DEVELOPMENT OF GENOMIC MICROSATELLITE MARKERS IN *Commiphora wightii*

Dissertation submitted to the Central University of Punjab

For the award of
Master of Science
In
Biosciences

BY

Gaganpreet Kaur

Supervisor
Dr. Pankaj Bhardwaj

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

July, 2014

CERTIFICATE

I declare that the dissertation entitled “DEVELOPMENT OF GENOMIC MICROSATELLITE MARKERS IN *Commiphora wightii*” 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.

Gaganpreet Kaur

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

Date:

CERTIFICATE

I certify that Gaganpreet Kaur has prepared her dissertation entitled "DEVELOPMENT OF GENOMIC MICROSATELLITE MARKERS IN *Commiphora wightii*", for the award of M.Sc. 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 of genomic microsatellite markers in *Commiphora wightii*

Name of student: Gaganpreet Kaur
Registration number: CUPB/MSc/SBAS/BIO/2012-13/15
Degree for which submitted: Master of Science
Name of supervisor : Dr. Pankaj Bhardwaj
Name of centre: Centre for Biosciences
Name of school: School of Basic and Applied Sciences
Keywords: *Commiphora wightii*, microsatellite markers, repeat regions, conservation and marker development

Commiphora wightii is an important medicinal plant, growing in arid to semi-arid conditions. It is widely distributed in tropical regions of Africa, Madagascar and Asia. In India, the specie is found in South- Western India and parts of Central India. The plant on incision secretes an oleo gum resin, which is used to treat various ailments since ancient times. The oleo gum resins of plant contain guggulsterones which have great medicinal value. The other uses of gum are in perfumery, calicoprinting, dyeing silk and cotton, fumigation and incense for which the plant is getting exploited by unfair tapping methods. This plant has been listed in red data book so conservation of this plant is the need of present. For this it is necessary to know about the genetic diversity and population structure of the plant. Microsatellite markers are markers of choice for such studied because they are reproducible, co-dominant and show high levels of polymorphism. In the present study genomic SSR markers were developed. A total of 22 primer pairs were designed for which, 1913 clones were analysed for fragments containing microsatellite regions and 338 clones were found positive and are selected for sequencing. Sequencing results show a total of 212 fragments with perfect repeats, interrupted repeats and compound repeats (> 5 repeats). These primer pairs can be used in genetic diversity analysis of naturally growing populations of *Commiphora wightii*. They can also be used in MAS, genome selection during gene introgression in plant breeding, genome mapping and gene tagging. This will help in planning the conservation strategies as well as rational utilization of this endangered plant.

Signature of Student Signature of Supervisor

ACKNOWLEDGEMENTS

The first and foremost thanks is 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, Harmeen Kaur, Shruti Choudhary, Sapna Thakur and Richa Mehra. I am also thankful to technical staff for their help and support.

Warmest regards to my family for their blessings.

Gaganpreet Kaur

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF APPENDICES

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.	Hour(s)	hr(s)
11.	Magnesium Chloride	$MgCl_2$
12.	Melting temperature	T_m
13.	Microgram	µg
14.	Microliter	µl
15.	Milligram	mg
16.	Milliliter	ml
17.	Millimolar	mM
18.	Minute	min
19.	Nanogram	ng
20.	Polymerase Chain Reaction	PCR
21.	Potassium Chloride	KCl
22.	Randomly Amplified Polymorphic DNA	RAPD
23.	Restriction Fragment Length Polymorphism	RFLP
24.	Second	sec
25.	Simple Seqeunce Repeats	SSR
26.	Sodium- Saline Citrate	SSC
27.	SSR Identification Tool	SSRIT
28.	Thymine	T

CHAPTER 1

Introduction

Commiphora wightii (Guggul) is a member of the family Burseraceae. It is a slow growing plant with profuse branching and is adapted to grow in arid to semi-arid conditions. The plant has extensive use as natural herb (Haque *et al.*, 2010). This plant secretes oleo gum resins containing a steroidal compound guggulsterone, which is used in various medicines and fixative in perfumery (Gupta *et al*., 1996). Its extracts are used in the treatment of anti-inflammatory, antirheumatic, hypocholesterolemic, hypolipidemic and infertility activities (Samantaray *et al.*, 2010). Now it is well known that the resin contain two bioactive isomers of guggulsterone *i.e.* guggulsterone E and guggulsterone Z which are responsible for its lipid and cholesterol-lowering activities (http://afri.icfre.org). It is extracted from the main stem by an intrusive process of 'tapping' (Jain and Nadgauda, 2013). Due to more demand for drugs, plant is subjected to destructive procedures and is overexploited for its economically and pharmaceutically important gum. IUCN have included this plant in red data list *(http://www.iucnredlist.org/details/31231/0)*. It is very important to conserve the germplasm of this plant to ensure sustained supply of guggulsterone.

To conserve the germplasm of the plant overall understanding of the diversity and population study is required (Bhuyan *et al.*, 2003). Various DNA based molecular marker techniques are available for such kind of study *viz.* 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). Out of these RAPD and AFLP are dominant markers, while RFLP and SSRs are co-dominant markers. Dominant markers are rather less important because one cannot depict the heterozygosity using these markers (Semagn *et al.*, 2006).

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 serve as an important tool to gather fine-scale ecological information (Gupta and Varshney, 2000). Microsatellites are dispersed throughout the most eukaryotic genomes (Hayden and Sharp, 2001). Microsatellite markers will help in better understanding of genetic structure with in population as well as diversity pattern of *Commiphora*

wightii, which in turn can help in selection of plants for breeding programmes for conservation of this plant.

Keeping in view the above mentioned facts, the present study was undertaken following specific objectives:

1. Construction of genomic microsatellite enriched library
2. Development of genomic microsatellite markers for *commiphora wightii*

CHAPTER 2

Review of Literature

Commiphora wightii is a very important medicinal plant in. 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.

2.1 Classification/ Taxonomic rank

Commiphora wightii belongs to family Burseraceae. It is a dicotyledonous plant and its complete classification is given below (http://www.ncbi.nlm.nih.gov).

Kingdom : Plantae
Phylum : Tracheophyta
Class : Magnoliopsida
Order : Spindales
Family : Burseraceae
Genus : *Commiphora*
Species : *wightii*

2.2 Distribution

The genus *Commiphora* has about 165 to 185 species. This genus is widely distributed in tropical regions of Africa, Madagascar and Asia (Jain and Nadgauda, 2013). In Indian subcontinent, *Commiphora wightii* occur in India, Pakistan, and Baluchistan etc. In India, the species is found in South- Western India and parts of Central India including Kerala, Karnataka, Tamilnadu, Andhra Pradesh, Maharashtra, Madhya Pradesh, Gujarat and Rajasthan. *Commiphora wightii* is widely distributed in Gujarat and Rajasthan (http://afri.icfre.org). It is shrubby, grows very slowly and is adapted to arid, semi-arid and poor soil *(http://www.iucnredlist.org/details/31231/0)*.The most suitable habitats for this species are undulating and hilly areas (Dixit and Rao, 2000).

Commiphora wightii (Guggul) is a traditional medicinal plant. Recent scientific advances in understanding the medicinal properties of this plant have increased its demand in international market. The plant is dimorphic in nature, one with bisexual male flowers and the other with female flowers. There are evidences for the occurrence of apomixis, poly-embryony and autonomous endosperm development in *Commiphora wightii*. Heterozygous population is the result of seeds produces by cross pollination (Jain and Nadgauda, 2013). Occurrence and

extent of sexual reproduction in *Commiphora wightii* was detected by histology, controlled pollination, flow cytometry and RAPD analyses. Apomixes is prevalent but sexual reproduction also happen which gives rise to diversity in population of *Commiphora wightii* (Geetha *et al.*, 2013). Two putative sex linked RAPD markers OPN 06_{1280} and OPN 16_{400} were produced by using decamer RAPD markers, which can be used in various breeding programmes (Samantaray *et al.*, 2010).

The gum resin exudates Guggulu form *Commiphoa wightii* is extracted by the method of tapping, in which incision is made in the bark of the plant and the gum exudates are collected. In the scientific world guggul was first introduced by an Indian Medical Researcher, G. V. Satyavati in 1966. According to different studies guggul gum is known to be hypolipidemic, hypocholestrolemic and anti-obesity, astringent and antiseptic, anti-arthritic, anti-microbial, anti-inflammatory and anti – cancerous. The gum can also be used for the treatment of thrombosis, chronic bronchitis, nodulocystic acne, spongy gums, chronic tonsillitis and teeth carries. Perfumery, calicoprinting, dyeing silk and cotton, fumigation and incense are the other uses of gum (Jain and Nadgauda, 2013).

The gum resin guggul contains about 150 compounds of which, Guggulsterone E and Guggulsterone Z are believed to be hypolipidemic and the most important component (Gajbhiye *et al.*, 2011).

Its anti-cancerous property is shown only recently. In a study guggulsterone is shown to inhibit the PI3K/Akt pathway that plays a critical role in head and neck squamous cell carcinoma development. It nullified the effect of smokeless tobacco and nicotine on PI3K/Akt pathway and inhibited the proliferation of head and neck cancer cells (Macha *et al*., 2011).

2.3 Conservation of *Commiphora wightii*

This plant is listed in the red data list of threatened species by IUCN, but now it is becoming endangered (Jain and Nadgauda, 2013). A threatened condition is the stage when exploitation and destruction of specie is more than its natural regeneration and these result in serious decline in no. of individuals in a population (Dixit and Rao, 2000). The fragrant oleo gum resin from *Commiphora wightii* has high value to pharmaceuticals and perfumery, so the plant is being

exploited (Dixit and Rao, 2000). Slow growth, lack of cultivation, poor seed-setting and seed germination rate, unsustainable over-exploitation, excessive and unscientific tapping method and invasion of alien species (Jain and Nadgauda, 2013) are among the possible causes for its decline.

2.3.1 Conservation measures for the specie can be

1. Plantation in suitable habitats and *ex-situ* conservation (multiplication through nursery and tissue culture techniques).
2. Establishment of small reserve area with regulated biotic pressure (*in-situ* conservation).
3. Development of better gum extraction techniques.
4. Local communities should involve in the conservation of species (Dixit and Rao, 2000).

Conservational efforts needed to be done for its local and commercial use in sustainable approach. Its depletion can lead to loss of gene pool and limit the future scope of deriving natural drugs from this plant (Jain and Nadgauda, 2013).

2.4 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. Self-crossing population is less diverse and more differentiated than out-crosser (Charlesworth and Charlesworth, 1987). Knowledge of 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 poses negative impact on genetic diversity and survival of a specie (Leimu *et al.*, 2006). The use of DNA based molecular markers has enormously increased to study genetic diversity with the advent of technology like PCR. These markers are useful in certifying genetic and evolutionary relationships (Booy *et al.*, 2000). There are several techniques for DNA marking such as, RFLP, RAPD, AFLP, SSR and 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 (Semagn *et al.*, 2006).

2.5 Microsatellite markers

Microsatellites are the simple repeats of two or more nucleotide. They constitute the smallest class of such sequences (Semagn *et al.*, 2006). The term microsatellite was first coined by Litt and Lutty in 1989. Microsatellites are also known as simple sequence repeats (SSRs), short tandem repeats (STRs) and simple sequence length polymorphisms (SSLPs) (Semagn *et al.*, 2006). Dinucleotide repeats are the most abundant repeats (Gupta *et al.*, 1996). Microsatellites are present in both the coding and non – coding regions of the genome (Zane *et al.*, 2002). However the distribution of microsatellites is non-random in the genome (Li *et al.*, 2002). SSR markers are more frequent in dicots than in monocots (Wang *et al.*, 2014).

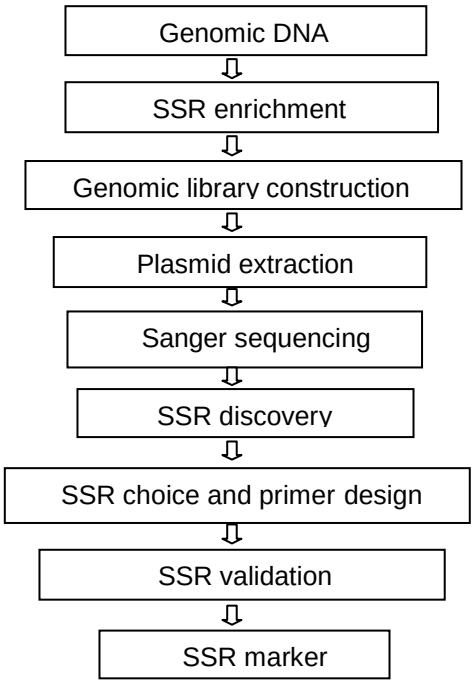
These repeats typically vary in length between 5 and 40, but longer repeat-sequences are also possible (Gupta and Varshney, 2000; Selkoe and Toonen, 2006). Dinucleotide repeats like {$(AC)_n$ and $(GA)_n$} have been worked out most commonly but, tri-nucleotide and tetra-nucleotide repeats like $(AAG)_n$ and $(AAT)_n$ are also present in plant genomes (Gupta and Varshney, 2000). Microsatellites show high levels of polymorphism. Variable number of repeat units within the microsatellite structure results in SSR allelic differences (Semagn *et al.*, 2006). Thus markers are co-dominant and are highly reproducible (Gupta *et al.*, 1996). Many changes of repeat numbers at SSR loci are caused by two mutational mechanisms: first involves DNA slippage during DNA replication and second is the recombination between DNA strands (Li *et al.*, 2002). Microsatellite markers are being used to map genomes, to quantify genetic diversity and to characterize accessions in plant germplasm collections (Gupta *et al.*, 1996). Cross species amplification can be done with microsatellite markers in which microsatellite primers developed for one species can be used to find out polymorphism in related species at homologous sites, provided the repeat sequence and the flanking regions containing the selected primer sites must be conserved across taxa (Wilson *et al.*, 2004).

2.5.1 Applications or uses of microsatellite markers

1. MAS (Marker Assisted Selection in plant breeding)
2. Genome selection during Gene Introgression in plant breeding
3. Genome mapping
4. Gene tagging
5. Cultivar identification, estimation of genetic relatedness and germplasm conservation (Gupta *et al.*, 1996)

2.5.2 Strategies of microsatellite marker development

A conventional strategy to develop microsatellite markers using library screening method 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 (Figure 1) (Glenn and Schable, 2005; Zalapa *et al.*, 2012).



modified (Zalapa *et al.*, 2012)

Figure 1: General scheme for microsatellite marker development

2.5.3 Microsatellite markers for population studies

Microsatellite markers have properties like high polymorphism, reproducibility, co-dominance and availability throughout the genome, owing to which, they have emerged as the markers of choice for population structure analysis (Gupta *et al.*, 1996).

Table 1: List of various studies involving Microsatellite development

Species	**Population Size**	**Fragments sequenced**	**No. of motifs found**	**Total no. of marker developed**	**No. of polym-orphic marker**	**References**
Camellia assamica ssp. assamica	29	1297	549	470	185	(Bhardwaj *et al*., 2013)
Glycine soja	294	NA	NA	36	NA	(Wang *et al*., 2014)
Arctium minus	134	352	NA	42	16	(López-Vinyallonga *et al*., 2010)
Oryza rufipogon	232	NA	NA	23	NA	(Song *et al*., 2003)
Dipteryx alata	94	1013	58	28	8	(Soares *et al*., 2012)
Cicer arietinumL.	60	NA	NA	20	14	(Ghaffari *et al*., 2014)
Taihangia rupestris	40	36	25	18	10	(Wang *et al*., 2010)
Daucus carota L.	NA	NA	NA	300	123	(Cavagnaro *et al*., 2011)
Tapiscia sinensis	102	323	111	36	11	(Zhang *et al*., 2014)
Santiria trimera	4	6611	454	26	7	(Koffi *et al*., 2012)

CHAPTER 3

Materials and Methods

3.1 Plant material and microsatellite enriched library construction

Leaf samples of *Commiphora wightii* were collected from SKN College of Agriculture Jobner, Rajasthan (26.972128°N, 75.380348°E) in October, 2013.

3.1.1 DNA isolation and Purification

Genomic DNA was isolated from young leaves using Haque *et al.*, 2008 protocol with some modifications (Appendix A and B). This is a single step method for isolation and purification of DNA (Haque *et al.*, 2008). The concentration of DNA was measured on NanoDrop 2000 spectrophotometer and purity of DNA was checked on 0.8% agarose gel electrophoresis.

3.1.2 Restriction Digestion and gel elution

Restriction digestion of purified genomic DNA was done using *Eco* R1 and *Mse* 1 enzymes. For this reaction mixture was prepared (Appendix A) and incubated at 37°C for 10 minutes and then heated at 65°C for 10 minutes to stop enzyme activity. The restricted product was checked on 1.2% agarose gel electrophoresis. Fragments of desired range from 200 – 1200 bp were eluted from gel and purified using NucleoSpin Gel and PCR cleanup kit (Mershey - Negel).

3.1.3 Adapter ligation

Adapter ligation was done using following adapters:

*Eco*R1 (1):5’CTCGTAGACTGCGTACC3’,
(2): 5’AATTGGTACGCAGTCTAC3’

*Mse*1 (1): 5’GACGATGAGATCCTGAG3’,
(2):5’TACTCAGGACTCAT3’.

The reaction mixture was prepared (Appendix A) and incubated at 16°C overnight; later enzyme activity was stopped by heating at 65°C for 20 minutes. The product was amplified using PCR (Appendix A) and the results were checked on agarose gel electrophoresis. The PCR product was purified using NucleoSpin Gel and PCR Clean-up kit (Mershey - Negel).

3.1.4 Primary enrichment

A microsatellite $(AG)_n$ enriched genomic library was constructed using Biotin – streptavidin capture method (Bhardwaj *et al.*, 2013) with some modifications. Mixture of biotinylated probe and adapter ligated DNA was incubated for 4 hours at 65°C with 200 rpm. Magnetic streptavidin beads were prepared with 6X SSC buffer and stored at 4°C. After 4 hours DNA probe at 4°C was cooled to room temperature and mixed with magnetic streptavidin beads at room temperature for 20 minutes. After that repeated washing steps were followed in which first washing was done at 65°C for 10 minutes and then at room temperature for 10 minutes with different dilutions of SSC buffer starting from 2X to 0.5X. This increased the specificity of probe hybridizing to desired fragments containing microsatellite repeats. Suspended the product in 100µl of PCR water (milli Q) and subsequently incubated at 95°C for 15 minutes. PCR water containing enriched DNA was collected by snap freezing and magnetic bead separation.

3.1.5 Microsatellite capture PCR/ Amplifying enriched DNA

The enriched product was amplified by PCR (Appendix A) using the primer sequence binding to the adapter. The detail of the reaction mixture is given in appendix A. The amplified products were checked on a 1.5% agarose gel electrophoresis.

3.1.6 Ligation of amplified enriched DNA into vector

Ligation of amplified enriched DNA into vector is done by TA cloning. T*aq* DNA polymerase since it lacks proofreading activity hence adds extra adenine (A) base towards 3’ end, to amplified products. This makes them successful to be cloned into T vectors like pGEM-T (T overhang at 3’ end). The reaction mixture (Appendix A) was made and incubated at 4°C for 16 hours in water bath, stored at -20°C. The goal was to achieve a single amplified enriched DNA (with desired microsatellite region in it) ligated to a single vector and insert a single rDNA into a single host. Ligation at 4°C lowers the free (kinetic) energy of the molecule in the solution allowing ligase to hold the substrates for more time, hence providing efficient ligation.

3.1.7 Transformation

Later these recombinant plasmids were used to transform competent *Escherichia coli*. The competent *Escherichia coli* cells were grown at lower temperature 18°C or 4°C as a result of which they have weak cell membranes, which favours the entry of foreign DNA into the cell. For transformation rDNA fragment (amplified enriched DNA ligated into vector) was to be incubated with culture of competent cells, followed by heat shock treatment for 40-60 seconds that allowed the weak cell wall of bacteria to become fluidic or transient pores were created on the membrane. The pores provided a way into the cell to the r DNA $\leftrightarrow Ca^{2+}$ (in transformation buffer) complex. Again incubated on ice, frozen the cell, closing their pores, transformation was complete. Clones positive for insert were confirmed by using blue-white screening (interrupted β- galactosidase gene), by providing IPTG (inducer) and X- gal (lactose analog) in Luria Bertani agar medium supplemented with ampicillin media in which white colonies were positive clones. Once the cells were transformed, the screening for positive clones was done by secondary enrichment.

3.1.8 Secondary enrichment

Positive clones containing the desired DNA fragment were later selected by secondary enrichment through PCR amplification (Appendix A). Secondary enrichment involves amplification with three primers (3 primers PCR). In addition to the primer corresponding to the ligated adapters of *Eco*R1 and *Mse*1 a third primer for probe $(AG)_{10}$ was also used. This probe primer aids in the selection of only the region containing the desired microsatellite within the enriched fragment. Positive clones were those, which show multiple bands on 1.5% agarose gel electrophoresis.

3.1.9 Plasmid isolation

These positive clones were grown overnight at 37°C in LB broth containing 100µg/ml ampicillin. Plasmid DNA from these clones was isolated using SureSpin Plasmid kit (Mershey - Negel) as per manufacturer's guidelines. The concentration of isolated plasmid DNA was checked on NanoDrop 2000 spectrophotometer.

3.1.10 Chain termination reaction (M13 PCR)

The reaction mixtures were prepared separately for forward M13 primer (5´-CTGGCCGTCGTTTTAC -3´) and reverse M13 primer (5´-CAGGAAACAGCTATGAC - 3´). Plasmid DNA was used according to desired concentration (Appendix A). Ready reaction mixture contained *Taq* polymerase, dNTPs as well as ddNTPs. The reaction needs to be terminated for sequencing after addition of single dNTP, so that each fragment generated successively was one bp longer than the previous one; therefore, it was necessary to maintain the ratio of dNTPs: ddNTPs (3:1).

3.1.11 DNA sequencing

DNA sequencing was done with ABI 3730 Xl DNA Analyzer using BigDye terminator cycle sequencing kit v3.1 (Applied Biosystems) as par manufacturer's procedures. Universal M13 forward and reverse primers are used for sequencing the vector and insert region.

3.2 Primer Designing

The softwares used for the designing of primers were Grammene [SSR identification tool (SSRIT)] to locate the SSR region within the sequence (Temnykh *et al.*, 2001), Gene runner (to trim the sequence), ClustalW (to align the primers, to check redundancy or duplicity; to locate *Eco*/*Mse*/plasmid, adapter/vector sequences) and Primer3 web version 4.0.0 (Untergasser *et al.*, 2012). The steps for designing were: 1) opened the seq file from the ABI sequence analysis (.txt) in notepad/ supported format 2) Checked whether the sequence/ fragment had microsatellite region by copied the sequence in FASTA (supported format) in Gramene (SSRIT). If the sequence had repeat motif (>5 was considered as repeat region), moved to next step, else dropped the sequence and checked for the next sequence 3) In case the sequence contained some repeat regions, checked whether primer can be designed *i.e.* see if the repeat motif was at optimum location shouldn't be at the end/beginning or plasmid/adapters were not too close 4) ClustalW was used to align with adapters (*Eco*R1 and *Mse*1) as well as plasmid (pGEM-T Easy) whose sequence is easily available online 5) Copied the sequences to gene runner (compatible with some windows only), cleaned the

sequence according to requirement 6) The cleaned sequence after that subjected to Primer3 web version 4.0.0, targets were given and the primer was designed.

CHAPTER 4

Results and Discussion

4.1 Genomic DNA isolation and microsatellite enriched library construction

Results and discussion concerning the microsatellite enriched library construction are given below:

4.1.1 DNA isolation

For microsatellite enriched library construction total genomic DNA was extracted from the fresh leaves of plant *Commiphora wightii*. The concentration and purity of isolated DNA was checked on NanoDrop 2000 spectrophotometer (621.3 ng/µl and 1.86 respectively) and also the DNA bands were visualized on 0.8% agarose gel electrophoresis using ethidium bromide staining and the image was taken in gel documentation system (BioRad) (Figure 2).

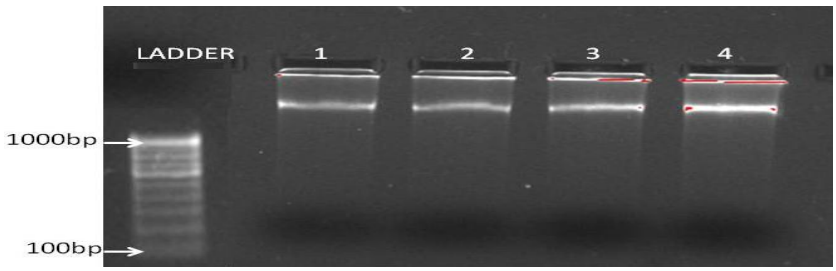


Figure 2: Lane no. 1, 2, 3 and 4 shows DNA

4.1.2 Restriction digestion

The digested product is also visualized in agarose gel electrophoresis. A smear of DNA fragments is there in 1.2% agarose gel. This smear indicated different sized fragments of DNA (Figure 3). The standard DNA marker was also loaded into the gel for size estimation. The DNA fragments from range 200 – 1200 bp were eluted from the gel for further process.

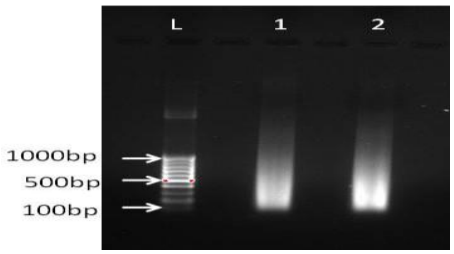


Figure 3: Lane no. L 100bp DNA ladder as size standard (Bangalore Genei^{TM}) shows ladder, lane no. 1 and 2 show restricted DNA

Two kinds of enzymes were used to restrict the genomic DNA (*Eco* R1 and *Mse*1). The selection and number of restriction enzyme(s) is critical to enrichment level, frequency and distribution of SSR- containing clones across the library (Hamilton and Fleischer, 1999; Vos *et al.*, 1995).

4.1.3 Adapter ligation and PCR amplification

Adapters provide the basis for primer binding with which amplification of adapter- ligated DNA fragments was carried out in order to recover losses during further processing. A continuous smear in 1.5% agarose gel indicated the successful ligation of adapters to the restricted fragments of DNA and amplification of those fragments (Figure 4).

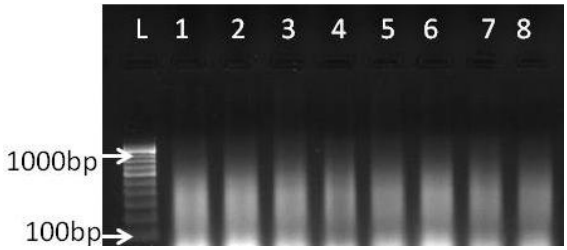


Figure 4: Lane no. L 100bp DNA ladder as size standard (Bangalore Genei^{TM}) shows ladder, lane no. 1 – 8 shows Adapter ligated amplified product

4.1.4 Primary enrichment

The fragments containing repeat regions or microsatellite region are captured using streptavidin coated magnetic beads. Exploiting the higher affinity of biotin for streptavidin, DNA fragments with repeat regions were bounded to biotin labeled probe which were subsequently captured by streptavidin magnetic beads and hence, were pooled out. The enriched product was amplified (Appendix A) and visualized on 1.5% agarose gel (Figure 5).

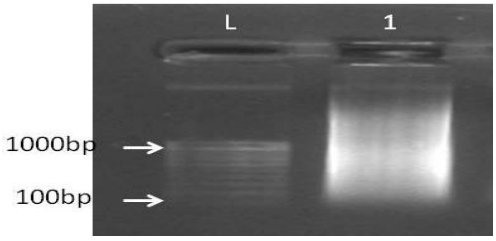


Figure 5: Lane no. L shows ladder and lane no. 1 shows amplified enriched DNA

In view of the high frequency of GA/AG repeats in plant genome (Gupta and Varshney, 2000), $(GA)_n$ oligonuleotide was employed for enrichment to develop novel microsatellites (Du *et al.*, 2013; Gao *et al.*, 2013; Liu *et al.*, 2013). A relatively low background of non-specific DNA can be achieved by affinity capture of clones containing repeat regions (Kandpal *et al.*, 1994).

4.1.5 Vector ligation and transformation

The enriched DNA was then ligated into pGEM-T easy vector and subsequently transformed the competent *E. coli* cells. The transformed cells were then grown on plates with LB agar medium provided with ampicillin, X-gal and IPTG. Two types of colonies appeared on plates, some white coloured and the other blue coloured. White colonies were the positive ones with desired DNA fragment. A total of 1913 colonies were taken and subculturing of those colonies was done for secondary enrichment.

4.1.6 Secondary enrichment

Secondary enrichment results were analyzed on 1.5% agarose gel to identify the positive clones containing desired DNA fragment (with microsatellite region) (Figure 6). In secondary enrichment three primers were used to amplify the fragments, two out of which contained complementary sequences to adapters and one was the probe primer with $(AG)_{10}$ repeat regions. Minimum two or multiple bands on the gel signify the fragments containing the repeat regions or microsatellite regions. The success rate (proportion of positive clones to total clones analyzed) is 17.66% in present study while it was 88.47% in *Lactuca sativa* L. (Rauscher and Simko, 2013), 06.86% in *Santiria* trimera (Koffi *et al*., 2012), 34.36% in *Tapiscia sinensis* (Zhang *et al*., 2014) *etc*.

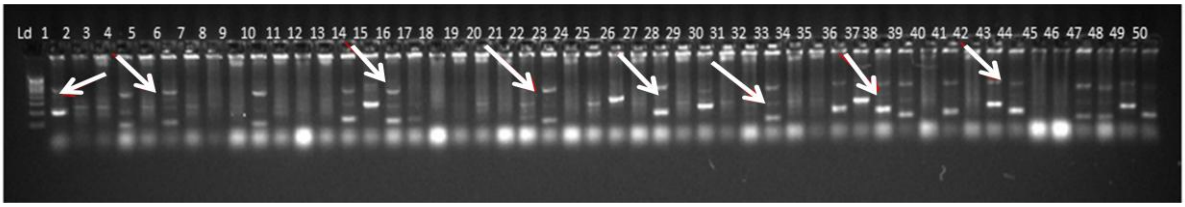


Figure 6: Lane 1-50 amplification profile generated using three-primer PCR; Ld: 100bp DNA ladder as size standard (Bangalore Genei^{TM}); lanes showing multiple bands (indicated by arrow) represent positive clones.

4.1.7 Sequencing and enriched- library assessment

For sequencing a total of 1913 clones were analyzed by secondary enrichment for positive clones (containing microsatellite region) out of which 338 (17.66%) were found positive (Table 2). Plasmid DNA isolation was done for these clones and according to concentration and purity, 338 plasmid chimera were selected for sequencing. Enrichment efficiency (proportion of SSR- clones to the total clones sequenced) was found to be 62.72% which was higher than peanut (Macedo *et al.*, 2012), *Monotrapa hypopytis* (Klooster *et al.,* 2009) and grasspea (Lioi and Galasso, 2013) *etc*. in other studies.

Table 2: Statistic for library enriched for genomic SSR development in *Commiphora wightii*

FACT	TOTAL
Total number of positive colonies obtained after blue- white screening	1913
Number of clones sequenced (percentage of the whole library)	338 (17.66%)
Clones containing SSR loci (percentage of clones sequenced)	212 (62.72%)
Non-suitable SSR clones (percentage of clones sequenced)	190 (89.62%)
Unique SSR clones (percentage of clones sequenced)	22 (06.50%)

Among the 212 sequences containing microsatellite regions, perfect repeats were found in 113 (53.30%) fragments, interrupted repeats in 77 (36.32%) fragments and compound repeats in 22 (10.37%) fragments (Table 3). The repeat regions contained mostly AG/ CT repeats because $(AG)_{10}$ probe was used for enrichment. The length of repeat motif was varied from 5 – 13. Only 10.37% of SSR containing sequences were suitable for primer designing, remaining were rendered unsuitable because of insufficiently long flanking regions or very close proximity of adapters and vector sequences. The success rate thus was less than 50.52% in *Lactuca sativa* L. (Rauscher and Simko, 2013), 13% in *Hippophae rhamnoides* (Wang *et al.*, 2008a), 27% in sugarcane (Cordeiro *et al.*, 1999), 51% in wheat (Pestsova *et al.*, 2000), 66% in sorghum (Bhattramakki *et al.*, 2000) and 90% in sugarcane (Parida *et al.*, 2009).

4.2 Primer designing and evaluation

A total of 22 primer pairs were designed successfully. Four softwares were used for designing the primers. These were Gramene (SSRIT) which was used to locate the microsatellite region in sequence, ClustalW used to align the sequence with adapter and vector sequences. Sequences of plasmid and adapters were cleaned form raw sequence using Gene runner. Primer3 web version 4.0.0 was used to pick primers. The detail of 22 primers pairs is given in table 3. The length of primer varied from 18 – 26bp. Melting temperature of the primers varied between 57.10°C – 60.98°C. The length of repeat motifs (AG/CT) varied from 5 – 13. Expected size of the product ranged between 100 – 186 bp. The number of designed primers is quite good as 26 microsatellite markers have been developed in *Hippophae rhamnoides* (Wang *et al.*, 2008a), 13 microsatellite markers in *Incarvillea mairei* (Ai *et al.*, 2009) and 12 microsatellites loci in *Michelia coriacea* (Zhao *et al.*, 2009).

These designed primers will help to amplify the microsatellite containing regions in the genome of plant and thus can be used to study the genetic diversity and population structure of the plant. They can also be used in MAS, genome selection during gene introgression in plant breeding, gene tagging and fingerprinting (Chagné *et al.*, 2004, Singh *et* al., 2010, Tamilkumar et al., 2009). The markers in turn can also be used to construct genetic maps. These marker dense genetic maps greatly help us to understand the evolutionary processes (Sharopova *et al.*, 2002).

Table 3: List of 22 designed primer pairs

S. No.	Primer name	Primer sequence	Total length (bp)	Repeat motif	Expected size (bp)	Melting temp (°C)
1	CWB1	F 5' GTCCTGCGTACCAATTCACTC3' R5' ATGGACAGTGCAGAGAGAGA3'	21 20	$(CT)_{10}$, $(TC)_{12}$	108	59.00 57.77
2	CWB2	F 5' GAGCCCGTTCGCCCTATG3' R5' TGCGTACCAATTCCTTTCTCG3'	18 21	$(AG)_9$	100	60.28 58.66
3	CWB3	F 5' CATACGAATTCCCGCTGACTG3' R5' ACCCTATGAATGAATACCCCTGA3'	21 23	$(TC)_{11}$	104	59.14 58.42
4	CWB4	F 5' GACTGCGTACCAATTCCACC3' R5' CGGGGACTCGATTGATGAGT3'	20 20	$(CT)_7$, $(CT)_5$	101	58.92 59.25
5	CWB5	F 5' CGCGGGAATTCGATTTCACT3' R5' TGATTACTGCGTACCAATTCACT3'	20 23	$(AG)_{10}$	151	58.99 58.17
6	CWB6	F 5' CGAGATTGTGAATTGGTACGC3' R5' ATTCACAGCGCTCTCTCTCT3'	21 20	$(AG)_5$, $(AG)_{19}$	138	57.70 58.53
7	CWB7	F 5' CGCGGGAATTCGATTGAGAG3' R5' CCAATTCAACATATGGCTTCTGT3'	20 23	$(GA)_5$, $(AG)_{10}$	130	59.14 57.14
8	CWB8	F 5' TCAGGACTCATCAATCGAATTCC3' R5' CGAGGGTTCAAAAGGGTTCG3'	23 20	$(CT)_{10}$	186	58.30 59.13
9	CWB9	F 5' CAATCGAATTCCCGCGGC3' R5' TGATGAGTCCTGAGTAAGTCACA3'	18 23	$(GA)_5$, $(AG)_{10}$	154	59.67 58.59
10	CWB10	F 5' TTTTGGTGGGTGGAATTGGT3' R5' ACCAATCAAATTTCACAGCGC3'	20 21	$(GA)_5$, $(GA)_9$	150	57.82 58.32
11	CWB11	F 5' ACAGCGCTCTCTCTCTCTTT3' R5' TGTTCTTTCTTTTCACCGGGT3'	20 21	$(CT)_{13}$, $(TC)_6$	154	58.46 57.99
12	CWB12	F 5' GTTGTCACGTGCAAAGAGAGT3' R5' GATGACAGCACACCACGTAG3'	21 20	$(GA)_6$	113	59.07 58.65

13	CWB13	F 5’ AGAGAGAGAGAGAGATCAAGAATTGG3’ R5’ TACCAATTCCACCCACCAAA3’	26 20	$(AG)_{10}$	100	60.04 60.98
14	CWB14	F 5’ TGCTGAAGGCATTAGACACG3’ R5’ ACTGAATTGCTAAGCAGTCACTC3’	20 23	$(CT)_{8}$	122	60.01 57.84
15	CWB15	F 5’ GGCCGCGGGAATTCGATT3’ R5’ CCACGTGTAGGCATTTATGGT3’	18 21	$(AG)_{5}$, $(GA)_{5}$	105	60.90 58.36
16	CWB16	F 5’ GATGACAGCACACCACGTAG3’ R5’ ACGTGCACAAGAGAGAGAGA3’	20 20	$(TC)_{6}$	111	58.65 58.39
17	CWB17	F 5’ CGCCATGTAACCACGTGTAG3’ R5’ GGAGGGAGGAGAGGAGAGTG3’	20 20	$(CT)_{10}$	159	60.05 60.34
18	CWB18	F 5’ TCCAGTGATGAGAAACTAACCTG3’ R5’ AGAGAGAGAGAGAGAGCGCTAGA3’	23 23	$(TC)_{10}$	100	58.38 58.57
19	CWB19	F 5’ GAGAGTAACATTGACGATAGAGAGAG3’ R5’ TTTCTTCGCCATTCCTTTTG3’	26 20	$(GA)_{9}$, $(GA)_{6}$	117	57.10 60.18
20	CWB20	F 5’ AAGTGCTAACATTGACTGTAGAGAGA3’ R5’ TTTCTTCGCCATTCCTTTTG3’	26 20	$(GA)_{7}$, $(GA)_{6}$	115	57.97 60.18
21	CWB21	F 5’ ACAATTGGGGATGTTTGTGC3’ R5’ TTGACTGCGTACCAATTCAAAC3’	20 22	$(GA)_{9}$, $(GA)_{5}$	152	60.62 60.04
22	CWB22	F 5’ CCAATTCAAGCAATAAACTTGAG3’ R5’ GGTAGGTGGAATTGGTAAGCA3’	23 21	$(TC)_{11}$	170	58.02 58.96

Summary

Commiphora wightii is growing in arid to semi – arid conditions and is adapted to slow growth. The plant has great medicinal value due to its oleo – gum resins which is known to be hypocholestrolemic and anti-obesity, astringent and antiseptic, anti-arthritic, anti-microbial, anti-inflammatory and anti – cancerous. Guggulsterones are present in its gum, which are believed to be hypolipidemic. The gum is extracted from plant by tapping method which causes harm to plant and hence the plant has become endangered nowadays. It is advised to conserve the germplasm of the plant for which detailed study of its population structure and genetic diversity is required. Several techniques for this kind of studies are there one of which is DNA based molecular markers. SSR markers are the markers of choice due to their properties like co-dominant inheritance, polymorphism and reproducibility, random and wide genomic distribution.

Microsatellites are simple repeats of two or more nucleotides among which dinucleotide repeats are more common. Their length varies from 5 – 40. This variable number of repeat units results in SSR allelic differences. Thus markers are co-dominant and are highly reproducible. Microsatellite markers can help in cross species amplification in which primers designed for one species can be used to check polymorphism in related species at homologous sites, but the repeat sequence and flanking regions containing selected primers must be conserved among taxa.

A total of 22 primer pairs are designed in this study, for which 1913 clones were checked for microsatellite region. Out of these 338 plasmid chimera were sequenced. A total of 212 sequences showed repeat regions containing perfect repeats, interrupted repeats and compound repeats. These primers will be used to amplify the microsatellite regions in the plant and hence to study the population structure and genetic diversity of the plant. Microsatellite markers can also be help in MAS, genome selection during gene introgression in plant breeding, genome mapping and gene tagging. This will help in the conservation of this endangered plant.

References

Ai, H.-L., Wang, H., Li, D.-Z. and Yang, J.-B. (2009). Isolation and characterization of 13 microsatellite loci from *Incarvillea mairei* (Bignoniaceae), an endemic species to the Himalaya-Hengduan mountains region. *Conservation genetics* **10**: 1613-1615.

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.

Bhattramakki, D., Dong, J., Chhabra, A.K. and Hart, G.E. (2000). An integrated SSR and RFLP linkage map of *Sorghum bicolor* (L.) Moench. *Genome* **43**: 988-1002.

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.

Cavagnaro, P.F., Chung, S.-M., Manin, S., Yildiz, M., Ali, A., Alessandro, M.S., Iorizzo, M., Senalik, D.A. and Simon, P.W. (2011). Microsatellite isolation and marker development in carrot-genomic distribution, linkage mapping, genetic diversity analysis and marker transferability across Apiaceae. *BMC genomics* **12**: 386.

Chagné, D., Chaumeil, P., Ramboer, A., Collada, C., Guevara, A., Cervera, M., Vendramin, G., Garcia, V., Frigerio, J. and Echt, C. (2004). Cross-species transferability and mapping of genomic and cDNA SSRs in pines. *Theoretical and Applied Genetics* **109**: 1204-1214.

Cordeiro, G.M., Maguire, T.L., Edwards, K.J. and Henry, R.J. (1999). Optimisation of a microsatellite enrichment technique in *Saccharum spp*. *Plant Molecular Biology Reporter* **17**: 225-229.

Dixit, A. and Rao, S.S. (2000). Observation on distribution and habitat characteristics of Gugal (*Commiphora wightii*) in the arid region of Kachchh, Gujarat, India. *Tropical Ecology* **41**: 81-88.

Du, F.K., Xu, F., Qu, H., Feng, S., Tang, J. and Wu, R. (2013). Exploiting the transcriptome of Euphrates Poplar, *Populus euphratica* (Salicaceae) to develop and characterize new EST-SSR markers and construct an EST-SSR database. *PLoS One* **8**: e61337.

Gajbhiye, N.A., S. Makasana, J. and Geetha, K.A. (2011). Chemical screening of guggul (*Commiphora wightii*) accessions collected from different natural habitats of Gujarat. *Open Access Journal of Medicinal and Aromatic Plants, North America* **1**: TS2-P31 (Abstr.).

Gao, Z., Wu, J., Wang, L., Ren, H. and Shu, Q. (2013). Rapid microsatellite development for tree peony and its implications. *BMC genomics* **14**: 886.

Geetha, K.A., Kawane, A., Bishoyi, A., Phurailatpam, A., Ankita, C., Malik, S.K., Srinivasan, R. and Bhat, S.R. (2013). Characterization of mode of reproduction in *Commiphora wightii* [(Arnot) Bhandari] reveals novel pollen–pistil interaction and occurrence of obligate sexual female plants. *Trees* **27**: 567-581.

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* **20**: 225-233.

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

Gupta, P., Balyan, H., Sharma, P. and Ramesh, B. (1996). Microsatellites in plants: a new class of molecular markers. *Current Science* **45**: 45-54.

Gupta, P.K. 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.

Gupta, S., Kumari, K., Muthamilarasan, M., Subramanian, A. and Prasad, M. (2013). Development and utilization of novel SSRs in foxtail millet [*Setaria italica* (L.) P. Beauv.]. *Plant Breeding* **132**: 367-374.

Hamilton, M. and Fleischer, R. (1999). Brief communication. Cloned microsatellite repeats differ between 4-base restriction endonucleases. *Journal of Heredity* **90**: 561-563.

Haque, I., Bandopadhyay, R. and Mukhopadhyay, K. (2008). An optimized protocol for fast genomic DNA isolation from high secondary metabolites and gum containing plants. *Asian J Plant Sci* **7**: 304-308.

Haque, I., Bandopadhyay, R. and Mukhopadhyay, K. (2010). Population genetic structure of the endangered and endemic medicinal plant *Commiphora wightii*. *Molecular biology reports* **37**: 847-854.

Hayden, M.J. and Sharp, P. (2001). Targeted development of informative microsatellite (SSR) markers. *Nucleic acids research* **29**: e44-e44.

Jain, N. and Nadgauda, R.S. (2013). *Commiphora wightii* (Arnott) Bhandari—A Natural Source of Guggulsterone: Facing a High Risk of Extinction in Its Natural Habitat. *American Journal of Plant Sciences* **4**: 57.

Kandpal, R.P., Kandpal, G. and Weissman, S.M. (1994). Construction of libraries enriched for sequence repeats and jumping clones, and hybridization selection for region-specific markers. *Proceedings of the National Academy of Sciences* **91**: 88-92.

Klooster, M., Hoenle, A. and Culley, T. (2009). Characterization of microsatellite loci in the myco-heterotrophic plant *Monotropa hypopitys* (Ericaceae) and amplification in related taxa. *Molecular ecology resources* **9**: 219-221.

Koffi, K.G., Heuertz, M., Jans, R., Hardy, O.J., Vendramin, G.G. and Duminil, J. (2012). Characterization of new microsatellite loci isolated from *Santiria trimera* (Burseraceae). *American Journal of Botany* **99**: e334-e336.

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.

Li, Y.C., Korol, A.B., Fahima, T., Beiles, A. and Nevo, E. (2002). Microsatellites: genomic distribution, putative functions and mutational mechanisms: a review. *Molecular ecology* **11**: 2453-2465.

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.

Liu, S.-R., Li, W.-Y., Long, D., Hu, C.-G. and Zhang, J.-Z. (2013). Development and Characterization of Genomic and Expressed SSRs in Citrus by Genome-Wide Analysis. *PLoS One* **8**: e75149.

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.

Macha, M.A., Matta, A., Chauhan, S.S., Siu, K.W.M. and Ralhan, R. (2011). Guggulsterone Targets Smokeless Tobacco Induced PI3K/Akt Pathway in Head and Neck Cancer Cells. *PLoS ONE* **6**: e14728.

Macedo, S.E., Moretzsohn, M.C., Leal-Bertioli, S.C., Alves, D.M., Gouvea, E.G., Azevedo, V.C. and Bertioli, D.J. (2012). Development and characterization of highly polymorphic long TC repeat microsatellite markers for genetic analysis of peanut. *BMC research notes* **5**: 86.

Parida, S.K., Kalia, S.K., Kaul, S., Dalal, V., Hemaprabha, G., Selvi, A., Pandit, A., Singh, A., Gaikwad, K. and Sharma, T.R. (2009). Informative genomic microsatellite markers for efficient genotyping applications in sugarcane. *Theoretical and Applied Genetics* **118**: 327-338.

Pestsova, E., Ganal, M. and Röder, M. (2000). Isolation and mapping of microsatellite markers specific for the D genome of bread wheat. *Genome* **43**: 689-697.

Rauscher, G. and Simko, I. (2013). Development of genomic SSR markers for fingerprinting lettuce (*Lactuca sativa* L.) cultivars and mapping genes. *BMC plant biology* **13**: 11.

Samantaray, S., Geetha, K., Hidayath, K. and Maiti, S. (2010). Identification of RAPD markers linked to sex determination in guggal [*Commiphora wightii* (Arnott.)] Bhandari. *Plant Biotechnology Reports* **4**: 95-99.

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.

Semagn, K., Bjørnstad, Å. and Ndjiondjop, M. (2006). An overview of molecular marker methods for plants. *African Journal of Biotechnology* **5**.

Sharopova, N., McMullen, M.D., Schultz, L., Schroeder, S., Sanchez-Villeda, H., Gardiner, J., Bergstrom, D., Houchins, K., Melia-Hancock, S. and Musket, T. (2002). Development and mapping of SSR markers for maize. *Plant molecular biology* **48**: 463-481.

Singh, H., Deshmukh, R.K., Singh, A., Singh, A.K., Gaikwad, K., Sharma, T.R., Mohapatra, T. and Singh, N.K. (2010). Highly variable SSR markers suitable for rice genotyping using agarose gels. *Molecular breeding* **25**: 359-364.

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.

Song, Z., Xu, X., Wang, B., Chen, J. and Lu, B.-R. (2003). Genetic diversity in the northernmost *Oryza rufipogon* populations estimated by SSR markers. *Theoretical and Applied Genetics* **107**: 1492-1499.

Tamilkumar, P., Jerlin, R., Senthil, N., Ganesan, K., Jeevan, R. and Raveendran, M. (2009). Fingerprinting of rice hybrids and their parental lines using microsatellite markers and their utilization in genetic purity assessment of hybrid rice. *Research Journal of Seed Science* **2**: 40-47.

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.

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.

Vos, P., Hogers, R., Bleeker, M., Reijans, M., van De Lee, T., Hornes, M., Friters, A., Pot, J., Paleman, J. and Kuiper, M. (1995). AFLP: a new technique for DNA fingerprinting. *Nucleic acids research* **23**: 4407-4414.

Wang, A., Zhang, Q., Wan, D. and Liu, J. (2008). Nine microsatellite DNA primers for *Hippophae rhamnoides ssp. sinensis* (Elaeagnaceae). *Conservation genetics* **9**: 969-971.

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.

Wang, Z., Liu, H., Liu, J., Li, Y., Wu, R. and Pang, X. (2014). Mining new microsatellite markers for Siberian apricot (*Prunus sibirica* L.) from SSR-enriched genomic library. *Scientia Horticulturae* **166**: 65-69.

Wilson, A.C., Massonnet, B., Simon, J.C., Prunier-Leterme, N., Dolatti, L., Llewellyn, K.S., Figueroa, C.C., Ramirez, C.C., Blackman, R.L. and Estoup, A. (2004). Cross-species amplification of microsatellite loci in aphids: assessment and application. *Molecular Ecology Notes* **4**: 104-109.

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.

Zane, L., Bargelloni, L. and Patarnello, T. (2002). Strategies for microsatellite isolation: a review. *Molecular ecology* **11**: 1-16.

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**.

Zhao, X.-F., Sun, W.-B., Yang, J.-B. and Meng, J. (2009). Isolation and characterization of 12 microsatellite loci for *Michelia coriacea* (Magnoliaceae), a critically endangered endemic to Southeast Yunnan, China. *Conservation genetics* **10**: 1583-1585

Appendices

APPENDIX A

COMPOSITION OF CHEMICALS

I. DNA EXTRACTION

DNA EXTRACTION BUFFER

Reagent/materials	Concentration/ pH	Quantity
TrisCl	1 M / 8.0	1.5 ml
Na.EDTA	0.5 M / 8.0	0.5 ml
NaCl	5 M	3.0 ml
β- mercaptoethanol		0.03 ml
CTAB	20%	1.75 ml
PVP	10%	3.0 ml
Autoclaved dist. Water		0.22 ml
Total		10.0 ml

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 DIGESTION

Material	Concentration	Quantity
Genomic DNA	5 µg	8.0µl
*Eco*RI (NEB)	10 units/ µg of DNA	5.0 µl
*Mse*I (NEB)	10 units/ µg of DNA	5.0 µl
Cut Smart Buffer(NEB)	1X	10.0 µl
Milli Q water		72.0 µl
Total		100.0 µl

Incubated at 37°C for 10-15 min. stopped the enzymatic activity by incubating at 65°C for 10 min.

III. ADAPTER LIGATION

Materials	Concentrations	Quantity
Restricted DNA		40.0 µl
*Eco*RI adapters	20 pmol	2.0 µl
*Mse*I adapters	20 pmol	2.0 µl
Ligase	10 units/ µg of DNA	0.5 µl
Buffer	1X	5.0 µl
Milli Q (H_2O)		0.5 µl
Total		50.0 µl

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

IV. PCR AMPLIFICATION OF LIGATED PRODUCT

Materials	Quantity
Ligated DNA	10.0 µl
Eco primer	1.5 µl
Mse primer	1.5 µl
dNTPs	3.0 µl
$MgCl_2$	3.0 µl
Taq pol	2.0 µl
Buffer	3.0 µl
Milli Q (H_2O)	6.0 µl
Total	30.0 µl

Thermocycling conditions for adapter ligated DNA

Steps	Temperature (°C)	Time (minutes)
Initial denaturation	94	3

Denaturation	94	1
Primer anneling	55	1
Primer extension	72	2
	Step 2 – 4 repeated 20 times	
Final extension	72	10

V. AMPIFYING ENRICHED DNA

Materials	Quantity
Enriched DNA	10.0 µl
Eco primer	1.0 µl
Mse primer	1.0 µl
dNTPs	5.0 µl
$MgCl_2$	3.0 µl
Taq pol	1.5 µl
Buffer (1X)	3.0 µl
Milli Q (H_2O)	5.5 µl
Total	30.0 µl

Thermocycling conditions for enriched DNA

Steps	Temperature (°C)	Time (minutes)
Initial denaturation	94	3
Denaturation	94	1
Primer anneling	55	1
Primer extension	72	2
	Step 2 – 4 repeated 20 times	
Final extension	72	10

VI. VECTOR LIGATION

Materials	Concentrations	Quantity
Amplified enriched DNA		3.0 µl
Vector	20 pmol	1.0 µl
Ligase	10 units/ µg of DNA	1.0 µl

Buffer	1X	5.0 µl
Milli Q (H_2O)		1.0 µl
Total		11.0 µl

Incubated at 4°C for 16 hrs in water bath, after that stored at -20°C.

VII. SECONDARY ENRICHMENT

Materials	Quantity
Buffer	2.0 µl
Eco primer	1.0 µl
Mse primer	1.0 µl
Probe primer	1.0 µl
dNTPs	2.0 µl
$MgCl_2$	1.0 µl
Taq pol	1.0 µl
Colony suspension	2.0 µl
Milli Q (H_2O)	9.0 µl
Total	20.0 µl

Thermocycling conditions for Secondary enrichment

Steps	Temperature (°C)	Time (minutes)
Initial denaturation	95	5
Denaturation	95	1
Primer anneling	52	1
Primer extension	72	1
	Step 2 – 4 repeated 35 times	
Final extension	72	10

VIII. CHAIN TERMINATION REACTION

Material	Quantity
Buffer	2.0µl
Ready reaction mixture	1.0 µl

Primer M13 (F/R)	1.0 µl
DNA (according to conc.)	6.0 µl/ 3.0 µl/ 2.5 µl
Milli Q water	0.0 µl/ 3.0 µl/ 3.5 µl
Total	10.0 µl

Thermocycling conditions for Chain termination reaction

Steps	Temperature (°C)	Time
Initial denaturation	96	5 minutes
Denaturation	96	30 seconds
Primer anneling	52	40 seconds
Primer extension	60	4 minutes
Step 2 – 4 repeated 45 times		
Hold	4	Infinite

IX. SEQUENCES FOR OLIGONUCLEOTIDES

a. ADAPTERS

Eco RI – 1 : CTCGTAGACTGCGTACC

Eco RI - 2 : AATTGGTACGCAGTCTAC

Mse – 1 : GACGATGAGATCCTGAG

Mse – 2 : TACTCAGGACTCAT

b. PRIMERS

Eco RI : GACTGCGTACCAATTC

*Mse*I : GATGAGTCCTGAGTAA

Probe : GAGAGAGAGAGAGAGAGAGA

APPENDIX- B

DNA isolation and purification (1 step method)

DNA was isolated by using following protocol (Haque *et al.*, 2008):

1. De-veinded 0.4 g of leaves and homogenized the leaves using liquid nitrogen. Put powder in 3 ml of pre-warmed DEB (65°C). Dispended 750 µl into 1.5 ml eppendorf. Added 5 µl RNase A to each tube.
2. Incubated for 45 min. at 65°C. Cooled to room temperature. Added equal volume of C: I (24:1) and inverted gently to mix.
3. Centrifuged for 10 min. at 13,000g at room temperature (RT). Separated aqueous phase and added 2/3rd (0.6) volume of isopropanol. Mixed and incubated at room temperature for 20 – 30 min. Centrifuged at 13,000g for 15 min.
4. Washed the pallet with 750 µl of 70% ethanol for 5 minutes only. Air dried the pallets for 30 min. or in desiccator for only 10 min.
5. Resuspended in TE buffer (30 µl) and allowed to dissolve completely. Stored at -20°C until use.
6. Concentration and purity of DNA was checked on NanoDrop 2000 spectrophotometer.