African Journal of Biotechnology Vol. 5 (25) pp. 2540-2568 29 December, 2006
Available online at http://www.academicjournals.org/AJB
ISSN 1684–5315 © 2006 Academic Journals

Review

An overview of molecular marker methods for plants

K. Semagn[1]*, Å. Bjørnstad[2] and M. N. Ndjiondjop[1]

[1]Africa Rice Center (WARDA), 01 BP 2031, Cotonou, Benin.
[2]Norwegian University of Life Sciences, Department of Plant and Environmental Sciences, P.O. Box 5003, N-1432, Ås, Norway.

Accepted 23 November, 2006

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. According to the kind of study to be undertaken, one can choose among the variety of molecular techniques, each of which combines at least some desirable properties. This article provides detail review for 11 different molecular marker methods: restriction fragment length polymorphism (RFLP), random amplified polymorphic DNA (RAPD), amplified fragment length polymorphism (AFLP), inter-simple sequence repeats (ISSRs), sequence characterized regions (SCARs), sequence tag sites (STSs), cleaved amplified polymorphic sequences (CAPS), microsatellites or simple sequence repeats (SSRs), expressed sequence tags (ESTs), single nucleotide polymorphisms (SNPs), and diversity arrays technology (DArT).

Key words: AFLP, DArT, DNA markers, hybridization, ISSR, polymerase chain reaction, RAPD, RFLP, SNP, SSR.

INTRODUCTION

The differences that distinguish one plant from another are encoded in the plant's genetic material, the deoxyribonucleic acid (DNA). DNA is packaged in chromosome pairs (strands of genetic material; Figure 1), one coming from each parent. The genes, which control a plant's characteristics, are located on specific segments of each chromosome. All of the genes carried by a single gamete (i.e., by a single representative of each of all chromosome pairs) is known as genome (King and Stansfield, 1997). Although the whole genome sequence is now available for a few plant species such as *Arabidopsis thaliana* (The Arabidopsis Genome Initiative, 2000) and rice (The Rice Genome Mapping Project, 2005), to help identify specific genes located on a particular chromosome, most scientists use an indirect method called genetic markers. A genetic marker can be defined in one of the following ways: (a) a chromosomal landmark or allele that allows for the tracing of a specific region of DNA; (b) a specific piece of DNA with a known position on the genome (Wikipedia-the free encyclopedia; http://en.wikipedia.org/wiki/Genetic_marker); or (c) a gene whose phenotypic expression is usually easily discerned, used to identify an individual or a cell that carries it, or as a probe to mark a nucleus, chromosomes, or locus (King and Stansfield, 1990). Since the markers and the genes they mark are close together on the same chromosome, they tend to stay together as each generation of plants is produced. As scientists learn where markers occur on a chromosome, and how close they are to specific genes, they can create a genetic linkage map. Such genetic maps serve several purposes, including detailed analysis of associations between economically important traits and genes or quantitative trait loci (QTLs) and facilitate the introgression of

*Corresponding author. E-mail: k.semagn@cgiar.org or semagnk@yahoo.com; Fax (229) 21350556; Tel (229)21350188.

desirable genes or QTLs through marker-assisted selection.

Genetic markers fall into one of the three broad classes: those based on visually assessable traits (morphological and agronomic traits), those based on gene product (biochemical markers), and those relying on a DNA assay (molecular markers). The idea of using genetic markers appeared very early in literatures (Sax, 1932; Wexelsen, 1933) but the development of electrophoretic assays of isozymes (Markert and Moller, 1959) and molecular markers (Botstein et al., 1980; Nakamura et al., 1987; Welsh and McClelland, 1990; Williams et al., 1990; Adams et al., 1991; Caetano-Anolles et al., 1991; Akkaya et al., 1992; Akopyanz et al., 1992; Jordan and Humphries, 1994; Zietkiewicz et al., 1994; Vos et al., 1995; Jaccoud et al., 2001) have greatly improved our understanding in biological sciences. Molecular markers should not be considered as normal genes, as they usually do not have any biological effect, and instead can be thought of as constant landmarks in the genome. They are identifiable DNA sequences, found at specific locations of the genome, and transmitted by the standard laws of inheritance from one generation to the next. The existence of various molecular techniques and differences 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, and advantages and disadvantages of the most widely used molecular markers for genetic diversity studies, genetic mapping, marker-trait association studies, and marker assisted selection programs.

DNA Extraction is the Beginning of Molecular Markers Analysis

Extraction (isolation) of DNA (nuclear, mitochondrial, and/or chloroplast DNA) from sample to be studied is the first step for all molecular marker types. DNA can be extracted either from fresh, lyophilized, preserved or dried samples but fresh material is ideal for obtaining good quality DNA. There are many alternative protocols for DNA extraction and the choice of a protocol depends on the quality and quantity of DNA needed, nature of samples, and the presence of natural substances that may interfere with the extraction and subsequent analysis. DNA extraction protocols vary from simple and quick ones (e.g., Clancy et al., 1996; Ikeda et al., 2001; Dayteg et al., 1998; von Post et al., 2003) that yields low quality DNA but nevertheless good enough for routine analyses to the laborious and time-consuming standard methods (e.g. Murray and Thompson, 1980; Dellaporta et al., 1983; Saghai-Maroof et al., 1984) that usually produce high quality and quantity of DNA (Figure 2). The most commonly used DNA extraction protocols involve breaking (through grinding) or digesting away cell walls and membranes in order to release the cellular constituents. Removal of membranes lipids is facilitated by using detergents such as sodium dodecyl sulphate (SDS), Cetyl trimethylammonium bromide (CTAB) or mixed alkyl trimethyl-ammonium bromide (MTAB). The released DNA should be protected from endogenous nucleases and EDTA is often included in the extraction buffer to chelate magnesium ions that is a necessary co-factor for nucleases. DNA extracts often contain a large amount of RNA, proteins, polysaccharides, tannins and pigments, which may interfere with the extracted DNA. Most proteins are removed by adding a protein degrading enzyme (proteinase-K), denaturation at 65 ^{o}C and precipitation using chloroform and isoamyl alcohol. RNAs are normally removed using RNA degrading enzyme called RNase A. Polysaccharide-like contaminants are, however, more difficult to remove. NaCl, together with CTAB is known to remove polysaccharides (Murray and Thompson, 1980; Paterson et al., 1993). Some protocols replace NaCl by KCl (Thompson and Henry, 1995).

As DNA will be released along with other compounds (lipids, proteins, carbohydrates, and/or phenols), it needs to be separated from others by centrifugation. The DNA in the aqueous phase will then be transferred into new tubes and precipitated in salt solution (e.g. sodium acetate) or alcohol (100% isopropanol or ethanol), re-dissolved in sterile water or buffer. Finally, the concentration of the extracted DNA needs to be measured using either 1% agarose gel electrophoresis or spectrophotometer. Agarose gel is useful to check whether the DNA is degraded or not (Figure 2) but estimating DNA concentration by visually comparing band intensities of the extracted DNA with a molecular ladder of known concentration is too subjective. Spectrophotometer measures the intensity of absorbance of DNA solution at 260 nm wavelength, and also indicates the presence of protein contaminants but it does not tell whether the DNA is degraded or not. There are three possible outcomes at the end of any DNA extraction:

a) There is no DNA.
b) The DNA appears as sheared (too fragmented), which is an indication of degradation for different reasons.
c) DNA appears as whitish thin threads (good quality DNA) or brownish thread (DNA in the presence of oxidation from contaminants such as phenolic compounds).

The researcher, therefore, needs to test different protocols in order to find out the best one that works for the species under investigation.

Types of Molecular Markers

Various types of molecular markers have been described in the literature, which are listed in alphabetical order as follows: allele specific associated primers (ASAP; Gu et al., 1995), allele specific oligo (ASO; Beckmann, 1988), allele specific polymerase chain reaction (AS-PCR; Land-

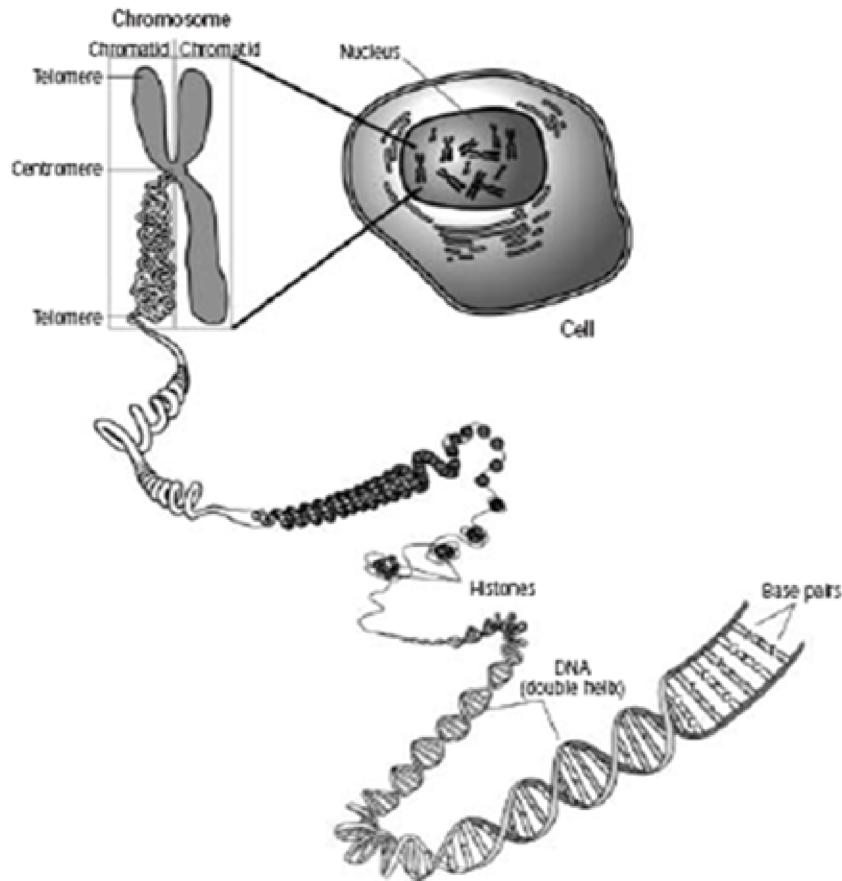

Figure 1. Structure of DNA after unpacking it from the chromosome.

egren et al., 1988) amplified fragment length polymorphism (AFLP; Vos et al., 1995), anchored microsatellite primed PCR (AMP-PCR; Zietkiewicz et al., 1994), anchored simple sequence repeats (ASSR; Wang et al., 1998), arbitrarily primed polymerase chain reaction (AP-PCR; Welsh and McClelland, 1990), cleaved amplified polymorphic sequence (CAPS; Akopyanz et al., 1992; Konieczny and Ausubel, 1993), degenerate oligonucleotide primed PCR (DOP-PCR; Telenius et al., 1992), diversity arrays technology (DArT; Jaccoud et al., 2001), DNA amplification fingerprinting (DAF; Caetano-Anolles et al., 1991), expressed sequence tags (EST; Adams et al., (1991), inter-simple sequence repeat (ISSR; Zietkiewicz et al., 1994), inverse PCR (IPCR; Triglia et al., 1988), inverse sequence-tagged repeats (ISTR; Rohde, 1996), microsatellite primed PCR (MP-PCR; Meyer et al., 1993), multiplexed allele-specific diagnostic assay (MASDA; Shuber et al., 1997), random amplified microsatellite polymorphisms (RAMP; Wu et al., 1994), random amplified microsatellites (RAM; Hantula et al., 1996), random amplified polymorphic DNA (RAPD; Williams et al., 1990), restriction fragment length polymorphism (RFLP; Botstein et al., 1980), selective amplification of microsatellite polymorphic loci (SAMPL; Morgante and Vogel, 1994), sequence characterized amplified regions (SCAR; Paran and Michelmore, 1993), sequence specific amplification polymorphisms (S-SAP; Waugh et al., 1997), sequence tagged microsatelite site (STMS; Beckmann and Soller, 1990), sequence tagged site (STS; Olsen et al., 1989), short tandem repeats (STR; Hamada et al., 1982), simple sequence length polymorphism (SSLP; Dietrich et al., 1992), simple sequence repeats (SSR; Akkaya et al., 1992), single nucleotide polymorphism (SNP; Jordan and Humphries 1994), single primer amplification reactions (SPAR; Gupta et al., 1994), single stranded conformational polymorphism (SSCP; Orita et al., 1989), site-selected insertion PCR (SSI; Koes et al., 1995), strand displacement amplification (SDA; Walker et al., 1992), and variable number tandem repeat (VNTR; Nakamura et al., 1987).

Although some of these marker types are very similar (e.g., ASAP, ASO and AS-PCR), some synonymous (e.g., ISSR, RAMP, RAM, SPAR, AMP-PCR, MP-PCR, and ASSR; Reddy et al., 2002), and some identical (e.g., SSLP, STMS, STR and SSR), there are still a wide range of techniques for researchers to choose upon. One of the main challenges is, therefore, to associate the purpose(s) of a specific project with the various molecular marker types. The various molecular markers can be classified into different groups based on:

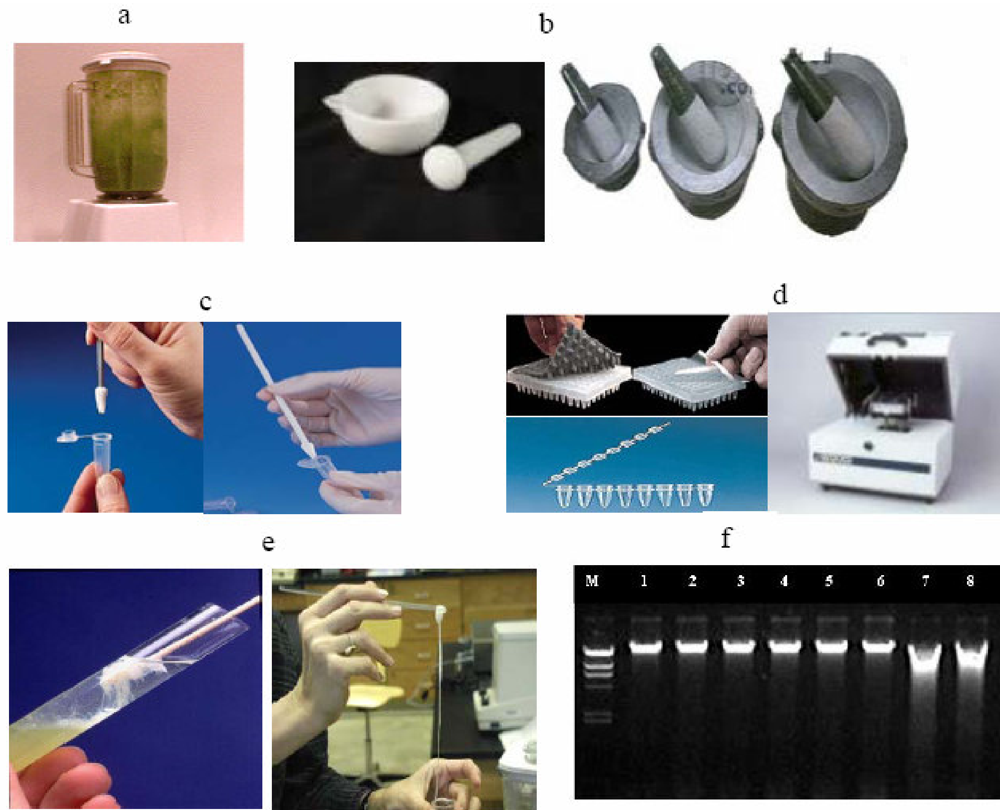

Figure 2. The different methods of grinding samples for DNA extraction: (a) large scale DNA extraction using juice-maker; (b) medium scale DNA extraction using mortal and pestle; (c) mini-scale DNA extraction using 1.5 ml tubes and small grinding pestle (even a nail or driller); (d) mini-scale DNA extraction using 96-well format microtiter plates or strip tubes (left) and grinding machine (right); (e) DNA after precipitation in 100% isopropanol. Note that the DNA is whitish and tiny thread like structure, which is an indication of good quality that can easily be fished using glass hooks); (f) agarose gel (1%) showing DNA concentration for 8 samples and a marker (M) with a known concentration. The DNA for sample number 7 and 8 shows partial degradation compared to the other samples.

a) Mode of transmission (biparental nuclear inheritance, maternal nuclear inheritance, maternal organelle inheritance, or paternal organelle inheritance).
b) Mode of gene action (dominant or codominant markers).
c) Method of analysis (hybridization-based or PCR-based markers).

The next section provides detail reviews for the latter.

Hybridization-based molecular markers

RFLP is the most widely used hybridization-based molecular marker. RFLP markers were first used in 1975 to identify DNA sequence polymorphisms for genetic mapping of a temperature-sensitive mutation of adeno-virus serotypes (Grodzicker et al., 1975). It was then used for human genome mapping (Botstein et al., 1980), and later adopted for plant genomes (Helentjaris et al., 1986; Weber and Helentjaris, 1989). 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 due to one or more of the following causes: 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. Hence, digestion of DNA with restriction enzymes results in fragments whose number and size can vary among individuals, populations, and species. The procedures and principles of RFLP markers are summarized in Figure 3:

a) Digestion of the DNA with one or more restriction enzyme(s).
b) Separation of the restriction fragments in agarose gel.
c) Transfer of separated fragments from agarose gel to a filter by Southern blotting.
d) Detection of individual fragments by nucleic acid hybridization with a labeled probe(s)
e) Autoradiography (Perez de la Vega, 1993; Terachi, 1993; Landry, 1994).

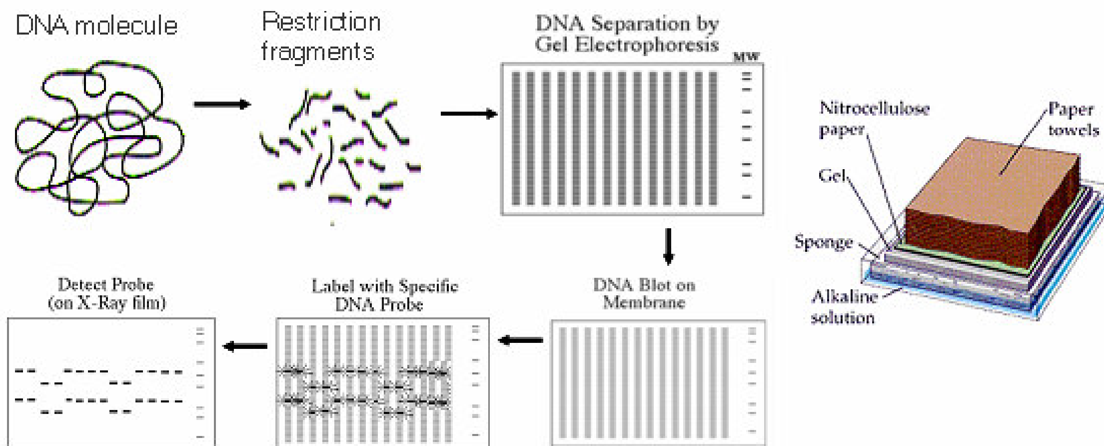

Figure 3. Outline of the different steps of restriction fragment length polymorphism (RFLP) markers. Double-stranded DNA fragments generated by restriction enzymes are separated according to length by gel electrophoresis. A sheet of either nitrocellulose or nylon paper (membrane) is laid over the gel, and the separated DNA fragments are transferred to the sheet by blotting (Southern transfer). The gel is supported on a layer of sponge in a bath of alkali solution, and the buffer is sucked through the gel and the nitrocellulose paper by paper towels stacked on top of the nitrocellulose. As the buffer is sucked through, it denatures the DNA and transfers the single-stranded fragments from the gel to the surface of the nitrocellulose sheet, where they adhere firmly. This transfer is necessary to keep the DNA firmly in place while the hybridization procedure is carried out. The nitrocellulose sheet containing the bound single-stranded DNA fragments is carefully peeled off the gel and placed in a sealed plastic bag that contained a radioactively labeled DNA probe for hybridization. The sheet is removed from the bag and washed thoroughly, so that only probe molecules that have hybridized to the DNA on the paper remain attached. After autoradiography, the DNA that has hybridized to the labeled probe will show up as bands on the autoradiograph.

Restriction enzymes (endonucleases) are bacterial enzymes (e.g., MseI, EcoRI, PstI, etc.) that recognize specific four, six or eight base pair (bp) sequences in DNA, and cleave double-stranded DNA whenever these sequences are encountered. For example, EcoRI has six bp recognition sequence and it cuts between G and A whenever the sequences 5'…GAATTC…3' or 3'…CTTAAG…5' exit together. The choice between using enzymes recognizing four, six or eight bp can be made depending on the resolution required and the electrophoresis facility available. The greatest resolution is obtained by using 'four-cutters' (enzymes recognizing a four base pair sequence) because there are many such sites in the genome. The fragments produced will be relatively small, which provides a better chance of identifying single base alterations. Conversely, use of enzymes that recognizes an eight bp sequence will require fewer probes, because larger fragments of the genome are analyzed at one time. As a result of this, only large alterations of the DNA will be visualized and complex electrophoresis system must be used to resolve the fragments into discrete bands. As a compromise, 'six-cutter' enzymes are most often used for RFLP analysis because they are readily available, cheaper and they usually produce fragments in the size range of 200 to 20,000 bp, which can be separated conveniently on agarose gels (Table 1). After digestion by restriction enzymes, the DNA is present as a mixture of linear double-stranded molecules of various lengths. These are then separated by electrophoresis through agarose or polyacrylamide gels. The choice between agarose and polyacrylamide is based on the restriction enzymes chosen. Four-cutters produce fragments too small to be resolved by agarose gels; hence, polyacrylamide gels are required. Conversely, polyacrylamide gels can not normally be used to resolve the fragments produced by six-cutters so agarose gels must be used. These considerations have led to most workers use six-cutter enzymes, as agarose gels are much easier to handle (Potter and Jones, 1991).

DNA fragments separated by gel electrophoresis are then denatured to single strands and transferred onto a solid support ('membranes' or 'filter') using a technique referred to as 'Southern blotting' or 'Southern hybridization' (Southern, 1975). The basis of this is the transfer of the DNA from the gel to a solid support, thus preserving the position of the fragments as they were in the gel, yet enabling hybridization reactions to be performed. Filter-immobilized DNA is then allowed to hybridize to labeled probe DNA. The probes used for hybridization are preferably single locus and mostly species-specific probes of about 500 to 3000 bp in size (Staub and Serquen, 1996). Probes used to identify specific DNA fragments by hybridization are of two types: genomic clones (fragments of nuclear DNA) and cDNA clones (DNA copies of mRNA molecules). Genomic libraries are

Table 1. Recommended gel percentages for separation of linear DNA.

Agarose gel		Polyacrylamide gel	
Gel concentration (%)	**Range of separation (bp)**	**Gel concentration (%)**	**Range of separation (bp)**
0.5	1,000-30,000	3.5	100-1,000
0.7	800-12,000	5.0	80-500
1.0	500-10,000	8.0	60-400
1.2	400-7,000	12.0	40-200
1.4	200-4000	20.0	5-100
2.0	50-2,000		

easy to construct and will contain many repetitive probes because repetitive sequences constitute the largest proportion of plant genomic DNA. Such probes will hybridize into many fragments on the filters and produce very complex patterns. cDNA libraries are difficult to construct but they contain predominantly unique or low copy number sequences representing expressed genes and usually provide fewer bands on the filter (for details, see ESTs below). Although use of cDNA probes will help the identification of small changes, the proportion of the genome covered by each probe will be relatively small and many more probes must be used (Potter and Jones, 1991). Therefore, the selection of appropriate source for RFLP probe varies with the requirement of particular application under consideration. The denatured probe solution is left in contact with the filter to allow hybridization of the probe to target sequences. Hybridization is the process by which the labeled probe binds to complementtary DNA on the filter, enabling visualization of specific DNA fragments. Labeling has been traditionally achieved by means of radioactive nucleotides, but non-radioactive methods are now available (Holtke et al., 1995; Mansfield et al., 1995). Non-specific hybridization must then be washed off under more 'stringent' conditions than those used for initial hybridization. If radioactive probes are used, the filter is placed against photographic film, where radioactive disintegrations from the probe result in visible bands. An autoradiography of membrane will reveal the set of fragments complementary to the probe. With non-radio-active probes, such as digoxigenin, antibodies against the modified nucleotides and a coupled enzymatic reaction are used to show up the set of fragments directly on the membrane (Holtke et al., 1995).

The information obtained with the RFLP technique depends upon both the number of probes and restriction enzymes used. Each different probe hybridizes with a different set of genomic DNA fragments and each enzyme excises a segment of genomic DNA at different points (Perez de la Vega, 1993). The major strength of RFLP markers are high reproducibility, codominant inheritance, good transferability between laboratories, provide locus-specific markers that allow synteny (conserved order of genes between related organisms) studies, no sequence information required, and relatively easy to score due to large size difference between fragments. There are, however, several limitations for RFLP analysis:

I. It requires the presence of high quantity and quality of DNA (e.g., Potter and Jones, 1991; Roy et al., 1992; Young et al., 1992).
II. It depends on the development of specific probe libraries for the species.
III. The technique is not amenable for automation.
IV. The level of polymorphism is low, and few loci are detected per assay.
V. It is time consuming, laborious, and expensive (Yu et al., 1993).
VI. It usually requires radioactively labeled probes.

PCR-based markers

PCR is a molecular biology technique for enzymatically replicating (amplifying) small quantities of DNA without using a living organism. It is used to amplify a short (usually up to 10 kb), well-defined part of a DNA strand from a single gene or just a part of a gene. Since its invention by Kary Mullis in 1983, this technique enabled the development of various types of PCR-based techniques (and a Nobel Prize for Kary Mullins in 1993). However, the basic PCR procedure was described in 1968 by Kleppe and his co-workers in Khorana's group, and it has been discussed if the main contribution of Mullis was the thermostable DNA polymerase and if he actually knew this paper. This point has been important in challenging the PCR patent. The basic protocol for PCR is simple (Figure 4):

I. Double-stranded DNA is denatured at high temperature (92-95 °C) to form single strands (templates).
II. Short single strands of DNA (known as primers) bind at a lower annealing temperature to the single stranded complementary templates at ends flanking the target sequences.
III. The temperature is raised usually to 72 °C (sometimes 68 °C) for the DNA polymerase enzyme to catalyze the template-directed syntheses of new double-stranded DNA molecules that are identical in sequence to the starting material.

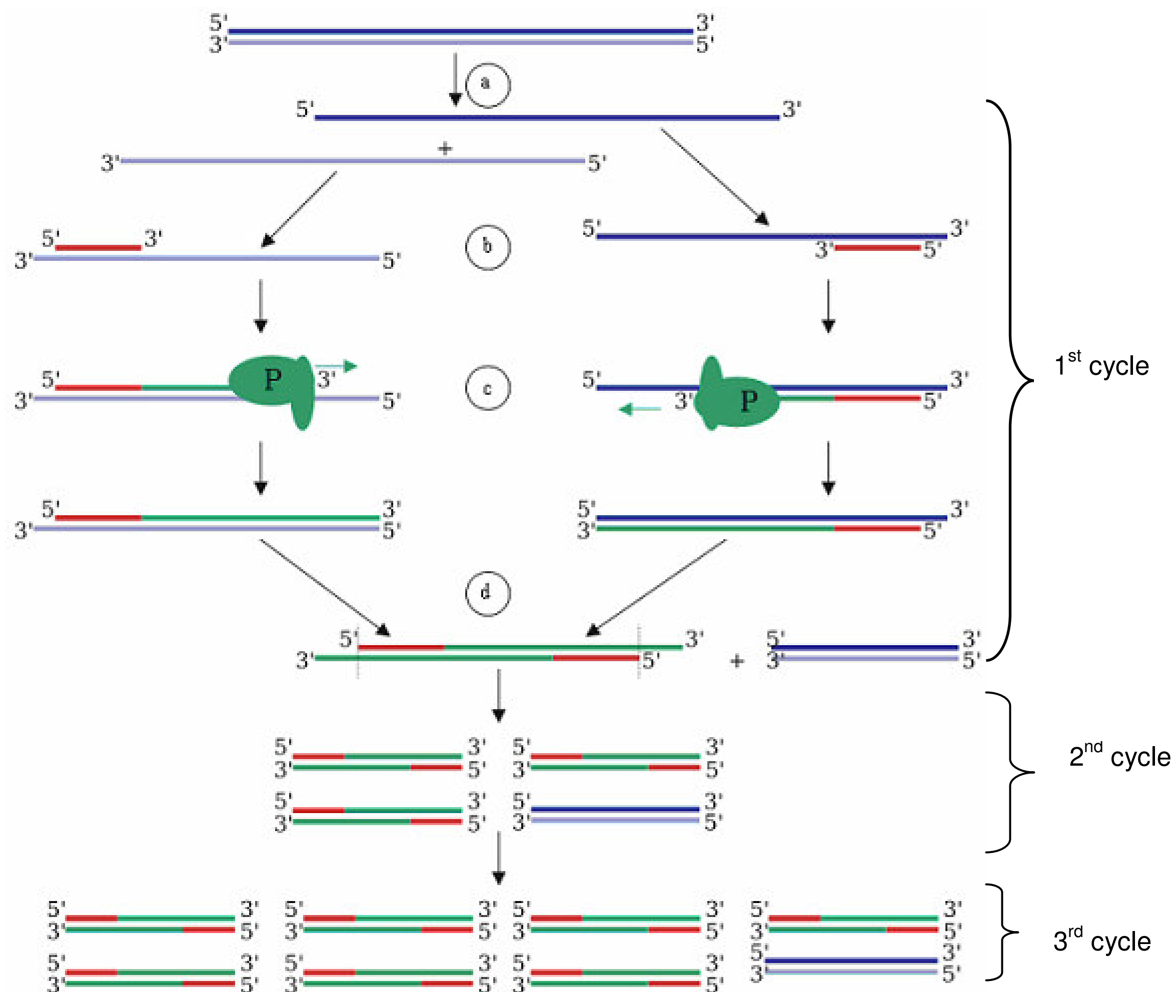

Figure 4. Schematic drawing of the different steps of polymerase chain reaction (PCR): (a) denaturing step at 92-95°C; (b) primer annealing step (37-68°C depending of the technique); (c) extension step at 72°C (P=Taq DNA polymerase), and (d) end of the first cycle with two copies of DNA strands. The two resulting DNA strands make up the template DNA for the next cycle, thus doubling the amount of DNA duplicated for each new cycle (Source: http://en.wikipedia.org/wiki/Polymerase_chain_reaction).

IV. The newly synthesized double-stranded DNA target sequences are denatured at high temperature, and the cycle is repeated.

The amplification of target DNA can be exponential in that every cycle has the potential to double the amount of target DNA from the previous cycle, provided there is sufficient amount of DNA polymerase, primers, and deoxynucleotide triphosphates (dNTPs) in the reaction solution. Although the basic protocol of PCR is straightforward, each application requires optimizing the various parameters for the species to be studied.

Back in the early days of PCR work, the DNA polymerase would need to be added fresh to the reaction at each temperature cycle because thermostable (high temperature tolerant) DNA polymerases were not commercially available. Of course, there were also no thermocyclers so moving the tubes from one temperature bath to another for several hours was a job that fells to graduate students and/or technicians. The discovery of Taq DNA polymerase, the DNA polymerase that is used by the bacterium *Thermus auquaticus* in hot springs, was decisive for the immense utility and popularity of PCR-based techniques. The original function of this enzyme was to facilitate the *in vivo* replication of DNA in the thermophilic bacteria, and thus it ables to operate at the high temperature required for the *in vitro* replication. This DNA polymerase is stable at high temperature needed to perform the amplification, whereas other DNA polymerases become denatured. Nowadays, the PCR technology is much more advanced with a wide range of thermostable DNA polymerases (such as Taq, Pfu or Vent polymerase) and automation of reactions can be done by a PCR machine (thermocycler) that has found its way into nearly every molecular biology lab in the world.

The major advantages of PCR techniques compared to hybridization-based methods include:

1. A small amount of DNA is required.
2. Elimination of radioisotopes in most techniques.
3. The ability to amplify DNA sequences from preserved tissues.
4. Accessibility of methodology for small labs in terms of equipment, facilities, and cost.
5. No prior sequence knowledge is required for many applications, such as AP-PCR, RAPD, DAF, AFLP and ISSR.
6. High polymorphism that enables to generate many genetic markers within a short time, and
7. The ability to screen many genes simultaneously either for direct collection of data or as a feasibility study prior to nucleotide sequencing efforts (Wolfe and Liston, 1998).

These advantages, however, can vary depending on the specific technique chosen by the researcher. The various PCR-based techniques are of two types depending on the primers used for amplification:

1) Arbitrary or semi-arbitrary primed PCR techniques that developed without prior sequence information (e.g., AP-PCR, DAF, RAPD, AFLP, ISSR).
2) Site-targeted PCR techniques that developed from known DNA sequences (e.g., EST, CAPS, SSR, SCAR, STS).

Arbitrarily Amplified DNA Markers

RAPD (random amplified polymorphic DNA), AP-PCR (arbitrarily primed PCR) and DAF (DNA amplification fingerprinting) have been collectively termed multiple arbitrary amplicon profiling (MAAP; Caetano-Annolles, 1994). These three techniques were the first to amplify DNA fragments from any species without prior sequences information. The difference among MAAP techniques include modifications in amplification profiles by changing primer sequence and length, annealing temperature (Caetano-Anolles et al., 1992), the number of PCR cycles used in a reaction (Caetano-Anolles et al., 1991; Welsh and McClelland, 1991; Micheli et al., 1993; Jain et al., 1994), the thermostable DNA polymerase used (Bassam et al., 1992), enzymatic digestion of template DNA or amplification products (Caetano-Anolles et al., 1993), and alternative methods of fragment separation and staining. These three techniques produce markedly different amplification profiles, varying from quite simple (RAPD) to highly complex (DAF) patterns. The key innovation of RAPD, AP-PCR and DAF is the use of a single arbitrary oligonucleotide primer to amplify template DNA without prior knowledge of the target sequence. The amplification of nucleic acids with arbitrary primers is mainly driven by the interaction between primer, template annealing sites and enzymes, and determined by complex kinetic and thermodynamic processes (Caetano-Anolles, 1997). A discrete PCR product is produced when, at an appropriate annealing temperature, the single primer binds to sites on opposite strands of the genomic DNA that are within an amplifiable distance (Figure 5), generally less than 3,000 base pairs. In all AP-PCR, DAF and RAPD, polymorphisms (band presence or absence) result from changes in DNA sequence that inhibit primer binding or interfere with amplification of a particular marker in some individuals; therefore, they can be simply detected as DNA fragments that are amplified from one individual but not from another.

The RAPD protocol usually uses a 10 bp arbitrary primer at constant low annealing temperature (generally 34 – 37 ^{o}C). RAPD primers can be purchased as sets or individually from different sources, such as the University of British Colombia (http://www.michaelsmith.ubc.ca/services/NAPS/Primer_Sets) and the Operon Biotechnologies (http://www.operon.com). Although the sequences of RAPD primers are arbitrarily chosen, two basic criteria indicated by Williams et al. (1990) must be met: a minimum of 40% GC content (50 - 80% GC content is generally used) and the absence of palindromic sequence (a base sequence that reads exactly the same from right to left as from left to right). Because G-C bond consists of three hydrogen bridges and the A-T bond of only two, a primer-DNA hybrid with less than 50% GC will probably not withstand the 72 ^{o}C temperature at which DNA elongation takes place by DNA polymerase. The resulting PCR products are generally resolved on 1.5-2.0% agarose gels and stained with ethidium bromide (EtBr); polyacrylamide gels in combination with either $AgNO_3$ staining (e.g., Huff et al., 1993; Vejl, 1997; Hollingsworth et al., 1998), radioactivity (e.g., Pammi et al., 1994), or fluorescently labeled primers or nucleotides (e.g., CorleySmith et al., 1997; Weller and Reddy, 1997) are sometimes used. Despite its low resolving power, the simplicity and low cost of agarose gel electrophoresis has made RAPD more popular and rapid than AP-PCR and DAF.

Most RAPD fragments result from the amplification of one locus, and two kinds of polymorphism occur: the band may be present or absent, and the brightness (intensity) of the band may be different (Figure 6). Band intensity differences may result from copy number or relative sequence abundance (Devos and Gale, 1992) and may serve to distinguish homozygote dominant individuals from heterozygotes, as more bright bands are expected for the former. However, some authors (Thormann et al., 1994) found no correlation between copy number and band intensity. The fact that fainter bands are generally less robust in RAPD experiments (Ellsworth et al., 1993; Heun and Helentjaris, 1993) suggests that varying degrees of primer mismatch may account for many band intensity differences. Since the source of band intensity differences is uncertain (copy number or primer mismatch), most studies disregard scoring differences in band intensity although some authors have used up to 7-state scale of band intensity (Demeke et al.,

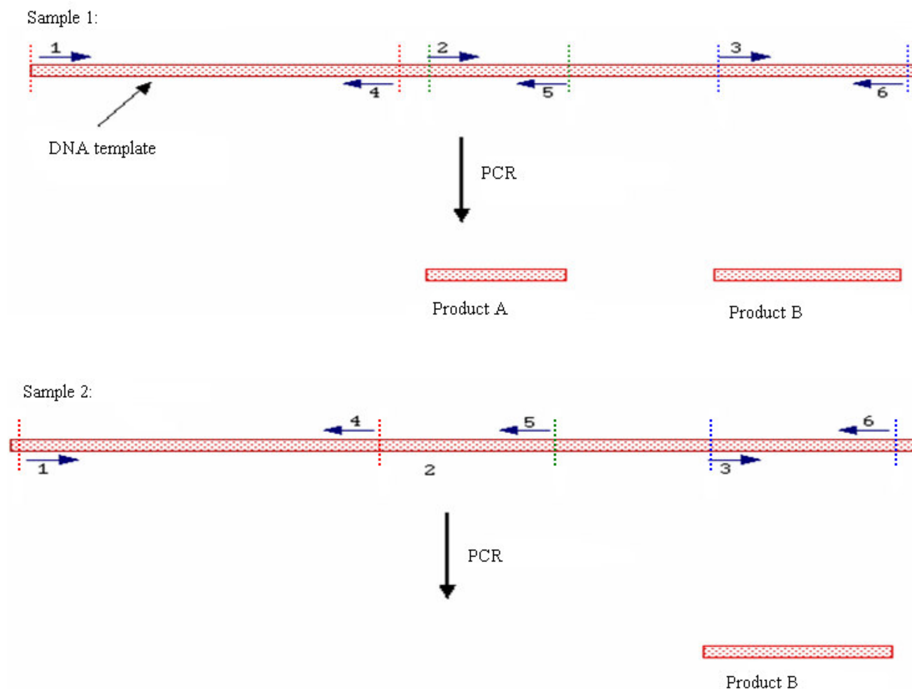

Figure 5. Schematic drawing for reaction conditions for random amplified polymorphic DNA (RAPD). The primers must anneal in a particular orientation (such that they point towards each other) and within a reasonable distance of one another. The arrows represent multiple copies of a single primer and the direction of the arrow indicates the direction in which DNA synthesis will occur. The numbers represent primer annealing sites on the DNA template. For sample 1, primers anneal to sites 1, 2, and 3 on the top strand of the DNA template and to sites 4, 5, and 6 on the bottom strand of the DNA template. In this example, only 2 RAPD products are formed for sample 1: (i) product A is produced by PCR amplification of the DNA sequence which lies in between the primers bound at positions 2 and 5; (ii) product B is the produced by PCR amplification of the DNA sequence which lies in between the primers bound at positions 3 and 6. No PCR product is produced by the primers bound at positions 1 and 4 because these primers are too far apart to allow completion of the PCR reaction. No PCR products are also produced by the primers bound at positions 4 and 2 or positions 5 and 3 because these primer pairs are not oriented towards each other. For sample 2, the primer failed to anneal at position 2 and PCR product was obtained only for primers bound at position 3 and 6.

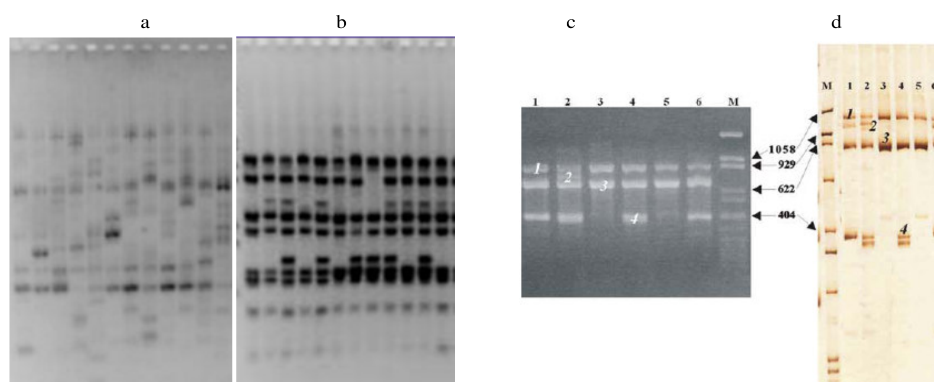

Figure 6. RAPD amplification products separated on 1.5% agarose gels with ethidium bromide staining (a-c) and 10% polyacrylamide gel with silver nitrate staining (d): (a) poor amplification, (b) good amplification, (c) the problem of homology of comigrating RAPD bands of 6 samples in agarose gel, and (d) each of the bands numbered from 1 to 4 in the agarose gel appeared to be two comigrating bands in the polyacrylamide gels.

1992; Adams and Demeke, 1993).

RAPD has three limitations:

1) Reproducibility
2) Dominant inheritance

3) Homology

Several factors have been reported to influence the reproducibility of RAPD reactions: quality and quantity of template DNA, PCR buffer, concentration of magnesium chloride, primer to template ratio, annealing temperature, Taq DNA polymerase brand or source, and thermal cycler brand (Wolff et al., 1993). The concern about reproducibility of RAPD markers, however, could be overcome through choice of an appropriate DNA extraction protocol to remove any contaminants (Micheli et al., 1994), by optimizing the parameters used (Ellsworth et al., 1993; Skroch and Nienhuis, 1995), by testing several oligonucleotide primers and scoring only the reproducible DNA fragments (Kresovich et al., 1992; Yang and Quiros, 1993), and by using appropriate DNA polymerase brand. The presence of artifactual bands (false positives) corresponding to rearranged fragments produced by nested primer binding sites (Schierwater et al., 1996; Rabouam et ,al. 1999) and intrastrand annealing and interactions during PCR (Hunt and Page, 1992; Caetano-Anolles et al., 1992) have also been reported to influence the reliability of RAPD data. The presence of both false negatives and false positives may, if frequent, seriously restrict the reliability of RAPDs for various purposes, including genetic diversity and mapping studies. All pair wise comparison of RAPD fragments along samples begins with the assumption that co-migrating bands (i.e., bands that migrate equal distance) represent homologous loci. However, as in any study based on electrophoretic resolution, the assumption that equal length equals homology may not be necessarily true, especially in polyploid species. For example, some RAPD bands scored as identical (equal length) have been found not to be homologous (e.g., Thormann et al., 1994; Pillay and Kenny, 1995; Figure 6); more accurate resolution of fragment size using polyacrylamide gels and $AgNO_3$ staining have been reported to reduce such errors (e.g., Huff et al., 1993). The other limitation of RAPD markers is that the majority of the alleles segregate as dominant markers, and hence the technique does not allow identifying dominant homozygotes from heterozygotes. The RAPD assays produce fragments from homozygous dominant or heterozygous alleles. No fragment is produced from homozygous recessive alleles because amplification is disrupted in both alleles.

The original DAF protocol is mainly different from RAPD in that it uses short primers (at least 5 bp), higher primer concentrations, two-temperature cycles in stead of three-temperature cycles, and detection of amplification product on $AgNO_3$ stained polyacrylamide gel. The main characteristics of AP-PCR technique in comparison to RAPD and DAF are:

a) The amplification reaction is divided into three steps, each with different stringencies and concentrations of constituents.
b) High primer concentrations are used in the first cycles.
c) Primers of 20 or more nucleotides, originally designned for other purposes (e.g., sequencing primers) are chosen arbitrarily.
d) Detection of amplification products involve radioactivity and autoradiography (Weising et al., 1995).

AFLP (amplified fragment length polymorphism)

AFLP technique combines the power of RFLP with the flexibility of PCR-based technology by ligating primer-recognition sequences (adaptors) to the restricted DNA (Lynch and Walsh, 1998). The key feature of AFLP is its capacity for “genome representation”: the simultaneous screening of representative DNA regions distributed randomly throughout the genome. AFLP markers can be generated for DNA of any organism without initial investment in primer/probe development and sequence analysis. Both good quality and partially degraded DNA can be used for digestion but the DNA should be free of restriction enzyme and PCR inhibitors. Details of the AFLP methodology have been reviewed by various authors (e.g., Blears et al., 1998; Mueller and Wolfenbarger, 1999; Ridout and Donini, 1999). The first step in AFLP analysis involves restriction digestion of genomic DNA (about 500 ng) with a combination of rare cutter (EcoRI or PstI) and frequent cutter (MseI or TaqI) restriction enzymes (Figure 7). Double-stranded oligonucleotide adaptors are then designed in such a way that the initial restriction site is not restored after ligation. Such adaptors are ligated to both ends of the fragments to provide known sequences for PCR amplification.

As described by Vos et al. (1995), PCR amplification will only occur where the primers are able to anneal to fragments which have the adaptor sequence plus the complementary base pairs to the additional nucleotides called selective nucleotides. An aliquot is then subjected to two subsequent PCR amplifications under highly stringent conditions with primers complementary to the adaptors, and possessing 3` selective nucleotides of 1 - 3 bases (Figure 7). The first PCR (preamplification) is performed with primer combinations containing a single-bp extension while final (selective) amplification is performed using primer pairs with up to 3-bp extension. Because of the high selectivity, primers differing by only a single base in the AFLP extension amplify a different subset of fragments. A primer extension of one, two or three bases reduces the number of amplified fragments by factors of 4, 16 and 64, respectively. Ideal primer extension lengths will vary with genome size of the species and will result in an optimal number of bands: not too many bands to cause smears or high levels of band comigration during electrophoresis, but sufficient to provide adequate polymorphism (Vos et al., 1995).

AFLP fragments are visualized either on agarose gel or on denaturing polyacrylamide gels with autoradiography,

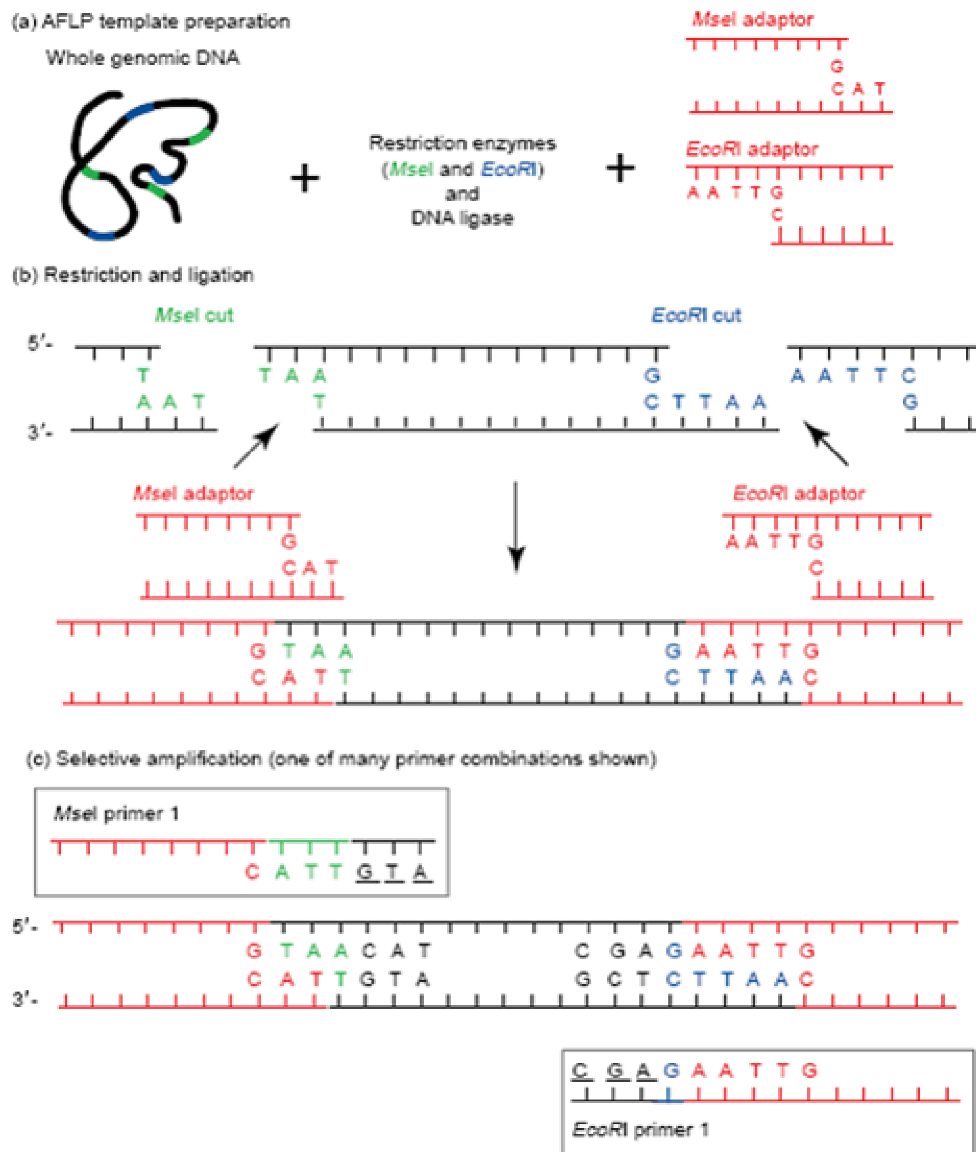

Figure 7. Steps in AFP analysis: small amounts of DNA are digested with two restriction enzymes (a), and adaptors are joined (ligated) to these ends (b). The end sequences of each adapted fragment consist of the adaptor sequence (in red) and the remaining part of the restriction sequence (in blue and green). To achieve amplification of a subset of these fragments, primers are extended into the unknown part of the fragments [underlined base pairs (bp)], usually one to three arbitrarily chosen bases beyond the restriction site (c, in black). The first PCR (preamplification) is performed with a single-bp extension, followed by a more selective primer with up to a 3-bp extension (Mueller and Wolfenbarger, 1999).

$AgNO_3$ staining (Figure 8) or automatic DNA sequencers. Polyacrylamide gel electrophoresis (PAGE) provides maximum resolution of AFLP banding patterns to the level of single-nucleotide length differences, whereas fragment length differences of less than ten nucleotides are difficult to score on agarose gels. For automatic AFLP product separation using fluorescent detection systems on DNA sequencers, such as ABI Prism, one of the selective primers must be labeled with different colored dyes (fluorophore) at the 5` end such as 6-carboxy-fluorescine (6-FAM), hexachloro-6-carboxy-fluorescine (HEX) or tetrachloro-6-carboxy-fluorescine (TET), etc. Only fragments containing a priming site complementary to the fluorophore labeled primer will be detected by the sequencers. There are four essential elements of fluorescence detection system:

a) An excitation source
b) A fluorophore
c) Wavelength filters to isolate emission photons from excitation photons.
d) A detector that registers emission photons and produces a recordable output, usually as an electrical signal or a photographic image.

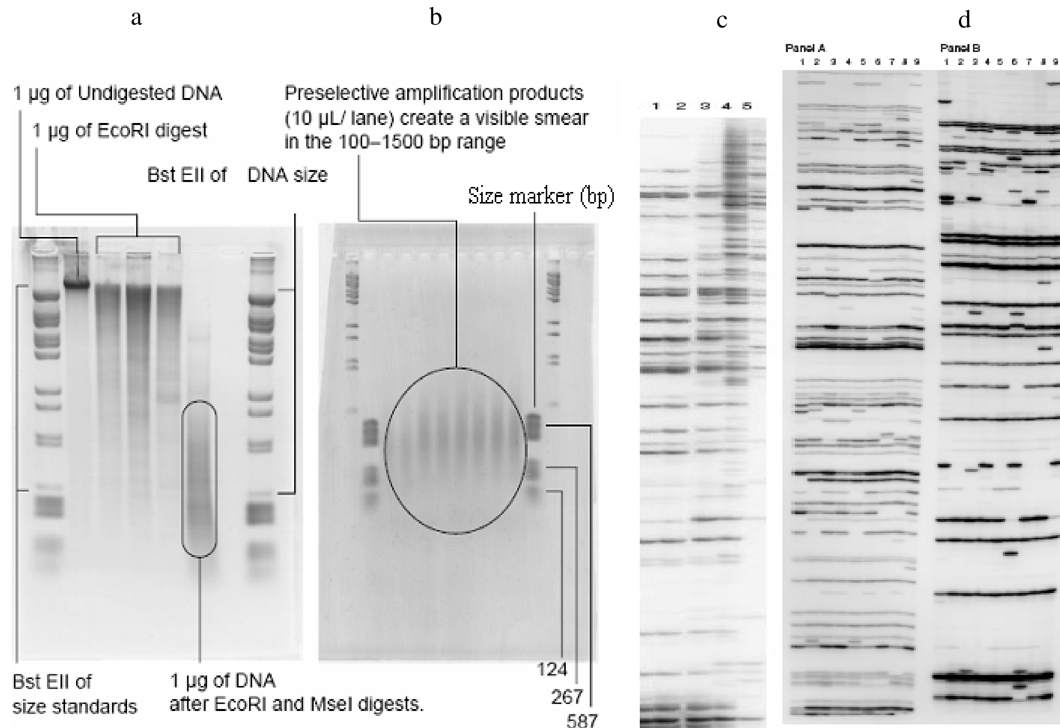

Figure 8. An example of AFLP analysis in plants. Agarose gel showing digested and undigested DNA (a), and preselective amplification products (b) for AFLP analysis; (c) the effect of partial digestion on selective amplification (the same sample was digested with five different concentrations of *EcoRI* and *MseI* enzymes, with lane 1 and lane 5 representing amplification products after digestion with the highest and lowest enzyme concentrations, respectively. Banding pattern for lanes 1 and 2 are the correct ones but not that of lanes 4 and 5); (d) selective amplification products for nine samples (1-9) and two primer pairs (E-AA/M-CAC for Panel A and E-AC/MCAC for Panel B).

Regardless of the application, compatibility of these four elements is essential for optimizing fluorescence detection. For high throughput analysis, three to nine different reactions labeled with different dyes can be multiplexed and loaded in a single lane (e.g., ABI Prism® 377 DNA Sequencer) or in a single injection (e.g., ABI Prism 3730 DNA Analyzer). An internal size standard labeled in a different color need to be loaded to estimate the size of AFLP amplification fragments using computer programs (e.g., GeneScan and GeneMapper software's from PE Applied Biosystems; http://www.app-liedbiosystems.com).

AFLP analysis is not as easy to perform as RAPDs, but is more efficient than using RFLPs. The advantages of AFLP include:

1) It is highly reliable and reproducible (Mueller et al., 1996; Lin et al., 1996; Powell et al., 1996; Jones et al., 1997).
2) It does not require any DNA sequence information from the organism under study.
3) It is information-rich due to its ability to analyze a large number of polymorphic loci simultaneously (effective multiplex ratio) with a single primer combi-nation on a single gel as compared to RAPDs, RFLPs and microsate-llites (Powell et al., 1996; Milbourne et al., 1997; Russell et al., 1997).
4) Co-migrating AFLP amplification products are mostly homologous and locus specific (vanderVoort et al., 1997; Waugh et al., 1997; Qi et al., 1998), with exceptions in polyploidy species.

In contrast to RAPD, the limitations of AFLP include:

a) It requires more number of steps to produce the result.
b) It requires template DNA free of inhibitor compounds that interferes with the restriction enzyme.
c) The technique requires the use of polyacrylamide gel in combination with $AgNO_3$ staining, radioactivity, or fluore-scent methods of detection, which will be more expensive and laborious than agarose gels.
d) It involves additional cost to purchase both restriction and ligation enzymes as well as adapters.
e) Like RAPD, most AFLP loci are dominant, which does not differentiate dominant homozygotes from heterozygotes. This reduces the accuracy of AFLP markers in population genetic analysis, genetic map-ping, and marker assisted selection.

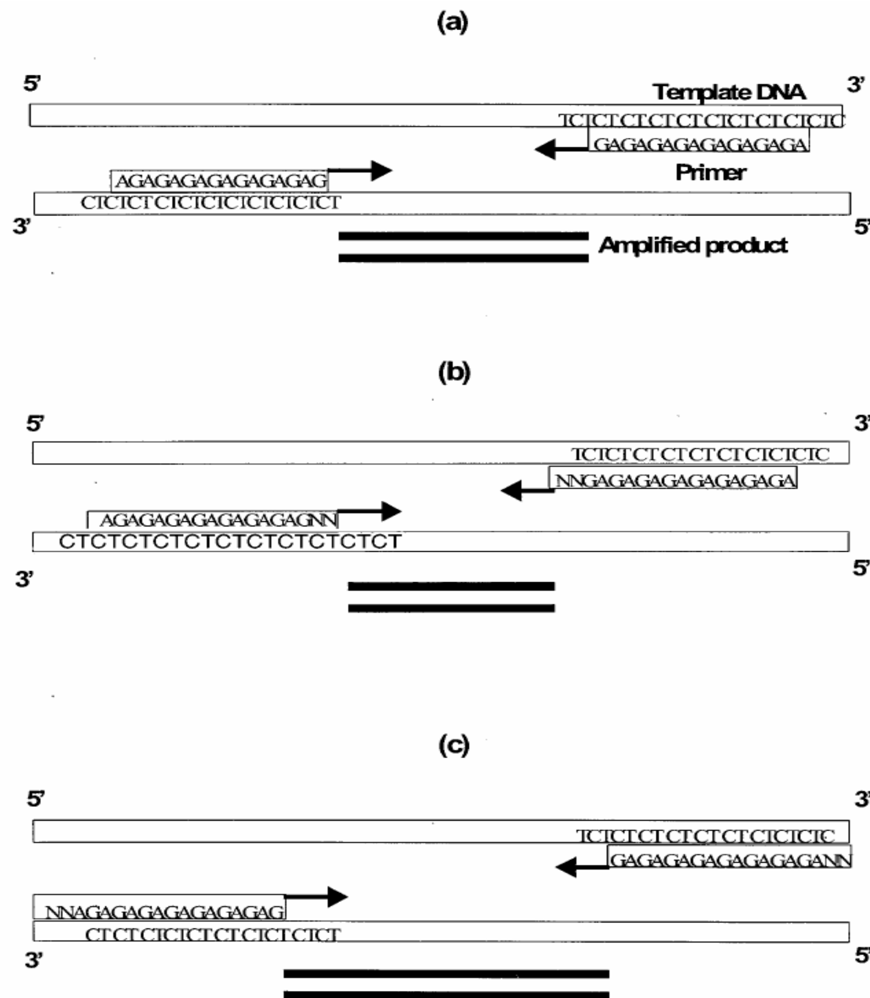

Figure 9. A schematic representation of ISSR-PCR with a single primer $(AG)_8$, unanchored (a), 3'-anchored (b) and 5'-anchored (c) targeting a $(TC)_n$ repeat used to amplify inter simple sequence repeat region flanked by two inversely oriented $(TC)_n$ sequences. (a) Unanchored $(AG)_n$ primer can anneal anywhere in the $(TC)_n$ repeat region on the template DNA leading to slippage and ultimately smear formation; (b) $(AG)_n$ primer anchored with 2 nucleotides (NN) at the 3' end anneals at specific regions on the template DNA and produces clear bands; (c) $(AG)_n$ primer anchored with 2 nucleotides (NN) at the 5' end anneals at specific regions and amplifies part of the repeat region also leading to larger bands (Reddy et al., 2002).

ISSR (inter-simple sequence repeat)

ISSR involves amplification of DNA segments present at an amplifiable distance in between two identical microsatellite repeat regions oriented in opposite direction (Figure 9). The technique uses microsatellites as primers in a single primer PCR reaction targeting multiple genomic loci to amplify mainly inter simple sequence repeats 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 (Meyer et al., 1993; Gupta et al., 1994; Wu et al., 1994) or more usually anchored at 3` or 5` end with 1 to 4 degenerate bases extended into the flanking sequences (Zietkiewicz et al., 1994; Figure 9). ISSRs use longer primers (15–30 mers) as compared to RAPD primers (10 mers), which permit the subsequent use of 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 °C. The amplified products are usually 200–2000 bp long and amenable to detection by both agarose and polyacrylamide gel electrophoresis.

ISSRs exhibit the specificity of microsatellite markers, but need no 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 usually show high polymorphism (Kojima et al., 1998) alt-

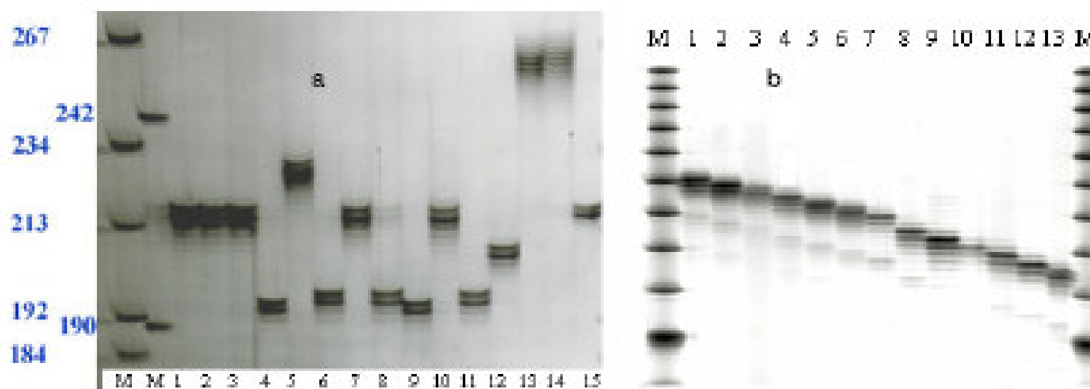

Figure 10. Microsatellites PCR products separated on polyacrylamide gels with silver nitrate staining. (a) Banding patterns of 15 rice cultivars (number 1 to 15) amplified using RM252 microsatellite primer; M is a molecular size ladder, with the size of each band indicated on the left side of the gel (http://www.gramene.org). (b) Gel showing microsatellites allelic diversity for gwm389 primer and 13 hexaploid wheat samples, with each sample possessing a different allele; M is a 10-bp molecular weight ladder and placed on either side of the gel (Liu and Anderson, 2003).

hough the level of polymorphism has been shown to vary with the detection method used. Polyacrylamide gel electrophoresis (PAGE) in combination with radioactivity was shown to be most sensitive, followed by PAGE with $AgNO_3$ staining and then agarose gel with EtBr system of detection. Like RAPDs, reproducibility, dominant inheritance and homology of comigrating amplification products are the main limitations of ISSRs. Fang and Roose (1997) reported a reproducibility level of more than 99% after performing repeatability tests for ISSR markers by using DNA samples of the same cultivar grown in different locations, DNA extracted from different aged leaves of the same individual, and by performing separate PCR runs. In other cases, the reproducibility of ISSRs amplification products ranged from 86 to 94%, with the maximum being when polyacrylamide gel electrophoresis and $AgNO_3$ staining were used and weak bands excluded from scoring (Moreno et al., 1998). ISSRs segregate mostly as dominant markers (Gupta et al., 1994; Tsumura et al., 1996; Ratnaparkhe et al., 1998; Wang et al., 1998), although co-dominant segregation has been reported in some cases (Wu et al., 1994; Akagi et al., 1996; Wang et al., 1998; Sankar and Moore, 2001). There is also a possibility as in RAPD that fragments with the same mobility originate from non-homologous regions (Sanchez et al., 1996).

Microsatellites

The genomes of higher organisms contain tree types of multiple copies of simple repetitive DNA sequences (satellite DNAs, minisatellites, and microsatellites) arranged in arrays of vastly differing size (Armour et al., 1999; Hancock, 1999). Microsatellites (Litt and Luty, 1989), also known as simple sequence repeats (SSRs; Tautz et al., 1986), short tandem repeats (STRs) or simple sequence length polymorphisms (SSLPs; McDo-nald and Potts, 1997), are the smallest class of simple repetitive DNA sequences. Some authors (e.g. Armour et al., 1999) define microsatellites as 2–8 bp repeats, others (e.g., Goldstein and Pollock, 1997) as 1–6 or even 1–5 bp repeats (Schlotterer, 1998). Chambers and MacAvoy (2000) suggested following a strict definition of 2–6 bp repeats, in line with the descriptions of the original authors. Microsatellites are born from regions in which variants of simple repetitive DNA sequence motifs are already over represented (Tautz et al., 1986). It is now well established that the predominant mutation mechanism in microsatellite tracts is ‘slipped-strand mispairing’ (Levinson and Gutman, 1987). This process has been well described by Eisen (1999). When slipped-strand mispairing occurs within a microsatellite array during DNA synthesis, it can result in the gain or loss of one, or more, repeat units depending on whether the newly synthesized DNA chain loops out or the template chain loops out, respectively. The relative propensity for either chain to loop out seems to depend in part on the sequences making up the array, and in part on whether the event occurs on the leading (continuous DNA synthesis) or lagging (discontinuous DNA synthesis) strand (Freudenreich et al., 1997). SSR allelic differences are, therefore, the results of variable numbers of repeat units within the microsatellite structure. The repeated sequence is often simple, consisting of two, three or four nucleotides (di-, tri-, and tetranucleotide repeats, respectively). One common example of a microsatellite is a dinucleotide repeat $(CA)_n$, where n refers to the total number of repeats that ranges between 10 and 100. These markers often present high levels of inter- and intra-specific polymorphism, particularly when tandem repeats number is ten or greater (Queller et al., 1993).

PCR reactions for SSRs is run in the presence of forw-

ard and reverse primers that anneal at the 5` and 3` ends of the template DNA, respectively. PCR fragments are usually separated on polyacrylamide gels in combination with $AgNO_3$ staining (Figure 10), autoradiography or fluorescent detection systems. Agarose gels (usually 3%) with EtBr can also be used when differences in allele size among samples is larger than 10 bp. However, the establishment of microsatellite primers from scratch for a new species presents a considerable technical challenge. Several protocols have been developed (Bruford et al., 1996; McDonald and Potts, 1997; Hammond et al., 1998; Schlotterer, 1998) and details of the methodologies are reviewed by different authors (e.g., Chambers and MacAvoy, 2000; Zane et al., 2002; Squirrell et al., 2003). A review by Zane et al. (2002) describes some of the technical advances that have been made in recent years to facilitate microsatellite development. They cover a range of methods for obtaining sequences rich in microsatellite repeats (some of which can be undertaken in a matter of days), and also highlight the availability of companies who will undertake the construction of enriched microsatellite libraries as a commercial service. The development of microsatellite markers involves several distinct steps from obtaining the library to developing a working set of primers that can amplify polymorphic microsatellite loci. These include:

1) Microsatellite library construction.
2) Identification of unique microsatellite loci.
3) Identifying a suitable area for primer design.
4) Obtaining a PCR product.
5) Evaluation and interpretation of banding patterns.
6) Assessing PCR products for polymorphism (Roder et al., 1998).

SSR primers are developed by cloning random segments of DNA from the target species. These are inserted into a cloning vector, which is in turn, implanted into *Escherichia coli* bacteria for replication. Colonies are then developed, and screened with single or mixed simple sequence oligonucleotide probes that will hybridize to a microsatellite repeat, if present on the DNA segment. If positive clones for microsatellite are obtained from this procedure, the DNA is sequenced and PCR primers are chosen from sequences flanking such regions to determine a specific locus. This process involves significant trial and error on the part of researchers, as microsatellite repeat sequences must be predicted and primers that are randomly isolated may not display polymorphism (Queller et al., 1993; Jarne and Lagoda, 1996).

The next step is to select the best candidate markers and then to optimize conditions for their amplification. Optimization of microsatellite systems involves a more or less comprehensive survey of PCR conditions for amplification of candidate loci. The objective here is to adequately balance the often conflicting requirements for high specificity and high intensity of amplification products. Thus, the issue of signal strength and purity remains the primary focus. Other considerations include obtaining products from various loci with non-overlapping ranges of allele sizes, which can be amplified with similar efficiency under a standard set of conditions and enables multiplexing for high throughput analysis (Schlotterer, 1998). Microsatellite loci are more common in some organisms than in others, and screening may produce few useful loci in some species (Cooper, 1995). The efficiency of microsatellite marker development depends on the abundance of repeats in the target species and the ease with which these repeats can be developed into informative markers. When researchers are isolating plant microsatellites, about 30% of the sequenced clones, on average, can be lost due to the absence of unique microsatellites. Of those sequences that contain unique microsatellites, a number of the clones in a library can contain identical sequences (and hence there is a level of redundancy) and/or chimeric sequences (i.e., one of the flanking regions matches that of another clone). At each stage of SSR development, therefore, there is the potential to ‘lose’ loci, and hence the number of loci that will finally constitute the working primer set will be a fraction of the original number of clones sequenced (Squirrell et al., 2003). The conversion of microsatellite-containing sequences into useful markers can be quite difficult, especially in species with large genomes (Smith and Devey, 1994; Kostia et al., 1995; Roder et al., 1995; Pfeiffer et al., 1997; Song et al., 2002). The low conversion rates of primer pairs to useful markers in these species are due to the high level of repetitive DNA sequences in their genomes. The recovery rate for useful SSR primers is generally low due to different reasons:

a) The primer may not amplify any PCR product.
b) The primer may produce very complex, weak or non-specific amplification patterns.
c) The amplification product may not be polymorphic.

Investigators often prefer to work with loci containing tri- and tetra-nucleotide repeat arrays rather than di-nucleotide arrays because the former frequently give fewer “stutter bands” (multiple near-identical ‘ladders’ of PCR products which are one or two nucleotides shorter or longer than the full length product; Figure 10 (Hearne et al., 1992; Diwan and Cregan, 1997). Thus, allele sizing is less error prone using tri- and tetra-nucleotide repeats than di-nucleotide repeats (Diwan and Cregan, 1997). However, this idea must be balanced against practical considerations. Di-nucleotide repeat arrays occur much more frequently than tri- or tetra-nucleotide repeat arrays, and it is easier to run combinatorial screens for them.

SSRs are now the marker of choice in most areas of molecular genetics as they are highly polymorphic even between closely related lines (Figure 10), require low amount of DNA, can be easily automated for high throughput screening, can be exchanged between laboratories, and are highly transferable between populations

(Gupta et al., 1999). For example, a total of 18,828 SSR sequences have been detected in the rice genome (The Rice Genome Mapping project, 2005), of which only 10 - 15% have yet been used, suggesting the high potential available for such marker systems. SSRs are mostly co-dominant markers, and are indeed excellent for studies of population genetics and mapping (Jarne and Lagoda, 1996; Goldstein and Schlotterer, 1999). The use of fluorescent primers in combination with automatic capillary or gel-based DNA sequencers has got its way in most advanced laboratories and SSR are excellent markers for fluorescent techniques, multiplexing and high throughput analysis.

The major constraint of using SSR markers from genomic libraries is the high development cost and effort required to obtain working primers for a given study species. This has restricted their use to only a few of the agriculturally important crops. A more widespread use of genomic SSRs in plants would also be facilitated if such loci were transferable across species. Recently, a new alternative source of SSRs development from expressed sequence tag (EST) databases has been utilized (Kota et al., 2001; Kantety et al., 2002; Michalek et al., 2002). With the availability of large numbers of ESTs and other DNA sequence data, development of EST-based SSR markers through data mining has become a fast, efficient, and relatively inexpensive compared with the development of genomic SSRs (Gupta et al., 2003). This is due to the fact that the time-consuming and expensive processes of generating genomic libraries and sequencing of large numbers of clones for finding the SSR containing DNA regions are not needed in this approach (Eujayl et al., 2004). However, the development of EST-SSRs is limited to species for which this type of database exists. Furthermore, the EST-SSR markers have been reported to have 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).

Differences in SSR allele size is often difficult to resolve on agarose gels and high resolutions can be achieved through the use of polyacrylamide gels in combination with $AgNO_3$ staining. The cost of polyacrylamide gels is higher than agarose gels and it is not also as rapid as the latter. The establishment and running cost for an automatic DNA sequencer is not affordable for researchers at the national research systems and universities in developing countries. The other technical problem with microsatellites is the fact that it is not always possible to compare data produced by different laboratories, due to the eventuality of inconsistencies in allele size calling. Such inconsistencies are mainly due to the large variety of automatic sequencing machines used, each providing different gel migration, fluorescent dyes, allele calling software's, and PCR reaction. For the later, the enzyme used for DNA synthesis (Taq DNA polymerase) catalyses the addition of an extra base (usually an adenine) at the end of the PCR product. The proportion of fragments with this extra base may vary from none to 100%, inducing one base pair size differences and complicating data analysis. Although biochemical treatments after PCR or modification of PCR primers can circumvent this problem (Brownstein et al., 1996; Ginot et al., 1996), they are seldom used.

EST (expressed sequence tags)

Each gene must be converted or transcribed into messenger RNA (mRNA) that serves as a template for protein synthesis. The resulting mRNA then guides the synthesis of a protein through a process called translation. The problem, however, is that mRNA is very unstable outside of a cell; therefore, 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 the procedure uses mRNA as a template rather than DNA. cDNA is a much more stable compound and it represents only expressed DNA sequence because it is generated from mRNA that represents exons by excising (splicing) introns. Once cDNA representing an expressed gene has been isolated, scientists can then sequence a few hundred 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 therefore tend to exhibit less cross-species conservation than do coding sequences. The challenge associated with identifying genes from genomic sequences varies among organisms and is dependent upon genome size as well as the presence or absence of introns, which are the intervening DNA sequences interrupting the protein coding sequence of a gene.

The production of ESTs starts with the construction of cDNA libraries. The identification of ESTs has proceeded rapidly, with over 6 million ESTs now available in computerized databases. ESTs were originally intended as a way to identify gene transcripts, but have since been 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 have been used for designing probes for DNA microarrays that is used to determine gene expression. ESTs also allow the efficient development of single or low-copy RFLP markers. RFLP markers developed from ESTs (EST-RFLP) have been extensively used for the construction of high-density genetic linkage maps (e.g., Harushima et al., 1998; Davis et al., 1999) and physical maps (e.g., Kurata et al., 1997). 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.

ESTs also allow a computational approach to the development of SSR and SNP markers (Eujayl et al., 2001; Cho et al., 2000) for which previous development strategies have been expensive. Pattern-finding programs can be employed to identify SSRs in ESTs. The available sequence information allows the design of primer pairs, which can be used to screen cultivars of interest for length polymorphisms. A modest 1% to 5% of the ESTs in various plant species have been found to have SSRs of suitable length (20 bp or more) for marker development (Kantety et al., 2002). It should be possible to find a large number of these SSRs in an organism for which a great number of ESTs have been generated. For example, 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% ESTs contained SSRs. EST-SSRs are generally anchored within more conserved transcribed regions across species than those from the untranscribed regions (Caudrado and Schwarzacher, 1998), and hence they are expected to be more transferable to closely related genera (Cordeiro et al., 2001; Hempel and Peakall, 2003; Decroocq et al., 2003). EST-SSRs also have a higher probability of being functionally associated with differences in gene expression than the genomic SSRs (Gao et al., 2004). 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; Thiel et al., 2003; Varshney et al., 2002), although ESTs of a few dicot species have been explored for SSR mining (Eujayl et al. 2004; Morgante et al. 2002; Qureshi et al. 2004; Saha et al. 2003; Scott et al., 2000; 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.

CAPS (cleaved amplified polymorphic sequence)

CAPS is a combination of the PCR and RFLP, and it was originally named PCR-RFLP (Maeda et al., 1990). The technique involves amplification of a target DNA through PCR, followed by digesting with restriction enzymes (Konieczny and Ausubel, 1993; Jarvis et al., 1994; Michaels and Amasino, 1998). Hence, CAPS markers rely on differences in restriction enzyme digestion patterns of PCR fragments caused by nucleotide polymerphism between samples. Critical steps in the CAPS marker approach include DNA extraction, PCR conditions, and the number or distribution of polymorphic sites. CAPS markers have several advantages. First, since analysis of restriction fragment length polymorphisms is based on PCR amplification, it is much easier and less time-consuming than analyzing alternative types of markers that require southern hybridizations. Second, CAPS primers developed from ESTs are more useful as genetic markers for comparative mapping study than those markers derived from non-functional sequences such as genomic microsatellite markers. Third, CAPS markers are inherited mainly in a co-dominant manner (Matsumoto and Tsumura, 2004). However, the ability of CAPS to detect DNA polymorphism is not as high as SSRs and AFLPs because nucleotide changes affecting restriction sites are essential for the detection of DNA polymorphism by CAPS. Furthermore, the development of CAPS markers is only possible where mutations disrupt or create a restriction enzyme recognition site. Some researchers developed an alternative marker called derived-CAPS (dCAPS) that eliminate the problems related with CAPS markers by generating mismatches in a PCR primer, which are subsequently used to create a polymorphism based on the target mutation (Michaels and Amasino, 1998; Neff et al., 1998).

SCAR (sequence characterized amplified region)

A SCAR marker is a genomic DNA fragment that is identified by PCR amplification using a pair of specific oligonucleotide primers (Paran and Michelmore, 1993; McDermott et al., 1994). 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). SCARs are advantageous over RAPD markers as they detect only a single locus, their amplification is less sensitive to reaction conditions, and they can potentially be converted into codominant markers (Paran and Michelmore, 1993).

STS (sequence tagged site)

STS was first developed by Olsen et al. (1989) as DNA landmarks in the physical mapping of the human genome, and latter adopted in plants. STS is a short, unique sequence whose exact sequence is found nowhere else in the genome. Two or more clones containing the same STS must overlap and the overlap must include STS. Any clone that can be sequenced may be used as STS provided it contains a unique sequence. In plants, STS is characterized by a pair of PCR primers that are designed by sequencing either an RFLP probe

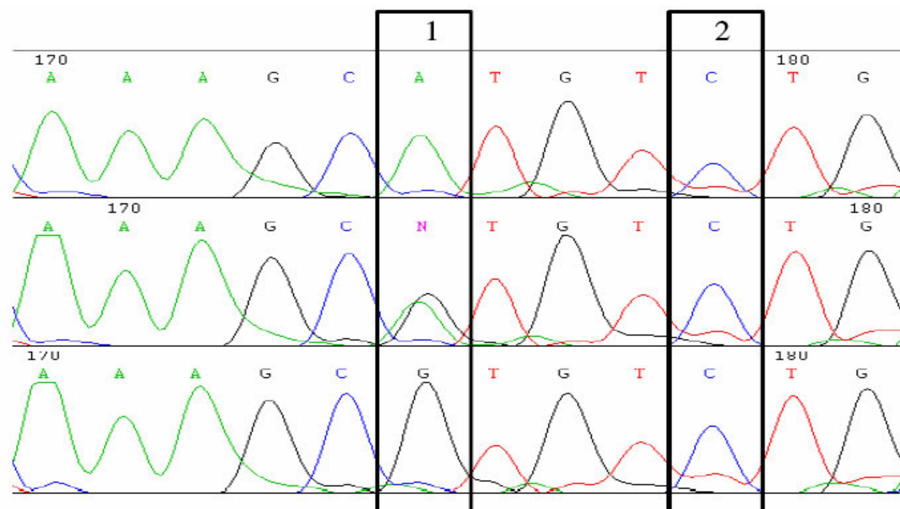

Figure 11. SNP discovery by alignment of sequence traces obtained from direct sequencing of genomic PCR products. It is not always possible to distinguish between sequence artefacts and true polymorphism, when two peaks are present at one position. Box 1: top sequence homozygote AA, middle sequence heterozygote AG, bottom sequence homozygote GG. Box 2: The polymorphism detection software has considered the top and bottom sequences as heterozygote CT and the middle one as homozygote CC. Clonal sequence removes many of such ambiguities, since any double peak is a sequence artefact (Vignal et al., 2002).

representing a mapped low copy number sequence (Blake et al., 1996) or AFLP fragments. Although conversion of AFLP markers into STS markers is a technical challenge and often frustrating in polyploids such as hexaploid wheat (Shan et al., 1999; Prins et al., 2001), it has been successful in several crops (Meksem et al., 1995, 2001; Qu et al., 1998; Shan et al., 1999; Decousset et al., 2000; Parker and Langridge, 2000; Prins et al., 2001; Guo et al., 2003). The primers designed on the basis of a RAPD have also sometimes been referred to as STSs (Naik et al., 1998), although they should be more appropriately called SCARs. STS markers are codominant, highly reproducible, suitable for high throughput and automa-tion, and technically simple for use (Reamon-Buttner and Jung, 2000).

SNP (single nucleotide polymorphism)

Public accessibility to the genome sequences of several organisms has enabled the study of sequence variations between individuals, cultivars, and subspecies. These studies revealed that single nucleotide polymorphisms (SNPs) and insertions and deletions (InDels) are highly abundant and distributed throughout the genome in various species including plants (Garg et al., 1999; Drenkard et al., 2000; Nasu et al., 2002; Batley et al., 2003a). By comparing sequences from a japonica rice cultivar to those from an indica cultivar, for example, Yu et al. (2002) identified, on average, one SNP every 170 bp and one InDel every 540 bp. The abundance of these polymorphisms in plant genomes makes the SNP marker system an attractive tool for mapping, marker-assisted breeding and map-based cloning (Gupta et al., 2001; Rafalski, 2002a; Batley et al., 2003b). As suggested by the acronym, a SNP marker is just a single base change in a DNA sequence, with a usual alternative of two possible nucleotides at a given position (Figure 11). Hence, in contrast to all previous methods, allele discrimination can not be based on size differences on a gel. Over the past years, a large number of different SNP genotyping methods and chemistries have been developed based on various methods of allelic discrimination and detection platforms (see Rafalski, 2002b; Vignal et al., 2002; Sobrino et al., 2005; Tost and Gut, 2005 for detailed review). All methods for SNP genotyping combine two elements: first, the generation of an allele-specific product, and second the analysis thereof.

Vignal et al. (2002) classified the various SNP detection methods into two broad groups:

1) Direct hybridization techniques.
2) Those techniques that involve the generation and separation of an allele-specific product (restriction enzyme cutting, single strand DNA conformation and hetero-duplexes, primer extension, oligonucleotide ligation assay, pyrosequencing, exonuclease detection or Taq-Man, invasive cleavage of oligonucleotide probes or invader assay).

Sobrino et al. (2005) assigned the majority of SNP genotyping assays to one of four groups based on molecular mechanism: allele specific hybridization, primer extension, oligonucleotide ligation and invasive cleavage (Figure 12). Allele specific hybridization, also known as allele specific oligonucleotide hybridization (ASO), is based on distinguishing between two DNA targets differring at one nucleotide position by hybridization (Wallace et al., 1979). Two allele-specific probes are designed,

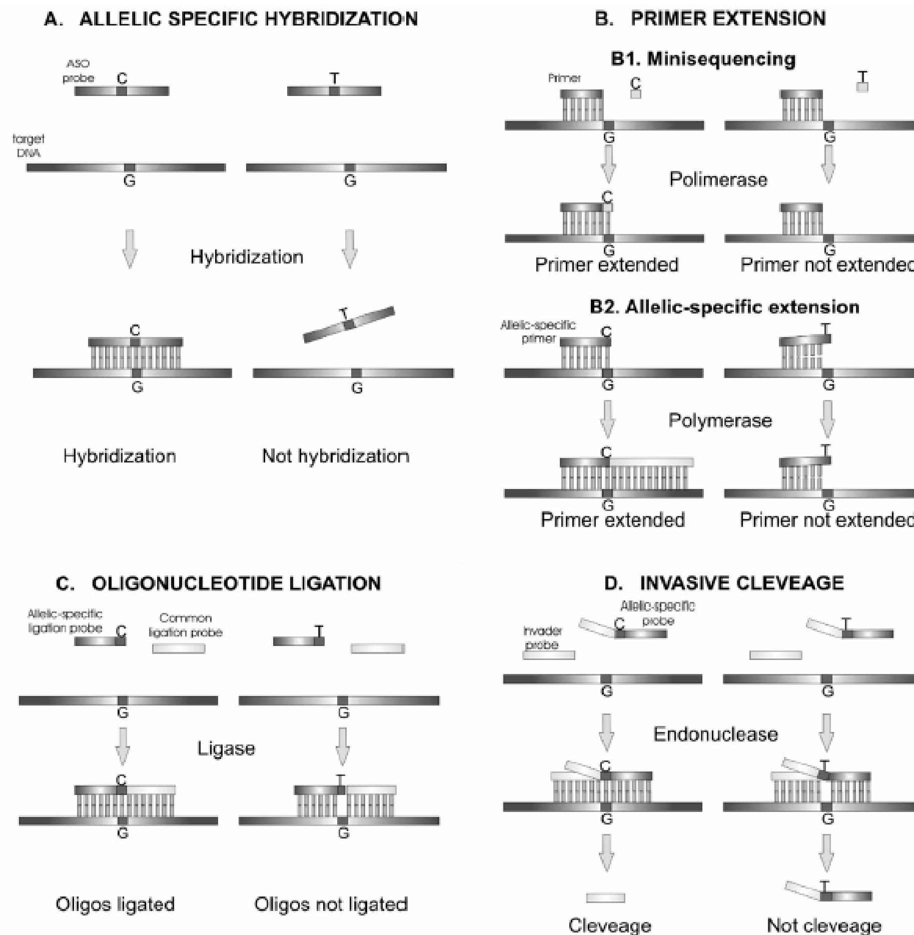

Figure 12. Illustration of the four allelic discrimination reactions. (A) Hybridization with allelic specific oligonucleotides (ASO): two ASO probes are hybridized with the target DNA that contains the SNP. Under optimized conditions, only the perfectly matched probe-target hybrids are stable. (B) Primer extension reactions: minisequencing (B1) and allelic-specific extension (B2). (B1) minisequencing: a primer anneals to its target DNA immediately upstream to the SNP and is extended with a single nucleotide complementary to the polymorphic base. (B2) allelic-specific extension: the 3' end of the primers is complementary to each allele of the SNP. When there is a perfect match the primer is extended. (C) oligonucleotide ligation assay (OLA): two allelic-specific probes and one common ligation probe are required per SNP. The common ligation probe hybridized adjacent to the allelic-specific probe. When there is a perfect match of the allelic-specific probe, the ligase joins both allelic-specific and common probes. (D) Invasive cleavage: the oligonucleotides required called invader probe and allelic-specific probes, anneal to the target DNA with an overlap of one nucleotide. When the allelic-specific probe is complementary to the polymorphic base, overlaps the 3' end of the invader oligonucleotide, forming the structure that is recognized and cleaved by the Flap endonuclease, releasing the 5' arm of the allelic-specific probe (Sobrino et al., 2005).

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 are unstable. Most hybridization techniques are derived from the Dot Blot, in which DNA to be tested (either genomic, cDNA or a PCR reaction) is fixed on a membrane and hybridized with a probe, usually an oligonucleotide. In the Reverse Dot Blot technique, it is the oligonucleotide probes that are immobilised. When using allele specific oligonucleotides, genotypes can be inferred from hybridization signals. However, hybridization techniques are error prone and need carefully designed probes and hybridization protocols (Pastinen et al., 1997). The latest improvements of this family of techniques, is the 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 synthesised 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, high sensitivity and high throughput.

Primer extension is based on the ability of DNA polymerase to incorporate specific deoxyribonucleotides complementary to the sequence of the template DNA. There are several variations of the primer extension reaction, which can be divided into three main types of react-

Table 2. The different detection methods that can be used for each allelic discrimination reaction (Sobrino et al., 2005).

Allelic Discrimination Reaction \ Detection Method	Electrophoresis (fluorescence)	FRET*	Fluorescence polarization	Arrays (fluorescence)	Mass spectrometry	Luminescence
Allelic-specific hybridization		✓	✓	✓		
Primer extension	✓	✓	✓	✓	✓	✓
Oligonucleotide Ligation	✓	✓		✓		
Invasive cleavage		✓	✓		✓	

*FRET: fluorescence resonance energy transfer.

ion:

a) The minisequencing 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.
c) Pyrosequencing.

Allele specific oligonucleotide ligation is a method for SNP typing based on the ability of ligase to covalently join two oligonucleotides when they hybridize next to one another on a DNA template. The invader assay is based on the specificity of recognition, and cleavage, by a flap endonuclease, of the three-dimensional structure formed when two overlapping oligonucleotides hybridize perfectly to a target DNA (Kaiser et al., 1999; Lyamichev et al., 1999).

There are several detection methods for analyzing the products of each type of allelic discrimination reaction (Table 2; Figure 13): gel electrophoresis, fluorescence resonance energy transfer (FRET), fluorescence polarization, arrays or chips, luminescence, mass spectrophotometry (Matrix-assisted laser desorption/ionisation time-of-flight mass spectrometry or MALDI-TOF), chromatography, etc. There are two different categories related with the assay format or detection: homogenous reactions, when they occur in solution, and reactions on solid support such as glass slide, a chip, a bead, etc. In technologies based on homogenous hybridization with FRET detection (such as the Light Cycler, TaqMan and Molecular Beacon), PCR and allelic discrimination reaction are performed in the same reaction. This advantage avoids further manipulation steps, favoring the automation and throughput of the process, especially when the high-throughput equipments for TaqMan assays are used. In general, homogenous reactions are more amenable to automation because there are not separation or purification steps after the allele discrimination reaction. However, the major drawback is the limited multiplex capability. In contrast, reactions on solid-support have greater multiplex capability but further manipulations are required.

A large number of different SNP typing protocols are available for researchers, and there is no single protocol that meets all research needs. Different aspects should be taken into account to determine the best suitable technology in terms of sensitivity, reproducibility, accuracy, capability of multiplexing for high throughput analysis, 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. It is difficult to predict if one technique will emerge in the future as a standard, especially since the needs will vary quite a lot between the academic laboratory performing medium-scale studies and comercial companies or genome centers aiming at very high throughput. The selection of technique is also dictated by the type of project envisaged, since it is quite different to perform genotypes with a limited number of SNPs on very large population samples, or a large number of SNPs on a limited number of individuals. For studies involving large sets of samples, the use of primer extension techniques analyzed by MALDI-TOF mass spectrometry hold high promises in terms of automation

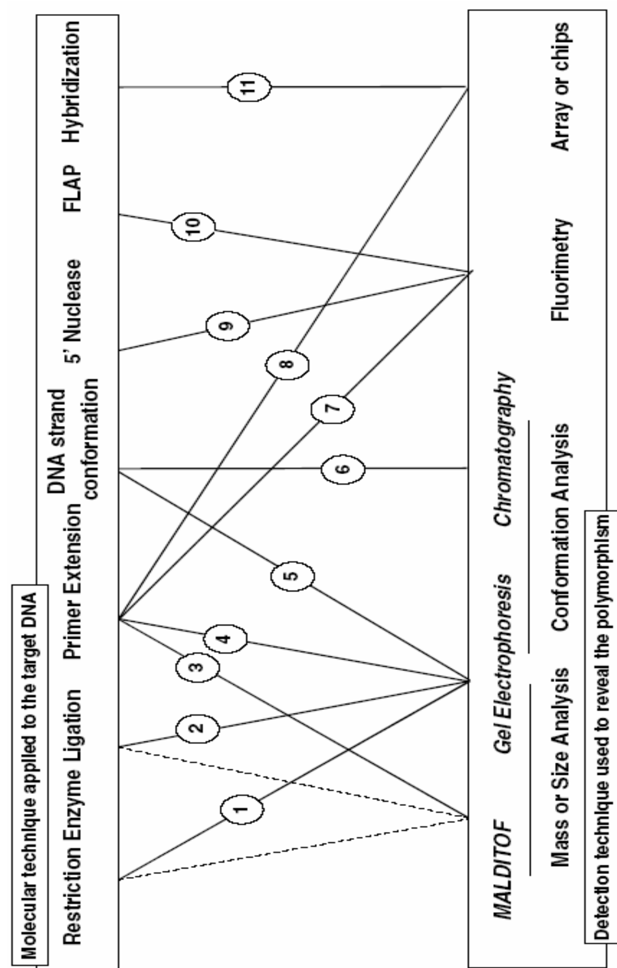

Figure 13. Links between the signal generation and detection. Many of the products generated by the allele-specific reactions can be detected with different methods. 1: PCR-RFLP, 2: Oligonucleotide ligation assay (OLA), 3: Good Assay, 4: Minisequencing techniques, 5: single stranded conformation polymorphism (SSCP) or denaturing gradient gel electrophoresis (DGGE), 6: denaturing high performance liquid chromatography (DHPLC), 7: Pyrosequencing, 8: SNP it, 9: exonuclease detection (Taqman), 10: Invader Assay, 11: Microarray or DNA chips (Vignal et al., 2002).

accuracy, throughput and price (Tost and Gut, 2002). Mass spectrometry-based methods for SNP genotyping have been continuously improved and matrix-assisted laser desorption/ionization mass spectrometry (MALDI-MS) is now one of the most automated and efficient detection platforms, price competitive when used at high throughput, and deliver results of highest accuracy and reliability (Tost and Gut, 2002). Mass spectrometry assays based on primer extension reactions out-compete most other methods in terms of robustness, accuracy, reproducibility, and success rates (Le Hellard et al., 2002; Luo et al., 2004). The same platform can also be applied to the analysis of DNA methylation (Tost et al., 2003), expression profiling (Ding and Cantor, 2003), and proteomics (Aebersold and Mann, 2003), making the mass spectrometer one of the most versatile tools in the post-genome sequencing era. The four allele discrimination methods (hybridization, ligation, cleavage, and primer extension) have been combined with MALDI analysis. The interpretation of the mass spectrometric data is largely automated and comes as part of the software with most of the commercially available mass spectrometers. For DNA analysis, peak identification (Figure 11) is carried out followed by labeling and assignment. Pyrosequencing is also a very promising technique, with prices and throughput that might reach those of MALDI-TOF. However, pyrosequencing has limitation for automation with very low multiplex capability because several steps need to be performed before the detection. The main advantage is the possibility of quantification, the contribution of each allele, a very useful feature in the analysis of mixtures profiles (Sobrino et al., 2005).

DArT (diversity arrays technology)

DArT is one of the recently developed molecular techniques and it has only been used in rice (Jaccoud et al., 2001), barley (Wenzl et al., 2004), eucalyptus (Lezar et al., 2004), Arabidopsis (Wittenberg et al., 2005), cassava (Xia et al., 2005), wheat (Akbari et al., 2006; Semagn et al., 2006), and pigeon-pea (Yang et al., 2006). The inventors promote it as 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 (www.cambia.org or http://www.diversityarrays.com for information). DArT is a microarray hybridization-based technique that enables the simultaneous typing of several hundred polymorphic loci spread over the genome (Jaccoud et al., 2001; Wenzl et al., 2004). Details of the methodology for DArT was first described by Jaccoud et al. (2001; Figure 14). For each individual DNA sample being typed, genomic representations are prepared by restriction enzyme (e.g., PstI and TaqI) digestion of genomic DNA followed by ligation of restriction fragments to adapters. The genome complexity is then reduced by PCR using primers with complementary sequences to the adapter and selective overhangs. The fragments from representations are cloned, and cloned inserts are 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.

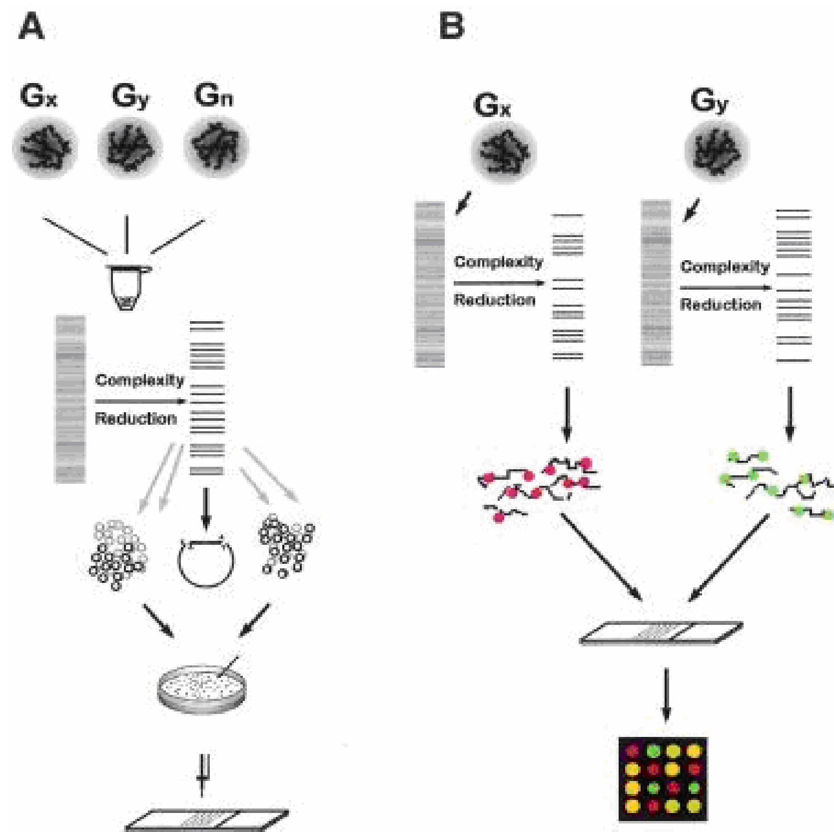

Figure 14. Schematic representation of DArT. (A) Generation of diversity panels. Genomic DNAs of specimens to be studied are pooled together. The DNA is cut with a chosen restriction enzyme and ligated to adapters. The genome complexity is reduced in this case by PCR using primers with selective overhangs. The fragments from representations are cloned. Cloned inserts are amplified using vector-specific primers, purified and arrayed onto a solid support. (B) Contrasting two samples using DArT. Two genomic samples are converted to representations using the same methods as in (A). Each representation is labeled with a green or red fluorescent dye, mixed and hybridized to the diversity panel. The ratio of green:red signal intensity is measured at each array feature. Significant differences in the signal ratio indicate array elements (and the relevant fragment of the genome) for which the two samples differ (Jaccoud et al., 2001).

DArT technique has a number of advantages:

I. It does not need prior sequence information for the species to be studied; this makes the method applicable to all species regardless of how much DNA sequence information is available for that species.
II. It is a high throughput, quick, and highly reproducible method.
III. It is cost effective, with an estimated cost per data point tenfold lower than SSR markers (Xia et al., 2005).
IV. The genetic scope of analysis is defined by the user and easily expandable.
V. It is not covered by exclusive patent rights, but on the contrary open-source (i.e., it is designed for open use and shared improvement).

This technique, however, has also its own limitations:

I. DArT is a microarray-based technique that involves several steps, including preparation of genomic representation for the target species, cloning, and data management and analysis. The latter requires dedicated software's such as DArTsoft and DArTdb. The establishment of DArT system, therefore, is highly likely to demand an extensive investment both in laboratory facility and skilled manpower.
II. DArT assays for the presence (or amount) of a specific DNA fragment in a representation. Hence, DArT markers are primarily dominant (present or absent) or differences in intensity, which limits its value in some applications.
III. The technology has been used in few species primarily by the team that developed it (who has setup a quite economical commercial service for some species); only a single independently group has so far successfully established the methodology to *Eucalyptus grandis* in South Africa (Lezar et al., 2004).

Which Markers For Which Purpose?

In the previous sections, we have reviewed the principle and methodology of 11 different types of molecular markers that have been used for different purposes in plants.

Table 3. Comparison of the five most widely used DNA markers in plants.

	RFLP	Microsatellite	RAPD	AFLP	ISSR
Genomic abundance	high	medium	very high	very high	medium
Part of genome surveyed	low copy coding regions	whole genome	whole genome	whole genome	whole genome
Amount of DNA required	high	low	low	medium	low
Type of polymorphism	single base changes, insertion, deletion	changes in length of repeats	single base changes, insertion, deletion	single base changes, insertion, deletion	single base changes, insertion, deletion
Level of polymorphism[a]	medium	high	high	very high	high
Effective multiplex ratio[b]	low	medium	medium	high	medium
Marker index[c]	low	medium	medium	high	medium
Inheritance	codominant	codominant	dominant	dominant	dominant
Detection of alleles	yes	yes	no	no	no
Ease of use	labour intensive	easy	easy	difficult initially	easy
Automation	low	high	medium	medium	medium
Reproducibility (reliability)	high	high	intermediate	high	medium to high
Type of probes/primers	low copy genomic DNA or cDNA clones	specific repeat DNA sequence	usually 10 bp random nucleotides	specific sequence	specific repeat DNA sequence
Cloning and/or sequencing	yes	yes	no	no	no
Radioactive detection	usually yes	no	no	yes/no	no
Development/start-up costs	high	high	low	medium	medium
Utility for genetic mapping	species specific	species specific	cross specific	cross specific	cross specific
Proprietary rights status	No	No (some are licensed)	licensed	licensed	no

[a] Level of polymorphism (average heterozygosity) is an average of the probability that two alleles taken at random can be distinguished
[b] Effective multiplex ratio is the number of polymorphic loci analysed per experiment in the germplasm tested.
[c] Marker index is the product of the average expected heterozygosity and the effective multiplex ratio.

The main challenges for researchers would, therefore, be selecting one or more of these markers for their specific project. The desirable properties of molecular markers are high polymorphism, codominant inheritance, frequent occurrence and even distribution throughout the genome, selectively neutral behavior, easy access, easy and fast assay, low cost and high throughput, high reproducibility, and transferability between laboratories, populations and/or species. No molecular markers are available yet that fulfill all these requirements. However, according to the kind of study to be undertaken, one can choose among the variety of molecular marker systems, each of which combines at least some of these desirable properties. A number of factors need to be considered in choosing one or more of the various molecular marker types:

a) Marker system availability.
b) Simplicity of the technique and time availability.
c) Anticipated level of polymorphism in the population.
d) Quantity and quality of DNA available.
e) Transferability between laboratories, populations, pedigrees and species.
f) The size and structure of the population to be studied
g) Availability of adequate skills and equipment
h) Cost per data-point and availability of sufficient funding.
i) Marker inheritance (dominant versus codominant) and the type of genetic information sought in the population (Staub and Serquen, 1996; Karp et al., 1997; Wolfe and Liston, 1998; Mackay, 2001; Rungis et al., 2005).

DArT is a recent technique and it remains to be thoroughly tested in various species. SCAR and STS markers would be developed by sequencing fragments associated with economically important traits and they are not available if one starts from scratch. SNPs seem very exciting markers but they require extensive investment in equipment and manpower. Hence, SNPs are highly 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. Table 3 compares the pros and cons of these markers.

REFERENCES

Adams MD, Kelley JM, Gocayne JD, Dubrick M, Polymeropoulos MH, Xiao H, Merril CR, Wu A, Olde B, Moreno RF, Kerlavage AR, McCombie WR, Venter JC (1991). Complementary DNA sequencing: expressed sequence tags and human genome project. Science 252:1651–1656.

Adams RP, Demeke T (1993). Systematic relationships in Juniperus based on random amplified polymorphic DNA (RAPDs). Taxon 42:553-571.

Aebersold R, Mann M (2003). Mass spectrometry-based proteomics. Nature 422:198–207.

Akagi H, Yokozeki Y, Inagaki A, Nakamura A, Fujimura T (1996). A co-dominant DNA marker closely linked to the rice nuclear restorer gene, Rf-1, identified with inter-SSR fingerprinting. Genome 39:1205-1209.

Akbari M, Wenzl P, Caig V, Carlig J, Xia L, Yang S, Uszynski G, Mohler V, ehmensiek A, Howes N, Sharp P, Huttner E, Kilian A (2006). Diversity arrays technology (DArT) for highthroughput profiling of the hexaploid wheat genome. Theor. Appl. Genet.113:1409-1420.

Akkaya MS, Bhagwat AA, Cregan PB (1992). Length polymorphisms of simple sequence repeat DNA in soybean. Genetics 132:1131-1139.

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.

Armour JAL, Alegre SA, Miles S, Williams LJ, Badge RM (1999). Minisatellites and mutation processes in tandemly repetitive DNA. In: Goldstein DB, Schlotterer C (eds) Microsatellites: evolution and applications. Oxford University Press, Oxford, pp. 24–33.

Bassam BJ, Caetano-Anolles G, Gresshoff PM (1992). DNA amplification fingerprinting of bacteria. Appl. Microbiol. Biotechnol. 38:70-76.

Batley J, Barker G, O'Sullivan H, Edwards KJ, Edwards D (2003a). Mining for single nucleotide polymorphisms and insertions/deletions in maize expressed sequence tag data. Plant Physiol. 132:84–91.

Batley J, Mogg R, Edwards D, O'Sullivan H, Edwards KJ (2003b). A high-throughput SNuPE assay for genotyping SNPs in the flanking regions of Zea mays sequence tagged simple sequence repeats. Mol. Breed. 11:111–120.

Beckmann JS (1988). Oligonucleotide polymorphisms: A new tool for genomic genetics. Biotechnol. 6: 161-164.

Beckmann JS, Soller M (1990). Toward a unified approach to genetic mapping of eukaryotes based on sequence tagged microsatellite sites. Bio/Technology 8: 930-932.

Blake TK, Kadyrzhanova D, Shepherd KW, Islam AKMR, Langridge PL, McDonald CL, Erpelding J, Larson S, Blake NK, Talbert LE (1996). STS-PCR markers appropriate for wheat-barley introgression. Theor. Appl. Genet. 82:715-721.

Blears MJ, De Grandis SA, Lee H, Trevors JT (1998). Amplified fragment length polymorphism (AFLP): a review of the procedure and its applications. J. Industrial Microbiol. Biotechnol. 21:99–114.

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.

Brownstein MJ, Carpten JD, Smith JR (1996). Modulation of non-templated nucleotide addition by Taq DNA polymerase: primer modifications that facilitate genotyping. Biotechniques 20:1004-1010.

Bruford MW, Cheesman DJ, Coote T, Green HAA, Haines SA, O'Ryan C, William TR (1996). Microsatellites and their application to conservation genetics. In: Smith TB, Wayne RK (eds) Molecular genetic approaches in conservation. Oxford University Press, New York, pp. 278–297.

Caetano-Anolles G (1994). MAAP- A versatile and universal tool for genome analysis. Plant Mol. Biol. 25:1011-1026.

Caetano-Anolles G (1997). DNA amplification fingerprinting. In: Micheli MR, Bova R (eds) Fingerprinting methods based on arbitrarily primed PCR. Springer Lab Manual, pp.65-80.

Caetano-Anolles G, Bassam BJ, Gresshoff PM (1991). DNA amplification fingerprinting using very short arbitrary oligonucleotide primers. Biotechnol. 9: 553-557.

Caetano-Anolles G, Bassam BJ, Gresshoff PM (1992). Primer-template interactions during DNA amplification fingerprinting with single arbitrary oligonucleotides. Mol. General Genet. 235:157-165.

Caetano-Anolles G, Bassam BJ, Gresshoff PM (1993). Enhanced detection of polymorphic DNA by multiple arbitrary amplicons profiling of endonuclease digested DNA: identification of markers linked to the supernodulation locus in soybean. Mol. General Genet. 241: 57-64.

Caudrado A, Schwarzacher T (1998). The chromosomal organization of simple sequence repeats in wheat and rye genomes. Chromosome 107:587–594.

Chabane K, Ablett GA, Cordeiro GM, Valkoun J, Henry RJ (2005). EST versus genomic derived microsatellite markers for genotyping wild and cultivated barley. Genet. Res. Crop Evol. 52:903–909.

Chambers GK, MacAvoy ES (2000). Microsatellites: consensus and controversy. Comparative Biochem. Physiol. (Part B) 126:455–476.
Cho YG, Ishii T, Temnykh S, Chen X, Lipovich L, McCouch SR, Parl WD, Ayers N, Cartinhour S (2000). Diversity of microsatellites derived from genomic libraries and GenBank sequences in rice (Oryza sativa L.). Theor. Appl. Genet. 100:713–722.
Clancy JA, Jitkov VA, Han F, Ullrich SE (1996). Barley tissue as direct template for PCR: a practical breeding tool. Mol. Breed. 2:181-183.
Cooper G (1995). Analysis of genetic variation and sperm competition in Dragon flies. D. Phil. Thesis, Oxford University.
Cordeiro GM, Casu R, McIntyre CL, Manners JM, Henry RJ (2001). Microsatellite markers from sugarcane (Sacharum spp.) ESTs cross transferable to erianthus and sorghum. Plant Sci. 160:115–1123.
CorleySmith GE, Lim CJ, Kalmar GB, Brandhorst BP (1997). Efficient detection of DNA polymorphisms by fluorescent RAPD analysis. BioTechniques 22: 690-692.
Davis GL, Mcmullen MD, Baysdorfer C, Musket T, Grant D, et al. (1999). A maize map standard with sequenced core markers, grass genome reference points, and 932 expressed sequence tagged sites (ESTs) in a 1736 locus map. Genetics 152: 1137–1172.
Dayteg C, Vos Post L, Lund R, Tuvesson S (1998). Quick DNA extraction method for practical plant breeding programmes. Plant and Animal Genome VI. San Diego, Ca., USA, Jan 18–22, p. 39.
Decousset L, Griffiths S, Dunford RP, Pratchett N, Laurie DA (2000). Development of STS markers closely linked to the Ppd-H1 photoperiod response gene of barley (*Hordeum Vulgare* L.). Theor. Appl. Genet. 101:1202–1206.
Decroocq V, Fave MG, Hagen L, Bordenave L, Decroocq S (2003). Development and transferability of apricot and grape EST microsatellite markers across taxa. Theor. Appl. Genet. 106:912–922.
Dellaporta SL, Wood J, Hicks JB (1983). A plant DNA minipreparation: version II. Plant Mol. Biol. Rep. 1:19-21.
Demeke T, Adams RP, Chibbar R (1992). Potential taxonomic use of random amplified polymorphic DNA (RAPD): a case study in Brassica. Theor. App. Gene. 84:567-572.
Devos KM, Gale MD (1992). The use of random amplified polymorphic DNA markers in wheat. Theor. Appl. Genet. 84:567-572.
Dietrich W, Katz H, Lincoln SE, Shin HS, Friedman J, Dracopoli NC, Lander ES (1992). A genetic map of the mouse suitable for typing intraspecific crosses. Genetics 131:423-447.
Ding C, Cantor CR (2003). A high-throughput gene expression analysis technique using competitive PCR and matrix-assisted laser desorption ionization time-of-flight MS. Proc. Natl. Acad. Sci. USA 100:3059–3064.
Diwan N, Cregan PB (1997). Automated sizing of fluorescent-labeled simple sequence repeat (SSR) markers to assay genetic variation in soybean. Theor. Appl. Genet. 95:723–733.
Drenkard E, Richter BG, Rozen S, Stutius LM, Angell NA, Mindrinos M, Cho RJ, Oefner PJ, Davis RW, Ausubel FM (2000). A simple procedure for the analysis of single nucleotide polymorphisms facilitates map-based cloning in Arabidopsis. Plant Physiol. 124:1483–1492.
Eisen JA (1999). Mechanistic basis for microsatellilte instability. In: Goldstein DB, Schlotterer, C (eds) Microsatellites: evolution and applications. Oxford University Press, Oxford, pp. 34–48.
Ellsworth DL, Rittenhouse D, Honeycutt RL (1993). Artifactual variation in random amplified polymorphic DNA banding patterns. Biotechniques 14: 214-217.
Eujayl I, Sledge MK, Wang L, May GD, Chekhovskiy K, Zwonitzer JC, Mian MAR (2004). Medicago truncatula EST-SSRs reveal cross-species genetic markers for Medicago spp. Theor. Appl. Genet. 108:414–422.
Eujayl I, Sorrells M, Baum M, Wolters P, Powell W (2001). Assessment of genotypic variation among cultivated durum wheat based on EST-SSRs and genomic SSRs. Euphytica 119:39–43.
Eujayl I, Sorrells ME, Wolters P, Baum M, Powell W (2002). Isolation of EST-derived microsatellite markers for genotyping the A and B genomes of wheat. Theor. Appl. Genet. 104:399–407.
Fang DQ, Roose ML (1997). Identification of closely related citrus cultivars with inter-simple sequence repeat markers. Theor. Appl. Genet. 95: 408–417.
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.
Gao LF, Jing LR, Huo NX, Li Y, Li XP, Zhou RH, Chang XP, Tang JF, Ma ZY, Jia JZ (2004). One hundred and one new microsatellite loci derived from ESTs (EST-SSRs) in bread wheat. Theor. Appl. Genet. 108:1392–1400.
Garg K, Green P, Nickerson DA (1999). Identification of candidate coding region single nucleotide polymorphisms in 165 human genes using assembled expressed sequence tags. Genome Res. 9:1087–1092.
Ginot F, Bordelais I, Nguyen S, Gyapay G (1996). Correction of some genotyping errors in automated fluorescent microsatellite analysis by enzymatic removal of one base overhangs. Nucleic Acids Res. 24: 540-541.
Goldstein DB, Pollock DD (1997). Launching microsatellites: a review of mutation processes and methods of phylogenetic inference. J. Hered. 88:335–342.
Goldstein DB, Schlotterer C (1999). Microsatellites, evolution and applications. Oxford University Press, New York.
Grodzicker T, Williams J, Sharp P, Sambrook J (1975). Physical mapping of temperature sensitive *mutants of adenovirus*. Cold Spring Harbor Symp Quant Biol 39:439-446.
Gu WK, Weeden NF, Yu J, Wallace DH (1995). Large-scale, cost-effective screening of PCR products in marker-assited 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.
Gupta PK, Roy JK, Prasad M (2001). Single nucleotide polymorphisms: a new paradigm for molecular marker technology and DNA polymorphism detection with emphasis on their use in plants. Curr. Sci. 80:524–535.
Gupta PK, Rustgi S, Sharma S, Sing R, Kumar N, Balyan HS (2003). Transferable EST-SSR markers for the study of polymorphism and genetic diversity in bread wheat. Mol. Genet. Genomics 270:315–323
Gupta PK, Varshney RK, Sharma PC, Ramesh B (1999). Molecular markers and their applications in wheat breeding. Plant Breed. 118:369–390.
Hamada H, Petrino MG, Kakunaga T (1982). A novel repeat element with Z-DNA-forming potential is widely found in evolutionarily diverse eukaryotic genomes. Proc. Natl. Acad. Sci. USA 79:6465-6469.
Hammond RL, Saccheri IJ, Ciofi C, Coote T, Funk SM, McMillan WO, Bayes MK, Taylor E, Bruford MW (1998). Isolation of microsatellite markers in animals. In: Karp A, Issac PG, Ingram DS (eds) Molecular tools for screening biodiversity. Chapman and Hall, London, pp. 279–287.
Hancock JM (1999). Microsatellites and other simple sequences: genomic context and mutational mechanisms. In: Goldstein DB, Schlotterer C (eds) Microsatellites: evolution and applications. Oxford University Press, Oxford, pp. 1–9.
Hantula J, Dusabenygasani M, Hamelin RC (1996). Random amplified microsatellites (RAMS) - a novel method for characterizing genetic variation within fungi. Eur. J. For. Path. 26:15-166.
Harushima Y, Yano M, Shomura A, Sato M, Shimano T, Kuboki Y, et al. (1998). A high-density rice genetic linkage map with 2275 markers using a single F2 population. Genetics 148:479–494.
Hearne CM, Ghosh S, Todd JA (1992). Microsatellites for linkage analysis of genetic traits. Trends Genet. 8:288–294.
Helentjaris T, Slocum M, Wright S, Schaefer A, Nienhuis J (1986). Construction of genetic linkage maps in maize and tomato using restriction fragment length polymorphisms. Theor. Appl. Genet. 61:650-658.
Hempel K, Peakall R (2003). Cross-species amplification from crop soybean Glycine max provides informative microsatellite markers for the study of inbreeding wild relatives. Genome 46:382–393.
Heun M, Helentjaris T (1993). Inheritance of RAPDs in F1 hybrids of corn. Theor. Appl. Genet. 85:961-968.
Hollingsworth WO, Christie CB, Nichols MA, Neilson HF (1998). Detect-

ion of variation among and within asparagus hybrids using random amplified DNA (RAPD) markers. New Zealand J. Crop Hortic. Sci. 26:1-9.

Holtke HJ, Ankenbauer W, Muhlegger K, Rein R, Sagner G, Seibl R, Walter T (1995). The digoxigenin (DIG) system for non-radioactive labelling and detection of nucleic acids - an overview. Cell. Mol. Biol. 41: 883-905.

Huff DR, Peakall R, Smouse PE (1993). RAPD variation within and among natural populations of outcrossing buffalograss [(*Buchloe dactyloides* (Nutt.) Engelm]. Theor. Appl. Genet. 86:927-934.

Hunt GJ, Page RE (1992). Pattern of inheritance with RAPD molecular markers reveal novel types of polymorphism in honey bee. Theor. Appl. Genet. 85:15-20.

Ikeda N, Bautista NS, Yamada T, Kamijima O, Ishii T (2001). Ultra-simple DNA extraction method for marker-assisted selection using microsatellite markers in rice. Plant Mol. Biol. Rep. 19:27–32

Jaccoud D, Peng K, Feinstein D, Kilian A (2001). Diversity arrays: a solid state technology for sequence information independent genotyping. Nucleic Acids Res. 29:e25.

Jain A, Bhatia S, Banga S, Prakash S (1994). Potential use of random amlified polymorphic DNA (RAPD) technique to study the genetic diversity in Indian mustard (*Brassica juncea*) and its relationship to heterosis. Theor. Appl. Genet. 88: 116-122.

Jarne P, Lagoda P (1996). Microsatellites, from molecules to populations and back. Trends Ecol. Evol. 11: 424–429.

Jarvis P, Lister C, Szabo V, Dean C (1994). Integration of CAPS markers into the RFLP map generated using recombinant inbred lines of *Arabidopsis thaliana*. Plant Mol. Biol. 24:685–687.

Jones CJ, Edwards KJ, Castaglione S, Winfield MO, Sala F, van de Wiel C, et al. (1997) Reproducibility testing of RAPD, AFLP, and SSR markers in plants by a network of European laboratories. Mol. Breed. 3:381-390.

Jongeneel CV (2000). Searching the expressed sequence tag (EST) databases: Panning for genes. Briefings in Bioinformatics 1:76-92.

Jordan SA, Humphries P (1994). Single nucleotide polymorphism in exon 2 of the BCP gene on 7q31-q35. Hum. Mol. Genet. 3:1915.

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.

Kaiser MW, Lyamicheva N, Ma W, Miller C, Neri B, Fors L, Lyamichev VI (1999). A comparison of eubacterial and archaeal structure-specific 5`-exonucleases. J. Biol. Chem. 274 :21387–21394.

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.

Karp A, Kresovich S, Bhat KV, Ayad WG, Hodgkin T (1997). Molecular tools in plant genetic resources conservation: a guide to the technologies. IPGRI Technical Bulletin No. 2, Rome, Italy.

King RC, Stansfield WD (1990). A dictionary of genetics. 4th ed., Oxford University Press, New York-Oxford, pp 188.

Koes R, Souer E, van Houwelingen A, Mur L, Spelt C, Quattrocchio F, et al. (1995). Targeted gene inactivation in petunia by PCR-based selection of transposon insertion mutants. Proc. Natl. Acad. Sci. 92:8149-8153.

KojimaT, Nagaoka T, Noda N, 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, Ausubel FM (1993). A procedure for mapping Arabidopsis mutations using co-dominant ecotype-specific PCR-based markers. Plant J. 4: 403–410.

Kostia S, Varvio SL, Vakkari P, Pulkkinen P (1995). Microsatellite sequences in a conifer, Pinus sylvestris. Genome 38:1244–1248.

Kota R, Varshney RK, Thiel T, Dehmer KJ, Graner A (2001). Generation and comparison of EST-derived SSRs and SNPs in barley (*Hordeum vulgare* L.). Hereditas (Lund) 135:145–151.

Kresovich S, Williams JGK, McFerson JR, Routman EJ, Schaal BA (1992). Characterization of genetic identities and relationships of *Brassica oleracea* L. via a random amplified polymorphic DNA assay. Theor. Appl. Genet. 85: 190-196.

Kurata N, Umehara Y, Tanoue H, Sasaki T (1997). Physical mapping of the rice genome with YAC clones. Plant Mol. Biol. 35:101–113.

Landegren U, Kaiser R, Sanders J, Hood L (1988). DNA diagnostics. Molecular techniques and automation. Science 241:1077-1080.

Landry BS (1994). DNA mapping in plants. In: Glick BR, Thompson JE (eds) Methods in plant molecular biology and biotechnology. CRC Press, Boca Raton-Ann Arbor-London-Tokyo, pp 269-285.

Le Hellard S, Ballereau SJ, Visscher PM, Torrance HS, Pinson J, Morris SW, et al. (2002) SNP genotyping on pooled DNAs: comparison of genotyping technologies and a semi automated method for data storage and analysis. Nucleic Acids Res. 30:e74.

Levinson G, Gutman GA (1987). Slipped-strand mispairing: a major mechanism for DNA sequence evolution. Mol. Biol. Evol. 4:203–221.

Lezar S, Myburg AA, Berger DK, Wingfield MJ, Wingfield BD (2004). Development and assessment of microarray-based DNA fingerprinting in *Eucalyptus grandis*. Theor. Appl. Genet. 109:1329–1336.

Lin JJ, Kuo J, Ma J, Saunders JA, Beard HS, Macdonald MH, Kenworthy W, Ude GN, Mathews BF (1996). Identification of molecular markers in soybean comparing RFLP, RAPD and AFLP DNA mapping techniques. Plant Mol. Biol. Rep. 14: 156-169.

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.

Liu S, Anderson JA (2003). Marker assisted evaluation of Fusarium head blight resistant wheat germplasm. Crop Sci. 43:760–766.

Luo C, Deng L, Zeng C (2004). High throughput SNP genotyping with two mini-sequencing assays. Acta Biochim. Biophys. Sin. 36:379–84.

Lyamichev V, Mast AL, Hall JG, Prudent JR, Kaiser MW, Takova T, et al. (1999). Polymorphism identification and quantitative detection of genomic DNA by invasive cleavage of oligonucleotide probes. Nat. Biotechnol. 17: 292–296.

Lynch M, Walsh B (1998). Genetics and analysis of quantitative traits. Sinauer Associates, Sunderland, MA.

Mackay TFC (2001). The genetic architecture of quantitative traits. Ann. Rev. Genet. 35:303-339.

Maeda M, Uryu N, Murayama N, Ishii H, Ota M, Tsuji K, Inoko H (1990). A simple and rapid method for HLA-DP genotying by digestion of PCR-amplified DNA with allele specific restriction endonucleases. Hum. Immunol. 27:111–121.

Mansfield ES, Worley JM, McKenzie SE, Surrey S, Rappaport E, Fortina P (1995). Nucleic acid detection using non-radioactive labeling methods. Mol. Cell. Probes 9:145-156.

Markert CL, Moller F (1959). Multiple forms of enzymes: Tissue, ontogenetic and species specific patterns. Proc. Natl. Acad. Sci. USA 45:753-763.

Matsumoto A, Tsumura Y (2004). Evaluation of cleaved amplified polymorphic sequence markers. Theor. Appl. Genet. 110: 80–91.

McDermott JM, Brandle U, Dutly F, Haemmerli UA, Keller S, Muller KE, Wolf MS (1994). Genetic variation in powdery mildew of barley: Development of RAPD, SCAR and VNTR markers. Phytopathology 84:1316-1321.

McDonald DB, Potts WK (1997). DNA microsatellites as genetic markers for several scales. In: Mindell DP (ed.) Avian molecular evolution and systematics. Academic Press, San Diego, pp. 29–49.

Meksem K, Leister D, Peleman J, Zabeau M, Salamini F, Gebhardt C (1995). A high resolution mapping of the vicinity of the R1 locus on chromosome V of potato based on RFLP and AFLP markers. Mol. Gen. Genet. 249:74–81.

Meksem K, Ruben E, Hyten D, Triwitayakorn K, Lightfoot DA (2001). Conversion of AFLP bands into high-throughput DNA markers. Mol. Genet. Genomics 265:207–214.

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.

Michaels SD, Amasino RM (1998). A robust method for detecting single-nucleotide changes as polymorphic markers by PCR. Plant J. 14:381–385.

Michalek V, Weschke W, Pleisner KP, Graner A (2002). EST analysis in barley defines a unigene set comprising 4000 genes. Theor. Appl. Genet. 104: 97–103.

Micheli MR, Bova R, Calissano P, D'Ambrosio E (1993). Randomly

amplified polymorphic DNA fingerprinting using combinations of oligonucleotide primers. Biotechniques 15:388-390.
Micheli MR, Bova R, Pascale E, D'Ambrosio E (1994). Reproducible DNA fingerprinting with the random amplified polymorphic DNA (RAPD) method. Nucleic Acids Res. 22:1921-1922.
Milbourne D, Meyer R, Bradshaw JE, Baird E, Bonar N, Provan J, Powell W, Waugh R (1997). Comparison of PCR-based marker systems for the analysis of genetic relationships in cultivated potato. Mol. Breed. 3:127-136.
Moreno S, Martin JP, Ortiz JM (1998). Inter simple sequence repeats PCR for characterization of closely related grapevine germplasm. Euphytica 101:117–125.
Morgante M, Hanafey M, Powell W (2002). Microsatellites are preferentially associated with nonrepetitive DNA in plant genomes. Nat. Genet. 30: 194–200.
Morgante M, Vogel J (1994). Compound microsatellite primers for the detection of genetic polymorphisms. US patent application no. 08/326456.
Mueller UG, Lipari SE, Milgroom MG (1996). Amplified fragment length polymorphism (AFLP) fingerprinting of symbiotic fungi cultured by the fungus-growing ant Cyphomyurmex minutus. Mol. Ecol. 5:119-122.
Mueller UG, Wolfenbarger LL (1999). AFLP genotyping and finger-printing. Trends Ecol. Evol. 14:389-394.
Murray MG, Thompson WF (1980). Rapid isolation of high molecular weight plant DNA. Nucleic Acids Res. 8:4321-4325.
Naik S, Gill KS, Prakasa Rao VS, Gupta VS, Tamhankar SA, Pujar S, Gill BS, Ranjekar PK (1998). Identification of a STS marker linked to the *Aegilops speltoides*-derived leaf rust resistance gene Lr17 in wheat. Theor. Appl. Genet. 86:424-439.
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.
Nasu S, Suzuki J, Ohta R, Hasegawa K, Yui R, Kitazawa N, Monna L, Minobe Y (2002). Search for and analysis of single nucleotide polymorphisms (SNPs) in rice (Oryza satva, Oryza rufipogon) and establishment of SNP markers. DNA Res. 9:163-171.
Neff MM, Neff JD, Chory J, Pepper AE (1998). dCAPS, a simple technique for the genetic analysis of single nucleotide polymorphisms: Experimental applications in *Arabidopsis thaliana* genetics. Plant J. 14:387–392.
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.
Pammi S, Schertz K, Xu G, Hart G, Mullet JE (1994). Random-amplified-polymorphic DNA markers in sorghum. Theor. Appl. Genet. 89: 80-88.
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.
Parker GD, Langridge P (2000). Development of an STS marker linked to a major locus controlling flour colour in wheat (Triticum aestivum L.). Mol. Breed. 6:169–174.
Pastinen T, Kurg A, Metspalu A, Peltonen L, Syvanen AC, Minisequencing (1997). A specific tool for DNA analysis and diagnostics on oligonucleotide arrays. Genome Res. 7:606-614.
Paterson AH, Brubaker CL, Wendel JF (1993). A rapid method for extraction of Cotton (Gossypium spp) genomic DNA suitable for RFLP or PCR analysis. Plant Mol. Biol. Rep. 11:122-127.
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.
Perez de la Vega MP (1993). Biochemical characterization of populations. In: Hayward MD, Bosemark NO, Romagosa I (eds) Plant breeding: Principles and prospects. Chapman and Hall, London, pp 182-201.
Pfeiffer A, Olivieri AM, Morgante M (1997). Identification and characterization of microsatellites in Norway spruce (Picea abies K.). Genome 40:411–419.
Pillay M, Kenny ST (1995). Anomalies in direct pairwise comparisons of RAPD fragments for genetic analysis. Biotechniques 19: 694-698.
Potter RH, Jones MGK (1991). Molecular analysis of genetic stability. In: Dodds JH (eds) In vitro methods for conservation of plant genetic resources. Chapman and Hall, London pp 71-91.
Powell W, Morgante M, Andre C, Hanafey M, Vogel J, Tingey S, Rafalski A (1996). The comparison of RFLP, RAPD, AFLP and SSR (microsatellite) markers for germplasm analysis. Mol. Breed. 2: 225-238.
Prins R, Groenewald JZ, Marais GF, Snap JW, Koebner RMD (2001). AFLP and STS tagging of Lr19, a gene conferring resistance to leaf rust in wheat. Theor. Appl. Genet. 103:618–624.
Qi X, Stam P, Lindhout P (1998). Use of locus-specific AFLP markers to construct a high-density molecular map in barley. Theor. Appl. Genet. 96:376-384.
Qu LJ, Foole TN, Roberts MA, Money TA, Aragon-Alcaide L, Snape JW, Moore G (1998). A simple PCR-based method for scoring the ph1b deletion in wheat. Theor. Appl. Genet. 96: 371–375.
Queller DC, Strassman JE, Hughes CR (1993). Microsatellites and Kinship. Trends Ecol. Evol. 8: 285-288.
Qureshi SN, Saha S, Kantety RV, Jenkins JN (2004). EST-SSR: a new class of genetic markers in cotton. J. Cotton Sci. 8:112–123.
Rabouam C, Comes AM, Bretagnolle V, Humbert JF, Periquet G, Bigot Y (1999). Features of DNA fragments obtained by random amplified polymorphic DNA (RAPD) assays. Mol. Ecol. 8: 493-503.
Rafalski A (2002a). Applications of single nucleotide polymorphisms in crop genetics. Curr. Opin. Plant Biol. 5:94–100.
Rafalski JA (2002b). Novel genetic mapping tools in plants: SNPs and LD-based approaches. Plant Sci.162:329–333.
Ratnaparkhe MB, Tekeoglu M, Muehlbauer FJ (1998). Inter simple sequence-repeat (ISSR) polymorphisms are useful for finding markers associated with disease resistance gene clusters. Theor. Appl. Genet. 97: 515–519.
Reamon-Buttner SM, Jung C (2000). AFLP-derived STS markers for the identification of sex in Asparagus officinalis L. Theor. Appl. Genet. 100:432–438
Reddy MP, SarlaN, Siddiq EA (2002). Inter simple sequence repeat (ISSR) polymorphism and its application in plant breeding. Euphytica 128: 9–17.
Ridout C, Donini P (1999). Use of AFLP in cereals research. Trends Plant Sci. 4:76-79.
Roder MS, Korzun V, Wendehake K, Plaschke J, Tixier MH, Leroy P, Ganal MW (1998). A microsatellite map of wheat. Genetics 149: 2007–2023.
Roder MS, Plaschke J, Konig SU, Borner A, Sorrells ME, Tanksley SD, Ganal MW (1995). Abundance, variability and chromosome location of microsatellites in wheat. Mol. Gen. Genet. 246:327–333.
Rohde W (1996). Inverse sequence-tagged repeat (ISTR) analysis, a novel and universal PCR-based technique for genome analysis in the plant and animal kingdom. J. Genet. Breed. 50: 249-261.
Roy A, Frascaria N, Mackay J, Bonsquet J (1992). Segregating random amplified polymorphic DNA (RAPDs) in Betula alleghaniensis. Theor. Appl. Genet. 85:173-180.
Rungis D, Hamberger B, Berube Y, Wilkin J, Bohlmann J, Ritland K (2005). Efficient genetic mapping of single nucleotide polymorphisms based upon DNA mismatch digestion. Mol. Breed. 16: 261–270.
Russell JR, Fuller JD, Macaulay M, Hatz BG, Jahoor A, Powell W, Waugh R (1997). Direct comparison of levels of genetic variation among barley accessions detected by RFLPs, AFLPs, SSRs and RAPDs. Theor. Appl. Genet. 95:714-722.
Saghai-Maroof MA, Soliman KM, Jorgensen RA, Allard RW (1984). Ribosomal DNA sepacer-length polymorphism in barley: Mendelian inheritance, chromosomal location, and population dynamics. Proc. Natl. Acad. Sci. USA 81: 8014-8019.
Saha S, Karaca M, Jenkins JN, Zipf AE, Reddy UK, Kantety RV (2003). Simple sequence repeats as useful resources to study transcribed genes of cotton. Euphytica 130:355–364.
Sanchez MP, Davila JA, Loarce Y, Ferrer E (1996). Simple sequence repeat primers used in polymerase chain reaction amplifications to study genetic diversity in barley. Genome 39: 112-117.
Sankar AA, Moore GA (2001). Evaluation of inter-simple sequence repeat analysis for mapping in Citrus and extension of genetic linkage map. Theor. Appl. Genet. 102: 206–214.

Sax K (1932). The association of size differences with seed-coat pattern and pigmentation in Phaseolus vulgaris. Genetics 8:552-560.

Schierwater B, Metzler D, Kruger K, Strit B (1996). The effect of nested primer binding sites on the reproducibility of PCR: mathematical modelling and computer simulation studies. J. Computational Biol. 3:235-251.

Schlotterer C (1998). Microsatellites. In: Hoelzel AR (ed) Molecular genetic analysis of populations: A practical approach. IRL Press, Oxford, pp. 237–261.

Scott KD, Eggler P, Seaton G, Rosetto M, Ablett EM, Lee LS, Henry RJ (2000). Analysis of SSRs derived from grape ESTs. Theor. Appl. Genet. 100:723–726.

Semagn K, Bjørnstad Å, Skinnes H, Marøy AG, Tarkegne T, William M (2006). Distribution of DArT, AFLP, and SSR markers in a genetic linkage map of a doubled haploid hexaploid wheat population. Genome 49:545–555.

Shan X, Blake TK, Talbert LE (1999). Conversion of AFLPs to sequence-tagged-site PCR markers in barley and wheat. Theor. Appl. Genet. 98:1072-1078.

Shuber AP, Michalowsky LA, Nass GS, Skoletsky J, Hire LM, Kotsopoulos SK, Phipps MF, Barberio DM, Klinger KW (1997). High throughput parallel analysis of hundreds of patient samples for more than 100 mutations in multiple disease genes. Hum. Mol. Genet. 6: 337B347.

Skroch P, Nienhuis J (1995). Impact of scoring error and reproducibility of RAPD data on RAPD based estimates of genetic distance. Theor. Appl. Genet. 91:1086-1091.

Smith DN, Devey ME (1994). Occurrence and inheritance of microsatellites in Pinus radiata. Genome 37:977–983.

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

Song QJ, Fickus EW, Cregan P (2002). Characterization of trinucleotide SSR motifs in wheat. Theor. Appl. Genet. 104:286–293.

Southern EM (1975). Detection of specific sequences among DNA fragments separated by gel electrophoresis. J. Mol. Biol. 98:503-517.

Squirrell J, Hollingsworth PM, Woodhead M, Russell J, Lowe AJ, Gibby M and Powell W (2003). How much effort is required to isolate nuclear microsatellites from plants? Mol. Ecol. 12:1339–1348.

Staub JE, Serquen FC (1996). Genetic markers, map construction, and their application in plant breeding. HortScience 31:729-740.

Tautz D, Trick M, Dover GA (1986). Cryptic simplicity in DNA is a major source of genetic variation. Nature 322:652-656.

Telenius H, Carter NP, Bebb CE, Nordenskjold M, Ponder BAJ, Tunnacliffe A (1992). Degenerate oligonucleotide-primed PCR: general amplification of target DNA by a single degenerate primer. Genomics 13:718-725.

Terachi T (1993). The progress of DNA analyzing techniques and its impact on plant molecular systematics. J. Plant Res. 106:75-79.

The Arabidopsis Genome Initiative (2000). Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. Nature 408:796-815.

The Rice Genome Sequencing Project (2005). The map-based sequence of the rice genome. Nature 436:793-800.

Thiel T, Michalek RK, Varshney RK, 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.

Thompson D, Henry RJ (1995). Single step protocol for preparation of plant tissue for analysis by PCR. Biotechniques 19:394-397.

Thormann CE, Ferreira ME, Camargo LEA, Tivang JG, Osborn TC (1994). Comparison of RFLP and RAPD markers to estimating genetic relationships within and among cruciferous species. Theor. Appl. Genet. 88: 973-980.

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

Tost J, Gut IG (2005). Genotyping single nucleotide polymorphisms by MALDI mass spectrometry in clinical applications. Clinical Biochem. 38:335-350.

Tost J, Schatz P, Schuster M, Berlin K, Gut IG (2003). Analysis and accurate quantification of CpG methylation by MALDI mass spectrometry. Nucleic Acids Res. 31:e50.

Triglia T, Peterson MG, Kemp DJ (1988). A procedure for in vitro amplification of DNA segments that lie outside the boundaries of known sequences. Nucleic Acids Res. 16:8180-8186.

Tsumura Y, Ohba K, Strauss SH (1996). Diversity and inheritance of inter-simple sequence repeat polymorphisms in Douglas-fir (Pseudotsuga menziesii) and sugi (Cryptomeria japonica). Theor. Appl. Genet. 92: 40–45.

vanderVoort JNAMR, vanZandvoort P, vanEck HJ, Folkertsma RT, Hutten RCB, Draaistra J, Gommers FJ, Jacobsen E, Helder J, Bakker J (1997). Use of allele specificity of co-migrating AFLP markers to align genetic maps from different potato genotypes. Mol. General Genet. 255:438-447.

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

Varshney RK, Thiel T, Stein N, Langridge P, Graner A (2002). In silico analysis on frequency and distribution of microsatellites in ESTs of some cereal species. Cell. Mol. Biol. Lett. 7: 537–546.

Vejl P (1997). Identification of genotypes in hop (Humulus lupulus L.) by RAPD analsysis using program Gel Manager for Windows. Rostlinna Vyroba 43:325-331.

Vignal A, Milana D, Sancristobala M And Eggenb A (2002). A review on SNP and other types of molecular markers and their use in animal genetics. Genet. Sel. Evol. 34:275-305.

von Post R, von Post L, Dayteg C, Nilsson M, Forster BP and Tuvesson S (2003). A high-throughput DNA extraction method for barley seed. Euphytica 130:255–260.

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.

Walker GT, Little MC, Nadeau JG, Shank DD (1992). Isothermal in vitro amplification of DNA by a restriction enzyme/DNA polymerase system. Proc. Natl. Acad. Sci. 89:392-396.

Wallace RB, Shaffer J, Murphy RF, Bonner J, Hirose T, Itakura K (1979). Hybridization of synthetic oligodeoxyribonucleotides to phi 174 DNA: the effect of single base pair mismatch. Nucleic Acids Res. 6:3543–3557.

Wang G, Mahalingan R, Knap HT (1998). (C-A) and (GA) anchored simple sequence repeats (ASSRs) generated polymorphism in soybean, Glycine max (L.) Merr. Theor. Appl. Genet. 96:1086-1096.

Waugh R, Bonar N, Baird E, Thomas B, Graner A, Hayes P, Powell W (1997). Homology of AFLP products in three mapping populations of barley. Mol. General Genet. 255:311-321.

Weber D, Helentjaris T (1989). Mapping RFLP loci in maize using B-A translocations. Genetics 121:583-590

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.

Weller JW, Reddy A (1997). Fluorescent detection and analysis of RAPD amplicons using the ABI PRISM DNA sequencers. In: Micheli MR, Bova R (eds) Fingerprinting methods based on arbitrarily primed PCR, Springer Lab Manual, pp 81-92.

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

Welsh J, McClelland M (1991). Genomic fingerprinting using arbitrarily primed PCR and a matrix of pairwise combinations of primers. Nucleic Acids Res. 19:5275-5279.

Wenzl P, Carling J, Kudrna D, Jaccoud D, Huttner E, Kleinhofs A, Kilian A (2004). Diversity Arrays Technology (DArT) for whole-genome profiling of barley. Proc. Natl. Acad. Sci. USA 101: 9915-9920.

Wexelsen H (1933) Linkage between quantitative and qualitative characters in barley. Hereditas (Lund) 17:323-341.

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.

Wittenberg AHJ, van der Lee T, Cayla C, Kilian A, Visser RGF, Schouten HJ (2005). Validation of the high-throughput marker technology DArT using the model plant *Arabidopsis thaliana*. Mol. Genet. Genomics 274:30-39.

Wolfe, AD, Liston A (1998). Contribution of PCR-based methods to plant systematics and evolutionary biology. In: Soltis DE, Soltis PS,

Doyle JJ (eds) Molecular systematics of plants II: DNA sequencing, Kluwer Academic Publishers, pp. 43-86.

Wolff K, Schoen ED, Peters-Van Rijn J (1993). Optimizing the generation of random amplified polymorphic DNA in chrysanthemum. Theor. Appl. Genet. 86:1033-1037.

Wu K, Jones R, Dannaeberger L, Scolnik PA (1994). Detection of microsatellite polymorphisms without cloning. Nucleic Acids Res. 22: 3257–3258.

Xia L, Peng K, Yang S, Wenzl P, de Vicente MC, Fregene M, Kilian A (2005). DArT for high-throughput genotyping of Cassava (Manihot esculenta) and its wild relatives. Theor. Appl. Genet. 110: 1092-1098.

Yang X, Quiros C (1993). Identification and classification of celery cultivars with RAPD markers. Theor. Appl. Genet. 86:205-212.

Yang S, Pang W, Ash G, Harper J, Carling J, Wenzl P, Huttner E, Zong X, Kilian A (2006). Low level of genetic diversity in cultivated Pigeon-pea compared to its wild relatives is revealed by diversity arrays technology. Theor. Appl. Genet. 113:585-595.

Young ND, Menancio-Hautea D, Fatokun CA, Danesh D (1992). RFLP technology, crop improvement, and international agriculture. In: Thottappilly G, Monti LM, Moham DR, Moore AW (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, et al. (2002). A draft sequence of the rice genome (Oryza sativa L. ssp. indica). Science 296:79–92.

Yu KF, Deynze AV, Pauls KP (1993). Random amplified polymorphic DNA (RAPD) analysis. In: Glick BR, Thompson JE (eds) Methods in plant molecular biology and biotechnology. CRC Press, Boca Raton-Ann Arbor-London-Tokyo, 360 pp.

Zane L, Bargelloni L, Patarnello T (2002). Strategies for microsatellite isolation: a review. Mol. Ecol. 11:1-16.

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