DEVELOPMENT AND CHARACTERIZATION OF GENOMIC MICROSATELLITE MARKERS IN *Melia azedarach*

Dissertation submitted to the Central University of Punjab

For the award of
Master of Philosophy
In
Biosciences

BY

Sapna Thakur

Supervisor
Dr. Pankaj Bhardwaj

Centre for Biosciences
School of Basic and Applied Sciences
Central University of Punjab, Bathinda

Sep, 2014

CERTIFICATE

I declare that the dissertation entitled “DEVELOPMENT AND CHARACTERIZATION OF GENOMIC MICROSATELLITE MARKERS IN *Melia azedarach*” has been prepared by me under the guidance of Dr. Pankaj Bhardwaj, Assistant Professor, Centre for Biosciences, School of Basic and Applied Sciences, Central University of Punjab. No part of this dissertation has formed the basis for the award of any degree or fellowship previously.

Sapna Thakur

Centre for Biosciences,
School of Basic and Applied Sciences,
Central University of Punjab,
Bathinda-151001

Date:

CERTIFICATE

I certify that Sapna Thakur has prepared her dissertation entitled “DEVELOPMENT AND CHARACTERIZATION OF GENOMIC MICROSATELLITE MARKERS IN *Melia azedarach*”, for the award of M.Phil. degree of the Central University of Punjab, under my guidance. She has carried out this work at the Centre for Biosciences, School of Basic and Applied Sciences, Central University of Punjab.

Dr. Pankaj Bhardwaj

Centre for Biosciences,
School of Basic and Applied Sciences,
Central University of Punjab,
Bathinda – 151001

Date:

ABSTRACT

Development and characterization of genomic microsatellite markers in *Melia azedarach*

Name of student: Sapna Thakur

Registration number: CUPB/MPh-PhD/SBAS/BIO/2012-13/01

Degree for which submitted: Master of Philosophy

Name of supervisor : Dr. Pankaj Bhardwaj

Name of centre: Centre for Biosciences

Name of school: School of Basic and Applied Sciences

Keywords: Indian Thar Desert, Genetic diversity, Population structure, Microsatellite marker, *Melia azedarach*.

Melia azedarach is ecologically imperative species known for its innumerable biological benefits such as antiviral, anthelminthic, antibacterial, etc. In this study, we developed 43 genomic microsatellite markers from (AG)n enriched library and subsequently employed 23 of them for genetic diversity and population structure analysis of *Melia azedarach* growing in Indian Thar desert. Fourteen populations encompassing 95 genotypes were selected for analysis and we found a moderate level of diversity (Na = 3.211, Ho = 0.558, He = 0.549, P = 94.41%) in them. Gene diversity (h) among population pairs varied from 0.566 to 0.714 with very low overall genetic differentiation (F = 0.021). The highest value of ΔK estimated using STRUCTURE indicated 2 subpopulations (K=2) and admixed cluster occupied maximum area (75.79%) under Bar plot. Genetic distance based UPGMA dendrogram also identified 2 major clusters among 14 *Melia azedarach* populations. UNJ tree based on genetic dissimilarity clustered genotypes from different population together. No significant correlation between geographical and genetic distance was found in present study (Rxy = 0.261, P = 0.18). Allele frequency distribution under "mode-shift" indicator was normal L-shaped, suggesting populations under study are not experiencing any recent bottleneck. This study laid the foundation for more precise inference about the biogeography and management of *M. azedarach* in the Indian Thar Desert.

Signature of Student Signature of Supervisor

ACKNOWLEDGEMENTS

The first and foremost thanks are to the “God Almighty” who bestowed me with his blessings, endowed me with strengths and fortitude.

I express the deepest gratitude to my supervisor, **Dr. Pankaj Bhardwaj**, Assistant Professor, Centre for Biosciences, Central University of Punjab for his inspiring guidance, constant encouragement, solid suggestions, through supervision and constructive criticism throughout the course of investigation and during the preparation of this manuscript.

I am grateful to **Prof. R. G. Saini**, Centre Coordinator, Biosciences, for his generous advice. My appreciations to **Prof. P. Ramarao**, Acting Vice- Chancellor and **Prof. Jai Rup Singh**, former Vice chancellor, Central University of Punjab, for providing infrastructure and other technical facilities. The cooperation of all the faculty members of Centre for Biosciences is highly acknowledged.

My special thanks to seniors and lab mates especially, Shruti Choudhary, Richa Mehra, Gaganpreet Kaur and Harmeen Kaur. I am also thankful to technical staff for their help and support.

Warmest regards to my family for their blessings.

I thank Indian Council of Medical Research (ICMR), New Delhi for awarding me fellowship.

Sapna Thakur

TABLE OF CONTENTS

S. No.	Contents	Page No.
1.	Introduction	1-3
2.	Review of Literature	4-9
3.	Material and Methods	10-13
4.	Results	14-24
5.	Discussion	25-27
6.	Summary	28-29
7.	References	30-36
8.	Appendix	37-43

LIST OF TABLES

Table No.	Table Description	Page No.
1.	List of studies validating various usage of *M. azedarach*	6
2.	List of various studies involving Microsatellite marker development and characterization	9
3.	Geographical locations of *M. azedarach* populations from Indian Thar Desert	11
4.	Features and evaluation details of 23 microsatellite markers	16-17
5.	Summary of F-statistics	18
6.	Summary of Analysis of Molecular Variance (AMOVA) for 14 *M.azedarach* populations	18
7.	Genetic diversity of populations	19
8.	Allele frequency based mutation drift equilibrium	20

LIST OF FIGURES

Fig. No.	Description of Figures	Page No.
1.	General scheme for microsatellite marker development	8
2.	Sites of sample collection (solid circles) in Indian Thar Desert	12
3.	Lane 1-50 amplification profile generated using three-primer PCR; Ld: 100bp DNA ladder as size standard (Bangalore Genei^{TM}); lanes showing multiple bands (indicated by arrow) represent positive clones	15
4.	Amplification profile generated with primer MAZ-41. Lanes 1 - 95 represent sampled individuals of *M. azedarach*; Ld: 100 bp DNA ladder as size standard (Bangalore GeneiTM)	15
5.	Graphic representation of proportion of alleles and their distribution in populations of *M.azedarach*	21
6.	Unrooted Neighbor Joining (UNJ) tree constructed from darwin v5.0.158; each branch represents single individual collected from corresponding inferred population	21
7.	Dendrogram based on Nei's genetic distance for 14 populations of *M.azedarach* based on 23 SSR loci	22
8.	Estimation of optimum K using the second order statistics (ΔK) given by Evanno *et al.*, (2005)	23
9.	Bar plot given by Structure with membership coefficient (Q) of individuals on y- axis. Two clusters inferred by Structure are combined from clustering of all individuals. Each color represents different genetic stock and each vertical line represents individual	23
10.	Relationship between genetic and geographic distances for 14 *M.azedarach* populations. Genetic distance is represented by pairwise F_{st} and geographic distance is given in kilometers.	24

LIST OF APPENDICES

Appendix Serial	Description of Appendices	Page No.
A.	Composition of chemicals	38-40
B.	DNA isolation and purification protocol	41-42
C.	Pairwise population F_{st} values	43

LIST OF ABBREVIATIONS

S. No.	Full Form	Abbreviation
1.	Adenine	A
2.	Base pair	bp
3.	Cetyl trimethylammonium bromide	CTAB
4.	Cytosine	C
5.	Degree Celsius	°C
6.	Deoxyribonucleic acid	DNA
7.	Deoxyribonucleotide triphosphates	dNTPs
8.	Gram	g
9.	Guanine	G
10.	Hardy- Weinberg Equilibrium	HWE
11.	Hour(s)	hr(s)
12.	Magnesium Chloride	$MgCl_2$
13.	Melting temperature	T_m
14.	Microgram	µg
15.	Microliter	µl
16.	Milligram	mg
17.	Milliliter	ml
18.	Millimolar	mM
19.	Minute	min
20.	Nanogram	ng
21.	Polyacrylamide gel electrophoresis	PAGE
22.	Polymerase Chain Reaction	PCR
23.	Polymorphic Information Content	PIC
24.	Potassium Chloride	KCl
25.	Randomly Amplified Polymorphic DNA	RAPD
26.	Restriction Fragment Length Polymorphism	RFLP
27.	Second	sec
28.	Simple Seqeunce Repeats	SSR
29.	Sodium- Saline Citrate	SSC
30.	SSR Identification Tool	SSRIT
31.	Thymine	T
32.	Unrooted Neighbor Joining	UNJ
33.	Unweighted Pair Group Method with Arithmetic Mean	UPGMA

CHAPTER 1

Introduction

Melia azedarach (Greek melia, the Ash and Persian azzadirackt, noble tree) is a deciduous tree of meliaceae family. Its vernacular names are bakain, mahanimba, pride of India, Persian lilac, bead tree, China berry, China tree and ghora nim etc. It is native to southeastern Asia, especially found in India and China but now grown in all other warmer parts of the world. Morphology of its different cultivars varies throughout the world but typically, it attains 10-15m height with about 60 cm diameter though wild trees of Asia can reach the height of about 40m. Its compound leaves are dark green in colour and resemble Ash tree. Its fruit is called drupe and is poisonous to man and pig (Hare *et al.*, 1998; Waggy, 2013).

It is planted as shade tree and commonly found around human habitation. Under optimal conditions it grows fast so its plantation is suitable for timber and ornamental purposes. It was also planted as locust barrier (Tourn *et al.*, 1999). Other uses of *M. azedarach* include leaves for fodder, timber “the white cedar” for furniture, oil for illumination and making necklaces, rosaries etc (Sen and Batra, 2011). Meliartenin, a limonoid obtained from the tree is known to have anti-feedant and insecticide properties which are comparable to azadirachtin (Carpinella *et al.*, 2003). Its seed extract is a potential rodenticide because of its inhibitory effect on folliculogenesis in albino rats (Roop *et al.*, 2005). Leaf extracts of *M. azedarach* have antibacterial effect against several gram positive and negative pathogens viz., *Staphylococcus aureus*, *Escherichia coli, Enterococcus faecalis* and *Pseudomonas aeruginosa* (Adamu et al., 2014*)*. *M. azedarach* has remained a source of several folk medicines and its potential as anthelmintic, antiviral, antiprotozoal, antinephrolithiasis and antibacterial agent is reasonably validated by scientific literatures. Leaf, fruit and seed extract of this plant has been tested to manage malarial mosquito, dengue mosquito, lice, ticks etc. under laboratory conditions (AL-Rubae, 2009; Maciel *et al.*, 2006; Wachsman *et al.*, 1982).

Genetic structure of plants is largely determined by its spatial or actual geographical distribution along with ecological factors like dispersion, mutation, natural selection and genetic drift etc. Geographically isolated and Small populations are comparatively more prone to genetic drift and limited gene flow than to their widely distributed counterparts (Loveless and Hamrick, 1984). Geographical barriers or distant spatial patterns lead to limited gene flow between populations which consecutively increases genetic differentiation among

populations (Fischer *et al.*, 2000). Genetic variation in a species will allow itself to adapt, evolve, respond to environmental stress and survive for a long term (Sheng *et al.*, 2005). Knowledge of genetic structure not only provides evidences regarding gene flow, genetic drift, mutation and migration but also provides basic information for restoration plans and rational exploitation (Doligez and Joly, 1997; Fischer *et al.*, 2000). Forest trees are widely important but their breeding is slow, so it is desirable to detect functionally important regions in their genome. Microsatellite markers are of particular interest in genetic diversity and population structure analysis. Given the attributes like co-dominance, high polymorphism and reproducibility, microsatellite markers serves as an important tool to gather fine-scale ecological information (Gupta and Varshney, 2000). Considering the significance present study "**Development and characterization of genomic microsatellite markers in *Melia azedarach***" was undertaken to investigate genetic diversity and population structure of *M.azedarach* in Indian Thar Desert with following objectives:

- Development of genomic SSR markers
- Characterization of developed markers in selected germplasm

CHAPTER 2

Review of Literature

Forests play a significant role in recycling nutrients and water in ecosystem and hence maintaining its balance. They harbour a great genetic diversity within them by providing habitat to various plant and animal species. *M.azedarach* is an important forest tree with pharmaceutical, insecticidal and pesticidal values.

Viewing the importance of microsatellite markers for diversity and population structure analysis present study was undertaken and literature pertaining to the research problem is presented under the following subheadings:

2.1 Melia azedarach

2.1.1 Habitat and distribution

2.1.2 Uses

2.2 Diversity assessment and Molecular markers

2.3 Microsatellite markers

2.3.1 Strategies of microsatellite marker development

2.3.2 Microsatellite marker for population studies

2.1 Melia azedarach

M. azedarach is one of the important forest tree found all over the world. It is most apt for reforestation plans owing to its fast growth and hardy nature. It is also widely planted along roadsides and fencerows. It is well adapted to temperature as low as -5°C to as high as 40°C. It reproduces both by seeds and vegetative means (Sheikh, 1993). As a response to physical injury it reproduces by production of sprouts from root. Because of its ability to reproduce vegetatively and wide adaptability it has become aggressive invader at its non-native places (Tourn *et al.*, 1999).

2.1.1 Habitat and distribution

Its habitat varies from wet to dry and cold to hot, also it can occur in either extreme. In India it is found in dry monsoon climate while in Australia it is a rainforest canopy species (Waggy, 2013). It is native to northern-India, lower Himalayas including Pakistan and Nepal, central and western China and tropical

Australia but is widely distributed in other warmer parts of world together with North America, Mexico and Argentina.

2.1.2 Uses

M.azedarach has remained a crucial source of folk medicines. It is widely planted around the world for timber, forage and ornamental purposes. Scientific literature pertaining to its uses is given in Table1.

Table 1: List of studies validating various usage of *M. azedarach*

Sr.No.	Title	Result	Reference
1	Ovicidal and larvicidal activity of *Melia azedarach* extracts on *Haemonchus contortus*	Seed ethanol extract completely inhibited egg hatching of *H. contortus* at the concentration of 1.56 mg/mL (LC_{50}=0.36 mg/mL). Leaf ethanol extract showed better inhibition of larval development (LC_{50}=9.18 mg/mL) than seed ethanol extract (LC_{50}=98.0 mg/mL).	(Maciel *et al.*, 2006)
2	Extracts of *Azadirachta indica* and *Melia azedarach* seeds inhibit folliculogenesis in albino rats	Seed extract significantly reduced number of normal single layered follicles and total normal follicles in albino rats.	(Roop *et al.*, 2005)
3	Antifeedant and Insecticide Properties of a Limonoid from *Melia azedarach* (Meliaceae) with Potential Use for Pest Management	*M. azedarach* fruit extract inhibited the feeding activity of 17 pest species of three different orders viz., Coleoptera, Orthoptera and Lepidoptera.	(Carpinella *et al.*, 2003)
4	The antibacterial activity, antioxidant activity and selectivity index of leaf extracts of thirteen South African tree species used in ethnoveterinary medicine to treat helminth infections	Cytotoxicity 0.145 mg/mL and selectivity index of leaf acetone extracts varied from 0.23 to 0.91against selective bacterial pathogens.	(Adamu *et al.*, 2014)
5	Antiparasitic activity of Melia azedarach growing in Argentina	Drupe/fruit extracts of *M. azedarach* was found active against both the tapeworm and the earthworm. In this study drupe extract exhibited better antiparasitic activity than piperazine phosphate against *Taenia solium*.	(Szewczuk *et al.*, 2003)
6	Larvicidal action of ethanolic extracts from fruit endocarps of *Melia azedarach* and *Azadirachta indica* against the dengue mosquito *Aedes aegypti*	Seed extracts were found lethal to for third to fourth instar larvae of *Aedes aegypti*. *M. azedarach* extract exhibited 1.6 times more efficiency than *Azadirachta indica* with fed larvae at 25°C for all measured dose-response experiments (LC_{50}, LC_{95} and LC_{99}).	(Wandscheer *et al.*, 2004)

2.2 Diversity assessment and Molecular markers

Genetic variations are basis for life's diversity and vital for a population to acclimatize to changing environment. Factors like genetic drift and inbreeding lead to loss of variation in population and jeopardize its adaptability. Self-crossing population is less diverse and more differentiated then out-crosser (Charlesworth and Charlesworth, 1987). Knowledge about population structure and genetic diversity is important as it helps in conserving diversity of species. Anthropogenic activities like, farming, road building, urbanization, overexploitation of a species for timber or other uses etc. influence population structure significantly (Charlesworth and Charlesworth, 1987; Francisco-Ortega *et al.*, 2000). Extreme reduction in population size and habitat fragmentation pose negative impact on genetic diversity and survival of a species (Leimu *et al.*, 2006).

With the advent of technology like PCR, use of DNA based molecular markers has enormously increased to study genetic diversity. These markers are useful in certifying genetic and evolutionary relationships (Booy *et al.*, 2000). There are several techniques for DNA marking such as restriction fragment length polymorphism (RFLP), random amplification of polymorphic DNA (RAPD), amplified fragment length polymorphism (AFLP) microsatellites or simple sequence repeats (SSR) markers and single nucleotide polymorphism (SNP). Dominant markers viz., AFLP and RAPD are less informative as they can't depict heterozygosity. Microsatellite markers are co-dominant and possess other desired traits like polymorphism, reproducibility, random and wide genomic distribution.

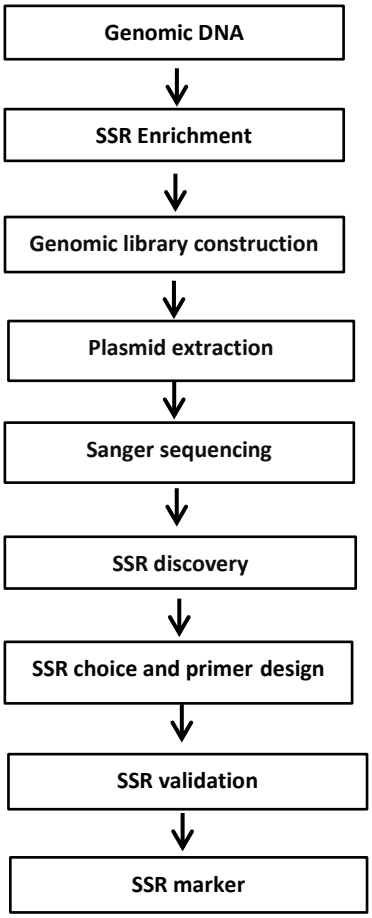
2.3 Microsatellite markers

The term Microsatellite was first coined by Litt and Luty in 1989 while working on cardiac muscle actin gene (Litt and Luty, 1989). Microsatellites or simple sequence repeats (SSR) are repeats of 1–6 nucleotides found randomly at high frequency in the genomes of most taxa but with varied frequencies of occurrence in different organisms. These repeats typically vary in length between 5 and 40, but longer repeat-sequences are also possible (Gupta and Varshney, 2000; Selkoe and Toonen, 2006). Dinucleotide repeats like $(AC)_n$ and $(GA)_n$ have been worked out most commonly but tri-nucleotide and tetra-nucleotide repeats like (AAG)n, and (AAT)n are also present in plant genomes (Gupta and Varshney,

2000). Polymorphism revealed by microsatellite markers is due to polymerase template slippage during replication. Polymerase template slippage in microsatellite region of genome can lead to incorrect dissociation and re-association (Kruglyak *et al.*, 1998)**.**

2.3.1 Strategies of microsatellite marker development

A conventional strategy to develop microsatellite markers using Sanger sequencing technology consists of DNA extraction, DNA digestion with a restriction enzyme, ligation of linkers to DNA fragments, enrichment for microsatellite-containing fragments, cloning and sequencing of products, microsatellite region identification, primer design and validation (Fig 1) (Glenn T.C. and Schable, 2005; Zalapa *et al.*, 2012).



modified (Zalapa *et al.*, 2012)

Fig 1: General scheme for microsatellite marker development

2.3.2 Microsatellite marker for population studies

Owing to their high polymorphism, reproducibility, co-dominance and availability throughout the genome, microsatellite markers have emerged as a marker of choice for population structure analysis (Table 2). These markers have high mutation rate which make them reasonably informative even in case of small populations (Selkoe and Toonen, 2006).

Table 2: List of various studies involving Microsatellite marker development and characterization

Species	**Popul-ation size**	**Fragments Sequenced**	**No of Motifs Found**	**Total no of Marker Develo-ped**	**No of Polymor-phic Markers**	**Reference**
Scutellaria baicalensis	94	125	89	38	21	(W. Zhang *et al.*, 2014)
Hevea brasiliensis	192	NA	NA	15	NA	(Souza, 2014)
Tapiscia sinensis	102	323	111	36	11	(P. Zhang *et al.*, 2014)
Camellia assamica ssp. *assamica*	29	1297	549	470	185	(Bhardwaj *et al.*, 2013)
*Cicer arietinum*L.	60	NA	NA	20	14	(Ghaffari *et al.*, 2014)
Vernicia fordii	81	400	196	78	40	(Yue *et al.*, 2013)
Lathyrus sativus L.	NA	410	119	104	7	(Lioi and Galasso, 2013)
Dipteryx alata	94	1013	58	28	8	(Soares *et al.*, 2012)
Arctium minus	134	352	NA	42	16	(López-Vinyallonga *et al.*, 2010)
Neolitsea sericea	46	110	80	37	10	(Zhai *et al.*, 2010)
Masdevallia solomonii	NA	178	160	26	10	(López-Roberts *et al.*, 2012)
Taihangia rupestris	40	36	25	18	10	(Wang *et al.*, 2010)
Pinus massoniana Lamb.	72	NA	NA	10	9	(Guan and Shiraishi, 2011)
Ophiorrhiza japonica	144	221	NA	17	9	(Nakamura *et al.*)
Dactylorhiza hatagirea	68	211	398	107	14	(Lin *et al.*, 2013)
Thuja plicata	44	96	NA	35	12	(O'Connell and Ritland, 2000)
Rhododendron arboreum	30	351	190	41	38	(Choudhary *et al.*, 2014)

CHAPTER 3

Materials and Methods

3.1 Plant material and microsatellite enriched library construction

A total of 14 populations consisting of 95 genotypes of *M. azedarach* have been collected from geographically isolated regions of Indian Thar Desert. Complete information regarding sample collection and location is given in Table 3 and Fig 2. Genomic DNA was isolated from young leaves by CTAB method (Doyle and Doyle, 1987) with minor modifications (Appendix A and B). A microsatellite (AG)n enriched genomic library was constructed using biotin-streptavidin capture method (Bhardwaj *et al.*, 2013) with some modifications. Enriched fragments were ligated to pGEM-T Easy (Promega) vector (Appendix A), later competent *Escherichia coli* were transformed with these recombinant plasmids. Clones positive for insert were confirmed by interrupted ß-galactosidase gene and secondary enrichment (Appendix A) through PCR amplification. Positive clones of secondary enrichment i.e. amplifying multiple bands were grown overnight at 37°C in a LB broth containing 100 µg/ml ampicillin. Plasmid DNA from these clones was isolated using NucleoSpin Plasmid miniprep kit (Macherey-Nagel) and sequenced with ABI 3730 xl DNA Analyzer using BigDye terminator cycle sequencing kit v3.1 (Applied Biosystems) as per the manufacturer's procedure.

Table 3: Geographical locations of *M. azedarach* populations from Indian Thar Desert

Population	Sample Size	Longitude (E)	Latitude (N)
Bathinda (BAT)	9	74°47'29.52"E	30°08'26.96"N
Hanumangarh (HAN)	10	74°17'32.04"E	29°36'48.33"N
Dabbali (DAB)	3	74°38'46.09"E	29°55'28.72"N
Sri Ganga nagar (SRI)	9	73°51'32.55"E	29°55'12.69"N
Kesrisinghpur (KES)	4	73°37'16.18"E	29°56'51.68"N
Padampur (PAD)	12	73°37'34.83"E	29°42'25.86"N
Jaisalmer (JAS)	7	70°21'48.26"E	26°52'23.50"N
Jodhpur (JOD)	3	72°59'57.12"E	26°13'17.32"N
Gharsana (GHA)	5	73°04'43.72"E	29°01'18.57"N
Anupgarh (ANP)	5	73°12'34.18"E	29°11'21.07"N
Raisinghnagar (RAI)	9	73°26'56.75"E	29°32'09.01"N
Karanpur (KAR)	5	76°58'14.85"E	26°10'55.44"N
Mahajan (MAH)	5	73°50'12.67"E	28°47'16.13"N
Lakhansar (LAK)	9	73°45'18.23"E	28°30'01.75"N
Total	95		

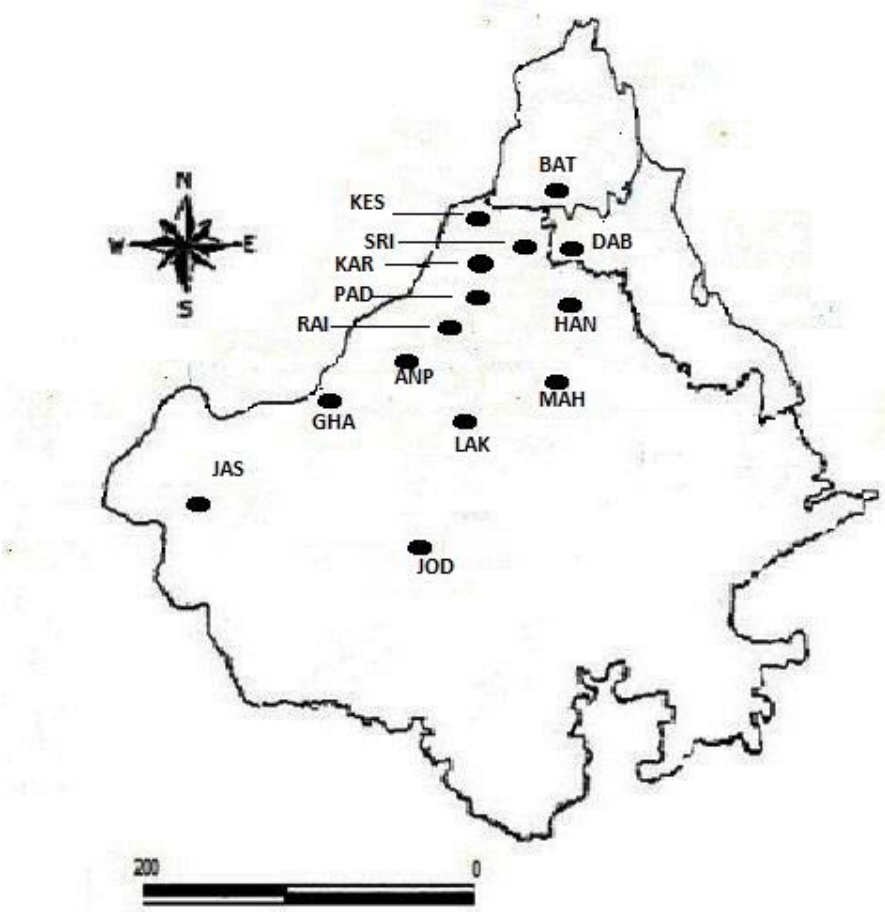


Fig 2. Sites of sample collection (solid circles) in Indian Thar Desert

3.2 Primer designing and characterization

SSRs containing sequences were identified using SSR Identification Tool (SSRIT) (Temnykh *et al.*, 2001). Primers were designed from region flanking SSRs using Primer3 web version 4.0.0 software (Untergasser *et al.*, 2012) after removing plasmid chimeras from sequenced fragments. For characterization, PCR reactions were carried out in a total volume of 10 µl, containing genomic DNA (25 ng), Taq DNA polymerase (0.3 U), PCR buffer (1X), dNTPs (2.5 mM, each), forward and reverse primer (5ng) in T100 Thermal Cycler (BIORAD). PCR profile comprised of an initial denaturation step at 94˚C for 3 min followed by 35 cycles of denaturation at 94˚C for 1 min, annealing at appropriate T_m for 1 min and elongation at 72˚C for 1 min and a final elongation at 72˚C for 8 min. Amplification reactions were mixed with equal volume of denaturing dye, heat denatured at 94˚C for 5 min and quantified by electrophoresis on 6% urea polyacrylamide gels

(Appendix A) at 60W for desired time depending on size of the fragment and visualized using silver staining.

3.3 Data Analysis:

Profiles for each of the Microsatellite primer pair were scored as a single band signifying a single allele. Genetic diversity estimation at parameters viz., effective number of alleles ($Ne_)$, observed and expected Heterozygosity (Ho, He), Shannon's information index (I) and UPGMA dendrogram construction, based on Genetic distance (Nei, 1978) was done using POPGENE software version 1.31 (Yeh *et al.*, 1999). The heterozygosity excess or deficiency was determined using Bottleneck (Luikart *et al.*, 1998).

The phylogenetic relationship among populations was depicted from DARwin v5.0.158 and STRUCTURE v2.3.4 (Pritchard *et al.*, 2000). Based on multi-locus Bayesian analysis, the hidden population clusters were distinguished. For estimation of optimum no. of populations (K), a simulation was conducted using parameters; K (1- 20) with a random start for each K value and 10 independent runs (Evanno *et al.*, 2005). Number of distinct population clusters (K) was determined using Structure Harvester (http://taylor0.biology.ucla.edu/ structureHarvester/), a web based program for visualizing STRUCTURE output and implementing the Evanno method (Earl, 2012).

Mantel's test to measure correlation between genetic and geographic distance and AMOVA were applied using GeneAlEx version 6.5 (Peakall and Smouse, 2012). Geographic distance (Km) among different populations was calculated from Latitude and longitude using 'haversine' formula (www.movable-type.co.uk/scripts/latlong.html).

CHAPTER 4

Results

4.1 Marker development and characterization

Out of 1912 recombinant clones from microsatellite $(AG)_n$ enriched genomic library, 573(30%) positive clones (Fig 3) were identified, 339 (59%) of which possessed SSRs with a range of 10-100bps. 291 sequences could not be used for designing primer pairs. 43 out of 48 remaining sequences were non redundant (NR) and utilized for primer designing. Developed primer pairs were used to amplify a pooled genomic DNA of 95 genotypes. Only 23(53.48%) primer pairs were used for further analysis as the remaining 20(46.52%) primer pairs led to weak or no amplification. Features and evaluation details of 23 successful primers pairs are given in Table 4.

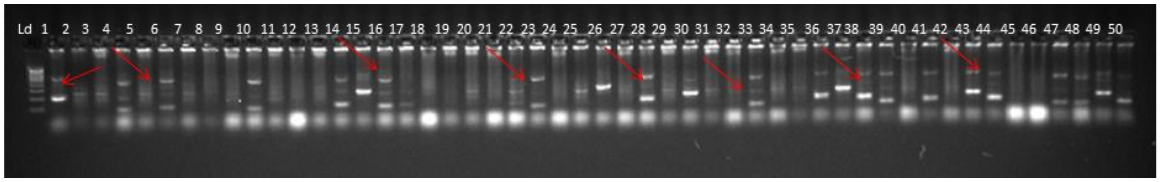


Fig 3: Lane 1-50 amplification profile generated using three-primer PCR; Ld: 100bp DNA ladder as size standard (Bangalore Genei^{TM}); lanes showing multiple bands (indicated by arrow) represent positive clones.

Polymorphic potential of 23 primer pairs was evaluated in a test array of 95 genotypes comprising 14 different populations from Thar Desert (Fig 4).

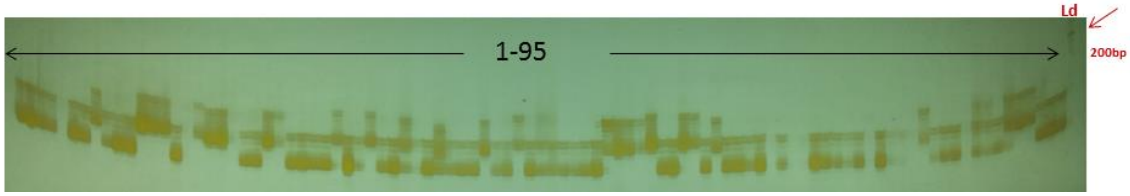


Fig 4: Amplification profile generated with primer MAZ-41. Lanes 1 - 95 represent sampled individuals of *M. azedarach*; Ld: 100 bp DNA ladder as size standard (Bangalore GeneiTM).

A total of 114 with average 4.96 alleles could be amplified by tested primer pairs in the range of 2-8 alleles per locus. Approximate size of amplified fragments in all tested primers ranged in between 90 - 392bp. Statistical analysis using software POPGENE version 1.31 revealed mean observed heterozygosity (Ho = 0.556) to be less than expected heterozygosity (He = 0.627). Polymorphic information content (PIC) varied from 0.131 to 0.832 with mean value of 0.572.

Table 4: Features and evaluation details of 23 microsatellite markers

Locus Name	Primer Sequence	Repeat Motif	Annealing Temperature	No. of Alleles	Heterozygosity*		PIC	I	Approximate Size Range (bp)	No. of Genotypes Amplified	Accession no.
					H_O	H_E					
MAZ-1	F5'AGGAAGAATGCCGCTGACTA R5'GGAAATGAAAACCGAAAGCA	$(CT)_5$	59	2	0	0.253	0.22	0.4181	214-216	95	KJ996073
MAZ-2	F5'GGGGAAGAGGGTCCAAGTT R5'TGAAAAACAATTATGTGATTTAGAAGA	$(CT)_{18}$	56	6	0.137***	0.735	0.684	1.4297	264-280	86	KJ996074
MAZ-3	F5'ACAATTGGGGAAGTCTGTGC R5'GCGCGAACTTCACTCTCTCT	$(AG)_{11},(AG)_6$	57	4	0.432***	0.709	0.651	1.2923	388-392	84	KJ996075
MAZ-5	F5'TCGTCATAACGCGAGAGTCA R5'CTTCGGCTTCTTCTGATTGG	$(CT)_{21}$	59	4	0.179***	0.67	0.596	1.135	140-160	63	KJ996076
MAZ-6	F5'TCCTGAGTAATTGCAGAATACACTAT R5'CCCCCACCCAATAAATACCT	$(AG)_{11},(AG)_9$	58	4	0.19***	0.566	0.484	0.9794	90-110	62	KJ996077
MAZ-7	F5'TTCTGGAAAACCAACCAACC R5'CCTGAGTAAAAGCTACTCTGAATGG	$(CT)_{16}$	58	5	0.947***	0.753	0.707	1.4692	176-186	90	KJ996078
MAZ-10	F5'TTAGGCATGGATCACAGAAAA R5'CAGATTGCTGCAAATTGGTAAA	$(AG)_{27}$	57	6	0.968***	0.613	0.532	1.1389	94-112	95	KJ996079
MAZ-12	F5'GGAAAGAGAGAAATGGTGCAA R5'GACGCGACTTGAACTCAAAA	$(AG)_{5\ ,}(AG)_{10}$	58	7	0.947***	0.778	0.739	1.6304	120-144	91	KJ996080
MAZ-13	F5'GGGTGTCTTTGGACGTGATT R5'CAACGCATGAAAGAGGAAAA	$(AG)_9,(AG)_{15}$	58	3	0.663	0.644	0.566	1.0587	180-190	95	KJ996081
MAZ-15	F5'AGCTCGAATCCATCCAGAAC R5'CCCTCTCTGTCTCTGACGCTAT	$(AG)_{8,}$ $(AG)_{4,}(AG)_4$	58	4	0.642*	0.518	0.438	0.8795	186-200	90	KJ996082
MAZ-17	F5'AATTGGTCTGATTTGACTCTCTCT R5'CCATGCCTCTATCTTGCTCTC	$(CT)_{17}$	56	4	0.8***	0.674	0.621	1.2411	188-198	84	KJ996083
MAZ-18	F5'TGGGCTACAAAATGAGAAAGG R5'TCAGGGGTATTCATTCATAGGG	$(AG)_{13},(AG)_{13}$	58	4	0.379	0.647	0.583	1.1607	146-160	79	KJ996084
MAZ-22	F5'TGAGTCCTGAGTAATGAAAAACAA R5'GTCCAAGTTGGCTCCTGGT	$(AG)_8,\ (AG)_7$	56	7	0.168***	0.752	0.71	1.5449	196-212	81	KJ996085
MAZ-41	F5'ACAACGCATGAAAGAGGAAAAG R5'CTTTGGACGTGATTGTGGGTATA	$(CT)_9,(CT)_{13}$	59	5	0.884***	0.751	0.707	1.4811	158-180	85	KJ996095
MAZ-37	F5'CCGTCAGGGACGTAATCCTA R5'CGGCTTCTTCTGATTGGTTC	$(CT)_{24}$	57	8	0.895	0.854	0.832	1.9793	176-220	90	KJ996091

MAZ-26	F5'TCCAAAGGCCACAATAACCA R5'TCACCAATGGAGAGGAAGCT	$(AG)_{16}$	59	5	0.042	0.275	0.262	0.608	172-190	88	KJ996086
MAZ-38	F5'ATTCCCACCTTTCCCATTTG R5'TTGCTTGAGTGCTTTTGTGG	$(CT)_{15}$	59	5	0.884***	0.583	0.493	1.0243	180-200	91	KJ996092
MAZ-29	F5'CCAGGCAGATAACGCAGAGA R5'GAGCCTAAACCCCACTTCCT	$(AG)_{11}$, $(AG)_5$	59	3	0.021	0.139	0.131	0.2933	170-180	89	KJ996088
MAZ-40	F5'TGAAAAACAATTATGTGATTTAGAAGA R5'TCACATGGAAGGAGTTTTGG	$(AG)_{19}$	56	7	0.779	0.814	0.784	1.7682	172-200	86	KJ996094
MAZ-35	F5'CCCCCTATCAAGACAAGCAA R5'GCTTGTCCTTTACTTTTGTTTTCC	$(AG)_{21}$	58	4	0.253***	0.689	0.633	1.2629	100-120	70	KJ996090
MAZ-30	F5'GGAAAGAAGAAGGTGGGTTCC R5'ACAAATAGAGGCGGCTGAGA	$(CT)_{14}$	58	5	0.79***	0.596	0.512	1.0332	185-200	86	KJ996089
MAZ-39	F5'AGGTCTTCGGCTTCTTCTGA R5'CCATCCCCGCCAAAATTCAA	$(AG)_{20}$	62	8	0.926	0.84	0.815	1.8893	150-190	89	KJ996093
MAZ-28	F5'ATTGGGGATGTTTGTGCCAG R5'ACTTGTTTTCTTCGCCATTCCT	$(AG)_{11,}(AG)_6$	60	4	0.863***	0.559	0.459	0.9222	185-195	91	KJ996087
Mean				4.956	0.556	0.627	0.572	1.2017			
SD				1.609	0.358	0.185	0.184	0.4279			

H_O= Observed Heterozygosity; H_E= Expected Heterozygosity; PIC= Polymorphic information content; I= Shannon's Information index; bp= base pairs; Significant deviation from Hardy-Weinberg equilibrium at *p=0.05, ***p=0.001

Table 5: Summary of F-statistics

Locus	Fis	Fit	Fst	Nm*
MAZ-1	1	1	0.1836	1.1115
MAZ-2	0.6779	0.7447	0.2072	0.9566
MAZ-3	0.2645	0.3911	0.1722	1.202
MAZ-5	0.7252	0.7779	0.1919	1.0524
MAZ-6	0.5135	0.6595	0.3002	0.5829
MAZ-7	-0.3239	-0.2373	0.0655	3.5694
MAZ-10	-0.6804	-0.5786	0.0606	3.8746
MAZ-12	-0.2863	-0.1767	0.0852	2.6844
MAZ-13	-0.2862	-0.0128	0.2125	0.9264
MAZ-15	-0.3354	-0.2479	0.0655	3.5693
MAZ-17	-0.294	-0.1903	0.0801	2.8702
MAZ-18	0.3026	0.3734	0.1016	2.2112
MAZ-22	0.7131	0.7614	0.1685	1.2337
MAZ-41	-0.2429	-0.1415	0.0816	2.8151
MAZ-37	-0.192	-0.0604	0.1104	2.0155
MAZ-26	0.8059	0.8718	0.3394	0.4865
MAZ-38	-0.6138	-0.4947	0.0738	3.137
MAZ-29	0.886	0.9047	0.1641	1.2736
MAZ-40	-0.0788	0.0301	0.101	2.2256
MAZ-35	0.542	0.6397	0.2134	0.9214
MAZ-30	-0.5229	-0.4431	0.0524	4.5187
MAZ-39	-0.2271	-0.1014	0.1024	2.1914
MAZ-28	-0.5437	-0.4729	0.0459	5.2002
Mean	-0.0204	0.1124	0.1302	1.6706

* Nm = Gene flow estimated from Fst = 0.25(1 - Fst)/Fst.

Table 6 Summary of Analysis of Molecular Variance (AMOVA) for 14 *M.azedarach* populations

Source of Variation	d.f.	Sum of Squares	Variance component	Percentage of Variation
Among Population	13	151.699	0.35767	4.94
Within Population	176	1210.28	6.87659	95.06
Total	189	1361.979	7.23427	

d.f= degrees of freedom

Summary of F-Statistics and Gene Flow (Table 5) showed mean values for Fis, Fit, Fst, and Nm were -0.020, 0.112, 0.130, and 1.670, respectively. Variation among populations was found to be less than that of within population as revealed by Analysis of Molecular Variance (AMOVA) across 14 tested populations (Table 6).

4.2 Genetic diversity

Genetic diversity within and among 14 populations was analysed by eight parameters viz., No. of different alleles (Na), No. of effective alleles (Ne), Gene diversity (h), Shannon's Information index (I), Observed heterozygosity (Ho), Expected heterozygosity (He), Fixation index (F), Percentage of polymorphic loci (P); their highest values were 3.826 (BAT), 2.934 (KAR), 0.714 (JOD), 1.059 (RAI), 0.641 (KES), 0.588 (KAR), 0.362 (KAR), 100%, respectively (Table 7).

Table 7: Genetic diversity of populations

Population	Na	Ne	I	Ho	He	F	h	P
Jaisalmer (JAS)	3.261	2.842	0.983	0.559	0.555	-0.035	0.600	86.96
Jodhpur (JOD)	2.696	2.509	0.903	0.478	0.556	0.154	0.714	95.65
Bathinda (BAT)	3.826	2.905	1.059	0.565	0.572	0.006	0.609	91.30
Hanumangarh (HAN)	3.565	2.622	1.001	0.587	0.557	0.0150	0.587	95.65
Dabwali (DAB)	2.478	2.266	0.786	0.638	0.493	-0.290	0.580	91.30
Sriganganagar (SRI)	3.261	2.569	0.97	0.609	0.552	-0.014	0.582	100.00
Kesarisinghpur (KES)	2.913	2.543	0.926	0.641	0.553	-0.132	0.630	95.65
Ghasana (GHA)	3	2.504	0.89	0.617	0.515	-0.196	0.566	86.96
Anupgarh (ANP)	3.043	2.698	0.965	0.539	0.563	0.094	0.636	95.65
Raisinghnagar (RAI)	3.652	2.842	1.05	0.585	0.58	0.005	0.616	100.00
Padampur (PAD)	3.652	2.704	1.023	0.471	0.564	0.253	0.594	100.00
Karanpur (KAR)	3.348	2.934	1.043	0.4	0.588	0.362	0.685	100.00
Mahajan (MAH)	2.826	2.475	0.869	0.565	0.51	-0.045	0.566	86.96
Lakhansar (LAK)	3.435	2.511	0.954	0.56	0.537	0.047	0.569	95.65
Total	3.211	2.637	0.959	0.558	0.549	0.021	0.61	94.41

Na = No. of Different Alleles; Ne = No. of Effective Alleles = 1 / (Sum pi^2); I = Shannon's Information Index = -1* Sum (pi * Ln (pi)); Ho = Observed Heterozygosity = No. of Hets / N; He = Expected Heterozygosity = 1 - Sum pi^2, F = Fixation Index = (He - Ho) / He = 1 - (Ho / He); h= gene Diversity; P= Percentage of polymorphic loci(%); where pi is the frequency of the ith allele for the population & Sum pi^2 is the sum of the squared population allele frequencies

Overall Ho (0.558) was more than He (0.549). Total genetic differentiation among all the populations was low (0.021) as suggested by the values of Fixation Index (F). Shannon's information index, I, an alternative for quantifying biological diversity indicated moderate (0.959) diversity among the *M.azedarach* populations in Indian Thar Desert.

To test any significant heterozygosity deficiency that is experienced by recently bottlenecked populations, mutation models IAM and SMM were applied using Bottleneck version 1.2.02 (Table 8). Observed and expected heterozygosity excess under sign test in IAM and SMM are 8, 12.23 (p=0.058); 5, 13.38 (p=0.004), respectively. Similarly, in standardized difference test T_2 exhibited negative values (-6.658, -19.28) with p=0.000 under both IAM and SMM, while probability of heterozygosity excess under wilcoxen rank test is 0.98 and 0.99 respectively.). Allele frequency distribution under "mode-shift" indicator was normal L-shaped (Fig 5).

Table 8: Allele frequency based mutation drift equilibrium

Mutation Model	Sign Test	Standardized Difference Test	Wilcoxen Test
IAM	Hee = 12.23 Hd = 15 He = 8 p = 0.05821	T_2= -6.658 p=0.000	Probability (one tail for Hd): 0.01633 Probability (one tail for He): 0.98496 Probability (two tails for He and Hd): 0.03267
SMM	Hee = 13.38 Hd = 18 He = 5 p = 0.0004	T_2= -19.2888 p=0.000	Probability (one tail for Hd): 0.00056 Probability (one tail for He): 0.99951 Probability (two tails for He or Hd): 0.00112

IAM= Infinite allele model; SMM= Stepwise mutation model; Hee= Expected Heterozygosity excess; Hd= Heterozygosity deficiency; He= Heterozygosity excess

4.3 Population structure

Unweighted Neighbor-joining (UNJ), tree constructed using DARwin v5.0.158 determines phylogenetic relationship between the individuals on the basis of genetic dissimilarity. Genotypes from different populations were clustered together under UNJ and there was no clear clustering according to geographical locations (Fig 6).

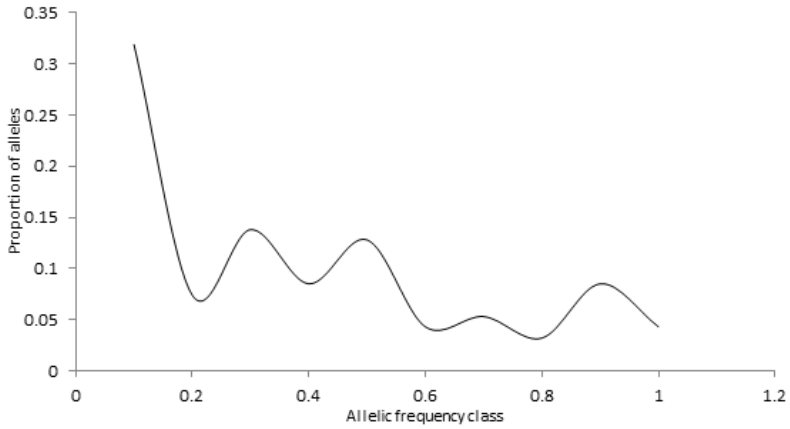


fig 5. Graphic representation of proportion of alleles and their distribution in populations of *M.azedarach*

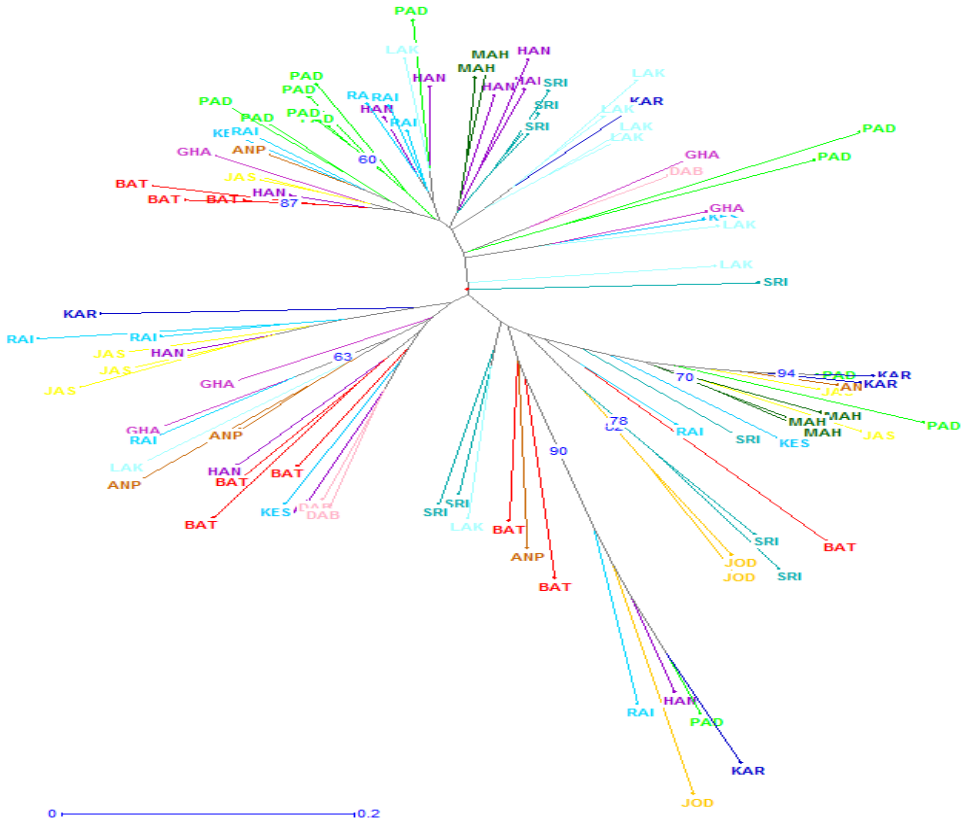


Fig 6. Unrooted neighbor joining (UNJ) tree constructed from DARwin v5.0.158; each branch represents single individual collected from corresponding inferred population

STRUCTURE v2.3.4 based on Bayesian approach was applied to search for hidden population structure among genotypes of different populations of *M. azedarach.* The relative distinctiveness of sampled populations was determined by assigning individuals on the basis of genetically homogeneous groups rather than on geographical locations of sampling. The highest value of ΔK identified 2 subpopulations or genetic stocks that were present among 14 sampled populations. This grouping is consistent with UNJ tree and UPGMA dendrogram (Fig 7).

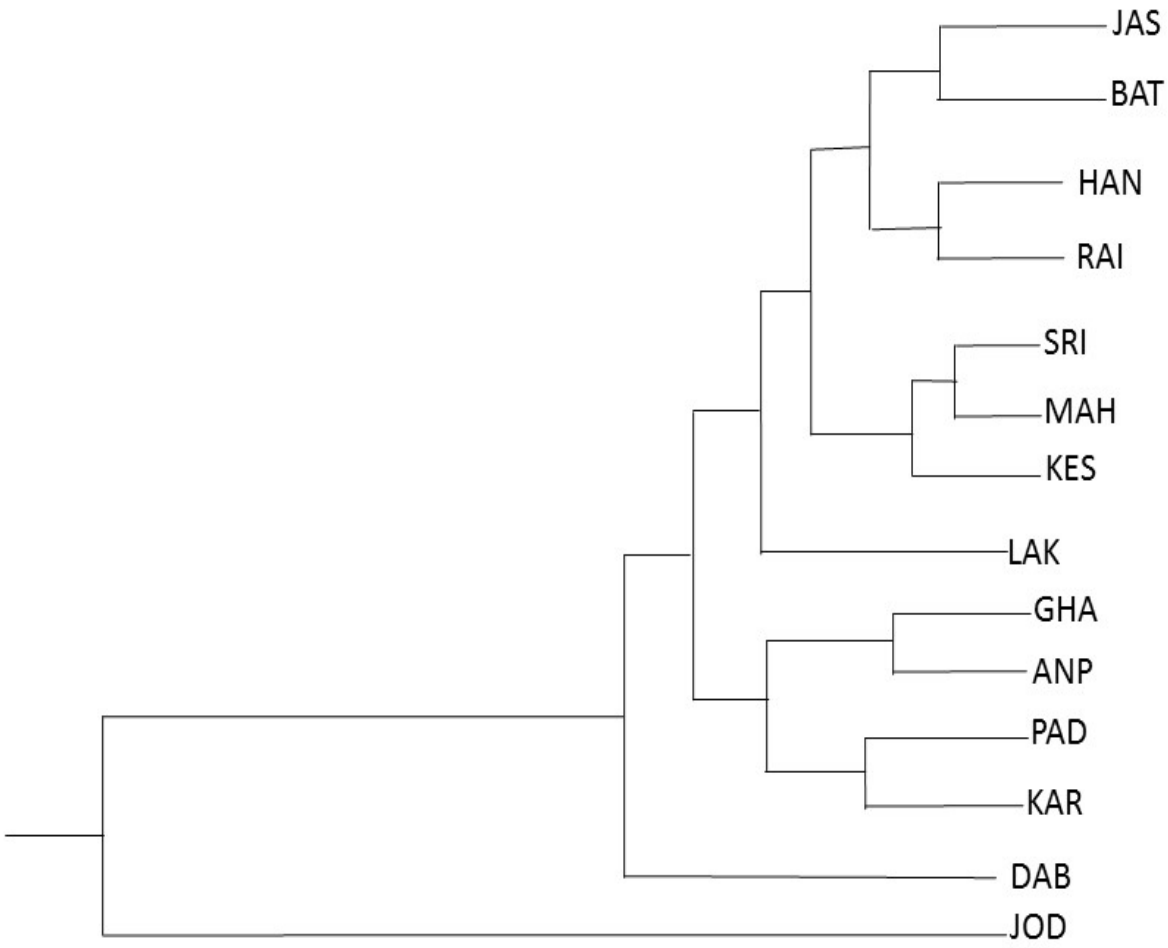


Fig 7. Dendrogram based on Nei's genetic distance for 14 populations of *M.azedarach* based on 23 SSR loci

Graphical method used to detect true number of populations i.e. K indicated mostly the number of subpopulations at K=2 considering ΔK distribution (Fig 8).

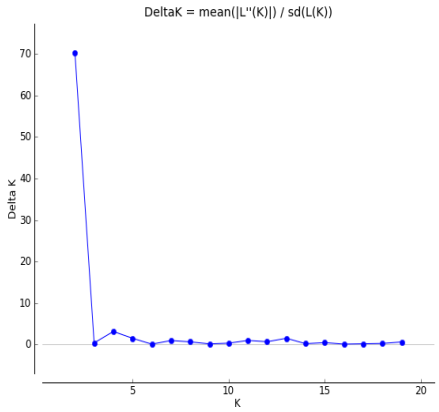


Fig 8: Estimation of optimum K using the second order statistics (ΔK) given by Evanno *et al.*, (2005)

Of the 95 individuals studied, first smaller genetic cluster occupied 7.37% (7 individuals), second larger genetic cluster occupied 16.84% (16 individuals) and the admixed cluster occupied 75.79% (72 individuals) (Fig 9).

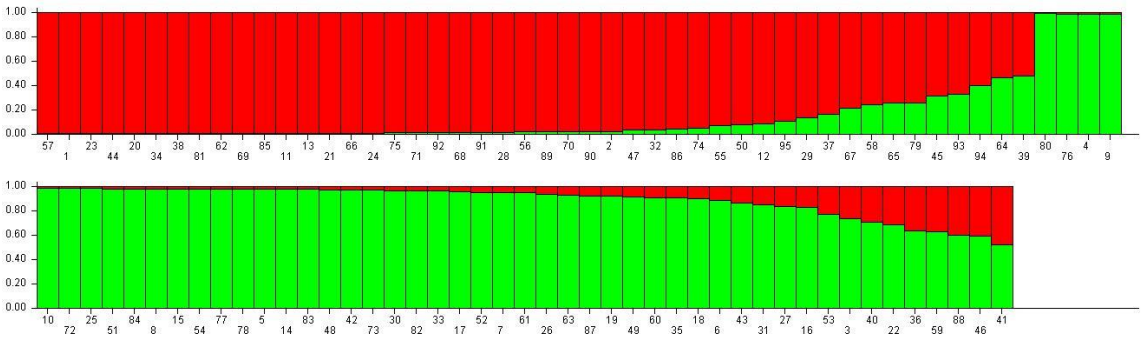


Fig 9: Bar plot given by Structure with membership coefficient (Q) of individuals on y- axis. Two clusters inferred by Structure are combined from clustering of all individuals. Each color represents different genetic stock and each vertical line represents individual

Genetic distances placed all the populations in two major clusters (I, II) one consisting of JOD (I) and other (II) remaining 13 populations according to generation of UPGMA Dendrogram (Fig 7) based on Nei's genetic distance. Sub-clusters, IIa and IIb consisted of one (DAB), and 12 remaining populations. Though the population pairs, BAT-JAS, HAN-RAI, and SRI-MAH are not closest geographically, they were clustered together in sub-subclusters of IIb. Two geographically close population pairs, PAD-KAR and GHA-ANP were clustered together in dendrogram.

To test if geographic distances have any relationship with genetic distances, data was further subjected to Mantel's test.

4.4 Isolation by distance

Mantel's test allows statistical testing for a correlation between genetic and geographical distance to assess the "isolation by distance" hypothesis. Pairwise population F_{st} ($F_{st}P$) values ranged from -0.024 to 0.178 for all the 14 populations (Appendix C). Though geographical distance between ANP, KAR is not closest the population pair observed minimum F_{st}. Scatter plots of pairwise F_{st} and geographical distance (Fig 10) showed a positive but non-significant correlation (Rxy = 0.261, P = 0.18).

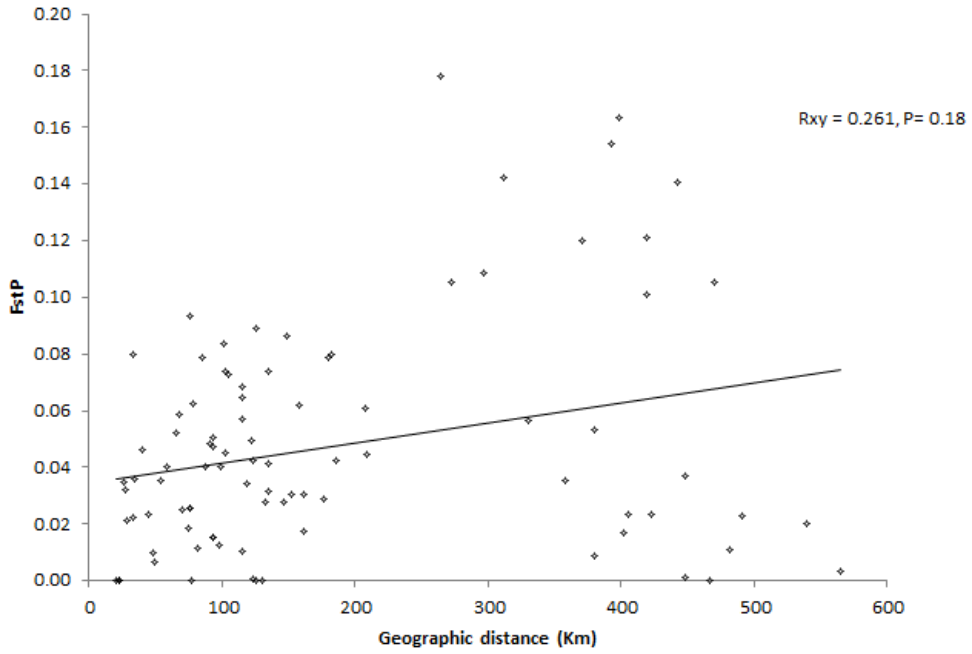


Fig 10: Relationship between genetic and geographic distances for 14 *M.azedarach* populations. Genetic distance is represented by pairwise F_{st} and geographic distance is given in kilometers.

CHAPTER 5

Discussion

Genetic diversity is the basic source of biodiversity and its maintenance plays an important role in conservation of species. Genomic SSR markers were developed in present study for diversity and population structure analysis of *M.azedarach.* Only 48 of 339 (14.15%) SSRs containing sequences could be utilised for primer designing. This low number of apt sequences for primer designing was due to presence of SSRs too close to one of the cloning sites. Primer pairs with redundant sequences and weak or monomorphic amplifications were not utilised and discarded from further analysis. Analysis of 23 successful primer pairs revealed mean observed heterozygosity (H_o) to be less than expected heterozygosity (H_e), although 11 of 23 (47.82 %) primer pairs showed significantly higher H_o than H_e. Mean heterozygosity per locus depicts the variation present in a population (Nei and Roychoudhury, 1974). A substantial number of allele (4.956) per locus were observed which could be attributed to high polymorphic nature of SSR markers and occurrence of cross-pollination in *M. azedarach* (Du *et al.*, 2012; Waggy, 2013). Polymorphic information content (PIC) quantifies the degree of polymorphism associated with a locus (Shete *et al.*, 2000). Sixteen of 23 (69.15%) tested markers were highly polymorphic (PIC>0.5). F-statistics is widely employed for genetic differentiation assessment (Goudet, 1995). Mean Fst (0.130) indicated very low global partitioning of genetic variability. Also, a moderate level on gene flow is detected over all the tested loci. Analysis of Molecular Variance (AMOVA) also supported F-statistics, in which only 4.94% variation was detected among populations while within population variation was found to be 95.06%. Allele frequency distribution under “Mode Shift” indicator is normal L-shaped which suggests that populations of *M.azedarach* in Indian Thar desert are not experiencing any recent bottlenecks.

Cluster analysis for determining the population structure of *M. azedarach* didn’t show any clear clustering pattern of geographically closer individuals. Bar plot at optimum ΔK (K=2), using STRUCTURE, clustered individuals of different populations together. This represented that a high level of gene flow has occurred which was in accordance with AMOVA and F-statistics results. High gene flow between populations which were geographically distant from each other could be attributed to effective cross-pollination, seed dispersal mechanisms and human interference (PospiSkova and Bartakova, 2004). Clustering of different populations

of *M. azedarach* in dendrogram based on Nei's genetic distance was not completely congruent with their geographic locations while it does cluster few geographically distant populations separately (JOD) and close populations together (PAD-KAR). We could neither establish nor disprove any relationship between genetic and geographic distances using cluster analysis in present study. So Mantel's test was carried out for further assigning of any correlation between genetic and geographic distances. Mantel's test compares two distances matrices with matching entries with a null hypothesis of no correlation. We found a positive but non-significant correlation (Rxy = 0.261, P = 0.18) between genetic and geographical distances. So, no significant correlation could be established which in turn along with presence of admixed cluster indicated occurrence of gene flow. Fourteen *M. azedarach* populations may have originated from two ancestral gene pools and different populations considered in the same genetic clusters, suggests recent common ancestry along with gene flow to be important factors. Present study has provided understanding of population structure as well as diversity pattern of *M.azedarach* in the Indian Thar desert, which in turn can help in selection of plants for breeding programmes and rational utilization of this plant.

Summary

Owing to wide importance of forest trees, detection of their genetic diversity and population structure analysis is desirable. *Melia azedarach* is an important forest tree known for its innumerable therapeutic uses and bio-control measures. Present study was aimed at investigating genetic diversity and population structure of *M.azedarach* in Indian Thar Desert. A set of 43 genomic microsatellite markers were developed and subsequently 23 polymorphic markers were employed for characterization across 14 populations encompassing 95 genotypes. Profiles of each primer pair were recorded as single band signifying single allele and afterwards subjected to statistical analysis using software viz., POPGENE v1.31, Bottleneck v1.2.02, STRUCTURE v2.3.4, structure harvester and GeneAlEx v6.5. Mean values of Observed heterozygosity (Ho), Expected Heterozygosity (He), Polymorphic information content (PIC) and Shannon's Information index(I) for 23 tested loci over 95 genotypes were 0.556, 0.627,0.572 and 1.2017, respectively. F-Statistics showed mean values for Fis, Fit, Fst, and Nm were -0.020, 0.112, 0.130, and 1.670, respectively. F-Statistics indicated very low global partitioning Fst (0.130) of genetic variability and a moderate level on gene flow over all the tested loci which was also supported by Analysis of Molecular Variance (AMOVA) in which only 4.94% variation was detected among populations while within population variation was found to be 95.06%. Sign test, standardized different test, Wilcoxen rank test and "Mode Shift" indicator revealed populations of *M.azedarach* in Indian Thar desert are not experiencing any recent bottlenecks.

Unweighted Neighbor-joining (UNJ) and UPGMA dendrogram were employed for determining phylogenetic relationship between the tested genotypes, no clear clustering pattern of geographically closer individuals was detected. The highest value of ΔK identified 2 subpopulations or genetic stocks that were present among 14 sampled populations. To test if geographic distances have any relationship with genetic distances, data was subjected to Mantel's test. No significant correlation (Rxy = 0.261, P = 0.18) could be established between geographical and genetic distance in this study. This study established the basis for more precise inference about the biogeography and rational utilization of *M. azedarach* in the Indian Thar Desert.

References

Adamu, M., Naidoo, V. and Eloff, J.N. (2014). The antibacterial activity, antioxidant activity and selectivity index of leaf extracts of thirteen South African tree species used in ethnoveterinary medicine to treat helminth infections. *BMC veterinary research* **10**: 52.

AL-Rubae, A.Y. (2009). The Potential Uses of *Melia Azedarach* L. as Pesticidal and Medicinal Plant, Review. *American-Eurasian Journal of Sustainable Agriculture* **3**.

Bhardwaj, P., Kumar, R., Sharma, H., Tewari, R., Ahuja, P.S. and Sharma, R.K. (2013). Development and utilization of genomic and genic microsatellite markers in Assam tea (*Camellia assamica* ssp. *assamica*) and related *Camellia* species. *Plant Breeding* **132**: 748-763.

Bhuyan, P., Khan, M. and Tripathi, R. (2003). Tree diversity and population structure in undisturbed and human-impacted stands of tropical wet evergreen forest in Arunachal Pradesh, Eastern Himalayas, India. *Biodiversity & Conservation* **12**: 1753-1773.

Booy, G., Hendriks, R., Smulders, M., Groenendael, J.V. and Vosman, B. (2000). Genetic diversity and the survival of populations. *Plant biology* **2**: 379-395.

Carpinella, M.C., Defago, M.T., Valladares, G. and Palacios, S.M. (2003). Antifeedant and insecticide properties of a limonoid from *Melia azedarach* (Meliaceae) with potential use for pest management. *Journal of Agricultural and Food Chemistry* **51**: 369-374.

Charlesworth, D. and Charlesworth, B. (1987). Inbreeding depression and its evolutionary consequences. *Annual review of ecology and systematics*: 237-268.

Choudhary, S., Thakur, S., Saini, R. and Bhardwaj, P. (2014). Development and characterization of genomic microsatellite markers in *Rhododendron arboreum*. *Conservation Genetics Resources*: 1-4.

Doligez, A. and Joly, H.I. (1997). Genetic diversity and spatial structure within a natural stand of a tropical forest tree species, *Carapa procera* (Meliaceae), in French Guiana. *Heredity* **79**: 72-82.

Doyle, J. and Doyle, J. (1987). A rapid DNA isolation procedure for small quantities of fresh leaf tissue. *Phytochemical Bulletin* 19: 11-15.

Du, Q., Wang, B., Wei, Z., Zhang, D. and Li, B. (2012). Genetic diversity and population structure of Chinese white poplar (*Populus tomentosa*) revealed by SSR markers. *Journal of Heredity* **103**: 853-862.

Earl, D.A. (2012). STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conservation Genetics Resources* **4**: 359-361.

Evanno, G., Regnaut, S. and Goudet, J. (2005). Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Molecular Ecology* **14**: 2611-2620.

Fischer, M., Husi, R., Prati, D., Peintinger, M., van Kleunen, M. and Schmid, B. (2000). RAPD variation among and within small and large populations of the rare clonal plant *Ranunculus reptans* (Ranunculaceae). *American Journal of Botany* **87**: 1128-1137.

Francisco-Ortega, J., Santos-Guerra, A., Kim, S.-C. and Crawford, D.J. (2000). Plant genetic diversity in the Canary Islands: a conservation perspective. *American Journal of Botany* **87**: 909-919.

Ghaffari, P., Talebi, R. and Keshavarzi, F. (2014). Genetic diversity and geographical differentiation of Iranian landrace, cultivars, and exotic chickpea lines as revealed by morphological and microsatellite markers. *Physiology and Molecular Biology of Plants*: 1-9.

Glenn T.C. and Schable, N.A. (2005). Isolating microsatellite DNA loci. *Methods in Enzymology* **395**: 202-222.

Goudet, J. (1995). FSTAT (version 1.2): a computer program to calculate F-statistics. *Journal of Heredity* **86**: 485-486.

Guan, L. and Shiraishi, S. (2011). Isolation and characterization of tetranucleotide microsatellite loci in *Pinus massoniana* (Pinaceae). *American journal of botany* **98**: e216-e217.

Gupta, P. and Varshney, R. (2000). The development and use of microsatellite markers for genetic analysis and plant breeding with emphasis on bread wheat. *Euphytica* **113**: 163-185.

Hare, W., Garland, T. and Barr, A. (1998). Chinaberry (*Melia azedarach*) poisoning in animals. *Toxic plants and other natural toxicants*: 514-516.

Kruglyak, S., Durrett, R.T., Schug, M.D. and Aquadro, C.F. (1998). Equilibrium distributions of microsatellite repeat length resulting from a balance between

slippage events and point mutations. *Proceedings of the National Academy of Sciences* **95**: 10774-10778.

Leimu, R., Mutikainen, P., Koricheva, J. and Fischer, M. (2006). How general are positive relationships between plant population size, fitness and genetic variation? *Journal of Ecology* **94**: 942-952.

Lin, P., Zeng, L., Yang, Z., Liu, R. and Zhong, Y. (2013). Development and characterization of polymorphic microsatellite marker for *Dactylorhiza hatagirea* (D. Don) Soo. *Conservation Genetics Resources*: 1-3.

Lioi, L. and Galasso, I. (2013). Development of genomic simple sequence repeat markers from an enriched genomic library of grass pea (*Lathyrus sativus* L.). *Plant Breeding* **132**: 649-653.

Litt, M. and Luty, J.A. (1989). A hypervariable microsatellite revealed by in vitro amplification of a dinucleotide repeat within the cardiac muscle actin gene. *American journal of human genetics* **44**: 397.

López-Roberts, M.C., Almeida, P.R., Oliveira, E.J. and van den Berg, C. (2012). Microsatellite marker development for the threatened orchid *Masdevallia solomonii* (Orchidaceae). *American journal of botany* **99**: e66-e68.

López-Vinyallonga, S., Arakaki, M., Garcia-Jacas, N., Susanna, A., Gitzendanner, M.A., Soltis, D.E. and Soltis, P.S. (2010). Isolation and characterization of novel microsatellite markers for *Arctium minus* (Compositae). *American journal of botany* **97**: e4-e6.

Loveless, M. and Hamrick, J. (1984). Ecological determinants of genetic structure in plant populations. *Annual Review of Ecology and Systematics* **15**: 65-95.

Luikart, G., Allendorf, F., Cornuet, J. and Sherwin, W. (1998). Distortion of allele frequency distributions provides a test for recent population bottlenecks. *Journal of Heredity* **89**: 238-247.

Maciel, M., Morais, S.M., Bevilaqua, C., Camurça-Vasconcelos, A., Costa, C. and Castro, C. (2006). Ovicidal and larvicidal activity of *Melia azedarach* extracts on *Haemonchus contortus*.*Veterinary parasitology* **140**: 98-104.

Nakamura, K., Ma, C.-J., Huang, C.-J., Ho, M.-J., Yang, H.-A. and Peng, C.-I. Development and characterization of microsatellite markers in the coffee-family herb *Ophiorrhiza japonica* (Rubiaceae). *Conservation Genetics Resources*: 1-3.

Nei, M. and Roychoudhury, A. (1974). Sampling variances of heterozygosity and genetic distance. *Genetics* **76**: 379-390.

O'Connell, L.M. and Ritland, C.E. (2000). Characterization of microsatellite loci in western redcedar (*Thuja plicata*). *Molecular ecology* **9**: 1920-1922.

Peakall, R. and Smouse, P.E. (2012). GenAlEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research—an update. *Bioinformatics* **28**: 2537-2539.

PospiSkova, M. and Bartakova, I. (2004). Genetic diversity of a black poplar population in the Morava river basin assessed by microsatellite analysis. *Forest genetics* **11**: 257.

Pritchard, J.K., Stephens, M. and Donnelly, P. (2000). Inference of population structure using multilocus genotype data. *Genetics* **155**: 945-959.

Roop, J., Dhaliwal, P. and Guraya, S. (2005). Extracts of *Azadirachta indica* and *Melia azedarach* seeds inhibit folliculogenesis in albino rats. *Brazilian journal of medical and biological research* **38**: 943-947.

Selkoe, K.A. and Toonen, R.J. (2006). Microsatellites for ecologists: a practical guide to using and evaluating microsatellite markers. *Ecology letters* **9**: 615-629.

Sen, A. and Batra, A. (2011). *Melia azedarach* L.–A Paradise tree. *Journal of Functional and Environmental Botany* **1**: 59-69.

Sheikh, M.I. (1993). *Trees of Pakistan*: Pictorial Printers.

Sheng, Y., Zheng, W., Pei, K. and Ma, K. (2005). Genetic variation within and among populations of a dominant desert tree *Haloxylon ammodendron* (Amaranthaceae) in China. *Annals of Botany* **96**: 245-252.

Shete, S., Tiwari, H. and Elston, R.C. (2000). On estimating the heterozygosity and polymorphism information content value. *Theoretical population biology* **57**: 265-271.

Soares, T.N., Melo, D.B., Resende, L.V., Vianello, R.P., Chaves, L.J., Collevatti, R.G. and de Campos Telles, M.P. (2012). Development of microsatellite markers for the Neotropical tree species *Dipteryx alata* (Fabaceae). *American journal of botany* **99**: e72-e73.

Souza, L.M. (2014). Population Genetic Structure Based on Microsatellite Markers in Rubber Tree (*Hevea brasiliensis*) Germplasm Collection. Paper presented at the Plant and Animal Genome XXII Conference.

Szewczuk, V.D., Mongelli, E.R. and Pomilio, A.B. (2003). Antiparasitic activity of *Melia azedarach* growing in Argentina. *Mol Med Chem* **1**: 54-57.

Temnykh, S., DeClerck, G., Lukashova, A., Lipovich, L., Cartinhour, S. and McCouch, S. (2001). Computational and experimental analysis of microsatellites in rice (*Oryza sativa* L.): frequency, length variation, transposon associations, and genetic marker potential. *Genome Research* **11**: 1441-1452.

Tourn, G., Menvielle, M., Scopel, A. and Pidal, B. (1999). Clonal strategies of a woody weed: *Melia azedarach*. *Plant and soil* **217**: 111-117.

Untergasser, A., Cutcutache, I., Koressaar, T., Ye, J., Faircloth, B.C., Remm, M. and Rozen, S.G. (2012). Primer3—new capabilities and interfaces. *Nucleic acids research* **40**: e115-e115.

Wachsman, M., Martino, V., Gutkind, G.O., Coussio, J.D. and Coto, C.E. (1982). Antiviral activity of a *Melia azedarach* plant extract. *Fitoterapia*.

Waggy, M. (2013). *Melia azedarach*. Fire effects information system. [cited 2014 May 10] Available: http://www. fs. fed. us/database/feis.

Wandscheer, C.B., Duque, J.E., da Silva, M.A., Fukuyama, Y., Wohlke, J.L., Adelmann, J. and Fontana, J.D. (2004). Larvicidal action of ethanolic extracts from fruit endocarps of *Melia azedarach* and *Azadirachta indica* against the dengue mosquito *Aedes aegypti*. *Toxicon* **44**: 829-835.

Wang, H.-W., Zhang, B., Wang, Z.-S., Cheng, Y.-Q. and Ye, Y.-Z. (2010). Development and characterization of microsatellite loci in *Taihangia rupestris* (Rosaceae), a rare cliff herb. *American journal of botany* **97**: e136-e138.

Yeh, F., Yang, R. and Boyle, T. (1999). POPGENE. Microsoft Windows-Based Freeware for Population Genetic Analysis: Release 1.31. *University of Alberta, Edmonton*.

Yue, P., Lei, P., LiAng, C., LingLing, Z., Nevo, E. and JunHua, P. (2013). Development of microsatellite markers in the oil-producing species *Vernicia fordii* (Euphorbiaceae), a potential biodiesel feedstock. *Applications in Plant Sciences* **1**.

Zalapa, J.E., Cuevas, H., Zhu, H., Steffan, S., Senalik, D., Zeldin, E., McCown, B., Harbut, R. and Simon, P. (2012). Using next-generation sequencing approaches to isolate simple sequence repeat (SSR) loci in the plant sciences. *American journal of botany* **99**: 193-208.

Zhai, S.-N., Yan, X.-L., Nakamura, K., Mishima, M. and Qiu, Y.-X. (2010). Isolation of compound microsatellite markers for the endangered plant *Neolitsea sericea* (Lauraceae). *American journal of botany* **97**: e139-e141.

Zhang, P., Wang, Z., Liu, Y., Tian, H., Yao, X. and Zhang, J. (2014). Development and Characterization of 11 Polymorphic Microsatellite Markers in *Tapiscia sinensis* (Staphyleaceae). *Applications in Plant Sciences* **1**.

Zhang, W., Yuan, Q., Jiang, D., Zhang, Y. and Huang, L. (2014). Development and characterisation of microsatellite markers for the medicinal plant *Scutellaria baicalensis* (Lamiaceae). *Biochemical Systematics and Ecology* **54**: 267-271.

Appendices

APPENDIX- A

COMPOSITION OF CHEMICALS

I. DNA EXTRACTION

DNA EXTRACTION BUFFER

1 M TrisCl (pH 8.0)	: 10.0 ml
0.5 M Na.EDTA (pH 8.0)	: 4.0 ml
5 M NaCl	: 35.0 ml
β- mercaptoethanol	: 0.2 ml (add just before use)
4% CTAB	: 15.0 ml
10% PVP	: 15.0 ml

Make final volume upto 100 ml by adding autoclaved distilled water

TAE (50X)

Tris	: 242 g
Glacial acetic acid	: 57.1 ml
0.5 M EDTA	: 100 ml

$T_{10}E_1$ Buffer

10mM Tris.Cl (pH 8.0)
1mM Na.EDTA (pH 8.0)

II. RESTRICTION

Genomic DNA	: 5 µg
EcoRI (NEB)	: 10 units/ µg of DNA
MseI (NEB)	: 10 units/ µg of DNA
Cut Smart Buffer (NEB)	: 1X
Total volume	: 40 µl (by adding MilliQ water)

Incubated at 37˚C for 10-15 min. Stopped the enzymatic activity by incubating at 65˚C for 20 min.

III. LIGATION

Restricted DNA	: 40 µl
EcoRI adapters	: 20 pmol
MseI adapters	: 20 pmol
Ligase	: 10 units/ µg of DNA
Buffer	: 1X
Total volume	: 50 µl (by adding MilliQ water)

Incubated at 16˚C for 12 hrs. Stopped the enzymatic activity by incubating at 65˚C for 20 min.

IV. SECONDARY ENRICHMENT

Colony Suspension	: 2 µl
EcoRI Primer	: 20 pmol
MseI Primer	: 20 pmol
Probe Primer	: 20 pmol
dNTPs	: 2.5 mM
$MgCl_2$	: 1.2 mM
Taq Polymerase	: 0.3 units
Buffer	: 1X
Total volume	: 20 µl (by adding MilliQ water)

PCR profile; Initial denaturation; 95°C for 15 min ; 35 cycles of denaturation at 95°C 1 min , annealing at 52°C for 1 min and elongation at 72°C for 1 min ; final elongation at 72°C for 10min.

V. DENATURING PAGE

1. 6% Gel Pouring Solution

Urea	: 45g
Acrylamide: bis-acrylamide (19:1)	: 30 ml
5XTBE	: 20 ml
APS (prepare fresh)	: 0.75 mg/ml
TEMED (v/v)	: 0.044% (add just before use)

Make final volume upto 100 ml by adding distilled water.

VI. 5XTBE

Tris base	: 54g
Boric acid	: 27.5 g
EDTA (pH 8.0)	: 3.72 g

Make final volume upto 1000 ml by adding distilled water

VII. Silver stain (prepare fresh)

Silver Nitrate	: 2 g
Formaldehyde	: 3 ml

Make final volume upto 2000 ml by adding distilled water

VIII. Developer Solution (prepare fresh)

Sodium carbonate	: 30 g
Formaldehyde	: 1.5 ml
Sodium thiosulfate	: 200 µl (add just before use)

Make final volume upto 1000 ml by adding distilled water

IX. Fixer (prepare fresh)

Glacial acetic acid	: 200 ml
Distilled water	: 800 ml

X. 10X Sample loading buffer

Bromophenol blue	: 25 mg
Xylene cyanol	: 25 mg
0.5M EDTA (pH 8.0)	: 200 µl
99% formamide	: 9.8 ml

Mixed PCR product and denaturing dye (1:1) and denatured for 3 min at 95 °C, snap cooled and loaded

XI. SEQUENCES FOR OLIGONUCLEOTIDES

a. ADAPTERS

EcoRI-1	: CTCGTAGACTGCGTACC
EcoRI-2	: AATTGGTACGCAGTCTAC
MSEI-1	: GACGATGAGATCCTGAG
MSEI-2	: TACTCAGGACTCAT

b. PRIMERS

EcoRI	: GACTGCGTACCAATTC
MseI	: GATGAGTCCTGAGTAA
Probe	: GAGAGAGAGAGAGAGAGAGA

APPENDIX- B

DNA isolation

DNA was isolated by using following protocol (Doyle, 1990):

1. About 5g of plant material was homogenized to powder form using liquid nitrogen in pre-chilled mortar and pestle.
2. The homogenized plant material was then added to 25 mL polypropylene centrifuge tube containing 10mL pre-warmed DNA extraction buffer and incubated at 65°C for 1hr with intermittent inversion.
3. The suspension was allowed to cool to room temperature and mixed with an equal volume of chloroform: iso-amyl alcohol (24:1) for 10 min. at room temperature by gentle inversion and later centrifuged at 15,000 rpm for 30min.
4. The upper aqueous phase was separated and transferred to a fresh centrifuge tube with a wide bore pipette tip. To it added 2/3rd volume of iso-propanol alcohol and mixed by gentle inversion. Centrifuged at 10,000 rpm for 15 min.
5. Discarded the supernatant and DNA pellet was washed twice with 70% alcohol
6. Washed DNA pellet was dried overnight at room temperature, later dissolved in TE buffer and stored at -20°C until use.

DNA purification:

Purification of isolated DNA is required to remove contaminants such as RNA, proteins and polysaccharides. RNA contamination is removed by the treatment of RNase A while proteins are removed by phenol: chloroform treatment. Following protocol was used to purify DNA:

1. To 500 µL of DNA sample added 10 µL RNase A (1µg/ µL) and incubated at 37°C for 1 hr.
2. To above solution added equal volume of phenol: chloroform: iso-amyl alcohol (25:24:1) centrifuged at 10,000 rpm for 5 min. Upper aqueous phase was then transferred to a new centrifuge tube and treated with equal volume of chloroform: iso-amyl alcohol (24:1).

3. The separated aqueous phase was transferred to new centrifuge tube and added 1/10th volume of 3M sodium acetate and two volumes of absolute alcohol. The DNA was pelleted by centrifuging at 13,000 rpm for 3 min.
4. DNA pellet was washed twice with 70% alcohol and dried overnight. The dried DNA pellet was dissolved in TE buffer.

APPENDIX- C

JAS	JOD	BAT	HAN	DAB	SRI	KES	GHA	ANP	RAI	PAD	KAR	MAH	LAK	
0.000														JAS
0.105	0.000													JOD
0.004	0.105	0.000												BAT
0.023	0.163	0.026	0.000											HAN
0.020	0.140	0.021	0.010	0.000										DAB
0.011	0.121	0.050	0.035	0.094	0.000									SRI
0.000	0.101	0.011	0.018	0.040	0.000	0.000								KES
0.035	0.142	0.061	0.074	0.080	0.089	0.065	0.000							GHA
0.009	0.056	0.043	0.057	0.017	0.074	0.015	0.000	0.000						ANU
0.023	0.120	0.028	0.012	0.043	0.040	0.007	0.059	0.023	0.000					RAI
0.037	0.154	0.049	0.052	0.083	0.080	0.032	0.015	0.025	0.035	0.000				PAD
0.001	0.023	0.028	0.079	0.069	0.046	0.000	0.013	0.000	0.036	0.000	0.000			KAR
0.017	0.109	0.029	0.045	0.087	0.000	0.000	0.062	0.026	0.048	0.073	0.001	0.000		MAH
0.053	0.178	0.045	0.031	0.079	0.062	0.031	0.040	0.047	0.034	0.042	0.030	0.022	0.000	LAK

Fig 1. Pairwise Population Fst Values