REVIEW OF LITERATURE
II. REVIEW OF LITERATURE

Efforts have been made to review the existing literature on the importance, origin, species, varieties and the distribution of Santalum album. Since meager work have been reported on the use of molecular markers in Santalum album in India, the literature pertaining to molecular markers, various DNA extraction protocols, polymerase chain reaction, gel electrophoresis, applications and limitations of RAPD markers (Rom et al., 1995) related to other Forest trees crops is also reviewed.

Genetic marker facilitates the study of inheritance and cloning of genes by genetic approaches. Molecular markers can detect differences in DNA sequences, and are thus less ambiguous than phenotypic markers, which require gene expression (Williams et al., 1993).

2.1 Importance of Santalum album

Santalum album, a terrestrial plant species of the Santalaceae family, is commonly known as a source of sandalwood. S. album has been the primary source of sandalwood and the derived oil. It has been acknowledged as "Vegetable Gold" and is highly valued in perfumery, cosmetic, medicine and handicraft industries (Srimathi and Kulkarni, 1995). Medicinally it has astringent and antipyretic properties. The wood is quite appreciated for carving, sculpting liturgical objects, cremation and incense as it has uniform straight close wood grain.

One of the oldest incense materials, Sandalwood has been in use for at least 4,000 years, which has been mentioned even in Sanskrit and
Chinese manuscripts. The oil was used in religious ritual, and many deities and temples were carved from its wood. It has a characteristic sweet, woody odour and has excellent blending properties due to the presence of santalols (90%), which makes it valuable as a fixative for other fragrances. Its ardent admirers have called Sandalwood oil "Liquid Gold," due to its precious nature.

The Chemical Constituents S. album oil is Sesquiterpenes; Sesquiterpenols; Sesquiterpenals; (includes 80 to 90% terpeniod alcohols including α and β-santalols (67%), (Shankaranarayanaya and Theagarajan, 2000) which is a mixture of two primary sesquiterpenic alcohols, santalic and teresantalic acid, aldehyde, pterocarpin and hydrocarbons, isovaleric aldehyde, santene, santenone. The sandalwood oil plays a vital role for the preparation of medicines. Sandal oil is a popular remedy in Gonorrhoea, chronic foetid bronchitis and cystitis, gleet, urethral haemorrhage. Oil is valuable in bronchial catarrh. A mixture of oil of sandal of cubebs and copaiba is generally recommended for gonorrhoea – along with seven drops in sugar. In remittent fevers this oil acts as a diaphoretic. It diminishes the rapidity of heart’s action. Externally, the oil is an excellent application in scabies in every stage and form. Sandal oil mixed with its double the quantity of mustard oil is a good application for pimples on nose. The chief use of sandalwood oil in medicine is in the symptomatic treatment of disuria. It is given in the sub – acute stages of cystitis and gonorrhoea for its action on the urinary passage during excretion. It is also used in diminishing the frequency of mictorition, which is so marked in the tuber culosis of the bladder. Sandal oil is employed as disinfectants for the urinogential tract and as expectorants in bronchittes.
and also used in the manufacture of traditional attars, scented tobacco (zarada), joss stick, soaps and detergents.

The predominant market for Indian sandalwood are France, USA, UK, and Middle East each year. It is estimated that 30,000kg of oil is exported to the Europe, the US, Asia and Middle East each year. Both domestic and international demand is increasing and expected to increase especially in new markets such as, aromatherapy and pharmaceutical application and Cosmetics (Jain et al., 2005). Annual World requirement for sandalwood oil is about 200 tones for the production of cosmetics, perfumery and medicines, which is equivalent to 10,000 tones of heart wood. India produces nearly 140 tones, about 60 per cent of which is exported, earning a foreign exchange of about 25 crores every year (Tandon, 1995).

2.2 Origin

The sandal tree, botanically known as *Santalum album* belongs to the family Santalaceae. The sandal tree grows almost exclusively in the forests of Karnataka, followed by Tamil Nadu, Kerala and Andhra Pradesh, Timor Islands of Indonesia etc. There are approximately 16 species of sandalwood (*Santalum album, S.spicatum, S. austrocaledonicum, S. yasi, S. lanceolatum, S. ellipticum, S.macgregorii, S. insulare*) occurring naturally throughout Australia, India, Indonesia, Papua New Guinea and the islands of the South Pacific. The most important commercial species today is *Santalum album* L, and is generally accepted that it is indigenous to Peninsular India (Gode, 1961). Some are of the view that it was introduced to India from Timor in Indonesia (Fischer 1938; Thirawat, 1955).
Fischer (1938) considered Timor island of East Indian Archipelago as the original home of *S. album* and from there it has been introduced to India some time in the remote past. Thirawat (1955) endorses Fischer's theory that sandal could be exotic to India as the same is substantiated by the natural occurrence of many species of *Santalum* in Malaysian islands and Australian mainland and availability of only one species in India. Boyce (1959) and Rajagopalshetty (1977) corroborate the above statement, and gave the following two evidences in support of the theory (a) Sandal has been found to invade into all the ecoclimatic zones in India (b) Occurrence of spike disease in India which is absent in Timor island (Indonesia) indicate that the immigrant species is adopting to the new ecosystem. However, sandal has been found to exist in India for more than 2500 years (Gode, 1961).

2.3 Distribution

*Santalum album* is tropical in distribution and is distributed in the south eastern region of the Southern hemisphere between 30° N and 40° S latitudes from Indonesia in the West to Juan Fernandez islands in the East and from Hawaiian Archipelago in the North to New Zealand in the South. In India, it is found distributed all over the country (9600 km²) (Shobha N.Rai, 1990), with over 90 per cent of the area in Karnataka and Tamil Nadu (Fig.1). The rest is distributed in other states like Andhra Pradesh, Kerala, Madhya Pradesh, Orissa, Maharashtra, Rajasthan, Uttar Pradesh, Bihar and Manipur (Venkatesan and Srimathi, 1981; Srinivasan *et al.*, 1992).
Figure 1—Distribution of sandalwood in India. Areas in square kilometers are as follows:

<table>
<thead>
<tr>
<th>State</th>
<th>Area (sq km)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Karnataka</td>
<td>5,245</td>
</tr>
<tr>
<td>Tamil Nadu</td>
<td>3,040</td>
</tr>
<tr>
<td>Andhra Pradesh</td>
<td>175</td>
</tr>
<tr>
<td>Nadhye Pradesh</td>
<td>33</td>
</tr>
<tr>
<td>Orissa</td>
<td>25</td>
</tr>
<tr>
<td>Maharashtra</td>
<td>8</td>
</tr>
<tr>
<td>Kerala</td>
<td>7</td>
</tr>
<tr>
<td>Uttar Pradesh</td>
<td>Less than 1</td>
</tr>
<tr>
<td>Private lands</td>
<td>500</td>
</tr>
<tr>
<td>Total</td>
<td>9,034</td>
</tr>
</tbody>
</table>


Figure 2. Santalum album Leaf, flower, fruit

Source: Recent Advances in Research and Management of Sandal (Santalum album L.) in India. Srimathi et al., 1995.
In Karnataka, it is spread over 5245 km\(^2\) mainly in southern parts of the state and sparsely in the north accounting for more than 50 per cent of the total area in the country; it is most plentiful in Shimoga, Chickamagalur, Coorg, Hassan, Mysore, Dharwad, KoJar, Bangalore, Belgaum, Uttar Kannada, Dakshina Karmada, Bellary and Tumkur. Natural regeneration in most of the populations is poor due to biotic and abiotic interference. Most of the growing populations are devoid of larger girth classes not only due to illicit felling, but also to grazing, browsing, hacking and other biotic interference (Swaminathan et al., 1997).

The sandalwood population in Tamil Nadu is distributed over 3040 km\(^2\) mainly in North Arcot (Javadi's, and Yelagiri hills), Salem, Periyar, Coimbatore and Vellore districts and sparsely in Nilgiris, Madurai and Trichy. Sandalwood trees in Andhra cover about 175 km\(^2\) of area and are mainly distributed in Chittoor, Kadapa, Tirumala hills, hyderabad, Kurnool and Arakku valley.

In Kerala, it is distributed over 15 km\(^2\) in Marayoor, Wynad and Thenmalain. The populations of Marayoor reserve of Anjanad valleys are dense and possess superior heartwood. The sandalwood tracts of Orissa occur in an area of 35 km in Rayagada, Kalahandi, Parlekmendi and Jeypore forest divisions, mostly concentrated in Koraput district. In Madhya Pradesh sandal trees are scattered in the forest divisions of Sehora, Sagar and Seoni covering 33 km\(^2\).

In Indonesia it was introduced from the Timor islands. The parasitic sandalwood trees (Santalum species) such as S. freycinetianum (Lanal sandalwood) and S. album (East Indian Sandalwood) became...
endemic to Southwest India, often hiding deep in the Southern forests (Sahni, 2000).

2.4 Description

*Santalum album*, a terrestrial plant species of the Santalaceae family, is commonly known as a source of sandalwood. It is a hemi-parasitic tree, occurring in semi-arid areas from India to the South Pacific and the northern coast of Australia. *Santalum album* is a small evergreen tree reaching a height of 7 to 20 meters and girth of up to 1.5 meters. Natural stands of Indian Sandalwood take 60 to 80 years to reach maturity and their greatest oil content. This is in great contrast to Indian Sandalwood trees grown under plantation in the ORIA, which can reach a height of approximately 8 meters in 15 years and may be harvested from year 13. These plantation trees are expected to weigh between 70 to 80 kilos at harvest. While the wood is highly sought after for sculpture, religious ceremonies and incense, it is the oil contained within the heartwood of the trees that is most coveted. The majority of the heartwood is contained within the central core of the trunk and roots. The oil itself is extremely viscous with a heavy, sweet, woody and fruity aroma that has been described as pungently balsamic. Besides these obvious fragrant characteristics, it contains high amounts of alpha and beta Santalol. These compounds are much prized for their fixative, antibacterial and sedative medicinal properties.

The tree flourishes best at an elevation of 700 m to 1200 m altitude, and in a rainfall of 600 mm to 1600 mm and in red ferruginous loam of varying fertility (Shashidhara et al., 2003). A recent survey has shown that sandalwood is capable of growing from sea level to an elevation of 1800 m, in various soils except in highly alkaline, saline
and swampy situations. Trees growing on stony or gravelly soils are known to have more scented wood (Srimathi et al., 1980); It is a polymorphic species showing considerable variation in leaf pattern (Ramaswamy, 1972). Leaves are opposite and decussate, sometimes show whorled arrangement. Simon et al., (2005) reported the occurrence of six morphological types like ovate, lanceolate, linear, elliptical, big and small types (Fig 2). Trees having ovate leaves occur most predominantly in nature, while those with lanceolate leaves occur to a small extent and those with elliptic or linear leaves occur at a low frequency. The colour of leaves varies from bluish/greenish yellow green (Srimathi et al., 1983).

Stem is initially green and tender, gradually turns brownish and becomes hard. The bark is reddish brown or dark brown and red inside. Stem is smooth in young trees, turns rough with deep vertical cracks in mature trees. Wood is hard and oily. Sapwood is white, scentless, while heartwood is yellowish to brown and strongly scented. Root is moderately long, delicate and contains scented oil. It develops nodular growth at an early stage, which is the first sign of haustoria. However, Nagaveni and Srimathi (1985) reported that a small number of sandalwood plants do exist without haustorial nodules even up to two years.

Flowers are purplish brown, unscented and are borne in auxiliary or terminal cymose panicles. Flowers are tetra to pentamerous, rarely hexamerous and hermaphrodite. Fruit is a drupe, purplish when fully matured and single seeded (Fig. 2). The shape of the fruit varies from round to oblong and sometimes show tapering ends (Nagaveni and Ananthapadmanabha, 1986).
2.5 Genetics

Cytological analysis of root tips indicated the diploid number of chromosomes (2n=20) in *S. album* (Darlington and Wylie, 1955). However, drift in chromosomal number in various tissues and parts of the tree have been reported. In the haustorium up to 40 chromosomes have been reported. A two to five fold increase in size of the chromosome was observed in many reports. This has been attributed to the, phenomenon of endopolyploidy and chromosomal degeneracy (Srimathi and Srinivasayya, 1962).

2.6 Genetic diversity

In the conservation of forest genetic resources, priority should be given to both the present rate of extinction as well as the need to ensure the availability of the resources for future use and adaptability to changing environment. Therefore, comparative genetic diversity studies are needed with the objective of saving gene pools by preventing loss of genotypes, genes and gene complexes. Such studies are particularly needed in tropical tree species like sandalwood because of the vicious process of destruction of this particular species has lead to the poor stocking in Southern India by the way of forest fire and by indiscriminate cutting by sandalwood smugglers.

Thus systematic effort to estimate the genetic diversity (Wang *et al.*, 1999), within species and among spatially isolated populations is the backbone of conservation of forest genetic resources in tropical and temperate species, both for present as well as for future development; Unlike most agricultural species that can generate crop varieties with multiple breeding cycles over a few years, forest tree...
breeders cannot rapidly produce new varieties nor can they quickly breed for new variations among populations. Therefore, the existing genetic diversity among sandalwood populations is fundamental and serves as a backbone of the genetic resources to conserve for both their local survival as well as for future development. These resources also form the basis for advanced breeding and, because of the long generation times, the structure and genetic characteristics of populations, they require very careful conservation strategies, perhaps even more than for most agricultural crops. Therefore, assessment of genetic diversity (Hintum et al., 1994), is vital in developing effective conservation strategies and sustainable management guidelines.

2.7 Genomic DNA Isolation

Isolation of DNA is the first step for the application of a molecular biology to any crop species. Extraction of highly purified genomic DNA from plant tissues is a difficult task due in part to their rigid cell wall which is composed of large amounts of complex carbohydrates (Hattori et al., 1987). Contamination by polysaccharides has been reported as the most common problem affecting plant DNA purity (Murray and Thompson, 1980). Therefore, purification of high molecular weight DNA from plant tissues is essential for many procedures used in molecular genetics for standardize the protocol.

Polymerase Chain Reaction (PCR) based methods are most versatile tools for genetic diversity and evolutionary studies. PCR is based on the efficient action of a thermostable polymerase such as Thermus aquaticus (Taq) DNA polymerase (Arneheim and Erlich, 1992). Several factors present in plant DNA preparations inhibit Taq polymerase activity. Some classes of polysaccharides reduce the
activity of polymerases, ligases and restriction endonucleases (Do
and Adams, 1991). As a result of contamination by polysaccharides
and/or other DNA binding substances, false negative polymorphic
bands have been observed in PCR based fingerprinting, which may
confound the interpretation of genetic differences between
individuals.

By using PCR in molecular genetics, several methods have
been developed for extraction of DNA from plant materials (Varadarajan
et al., 1991). Which allows the recovery of small amounts of DNA of
sufficient purity for PCR amplification. It can be achieved using
organic or inorganic solvents. However, inhibitions of PCR by traces
of organic solvents as well as their toxicity are the drawbacks of
protocols involving organic solvents, thus a critical factor in the isolation
of plant DNA is the efficient disruption of the plant cell wall and the
separation of DNA from other cell components without affecting the
integrity of DNA. Unfortunately, many techniques for breaking open
cells also shear DNA and thus any method must be a compromise
between DNA lengths and yield (Tapan et al., 2000).

2.7.1 DNA Isolation Protocols

Murray and Thompson, (1980) used a relatively simple
procedure for the rapid isolation of high molecular weight DNA
(50kb), which is free of contaminants, needed for various
applications such as PCR, cloning etc. This protocol can be used
for crops, which are rich in polysaccharides, phenols and other
secondary compounds. In this method nucleic acids can be
selectively precipitated with Cetyl Trimethyl Ammonium
Bromide (CTAB). RNA and DNA are soluble in CTAB and NaCl
(0.7 M), but precipitate when the salt is reduced below 0.4M.
However, many polysaccharides are insoluble over this salt range and are thus not solubilized. The CTAB precipitated DNA was again purified with CsCl (Cesium chloride of different density). This approach has been used to isolate DNA from carrot, wheat, oats, tobacco and peas. Results indicated that dry tissue can be effectively disrupted while the DNA is unhydrated and thus less susceptible to shear. The protocol yielded 20-70 μg of DNA per 100 mg of dry tissue.

Dellaporta et al., (1983) reported a rapid micro scale method for isolation of plant DNA without the use of ultra centrifugation with CsCl. This is a mini-preparation procedure that can be followed for crops less in polyphenols, polysaccharides, and other secondary compounds. In this method SDS is used to remove proteins and has been successfully used on Nicotiana tobacum, Lycopersicon sp, Amaranthus sp, Petunia hybrida and Glycine max. The DNA produced is of moderately high molecular weight (>50kb) and serves as a satisfactory substrate for most restriction enzymes and is suitable for genome blot analysis. DNA yields from leaf tissues of most species tested with this protocol were 50 to 100 μg per gram of fresh tissue and remarkably uniform from sample to sample.

Although the protocols involving the use of fresh frozen/lyophilized tissue yielded good quantities of high molecular weight DNA the use of liquid nitrogen / liophilizers presented some problems. To avoid such problems Tai and Tanksley (1990) have developed an inexpensive method for dehydration of plant tissue and isolation of DNA from dehydrated leaves of tomato by using protocol...
given by Dellaporta et al., (1983) in which drying was carried out by passing hot air over tissue in food dehydrator. Consistent result has been obtained from tissue dried at 45°C to 55°C (temperature measured initially during drying) for 12 to 24 hours. Subsequently tissue was powdered with the use of paint mixers. The DNA obtained was comparable to the one obtained fresh in respects to both quality and quantity.

DNA isolated from muskmelon cucumber, potato and geranium according to the protocol of Dellaporta et al., (1983) were often contaminated with large amounts of polysaccharides. Fang et al., (1992) demonstrated a quick and inexpensive method to remove polysaccharides contamination form such DNA preparations. The contamination was removed by re dissolving the DNA in TE with 0.5 to 3.0 M NaCl and then precipitation with two volumes of ethanol. Most of polysaccharides were removed effectively in a single high salt precipitation at 1.0 to 2.5M NaCl At 3.0M NaCl the salt precipitated out of solution whereas higher concentration resulted in salt precipitation that was difficult to be removed. Purified DNA was satisfactory as template for PCR.

Porebski et al., (1997) described a relatively quick, inexpensive and consistent protocol for extraction of DNA from expanded leaf material containing large amounts of polysaccharides, tannins and polyphenols. Mature strawberry leaves, which contain high levels of these secondary compounds, were used as study group. The method is a modified CTAB method involving high salt concentration to remove polysaccharides. They used PVP to remove polyphenols, extended RNAase treatment and phenol-chloroform step. The average yields from this protocol ranged from 20 to 48μg per gram mature leaf tissue for Assessment of Genetic diversity in Santalum album L. using Random Amplified Polymorphic DNA and development of specific primers S.A. AZEEZ
both octaploid and dioploid species. The resulted DNA was PCR amplifiable before and after extended storage.

A simple, fast and efficient method for DNA purification from mature leaf samples of four hard wood tree species (*Acer, Fraxinus, Purnus and Querous*) was described by Leofoprt and Douglas (1999). The protocol was modified CTAB method which involved combination of β-mercaptaethanol, PVP, SDS and LiCl including short centrifugation runs. The protocol yielded 950μg per gram fresh weight even when very matured leaves are processed (Bhattacharaya *et al.*, 1999).

### 2.7.2 Protocols for difficult to isolate crops

Problem to isolate quality DNA from the genus *Abelmoschus* was stickiness of leaves even after grinding the leaves in liquid nitrogen which was because of large amounts of polysaccharides produced during photosynthesis. Kolcko and Hamon (1990) used cotyledons of dark grown seedling to overcome this problem. About one gram of seven days old unopened cotyledons yielded 100 μg of DNA. The quality of DNA was confirmed by digesting it with several restriction endonucleases. Isolation of highly purified DNA is difficult particularly from plants rich in polyphenolic compounds because damage to the leaves causes significant buildup of polyphenolic complexes.

Jennifer and Paul (1990) developed a protocol for isolation of DNA from plants like cocoa which are high in polyphenols. This method relies on both concentrating the nuclei away from cytoplasmic...
components prior to lysis and strongly inhibiting, the formation of oxidized polyphenolic compounds in the extraction solution. The DNA obtained was relatively pure, yielding 20 µg of DNA per gram of starting material by adopting the technique given by Sambrook et al., (1989) the DNA had $A_{260}/A_{280}$ ratio of 1.76 concentration that the DNA was relatively pure.

Sweet potato and its related species of Convolvulaceae are characterized by several unusual secondary metabolites including iridoid compounds and flavonoids in addition to latex/mucilaginous secretion and some unidentified phenols. Vardarajan and Prakash (1991) adopted SDS method of minpreparation protocol given by Dellaporata et al., (1983) with three main departures in order to obtain DNA, which is free from such contaminants.

Extraction of high quality DNA from plant species like Opuntia and other cacti is difficult. These plants contain high levels of mucilages and complex polysaccharides that bind water present in the extraction buffer producing a gel like mixture, thus preventing isolation of quality DNA. Jacobo et al., (2000) developed a modified CTAB method for extraction of enzyme restrictable DNA from four Opuntia sp. The method involves adjusting the amount of tissue used according to species and age, followed by processing in an extraction buffer to separate coarse material. Extended centrifugation and digestion in a separation buffer with 2 per cent CTAB was found to maintain polysaccharides in solution and allowed easier recovery of the aqueous phase containing DNA. This method was successful in extracting DNA from tender tissue of other cacti species like Stenocereus sp., Cleistocactus sp., and
Echinocereus sp. The DNA obtained was suitable for PCR amplification.

2.7.3. Effect of tissue and method of isolation on PCR

Rogers et al., (1996) evaluated six methods of DNA extraction on PCR amplification, using leaf discs and other plant tissues such as seed, root and tubers from six plant species (sugar beet, sea beet, Brassica oleracea, rapeseed, potato and maize). Comparison using leaf material indicated differences among species in the PCR success rate and reliability of the tested methods. However, method two was found to be applicable to all the species tested, while other methods were found effective only for a particular species. Further, DNA extracted from non-leaf tissues (root, seed and tubers) showed low success rates of PCR amplification.

Boiteux et al., (1999) evaluated seven plant genomic DNA purification protocols for genetic fingerprinting analysis using six tissues obtained from inbred carrot lines. Evaluations included DNA yield, purity, DNA cleavage with Hind III, DNA integrity and DNA suitability for RAPD. Significant differences were observed among tissues and purification methods for the total amount of DNA. An extraction method using CTAB buffer + organic solvents gave the best results in DNA yield, purity and Hind III cleavage when compared with other six non-organic extraction methods of the tissues examined.

Differences in RAPD band intensity and number were observed across tissues and DNA extraction methods using identical PCR conditions for RAPD. Callus was the best type of tissue...
for RAPD based fingerprinting producing a consistently higher number of more intense amplicons compared to other tissues (Strim, 1990).

2.7.4 DNA quantification and quality assessment

DNA quantification can be done by flurometry, spectrometry and agarose gel electrophoresis by comparing with standard DNA concentrations (Boiteux et al., 1999), while, the purity is assessed based on digestion with restriction endonucleases (PstI, EcoRI, HindIII, BamHI etc.) and spectrophotometer readings. Quality refers to purity of DNA and to what extent it is free from secondary metabolites and other contaminants, which hinders its use in molecular genetics. A good quality DNA generally exhibits the following spectral properties; $A_{230}/A_{260}$, $A_{280}/A_{260}$ or $A_{260}/A_{280}$ ratios of less than 0.10, less than 0.45, less than 1.65 or more than 1.80, respectively (Shantha et al., 1998). If $A_{260}/A_{280}$ ratio is more than 1.80, then it indicates RNA contamination and if the ratio is less than 1.65, indicates the presence of contaminants such as proteins (Sambrook et al., 1989).

2.8 Genetic markers and their types

Molecular genetic technique have been applied to plant cultivar identification in the past decade by developing molecular markers that detect difference in DNA sequences between cultivars (Thomas et al., 1993). Highly specific marker profiles commonly know as DNA fingerprints, can be developed for each cultivar and used for its identification. Compared to isozymes and flavonoid markers (Nilson et al., 1985). DNA markers have many advantages. They are detectable in all tissues at all ages of the plant, there by enabling early identification.
They are virtually insensitive to epistatic pleotropic effect (Williams et al., 1992). They are developmentally stable and are not affected by the environment. Further more, the number of DNA markers available for examination is unlimited because genomes are composed of base pairs and many different types of molecular marker systems are available to compare these differences.

Trigiano and Caetano-Anolles (1998) described that the DNA fingerprinting is a bar code like DNA fragment pattern generated by multilocus probes after separation of genomic DNA fragments. They were considered to be the ultimate tool for biological individualization. Recently the term DNA fingerprinting profiling is used to describe the combined use of several single locus detection system and is being used as versatile tools for investigating various aspects of the plant genome. These include characterization of genetic variability, genome fingerprinting, genome mapping, gene localization, analysis of genome evolution, taxonomy and plant breeding (Weeden et al., 1992).

Joshi et al (1999) reported about the properties for ideal DNA markers. According to them they should be highly polymorphic in nature, co-dominantly inherited, frequently occur in the genome, selectively neutral in behavior, highly reproducible and easy exchangeable data between laboratories. Genetic polymorphism is classically of two or more discontinuous variant or genotypes.

Most higher plants have considerable portion of DNA as repetitive non-coding DNA that is not transcribed. Species with larger genomes normally have more repeated DNA and higher portion of repeated DNA are single copy DNA ( Tanksely and Pichesky, 1988). Thus only a small fraction of the total genetic variation at a DNA
nucleotide sequence level reveals itself as a distinct trait showing Mendelian inheritance. The degeneracy of the genetic code ensures that about one in three nucleotides changes will not affect the amino acid sequence of the protein. Thus the great bulk of genetic variation at the nucleotide level may have detectable expression at phenotypic level. It is this genetic variation that is expressed as DNA based polymorphism. Genetic marker can be defined as any stable and inherited variation that can be measured or detected by a suitable method, and can be used subsequently to detect the presence of a specific genotype or phenotype, which otherwise is very difficult to detect (Mitra et al., 1999). Such variations occurring at different levels, i.e., at the morphological, chromosomal, biochemical or DNA level can serve as the genetic markers (Marsolais et al., 1993).

2.8.1 Morphological markers

Phenotypic traits are the oldest and widely used genetic markers that can be used for the evaluation of diversity. They are very informative in germplasm management where the cultivars can be identified based morphological traits viz., seedling and leaf characters, flower types, fruit size, shape, colour etc. Morphological markers are simple and inexpensive which can be studied from herbarium specimen or other dead tissues. But the disadvantage of this marker is phenotypic characters are highly influenced by environmental conditions. Hence, the data obtained by such evaluations are not easily understood at genetic level, often resulting in maintenance of duplicate accessions. Moreover, in perennial crops it is required to grow for many years to observe growth and reproductive parameters, which is economically not feasible. For meaningful assessment of genetic diversity a large number
of polymorphic markers are required this limits the use of morphological markers that reveals a low degree of polymorphism (Mitra et al., 1999). The actual identity of some cultivars is still in question, because similar cultivars grown in different regions often have different names (Pathak and Ojha, 1993) because of the differences in the manifestation of morphological traits (Lakshminarayana, 1980). Cruz et al., (1999) estimated the genetic diversity among 326 accessions of 3 species of Dioscorea viz., D. alia, D. esculenta and D. hispida with 113 agromorphological characters from different regions of Philippines. Multivariate analysis was carried out and the PCA and dendrogram revealed the relationship between the three species. The PCA and dendrogram revealed distinct and tight clusters. Eight accessions clustered next to the cluster of D. alta and were reclassified as D. bulbifom. The lowest genetic diversity was observed in D. esculentum and highest in D. hispida. Morphological characterization of 35 accessions of Arachis comprising of 13 species was carried out with 32 qualitative (20 vegetative and 12 reproductive) and 33 quantitative (22 vegetative and 11 reproductive) characters (Chandran and Pandya, 2000). Thirteen clusters were formed on the dendrogram four of which were accessions from A. duranensis.

Morphologies reflect not only the genetic constitution of the cultivars, but also the interaction of the genotype with the environment (G X E) within which it is expressed. Due to G X E effects, it is clearly inappropriate to compare morphological data for varieties that have been collected across different years and/or locations.
2.8.2 Molecular markers

Recent developments in DNA technologies have made it possible to uncover a large number of genetic polymorphisms at the DNA level, and to use them as markers for evaluation of the genetic basis for the observed phenotypic variability. These markers which reveal the variations at the DNA level are referred to as molecular markers (Mitra et al., 1999). Ever since their development, they are constantly being modified to enhance their utility and to bring about automation in the process of genome analysis (Swati et al., 1999).

2.8.2.1 Biochemical markers

These are in popular usage because the variations for these markers are ubiquitous and these variations can be understood in genetic terms. Proteins are molecules with net electrical charges that are effected by pH. They can be separated by electrophorosis (Bietz., 1986). Isoenzymes also called as isozymes are multiple molecular forms of enzyme occurring within the same organism, having identical catalytic activities and separable by electrophoretic procedure (Feret and Bergmann, 1976). The isoenzyme technique can be used for identification of clones, hybrids or varieties of tree species, for early prediction of quality characters of forest Species, genetic analysis in breeding studies (Parthasarathi et al., 1985) and in studying gene flow in stands and seed orchards (Susan, 1993).

2.8.2.1.1 Isoenzyme analysis in sandalwood research

Among the spectrophoretic techniques, isoenzyme analysis is widely used for its efficiency in intra specific variability studies; Phenotypic characterization of sandalwood trees at cell level by Parthasarathi et al. (1985) revealed characteristic differences between leaf...
types, both at vegetative and flowering stages and their pattern of isoenzymes (POD-peroxidase and MDH-malate dehydrogenase). It is also helpful in diagnosing the deficiency symptoms of a particular element and restoring normalcy by providing that trace element at the proper time (Kamala et al., 1986). Apart from this, change in the pattern of POD and MDH in the diseased plant could be used to confirm the presence of spike disease in doubtful cases and even diagnosing at an early stage of onset of disease (Angadi and Ananthapadmanabha, 1988).

Isoenzyme analysis is of great use in identifying provenances (Egerton-Warburton, 1990), estimating generic distance between sandal plants with in the same population and between different population (Brand, 1994) and also in genetic resource mapping of highly distributed S. album population (Venu et al., 1997).

### 2.8 2.2 DNA markers

Study of DNA variation in forest genetics is not a new phenomenon, but the recent approach is to develop appropriate system of markers to examine the diversity, genetic differentiation and gene flow in fragmented populations. Information emerging from such studies is useful in evolutionary studies, linkage mapping, in identifying genes responsible for disease resistance, map based cloning and formulating proper conservation strategies (Joshi et al., 1999).

The most exciting feature of DNA based markers is the extent of detectable polymorphism. The DNA content in higher plants is highly variable. Aurumuganathan and Earle (1991), estimated the DNA content of more than 100 important crop species. The DNA content varied from 0.30 picograms (pg) per 2c nuclei or 145 million base pairs.
(mbp) in *Arabidopsis* to over 50 pg or 24.255 mbp. However the DNA content of most of the intensively mapped diploid species (tomato, rice, arabdopsis) is in the range from 0.30 to 1.0 pg. However in eucalyptus, which belongs to the same family as guava (Myrtaceae) has a DNA content of 530 mbp 0.75 pg.

DNA based markers can be classified into:

a) Restriction based (RFLP)

b) PCR based (RAPD)

c) Both restriction and PCR based (AFLP)

Among wide array of DNA markers, RFLP, RAPD and AFLP are most commonly used in crop plants. Eucalyptus belongs to the family ‘Myrtaceae’ therefore literature available on the same has been included.

2.8.2.2.1 Restriction Fragment Length Polymorphism (RFLP)

RFLP analysis involves digesting the DNA of the subject genome with cleaving restriction endonucleases (Grodzicker *et al.*, 1974). Fractionating the fragments electrophoretically and then preferentially visualizing fragments containing particular homologous sequences by hybridizing them with a specific DNA probe. RFLP reveals length differences.

Byrne *et al.*, (1998) assessed the genetic variability in the nuclear genome of Eucalyptus nitens using 40 anonymous RFLP loci. Information about the RFLP probes was available form the linkage map.
already developed for the genome. Several loci that were monomorphic in the mapping pedigree were also used to avoid bias towards the polymorphic loci. Variation was determined for 20 individuals from each of eight population covering the natural range of the species. The proportion of rare allele detected in the population ranged from 21 to 49 per cent. The level of diversity within the nuclear genome was high as 6.9 (Ps 92.5% and Hs 0.445) compared to other species where RFLP (Thormann et al., 1994), loci than for isozyme loci. But the patterns for genetic diversity were similar for both marker types. Population similarities based on RFLP analysis of the nuclear genome were not concordant with those based on the chloroplast.

Byrne (1999) estimated the diversity among three Oil malles taxa of eucalyptus, viz., Eucalyptus kochii, E. Planissima and E. hortistes with RFLP markers. The diversity was analyzed by using 20 individuals from 10 populations representing the three taxa. The level of diversity in the taxa was high when compared to other species where RFLP surveys have been conducted on eucalyptus species. The populations showed little differentiation, low genetic distance and high gene flow between them. Phylogenetic relationships revealed clustering of the population based on geographical distribution. This study revealed that the taxa represent a single widespread variable species with some provenance delineation.

RFLP analysis also has been used in DNA fingerprinting and estimation of diversity in grapes (Bowers et al., 1993), and (Retief et al., 2005). In apple for paternity analysis (Nybom and Schaal, 1999) and in the construction of genetic linkages maps in many crops like tomato (Tanksley, et al., 1988, Peterson et al., 1988), pepper (Tanksley et al., 1988) and other crops.
In species where surveys of isozyme and RFLP variations have been compared, RFLP loci tend to be more polymorphic (Harvey and Muelbaner, 1989). Some studies have found differences in genetic structure and partitioning of the variation between isozyme and RFLP loci (Zhang et al., 1993) and the differences may reflect different evolutionary pressures on isozyme loci versus non-coding regions of DNA.

RFLP is a co-dominantly inherited marker system (it can distinguish heterozygous from two parents). It is relatively reliable and reveals high allelic diversity, but it is time consuming, expensive, labor intensive and involves use of radioactive chemicals, which are harmful to the users. It is slow because it needs more time for the construction of probes, processing the plant materials for high quality and quantity (μg) DNA. Most RFLP probes are single copy probes that reveal one DNA fragment per probe. While recently evolved PCR based molecular markers (Reisch et al., 2000), like RAPD and AFLP reveal several bands per reaction. RFLP cannot be used in crops with very low level of restriction fragments or for which RFLP probe sets have been generally not available. It is also not best method for applications involving more individuals such as screening a breeding population with marker linked with disease resistance (Walton 1993).

2.8.2.2.1 Microsatellites and Minisatellite markers

The term microsatellite (SSRs - Simple Sequence Repeats) was coined by Litt and Luty (1989), while the term minisatellite (VNTRS - Variable Number of Tandem Repeats) was coined by Jeffreys et al., (1985). Both are multilocus probes creating complex
banding patterns and are usually non-species specific occurring ubiquitously. They form an ideal marker system for DNA fingerprinting in forensic studies as well as in genome mapping (Jeffreys et al., 1991). Fingerprints generated by these probes are also known as oligonucleotide fingerprints. The methodology has been derived from RFLP and the specific fragments are visualized by hybridization with a labeled micro/minisatellite probes. The DNA content in higher plants is highly variable. Anupam Dixit et al., (2005), estimated the DNA content in over 100 important crop species. DNA content varied from 0.30 picogram (pg) 2 diploid nuclei or 145 million base pairs (mbp) in Arabidopsis to over 50 pg or 24/255 ml p in leek. The genomic DNA content of sandalwood (Santalum album L.) was estimated by Flow Cytometry- by using leaf tissue by Dr. Arumuganathan, University of Nebraska, Lincoln, USA to be 0.4 pg 2 diploid nuclei. Higher plants have a considerable portion of DNA as repetitive in-coding DNA that is not transcribed. Species with larger genomes normally have more repeated DNA and a higher proportion of repeated DNA to single copy (Tanksley and Pichersky, 1988). Thus, only a small fraction of the total genetic variation at a DNA nucleotide sequence level reveals itself as a distinct trail, showing Mendelian inheritance. The degeneracy of the genetic code ensures that about one in three nucleotide changes will not affect the amino acid sequence of the protein produced. Thus, the great bulk of genetic variation at the nucleotide level may not have any detectable expression at phenotypic level. SSRs are used to detect polymorphisms in these tandem repeats. They can be hybridized, based, or PCR based where the primer(s) used flank-repeated sequences. SSRs are co-dominant and relatively easy to perform and are extensively used in marker studies.

Assessment of Genetic diversity in Santalum album L. using Random Amplified Polymorphic DNA and development of specific primers

S.A. AZEEZ
2.8.2.2.1.2 Sequence Tagged Sites (STSs)

RFLP probes specifically linked to a desirable trait can be converted into PCR based STS markers, based on nucleotide sequence of the probe giving polymorphic banding pattern to obtain specific amplicons. This technique overcomes the tedious hybridization procedure involved in RFLP analysis and is extremely useful for studying the relationship between normal different species (Bustos et al., 1999). When these markers are linked to some specific traits like stem rust resistance gene in barley, they can be easily integrated into plant breeding programmes for marker assisted selection of the trait of interest.

2.8.2.2.2 Randomly Amplified Polymorphic DNA (RAPD)

RAPD method of DNA analysis was developed independently by two different laboratories (Welsh and McClelland, 1990; Williams et al., 1990). This technique is now being used extensively to complement RFLP. This procedure detects nucleotide sequence polymorphism in DNA amplification assay using only a single primer of arbitrary nucleotide sequence using PCR. The major advantage of this assay over RFLP method is that (1) there is no prior requirement for DNA sequence information of the genome. (2) A universal set of primers can be used for species. (3) Only primer sequence is needed for information transfer. (4) The process can be automated (Williams et al., 1990). The protocol is relatively quick and easy to perform and uses fluorescence in place of radioactivity (Williams et al., 1992). As the RAPD is amplification based assay only nanograms quantities of DNA are required.
The PCR Program used for RAPD analysis is a typical PCR program except the stringency during primer annealing is much lower than usual because of the primer used in the RAPD procedure is relatively short and has an arbitrary sequence. For random 10-mers, Primer annealing at 36°C works well for many organisms, including plants. The procedure involves isolation of DNA, PCR amplification of DNA and scoring the presence or absence of bands, which are, resolved agarose gel. The limitation of the use of RAPD markers is that they are dominant. (Xu et al., 1993). This can be overcome by using more than one closely linked marker (Williams et al., 1990).

2.8.2.2.3 Amplified fragment length polymorphism (AFLP)

AFLP combines the reliability of RFLP with power of PCR technique. It allows the reliable identification of over 50 loci in a single reaction. It can be applied for DNAs of any origin and complexity. Fingerprints can be produced without prior knowledge using limited sets of generic primers. It is robust and reliable because stringent reaction conditions are used for primer annealing (Vos et al., 1995).

It is a very recent technique and is based on the detection of genomic restriction fragments by PCR amplification. It involves digestion of DNA with two restriction enzymes (one frequent cutter and a rare cutter), enzyme specific oligonucleotide adopters of known sequence are ligated to each end. Restriction fragments are then amplified using enzyme adopter specific primers that are complimentary to adopter sequence and remains at the restriction site plus up to usually one to four random nucleotides at 3' end. The number of restriction fragments amplified will be determined by the complexity of the

Assessment of Genetic diversity in Santalum album L. using Random Amplified Polymorphic DNA and development of specific primers

S.A. AZEEZ
template and it can also be tuned by selection of specific primer sets with selective extension. It involves two amplifications, pre-amplification and selective amplification. Pre-amplification will be carried out using a primer set with one selective extension. Pre-amplified DNA will be used as template for second selective amplification, which will be carried out with primer sets with three to four nucleotide extension. The radiolabelled PCR products are size fractionated on a polyacrylamide gel (Hill et al., 1996).

It has been used in assessing genetic diversity of many crops like soybean (Maughan et al., 1996), tea (Paul et al., 1997), Eucalyptus (Marques et al., 1998), Azadiracta indica (Sing et al., 1999), Pinus (Lerceteau and Szmidt, 1999) and gooseberries (Lanham and Brennan, 1999).

For AFLP additional DNA purification steps are necessary, as it needs high quality and quantity DNA. It is somewhat lengthy process as it includes two amplifications. It is costly because it involves use of restriction endonucelases, enzyme specific adopters, enzyme adopter specific primers with selective extensions, ligators and use of polyacrylamide gel for electrophoresis (Vos et al., 1995). However it is highly informative and economic than any other molecular marker at present (Hill et al., 1996).
2.8.2.3 Application of RAPD markers

2.8.2.3.1 Estimating genetic diversity

Molecular classification of malus involving 20 wild species and 27 cultivars was done using RAPD markers (Dunemann et al., 1994). DNA fingerprints generated fragments and coefficients estimating genetic similarity were calculated on the basis of about 50 polymorphic RAPD loci in each set of genotypes. A high degree of genetic diversity was found at the molecular level among both different apple cultivars and wild species of the genus *Malus*. The results gave additional evidence for the hypothesis that *M. pumila* and *M. sylvestris* were involved in the origin of the cultivated apples.

In another study, Yae et al., (1995) classified *Malus domestica* cultivars using random amplified polymorphic DNA markers. Forty apple (*M. pumila*) cultivars were compared by RAPD analysis using 16 random primer. Of the 139 clear and reproducible bands, 106 were polymorphic. From these polymorphisms, the cultivars were divided into 6 groups by cluster analysis. Group I contained *Rall s janet* Fuji and the bud mutations of Fuji group II contained Sikaiichi. EarliBlaze. Delicious and its strains formed group IV: 'Jonathan', 'Jonared', 'kogetsu' and 'Mollies Delicious' formed group V; and group VI contained only 'Spur Golden Delicious'.

Graham and McNicol (1995) examined the ability of RAPD markers to determine the relationships within and between *Rubus* species. RAPD markers were generated from 13 different *Rubus* species in order to assess the degree of similarity between species from the important subgenera *Idaebats*, *Eubats* and *Anopl bats*. All ten primers revealed
scorable polymorphisms within both the closely related and the genetically diverse individuals. Some 372 markers were generated and scored from the material analysed. Estimates of similarity, dendrograms and principal coordinate analysis were calculated, with the results generally being in agreement with previous classifications of the species studied, confirming the validity and usefulness of the RAPD method. However, amongst the species studied, R. madtsei (subgenera Idaebats) proved more diverse and grouped in with both the subgenera Idaebats and Eubats at only 26 per cent similarity.

Xiang et al. (1996) analysed peach rootstock cultivars using RAPD markers using synthetic decamer oligonucleotide primers. Based on combined banding patterns, all 18 rootstocks using 40 markers produced a dendrogram of genetic relatedness which is in good agreement with their putative pedigrees. The first and major bifurcation in dendrogram divided these rootstock cultivars into two groups according to their resistance or susceptibility to root-knot nematodes.

Machado et al., (1996) detected genetic relationships among 39 Mediterranean mandarins (Citrus deliciosa tenora) genotypes using 21 random primers. Genotype-specific RAPD markers were found in mainly known hybrids. Un-weighted Pair Group Method with Arithmetic Average (UPGMA) cluster analysis revealed a low level of genetic variation between accessions of Mediterranean mandarins. Whereas, the hybrids with other Citrus species showed greater genetic dissimilarity (Ling, 2000). Twenty accessions yielded similar patterns, suggesting either they could be a single clone. Or that the technique was not able to detect genomic variation. However, for the other genetic

Assessment of Genetic diversity in Santalum albidum L. using Random Amplified Polymorphic DNA and development of specific primers
S.A. AZEEZ
variation between accessions was quite low. The large number of hybrids and the low level of polymorphism between accessions support the hypothesis that Mediterranean mandarins are all true hybrids of common mandarins (*Citrus reticulata*) (Sawazaki *et al.*, 1997).

Graham *et al.* (1996) estimated genetic diversity in eight strawberry cultivars from four breeding programmers around the world. Ten random primers (RAPD) successfully amplified DNA fragments from each cultivar and specific fingerprints were generated from the molecular marker data. The cultivars were traced back to founding clones and the relationships between the cultivars were examined from both the molecular and the pedigree data.

Steiner *et al.*, (1998) examined 37 clove taxa via PCR using 2 RAPD markers. Plants grown in greenhouse were also observed for inflorescence color and relative flowering time. Cluster analysis of pairwise means of RAPD marker distances and cultivar pedigree histories revealed similar genetic backgrounds among most cultivars. Flower color associated with RAPD marker grouping, was not related to flowering. Both red vs. non-red inflorescence color and early vs. late flowering traits could be predicted by specific RAPD bands (95 and 93% accuracy, respectively).

RAPD markers were used in cashew to assess the diversity among the varieties and hybrids released in India by Murali (1999). Of the one hundred and twenty random primers screened, ten were chosen to score for 157 bands. The results indicated that the diversity among cashew cultivars and hybrids was moderate. In an attempt to study genetic diversity among fifty mango varieties (Hemanth *et al.*, 2001),
eighty primers were screened, ten of which gave 139 bands. UPGMA analysis based on Jaccard's coefficient revealed a moderate to high degree of genetic diversity among mango cultivars.

Kim et al. (1998) used RAPD markers to assess the genetic diversity in East-Asian Pinus species. The cluster analysis indicated that East-Asian pines had higher levels of genetic diversity than pine trees from Europe and North America. Among the different species studied, Diploxyylon pines (natural hybrid of P. tabulaeformis and P. yumumensis) showed higher levels of genetic diversity than Haploxyylon pines (P. pitmiln, P. sibirica and P. koraiensis). Genetic variation at the polymorphic loci was partitioned such that 96 per cent of the diversity was found within the population, while 4 per cent accounted for population differentiation, indirectly indicating that gene flow among Asian pine tree populations is high enough to counteract the effect of genetic drift.

Neem is an evergreen, multipurpose tree of the tropics and is believed to be highly cross-pollinated. RAPD markers were used to study inter-provenance variation among 34 accessions of neem. Similarity matrix computed based on UPGMA method revealed that the similarities in RAPD profile amongst the different accessions was more than that expected due to the out crossing nature of neem. The results indicated that neem has a narrow genetic base and further improvement of this tree species may require the introduction of additional germplasm into breeding programmes (Nuzhat, RAPD markers were used to determine intraspecific variability among 15 accessions of...
Andrographis paniculate (a popular antipyretic and hepatoprotective drug) collected from India and south-cast Asia (Padmesh et al., 1999), Molecular analysis revealed a moderate level of variation within the species. Similarity measurement using UPGMA followed by cluster analysis resulted in five major groups based on geographical distribution that generally reflected expected trends between the genotypes. There were also important exceptions like accession AP-48 from Thailand showing close resemblance to AP-38 from Tamil Nadu and AP-29 from Assam significantly diverse from the rest of the native genotypes.

Stevens et al., (1999) assessed variation in genetic diversity among 23 local populations of populus seedlings with respect to altitude, substrate, plant competition and geographical location. One hundred and thirty two RAPD markers based on six primers were used to assess genetic relationships within and among the population. The results indicated that most of the variation occurred with in the population. Even genetically distant populations shared nearly 90 per cent of the same markers and there was no geographical differentiation among the population.

Thirty-four released cashew varieties and hybrids from India and a clone resistant to tea mosquito have been fingerprinted by using RAPD markers (Mneney et al., 2001). One hundred and fifty seven amplified products generated by ten random primers were used to estimate genetic relationship and the results indicated that the genetic base of the cashew cultivars is not narrow as reported earlier, but is moderate. The clustering of the cultivars was more or less similar to their geographical origin.
Genetic diversity among 12 genotypes of gooseberry (Ribes, was evaluated using RAPD, SSR and AFLP markers (Lanham and Brennan, 1999). The similarity matrices and dendrograms constructed revealed the genetic similarity between the different genotypes. Clustering varied with the type of marker used. The genetic base of the European cultivated gooseberry was reported to be narrow and the need for introduction of additional germplasm into breeding programmes was emphasized.

Lim et al., (1999) performed RAPD analysis to evaluate genetic relatedness among 12 species of Vanda. One hundred and fifty eight fragments generated by eight primers were used for statistical analysis. The dendrogram constructed by UPGMA analysis formed two major clusters. V. teres and V. hoakeriana formed one of the clusters while the other 10 species formed another cluster. The genus Ascocentrum appears to be closely related to the Vanda species.

Saikia et al. (2000) performed RAPD analysis to estimate genetic diversity between 24 yew trees (has anti tumor and anti cancer properties) from the Himalayan region. Forty-eight polymorphic bands generated by 13 primers were used for RAPD analysis. In the dendogram the accessions AP 151 and WB 214 clustered separately and were distinct from the others. The accessions from Arunachal Pradesh area were more diverse than the accessions from Darjeeling and West Bengal. Even though many of the accessions could be differentiated based on their RAPD profiles (Matsumoto et al., 1996), the accessions assessed are believed to have a narrow genetic base.
Yang and Li (2000) used RAPD markers to estimate genetic variability in *Eucalyptus microileica*, a native Australian species grown in the arid and semiarid zones. One hundred and two polymorphic bands generated by 18 primers were used to analyze twelve natural populations from widely separated locations. Gene diversity values for each population ranged between 0.176 (the population from southeastern Australia) and 0.232 (the population from Western Australia) with an average of 0.200. Total gene diversity for this species was 0.240, where 83.3 per cent of the variation was found within populations and 16.7 per cent between populations.

RAPD markers were used to evaluate genetic diversity and to identify interspecific hybrids among citrus germplasm collections from Embrapa-Mandioca (Yonezawa, 1995). Thirty accessions were characterized by using 20 random primers and a fingerprint of each of the cultivars was obtained. Cluster analysis showed five distinct groups of genotypes, which agreed with the data on their origin and taxonomic classification. In the hybrid analysis, six primers were available which showed polymorphism between Volkamer and Rangpur (Vilarinhos et al., 2000).

Gemas et al., (2000) performed RAPD analysis to characterize and to estimate genetic diversity among the three commercial varieties of *Olea europaea* L. viz., 'Galega Vulgar' (GV), 'Cordovil de Serpa' (CS), and 'Verdeal Alentejana' (VA). One hundred and fifty six amplified products generated by 20 decamer primer were used to assess diversity among 28 trees of CS, 28 trees of GV and 22 VA trees. The dendrogram constructed clearly formed three distinct clusters for CS, GV and VA types.
Ravishankar *et al.*, (2000) assessed the genetic relatedness among 18 commercial mango cultivars from India using RAPD markers. Out of 30 primers screened 19 of them produced a total of 178 bands (130 polymorphic and 48 monomorphic), which were used to study the genetic relatedness. The cultivars from western, northern and eastern India clustered together, while the south Indian cultivars clustered separately. In another study, fifty commercial mango cultivars from different parts of India were fingerprinted by using 139 RAPD markers generated by 10 decamer primers and the genetic distance among the cultivars was found to be moderate to high. The mango hybrids with a common parent clustered together while the regular ami irregular bearers clustered separately. The variety 'Mulgoa' was found to be distinct from all the cultivars studied and the south Indian cultivars were more diverse than the other cultivars (Kumar *et al.*, 2001).

Apple Simple Sequence Repeats (SSRs) were intergenerically applied to the characterization of 36 pear accessions, including 19 Japanese pear (*Pyrus pyrifolia*), 7 Chinese pear (*P. bretschneideri, P. ussuriensis*), 5 European pear (*P. communis*), 3 wild relatives (*P. calleryana*) and 2 hybrids. All the tested SSRs from apple produced discrete amplified products in pear that could be used to differentiate the accession and to estimate the diversity (Yamamoto *et al.*, 2001).

Kafkas and Perl-Treves (2001) studied the taxonomic relationships and genetic variation between *Pistacia vera* and wild species, *P. atlantic, P. terebinthus P. eurycarpa*, and *P. khinjuk* with RAPD markers. Forty wild *Pistacia* genotypes and two *P. vera* varieties generated 138 bands with 10 arbitrary primers. The
dendrogram revealed that *P. terebinthus* was the most divergent species and clustered separately, while *P. vera* and *P. eurycarpa* were closely related. Species-specific bands were identified for each of the four species. From their results they suggested that *P. eurycarpa* may be a hybrid between *P. khinjuki* and *P. atlantica*, and the close relationship of *P. vera* and *P. khinjuki* confirmed the research findings of the earlier workers who suggested that they are one species based on the analysis of chloroplast DNA. Pillay, *et al.* (2001) evaluated genetic diversity and phylogenetic relationships of 29 East African highland banana (*Musa* spp.) cultivars and 2 out group taxa, *M. acuminata* Calcutta-4 and Agbagba using RAPD markers. RAPDs revealed a narrow genetic base in highland germplasm of East Africa though a high level of morphological variability exists (Gallego, 1996). This disaccordance was attributed to the fact that the RAPD primers used did not anneal to areas of the genome responsible for morphological variation. The major cluster in the dendrogram consisted of all the AAA types with the two beer varieties (Isha and Ikigeregere) and two cooking varieties (Igisahira and Kibungo). The cultivars 'Calcutta 4' (AA), 'Kamaramas' (AB) and 'Kisubi' (AB) clustered separately from all the AAA types.

### 2.8.2.3.2 Genetic purity

A study was conducted to determine the genetic purity of hybrid seed in watermelon (*Citullus lanatus*) and tomato (*Lycopersicon esculentum*) using RAPD analysis (Hashizume *et al.*, 1993). Fifty-nine oligonucleotides were screened in inbred parental lines (Ha, HB) currently used for commercial seed production in watermelon. The PCR using nine primers resulted in polymorphism and the primer RAPD-12-12 (5'-ACCACCTGGGCTC-3') generated a
fragment specific to the F1 and male parent (HB), and thus enabling the female parent (HA) to be discriminated from the hybrid.

Rom et al., (1995) used RAPD technology to test purity control of commercial F1 hybrid tomato. DNA from these F1 hybrids ('Naama', 'TY20' and '5692') and their parental lines were subjected to RAPD analysis. Polymorphism between the parents generated paternal specific RAPD markers, enabling a clear distinction to be made between hybrids and their maternal parents. In addition, combination of the polymorphic DNA products generated by these primers exhibited hybrid-specific patterns, enabling each cultivar to be identified. In another study, RAPD analysis was performed to determine genetic purity of F1 hybrid seeds of Cichorium intybus L. by Bellamy et al. (1996). Comparison of PCR products obtained by using 100 arbitrary primers (10 bp) allowed identification of all the lines tested. Several primers produced linespecific RAPD markers, and the differences between the lines were confirmed both on individual heads and young seedlings.

Gardiner et al., 1996 found that information about a rare allele of phosphoglucomutase (PGM) that is shared by 'Braeburn' and 16 percent of cultivars in the New Zealand cultivars collection was combined with historical information about cultivars distribution to select a set of 15 cultivars for more detailed genetic relatedness to the key New Zealand apple 'Braeburn'. DNA from all 16 cultivars were examined by RFLP analysis using 14 probe enzyme combinations and also by RAPD analysis with 39 selected primers. The RFLP and RAPD data excluded a proposal that 'lady Hamilton' and Braeburn' were genetically identical.
Walker and Werner (1997) attempted to analyze the 'Cherokee' rose for its putative hybrids 'Silver Moon' and 'Anemone'. RAPD analysis carried out with sixteen primers produced forty reproducible polymorphic bands. On analyzing the shared bands it was evident that 'Anemone' shared 69.4 per cent of its bands with the 'Cherokee' rose types while 'Silver Moon' shared only 22 per cent of its bands. The results indicated that 'Anemone' is likely to be a progeny of 'Cherokee' rose type.

Heinkel et al. (2000) performed a parental analysis on the plum cultivars 'Cacaks Beauty', 'Cacaks Best', 'Cacaks Early' and 'Cacaks Fruitful' with RAPD markers. It has been earlier reported that cultivars 'Cacaks Beauty', 'Cacaks Best' and 'Cacaks Early' are from 'Wangenheim' x 'Pozegaca' and 'Cacaks Fruitful' from 'Stanley' x 'Pozegaca'. Twenty six selected primers amplified 158, 144 and 160 fragments in 'Wangenheim', 'Pozegaca' and 'Stanley' respectively. Out of these 102 were monomorphic and 18 specific to 'Wangenheim', 12 specific to 'Pozegaca' and 37 to 'Stanley'. On studying the inheritance of these markers it was evident that 'Stanley' and 'Waigenhein' showed high Conformity with 'Cacaks Beauty', 'Cacaks Best' and 'Cacaks Early' contrary to the earlier report on their origin.

### 2.8.2.3.3 Cultivar identification

RAPD markers have been used to characterise cacao clones representing the three main cultivated subpopulations: Criollo, Foresterro and Trinitario. The use of thirteen different primers of arbitrary sequence was effective to analyse the extent and nature of the polymorphism.
Nei's similarity matrix was used to highlight the distinction between the wild cocoa species and cultivars (Wilde et al., 1992).

Twenty-five accessions of apple representing eight cultivars ('Golden Delicious Gala Jonagold Jopnathan Florina Fior di Casia and Imperatore Dallago') have been characterized with RAPDs. The reliability of the method was tested by analyzing scions of the same clone and also by comparing different accessions of the same cultivar. Using two separate ten base primers, it was possible to obtain a distinctive fingerprint for each of the cultivars. The method is simple, rapid and should provide a useful system for documenting the identity of clonal material (Mucahy et al., 1992).

Identification of apple cultivars was made by using RAPD markers by Koller et al. (1993). Apple cultivars were differentiated and a key was proposed. Cheng et al., (1996) identified co-dominant RAPD markers tightly linked to fruit skin colour in apple. A simple genetic basis for the red yellow skin-colour polymorphism in apple was verified using DNA markers. Bulked segregant analysis identified one 10-mer base oligomer that generated different fragments in each of the bulks. After testing the primer in four populations. Two fragments were associated with red skin colour and another two fragments associated with yellow skin colour. In the 'Rome Beauty' X 'White Angel' population two fragments were associated with red skin colour. Both parents displayed an alternate fragment associated with yellow-skinned fruit.

Identification of avocado cultivars with three RAPD primers homologous to regions of the chromosomal DNA was carried out (Lewis, 1992). Variety 'Fuerte' and 'Edranol' were closely related.
compared to 'Hass'. Similarly identification of rashberry cultivars by RAPD analysis was carried out by Parent et al. (1993). A fast and simple test to accurately characterize the various cultivars of rashberry of the Quebec certification programme was developed based on RAPD analysis using minute quantities of plant DNA. Rashberry cultivars and amplified with 120 different primers. A combination of three primers was chosen for its ability to differentiate all the cultivars.

Novy et al., (1994) identified varietal misclassification and regional divergence in cranberry (*Vaccinium macrocarpon*) using RAPD technology. Twenty-two decamer primers, amplified 162 scorable DNA fragments, of which 66 (41%) were polymorphic. On the basis of the 66 silver-stained RAPD (ssRAPDs), 17 unique profiles were identified rather than the expected 22. Fourteen varieties had unique ssRAPDs profiles while the remaining eight were represented by three ssRAPDs profiles. Permutational analyses of the data suggested that the observed ssRAPDs profile duplications were examples of varietal misclassification. Further analyses identified two ssRAPD markers that were found only in Eastern varieties (from Massachusetts and New Jersey) and not in Wisconsin varieties. The varieties differing on an average by 22 bands, ssRAPDs were shown to be useful in varietal identification and the assessment of genetic diversity in cranberry.

RAPD analysis of 17 olive (*Olea europaea* L) cultivars including oil and table types originating from throughout the Mediterranean area was carried out by Fabbbri et al., (1995). Their results indicated a high degree of polymorphism in the germplasm examined. Out of 40 random decamer primers screened 17 produced reproducible polymorphic amplification fragments. Each of the 17
produced cultivars was discriminated with a few primers. Results were analyzed for similarity among the cultivars and a cluster analysis was performed. The analysis revealed two main groups one comprising primarily small-fruited cultivars grown mainly for oil production and the other characterized by having large fruit. There was no apparent clustering of olive cultivars according to their geographic origins.

Another attempt on molecular characterization of olive varieties using RAPD markers was carried out by Wiesman et al., (1998). Primers were identified and used in combination to discriminate between different varieties. Significant biodiversity was demonstrated among "Nabali' olive trees, suggesting that the grouping known as "Nabali' is actually a mixture of genetically distinct variants. On the other hand, RAPD profiles of selected variants of 'Souri' revealed a high degree of similarity indicating that these variants represent environmental phenotypes of the same genome. Molecular differences were demonstrated between the 'Nabali group variants and 'Souri'. Other more recently developed or introduced varieties showed individually distinct RAPD profiles (Nicese et al., 1998). Using a unique system of six doubled haploid peach parents and their seven resulting F1 hybrid populations. They found that 18 per cent of the 50 scorable bands from 25 primers did not follow expected inheritance patterns. Bands present in parents were not transmitted to the progeny, or non-parental bands appeared in the progeny. Differences in RAPD patterns were also observed between haploids and spontaneous diploid bud sports on the same tree. Because peach is a long-lived clonally propagated crop, it is possible that somatic rearrangements occur which would lead to this unexpected inheritance of RAPD markers. Such anomalous patterns
indicate that RAPD markers should be used cautiously in Pyrus species mapping and genetic diversity studies (Monte-corvo, 2000).

Identification of cultivators and validation of genetic relationships in *Mangifera indica* L. was carried out using RAPD markers (Schnell et al., 1995). Twenty-five accessions of mango were examined for genetic markers with 11 primers which are reproducible. The number of bands generated were primer and genotype dependent, and ranged from one to ten. No primer gave unique banding patterns for each of the 25 accessions. However, ten different combinations of two primer-banding patterns produced unique fingerprints for each accession. A maternal half-sib (MHS) family was included among the 25 accessions to see if genetic relationships could be detected. RAPD data were used to generate simple matching coefficients, which were analyzed phonetically, and by means of principal coordinate analysis (PCA) while the randomly selected accessions were scattered with no apparent pattern.

Rong et al., (1995) studied the effectiveness of RAPD analysis for cultivar identification of persimmons (*Diaspyras kaki*) by using 10 base primers. Fifteen cultivars tested were completely distinguishable from each other by RAPDs using OPA-6 or OPA-8. Furthermore, two bud mutants of cv. 'Hiratenenashi', 'Tonewase' and 'Sugitawase', showed different DNA patterns with a few additional minor bands with OPA-6 primer. In addition, polymorphisms among 11 Diospyros species were observed by RAPD, using OPA-10 primer. The same method revealed little polymorphism among intraspecific levels of *D. kaki* or *D. lotus*. 

---

Assessment of Genetic diversity in *Santalum album* L. using Random Amplified Polymorphic DNA and development of specific primers
S.A. AZEEZ
Damasco et al., (1996) detected dwarf off-types in micro propagated Cavendish (Musa spp. AAA) bananas. A RAPD marker specific to the dwarf off-type from micro propagation of Cavendish group cultivars 'New Guinea Cavendish' and Williams' was identified following an analysis of 57 normal and 59 dwarf plants generated from several micro propagation events. Of the 66 random decamer primers used in the initial screening 28.8 per cent revealed polymorphisms between normal and dwarf plants. Primer OPJ-4 (5'-CCGAACACGG-3) amplified a 1.5 kb band which was consistently present in normal but absent in all dwarf plants of both cultivars. Reliable detection of dwarf plants was achieved using marker, providing a suitable means of in vitro detection. Other micro propagation induced RAPD polymorphisms were not associated with the dwarf trait (Williams et al., 1990).

Levi and Rowland (1997) identified blueberry cultivars and evaluated their genetic relationships using RAPD technology. Fifteen high bush (or high bush hybrid) blue berry cultivars (Vaccinium corymbosum), two rabbit eye blueberry cultivars (V. ashei) and one southern low bush (V. darrowi) selection from the wild was examined using seventeen 10-base RAPD primers in polymerase chain reactions. Fifteen RAPD markers resulting from these reactions were chosen to construct a DNA fingerprinting table to distinguish among the genotypes included in this study. Similarity values were calculated based on 132 RAPD bands. And a dendrogram was constructed based on the similarity matrix. The estimates of relative genetic similarity between genotypes within the V. corymbosum group did not agree well with known pedigree data and thus indicated that RAPD data did not accurately assess the genetic relationships of cultivars within this species (Rossetto, 1997).
Ye et al. (1998) reported the DNA fingerprinting utilizing RAPD polymorphisms to investigate the relationship among 16 grapevine cultivars and sports thought to have arisen from these cultivars. From 53 primers, a total of 464 bands were generated, of which 29 per cent were common to all genotypes tested. Cluster analysis classified all tested cultivars into two main groups (\textit{Vitis vinifera} 1, and \textit{Vitis x labrusea} Bailey) as expected (Pooler and Scorza, 1995).

Identification of cultivars and validation of genetic relationships in \textit{Mangifera indica} L. was carried out using RAPD markers (Schnell et al., 1995). Twenty-five accessions of mango were examined for genetic markers with 11 primers, which are reproducible. The number of bands generated were primer and genotype dependent, and ranged from one to ten. No primer gave unique banding patterns from each of the 25 accessions, however, ten different combinations of two primer-banding patterns produced unique fingerprints for each accession. A maternal half-sib (MIIS) family was included among the 25 accessions to see if genetic relationships could be detected, which were analyzed phonetically and by means of principal coordinate analysis (PCA), while the randomly selected accessions were scattered with no apparent pattern.

Genetic characterization and relatedness among California Almond cultivars and breeding lines were detected by RAPD analysis by Bartolossi et al. (1998). The genetic relatedness among 17 almond genotypes and one peach genotype was estimated using 37 RAPD markers. Genetic diversity within almond was found to be limited despite its need for obligate out crossing. Three groupings of cultivar
origins could be distinguished by RAPD analysis. A similarity index based on the proportion of shared fragments showed relatively high levels of similarity (0.75 or greater) within the almond germplasm. The level of similarity between almond and the peach was 0.424, supporting the value of peach germplasm to future almond genetic improvement.

Neto. (1995) used RAPD markers for identification of cashew seedlings. RAPD profiles of four seedlings were obtained using six 10-mer arbitrary primers and they were compared with each other to clone specific amplified products. The size of DNA fragments ranged from 240 to 1780 base pairs. The number of bands on the profiles varied from 0 to 4, depending on the primers and seedlings tested. A total of 27 amplification products suitable for DNA fingerprinting were disclosed and two primers were selected for further identification of genome in cashew.

Genetic characterization of gooseberry germplasm was carried out by Lanham and Brennan (1999) using RAPD, ISSR and AFLP markers. AFLP generated unique profiles for each genotype. ISSR could not distinguish three of the genotypes. The results imply that the European cultivated gooseberry has a narrow genetic base and that further improvement to this crop species may require the introduction of additional germplasm into breeding programmes.
2.9 Polymerase Chain Reaction

The DNA amplification by thermal cycling called polymerase Chain Reaction (PCR) is an in vitro method that can be used to amplify a specific DNA segment from small DNA template/duplex into millions of copies. It is invented by Kary Mullis et al. (1986) and described by Saiki et al (1985). It is carried out in three steps at discrete temperatures heat denaturation (94-98°C), annealing of synthetic primers to template (35-55°C) followed by strand extension (72°C) from 3'hydroxyl end which is directed towards each other resulting in the exponential accumulation of discrete fragment whose termini are defined by 5'ends of primers. The exponential amplification is because of the products synthesized in cycle “n” function as a template for the other primer in cycle “n+1” cycle. The length of the product generated during the PCR is equal to the sum of the length of the two primers plus the distance in the target sequence (Erlich et al., 1991).

2.9.1 Steps involved in PCR

2.9.1.1 Heat denaturation

Innis and Gelfand (1990) reported that in the step of denaturation the double stranded DNA get separated into single strands. Typically denaturation condition is 95°C for 30 seconds or 97°C for 15 seconds. However higher temperatures may be appropriate, especially for G+C rich templates. It only takes few seconds to denature DNA at its strand separation temperature (Tss). Low temperature is the most likely cause for the PCR product, which allows the DNA strands to “snap back” and thus reduces the product yield. In contrast, denaturation step that has too high temperature and or too long lead to unnecessary loss of enzyme.
activity. It is better to perform PCR with temperature range of 94-97°C for denaturation and strand separation to avoid mispriming, misincorporation of incorrect nucleotides.

### 2.9.1.2 Annealing

Annealing is hybridization of primer to single stranded DNA. The temperature and the length of time required for primer annealing depend on the basic composition, length and concentration of the amplification primers. An applicable annealing temperature ($T_m$) is 5°C below the $T_m$ of the amplification primers. At typical primer concentration (0.2mM), annealing will require only few seconds (Innis et al., 1988).

Kim and Smithies (1988) reported that increasing the annealing temperature especially during first several cycles, will increase specificity of priming and reduces incorrect priming as well as misextension of incorrect nucleotides at 3' end of primers. For maximum specificity in the initial cycle *Taq* DNA polymerase should be added during primer annealing.

### 2.9.1.3 Primer extension

Primer extensions are traditionally performed at 72°C because this temperature was appropriate for extending primers on M-13 based model template. Low extension temperature favors misextension of primers and extension of misincorporated nucleotides. So it is better to perform PCR at high temperature for extension (Kims and Smithies, 1988).
Innis et al., (1988) noticed that extension time depends upon the length and concentration of target sequence and upon temperature. A reaction cycled between denaturation and the primer binding temperature generally allows sufficient time for extension of short PCR nucleotide products with polymerase.

An extension time of one minute at 72°C is considered sufficient for products up to 2 kb in length. However, longer extension times may be helpful in early cycles of the substrate concentration is very low. Estimates for the rate of nucleotide incorporation at 72°C vary from 35-100 nucleotides per second, depending upon the buffer, pH, salt concentration and nature of the DNA template (Saiki and Gelfand 1989).

2.9.1.4 Number of cycles

The optimum number of cycles depends mainly upon the starting concentration of target DNA, when other parameters are optimized. A common mistake is to execute too many cycles. Too many cycles can increase the amount and complexity of nonspecific background products. In contrast too few cycles give a low product yield. Innis and Gelfand (1990) suggest that when the number of target molecules are 3x10^5, 1.5x10^4, 1x10^3 and 5 number of cycles to be execute are 25-30, 30-35, 35-40 and 40-45 respectively (Sambrook et al., 1989).

2.9.2 Components of PCR

2.9.2.1 Primers

Innis and Gelfand (1990) reported that primer concentration between 0.1 to 0.5 μM is generally optimal. Higher concentration of primers may promote mispriming and accumulation of non-specific products and may increase the probability of generating a template independent
artifact termed primer-dimers. Non specific products and primer-dimer artifacts are themselves substances for PCR and complete with the desired product for enzyme, dNTPs and primers resulting in lower yield of the desired product.

For construction of random primers used in RAPD there is no need of any prior knowledge of genome. Arbitrary primers will be usually 9-10 base pair long with a GC content of 50-80 percent and do not have palindromic sequences. The number of DNA fragments that are amplified depends on the period and genomic DNA used (Williams et al., 1990).

Kang et al., (1998) reported that GC content of the primers affected their sensitivity to the length of time in programmed that was allowed for annealing. RAPD patterns obtained from primers with high GC content (70-80%) were affected by short annealing time, where as those obtained from primers with 50-60% GC content were reduced in intensity even with 30 seconds of annealing time.

2.9.2.2 Magnesium concentration

It is beneficial to optimize the magnesium ion concentration. The magnesium may affect primer annealing, strand dissociation temperatures of both template and PCR product, product specificity, formation of primer-dimer artifacts, enzyme activity and fidelity. Taq DNA polymerase requires free magnesium on top of that bound by template DNA, primers and dNTPs. Accordingly, PCR should contain 0.5 to 2.5 mM MgCl₂ over the total dNTP concentration. The presence of EDTA or other chelators in primer stock or template DNA may disturb this apparent magnesium optimum (Innis and Gelfand, 1990).
Baumforth et al., (1999) reported that storage of purified DNA in phosphate buffer or buffer containing high concentration of EDTA should be avoided because of their effect on magnesium ion concentration.

2.9.2.3 DNTPs

The stability of dNTPs during repeated cycles of PCR is such that approximately 50 per cent remains as dNTP after 50 cycles. Deoxynucleotide concentration between 20-200μM each results in the optimal balance among yield, specificity, and fidelity. The four dNTPs should be used at equivalent concentrations to minimize misincorporation errors. Low dNTPs concentration minimizes mispriming at non-target sites and reduces the likelihood of extending misincorporated nucleotides (Innis et al., 1988). One should decide on the lowest dNTPs concentration appropriate for the length and composition of target sequences, eg., 20μM each dNTPs in 100μl reaction is theoretically sufficient to synthesize 2.6μg of DNA or 10 pmol of a 400bp sequence.

Conditions that promote misincorporation include when deoxynucleotide concentrations are below the K_m(<1μM) or when the concentration of one nucleotide or dNTPs is relatively lower than other three (Innis et al., 1988).
2.9.2.4 DNA polymerases

The availability of thermostable Taq DNA polymerase enzyme has not only simplified the procedure for the PCR but also increased the specificity and yield of the amplification. Taq DNA polymerases have non 3'-5' exonuclease activity during polymerization but have 5'-3' exonuclease activity which helps in eliminating misincorporation and non-specific bands (Saiki et al., 2000).

Klenow fragment of Escherichia coli DNA polymerase is unstable at high temperatures. As a result there is a requirement for the addition of fresh enzyme after the denaturation step in each cycle. Where as the Taq DNA polymerase from Thermus aquaticus is thermostable even at very high temperatures. The incorporation of Taq DNA polymerase into PCR protocol that allows the primers to anneal and extend at much higher temperature which was impossible with Klenow fragment, eliminating the non-specific amplifications. Moreover long PCR products could be amplified from genomic DNA, probably due to a reduction in the secondary structure of template at elevated temperature used for primer extension. The upper size limits for amplification with Klenow fragment was only 400 bp. Where as fragments as large as 10 kb have been synthesized with Taq DNA polymerase and other thermostable enzymes. In recent new DNA polymerases isolated from the thermophilic bacteria (Thermococcus litoralis) called vent DNA polymerase has 3' – 5' exonuclease activity and may therefore have lower misincorporation rate. However, the capacity of this to degrade single stranded molecules (like oligonucleotide primers or PCR product prior to primer annealing) will pose problem for PCR amplification (Erlich et al., 1991).
Innis and Gelfand (1990) reported that the concentration range for Taq DNA polymerase is between 1-2.5 units per 100μl reaction when other parameters are optimum. However, enzyme requirement will be depends on individual target templates or primer. When optimizing a PCR they recommend testing enzyme concentration ranging from 0.5 to 5 units per 100μl and assaying the results by gel electrophoresis. If enzyme concentration is too high, nonspecific background products may occur and if too low an insufficient amount of desired product is made. Taq polymerase from different suppliers may behave differently because of different formulations, assay conditions and or unit definitions.

2.9.2.5 Buffers

A number of different buffers for PCR are currently used. Taq appears to have an optimum pH of 7.0 to 7.5 at 72°C. Normally a Tris buffer with pH 8.5 to 9.0 at 25°C is used because the pH of Tris buffer decreased by 0.03 pH units for each degree increase in temperature. Thus a buffer that is made to pH 8.8 at 25°C has only pH 7.3 at 72°C.

2.9.2.6 Detergents

Taq is highly hydrophobic protein and tends to precipitate form aqueous solution. The addition of non-ionic detergents (Triton X-100, NP-40, or Tween 20) at a final concentration of 0.01%) helps to maintain full activity, both in storage solution and in amplification reaction. Highly purified (low peroxide) preparations of detergents should be used (Baumforth et al., 1999).
2.9.2.7 Problems associated with PCR

Although high processivity of Taq makes production of long PCR products conceivable and some ten kb products have been detected by southern blotting. PCR is most useful for the amplification of DNA segments less than 2 kb in length, longer products are increasingly likely to be out competed in the reaction by products from mis-primed sites either internal to the amplified product or from elsewhere in the genome. Alternative priming makes it more difficult to amplify long products from complex genome (Baumforth et al., 1999).

With the repeated cycles of denaturation and annealing used in PCR the possibility exists for products only partially extended during one cycle to reanneal to a different template in a later cycle. If two alleles or repeated genes are present, it is possible to create “shuffle clones” which are recombinant products of the two sequences. These “shuffle clones” are frequent when PCR is performed with Klenow polymerase but may be rare when highly processive Taq is used (Mc Bride., 1989).

“Shuffle clone” may be more frequent in amplifications from highly damaged DNA. If the template DNA consists of fragments shorter than the desired amplification product, the fragments from different alleles may be assembled randomly by PCR (Paabo et al., 1990).

Because the PCR can generate trillions of DNA copies from a template sequence, contamination of the amplification reaction with products of previous PCR reaction (product carry over) exogenous DNA other cellular material can create problems both in research and diagnostic application. This can be avoided by using multiple controls (No template DNA added to the reaction) are necessary to monitor...
purity of reagents and to several contamination in genetic typing the contamination of a sample reaction can often detected by genotyping result with more than two alleles. Attention to lab procedures pre aliquoting reagents, the use of dedicated pipettes, positive displacement pipettes or tips with barriers preventing contamination of pipette barrel and the physical separation of the reaction preparation from the area of reaction product analysis minimizes the risk of contamination (Erlich et al., 1991).

2.9.2.7.1 DNA Fingerprinting

RAPDs are useful for DNA fingerprinting where there is a need to identify varieties of a crop or to determine parentage within breeding material. Wilde et al. (1992) reported that RAPD analysis has resulted in reliable discrimination both between and within groups of cocoa genotypes. Julie et al., (1994) reported the use of PCR for fingerprinting red raspberry cultivars. Nine out of the ten random primers used, successfully generated specific fingerprints for all the raspberry cultivars that could be used for cultivar identification.

RAPD fingerprinting of four dwarfed clones of cashew (Anacardium occidentale L.) was reported by Neto., (1995). Twenty-seven primers generated amplified products in the range of 240-1780 bp, few of which were found suitable to distinguish the four seedlings. The four dwarfed clones; CP 06, CP 09, CP 76 and CP 1001 could be most effectively differentiated from each other by using OPA 8, OPA 16 and OPB 15. The amplified products of six selected primers were used to study the similarity among the four clones. CP 06 and CP 76 were found to be most similar (77.8 %) followed by CP 06 and CP
1001, which were 66.7 per cent similar, and least similarity was reported between CP 76 and CP 1001 (59.3%). Similarly, Karihaloo et al., (2000) reported fingerprinting of 19 cashew accessions using decamer primer S-11.

In another investigation, Ye et al. (1998) reported DNA fingerprinting utilizing RAPD polymorphisms to determine the relationship among 16 grapevine cultivars and sports thought to have arisen from these cultivars. From 53 primers, a total of 464 bands were generated, of which 29 per cent were common to all the genotypes tested. Cluster analysis classified all tested cultivars into two main groups (Vitis vinifera L. and Vitis labrusca Bailey) as expected. No polymorphism was detected among known clones of 'Chardonnay' or 'Pinot Noir'.

Tyson et al., (1998) employed RAPD markers to determine if a hybrid copse between Eucalyptus risdonii and E. amygdalina was clonal. Based on lignotuber morphology this copse was composed of approximately 20 separate individuals. No variation in RAPD was observed for 67 bands scored from nine primers among the 20 individuals. In contrast, variation was observed between individuals sampled from outside the copse (average 28.6 bands differences between any two individuals) and within full sibling families (average 12.9 band differences between individuals).
2.9.3 Gel electrophoresis

Naturally all products in molecular biology require at some point fractionation of nucleic acids. Although, chromatography techniques are appropriate for some approaches, gel electrophoresis has much greater resolution than alternative methods and is generally the fractionation method of choice. Gel electrophoresis separation can be either analytical or preparative and can involve fragments with molecular weights ranging from <1000 daltons to >10^8 daltons. A variety of electrophoretic systems have developed to accommodate such a large range of applications.

The use of electrophoresis to separate nucleic acids is simpler than its application to resolve proteins. Nucleic acids are uniformly negatively charged and for double stranded DNA there is no complicating structural effects that affects its mobility. A variety of important variables affect migration of nucleic acids on gels. These include the confirmation of the nucleic acid, the pore size of the gel, the voltage gradient applied and the salt concentration of the buffer. The most basic of these variables is the pore size of the gel that deflects the size of the fragments that can be resolved (Ausubel et al., 1989).

In practice two types of gel electrophoresis are used to resolve fragments, namely agarose and polyacrylamide gel electrophoresis. Agarose electrophoresis is relatively simpler to perform and separate DNA fragments from 0.5 kb to 25 kb in size. However, polyacrylamide gel electrophoresis is the method of choice to separate DNA fragments less than 1kb (Kang et al., 1998). In agarose gels there are two types, larger pore agarose which are used to resolve fragments larger than
1000 bp and smaller pore agarose gels are used to resolve smaller fragments less than 1000 bp (Ausubel et al., 1989).

Agarose gels are cast by melting agarose in the desired volume of buffer by gentle boiling until the solution becomes clear. The gel is then cooled to about 55°C, the ethidium bromide is added at the rate of 0.4 mg/ml for better visualization. The gel is poured into a mould and is allowed to harden, upon hardening the agarose forms a matrix, then density of which is determined by the concentration of the agarose. When an electric field is applied across the gel, the DNA, which is negatively charged at neutral pH, migrates towards the anode. The rate of migration is determined by a number of parameters viz., molecular size of the DNA, applied voltage, direction of the electric field, base composition and temperature, presence of intercalating dyes, composition of the running buffer and apertures of the gel. Appropriate agarose concentrations are required for separating DNA fragments of various sizes. The concentrations commonly used are 0.5, 0.7, 1.0, 1.2 and 1.5 per cent to separate fragments of 30 to 1 kbp, 12 to 0.8, 10-0.5, 7-0.4 and 3-0.2 kbp in size (Sambrook et al., 1989).

2.9.4 DNA Sequencing

Wolfe et al. (1999) examined four populations of the rare, highly clonal grass *Calamagrostis porteri* ssp. *insperata* using allozymes and the two polymerase chain reaction (PCR)-based markers, random amplified polymorphic DNA (RAPD) and intersimple sequence repeat (ISSR) bands. Only one of the 15 allozyme loci was variable and two alleles were detected, both of which were found in two populations, while only one genotype was detected in the other two populations.
ISSR and RAPD markers detected more genotypes within populations than did allozymes. ISSR markers detected more diversity than RAPD markers in three of the four populations examined. In one population, no RAPD diversity was found whereas eight different genotypes were found among the 10 plants with ISSR markers. This diversity is present despite rare flowering, no documented occurrence of seed set in natural populations and very low seed set with experimental pollinations, all of which suggest that sexual reproduction rarely occurs. The subspecies is self-compatible, but seed initiation is lower in selfed ovules; also, there is high embryo abortion regardless of pollen source. Variation detected by RAPD and ISSR primers may reflect higher levels of sexual reproduction in the past, very rare sexual reproduction in extant populations, somatic mutations, or a combination of the three. Although the PCR-based markers identify several multilocus genotypes within populations (Milbourne et al., 1997). It is not known whether these all represent distinct genets generated by sexual reproduction or result from somatic mutations in the old, perennial and highly clonal plants (Anil et al., 2000).

Sequence characterized amplified regions (SCARs) were derived from eight random amplified polymorphic DNA (RAPD) markers linked to disease resistance genes in lettuce by Paran et al., (1992). SCARs are PCR-based markers that represent single, genetically defined loci that are identified by PCR amplification of genomic DNA with pairs of specific oligonucleotide primers; they may contain high-copy, dispersed genomic sequences within the amplified region. Amplified RAPD products were cloned and sequenced. The sequence was used to design 24-mer oligonucleotide primers for each end. All pairs of SCAR primers resulted in the amplification of single major
bands the same size as the RAPD fragment cloned. Polymorphism was either retained as the presence or absence of amplification of the band or appeared as length polymorphisms that converted dominant RAPD loci into codominant SCAR markers. This study provided information on the molecular basis of RAPD markers. The amplified fragment contained no obvious repeated sequences beyond the primer sequence. Five out of eight pairs of SCAR primers amplified an alternate allele from both parents of the mapping population; therefore, the original RAPD polymorphism was likely due to mismatch at the primer sites.

Dean et al. (2001) developed primer for objectively determining the genotype of cucurbit plants, particularly species of melon, with respect to resistance or susceptibility to Fusarium wilt infection. They used a polymerase chain reaction to amplify sample DNA using either an AM or FM oligonucleotide primer pair. The PCR product which results from either primer pair differs in size, depending upon whether the template DNA was obtained from a plant susceptible or resistant to Fusarium wilt, permitting easy and rapid identification of plant genotype.

Robin Beat Gasser et al. (2002) developed a PCR-based method for the identification of species of the genus Eimeria, (commonly known as coccidia), is described. The method is genus-specific and utilized either, or both, of two novel primer sets; designated WW1 (SEQ ID NO:31) and WW3r (SEQ ID NO:32), and, WW2 (SEQ ID NO:33) and WW4r (SEQ ID NO:34).

Assessment of Genetic diversity in Santalum album L. using Random Amplified Polymorphic DNA and development of specific primers
S.A. AZEIZ