GENETICS AND PLANT BREEDING

Genetic Diversity and Population Structure of *Pinus kesiya* through Trans-specific Amplification of Nuclear SSR Markers

Kirti Chamling Rai*, H.S. Ginwal and Romeet Saha

Division of Genetics and Tree Propagation, Forest Research Institute, Dehradun, India

*Corresponding author: kirtichamling@gmail.com (ORCID ID: 0000-0001-5241-4334)

Paper No. 737

Received: 20-07-2018

Accepted: 02-11-2018

ABSTRACT

De novo primer development is cost intensive and time-consuming, therefore using primers developed for other species on the target species is a more preferred alternative. In the present study, a total of 47 primer pairs from *P. taeda, P. merkusii, P. resinosa* and *P. densiflora* were used for trans-specific amplification of *P. kesiya*. It was observed that only 5 (10.6%) primer pairs out of 47 transferred in *P. kesiya* which may be due to the phylogenetic distance of the target species from the source species. The expected heterozygosity (*H*<sub>E</sub>) ranged from 0.490 to 0.603 with a mean of 0.540 and the observed heterozygosity (*H*<sub>O</sub>) ranged from 0.044 to 0.819 with a mean of 0.342. The study has shown that the nuclear SSR markers can be utilized for estimating the genetic structure of *P. kesiya* populations. Results of the present work will go a long way in implementing proper strategies for the better management and conservation of *P. kesiya* forests and initiating tree improvement programmes in this species.

Highlights

- Genetic diversity and variation was assessed using nuclear microsatellite markers in *P. kesiya* populations.
- The transferability of nuclear SSRs was very low with only 10.6% transfer rate.
- The populations showed moderate levels of genetic diversity.

Keywords: Cross-species amplification, Genetic diversity, Microsatellite markers, Nuclear SSR, *Pinus kesiya*

Environmental changes are unpredictable and it is of utmost importance that sufficient genetic diversity is secured to permit the species to continuously evolve in response to environmental pressure. Cross-species amplification is a reasonable approach for assessing the genetic diversity of populations which eliminates the need for de novo development of polymorphic primers as it is time-consuming and cost intensive. Microsatellite or simple sequence repeats (SSRs) are highly informative markers that are co-dominantly inherited. They are highly polymorphic and their transferability between closely related species makes them useful for the genetic studies of related species (Weising *et al.* 2005). SSR discovery from genomic libraries has proven problematic in conifers, with a low return for effort (Rajora *et al.* 2001; Hodgetts *et al.* 2001) due to the large size and repetitive nature of the conifer genome. This is why microsatellite transfer across the species is seen as a valued methodology (Ginwal *et al.* 2011).

*Pinus kesiya* commonly known as Khasi pine belongs to the family Pinaceae. The genus Pinus section Diploxyylon is characterized by hard timber and two xylem bundle. It is widely distributed between 30°N and 12°N in South East Asia. It occurs in Myanmar, India, Tibet, Laos, Vietnam, Thailand, the Philippines and the People's Republic of China (Hansen *et al.* 2003). In India, the species occur naturally in Khasi and Jaintia Hills and have been reported to grow at altitudes of between 800 and 2000 m amsl. Further to the east, it occurs throughout Manipur, Nagaland and Arunachal
Pradesh at elevations ranging from 1200 to 2000 m amsl (Chaudhary and Bhattacharyya, 2002). It is among the principal species in Meghalaya and constitutes 8.29% of the total forested area (FSI, 2015). 

*P. kesiya* is a commercially important species providing pulp, lumber and oleoresin. It is fast growing and has the capacity to adapt to various growing conditions and produces a high quality, long-fibered pulp (Hansen *et al.* 2003). In India, it is generally used for fuel wood and charcoal manufacture. The indigenous people of Meghalaya depend on *P. kesiya* for their traditional agricultural practices. They cultivate ginger (*Zingiber officinale*), turmeric (*Curcuma domestica*), paddy and vegetables under the Khasi pine based farming system. It also has medicinal properties and their young shoot is often used for treating children in case of a cough (Jeeva *et al.* 2006).

The pine forest of Meghalaya is under immense pressure due to overexploitation and getting denuded as a result of unplanned developmental activities and massive tree felling. The pine forest has become fragmented and damaged and is at a risk of reduced genetic diversity. Although Khasi pine is an ecologically and economically important species of the Northeast, very few studies have been carried out with regard to their genetic diversity and population genetic structure. The existence of Khasi pine depends on the sound conservation and management practices for which understanding of the gene diversity estimates will be of immense help and importance. The focus of the study was to examine the cross-species transferability of nuclear microsatellite markers and to use these markers to find out the genetic diversity and population structure of 10 *P. kesiya* populations.

**MATERIALS AND METHODS**

**Plant material**

A total of 250 individuals of *P. kesiya* from 10 geographical locations covering the entire distribution range in Northeast India were sampled for the study (Table 1). Samples were collected randomly and to avoid sampling individuals arising from the same parent, the distance between each individual was at least 100m.

**DNA extraction and PCR amplification**

The genomic DNA from the needles of the sample was extracted using the CTAB method with some modifications (Doyle and Doyle, 1990; Stange *et al.* 1998). The extracted DNA was qualified on 0.8% agarose gel and the concentration was measured using a Biophotometer. The samples of genomic DNA were then diluted to a final concentration of 15ng/µL as a working solution.

A total of 47 primers belonging to different *Pinus* species (Zhou *et al.* 2002; Change *et al.* 2004; Elsik *et al.* 2000; Nurtjahjaningsih *et al.* 2005; Watanabe *et al.* 2006; Boys *et al.* 2005) were tested for transferability.
in *P. kesiya* (Table 2). PCR amplification procedures were optimized on the basis of literature given by Vendramin with modifications such as the concentration of MgCl$_2$ and optimization of annealing temperatures, etc (Vendramin et al. 1996). The PCR reaction mixture consisted of 15 ng genomic DNA, 1x Taq buffer, 3.0 mM of MgCl$_2$, 0.2 mM of dNTP, 0.2 µM of forward and reverse primer and 5 unit of Taq Polymerase (Bangalore Genei Ltd. India). Samples were amplified using the following profile: initial denaturation (94°C, 5 min), followed by 35 cycles of denaturation (94°C, 1 min), annealing (locus-specific temperature, 1 min), extension step (72°C, 1 min) and a final extension (72°C, 8 min). The DNA fragments were visualized in a gel documentation imaging system (GelDoc-It System, UVP Ltd).

**Data analysis**

The bands obtained were scored manually which were then prepared into Input files for further analysis using various software. The software POPGENE version 1.32 (Yeh et al. 1999) was used to calculate the observed heterozygosity ($H_O$) and expected heterozygosity ($H_E$). The Wright’s $F$-statistics ($F_{st}$) and inbreeding co-efficient ($F_{Is}$) were calculated in FSTAT 2.9.3 (Goudet, 2002). Mantel test was performed to evaluate the genetic patterns of isolation by distance (IBD). A matrix of genetic distances among the population was built by calculating pairwise $F_{st}$ estimate which was then correlated with pairwise geographic distances (measured in kilometers) between populations (Diniz-Filho et al. 2013). The analysis was done with the help of XLSTAT, 2016 software. For the analysis of population structure, a model-based (Bayesian) cluster analysis was performed which was implemented in the software STRUCTURE version 2.2 (Pritchard et al. 2000a and 2000b). Ten independent STRUCTURE runs for each K (K = 1 to 10) were performed separately with 100,000 iterations and a burn-in period of 100,000. The value of K was detected by an ad hoc quantity based on the second order rate of change of the likelihood function with respect to K ($\Delta K$) (Evanno et al. 2005).

The phylogenetic tree using Unweighted Pair-Group Method with Arithmetic Average (Sneath and Sokal, 1973) was constructed using the POPTREE2 program (Takezaki et al. 2010). The population genetic structure was inferred by an analysis of variance framework (AMOVA analysis) according to Excoffier et al. (1992), using the Arlequin software version 3.11 (Excoffier et al. 2005).

### RESULTS AND DISCUSSION

#### Trans-specific amplification and Genetic Diversity

A total of 47 primers were tested for cross-species amplification in *P. kesiya* out of which only 5

---

| Locus       | Sequences (5’-3’)                  | Ta(°C) | Product size (bp) | Repeat motif | na | ne  | $H_O$ | $H_E$ | $F_{Is}$ |
|-------------|------------------------------------|--------|------------------|--------------|----|-----|-------|-------|---------|
| SSRPt ctg3754$^*$ | F: TCTTTGGGTTTCTGGAGTGG R: GCTGTTGCTTGTCTCTTGG | 60     | 421              | AGC          | 3.000 | 1.963 | 0.819 | 0.490 | -0.165  |
| SSRPt ctg3021$^*$ | F: CTCAGATTCTCTCCAAATGCG R: CATGCAACATATGCAAACCG | 60     | 234              | AGC          | 3.000 | 2.012 | 0.536 | 0.505 | 0.006   |
| RPTest6$^*$   | F: AGATTCGACAGCTACACC R: CAGACATGAGCCAGCTGTG | 60     | 147              | TGC          | 3.000 | 2.509 | 0.044 | 0.603 | 0.972   |
| RPTest1$^*$   | F: GATCGTTATTCCTCTGCA R: TTGATATCTGCTGG | 60     | 125              | AAT          | 3.000 | 2.483 | 0.246 | 0.599 | 0.845   |
| pm 07$^*$     | F: GAATCAAGCATAGAAATGAG R: CTTGTAATGCTACTATTTG | 55     | 284-309          | (AC)$_4$(AT)$_4$ | 2.000 | 1.999 | 0.067 | 0.501 | 0.404   |

(Adopted from $^*$ nSSRs Change et al. 2004, $^\Psi$ nSSRs Nurtjahjaningsih et al. 2005).

Ta, Annealing temperature; na, Observed number of alleles; ne, Effective number of alleles; $H_O$, Observed Heterozygosity; $H_E$, Expected Heterozygosity; $F_{Is}$, Fixation index.
primers amplified positively and indicated 10.6% transfer rate, the rest showing non-specific or no amplification. The primers from *P. merkusi* and *P. taeda* exhibited a low transfer rate of 20% and 13% respectively while primers from *P. densiflora* and *P. resinosa* did not show any positive amplification in the target species (Fig. 1). This low rate of transfer was due to the phylogenetic distance between the species. According to Celinski et al. (2013), the rate of transfer of microsatellites successfully transferred between species depends directly on their divergence time. Previous studies have also shown that pines have low transferability of microsatellite loci between hard and soft pines which belong to different subgenera of the genus *Pinus* (Bérubé et al. 2003).

The cluster dendrogram segregated all the 10 populations of *P. kesiya* into two major clusters (Fig. 2). The dendrogram did not show any clear demarcation of populations from different geographical locations.

![Fig. 1: Transfer rates of microsatellites from different species tested in *P. kesiya*](image1)

**Genetic structure and genetic differentiation**

A strong correlation between genetic and geographical distances (Mantel test: r=0.433; P<0.05) revealed a pattern of isolation-by-distance (IBD) across the distribution range of *P. kesiya* (Fig. 3). This pattern suggested that the gene flow is more likely to occur between neighbouring populations as their dispersal might be constrained due to the distance between them (Hutchison and Templeton, 1999). As a result, more closely situated populations tend to be more genetically similar to one another (Wright, 1943).

![Fig. 3: Relationship between pairwise $F_{ST}$ and geographic distance (r=0.433) for the 10 *P. kesiya* populations](image2)

A Bayesian analysis of the population structures which applies the Markov Chain Monte Carlo (MCMC) algorithm divided all the populations into four genetic clusters. Very few genotypes exhibited admixture (25.6%) while the rest were exclusively assigned to each single population (74.4%). The
Genetic Diversity and Population Structure of *Pinus kesiya* through Trans-specific...

Table 3: Analysis of molecular variance (AMOVA) for populations of *P. kesiya*

| Source of Variation                  | df  | Sum of squares | Variance components | Percentage of variation | Statistics |
|--------------------------------------|-----|---------------|---------------------|-------------------------|------------|
| 1. Among populations                 | 9   | 171.56        | 0.72 Va             | 39.64                   | $F_{ST} = 0.39^{***}$ |
| Within populations                   | 240 | 262.64        | 1.09 Vb             | 60.36                   |            |
| **Total**                            | 249 | **436.34**    | **1.81**            |                         |            |
| 2. Among groups                      | 4   | 115.63        | 0.36 Va             | 19.53                   | $F_{CT} = 0.19^{***}$ |
| Within populations within groups     | 5   | 55.93         | 0.40 Vb             | 21.69                   | $F_{SC} = 0.27^{***}$ |
| Within populations                   | 240 | 262.64        | 1.09 Vc             | 58.79                   | $F_{ST} = 0.41^{***}$ |
| **Total**                            | 249 | **434.20**    | **1.86**            |                         |            |

***Significant at 0.1% level of probability.

Fig. 4: (a) Bayesian posterior probability of data LnP(D) as function of K, where K=1-10; (b) Magnitude of ΔK as a function of K for SSR markers

low frequency of the admixture suggests that all the populations were distinct from each other possibility due to the limited gene flow between them.

Partitioning of the molecular variance was done without assuming hierarchical structure as well as assuming hierarchical structure. Both revealed that most of the variation was within population (60.36% and 58.79%, respectively) as is clear from Table 3. This is in agreement with a number of studies where it has been observed that conifers show high levels of genetic variation within populations and relatively little differentiation among populations (Yeh and El-Kassaby, 1980; Wheeler and Guries, 1982; Hiebert and Hamrick, 1983; Loveless and Hamrick, 1984; Kim et al. 1994; Mueller-Starck, 1995; Agundez et al. 1997). Variance estimates were based on 1000 permutations. Accordingly, the difference between the individuals within the populations was statistically significant (P<0.001).

**CONCLUSION**

In the last decade, genetic analysis of populations has evolved as a result of the technical advancement of molecular markers (Wan et al. 2004). Using DNA-based markers such as microsatellite markers (SSRs) has made it possible to analyze the genetic diversity both at the population as well as the species level. The de novo development of polymorphic microsatellite markers is difficult in the large highly duplicated conifer genome (Mariette et al. 2001). Therefore, transfer of microsatellites among related species is a reasonable approach for genetic diversity analysis (Celinski et al. 2013).

Pine is the most widely distributed species of the world having high commercial value. Its ecological significance and utility have made it the prime focus for many molecular evolutionary studies across the world. In India, a few significant studies have emerged on Pine species using the molecular marker techniques. However, extensive studies on *P. kesiya* through this approach in India have not been reported so far. The study indentified some useful microsatellite markers for *P. kesiya* which can be used for assessing the population genetic structure and diversity of this species. The study also highlighted some of the high genetic diversity forests from among the populations studied which
can be used to establish seed production areas in these forests and the seed from such areas should be used for plantations and infusing diversity in less diverse populations. For maintaining the existence of the Khasi pine forests in the Northeastern region, sound conservation and management practices are required. Results of the present work will go a long way in implementing proper strategies for conservation and management of *P. kesiya* forests and initiating tree improvement programs in this species.

**ACKNOWLEDGMENTS**

The authors are grateful and thank the Director, Forest Research Institute, Dehradun for providing the facilities to conduct the research work, Head of department, Division of Genetics and Tree Propagation, Dehradun and Forest department, Meghalaya for their valuable information. The authors are indebted to Dr. Lokopuni, IFS, for providing the samples from Manipur for the experiment.

**REFERENCES**

Agundez, D., Degen, B., Von Wuehlisch, G. and Alia, R., 1997. Genetic variation of Aleppo pine (*Pinus halepensis* MILL.) in Spain, *Forest Genetics*, 4: 201-209.

Bérubé, Y., Ritland, C. and Ritland, K. 2003. Isolation, characterization, and cross-species utility of microsatellites in yellow cedar (*Chamaecyparis nootkatensis*), *Genome*, 46: 353-361.

Boys, J., Cherry, M. and Dayanandan, S. 2005. Microsatellite analysis reveals genetically distinct populations of Red Pine (*Pinus resinosa*, *Pinaceae*), *American Journal of Botany*, 92: 833-841.

Celinski, K., Powlaczyk, E.M., W-Poltorak, A., Chudzinska, E. and P-Glowacki, W. 68. Cross-species amplification and characterization of microsatellite loci in Pinus mugo Turra, *Biologia*, 68: 621-626.

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

Chaudhary, V. and Bhattacharyya, B. 202. Suitability of *Pinus kesiya* in Shillong, Meghalaya for tree-ring analysis, *Current Science*, 83: 1010-1015.

Diniz-Filho, J.A.F., Soares, T.N., Lima, J.S., Dobrovolski, R., Landeiro, V.L., Telles, M.P.C., Rangel, T.F. and Bini, L.M. 2013. Mantel test in population genetics, *Genetics and Molecular Biology*, 36: 475-485.

Doyl, J.J. and Doyle, J.L. 1990. A rapid total DNA preparation procedure for fresh plant tissue, *Focus*, 12: 13-15.

Elisk, C.G., Minihan, V.T., Hall, S.E., Scarpa, A.M. and Williams, C.G. 2000. Low copy microsatellite markers for *Pinus taeda* L., *Genome*, 43: 550-555.

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

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

Hansen, C.P., Pedersen, A.P. and Graudal, L.O.V. 2003. International Series of Provenance Trials of *Pinus kesiya*: Field Assessment Manual, Danida Forest Seed Centre.

Hiebert, R.D. and Hamrick, J.L. 1983. Patterns and levels of genetic variation in Great Basin bristlecone pine, *Pinus longaeva*, *Evolution*, 37: 302-310.

Hodgetts, R.B., Aleksiuk, M.A., Brown, A., Clarke, C., MacDonald, E., Nadeen, S. and Khaza, D. 102. Development of microsatellite markers for white spruce (*Picea glauca*) and other related species, *Theoretical and Applied Genetics*, 102: 1252-1258.

Hutchinson, D.W. and Templeton, A.R. 1999. Correlation of pairwise genetic and geographical distance measure: inferring the relative influence of gene flow and drift on the distribution of genetic variability, *Evolution*, 53: 1898-1914.

Jeeva, S.R.D.N., Laloo, R.C. and Mishra, B. 2006. Traditional agricultural practices in Meghalaya, North East India, *Indian Journal of Traditional Knowledge*, 5: 7-18.

Kim, Z.S., Lee, S.W., Lim, J.H., Hwang, J.W. and Kwon, K.W. 1994. Genetic diversity and structure of natural populations of *Pinus koraiensis* (Sieb. et Zucc.) in Korea, *Forest genetics*, 1: 41-49.

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

Mariette, S., Change, D., Decroocq, S., Vendramin, G. G., Lalanne, C., Madura, D., Plomion, C. 2001. Microsatellite markers for *Pinus pinaster* Ait, *Annals of Forest Science*, 58: 203-206.

Mueller-Starck, G. 44. Genetic variation in high elevation populations of Norway spruce (*Picea abies* [L.] Karst.) in Switzerland, *Silvae Genetica*, 44: 356-362.
Muller, G. 1997. Untersuchungen über die natürliche Selbstbefruchtung in Bestanden der Fichte (Picea acies L.) Karst und Kiefer (Pinus sylvestris L.), Silvae Genetica, 26: 207-217.

Nei, M. 1975. Molecular population genetics and evolution, Amsterdam, North-Holland.

Nurtjahjaningsih, I.I.J., Saito, Y., Lian, C.L., Tsuda, Y. and Ide, Y. 2005. Development and characterization of microsatellite markers in Pinus merkusii, Molecular Ecology Notes, 5: 552-553.

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

Pritchard, J.K., Stephens, M., Rosenberg, N.A., Donnelly, P. 2000. Association mapping in structured population, American Journal of Human Genetics, 67: 170-181.

Rajora, O.P., Rahman, M.H., Dayanandan, S. and Moseler, A. 2001. Isolation, characterization, inheritance and linkage of microsatellite DNA markers in white spruce (Picea glauca) and their usefulness in other spruce species, Molecular and General Genetics, 264: 871-882.

Sneath, P.A. and Sokal, R.R. 1973. Numerical taxonomy, W.H. Freeman, San Francisco, C.A.

Stange, C. Prehn, D. and Johnson, P.A. 1998. Isolation of Pinus radiata genomic DNA suitable for RAPD analysis, Plant Molecular Biology Reporter, 16 (1998) 1-8.

Takezaki, N., Nei, M. and Tamura, K. 2010. POPTREE2: Software for constructing population trees from allele frequency data and computing other population statistics with Windows interface, Molecular Biology and Evolution, 27: 747-752.

Vendramin, G.G., Lellirr Rossi, P. and Morgante M., 1956. A set of primers for the amplification of 20 chloroplast microsatellites in Pinaceae, Molecular Ecology, 5: 595-598.

Watanabe, A., Iwaizumi, M.G., Ubukata, M., Kondo, T., Lian, C. and Hogetsu, T. 2006. Isolation of microsatellite markers from Pinus densiflora Sieb. et Zucc. using a dual PCR technique, Molecular Ecology Notes, 6: 80-82.

Weising, K., Nybom, H. and Wolff, K. 2005. DNA fingerprinting in plants principles, methods, and applications. 2nd edn. Taylor & Francis Group, CRS Press, New York.

Wheeler, N.C. and Guries, R.P. 1982. Population structure, genetic diversity and morphological variation in Pinus contorta Dougl, Canadian Journal of Forest Research, 12: 595-606.

Wright, S. 1943. Isolation by distance, Genetics, 28: 114-138.

Wan, Q.H., Wu, H., Fujihara, T. and Fang, S.G., Which genetic marker for which conservation genetics issue? Electrophoresis, 25: 2165-2176.

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

Yeh, F.C. and El-Kassaby, Y.A. 1980. Enzyme variation in natural populations of Sitka spruce (Picea sitchensis), Genetic variation patterns among trees from 10 provenances, Canadian Journal of Forest Research, 10: 415-422.

Zhou, Y., Bui, T., Auckland, L.D. and Williams, C.G. 2002. Undermethylated DNA as a source of microsatellites from a conifer genome, Genome, 45: 91-99.
