Development and Characterization of UGMS markers for Genetic Diversity Analysis in *Rhododendron arboreum.*

A Dissertation Submitted to the Central University of Punjab

For the award of

Master of Philosophy

In

Biosciences

By

Gagan Sharma

Administrative Guide: Prof.R.G.Saini
Dissertation Coordinator: Dr.Pankaj Bhardwaj

Centre for Biosciences
School of Basic and Applied Sciences
Central University of Punjab, Bathinda
November, 2013

CERTIFICATE

I declare that the dissertation entitled “Development and Characterization of UGMS markers for Genetic Diversity Analysis in Rhododendron arboreum” 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.

Gagan Sharma

Registration Number: CUP/MPH-PhD/SBAS/BIO/2011-12/05

Centre for Biosciences,

School of Basic and Applied Sciences,

Central University of Punjab,

Bathinda – 151001.

Date:

CERTIFICATE

I certify that Gagan Sharma has prepared his dissertation entitled “Development and Characterization of UGMS markers for Genetic Diversity Analysis in Rhododendron arboreum”, for the award of M.Phil. Degree of the Central University of Punjab, under my guidance. He has carried out this work in Molecular Genetics Lab. at the Centre for Biosciences, School of Basic and Applied Sciences, Central University of Punjab.

Dr. Pankaj Bhardwaj
Assistant Professor
Centre for Biosciences,
School of Basic and Applied Sciences,
Central University of Punjab,
Bathinda - 151001.

Date:

Prof.R.G.Saini

COC and Professor Incharge (Examination)

Centre for Biosciences,

School of Basic and Applied Sciences,

Central University of Punjab,

Bathinda-151001.

Date:

ABSTRACT

Development and Characterization of UGMS markers for Genetic Diversity Analysis in *Rhododendron arboreum.*

Name of student:	Gagan Sharma
Registration Number:	CUP/MPH-PhD/SBAS/BIO/2011-12/05
Degree for which submitted:	Master of Philosophy
Administrative Guide:	Prof.R.G.Saini
Dissertation Coordinator:	Dr.Pankaj Bhardwaj
Centre:	Centre for Biosciences
School of Studies:	School of Basic and Applied Sciences
Key words:	Rhododendron arboreum, Genetic diversity, UGMS, Dendrogram, EST-SSRs.

Rhododendron arboreum is an ecologically important species growing in Himalayan regions. It is widely popular due to its medicinal potential, edible, sacred & economic value, however very little is known about its genetic diversity. The development and characterization of UGMS markers will provide a major impact on genetic diversity analysis. Total 792 (615 singletons and 177 contigs) unigenes were predicted from 1,241 publicly available EST database from NCBI (http://www.ncbi.nlm.nih.gov/) site in R.catawbiense by clustering of 2 – 33 random EST sequences. From 50 SSR contained Unigenes, 36 UGMS primers were designed for the cross amplification approach and six pairs of microsatellite primers were produced clear PCR amplification. The cross related species/genera transferability rate was 16.66% from R.catawbiense to R.arboreum. The polymorphism information content (PIC) ranged from 0.2756 to 0.9212 with an average of 0.5765.The average observed heterozygosity (H_o) was 0.8666 i.e. greater than average expected heterozygosity (H_e) was 0.6792. The genetic diversity was found with an average of 0.6222 within population of R.arboreum and genetic diversity found with an average of 0.3436 among the population.This study revealed the insight of abundance & distribution of microsatellite in the expressed region of the Rhododendron arboreum genome.

(Gagan Sharma) (Dr.Pankaj Bhardwaj) (Prof.R.G.Saini)

ACKNOWLEDGEMENT

This dissertation is a milestone in my academic career. I am grateful to a number of people who have guided and supported me throughout my research work and provided assistance me from time to time. I would first like to thank my supervisor; Dr. Pankaj Baradwaj who guided me in selecting the final theme for this research. This dissertation would not have been possible without the unwavering support, encouragement and patience of him and for the help and advice given to me at right time and in right direction.

My sincere thanks to Prof. (Dr.) R.G. Saini, Examination Incharge and Centre coordinator (COC) of Centre for Biosciences. I am also extremely thankful to Prof. (Dr.) P. Ramarao, Dean, Academic Affairs and Prof. (Dr.) Jairup Singh, Vice Chancellor, Central University of Punjab for providing me research facilities to carry out my work.

I would like to thank Dr. Felix Bast, Dr. Sanjeev Thakur, Dr. Anil Kumar Mantha, Dr. Monisha Dhiman, Dr. Sandeep Singh and Dr. Preeti Khatarpal for their guidance and help during my experimental work. I am also thankful to Gagan Dhaliwal, Ashwani Sharma, Sona Rani, Gopal Soni, Bhadur Singh and Mandeep Singh to help & assist in lab during experimental & computer work.

I am extremely thankful to all my neighbor Rajnath Bhaiya and Best friends Gurvinder Singh, Munish Kumar, India Singh Sidhu, Anil Kumar, Navgeet Paul, Kamal Ahmed, Aijaz Ahmad Bhatt, Satej Bhushan, Balraj, V.J.Jithin, Kamal Preet Dhillon, Parminder Singh Sidhu, Shweta Thakur and Jatinder for their cooperation and support during M.phil, I am also thankful to all my seniors for their support and help.

I am extremely thankful to my Best Friend and Roommate Mr.Gurvinder Singh Sidhu, some Himachali friends Pawan, Babu, etc. and my Senior Shweta Thakur for their cooperation and support during the Sample collection process. Finally, my family has supported and helped me along the course of this dissertation by giving encouragement and providing the moral and emotional support I needed to complete my dissertation. To them, I am eternally grateful.

Gagan Sharma

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF APPENDICES

LIST OF ABBREVIATIONS

S.No	Full form	Abbreviation
1	UniGenes derived Microsatellite Markers	UGMS
2	Rhododendron Unigene derived Microsatellite Markers	RUGMS
3	Coefficient of Similarity	CS
4	Polymorphic Information Content	PIC
5	American Rhododendron Society	ARS
6	Toronto Region Rhododendron & Horticulture Society	TRR & HS
7	Expressed Sequence Taqs	ESTs
8	Simple Sequence Repeats	SSRs
9	Expressed Sequence Taq-Simple Sequence Repeats	EST-SSRs
10	Deoxyribonucleic acid	DNA
11	Polymerase Chain Reaction	PCR
12	Unweighted Pair Group Method Analysis	UPGMA
13	Effective number of alleles	Ne
14	National Centre for Biotechnology Information	NCBI
15	Inter-Simple Sequence Repeat	ISSR
16	base pair	bp
17	nanogram	ng
18	Expected Heterozygosity	H_e
19	Observed Heterozygosity	H_o

Chapter-I

INTRODUCTION

The genus Rhododendron belongs to family Ericaceae. The word Rhododendron is derived from two Greek words rhodon (rose) & Dendron (tree) meaning Rose tree (Hora, 1981). The genus with attractive & beautiful flowers is represented by 1000 species in the world. They are mostly distributed at higher elevations in the Sino-Himalayan region with maximum concentrations in Western China. In India, the species are mostly confined to the Himalayan region particularly in eastern Himalaya. It is very popular species in whole world, therefore for its improvement many countries have run Rhododendron horticultural societies, such as American Rhododendron Society (ARS), Toronto Region Rhododendron and Horticultural Society (TRR&HS), The Dutch Rhododendron Society, Fraser South Rhododendron Society, German Rhododendron Society, etc. *Rhododendrons* are mostly used by the local inhabitants of Indo-Himalayan regions due to its aesthetic & sacred values with some medicinal values also. It's anti-inflammatory & hepatoprotective activities are due to some neutraceuticals compounds such as Flavonoids, Saponins & Phyto-phenols present in it. It is used as an herb & treats to lack of appetite, coughing, and various skin disorders (Pradhan & Lachungpa, 1990). Some species are used to treat arthritis, Heart ailments & anti-tubercular properties (Chopra et al., 1958).

1.1 Ecological distribution

It occurs throughout the moist areas of northern hemisphere & into the southern hemisphere in southeastern Asia & northern Australia. The highest diversity is found in the Himalaya from Uttarakhand, Nepal, Sikkim to Yunnan & Sichuan & mountains of Indo-china, Korea, Japan & Taiwan (Nayar, M.P & Shastry, A., 1987). The genus Rhododendron, having about 50 species in India, is mainly distributed in the Himalayan region. With the shrinking of green cover almost everywhere, the Rhododendrons are also experiencing the impact of disturbed ecological systems. This is clearly visible in the Himalaya, where the ecological systems and land physiography are understood to be fragile and found to be easily disturbed. In the context of Himalaya, though heavy degradation is not evident, it surely shows certain signs of the menace entering the region.

The rise in population with demand on land for farming, increased animal husbandry practices, construction of roadways, hydel power stations, allied works, & lately the tourist influx have collectively resulted in the building up of considerable pressure on the availability of Rhododendron species. The major threats to Rhododendrons are deforestation and unsustainable extraction for firewood and incense by local people. Due to the presence of polyphenols and flavonoids in Rhododendrons make it excellent firewood that burns even under wet conditions (Sastry & Hajra, 2010).

1.2 Conservation & Biodiversity

The Indian Himalayan Region (IHR) occupies a special place in the mountain ecosystem of the world. The IHR is one of the most fragile mountain regions of the world & holds an enormous repository of biological diversity, which is increasingly under pressure from the man activities (Shekar, K.C., 2010).This region comprises a rich variety of flora, fauna & human communities estimated 8000 species of vascular plants in the Himalayan region around 3160 are endemic & 450 species are endangered. The rapid indiscriminate exploitations, destruction of habitats, spread of chemicals & introduction of alien species, a number of plants have been disappeared while other awaits a similar fate. Global efforts are being made to conserve the phytodiversity especially of rare, endangered & threatened species (Rodger & Panwar, 1998), which are known to be important component of biodiversity. Due to human interference the natural population of Rhododendrons in the entire Himalayan region is gradually diminishing. The major threats to Rhododendrons are deforestations & unsustainable extraction for firewood incense by local people. A set of Rhododendrons which are classified as rare/endangered may be wiped out from the biota in the near future, if proper conservation measures are not carried out.

1.3 Genetic markers and Biodiversity

Genetic markers are essential tools for the quick detection & characterization of genetic variation in the population. There are many types of genetic markers available for use in ecology & evolution; it is first suitable to reflect on the various applications of genetic markers and the desired attributes for such applications. Some of the more important applications of genetic markers include: (1) Describing mating systems, levels of inbreeding, and chronological and spatial patterns of genetic variation within stands. (2) Describing geographic patterns of genetic variation (3) inferring taxonomic and phylogenetic relationships among species (4) Fingerprinting and germplasm identification in breeding and propagation populations and (6) constructing genetic linkage maps. In present study the use of EST-SSR marker were developed for the analysis of genetic diversity in Rhododendron arboreum species. EST-SSRs are derived from the transcripts,where in flanking regions are expected to be more conserved and show homology with the related genera/species. Thus,they tend to show higher level of transferability in cross amplification and possibility of the markers (Varshney et al.,2005). EST-SSRs are expected to be more transferable to closely related genera since they are anchored within more conserved transcribed regions compared to genomic SSRs (Cordeiro et al.,2001). However polymorphism level of EST-SSR markers has beeen lower than geneomic SSRs (Barbara et al.,2007) but EST-SSRs also have a higher probability of being functionally associated with differences in gene expression than genomic SSRs (Guo et al.,1996).

1.3.1 Simple Sequence Repeat Markers

Simple sequence repeat (SSR, Short, tandemly-repeated sequences of two, three or four nucleotides) markers were first developed for use in genetic mapping in humans (Litt and Luty, 1989). SSRs are found throughout the genome and tend to be amongst the most polymorphic genetic marker among all the markers (Morgante and Olivieri, 1993). SSRs provide considerable polymorphism due to the variation in the number of

repeating units and have good genome coverage. They are codominantly inherited, easily reproducible, simple, PCR-based, extremely polymorphic and highly informative because of the number and frequency of detected alleles. SSR markers have the ability to distinguish between closely related individuals. However, they are expensive to develop, require sequence information and specific primers that usually do not work in other species (Morgante and Olivieri, 1993). SSRs are amplified by PCR using primers that complement the flanking regions (usually conserved), and they are co-dominantly inherited and highly polymorphic. The major constraint of using SSRs as molecular markers is the cost and effort required for their development (Gong et al., 2010). A more widespread use of SSRs in plants would be facilitated if such loci were transferable across species. The emphasis on SSRs has lately shifted towards the use of EST-SSRs, which are relatively inexpensive compared with the development of genomic SSRs (Chai et al., 2013). Nevertheless after collection of DNA sequence data which include expressed sequence tags (EST), this marker system has become more efficient for genomic analyses (Varshney et al., 2005). It is estimated that 2-5 % of all plant derived ESTs have SSRs (Kantety et al., 2002). Also, one of the most important benefits of using EST based SSRs markers are transferability of these markers among species, because they are from more conserved regions of the genome (Ellis et al., 2007).

1.4 Statement of the Problem

The genetics of the ecologically important species have been of great interest for both evolutionary biologists and conservation managers from long time (Claudia,C., 2008) Analysis of the genetic structure is necessary not only to fully evaluate the impact of the endangered status on genetic variation of the population, but also because knowledge on the genetic structure of the species can be applied to the preservation of the evolutionary potential of species, this is one of the conservation goals (Hamrick and Godt, 1996). Thus, molecular tools can play a valuable role for investigating the pattern of genetic diversity in these species, and clarifying demographic and ecological issues early in species management in order to plan long-term conservation or restoration programs (Lande.R., 1988). There is a need to

develop the EST-SSR markers related to Rhododendron arboreum to conserve such type of great biodiversity from human interference in their ecological niche and to understand its present status.

1.5 Research Objectives

- To collect the sample & targeting various population of Rhododendron arboreum from western Himalayan region.
- To study & analysis of various target population of western Himalayan region for their Genetic Diversity by UGMS markers.

Chapter-II

REVIEW OF LITERATURE

Biodiversity refers to variation found within the living world, while genetic diversity represents the heritable variation at the gene level of the species within & between populations. The extent & distrbutiion of genetic diversity in a plant species depends on its evolution, breeding system, ecological, geographical factors, past bottlenecks & by many types of human factors also (Caser, Akkak & Scariot, 2010). Genetic diversity analysis study will increase our knowledge & help us to determine what to conserve as well as where to conserve, how will improve the particular variety, origin & evolution of the species of interest.

2.1 Rhododendron arboreum and its Conservation

Rhododendron arboreum is one of the most stately and impressive Rhododendron species. It is extremely variable in stature, hardness, and flower color and leaf characteristics. It species name arboreum means tree like (Orwa et al., 2009). Conservation of biotic resources remains one of the basic needs in terms of preserving biotic diversity which, in time, translates into the richness of biotic wealth. Plant hunting expeditions by the earlier explorers from Europe into interior Asia (Delmas et al., 2011) since the turn of the century had set the trend for removal of Rhododendrons in large scale from the region (Singh, K., 2009). Therefore, it is a need of an hour to conserve this precious species in Himalayan regions of India. The work done on Rhododendron species in IHR is mainly based on extensive & deeply literature review made in different states of Himalayan regions for exploring the threat categories. Which have taken review from various published scientific papers (Javed et al., 1995a), Redlist documents (Nayar & Shastry, 1987), IUCN list (Rodger & Panwar, 1998), etc. For enumerations of species, the IHR is divided into two botanical regions namely the Western & Eastern Himalaya (Singh & Gurung, 2009). The western Himalayan region is the states of J&K, Himachal Pradesh & Uttaranchal. Similarly the Eastern Himalayan regions include (Arunachal Pradesh, Assam, Manipur, Meghalaya, Mizoram, Nagaland & Tripura). Sikkim & Darjeeling district of West Bengal also. We are taking the Western Himalayan region for the study of

genetic diversity analysis via UGMS markers. This can provide some proof for the classification & genetic diversity of Rhododendron & explore more about its phylogeny & its adaptations so as to protect & utilize Rhododendron resources efficiently.

Table 2.1: List of Research work done on Rhododendron arboreum

S.No	Topic of the Article/Journal	Results or Findings	References
1	Clonal Stability of RAPD Markers in Three Rhododendron . Species	Amplification profiles via RAPD to reveal genetic difference between three species of Rhododendron and analyzed that clonaly propagated plants showed less polymorphism than the other individual.	Javed et al. (1995a)
2	Notes on the Sikkim Himalayan Rhododendron : a taxa of great conservation importance.	The Rhododendron of Sikkim Himalayan region are under pressure & some species at the point of extinction.Therefore, it is necessary to prepare comprihensive management plans to conserve the Rhododendron species.	Singh,K.K . (2009)
3	Isolation & characterization of microsatellite loci *Rhododendron ferrugineum* (E ricaceae) using pyrosequencing technology.	Development of microsatellite primers for *Rhododendron ferrugineum* (Ericaceae) to evaluate the genetic diversity, genetic structure and its mating system.	Delmas et al.(2011)
4	Planning a wildlife protected area network in India	Development of Protected Area Management Network in India for the conservation of biodiversity along with increased empowerment & cumminity level responsibility to preserve these threaten species.	Rodgers, W. A. and Panwar, H. S. (1988)
5	Biodiversity & conservation of rhododendrons in Arunachal pradesh in the Indo-burman biodiversity hotspot.	About 98% of the Indian species found in Himalayan region & out of which 85% present in Arunachal Himalayan region. But due to deforestation & unsustainable extraction of firewood by local peoples, it is classified as rare, endangered & threatened may be wiped out in near future.	Paul et al. (2005)

Table *continued……*

6	Origin and evolution of invasive naturalized material of Rhododendron ponticum L. in the British Isles	About 89% of R.Ponticum accessions are almost derived from Spain, while 10% belongs to Portuguese. Therefore, these results showed an Iberian origin of this British material & 26 accessions are also evident introgression from R.catawbiense & an unidentified species.	Milne & Abbott (2000)
7	Conservation of Rhododendrons in Sikkim Himalaya: An Overview	*Rhododendron* are classified as rare & endangered species,it may be wiped out in near future,if some conservation steps are not taken by local community.	Singh et al. (2009)
8	Discovery of species wide EST-derived markers in Rhododendron by Intron flanking primer design	Total 127 sequences of cDNA out of 323 were used for intron spanning markers development.Which show 35% polymorphism generating efficiency & transferability was confirmed within the *Rhododendron* genus.	Keyser et al. (2008)
9	In Vitro Propagation of R. maddeni Hook. F. an Endangered Rhododendron species of Sikkim Himalaya	Developed a protocol for rapid & large scale propagation of an endangered Rhododendron species of Sikkim Himalayan by using in vitro culture techniques.	Singh & Gurung. (2009)
10	Are rhododendron hybrids distinguishable on the basis of morphology and Microsatellite polymorphism?	33 Rhododendron germplasm used for genetic diversity assessment & found an average genetic diversity were 0.724 via morphological traits & 0.174 via microsattelites marker. The microsatellites markers proved an efficient tool in assessing the genetic variability among old broad leaf Rhododendron genotypes.	Caser.Akkak & Scariot. (2010)

2.2 Unigene derived microsatellite markers

Microsatellites or Simple Sequence Repeats (SSRs) or Short Tandem Repeats (STRs) consist of tandem repeats between 1-6bp, repeated up to 60 times. Hamada and colleagues first demonstrated these domains during the early eighties (Tautz & Renz, 1984; Tautz, 1989). Repeat units may consist of $(A)_n$, $(TG)_n$, $(CA)_n$ or $(AAT)_n$ repeats. Microsatellites are highly polymorphic due to the variation in the number of repeats (Litt & lutty**,** 1989)**.** It is not uncommon to find up to 10 alleles per locus and

heterozygosity values of 60% in a relatively small number of samples (Goldstein & Polack, 1997). Microsatellites tend to mutate with mutation rates of up to 10^{-2} per generation (Bruford & Wayne, 1993). A large number of microsatellite markers have been mapped for various species, including humans, mice, fruit flies and farm animals such as sheep, pigs and chickens (ibid,1997). These elements are most valuable markers in studies on genetic variability, parentage verifications and genome mapping projects. Primers developed for amplification of microsatellite loci can often be used between related species, which further promotes the application of microsatellites. They play a special role in gene regulation & genome evolution (Provan et al., 2001). The nucleotides sequences of the target unigenes are available to facilitate primer designing for PCR amplification of desired microsatellites. SSRs are found abundantly in eukaryotic genome with high polymorphic rate inexpensive PCR based assay (Sharma et al., 2009), show co- dominant Inheritance & used as a genetic markers. Its application includes: (a) Fingerprinting (Curro et al., 2010) (b) Diversity Analysis (Hamrick and Godt, 1996) (c) Genome introgression (Castillo et al., 2008) (d) Construction of genetic maps (Varshney et al., 2005) (e) Gene localizations (Varshney, Graner and Sorrells, 2005). SSRs are mostly found within the coding regions its nearby locations. Approximately 1-5% of plant expressed sequence tags (ESTs) contain SSRs of 20 nucleotides. Their expansions or contractions in coding regions are responsible for the gain or loss of any type of function via frame- shift mutations (Yu et al., 2004). These processes lead to describe the origin, evolution & distribution pattern of SSRs in the genome **(**Dayanandan et al., 2005). Different types of variations are found within the untranslated regions of gene, SSRs can regulate gene expression & their distributions are varying from species to species, across exonic, intronic & intergenic regions of the genomes. These variations in abundance are the result of rapid mutations in SSRs. Therefore, they play a special role in genome evolution by creating & maintaining quantative genetic variations (Hamrick & Godt, 1996). The standard method for development of SSR markers via genomic DNA libraries followed by DNA hybridization with radioactive probes but these processes are very time consuming & laborious. Therefore, Expressed Sequence tags (ESTs) databases available at public domain such as

(http://www.ncbi.nlm.nih.gov) can serve as the alternative source for identification & development of microsatellites markers. These databases reduce time & cost for microsatellites development. However, non-availability of sufficient sequence information & redundancy, due to which multiple set of markers are present at the same locus are the major drawbacks of EST-derived microsatellite markers (ibid, 2009). To sort out this type of problems or drawbacks of EST-derived microsatellite markers recently developed unique gene sequences (unigenes) via clustering of overlapping EST sequences (ibid, 2009). They overcome the problem of redundancy in EST databases & detect variation in the functional genome with unique identity & position.

Table 2.2: List of Research work done on UGMS markers

S.No	Topic of the Article/Journal	Results or Findings	References
1	Simple sequences are ubiquitous repetitive components of eukaryotic genomes.	SSR hybridized synthetical simple sequence DNA to genome blots of phylogenetically different organisms & found that they arise by slippage replication & unequal crossover.	Tautz and Renz (1984)
2	EST derived SSR markers for comparative mapping in wheat and rice	The structural & functional relationship between the genomes of wheat & rice by using the EST-SSR linkage maps was revealed. 44% of the primer pairs designed specifically for wheat sequences were successfully amplified in both species.	Yu et al.(2004)
3	SSR and EST-SSR markers for Opuntia Spp. Fingerprinting and genetic diversity evaluation	The level of intraspecific genetic diversity and fingerprinting was analyzed in opuntia spp. by using SSR markers and revealed high level of genetic diversity in Mexican germplasm. It provides information to better understand the ancestry of the cultivated species.	Curro et al. (2010)
4	Microsatellites and their application to population genetic studies	The study reveals the importance of microsatellites in population genetic studies & given the mutation rates in these repeats.	Bruford & Wayne (1993)
5	Chloroplast microsatellites: new tools for studies in plant ecology and evolution	The chloroplast genome shares many features with animal mitochondrial DNA, such as its conserved gene order, the widespread availability of primers & a general lack of heteroplasmy & recombination have made it an attractive tool for phylogenetic studies of plants.	Provan et al**.** (2001**)**

Table continued……

6	PCR amplified microsatellite as marker in plant genetics	The study revealed that the length of repeats with a minimum number of 7-10 units for better results. They uncovered microsatellite in 34 species with a frequency of 1 in every 50 Kb	Morgante & olivieri (1993)
7	Identification, characterization and utilization of unigene derived microsatellite markers in tea (Camellia sinensis.)	The predictions of 1223 unigenes from 2181 EST of tea & identified 8.9% unigene, which contain 120 SSRs. Unigenes, are developed via clustering of overlapping EST sequences, due to which the problems of redundancy in EST database are overcome.	Sharma et al. (2009)
8	Microsatellite analysis reveals genetically distinct populations of red pine (pinus resinosa, pinaceae)	Characterizion of 13 nuclear microsatellite loci & found 5 polymorphic microsatellite loci with an average of 9 alleles per locus were identified. The population differentiation was high with 28-35% of genetic variation portioned among populations.	Dayanandan et al**.** (2005)
9	Effects of Life History Traits on Genetic Diversity in Plant Species	Seven two trait combinations such as breeding system & seed dispersal mechanism of five life history characteristics were used to analyze interspecific variation in the level & distribution of allozyme genetic diversity in seed plants.	Hamrick & Godt (1996)
10	Microsatellites in Brassica unigenes: relative abundance, marker design, and use in comparative physical mapping and genome analysis.	The average frequency of microsatellites in Brassica unigenes was one in every 7.25 Kb sequences.Development of 347 unigenes from 2794 ESTs & used them for comparative physical mapping of Brassica identified 85% unigenes as single copy given clues for the presence of conserved gene order.	Parida et al. (2010)

2.3 Cross-species transferability

Cross species transferrability is an important mechanism to be used in population genetic studies (Acuna et al., 2012)**.** This is done by the designing of primers for one species i.e known as source species are used to amplify homologous loci in related or target species (Castillo et al., 2008)**.** The phylogenetically close species will share a higher proportion of microsatellites or other types of markers & show high level of amplification via the PCR process (Parida et al., 2006)**.** Then genetic distance between these two different species will be used as a parameter to detect the rate of

successful cross species transferability between these two different taxonomic groups (Hendre et al., 2008). Although EST-SSR may be conserved after many generation & over the larger evolutionary distances (Gutierrez et al., 2008). Therefore, the transfer of EST-SSR beyond the genus level appears too limited often (Yu et al., 2010). To overcome this type of problems some new types of nuclear markers are developed (Yadav et al., 2010), which are easily transferable & show higher level of polymorphism than EST-SSRs (Chai et al., 2013). They belong to EST-derived single nucleotide polymorphism (SNP), exon primed (Singh et al., 2013) & intron spanning (Chandra et al., 2013) markers. EST-SSR show low level of polymorphism (Guo et al., 2006), when tested within the species due to conserved nature of coding sequences as compared to when tested within related genera. Parida et al., (2006) identified and characterized microsatellite motifs in the unigenes available in five cereal crops (rice, wheat, maize, sorghum, barley) and Arabidopsis. These unigene derived microsatellite (UGMS) markers are expected to possess high inter specific transferability (Sharma et al., 2011) as they belong to relatively conserved regions of the genome (Park et al., 2010).

Table 2.3: List of Research work done on Cross-species transferability

S.No	Topic of the Article/Journal	Results or Findings	References
1	Exploiting EST databases for the development and characterization of 3425 gene-tagged CISP markers in biofuel crop sugarcane and their transferability in cereals and orphan tropical grasses	3425 Novel gene tagged markers known as conserved intron scanning primers (CIPS)were developed by utilizing 180 ESTs of Saccharum officinarum & then compare them with the ESTs of Sorghum, Barley, Whole rice genome sequence.337 PPs i.e. 185 (55%) out of 3425 primers show cross transferability in different Saccharum species & related genera 22% with four major grass species & 33% with Oat. PIC value was 0.52, marker index (MI) 3.50 & genetic diversity ranging from 0.88-0.99.	Chandra et al. (2013)

Table continued……

2	Microsatellite markers in candidate genes for wood properties and its application in functional diversity assessment in Eucalyptus globulus	In this study 8 new non reported SSR were identified in 7 candidate gene responsible for wood quality or properties in Eucalyptus globulus. Its PIC value ranging from 0.617 - 0.855 & probability of identity (PI) ranging from 0.030-0.151.	Acuna et al. (2012)
3	Transferability and polymorphism of barley EST-SSR markers used for phylogenetic analysis in Hordeum chilense	82 Barely EST-SSRs showed 21 primer pairs (26%) polymorphism among H. chilense & then used to test its transferability & polymorphism in other poaceae family.	Castillo et al. (2008)
4	Unigene derived microsatellite markers for the cereal genomes	UGMS marker of rice showed 9.5% homology with maize & wheat genome for their evolutionary closeness with them. UGMS markers of cereals & Arabidopsis showed least homology, because they diverged from each other about 200 million year ago are validate due to cross transferability of UGMS-COS markers in these five cereals crops & their comparison with Arabidopsis.	Parida et al. (2006)
5	Development of new genomic microsatellite markers from robusta coffee (Coffea canephora) showing broad cross-species transferability and utility in genetic studies	From 58 primer pairs of coffee from which 44 were validated as a single locus marker in Arabica & robusta genotypes. Its PIC value range from 0.49 to 0.62 & they showed 92% transferability across related species & genera of the coffee.	Hendre et al. (2008)
6	Cross-species amplification of Medicago truncatula microsatellites across three major pulse crops	The 209 EST based & 33 BAC (Bacterial Artificial Chromosomes) based microsatellites from Medicago truncatula in three European legume pulses-Pea, faba bean & chickenpea. They found 40% transferability in faba bean, 36.3% in chickenpea & 37.6% in pea respectively.	Gutierrez et al. (2008)
7	Development, characterization, and cross-species/genera transferability of SSR markers for rubber tree (Hevea brasiliensis)	252 SSRs were cloned from rubber tree from which 126 PPs were designed & only 36 PPs showed cross species amplification in 12 accessions of cultivated species, 4 related species & 3 species of Euphorbiaceae family.	Yu et al. (2010)

Table continued……..

8	EST-derived SSR markers in Jatropha curcas L.: development, characterization, polymorphism, and transferability across the species/genera	The development of EST-SSR markers showed 57% to 95.6% transferability among five species of Jatropha & 47% transferability across genera in Ricinus Communis and concluded that the transferability rates were depends on its search criteria also.	Yadav et al. (2010)
9	Identification, characterization and utilization of unigene derived microsatellite markers in tea (Camellia sinensis L.)	Prediction of 1223 unigenes from 2181 EST of tea & identified 8.9% unigene, which contain 120 SSRs and its transferability rates were 83.6% into their related species & 60% sequences had shown similarity with known proteins of Arabidopsis thaliana via gene ontology characterization process. To assess the conservation of these UGMS loci in this species, they tested them on five other species of tea with 61 primer pairs & recorded 83.6% cross transferability rate.	Sharma et al. (2009)
10	Transferability, polymorphism and effectiveness for genetic mapping of the Pummelo (Citrus grandis Osbeck) EST-SSR markers	Evaluation of 212 Pummelo EST-SSR (CgEMS) for their cross transferability was carried out. 136 markers showed amplification & 99 were transfer across the genera. Transferability of CgEMS to C. sinensis, C. reticulata, C. lemon, Fortunella sp. and Poncirus sp. was 76%, 76%, 75%, 74% and 73%, respectively.	Chai et al. (2013)
11	Developing new SSR markers from ESTs of pea (Pisum sativum L.)	530 unigenes of pea were identified out of 586 microsatellites & only 49 SSRs were used in primer designing & from them 9 loci were amplified & showed PIC of 0.18 -0.58 with cross species transferability in faba beans.	Gong et al. (2010)
12	Development, cross-species/genera transferability of novel EST-SSR markers and their utility in revealing population structure and genetic diversity in sugarcane	Development of 351 EST-SSRs from 4085 non redundant EST sequences of two Indian cultivars of sugarcane & 134 out of 351 sequences showed polymorphism with ranged of PIC value of 0.12 to 0.99 & cross transferability ranged from 87% to 93.4% in Saccharum, 80% to 87% in allied genera & 76% to 80% in cereals.	Singh et al. (2013)

Table continued……

13	Cross-species transferability of Gossypium arboreum-derived EST-SSRs in the diploid species of Gossypium	They examined that 124 out of 207 Gossypium arboreum-derived EST-SSRs. 60% PPs showed amplification & their cross –species amplification in 22 diploid Gossypium species was 96.5%.	Guo et al. (2006)
14	Identification and Cross-species transferability of 112 novel unigene-derived microsatellite markers in tea (camellia sinensis)	The study revealed that 112 TUGMS loci showed 100% transferability in cultivated C. assamica and C. assamica subsp. lasiocalyx. A high rate of cross-transferability was also recorded in C. japonica (69; 61.60%), C. rosiflora (96; 85.71%), and C. sasanqua (99; 88.39%) & PIC value ranged from 0.020 to 0.924.	Sharma et al. (2011)
15	Rose (Rosa hybrid L.) EST-derived microsatellite markers and their transferability to strawberry (Fragaria spp.)	Development of 312 UGMS markers from Rosa ESTs out of which 183 PPs showed polymorphism & 273 PPs was showed cross transferability in 4 genotypes of strawberry, a genus member of the Rosaceae family.	Park et al. (2010)

2.4 Advantages of EST-SSR over Genomic SSR markers

The frequency of microsatellites was significantly higher than in EST-SSRs than in genomic SSRs shown across all species (Castillo et al., 2008). The EST-SSRs show less level of polymorphism than the genomic SSRs due to greater DNA sequence conservation in transcribed regions (Pashley et al., 2006). For the assessment of genetic diversity the genic markers are more informative than genomic SSRs. Recently (Ellis and Burke, 2007) search EST collection from 33 plants & found that an average of 9% ESTs contains at least one SSR. EST –SSR derived from inside the ORF regions show higher level of amplification (65%) compared to UTR regions (54.5%) (Park et al**.,** 2010). Many times the polymorphism depends on the searching criteria of EST-SSR. EST-SSR derived from 3'-ESTs were more superior than 5'-ESTs. SSR present in 5'-UTR affect the gene transcription & translation process, while present in 3'-UTR are responsible for gene silencing & slippage (Chagne et al., 2004). EST-SSRs are also efficient as genomic SSRs for distinguishing between the

closely related species (Echt et al., 2011). For assessment of functional diversity the genic markers are more superior than genomic SSRs, while due to higher level of polymorphism, they are better for fingerprinting & varietal identification processes (Varshney, Graner & Sorrells, 2005).

Table 2.4: List of Research work showed the advantages of EST-SSRs over genomic SSRs

S.No	Topic of the Article/Journal	Results or Findings	References
1	Transferability and polymorphism of barley EST-SSR markers used for phylogenetic analysis in Hordeum chilense	The cross species transferability of SSRs derived from ESTs databases is greater than that of SSRs derived from enriched genomic DNA libraries, because they are develop from expressed region & therefore they are more conserved across a number of related species than non coding regions.	Castillo et al. (2008)
2	EST-SSRs as a resource for population genetic analyses	The EST-SSRs are more transferable across taxonomic boundries than traditional one. Therefore; they are used for the analysis of functional diversity, genetic mapping & MAS in crop species.	Ellis and Burke (2007)
3	High genetic diversity in a rare and endangered sunflower as compared to a common congener	EST-SSRs located in coding regions are significantly more transferable than UTRs. However levels of polymorphism are same for both. EST-SSRs have more statistical power & they produce clear results for scoring due to fewer null alleles.	Pashley et al. (2006)
4	Microsatellite DNA in genomic survey sequences and Unigenes of loblolly pine	The SSR densities were 96 SSR/Mb in Genome survey sequences (GSS) and 385 SSR/Mb in Unigenes and design 120 genome SSRs and 315 EST-SSRs for the comparisons between them and found EST-SSRs produced better results than genomic SSRs.	Echt et al. (2011)
5	Genic microsatellite markers in plants: features and applications	The study revealed that EST-SSRs are provided more information than genomic SSRs, because they represent transcribed genes via putative homology search.	Varshney, Graner & Sorrells, (2005)

2.5 Assessments of synteny, orthologous & homologous genes, etc.

In every organism some conserved DNA is found in the sequences of orthologous genes over large part of chromosomes & these conserved DNA is helpful to found out that the organism descended from a common ancestor (Varshney et al., 2007). Comparative genome mapping is the study of the conservation of gene content within chromosomal regions (synteny) between organisms utilizing comparative sequence analysis & genetic mapping of orthologous markers (Tang et al., 2006). These studies identified conserved genomic regions, chromosomal rearrangement, duplication, etc. providing information on the evolution of these families & assess their transferability to other species (Acuna et al., 2012). Cos markers are designed specifically from single or low copy genes, which reduce the difficulty of distinguishing orthology from paralogy found when multigene families are used for markers development & use of coding sequence enhanced the sequence conservation of primer binding sites across greater phylogenetic distances. The location of RosCos primer binding sites (Bushkara et al., 2011) in the conserved exons increase the transferability among genera & the potential of amplifying a variable intron, which provide an advantage over nongene based SSR as mutations within the SSR primer binding sites can lead to a decrease in transferability between genera (Powel et al., 1996). Therefore RosCos & SNP marker have greater success rate for transferability as well as for comparative genetic mapping to evaluate the synteny between two phylogentically distant genera (ibid, 2011).

Table 2.5: The list of Research work done on Genome synteny

S.No	Topic of the Article/Journal	Results or Findings	References
1	A high density barley microsatellite consensus map with 775 SSR loci	The SSRs are more useful for measuring functional diversity & high level of transferability as well as syntenic regions of rye. When genic SSR marker use for diversity studies the contraction & expansion of SSR repeats in genes of known function can be tested for their association with phenotypic variation.	Varshney et al. (2007)

Table continued……..

2	Homologous analysis of SSR-ESTs and transferability of wheat SSR-EST markers across barley, rice and maize	From 423,611 wheat ESTs,101,299 were found to express commonly in rice, maize & barely ESTs, which accounted for 23.94% of wheat ESTs data. Then 1707 SSR containing ESTs were mined from the 101,299 homologous ESTs, which accounted for 8.8% of all the 19,434 wheat EST-SSRs.	Tang et al. (2006)
3	Microsatellite markers in candidate genes for wood properties and its application in functional diversity assessment in Eucalyptus globulus	The SSR markers derived from candidate genes (SSR-CG) can be used effectively in co-segregation studies & marker assisted diversity management.Mostly SSRs are found in 5'-3' UTRs but in the case of eucalyptus 72% of the SSRs found in ORF region, While 28% in the UTRs. These unexpected results might be caused by an overestimation of the ORF length.	Acuna et al. (2012)
4	Rosaceae conserved orthologous set (RosCos) markers as a tool to assess genome synteny between Malus and Fragaria	The 21 regions of genomic synteny between two distant species and able to found out the location of RosCos marker on apple linkage group (LG) & 7 LG of Strawberry was used to assess the ancestral relationship between two genera. In every organism some conserved DNA is found in the sequences of orthologous genes over large part of chromosomes & these conserved DNA is helpful to found out that the organism descended from a common ancestor.	Bushkara et al. (2011)
5	A linkage map for the B-genome of Arachis (Fabaceae) and its synteny to the A-genome	Construction of B-genome linkage map compliment to A-genome of Arachis.The map has 10 linkage groups with 149 loci & showed higher transferability -81.7%,when applied in other Arachis species.When B-genome compared with A-genome map using 51 common markers showing high level of synteny between both genome.	Moretzsohn et al. (2009)

2.6 Genetic diversity & Importance of SSR markers & their comparisons with other markers, etc.

The different types of life form & their breeding system had significantly influences their genetic diversity & distribution (Selkoe and Toonen, 2006). Genetic diversity serves as a way for poulations to adapt to changing environment. If more varation is found in population, its show that some individual in population will possess high varation of allells that are suited for environment (Caruso et al., 2008). Genetic diversity play important role in the survival & adaptability of a species. The species that show less gentic variation are at greater risk of extinction (Nerkar et al., 2012). Genetic diversity can not be determined directly from natural poulations because of some environmental influences on its quantative traits (Hamrick & Godt, 1996). It is also noted that out-crossing species are more diverse than other traits in their populations & also found that woody plants have less among population differentiation & more genetic diversity than non woody plants or herbaceous species (ibid,1996). The gene movement of trees should also affect their ability to maintain genetic diversity & found that new alleles in populations of out-crossing species have higher probability of being dispersed into other populations than novel genes introduced into populations of herbaceous species with limited gene dispersal potential (Dayanandan et al., 2005). New alleles have greater dispersal potential due to which they are less likely to be lost through drift or population extinction, while the novel alleles are more likely to be lost in species that experience less gene flow; therefore these species have fewer polymorphic loci & less genetic diversity. Chloroplast microsatellites have revealed much higher levels of diversity than those of chloroplast RFLPs & RAPD (Cotti, C., 2008). it showed wider applications such as in population genetics, it showed finer details as good indicators of historical bottleneck, founder effects & genetic drift, etc. (ibid, 2008). While some new improved types of markers such as SNP & Intron spanning, etc. (Jones et al., 2007) have show better results than SSRs also.

Table 2.6: The List of Research work done on genetic diversity, etc.

S.No	Topic of the Article/Journal	Results or Findings	References
1	Microsatellites for ecologists: a practical guide to using and evaluating microsatellite markers.	SSRs can be used to estimate many parameters of interest to ecologist such as migration rates, population size, bottleneck, kinship, etc. It is also used to find out hidden allelic diversity, such as alleles of the same size but different lineages a termed known as homoplasy.	Selkoe and Toonen (2006)
2	Generation of expressed sequence tags from carob flowers for gene identification and marker development	The generation of EST from Carob flowers for the identification of gene and found that some identified genes are responsible for reproduction, biotic & abiotic factors such as environmental stress, etc.	Caruso et al. (2008)
3	Unigene-derived Microsatellite Marker Based Variation Study of Phytophthora nicotianae Isolates Infecting Citrus	The assessment of genetic variation among 24 Phytophthora nicotianae isolates from 5 major citrus growing states of India by using UGMS markers & found that 62.27%polymorphism & these are functionally informative for studying genetics of population, epidemiology, ecology & evolution.	Nerkar et al. (2012)
4	A comparison of simple sequence repeat and single nucleotide polymorphism marker technologies for the genotypic analysis of maize (Zea mays L.)	The comparisons of the SSR & SNP marker technology in terms of their informativeness, levels of missing data, repeatability & the ability to detect expected alleles in hybrids & DNA pools and found that average expected heterozygosity values were 0.62 for SSRs, 0.43 for SNPs & 0.63.SNP data had lower level of missing data i.e. 2.1-3.1% compared with SSRs-13.8%.Data repeatability was higher for SNPs i.e. 91.7-99.3% compared to SSRs i.e. 91.7%. These results showed that SNP technology can provide increased marker data quality & quantity compared with SSRs	Jones et al. (2007)
5	Effects of Life History Traits on Genetic Diversity in Plant Species	New alleles have greater dispersal potential due to which they are less likely to be lost through drift or population extinction while the novel alleles are more likely to be lost in species that experience less gene flow; therefore these species have fewer polymorphic loci & less genetic diversity.	Hamrick & Godt (1996)

Table continued……

6	Comparison of RAPD, RFLP, AFLP and SSR markers for diversity studies in tropical maize inbred lines	The comparision of 774 loci for AFLPs,262 loci for RAPDs,185 loci for RFLPs & 68 loci for SSRs and revealed the genetic distance estimation,the AFLP & RAPD gave the most correlated value,coefficient of variation (CV) given by codominant markers was higher than the dominant markers (AFLP & RAPD) & finally concluded that AFLP is best suited molecular assay for fingerprinting & assessing genetic relationship.	Garcia et al. (2004)
7	Molecular markers for the assessment of genetic variability in threatened plant species	Investigation of the phylogenetic relationship, population structure & gene flow. The presence of null alleles are due to mutations in one or both primer binding sites & these mutations can prevent PCR amplifications in SSRs.	Cotti, C. (2008)
8	Cross species transferability and mapping of genomic and cDNA SSRs in pines	The identification of 419 SSRs from which 12.8% overlapped between the two sets.They assesed that the genomic SSR perform better in terms heterozygosity while EST-SSR better in term of PIC and transferabilty.	Chagne et al. (2004)
9	Development of a molecular linkage map of pearl millet integrating DArT and SSR markers	The construction of linkage map by integrating DArT and SSR tecnology.The map contain 321 loci which comprises 258 DArT and 63 SSRs.DArT provide cost effective selection in MAS and QTL.	Supriya et al. (2010)
10	Genetic diversity within and among pinus populations:comparison between AFLP and microsatellite markers	The comparisions of 23 population of pinus at 3 microsatellite loci and 122 AFLP loci and revealed that microsatellite give higher within genetic diversity than AFLP.	Mariette et al. (2001)

Chapter-III
MATERIAL AND METHODS

3.1 Sample collection and Study Sites

The study sites are located in five different parts of Himachal Pradesh in India. Young, disease free and fresh leaves of Rhododendron arboreum from Western Himalayan region of Himachal Pradesh were taken as a sample from 47 individual trees. The samples were collected and transported to the laboratory for genomic DNA extraction. The list of these samples is given in Table no: 1.

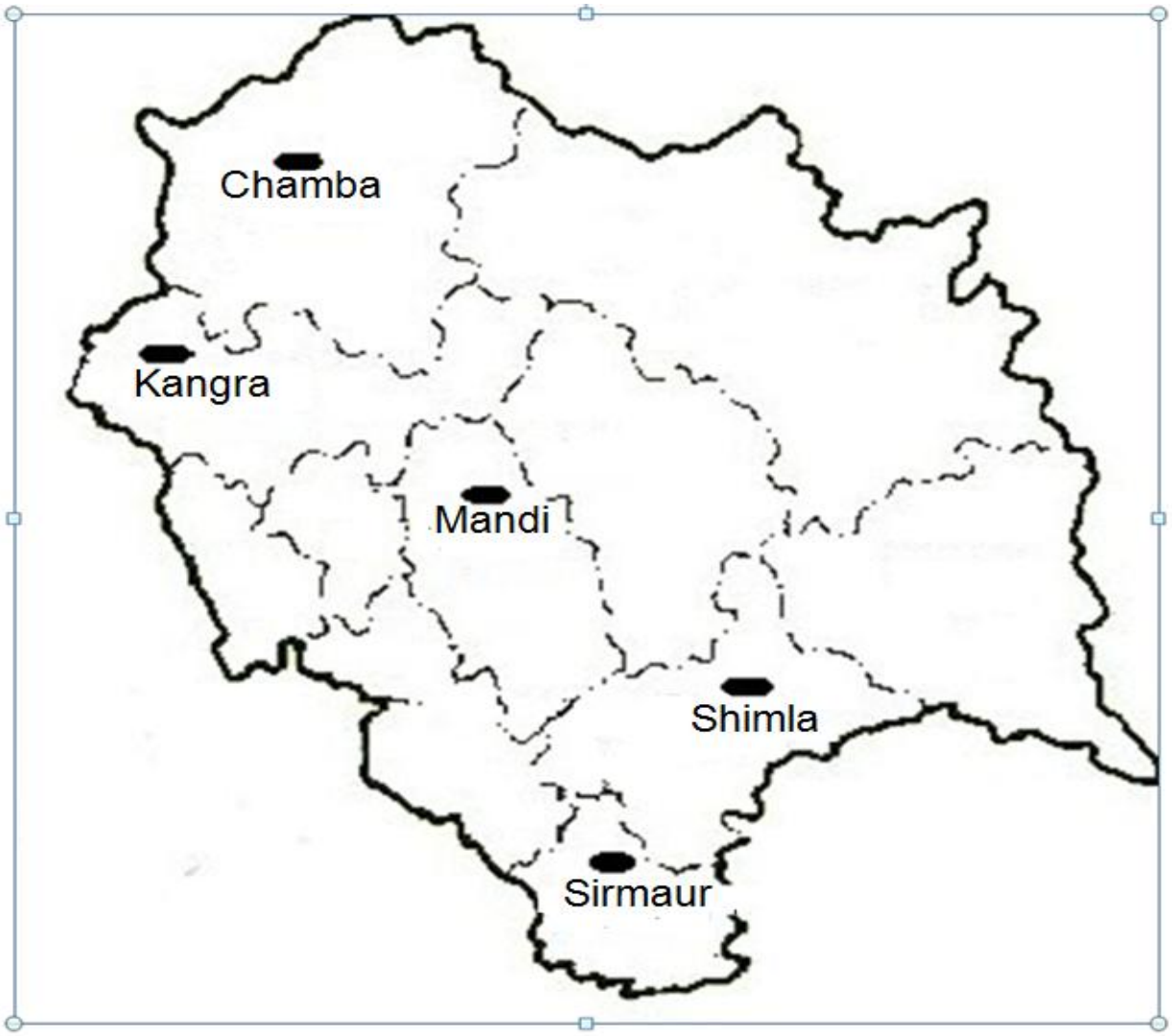

Fig-3.1: Different locations of Sample collection from 5-different parts of Himachal Pradesh

Table 3.1: Details of sample collection with location

S.No.	Location	Latitude	Longitude	Elevation
1	Chowari	32°25´59.83´´N	76°01'00.29"E	3533ft
2	Dhanoti village,chowari	32°29´59.83´´N	76°02'00.29" E	3895ft
3	Dhanoti bus stand	32°33´59.83´´N	76°06'00.29" E	3895ft
4	Chowari to Jot tracking route	32°56´59.83´´N	76°25'00.29" E	4045ft
5	Jot to Delhousie tracking route	32°48´48.33´´N	76°35´04.07´´ E	16239ft
6	Jot to Delhousie tracking route	32°27´48.33´´N	76°22´04.07´´E	16252ft
7	Jot to Delhousie tracking route	32°48´36.86´´N	76°33´52°49´´ E	16553ft
8	Jot Picnic spot	32°48´36.64´´N	76°28´48.58´´ E	17157ft
9	Baihali Jot near highest picnic spot	32°50´04.33´´N	76°37´00.21´´ E	20144ft
10	Jot bus stand	32°02´04.33´´N	76°06´00.21´´E	19857ft
11	Sadar Bazar,Delhousie	32°32´03.46´´N	75°57´57.42´´E	6437ft
12	Garam Sadak,Delhousie	32°32´03.05´´N	75°58´46.94´´E	6765ft
13	Subash Chowk,Delhousie	32°32´06.87´´N	75°58´03.29´´E	6647ft
14	Delhousie bus stand	32°32´03.46´´N	75°59´00.62´´E	5565ft
15	Swas boli,Delhousie	32°32´05.42´´N	75°59´01.57´´E	5578ft
16	Tikka lane, Delhousie Cantt	32°32´41.55´´N	75°57´19.74´´E	5527ft
17	Chandri Ghat,Delhousie	32°32´00.47´N	75°58´06.47´´E	5569ft
18	Subash Bowli,Delhousie	32°32´08.45´´N	75°59´07.53´´E	5582ft
19	Panchpula,Delhousie	32°32´03.05´´N	75°58´46.94´´E	4243ft
20	Delhousie Cantt	32°32´40.55´´N	75°57´18.74´´E	5528ft
21	Beedh to Gharnola road,dist. Kangra	32°02´11.18´´N	76°42´54.06´´E	4695ft
22	Beedh near mata mandir	32°02´10.15´´N	76°42´50.62´´E	4652ft
23	Beedh Forest Dept.	32°02´10.16´´N	76°42´57.66´´E	4673ft
24	Gharnola village	32°02´11.05´´N	76°42´50.08´´E	4689ft
25	Gharnola tracking route	32°02´11.09´´N	76°42´50.11´´E	4691ft
26	Billing Chowk	32°04´23.80´´N	76°47´48.22´´E	7868ft
27	Dragna Village, ,Dist-Mandi	32°04´07.49´´N	76°51´41.59´´E	7215ft

28	Dragna Village, ,Dist-Mandi	32°04´07.49´´N	76°51´41.59´´E	7215ft
29	Barot, Tehsil-Padhar,Dist-Mandi	32°08´01.49´´N	75°55´28.51´´E	1639ft
30	Near Barot fishery firm	32°08´01.49´´N	75°55´28.51´´E	1639ft
31	Barot, Tehsil-Padhar,Dist-Mandi	32°08´01.49´´N	75°55´28.51´´E	1639ft
32	Multhan,Tehsil-Multhan,Dist-kangra	32°04´28.01´´N	76°52´06.94´´E	7730ft
33	Multhan,Tehsil-Multhan,Dist-kangra	32°04´28.01´´N	76°52´06.94´´E	7730ft
34	Dayot Village, Tehsil-Multhan,Dist-Kangra	32°28´35.07´´N	76°55´52.06´´E	7781ft
35	Garla Village, Tehsil-Multhan,Dist-Kangra	32°04´47.15´´N	76°52´27.48´´E	9072ft
36	Bada Gaon, Tehsil-Multhan,Dist-Kangra	32°06´25.09´´N	76°55´40.45´´E	13299ft
37	Bada Gaon, Tehsil-Multhan,Dist-Kangra	32°06´25.09´´N	76°55´40.45´´E	13299ft
38	Jhaku Mandir,Shimla	31°06´04.47´´N	77°11´02.12´´E	8016ft
39	Jhaku Mandir,Shimla	31°06´04.47´´N	77°11´02.12´´E	8016ft
40	Totu market,Shimla	31°06´04.48´´N	77°08´31.81´´E	6309ft
41	Totu market,Shimla	31°06´04.48´´N	77°08´31.81´´E	6309ft
42	Near ISBT,Shimla	31°05´52.87´´N	77°09´03.45´´E	6413ft
43	Sanjoli,Shimla	31°06´09.71´´N	77°11´29.33´´E	7364ft
44	Near Summer hill RLY.Station	31°06´50.27´´N	77°07´24.37´´E	3937ft
45	Kufri Water Catchment area	31°06´10.46´´N	77°15´78.93´´E	8678ft
46	Kufri forest department	31°06´04.77´´N	77°15´51.03´´E	8753ft
47	Munda ghat,Chail	31°04´00.71´´N	77°15´54.22´´E	7495ft

3.2 DNA isolation

Genomic DNA was extracted from young leaf tissues of each genotype using CTAB method given by Doyle and Doyle (1990). The quality and quantity of DNA were estimated on 0.8% agarose gel and checked by Nano drop 2000 spectrophotometer.The detail of DNA isolation and purification protocol has been given in Appendix-A .

3.3 EST data mining, Unigene prediction and SSR Detection

A total of 1,241 FASTA formatted EST sequences in Rhododendron catawbiense were retrieved on March 24, 2012 from the National Centre for Biotechnology Information (NCBI;http://www.ncbi.nlm.nih.gov/entrez) for subsequent data mining. This dataset was scanned and assembled using SeqMan DNA Star lasergene version 10.0 (DNASTAR Inc, Madison, WI) and predicted potential unigenes that contained contigs and singletons from all the EST sequences with parameters (match size: 5, minimum match percentage: 80, match spacing: 150, gap penalty: 0.00, gap length penalty: 0.70, maximum mismatch bases: 15). Further, gaps in the aligned sequences due to limited dataset were removed on the basis of probability function of nucleotide occurring at the particular position using Gene runner. All the unigenes were subsequently searched individually for the presence of SSRs with help of Repeat masker (http://www.repeatmasker.org/cgi-bin/WEBRepeatMasker) and SSRs with a minimum length of ≥ 18 bp (di & tri) and ≥ 15 bp (tetra, penta & hexa) were masked. These parameters were chosen to identify SSRs with high polymorphic rate and these assemblies were used as data on Primer3 (http://bioinfo.ut.ee/primer3-0.4.0/) primer design input site for the designing of primers. The primers were design with following criteria:- (I) nucleotide length of 18 – 22 base pairs, (II) a Tm value of 50°C to 60°C, (III) the 3' end base with a G or C, preferably and (IV) an amplified fragment size of 100 – 350 bp. The formation of secondary structure and primer dimmers were critically monitored to get success of the primers. The names of the primers were prefixed as RUGMS (Rhododendron Unigene derived microsatellite) markers as the source is from Rhododendron catawbiense Unigene database.

3.4 PCR amplification

We standardized the protocol for Polymerase chain reaction (PCR) for the DNA amplification by assessing DNA concentration, Primer concentration, different brands and concentrations of Taq polymerase as well as temperature. PCR reactions is carried out in a total of 20 µl volume containing 25 ng template DNA, 2.0 µl of each forward and reverse primer, 4.0 µl of 100mM of dNTPs (Himedia), 2.5 µl of Taq DNA

polymerase (Himedia), 2.0 µl of 10X PCR buffer (Himedia) and 2.0 µl of 2.5 mM of MgCl2 & 2.5 µl of autoclaved distilled water in each PCR tubes. Amplifications are performing in a thermal cycler (Applied Biosystems Veriti 96 well) with fallowing parameters: initial denaturation at 94°C for 5 min. followed by amplifications. Each amplification cycle is initially at 94°C for 1 min for 35 cycles followed by annealing temperature (Ta) for 1 min. i.e. 50°C and then 72°C for 2 min; final extension at 72°C for 8 minute. The amplified products were stored at 4°C. The amplification products of PCR were mixed with denaturing dye (98% Formamide) and run on 7.8% PAGE in 1x TBE buffer. The gels photographs are taken by placing gel slides on X-ray illuminator view box.

3.5 Data analysis

The DNA fragments size were estimated by comparing with DNA marker (100bp DNA ladder, Himedia) run on the same PAGE gel and DNA fragment size estimated. The visualized intense bands were scored and recorded for statistical analysis. Polymorphism information content (PIC), Number of amplified bands, allele frequencies and expected (He) and observed heterozygosity (Ho) was calculated from observed genotypes. This gives estimate about discrimination ability of a marker by considering number of alleles at a locus along with their relative frequencies. Various type of statistical analysis software's such as Popgene software version 1.32 (Yeh et al., 1997) & NTSYC (Rohlf, 2000) are available on the internet as free of cost are used for analyzing the Genetic diversity in the provided sample for inferring the differences in the sample population in the area of population genetics.

The POPGENE software version 1.32 (Yeh et al.,1997) was empolyed to determine different parameters which include number of polymorphic loci and their percentage, observed number of loci, effective number of alleles (N_e) was estimated using the formula Ne = Mix*i*2, where, xi is the frequency of the /th allele for each locus (Crow and Kimura, 1964). Each band amplified by each primer was first scored as allelic data and then changed into a binary matrix as (1) for presence of band and (0) for absence of band were used to generate jaccard's coefficient of similarity by using distance matrix formed by NTSYSpc version 2.1 software (Rohlf,2000). The similarity

coefficient was used to construct a dendrogram by the unweighted pair group method with arithmetic averages UPGMA (Rohlf, 1993). The Polymorphism Information Content (PIC) value was described by Botstein et al. (1980) & modified by Anderson et al. (1993) for self pollinated species was calculated as follows:

$$PIC = 1 - \sum_{j=1} P_{ij}^2$$

Where,Pi is the frequency of the i/th allele & Pj is the frequency of the allele. The above mentioned methods are used for estimating the results.

Chapter-IV

RESULTS

Results

4.1 ESTs/Unigenes data set

A total 792 (615 singletons and 177 contigs) unigenes were predicted from 1,241 publicly available EST database from NCBI (http://www.ncbi.nlm.nih.gov/) site in Rhododendron catawbiense by clustering of 2 – 33 random EST sequences. Non-redundant (NR) sequence data set represented~184.41 kb expressed genome of Rhododendron catawbiense.

4.2 Abundance and distribution of SSRs

All 792 potential unigenes were searched for the presence of microsatellites. A total of 177 (22.3%) unigenes containing 50 SSRs with motif length ranging from 2 to 6 bp were identified. One sequence contained six SSRs and three sequences contained two SSRs each. Six SSRs were of compound types. One SSR was detected for every 27.17 kb of the EST sequences. Further analysis of SSR containing Unigene sequence data revealed that majority of them (72%) were perfect repeat and/or class I (≥ 20 nucleotides; nts length). However, remaining 28% (comprising of 17.9% di repeats and 3.3% each of tri repeats, tetra and penta repeats) were found to be of class II types (≥ 12 nts and < 20 nts length). In present study, only class I UGMS markers were selected for primer designing.

Data analysis of SSR motifs in unigenes revealed 17 di repeats (34%), 23 tri repeats (46%), 2 tetra repeats (4%), 6 penta repeats (12%) and 2 hexa repeats (4%) (Table no: 2).Among the di-nucleotide repeats the GA/TC (94.11%). motifs were most abundant (94.11%) followed by TG/CA. Among the microsatellites containing tri-repeats, which were followed by TTC/GAA (34.78%), CCA/TGG (26.08%), CAT/ATG, TCC/GCA, and CGA/TCG were same (4.34%). Abundance of other tri repeat containing SSRs were more or less in the similar range. Frequency of tetra, penta and hexa repeat containing SSRs was less. In tetra nucleotide microsatellite the repeats were CATA/TATG (50%) & TTTC/GAAA (50%). In Penta nucleotide microsatellite the repeats were TTCTC/GAGAA (33.33%), CCCCG/CGGGG (16.

66%), TTTAG/CTAAA (16.66%), GCATG/CGTAC (16.66%) and CAAAA/TTTTG (16.66%). In Hexa nucleotide microsatellite the repeats were GGAGGA/TTCTCC (50%) & TAAAAA/TTTTTA (50%) and no GC/CG microsatellites were identified. All these repeats types with their distribution and frequencies were shown in below mention table with their frequencies.

Table no: 4.2 showing different types of SSR predictions in 50 unigenes of Rhododendron Catawbiense & its types of repeats are following:-

S.No.	Type of repeats	S.No. of Repeat Types	Repeats	No. of each Repeats	Total
1	Di-Repeats	1	TG/CA	1	17
		2	GA/TC	16	
2	Tri-Repeats	1	CCA/TGG	6	23
		2	CTG/CAG	3	
		3	TTC/GAA	8	
		4	CCG/CGG	3	
		5	CAT/ATG	1	
		6	TCC/GGA	1	
		7	CGA/TCG	1	
3	Tetra-Repeats	1	CATA/TATG	1	2
		2	TTTC/GAAA	1	
4	Penta-Repeats	1	TTCTC/GAGAA	2	6
		2	CCCCG/CGGGG	1	
		3	TTTAG/CTAAA	1	
		4	GCATG/CGTAC	1	
		5	CAAAA/TTTTG	1	
5	Hexa-Repeats	1	GGAGAA/TTCTCC	1	2
		2	TAAAAA/TTTTTA	1	

4.3 UGMS primer designation

Of the 177 NR unigenes containing one or more SSRs, 50 (28.24%) were amenable to design flanking oligonucleotide primer pairs. Thirty six UGMS primer pairs (25 from singletons and 9 from clusters) flanking to different repeat motifs could be designed. Primer pairs flanking tri repeats (46%) were the most abundant followed by di (34%), penta (12%), tetra (4%) and hexa (4%) repeats containing microsatellites. Primers could not be designed for the rest fourteen (7.91%) SSR containing unigenes because of either insufficient flanking sequence (occurrence of SSR near or/at either end of the Unigene) or inability to fulfill the criteria for primer design.

4.4 Cross species Marker evaluation and polymorphism detection

Thirty six primer pairs designed in this study from the EST databases of R. catawbiense were used to amplify DNA from a panel of 47 accessions of R. arboreum. Of these, 6 (16.66%) primer pairs produced repeatable and reliable amplifications in at least forty seven accessions of R.arboreum, while 30 (83.33%) primer pairs either completely failed or led to weak amplifications and thus were excluded from further analysis (Appendix-B). Marker evaluation details are given in table no: 4.4. These six amplified primers were used for the scoring & genotyping of the R.arboreum samples. These six primers were further used for the assessment of genetic diversity of R.arboreum in Western Himalayan region of India. PCR products of the expected size were obtained in all the cases except in one UGMS primer (RUGMS-39) that had amplified larger size additional amplicons in some cases. Multi-locus amplifications were recorded in case of RUGMS-212, RUGMS-282 and RUGMS-39. Over all, amplification success rate was the maximum in case of RUGMS primer pairs containing tri repeats (66.66%), followed by di-repeat (16.66%) and penta-repeat (16.66%). Six polymorphic primer pairs namely RUGMS-212, RUGMS-282, RUGMS-619, RUGMS-204, RUGMS-320 and RUGMS-39 gave amplification in all the tested genotypes irrespective of species and hence can be

utilized as universal markers for molecular analysis in R.arboreum. However, these markers need to be validated in a larger panel of Rhododendron species. Six primer pairs amplified 29 alleles of which 20 (68.96%) were found to be polymorphic. All the UGMS markers identified in the present study remained highly polymorphic. The number of alleles detected in the present case ranged from 4 to 5 alleles per locus. The UGMS markers namely RUGMS-619, RUGMS-212, RUGMS, RUGMS-282 and RUGMS-39 recorded a maximum of 5 alleles and RUGMS-204 with 4 alleles respectively. A high level of polymorphism has been observed at the species level. The expected heterozygosity (H_e) and observed heterozygosity (H_o) ranged from 0.6338 to 0.7109 (with an average of 0.6783) and 0.8000 to 0.9250 (with an average of 0.8666), respectively (Table 4.4). All the UGMS markers showed a significant departure from Hardy-Weinberg equilibrium (HWE) at $P < 0.001$ level. The polymorphism information content (PIC) ranged from 0.2756 to 0.9212 with an average of 0.5765. There was significant difference in the average PIC values was recorded in UGMS locus harboring different repeat types. Average PIC values ranged from 0.4999 (penta repeats) to 0.9212 (tri repeats). However, an average of 0.5885 PIC values were recorded in RUGMS primers with di repeats respectively (Table 4.4). Of the 6 UGMS primer pairs with PIC values ≥ 0.50, 3 (50%) namely RUGMS-619, RUGMS-320 and RUGMS-204 recorded amplification in 47 accessions were identified as informative and thus would be useful in future marker assisted studies in Rhododendron. Further, the others 3 (50%) UGMS primer pairs with average PIC values ranged from above 0.50 to 1.0 were identified, namely RUGMS-212, RUGMS-282 and RUGMS-39 recorded amplification in 47 samples which may categorized as very informative primers after their validation in a larger panel of Rhododendron arboreum accessions. The approximate size of the PCR amplified bands were ranged from 200bp to 500bp. The highest number of genotypes amplified at locus RUGMS-204 and lowest at locus RUGMS-212 and RUGMS- 39.

Table no: 4.4 Marker validation and features of new 6 UGMS markers of Rhododendron arboreum

Locus name	Primer sequence	Repeat motif	Annealing temperature	No. of alleles	Heterozygosity		PIC	Approximate size range(bp)	No. of genotypes amplified
					H_o	H_e			
RUGMS-619	F5'-CCTTGAATGGAGGTGGAGAA R5'- CCAACCAGCAAATAGGCATT	$(CCA)_{20}$	50°C	5	0.8000	0.7109	0.3463	200bp	38
RUGMS-212	F 5'-TGGCTAGCCCAAGCACTAAT R 5'-CCAAATGGAAATGGGTTTTG	$(TTC)_9$	50°C	5	0.9250	0.6572	0.9212	250bp	19
RUGMS-282	F 5'-TCTACTACTGCGCCAAAGCA R 5'-ATTGGTCCCCTTCTGGTGAT	$(GA)_{11}$	50°C	5	0.8500	0.7012	0.5885	300-400bp	28
RUGMS-320	F 5'-CTCCGTTGTGTTTGGTTGAG R 5'-TTCTCGAAGATAGGGGCAGA	$(TTTAG)_6$	50°C	5	0.8250	0.7050	0.4999	300-400bp	33
RUGMS-204	F 5'-GATGCTTCAAAACCCCAAAA R 5'-CCGAAGGGTTCATGAAGAAA	$(CCG)_8$	50°C	4	0.9000	0.6338	0.2756	250bp	40
RUGMS-39	F 5'-GAATTGATTGCTTGGGGAAA R 5'-TTCGGGTTTCAACAGCTAGG	$(CTG)_2$	50°C	5	0.9000	0.6672	0.8279	300-500bp	19

4.5 Genetic variation in populations

Four parameters were used to assess the genetic variation within populations and they included number of loci with an average alleles per locus, number and percentage of polymorphic loci, actual or observed number of alleles (Na), effective number of alleles (Ne), observed heterozygosity (H_o), expected heterozygosity (H_e). The values of all these variables were shown in above mention table. The 47 samples analyzed from 5 different populations of R.arboreum, 5 different alleles were detected at six loci, with an average of 4-5 alleles per locus. The alleles revealed by SSR markers showed a high degree of polymorphism; with as many as 3 primers out of total 6 primers produced 100% polymorphic bands. A total of 177 bands were obtained from 6 SSR primers, of which 66.66% were polymorphic, with an average of 62.76 bands per primer. The PIC values, derived from allelic diversity and frequency among the genotypes, were not uniform for all of the SSR loci tested. The Polymorphism Information Content (PIC) value is also used as a measure of genetic variation. The population diversity & allelic variability is indicated by PIC. It's value always between 0 and 1. This is use to assess the diversity of a gene or DNA segment in a population, which will throw light on the evolutionary pressure on the allele & the mutation on the locus might have undergone over a period of time. The PIC value for the SSR loci in present study ranged from 0.2756 to 0.9212. The highest PIC value present at locus RUGMS-212 & lowest PIC value present at locus RUGMS-204. The average PIC value was 0.5765, which show moderate level of genetic diversity within the population. The PIC value refers to the markers informativeness for detecting polymorphism within population. The average of the highest polymorphic loci was 6.66% at locus RUGMS-204 and lowest at RUGMS-212 and 39. The PIC value in present study shows that SSR marker is very informative for measuring polymorphism within the population. In each population, the mean of effective number of alleles (Ne) was 3.13 and observed number of alleles (Na) was 4.83 compared within the population.

The mean effective number of alleles is the minimum number of alleles required to maintain the current level of heterozygosity in the population if all allele frequencies were equal. Most populations showed a higher number of observed alleles than the effective number of alleles in the table shown below.

Table no. 4.5: Genetic variation of Rhododendron arboreum population by UGMS marker.

Population	Sample size	Na	Ne
Chamba	10	5.0000	3.4595
Kangra	10	5.0000	2.9170
Mandi	10	5.0000	3.3898
Shimla	10	4.0000	2.7304
Sirmaur	07	5.0000	2.9602
Total	**47**		

Na = Observed number of alleles, Ne = Effective number of alleles

Over all populations of R. arboreum the observed heterozygosity (H_o) was highest (0.92) at locus RUGMS-212 followed by (0.90) at RUGMS-204 & RUGMS-39, (0.85) & (0.82) at locus RUGMS-282 & RUGMS-320 & (0.80) at locus RUGMS-619 loci respectively. Expected heterozygosity (H_e) was highest (0.71) at locus RUGMS-619 followed by (0.705) at locus RUGMS-320, (0.701) at locus RUGMS-282, (0.66) at locus RUGMS-39, (0.65) at locus RUGMS-212 & (0.63) at RUGMS-204 loci respectively. According to our results the observed heterozygosity is greater than the expected heterozygosity in each population. Jain et al. (2000) also reported a high level of observed heterozygosity for R.arboreum. The high level of observed heterozygosity could be attributable to predominant outcrossing in the species (Jain et al., 2000). This species is thought to be predominantly pollinated by insects, which supports high outcrossing rates as in other insect pollinated plants (Hamrick and Godt, 1996).

The genetic variation among the population was also investigated by using four parameters, which included average observed (H_o) and expected heterozygosity (H_e), coefficient of genetic similarity and dissimilarity, Nei's genetic identity and genetic distance and allele frequency. The frequency distribution of alleles varied over populations. A few alleles were rare and found in some populations with a frequency as low as 0.05 similarly few alleles were restricted to only one or two populations only. Nei's (1978) standard genetic distances and genetic identity between all pairs of populations ranged highest distance from (0.0065) to (0.7697) in population of chamba and mandi respectively at different locus. The observed heterozygosity (H_o) was highest than expected heterozygosity (H_e) in each population. The average observed heterozygosity (H_o) was 0.8666 and average expected heterozygosity (H_e) was 0.6792. According to present results the genetic variation among the population is also less than genetic variation within population. By using jaccard's coeffcient of similarity obtained by using distance matrix table formed by NTSYSpc version 2.1 software (Rohlf, 2000), the level of genetic similarity and dissimilarity were calculated. The genetic diversity was found with an average of 0.6222 within population of R.arboreum and genetic diversity found with an average of 0.3436 among the population. The results revealed that high level of gentic diversity was found within the population of R.arboreum was due to out-crossing species. The across related species/genera transferabilty rate from *R*.catawbiense to *R.arboreum* was found 16.66% i.e higher than sugarcane (Cordiero et al., 2001) was 15% only and lower than many studies as in robusta coffee (Hendre et al., 2008) was 92% transferability rate. The lower level of transferability show less homology betweeen the related genera. Hence, within population were showing higher level of genetic diversity than among the population.

4.6 Cladistic Analysis

A dendrogram was constructed by UPGMA clustering based on Jaccard's coefficient of similarity by using distance matrix table formed by NTSYSpc version 2.1 software

(Rohlf, 2000) to visualize the relationships among the 5 populations of R.arboreum. The clustering of the SSR based genetic similarity (GS) was shown in the dendrogram divided into four main clusters including some sub-clusters also. Which present some degree of similarity between them. The cluster I have average genetic similarity 0.55 and comprises of 4 different genotypes, cluster II have average genetic similarity of 0.45 and comprises of 21 different types of genotypes, cluster III have average genetic similarity of 0.39 and comprises of 18 different types of genotypes and cluster IV have an average genetic similarity of 0.20 and comprises of 4 different types of genotypes, respectively. Cluster analysis showed a significant genetic variation among the genotypes, with average similarity coefficient ranging from 0.20 to 0.55 respectively. The highest values of Nei's genetic distance were found between the populations of Chamba (0.0065) and Mandi (0.7697) respectively.

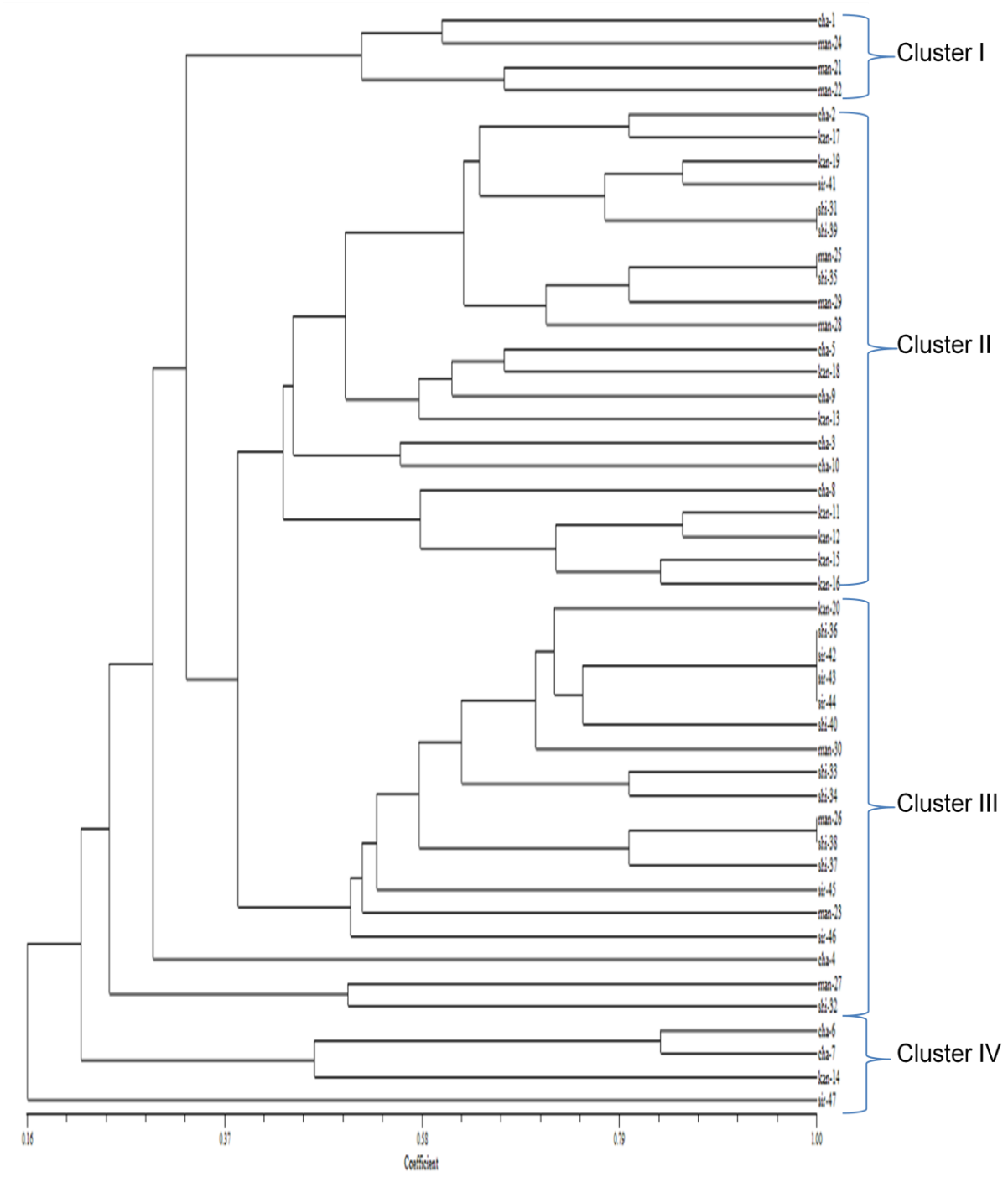

Fig. 4.6 Dendrogram showing clustering of 47 samples of five populations of *Rhododendron arboreum* constructed by using UPGMA cluster analysis of Jaccard's genetic similarity based on SSR data.

Chapter-V

DISCUSSION

Discussion

Biodiversity is the variation of taxonomic life forms within a given ecosystem, biome or for the entire earth. While the genetic diversity is the combination of different genes found within a population of a single species & the pattern of variation found within different populations of the same species. Environmental change is a continuous process & genetic diversity is required for population to evolve & adapt to such changes. Loss of genetic diversity is often associated with inbreeding & reduction in reproductive fitness (Schoettle et al., 2011). Genetic diversity is usually thought of as the amount of genetic variablity among individuals of a variety, or population of a species (Brown, 1983). The four main component of genetic diversity can be usefully measured; the number of different forms (alleles), their distribution and their effect on performance. The variation that raise genetic diversity arises from mutation & recombination. Genetic diversity is the basis for survival & adaptation & make it possible to continue & advance the adaptive processes on which evolutionary success & to some extent human survival depends. The data collected is used for the assessment of genetic diversity and that is important for the conservation of R.arboreum species in western Himalayan region.

EST-SSR has been used to assess the genetic diversity within and among the populations of sugarcane germplasm (Cordeiro et al., 2003), Hordeum vulgare L. (Meszaros et al., 2006), pigeonpea (Datta et al., 2013), Elite Chilli Pepper Lines (Dhaliwal et al., 2013), Oryza sativa L. (Garris et al., 2004), Grapes (Emanuelli et al., 2013) and many others. In present study, the polymorphic information content (PIC) value ranged from 0.2756 to 0.9212 indicate high level of genetic diversity within population due to geographic isolation of the two species. The highest PIC value present at locus RUGMS-212 & lowest PIC value present at locus RUGMS-204 with an average PIC value of 0.6326. When the PIC value compared to other studies such as in sugarcane by (Singh et al., 2013) the PIC ranged from 0.12 to 0.99, Sesame (Bin et al., 2008) average (PIC) was 0.390, ranged from 0.105 to 0.844, C*amellia* S*inensis* (Sharma et al., 2011) average (PIC) was average 0.392, ranged from 0.020 to 0.924 and many others species show low as well as high value than present study.

An average PIC value 0.6326 shows that the RUGMS markers used in the present study are very informative and it also estimate the discriminatory power of the RUGMS marker locus to reveal the genetic diversity. According to present study in *R.arboreum* the higher level of genetic diversity was observed within population (H_S = 0.6222) compared to among population (H_T =0.3436). This is in accordance to Jain et al., 2000, which found the genetic diversity of Rhododendron within population (0.40) and between the population (0.18) respectively and it could be attributable to predominant outcrossing in the species. R.arboreum is thought to be predominantly pollinated by insects, which supports high outcrossing rates as in other insect pollinated plants (Hamrick and Godt, 1996). The investigations regarding genetic diversity revealed low levels of genetic diversity within populations in case of populations of geographically restricted plant species but high levels of genetic diversity within population was observed in case of out-crossing species (Persson & Nybom,1998). When genetic diversity is classified into within and among populations genetic diversities, selfing species exhibit low levels of genetic diversity within population, but a considerable high genetic diversity among populations (Wang et al., 1996). Small, localized populations are more susceptible to genetic drift and limited gene flow as compared to widely distributed species resulting in an increased genetic diversity between populations and decreased genetic diversity within population (Hamrick and Loveless, 1984).

The conserved genome synteny can facilitate the cross-species transferability among related species/genera. The degree to which synteny can facilitate cross-species analysis of gene function will depend both on the conservation of gene order and contents, as well as on the frequency with which similar traits have a common genetic basis in different species (Tang et al., 2006)). The phylogenetically close species will share a higher proportion of microsatellites or other types of markers & show high level of amplification via the PCR process (Parida et al., 2006). In present study, the across related species/genera transferabilty rate from *R.*catawbiense to *R.arboreum* was found 16.66% i.e higher than sugarcane (Cordiero et al., 2001) was 15% only and lower than many studies as in robusta coffee (Hendre et al., 2008) was 92% transferability rate. The lower level of transferability show less homology betweeen

the related genera. It show the divergence of the lineage between the R.catawbiense to R.arboreum. The level of synteny between the species depends upon the genomic restrucuring events that occurred since their evolurtionary divergence. According to many studies EST-SSRs shows better transferability of conserved (orthologous) sequences across different species (Varshney, Graner & Sorrells, 2005). While some other studies observed the loss of sequence homology when markers developed from one species were screened on distantly related species (Sim et al., 2009). The loss of sequence homology is due to loss of specificity of the primers to the targeted sequence in distantly related species (ibid, 2009).

Breeding system and geographical range have greatly influence on the genetic variation, both at the species level as well as at population level. There is a strong relationship or association between gene diversity and distribution size (Hamrick and Godt, 1989). Widespread species having a higher level of gene diversity than more narrow region distrbuted species. Which means rare and geographically restricted plant or tree species show low level of genetic diversity (ibid,1996; Slotis et al., 1992). While several studies also show high level of genetic variation in rare species (Ranker, T.A., 1994). The association between gene diversity and breeding system with predominantly outcrossing species having a higher level of intrapopulation genetic diversity (Templeton et al., 1986). Sometimes rare or endemic species in small population may affect the level of intrapopulation variation due to fluctations in allelic frequencies that is known as genetic drift and inbreeding via selfing (Karron, J. D., 1987). Genetic drift results in decreased the variation level of within population and on the other hand increased the variation level of among population (Ellstrand and Elam,1993). Therefore, the high level of genetic diversity or variability was found intrapopulation than interpopulation in the present study. Conservation geneticists can identify the cases via the present types of studies in which variation has occur or being lost, investigate the main reasons behind this type of loss and provide suggestions and recommendations to counter the ultimate effect for their conservation.

SUMMARY

The genus Rhododendron belongs to family Ericaceae. The word Rhododendron is derived from two Greek words rhodon (rose) & Dendron (tree) meaning Rose tree. Rhododendron arboreum is an ecologically important species for the ecosystem of Himalayan region. This plant holds the Guinness Record for World Largest Rhododendron & widely popular for its medicinal benefits, edible, sacred & economic value, however very little is known about its genetic diversity. Genetic diversity refers to the variation of genes within species, that is, the heritable variation within and between populations of organisms. Genetic diversity provides the raw material for evolution by natural selection. It is an essential requirement for the improvement and conservation of the plant germplasm.

In present study the development of EST-SSR markers were designs by using conserved genes in the form of Expressed Sequence Tags (ESTs) of species Rhododendron catawbiense from the NCBI site. A total 792 (615 singletons and 177 contigs) unigenes were predicted from 1,241 publicly available EST database from NCBI (http://www.ncbi.nlm.nih.gov/) site in Rhododendron catawbiense by clustering of 2 – 33 random EST sequences. Non-redundant (NR) sequence data set represented~184.41 kb expressed genome of Rhododendron catawbiense. All 792 potential unigenes were searched for the presence of microsatellites. A total of 177 (22.3%) unigenes containing 50 SSRs with motif length ranging from 2 to 6 bp were identified. From 50 Unigenes, 36 UGMS primers were designed for the cross amplification approach from Rhododendron Catawbiense to Rhododendron arboreum samples and six pairs of microsatellite primers were produced clear PCR amplification products. These six amplified primers were used for the scoring & genotyping of the R.arboreum samples. These six primers were further used for the assessment of genetic diversity of R.arboreum in Western Himalayan region of India. The cross related species/genera transferability rate was 16.66% from R.catawbiense to R.arboreum.

The polymorphism information content (PIC) ranged from 0.2756 to 0.9212 with an average of 0.5765. The high values of polymorphism for all the primers indicated that the main cause for high genetic diversity within population was due to geographic

isolation which regulates the process of genetic diversity. The average observed heterozygosity (H_o) was 0.8666 and average expected heterozygosity (H_e) was 0.6792. The mean of effective number of alleles (Ne) was 3.13 and observed number of alleles (Na) was 4.83 compared within the population. A few alleles were rare and found in some populations with a frequency as low as 0.05 similarly few alleles were restricted to only one or two populations only. Nei's (1978) standard genetic distances and genetic identity between all pairs of populations ranged highest distance from 0.0065 to 0.7697 in population of Chamba and Mandi respectively at different locus. The genetic diversity was found with an average of 0.6222 within population and with an average of 0.3436 among the population of R. arboreum.The results revealed that high level of gentic diversity was found within the population of R.arboreum was due to out-crossing species. The results shows that their is very less homology was found between the conserved genes of R.catawbiense and R.arboreum due to wide range of geographical isolation and different lineage system.

This study was revealed the insight of abundance & distribution of microsatellite in the expressed region of the R. arboreum genome. These primers are highly polymorphic in nature & ability to differentiate the inter & intra specific levels of Rhododendron arboreum species. The main advantage of UGMS markers over other molecular markers is accurate reflection of density of SSRs in the transcribed region of the genome. It is also helpful in establishing phylogenetic relationships among different type of Rhododendron species in Indian sub-continent region especially in Western Himalayan region.

Suggestions for Further Studies

Present investigation suggests the following areas for further studies:

1. Studies can be undertaken on a large sample to have more reliable results.
2. Sample can be raised from more than one region of Indian sub-continent.
3. Investigation can be done with more topics of Rhododendron arboreum.
4. Genetic diversity analysis will be done via using other types of molecular markers.

BIBLIOGRAPHY

Akkak, A., Caser, M. and Scariot, V. (2010). Are Rhododendron hybrids distinguishable on the basis of morphology and microsatellite polymorphism? Scientia Horticulturae 469-476.

Albertini, E., Porceddu, A., Marconi, G., Barcaccia, G., Pallottini, L. and Falcinelli, M. (2003). Microsatellite-AFLP for genetic mapping of complex polyploids. Genome 824-832.

Anderson, A. J., Churchil, A. G., Autrique, E. J., Tankslay, E. M. and Sorrells, D. (1993). Optimizing parental selection for genetic linkage maps. Genome 36: 181-186.

Botstein, D., White, L. R., Skolnick, M. and Davis, W. R. (1980). Construction of a genetic linkage map in man using restriction fragment length polymorphism. American Journal of Human Genetics 32: 314-318.

Bin, L. W., Yang, H. Z., Zhan, Y. Z., Zhen, W. G. and Zhen, T. Z. (2008). Developing EST-derived microsatellite in Sesame (Sesamum Indicum L.). Acuta Agronomica Sinica 34: 2077-2084.

Balestre, M., Pinho, V., Souza, J. C., Lima, J. L., (2008). Comparison of maize simmilarity and dissimilarity genetic coefficients based on microsatellite markers. Genetics and Molecular Research 695-705.

Barbara, T., Palma, C., Paggi, M. G., Bered, F., Fay, F. M. and Lexer, C. (2007). Cross species transfer of nuclear microsatellite markers: potential and limitations. Molecular Ecology 3759-3767.

Biswas, K. M., Chai, L., Mayer, C., Xu, Q., Guo, W. and Deng, X. (2012). Exploiting BAC-end sequences for the mining,characterization and utility of new short sequences repaet (SSR) markers in citrus. Molecular Biology Reports 5373-5386.

Bombarley, A., Merchante, C., Csukasi, F., Caballero, L. J., Escobar, M. N., Portales, B. R., Botella, A. M., Blanco, M. J., Sevilla, S. F. J. and Valpuesta, V. (2010). Generation and analysis of ESTs from strawberry (Fragaria xananassa) fruits and evaluation of their utility in genetic and molecular studies. BMC Genomics 320-337.

Bushakra, M. J., Sargent, J. D., Cabrera, A., Crowhurst, R., Girona, L. E., Veasco, R., Symonds, V. V., Knaap, V. E., Troggio, M., Gardiner, E. S. and Chagne, D. (2011). Rosaceae conserved orthologous set (RosCOS) markers as a tool to assess genome synteny between Malus and Fragaria. Tree Genetics & Genomes 111-120.

Claudia, C., (2008). Molecular markers for the assessment of genetic variability in threatened plant species. Euphytica 656-678.

Caruso, M., Distefano, G., Ye, X., Malfa, L. S., Gentile, A., Tribulato, E and Roose, L. M. (2008). Generation of expressed sequence tags from carob (Ceratonia siliqua L.) flowers for gene identification and marker development. Tree Genetics & Genomes 869-879.

Chagne, D., Chaumeil, P., Ramboer, A., Collada, C., Guevara, A., Cervera, T. M., Vendramin, G. C., Gracia, V., Frigeria, M .J., Echt, C., Richardson, T. and Plonion, C. (2004). Cross-species transferability and mapping of genomic and cDNA SSRs in pines. Theoretical and Applied Genetics 1204-1214.

Curro, S., Caruso, M., Casas, L. G., Malfa, L. S. and Gentile, A. (2010). SSR and EST SSR markers for Opuntia Spp. fingerprinting and genetic diversity evaluation. ISHS Acta Horticulturae, 940: 42-54.

Chai, L., Biswas, K. M., Yi, H., Guo, W. and Deng, X. (2013). Transferability,polymorphism and effectiveness for genetic mapping of the Pummelo EST-SSR markers. Scientia Horticulturae 85-91.

Cordeiro, G. M., Pan, B. Y. and Henry, J. R. (2003). Sugarcane microsatellite for the assessment of genetic diversity in sugarcane germplasm. Plant Sciences 181-189.

Crow, J. F. and Kimura, M. (1964). The theory of genetic loads. Proceeding of the XIth International Congress of genetics 2: 495-505.

Chopra, R. N. and Chopra, C. I. (1958). Indigenous Drugs of India. Kolkata: Academic Publisher.

Chauhan, N. S. (1999). Medicinal & Aromatic plants of Himachal Pradesh. New Delhi: Indus Publishing Company.

Datta, S., Singh, P., Sahil, M., Prakash, G. P., Chaudhary, K. A., Agbagwa, O. I., and Nadarajan, N. (2013). Novel genic microsatellite markers from Cajanus scarabaeoides and their comparative efficiency in revealing genetic diversity in pigeonpea. Journal of Genetics 24-30.

Dayanandan, S., Cherry, M. and Boys, J. (2005). Microsatellite analysis reveals genetically distinct populations of red pine (Pinus resinosa Pinaceae). American Journal of Botany 833-841.

Delmas, L. E. C., Lhuillier, E., Pornon, A. and Escaravage, N. (2011). Isolation and characterization of microsatellite loci in Rhododendron Ferruginaum (Ericaceae) using pyroseqyencing technology. American Journal of Botany 120-122.

Doyle and Doyle. (1993). Isolation of plant DNA from fresh tissue. Focus 12: 13-15.

Echt, S. C., Saha, S., Deemer, L. D. and Nelson, D. C. (2011). Microsatellite DNA in genomic survey sequences and unigenes of loblolly pine. Tree Genetics & Genomes 773-780.

Ellis, J. R. and Burke, J. M. (2007). EST-SSRs as a resource for population genetics analysis. Heredity 125-132.

Ellstrand, N. C. and Elam, D. R. (1993). Population genetic consequences of small population size:Implication for plant conservation. Annual Review of Ecology and Systmatics 24: 217-242.

Ferguson, M., Rabbi, I., Kim, J. D., Gedil, M., Lavalle, L.B.A.L. and Okogbenin, E. (2012). Molecular markers and their application to Cassava breeding: Past,Present and Future. Tropical Plant Biology 95-109.

Gong, M. Y., Xu, C. S., Mao, H. W., Ho, Z. Q., Zhang, W. G., Ding, J. and Li, D. Y. (2010). Developing new SSR markers from ESTs of pea (Pisum sativum L.). Journal of Zhejiang University-Science B (Biomedicine & Biotechnology) 702-707.

Garris, J. A., Tai, H. T., Coburn, J., Kresovich, S., and McCouch, S. (2005). Genetic structure & diversity in Oryza sativa L. Genetics 169: 1631-1638.

Gong, M. Y., Xu, C. S., Mao, H. W., Hu, Z. Q., Zhang, W. G., Ding, J. and Li, Y. Z. (2010). Generation and characterization of 11 novel EST derived microsatellite from Vicia faba. American Journal of Botany 69-71.

Guo, W., Wang, W., Zhou, B. and Zhang, T. (2006). Cross species transferabillity of G.arboreum derived EST-SSRsin the diploid species of Gossypium. Theoretical and Applied Genetics 1573-1581.

Goldstein, D. B. and Pollock, D. D. (1997). Launching microsatellites:Review of mutation processes and method of phylogenetic inference. Journal of Heredity 335-342.

Graner, A., Varshney, R. K. and Sorrells, M. E. (2005). Genic microsatellite markers in plants: feature and applications. Trends in Biotechnology 1016-1022.

Hamrick, J. L. and Godt, M. J. W. (1996). Effects of life history traits on genetic diversity in plant species. Biological Sciences 1291-1298.

Hou, Y. and Lou, A. (2011). Population genetic diversity & structure of naturally isolated plant species, Rhodiola dumulosa (Crasulaceae). Public Library of Science 6: 177-189.

Hearnden, R. P., Eckermann, J. P., McMichael, L. G., Hayden, J. M., Eglinton, K. J. and Chalmers, J. K. (2007). A genetic map of 1,000 SSR and DArT markers in a wide variety barley cross. Theoretical and Applied Genetics 383-391.

Hendre, S. P., Phanindranath, R., Annapurna, V., Laremruata, A. and Aggarwal, K. R. (2008). Development of new genomic microsatellite markers from robusta coffee (Coffee canephora) showing genetic studies. BMC Plant Biology 1186-1205.

Hora, B. (1981). The Oxford Encyclopedia of Trees of the World. New York: Oxford University Press.

Ingvarsson, P. (2002). A metapopulation prespective on genetic diversity and differentiation in partially self-fertilizing plants. Evolution 2368-2373.

Jones, S. E., Sullivan, H., Bhattramakki, D. and Smith, C. S. J. (2007). A comparison of simple sequence repeat and single nucleotide marker technologies for the genotypic analysis of maize (Zea mays L.). Theoretical and Applied Genetics 361-371.

Jain, A., Pandit, M. K., Elahi, S., Bhaskar, A. and Kumar, V. (2000). Reproductive behaviour and genetic variability in geographically isolated populations of R.arboreum (Ericaceae). Current Science 79 (9):, pp. 1377-1381.

Jaccard, P. (1908). Nouvelles researcher surla distribution florale. Bulletin de la societe vaudoise des Sciences Naturelles 44: 223-270.

Keller, L. F. and Waller, D. W. (2002). Inbreeding effects in wild populations. Trends in Ecology & Evolution 17: 230-241.

Karron, J. D. (1987). The pollination ecology of co-occuring geographically restricted and widespread species of Astragalus (Fabaceae). Biological Conservation 39:179-193.

Kalia, K. R., Rai, K. M., Kalia, S., Singh, R. and Dhawan, K. A. (2011). Microsatellite markers: An overview of the recent progress in plants. Euphytica 309-334.

Lima, M. L., Garcia, A. A., Oliveria, K. M., Matsuoka, S., Arizono, H., DeSouza, C. L. and DeSouza, A. P. (2002). Analysis of genetic similarity detected by AFLP and coefficient of parentage among genotypes of Sugarcane (Saccharum spp.). Theoretical and Applied Genetics 30-38.

Lande, R. (1998). Genetics and demography in biological conservation. Science 241: 1455-1460.

Litt, M. and Lutty, 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: 397-401.

Morgante, M. and Oliviera , A.M. (1993). PCR-amplified microsatellite as markers in plant genetics. The Plant Journal 175-182.

Ma, Q. J., Zhou, H. Y., Ma, L. C., Yao, Z. M., Jin, Q. J., Wang, C. X. and Chen, L. (2010). Identification and characterization of 74 novel polymorphic EST-SSR markers in the tea plant (Camellia Sinennsis L.). American Journal of Botany 153-156.

Mariette, S., Chagne, D., Patrick, L. C., Raffin, A., Plomion, C. and Kremer, A. (2001). Genetic diversity within and among Pinus pinaster populations:comparisons between AFLP and microsatellite markers. Heredity 469-479.

Moretzsohn, M. Barbosa, VG. A., Freitas, A. D., Teixeira, C., Bertoli, L. S., Guimaraes, M. P., Pereira, W. R., Lopes, R. C., Carvallari, M. M., Valles, J. and Gimenes, A. (2009). A linkage map for the B-genome of Arachis (Fabaceae) and its synteny to the A genome. BMC Plant Biology 733-739.

Milne, R. I. and Abbott, R. I. (2000). Origin and Evolution of invasive naturalised material of Rhododendron ponticum L. in the British Isles. Molecular Ecology 541-556.

Nunziata, A., Ruggieri, V., Greco, N., Frusciante, L. and Barone, A. (2010). Genetic diversity within wild potato species (Solanum spp.) revealed by AFLP and SCAR markers. American Journal of Plant Sciences 95-103.

Nayar, M. P. and Shastry, A. (1987). Key words to the taxonomy of flowering plants of India. Calcutta: Botanical Survey of India.

Nei, M. (1972). Genetic distance between populations. American Naturalist 106: 1943-1957.

Nei, M. (1978). Estimation of average heterozygosity and genetic distance from a small number of individuals. Genetics 147: 1943-1957.

Orwa, C., Mutua, A., Kindt, R., Jamnadass, R. and Simons, A. (2009). Agroforest tree database. Dehradun: Forest department.

Parida, K. S., Kumar, R. A. K., Dalal, V., Singh, K. N. and Mohapatra, T. (2006). Unigene derived microsatellite markers for the cereal genomes. Theoretical and Applied Genetics 808-817.

Park, H. Y., Ahn, G. S., Choi, M. Y., Oh, J. H., Ahn, C. D., Kim, G. J., Kang, S. J., Choi, W. Y. and Jeong, R. B. (2010). Rosa (Rosahybrida L.) EST-derived microsatellite markers and their transferability to strawberry (fragaria spp.). Scientia Horticulturae 733-739.

Paul, A., Khan, L. M., Arunachalam, A. and Arunachalam, K. (2005). Biodiversity and conservation of Rhododendrons in Arunachal pradesh in the Indo-Burma biodiversity hotspot. Current Science 52-63.

Provan, J., Powell, W. and Hollingsworth, P. M. (2001). Chlorplast microsatellite: new tools for studies in plant ecology and evolution. Trends in Ecology & Evolution 142-147.

Pradhan, U. C. and Lachungpa, T. C. (1990). Sikkim Himalayan Rhododendrons. Kalimpong: Primulaceae books.

Paden, D. W., Rayburn, A. L. and Iqbal, M. J. (1995). Clonal stability of RAPD markers in three Rhododendron species. Journal of Environmental Horticulture 43-46.

Riek, J. D., Bockstaele, E. V. and Keysar, E. D. (2009). Discovery of species-wide EST-derived markers in Rhododendron by intron flanking primer design. Molecular Breeding 171-178.

Rai, L. K., Gurung, B. and Singh, K. K. (2009). Conservation of Rhododendron in Sikkim Himalaya: An overview. World Journal of Agricultural Sciences 284-296.

Ram, G. S., Parthiban, T. K., Kumar, S. R., Thiruvengadam, V. and Paramathma, M. (2008). Genetic diversity among Jatropha species as revealed by RAPD markers. Genetic Resources and Crop Evolution 803-809.

Rodgers, W. A. and Panwar, H. S. (2003). Planning a wildlife protected area network in India. Molecular Ecology 267-339

Ranker, T. A. (1994.). Evolution of high genetic variability in the rare Hawaiian fern Adenophorus periens and implications for conservation management. Biological Conservation 70: 19-24.

Rowe, B. D., Blazich, A. F., Warren, L. S. and Ranney, G. T. (1994). Seed germination of three provinces of Rhododendron catawbiense: Influence of light and temperature. Journal of Environmental Horticulture 12(3): 155-158.

Singh, K. R., Jena, N. S., Khan, S., Yadav, S., Banarjee, N., Raghuvanshi, S., Bhardwaj, V., Dattamajumder, K. S., Kapur, R., Solomon, S., Swapna, M., Srivastava, S. and Tyagi, K. A. (2013). Development & cross species/genera transferability of novel EST-SSR marker & their utility in revealing population structure & genetic diversity in Sugarcane. Gene 524:309-329.

Simko, I. (2009). Development of EST-SSR markers for the study of population structure in Lettuce (Lactuca sativa L.). Journal of Heredity 256-262.

Singh, K. K. and Gurung, B. (2009). In vitro propagation of R. maddeni Hook. F. an endangered Rhododendron species of Sikkim Himalayan. Notulae Botanicae Horti Agrobotanici Cluj-Napoca 79-83.

Sharma, H., Kumar, R., Sharma, V., Kumar, V., Bhardwaj, P., Ahuja, S. P. and Sharma, K. R. (2011). Identification and cross-species transferability of 112 novel unigene-derived microsatellite markers in tea (Camellia Sinensis). American Journal of Botany 33-38.

Singh, A., Chaudhaury, A., Srivastava, S. P. and Lakshmikumaran, M. (2002). Comparison of AFLP and SAMPL markers for assessment of intra-population genetic variation in Azadirachta indica Plant Science 17-25.

Slotis, P. S., Soltis, D. E. and Norvell T. L., (1997). Genetic diversity in rare and widespread species of Lomatium (Apiaceae). Madrono 44: 59-73.

Sim, C. S., Yu, K. J., Jo, K. Y., Sorrells, E. M. and Jung, G. (2009). Transferability of cereal EST-SSR markers to ryegrass. Genome 52: 431-437.

Singh, K. R., Singh, B. R., Singh, P. S. and Sharma, L. M. (2012). Genes tagging and molecular diversity of red rot suspectible/tolerant sugarcane hybrids using c-DNA and unigene derived markers. World Journal of Microbial Biotechnology 1669-1679.

Supriya, A., Senthilvel, S., Nepolean, T., Eshwar, K., Rajaram, V., Shaw, R., Hash, T. C., Killian, A., Yadav, C. R. and Narasu, L. M. (2011). Development of a molecular linkage map of pearl millet integrating DArT and SSR markers. Theoretical and Applied Genet ics 239-250.

Szira, F., Andreas, B., Kerstin, N., Khalil, N., Gabor, G. and Andras, B. (2011). Could EST-based markers be used for the marker-assisted selection of drought tolarent barley (Hordeum vulgare) lines? Euphytica 373-391.

Sastry, A. R. K. and Hajra, P. K. (2010.). Rare and endemic species of Rhododendron in India. Kolkatta.: Botanical survey of India.

Slatkin, M. (1995). A measure of population subdivision based on microsatellite allele frequencies. Genetics 139: 457-462.

Tamura, K., Dudley, J., Nei, M. and Kumar, S. (2007). Molecular Evolutionary Genetics Analysis (MEGA) Software Version 4.0. Molecular Biology and Evolution 1596-1599.

Templeton, R. A., Hemmer, H., Mace, G., Seal, S. U., Shields, M. W. and Woodruff, S. D. (1986). Local adaptation, coadaptation and population boundaries. Zoo Biology 5: 115-125.

Tienderen, V. H. P., Haan, A. A., Linden, V. D. G. and Vosman, B. (2002). Biodiversity assessment using markers for ecologically important traits. Trends in Ecology and Evolution 17:577-582.

Varshney, K. R., Sigmund, R., Borner, A., Korzun, V., Stein, N., Sorrells, E. M., Langridge, P. and Graner, A. (2005). Interspecific transferability and comparative mapping of barley EST markers in wheat, rye and rice. Plant Science 195-202.

Varshney, K. R., Mahendar, T., Aggarwal, K. R. and Borner, A. (2007). Genic molecular markers in plants: Development and Applications. Genomics Approaches and Platforms 13-29.

Wang, C. X., Guo, L., Shangguan, F. L., Wang, C., Yang, G., Qu, C. S. and Fang, G. J. (2012). Analysis of expressed sequence tags from grapevine flower and fruit and development of simple sequence repeat markers. Molecular Biology Reports 6825-6834.

Wang, S., Wang, X., He, Q., Liu, X., Li, L., Gao, J. and Wang, F. (2012). Transcriptome analysis of the roots at early and late stages using Illumina paired-end sequencing and development of EST-SSR markers in radish. Plant Cell Reports 1233-1245.

Wright, S. (1951). The genetical structure of populations. Annals of Eugenetics 139: 457-462.

Wayne, K. R., Morin, A. P., Luikart, G. and Jolla, L. (2004). SNPs in ecology, evolution & conservation. Trends in Ecology & Evolution 19: 208-216.

Yeh, F. C., Yang, F. C., Boyle, T. B. Z., Ye, H. Z. and Mao, X. J. (1997). POPGENE, the user-friendly shareware for population genetic analysis. Molecular Biology and Biotechnology Centre, . University of Alberta, Edmonton, Canada.

Yadav, D.K., Mohapatra, T. and Parida, S. K. (2010). Microsatellite in Brassica unigenes: Relative, abundance, marker design, use in comparative physical mapping and genome analysis. Genome 55-67.

Yadav, K. H., Ranjan, A., Asif, M., Mantri, S., Sawant, S. and Tuli, R. (2011). EST-derived SSR markers in Jatropha curcas L. development, characterization, polymorphism and transferability across the species/genera. Tree Genetics & Genome 207-219.

Yu, K. I., Rota, L. M., Kantety, V. R. and Sorrells, E. M. (2004). EST derived SSR markers for comparative mapping in wheat and rice. Molecular Genetics and Genomics 742-751.

Zhang, J., Liang, S., Duan, J., Wang, J., Chen, S., Cheng, Z., Zhang, Q., Liang, X. and Li, Y. (2012). De Novo assembly & characterization of the transcriptome during seed development, & generation of genic SSR-markers in Peanut (Arachis hypogaea L.). BMC Genomic 13: 90-95

APPENDICES

Appendix-A

Details of DNA isolation protocol:

DNA was isolated following the protocol of Doyle and Doyle (1993); as:

- 2-4g of plant material were homogenized in liquid nitrogen with help of pre-cooled mortar and pestles and mixed with pre-heated DEB at 65^{o}C.
- The powder was transferred to 50 ml polypropylene centrifuge tube containing 10 ml of pre-warmed (65^{o}C) DNA extraction buffer and suspension was incubated for I hour at 65^{o}C and cool it to room temperature.
- The mixture was emulsified with an equal volume of chloroform: Iso-amylalcohol (24:1) for 5 min by gentle inversion.
- Make the final volume in each polypropylene centrifuge tube equal by adding CIA.
- The mixture was centrifuged at 15,000 rpm for 30 min.
- The aqueous phase (supernatant) was transferred to a fresh centrifuge tube with a wide bore pipette and add 2/3 volume of iso-propanol to it by quick and gentle inversion.
- The precipitated DNA was spooled out using disposable pipette tip.
- Washed twice with 70% alcohol at 13,000rpm for 5 min.
- The pellet was dried under vacuum and dissolved in 1 ml of $T_{10}E_1$ buffer and incubated at 37ºC.

Details of DNA purification protocol:

- Purification of DNA is needed to remove RNA, proteins and polysaccharides, which are considered to be the major contaminants in the DNA precipitates. RNA was removed by RNAse treatments and proteins were removed by phenol-chloroform extraction.
- 10µl RNase A (1µg/1µl) is added to the DNA sample (500µl) and incubated at 37°C for 1 hr.
- After 1 hr equal volume of phenol: chloroform: iso-amyl alcohol (25:24:1) was added and tube were spun at 10,000 rpm for 5 min at the room temperature. The aqueous phase was separated into fresh micro centrifuge tube.
- The upper aqueous phase was collected and extracted with chloroform–Isoamyl alcohol for the second time.
- Extracted aqueous phase is further purified by adding equal volume of Chloroform: Isoamyl (24:1) and centrifuged at10, 000 rpm for 5 min.
- The above step is repeated again.
- The separated upper aqueous phase was collected after centrifugation and mixed with 1/10th volume of 3 M sodium acetate.
- DNA was precipitated by adding two volumes of absolute alcohol, pelleted by centrifugation at 13,000 rpm for 3 minutes.
- The pellet is then washed with 70% alcohol twice at 10,000 for 3 min. and dried in vacuum and dissolved in TE buffer.

APPENDIX-B

List of 30 primers with their sequence, length, annealing Temperature, Expected size and Number of ESTs included:

S. No	Primer No.	Primer sequence	Total length (bp)	Temp-annealing	Expected size(bp)	No. of Contributing ESTs
1	16	F 5'- GCCCATTTGGGTTTTAGCTT R 5'- TGGCCCACTTCTTTTACTGG	20 20	50-55°c	850-900	2
2	35	F 5'- CGCTTTGTTTCCAACAACAAT R 5'- TTGATGATGTCGTCCAGCAC	21 20	50-55°c	275-298	3
3	49	F 5'- AGTCACCAGTTCCGGCTAGA R 5'- ACTTTTCCCAACGCTCCTCT	20 20	50-55°c	1286-1316	10
4	60	F 5'- TTGAGCATGCGAGAAGAGAA R 5'- AGGCTTTGCTTTTCCACCTT	20 20	50-55°c	258-291	2
5	95	F 5'- GGCAATGAGGGACATGAGAT R 5'- ACGAGGCTTCTCCTTCTTCC	20 20	50-55°c	656-693	2
6	112	F 5'- GACCAGTGACGGGTTGAGTT R 5'- CCACACAAAAATTCCCCAGA	20 20	50-55°c	544-563	2
7	131	F 5'- CCACCTCATGAGTGTATTTTTGG R 5'- GTTAATGGGTGGAGCATTGG	23 20	50-55°c	140-240	2
8	137	F 5'- GGCCGTGGTGGTAGACAT R 5'- TGGGTATCAGTAGAAATACCTCGT	18 24	50-55°c	400-536	5
9	146	F 5'-TTTTCCTACAATATAAGCTGTAACTGA R 5'- TGGTAGGCTACGCGAGAGAT	27 20	50-55°c	78-104	2
10	156	F 5'- ATCCGATTCCGGTTGCTTAT R 5'- AATCAGTGGCAATGGAGGAG	20 20	50-55°c	112-152	2
11	292	F 5'- TGCAGACATTCACAAATCAGAA R 5'- CTTGGCTTTTGCTTCCCTTA	22 20	50-55°c	156-191	1
12	301	F 5'- CCTTGATCGTCGCTATCCTC R 5'- CCAGAAGCTCGTCGAGTACC	20 20	50-55°c	403-447	1
13	324	F 5'- GCAACCTCCCAATTTCACAT R 5'- GTGAGGAAGCTCAAGGCTGT	20 20	50-55°c	76-120	1
14	325	F 5'- TGGATGTGGTTGTTGTTGCT R 5'- ATTACCCACCTCCCTCGAAA	20 20	50-55°c	393-412	1

Table continued……..

15	344	F 5'- GATGGACTTCTTCCCGACAA R 5'- GATGGACTTCTTCCCGACAA	20 20	50-55°c	103-152	1
16	347	F5'- GCTCCCCAACAGAAGGAGAT R5'- ACAGAGGGTTTAGGCCTGGT	20 20	50-55°c	100-150	1
17	395	F5'- GCGTGGCCAACTATTCATAA R5'- ACCCATCTTTCCCTTTCACC	20 20	50-55°c	110-200	1
18	418	F5'- CGCTTCGAAGTTCAAACCAT R5'- GGGACCCATTTCTCTTCCTC	20 20	50-55°c	150-200	1
19	507	F5'- AACGACATCGATTTGGAAGC R5'- GTGGAGTTGCGTTGGAGAAT	20 20	50-55°c	120-150	1
20	522	F5'- CGCGCGAAGAATTGAAAAT R5'- AATAGATTTGAGATTTATCTGTGGAAA	19 27	50-55°c	100-160	1
21	540	F5'- TATGGCTCAGACCATGTTGC R5'- CTTTATTGGGGCAGCCTTCT	20 20	50-55°c	150-180	1
22	571	F5'- AGCGTTTGGTTAGGAGCGTA R5'- AGCTTGGCCATTGATGTACC	20 20	50-55°c	100-150	1
23	601	F5'- AGGAGCTGCGAGGAATACAA R5'- CTGCGTCACTGAACCCACTA	20 20	50-55°c	340-400	1
24	622	F5'- CCTCCTTCATCTTCCCCTCT R5'- GAATCTTCATCTCCGCTCCA	20 20	50-55°c	298-381	1
25	656	F5'- GGAGACCGATCGACAAGAAC R5'- AACAAAACAGCAGCAGCAGA	20 20	50-55°c	309-346	1
26	691	F5'- ACATGGCCAACTGCCTAATC R5'- AAAGGAGGGCTGGTAATCGT	20 20	50-55°c	401-428	1
27	716	F5'- AGGCTCTTCCCCTCTGTCTC R5'- TTGAGGCGATTCAAGGTAGG	20 20	50-55°c	100-140	1
28	727	F5'- CGAATCCCAACCTTCTCAAC R5'- AGGACCAGCCAGAAGATTGA	20 20	50-55°c	145-175	1
29	781	F5'- CAAACAACGCCCATAGCTG R5'- GGTCGGTACACGTTCGATCT	19 20	50-55°c	100-150	1
30	393	F5'- CACTCCTCAACAGCCCCTAC R5'- GCCCTAGAATCCGAAACTCC	20 20	50-55°c	100-150	1

Table no. 4.4: Marker validation and features of new 6 UGMS markers of Rhododendron arboreum

Locus name	Primer sequence	Repeat motif	Annealing temperature	No. of alleles	Heterozygosity		PIC	Approximate size range(bp)	No. of genotypes amplified
					H_o	H_e			
RUGMS-619	F5'- CCTTGAATGGAGGTGGAGAA R5'- CCAACCAGCAAATAGGCATT	$(CCA)_{20}$	50°C	5	0.8000	0.7109	0.3463	200-300bp	38
RUGMS-212	F 5'- TGGCTAGCCCAAGCACTAAT R 5'- CCAAATGGAAATGGGTTTTG	$(TTC)_9$	50°C	5	0.9250	0.6572	0.9212	250-300bp	19
RUGMS-282	F 5'- TCTACTACTGCGCCAAAGCA R 5'- ATTGGTCCCCTTCTGGTGAT	$(GA)_{11}$	50°C	5	0.8500	0.7012	0.5885	300-400bp	28
RUGMS-320	F 5'- CTCCGTTGTGTTTGGTTGAG R 5'-TTCTCGAAGATAGGGGCAGA	$(TTTAG)_6$	50°C	5	0.8250	0.7050	0.4999	300-400bp	33
RUGMS-204	F 5'- GATGCTTCAAAACCCCAAAA R 5'- CCGAAGGGTTCATGAAGAAA	$(CCG)_8$	50°C	4	0.9000	0.6338	0.2756	250-300bp	40
RUGMS-39	F 5'- GAATTGATTGCTTGGGGAAA R 5'- TTCGGGTTTCAACAGCTAGG	$(CTG)_2$	50°C	5	0.9000	0.6672	0.8279	300-500bp	19