Evaluation of Allelic Variation in *Cedrus deodara* (Roxb.) G. Don to identify Region Specific Nuclear Microsatellite Markers (nSSRs) in Uttarakhand, India

Kumar Akhilesh*, Barthwal Santan and Ginwal H.S.
Division of Genetics and Tree Propagation, Forest Research Institute, Dehradun-248195, INDIA
*akhibiochemist06@gmail.com

Abstract

Cedrus deodara is an excellent timber species extensively prone to illicit felling and smuggling of timber. Knowledge of molecular marker techniques can help in tracing the origin of the stolen timber back to the forest area for legal purposes. The present study demonstrates the use of nSSR markers for timber tracking in Cedrus deodara. In this research twenty-five individual trees were randomly selected from two geographically distinct populations Kathiyan and Jageshwar forest of Uttarakhand. Out of 47 nuclear SSR primers, only nine showed positive amplification out of which five (pdms009, pm01, pm05, SSRPtctg4698 and RPTest9) were finally selected to reveal allelic variation. Total of 10 alleles of individuals were detected at all five loci for both populations. For Kathiyan, population effective number of alleles ranged from 1.26 to 2 while for Jageshwar, it ranged from 1.14 to 1.98. Expected heterozygosity (H_E) for Chakrata population ranged from 0.21 to 0.51 while for Jageshwar it was between 0.13 to 0.5.

Analysis of molecular variance (AMOVA) shows that most of the variations were found within a population (97.81 %) as compared to among the population (2.19 %) with an F_{ST} value of 0.02. In this research, we found that the same primer pairs gave a different product length for both geographically distinct populations of C. deodara, for Kathiyan population locus pdms009, pm01, pm05, SSRPtctg4698 and RPTest9 product lengths for both alleles a and b were 197 and 125, 304 and 221,402 and 317,261 and 208, 332 and 246 base pair respectively, while for Jageshwar population, it were 339 and 238, 319 and 203, 456 and 283, 245 and 197, 309 and 235 base pair respectively.

Keywords: *Cedrus deodara,* nSSR, AMOVA, F_{ST}, Allelic variation.

Introduction

C. deodara is native to the western Himalayas in North-Central India. It occurs at 1200-3050 m altitudes where the mean annual temperature is 12–17ºC and the mean annual rainfall is 200-1800 mm[11]. Deodar is mainly known for its high valued timber, hard and durable wood obtained from it which is white to light yellowish-brown in colour with a characteristic odor and oily feel. The wood of deodar is very strong with an average weight of 560 kg/m^3. The deodar wood is extensively used for building, carriage, railway wagon, house building, beams, floorboards, door and window frames, furniture and general carpentry.

It also produces quality plywood[10]. *C. deodara* has been proven to have great pharmacological potential with great utility and usage as folklore medicine[4]. *C. deodara* possesses many qualities including anti-inflammatory, antitumor, immunomodulatory properties as well as exerting an influence on the nervous system. The oleoresin and the dark coloured oil obtained from the wood of deodar are valued for their application for ulcers and skin diseases[13]. The multi-purpose usefulness of this species has led to its overexploitation in the recent decades.

The main objective of the present research is to explore the ability and potential of available DNA-based techniques for tracing timber to its origin. Nowadays, illegal logging is one of the main reasons for deforestation in natural forests and is the cause of high ecological and economic damage.

Material and Methods

The plant materials for the present research were collected from the two forest sites viz. Chakrata forest division (Kathiyan) and Civil Soyam Almora forest division (Jageshwar), Uttarakhand, India. These two forest sites are widely separated from each other (Figure 1).

Plant Material: Needles were collected from 25 individuals per population, which were sampled randomly from 50 meter distance. The samples were properly labelled and stored in the freezer at -80ºC prior to DNA isolation.

DNA isolation and PCR amplification: The total genomic DNA was extracted from young needles of *C. deodara* by using a modified protocol of the CTAB method[7]. The extraction buffer consisted of 100mM Tris-HCl (pH-8.0), 20mM ethylenediaminetetraacetic acid (EDTA) (pH-8.0), 1.4mM NaCl, 3% (w/v) CTAB, 2.5% (w/v) polyvinylpyrrolidone (PVP), 5mM ascorbic acid and 0.2% (v/v) β-mercaptoethanol. The DNA concentration and purity ware quantified by Biophotometer at A_{260}/A_{280} nm wavelength and the purity of DNA was checked by running the samples on 0.8% agarose gel. For PCR reaction, the genomic DNA was diluted to a final concentration of 15ng/µl. Microsatellite screening and PCR amplification were performed with a 20µl reaction mixture containing 1X

Taq buffer, 3mM $MgCl_2$, 0.2mM each of dNTPs, 0.2µM both forward and reverse primers and 1 unit of Taq polymerase and 1µl template DNA.

The amplification was carried out using 1cycle of 95°C for 5 minutes, 35 cycles of 1 minute at 94°C, annealing temperature was varied between 58-58.5°C for 1 minute, extension temperature was 72°C for 1 minute and final extension was 1cycle of 8 minutes at 72°C. The amplified PCR products were separated in 2% (w/v) metaphor agarose gel with 1X TBE buffer, stained with ethidium bromide as an intercalating agent. The amplified product with 3µl of 10X agarose gel loading dye was loaded onto the horizontal gel electrophoresis apparatus. The gel was running at the constant voltage of 5V/cm of the length of the gel till the bands get separated. After the 3/4th running of the gel, it was photographed with a gel documentation imaging system. The size of the amplified band was compared with the 100 base pair DNA ladder which was used as a marker to know the amplified DNA fragments.

Determination of product length of amplified allele: 100 base pair ladders were plotted against migration distances of the ladders and with the help of MS excel 'chart wizard', a trend line was applied to the scatter plot. Polynomial curves with powers from two to four were used to produce the closest fit to the marker curve, but higher polynomial powers led to a partial misfit of the trend line to the curve. The polynomial coefficients for calculating the formula were derived from the regression equation of the trend line displayed on the same chart[8]. The formula was then used to calculate the relative mobility of the amplified DNA fragments in all the individuals representing both site populations and convert migration distances of DNA fragments into molecular sizes. For accuracy, molecular marker distance migrations were used to back calculate the marker sizes.

Nuclear SSR analysis for population specific markers study: Geographically distinct populations existing in nature differ in their allele/gene frequencies and they are often of interest to determine the degree to which genetic variation in a region is distributed within and among populations. This information is found useful for understanding the degree to which migration counters population subdivision due to selection genetic drift[15]. Allele/ gene diversity over all the 5 nuclear SSR loci was assessed by calculating the observed number of alleles (na), effective number of alleles (ne), observed heterozygosity (H_O) and expected heterozygosity (H_E) by using POPGENE version1.32[9,17]. The molecular weight of both alleles was also calculated and compared.

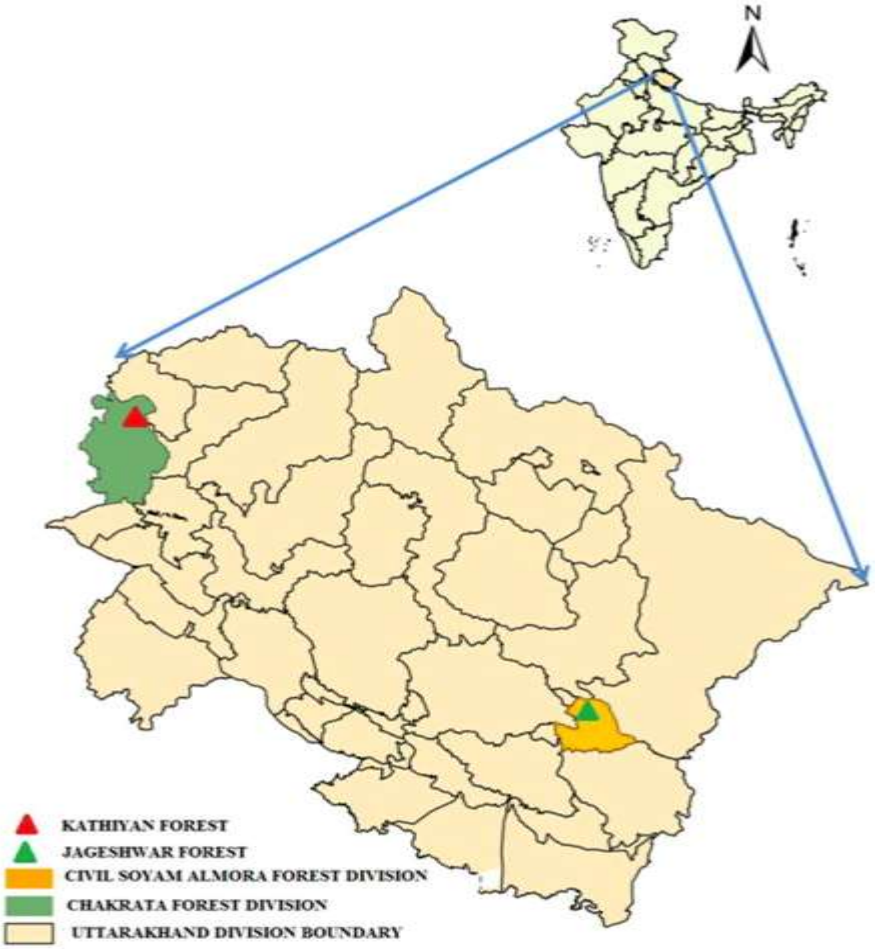


Figure 1: Map showing the location where sample collection was done.

AMOVA analysis: The population genetic structure was inferred by an analysis of variance framework (AMOVA) using the Arlequin software version 3.11[2,3]. The type of hierarchial AMOVA implemented here was with genotypic data, one group of populations and number within the individual level. This technique treats genetic distances as deviations from a group mean position and squared deviations as variances. The significance of covariance components was tested using 1000 permutations. The total sum of squares of genetic distances can then be partitioned into components that represent the within population and among population mean squares.

Results and Discussion

In this research, we mainly focused on identifying population specific molecular markers. That means those markers which can help in relating individuals to its population. Such markers would be useful for tracing timber origin in case of illegal logging of trees. Illegal logging is one of the main factors contributing to forest destruction[1]. In this study, five nuclear microsatellite primer pairs were used to amplify under standard conditions and all the 5 nSSRs loci were found polymorphic in *C. deodara*.

Characterization of populations using nuclear SSRs: Polymorphism involves the existence of different forms (alleles) of the same gene in plants or a population of plants. These differences are tracked as molecular markers to identify desired genes and the resulting traits. For population specific marker study, nuclear SSR markers were used. Out of 47 nuclear SSR primers, only nine showed positive amplification out of which five were finally selected based on polymorphism. Figure 2 shows the allelic variation detected at locus RPTest9 for Chakrata and Jageshwar populations. Details of nuclear primer pairs used in the study are given in table 1.

For the Chakrata population, a total of 10 alleles was detected at all the five loci with a mean of 2. The effective number of alleles ranged from 1.26 to 2 with a mean of 1.84 ± 0.32. Expected heterozygosity (H_E) for the Chakrata population ranged from 0.21 to 0.51 with a mean of 0.45 ± 0.13. Observed heterozygosity (H_O) ranged from 0.14 to 0.95 with 0.72 ± 0.34 at marker level. Minimum observed heterozygosity was observed at locus pm05 while maximum observed heterozygosity was observed at locus RPTest9 (Table 2).

For the Jageshwar population also, 10 alleles were detected at all the five loci with a mean of 2. The effective number of alleles ranged from 1.14 to 1.98 with a mean of 1.79 ± 0.36. Expected heterozygosity (H_E) for the Jageshwar population ranged from 0.13 to 0.50 with a mean of 0.42 ± 0.16. Observed heterozygosity (H_O) ranged from 0.13 to 0.91 with 0.69 ± 0.31. Minimum observed heterozygosity was observed at locus pm05 while maximum observed heterozygosity was observed at locus RPTest9 (Table 3).

Table 1
List of nuclear microsatellite primer pairs used in study.

Locus	Sequence (5'-3')	Annealing Temperature (°C)	Number of repeat units
pdms009[15]	F-CAATGAGTAGAAGATCATGGTGG R- CTAGGGAGCCGCATTTACAC	60	$(CT)_{31}(CA)_{20}$
pm01[10]	F- AGAGAAGGCACGATTTTGTC R- TCCCACTAATCACTTTGAAAG	56	$(TG)_{12}$
pm05[10]	F- GAGTCTAATTGCAAACCCCA R-TGGAGATCTACCACTTTTTC	52	$(TG)_9$
SSRPtctg4698[18]	F- CGAAAAGGTGGTTCTGATGG R- TTTTCCGCTGGATTTACCAC	49	ATC
RPTest9[18]	F- CCAGACAACCCAAATGAAGG R- GCCTGCTATCGAATCCAGAA	51	AGC

Table 2
Summary of genetic variation and heterozygosity statistics for all loci of Chakrata population

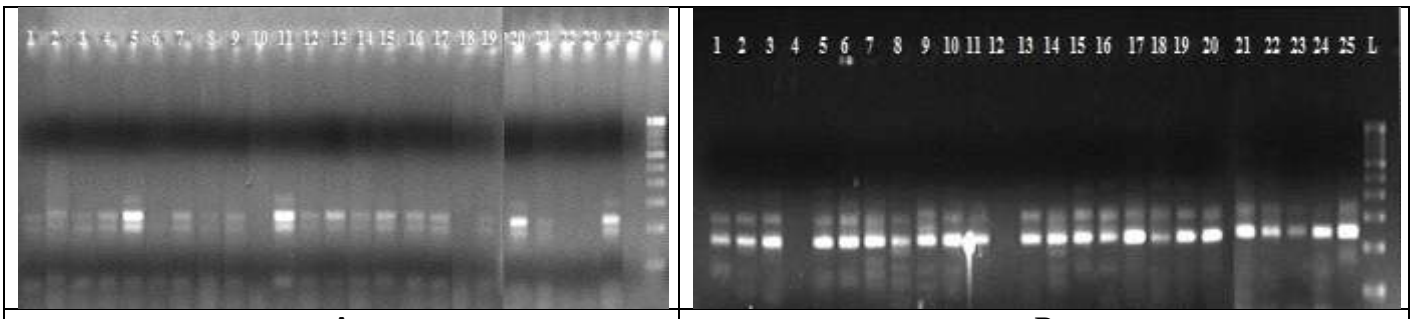
Locus	Na	ne	H_O	H_E
pdms009	2.00	1.99	0.93	0.51
pm01	2.00	1.99	0.94	0.51
pm05	2.00	1.26	0.14	0.21
SSRPtctg4698	2.00	2.00	0.66	0.51
RPTest9	2.00	1.99	0.95	0.50
Mean	**2**	**1.84**	**0.72**	**0.45**
St. Dev.		0.32	0.34	0.13

St. Dev.= Standard deviation, Na= Observed number of allele, ne= Effective number of alleles, H_O= Observed heterozygosity, H_E= Expected heterozygosity

The percent of polymorphic loci was 100%. Heterozygosity describes to us a great deal about the structure and even the history of a population and plays an important role in genetic applications in tree breeding as well as gene conservation[19]. The expected heterozygosity of a population represents the chance that two copies of a locus that are sampled randomly from a gene pool have different allelic states. It is biologically the most meaningful measure of genetic diversity[12].

Variation in nuclear SSR for identification of population specific markers: After the estimation of the product length of both alleles, we found the same primer pairs give a different product length for both geographically different populations of *C. deodara.* For Chakrata population locus pdms009, pm01, pm05, SSRPtctg4698 and RPTest9 product lengths for both alleles a and b were 197 and 125, 304 and 221,402 and 317,261 and 208, 332 and 246 base pair respectively while of Jageshwar population, they were 339 and 238, 319 and 203, 456 and 283, 245 and 197, 309 and 235 base pair respectively (Table 4). These results suggest that environmental factors can affect genetic/allelic variation among populations[5].

Analysis of molecular variance (AMOVA) using nuclear SSR markers: The hierarchical level of molecular divergence among populations was confirmed by analysis of molecular variance. In AMOVA, the percentage of variation among the population was 2.19 and genetic differentiation (F_{ST}) was 0.02 (Table 5). The AMOVA analysis revealed significant differences ($p<0.001$) among the populations and among individuals within a population. The AMOVA analysis shows that most of the variation in *C. deodara* lies within the population, a result compatible with those from cpSSR study in *C. deodara*[6].



A **B**

Figure 2: nSSR profile of *C. deodara* genotype (Lane 1-25) generated using primer RPTest9 for (A) Chakrata and (B) Jageshwar population. Photograph showing the amplified polymorphic bands. Lane L is the 100 base pair DNA ladder.

Table 3
Summary of genetic variation and heterozygosity statistics for all loci of Jageshwar population

Locus	na	ne	H_O	H_E
pdms009	2.00	1.96	0.86	0.50
pm01	2.00	1.94	0.82	0.49
pm05	2.00	1.14	0.13	0.13
SSRPtctg4698	2.00	1.95	0.73	0.50
RPTest9	2.00	1.98	0.91	0.50
Mean	**2**	**1.79**	**0.69**	**0.42**
St. Dev.		0.36	0.31	0.16

St. Dev.= Standard deviation, Na= Observed number of allele, ne= Effective number of alleles, H_O= Observed heterozygosity, H_E= Expected heterozygosity

Table 4
Product length (base pair) for Chakrata and Jageshwar population

Locus name	Chakrata forest		Jageshwar forest	
	allele a	allele b	allele a	allele b
pdms009	197	125	339	238
pm01	304	221	319	203
pm05	402	317	456	283
SSRPtctg4698	261	208	245	197
RPTest9	332	246	309	235

Table 5
Analysis of molecular variance (AMOVA) based on geographical regions

Source of Variation	d.f.	S.S.	Variance components	% of Variation	F_{ST}	*P Value*
Among populations	1	0.72	0.01	2.19	0.02	0.000
Within populations	48	22.16	0.46	97.81		
Total	49	22.88	0.47			

d.f. = Degree of freedom, S.S. = Sum of squares, F_{ST}= Fixation indices, Significant at level of p<0.001

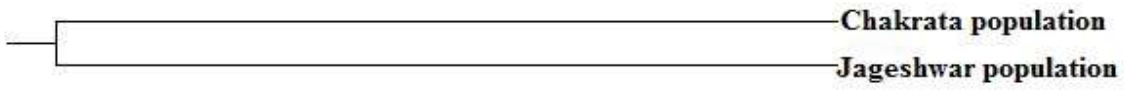


Figure 3: UPGMA dendrogram showing relationship between populations of *C. deodara.*

Genetic relationship between populations with nSSR markers: The genetic relationship between the populations based on a dendrogram showing the similarities between populations was established using the Unweighted Pair-Group method with Arithmetic Averages (UPGMA). The dendrogram based on Nei's genetic identity discriminated that both populations are grouped into one cluster (Figure 3).

Conclusion

Cedrus deodara is an important conifer of the western Himalayas. It has a high economic value and due to its high demand, the natural population of this species is gradually declining due to illegal felling and smuggling for timber. In this study, we reported five nuclear microsatellite markers (pdms009, pm01, pm05, SSRPtctg4698 and RPTest9) which gave different banding patterns for two different populations. The development of markers specific to populations would help in the prevention of illegal felling and smuggling of the timber from the forest area.

Acknowledgement

The authors are grateful and thank the Director, Forest Research Institute, Dehradun for providing the facilities to conduct the research work and to the Forest department, Uttarakhand for their priceless help.

References

1. Brack D., Illegal logging and the illegal trade in forest and timber products, *International Forestry Review*, **5(3),** 195-198 **(2003)**

2. Excoffier L., Laval G. and Schneider S., Arlequin ver. 3.0: An integrated software package for population genetics data analysis, *Evolutionary Bioinformatics Online*, **1,** 47-50 **(2005)**

3. Excoffier L., Smouse P.E. and Quattro J.M.., Analysis of molecular variance inferred from metric distances among DNA haplotypes; Application to human mitochondrial DNA restriction data, *Genetics*, **131,** 479-491 **(1992)**

4. Gupta S., Walia A. and Malan R., Phytochemistry and Pharmacology of *Cedrusdeodera*: an overview, *International Journal of Pharmaceutical Sciences and Research*, **2(8),** 2012-2018 **(2011)**

5. Huang W., Zhao X., Zhao X., Li Y. and Lian J., Effects of environmental factors on genetic diversity of *Caraganamicrophylla* in Horqin sand land, northest China, *Ecology and Evolution,* **6(22),** 8256-8266 **(2016)**

6. Kumar A., Barthwal S and Ginwal H.S., Molecular Genetic Analysis of *Cedrus deodara* (Roxb.) G. Don Age Class Population using cpSSR Markers, *Research Journal of Biotechnology*, **15(12),** 137-142 **(2020)**

7. Kumar A., Barthwal S. and Ginwal H.S., Genomic DNA isolation from old needles of *Cedrus deodara* (Roxb.) G.Don and amplification of nuclear microsatellite markers, *Indian Journal of Forestry*, **38(3),** 215-217 **(2015)**

8. Lorenz E., Leeton S. and Owen R.J., A simple method for sizing large fragments of bacterial DNA separated by PAGE, *Cabios*, **13**, 485-486 **(1997)**

9. Nei M. and Li W.H., Mathematical model for studying genetic variation in terms of restriction endonucleases, Proc Natl Acad Sci USA, 5269–5273**(1979)**

10. Nurtjaningsih I.L.G., Satio Y., Lian C.L., Tsuda Y. and Ide Y., Development and characterization of microsatellite markers in *Pinusmerkussi*, *Mol. Ecol. Notes*, **5,** 552-553 **(2005)**

11. Orwa C., Mutua A., Kindt R., Jamnadass R. and Simons A., Agro forest tree database: a tree reference and selection guide version 4.0 **(2009)**

12. Templeton A.R., Biodiversity at the molecular genetic level: experiences from disparate macro-organisms, *Phil. Trans. R. Soc.*, **345,** 59–64 **(1994)**

13. Tewari D.N., A monograph on deodar (Cedrusdeodara (Roxb.) G. Don), Dehradun, India, International Book Distributors, 212 **(1994)**

14. Tiwari A.K., Srinivas P.K., Kumar S.P. and Rao J.M., Free Radical Scavenging Active Components from *Cedrusdeodara*, *Journal of Agricultural and Food Chemistry*, **49,** 4642-464 **(2001)**

15. Watanabe A., Iwaizumi M.G., Ubukata-Kondo M.T., Lian C. and Hogetsu T., Isolation of microsatellite markers from PinusdensifloraSieb. Et Zucc using a dual PCR technique, *Mol. Ecol. Notes*, **6,** 80-82 **(2006)**

16. White B.J., Santolamazza F., Kamau L., Pombi M. and Grushko O., Molecular karyotyping of the 2 La inversion in Anopheles gambiae, *American Journal of Tropical Medicine and Hygiene*, **76,** 334-339 **(2007)**

17. Yeh F.C., Yang R.C. and Boyle T.B.J., PopeGene Version 1.31: Microsoft windows based Freeware for Population Genetic Analysis, University of Albert, Edmonton, http://www. Ualberta. ca/fyeh/ **(1999)**

18. Zhou Y., Bui T., Auckland L.D. and Williams C.G., Under-methylated DNA as a source of microsatellites from a conifer genome, *Genome*, **45,** 91-99 **(2002)**

19. Ziehe M. and Hattemer H.H., The significance of heterozygosity in tree breeding and conservation, *For. Tree Improvement*, **26,** 3-25 **(1998)**.

(Received 25th July 2021, accepted 07th October 2021)