Genome analysis of neem

Chapter - 1

Preface
Neem (Azadirachta indica), is one of the trees known in the Indian sub-continent since antiquity. The ancient Indian found many therapeutic uses for the tree and also observed that the tree could survive in very dry and arid conditions. With the isolation and characterization of azadirachtin and a number of other chemicals from seeds, the neem tree is receiving global attention. The anti feedant and repellent properties of azadirachtin have now been firmly established world over and it has become an important chemical in integrated pest management. With the growing neem-based industry, the demand for quality neem is expected to rise sharply.

Neem tree is commonly found in South Asia and parts of Africa. It is a fast growing tree. As a rule it is a hardy evergreen tree, but under extreme conditions, such as extended dry periods, it may shed most or nearly all of its leaves. Neem leaves are imparipinnately compound, alternate and crowned at the end of the branches. The young leaves are reddish-green, gradually turning into lush green or dark green. The branches are widely spread and the trunk is relatively short and may reach a girth of 1.5-3.5m (Schmutterer, 1995). According to a recent survey (Randhawa and Parmar, 1993), there were about 14 million neem trees in India, of which more than half were in Uttar Pradesh state alone and the rest in the states of Tamil Nadu, Karnataka, Andhra Pradesh, Maharastra and Gujarat.

Even though neem is indigenous and has many reputable properties and is a multi-purpose tree, it has generally been,
neglected by plant scientists in India. No serious attempts seem to have been made for genetic analysis. The wide distribution of neem tree is indicative of genetic variability that could be exploited for selection and improvement of neem. The pattern of inter- as well as intra-specific variation needs to be studied and evaluated for superior germplasm selection for improvement. Most published research has dealt with the isolation, identification, formulation and field testing of pest control agents derived from the seed kernels and leaves. Emphasis needs to be placed on both the species diversity and intraspecific variability prevalent in the natural range of *Azadirachta indica*, which is widely distributed, and in case of *Azadirachta siamensis*, which is mostly confined to Indo-China and Thailand. A number of studies have been reported on morphological and phenological variation in neem (Gogate and Gujar, 1993), however, there is virtually no data about genetic basis for this variation. As for the DNA analysis not even a single report was published to our knowledge when this work was initiated in 1994.

Genetic studies of plants, be they as food plants or as model systems, have until recent years depended on the phenotypic trait and all factors governing it have an important role to play in determination of the genotype. This has several times in past led to situations where either the genotype is wrongly identified or to situations where the genotype could not be identified. With the advancements in DNA-based technologies, the concept of "molecular markers" came into being and within a short period has dominated all genetic markers. Three major DNA-based technologies have been most frequently employed in carrying out genetic analysis in plants (Ranade and Sane, 1995). These include (i) Restriction Fragment Length Polymorphism (RFLP) analysis, (ii) DNA Fingerprint Profiling (DFP) analysis and (iii)
Multiple Arbitrary Amplicon Profiling (MAAP). RFLP is the most extensively employed molecular technique where sufficiently contrasting genotypes are present. RFLPs have been very elegantly employed in determining molecular linkage maps, in identifying genes linked to characters of agronomic importance and in mapping genes of QTLs or Quantitative trait loci. The RFLPs are generally not very informative in identification of genotypes within random mating or non-pedigreed populations, since the RFLP probes are invariably specific to a single or few genetic loci. Hence the RFLPs are considered to be of limited application in population genetics. The DFP technology is more powerful than the RFLP since, an individual genotype can be identified at the molecular level on the basis of an extremely high level of polymorphism in the sequence of its DNA (Ranade and Sane 1995).

The exploitation of DNA polymorphism by an ever-increasing number of molecular marker technologies has begun to have an impact on plant genome research. A limiting feature of nucleic acid typing techniques is the requirement of prior knowledge of nucleotide sequence or availability of cloned and characterized hybridization probes. The ability to produce characteristic signature from virtually any nuclei acid without these constraints has been made possible by techniques where nucleic acids are amplified using arbitrary primers. These primers target multiple anonymous sites (amplicons) in nucleic acid templates producing signatures (fingerprints) composed of arbitrary collections of amplification products (Williams et al., 1990, Welsh and McClelland, 1990, Caetano-Anolles et al., 1991). The targeted sites are amplified with one or more short oligonucleotide primers either completely arbitrary in nature or derived from known terminal, unique, repetitive or dispersed sequences. These primers can and do hybridize to, perfectly or
even partially complementary sites at several loci. The overall strategy thus differs considerably from the typical polymerase chain reaction (Mullis and Faloona, 1987) in which there is an absolute requirement of sequence specificity and primer complementarity to defined template regions. In fact this requirement excludes the use of typical PCR in cases where nucleic acid sequence information is unavailable, as for example in case of neem.

MAAP is a versatile and universal approach for the analysis of nucleic acids as demonstrated by many applications and a wide range of organisms studied (Caetano-Anolles 1994; Williams et al., 1993, Tingey and delTufo, 1993). Its use has been reported in about 300 plant species evenly spread throughout the plant kingdom (Wiesing et al., 1994), including many of commercial importance. Its universal nature has made this molecular tool invaluable for the survey of diversity and in the study of populations. Furthermore, its simplicity and speed has extended the analysis to those applications where high throughput is needed.

DNA-based markers are in general, a good choice for studies on genetic variation because,
i) these enable, a more accurate identification of the genotype in a given population,
ii) the detection and analysis of such markers is unaffected by environment,
iii) they enable a better assessment of genetic diversity by scoring the occurrence of a larger number of alleles than that is possible with the phenotypic markers,
iv) rapid identification of the "elite" or "desired" genotype amongst the test population is also possible and
v) the number of possible DNA markers is large thereby enabling easy saturation of the linkage maps which in turn facilitates gene-tagging and marker-assisted gene isolation.

In the present studies, DNA-based markers were selected for genome analysis of neem. The polymorphism resulting from the accumulation of various mutations and changes in the genome is the rationale behind the various techniques employed in the present studies. The changes include sequence insertions, deletions, transposition, point mutations, errors in DNA replication, duplications, meiotic and other recombinations and gene conversions. These changes are detected as polymorphism of DNA fragments revealed both by restriction endonuclease digestion and DNA probing as well as by DNA amplification.

Of the various molecular methods discussed earlier, in the present studies we have extensively depended upon the PCR-based methods, primarily because of the rapidity and facility with which these methods lead to easily interpretable data. The PCR-based method used include RAPD, MP-PCR and DAMD and have been discussed in detail in the respective chapters.

1.2 OBJECTIVES

The present work as embodied in this thesis was initiated in 1994 with the broad objectives of, i) characterization of neem DNA, ii) identification and isolation of sequence specific families and iii) determination of the nature and extent of DNA variability.

Within the framework of these broad objectives, the specific objectives for the present thesis were defined as follows:
1. Optimization of RAPD-PCR
2. Determination and analysis of genetic diversity in neem using RAPD technique
3. PCR based fingerprinting using minisatellite and microsatellite sequences as primers
4. Determination of inter- as well as intra-geographical group relationship amongst neem provenances.

1.3 SCOPE OF THE THESIS

The present thesis entitled "Molecular analysis of the genome of neem (Azadirachta indica)" embodies results of our investigations on genetic relationship amongst various provenances of neem. The entire thesis is divided into 7 chapters, as described below:

Chapter 1 The present chapter, entitled "Preface" includes the introduction to the thesis, objectives and scope of the thesis.

Chapter 2 This chapter is entitled "PCR-based determination and analysis of genetic variability in trees" is an attempt to review and summarize bulk of the published data on the above theme.

Chapter 3 "Standardization of DNA isolation and RAPD-PCR conditions" describes our experiments and results thereof towards standardization of experimental protocols for neem DNA isolation and its use as a template in RAPD-PCR.
Chapter 4 "Analysis of genetic diversity in neem using RAPD technique" describes the determination of overall genetic diversity in two populations of neem.

Chapter 5 "PCR based DFP using micro- and mini-satellite sequences" deals with use of short tandem repeat and minisatellite core sequences as primers in PCR.

Chapter 6 "Inter- as well as intra-geographical group relationship amongst neem provenances" describes experiments and results thereof for the determination of genetic variability in small subsets of neem provenances.

Chapter 7 The overall achievements in the present thesis are discussed and the future prospects for this work have been outlined, in this chapter "General Discussion".

References, cited in the text in each of the chapters 1 through 7, are listed at the end of the respective chapters.

1.4 REFERENCES


