 983-989

Tanksley SD, Ganal MW, Martin GB. 1995. Chromosome landing: A paradigm for map based gene cloning in plants with large genomes. Trends Genet. 11: 63-68

Tost J, Gut IG. 2002. Genotyping single nucleotide polymorphisms by mass spectrometry. Mass Spectrom Rev. 21: 388-418

Varshney RK, Graner A, Sorrells ME. 2005. Genic microsatellite markers in plants: features and applications. Trends Biotechnol. 23: 48-55

Virk PS, Newbury HJ, Jackson MT, Ford-Loyd BV. 1997. In JA Callow, HJ Newbury, eds, Biotechnology and Plant Genetic Resources: Conservation and Use, pp 21-23

Vos P, Hogers R, Bleeker M, Reijans M, Van De Lee T, Hornes M, Frijters A, Pot J, Peleman J, Kuiper M, Zabeau M. 1995. AFLP: a new technique for DNA fingerprinting. Nucleic Acids Res. 23: 4407-4414

Weising K, Nybom H, Wolff K, Meyer W 1995. DNA fingerprinting in plants and fungi. CRC Press, Boca Raton, Ann Arbor, London, Tokyo, 321 pp

Welsh J, McClelland M. 1990. Fingerprinting Genomes using PCR with arbitrary primers. Nucleic Acids Res. 18: 7213-7218

Williams JGK, Kublelik AR, Livak KJ, Rafalski JA, Tingey SV. 1990. DNA polymorphism's amplified by arbitrary primers are useful as genetic markers. Nucleic Acids Res. 18: 6531-6535

Williams MNV, Pande N, Nair S, Mohan M, Bennett J. 1991. Restriction fragment polymorphism analysis of polymerase chain reaction product amplified from mapped loci of rice genomic DNA. Theor. Appl. Genet. 82: 489-498

Williamson VM, Ho J-Y, Wu FF, Miller N, Kaloshian I. 1994. A PCR-based marker tightly linked to the nematode resistance gene, *Mi*, in tomato. Theor. Appl. Genet. 87: 757-763

Winter P, Kahl G. 1995. Molecular marker technologies for plant improvement. World J. Microbiol. Biotechnol. 11: 438-448

Wu J, Krutovskii KV, Strauss SH. 1998. Abundant mitochondrial genome diversity, population differentiation, and convergent phenotypic evolution in pines. Genetics 150: 1605-1614

Xiao J, Li J, Grandillo S, Ahn SN, Yuan L, Tanksley SD, McCouch SR. 1998. Identification of trait-improving quantitative trait loci alleles from a wild rice relative, *Oryzarufipogon*. Genetics 150: 899-909

Yang GP, Saghai Maroof MA, Xu CG, Zhang Q, Biyashev RM. 1994. Comparative analysis of microsatellite DNA polymorphism in landraces and cultivars of rice. Mol. Gen. Genet. 245: 187-194

Young ND, Menancio-Hautea D, Fatokun CA, Danesh D 1992. RFLP technology, crop improvement, and international agriculture. In G Thottappilly, LM Monti, DR Moham, AW Moore, eds, Biotechnology: Enhancing research on tropical crops in Africa. Technical Center for Agriculture and Rural Cooperation, International Institute of Tropical Agriculture, pp 221-230

Yu J, Hu, S, Wang J, Wong GKS, Li S, Liu B. 2002. A draft sequence of the rice genome (*Oryza sativa* L. ssp. *indica*). Science 296: 79-92

Yu YG, Saghai Maroof MA, Buss GR, Maughan PJ, Tolin SA. 1994. RFLP and Microsatellite Mapping of a Gene for Soybean Mosaic Virus Resistance. Phytopathol. 84: 60-64

Zhao X, Kochert G. 1992. Characterisation and genetic mapping of short, highly repeated interspersed DNA sequence from rice (*Oryzae sativa* L). Mol. Gen. Genet. 2311: 353-359

Zietkiewicz E, Rafalski A, Labuda D. 1994. Genome fingerprinting by simple sequence repeats (SSR)-anchored PCR amplification. Genomics 20: 176-183