Indian Journal of Biotechnology
Vol. 17, July 2018, pp 538-540

Short communication

Identification of allele specific primers in chir pine (*Pinus roxburghii* Sarg.) through data mining

Swati Allen*, H S Ginwal and Santan Barthwal

Division of Genetics and True Propagation,
Forest Research Institute, PO-IPE,
New Forest, Dehradun, 248195, Uttarakhand, India

Received 17 November 2015; revised 18 May 2016; accepted 18 May 2016

Pinus roxburghii Sarg. commonly known as Chir pine, is a commercially exploited species for resin in India. Resin is a phenotypic trait which is expressed only in mature trees. Due to this a large number of trees are cut every year which pose a serious ecological threat. Recent advancements in bioinformatics and molecular biology have enabled scientists to identify high resin yielding pine genotypes at nursery stage using molecular markers. In the present study, investigation, 337 terpene coding sequences from the database were analyzed using data mining and 45 microsatellites allele specific primers were designed for marker assisted selection. Out of these 45 primers, only 22.2% primer pairs showed polymorphism within the 53 genotypes of Chir pine examined. A wide range of fragment size was observed from 100 - 600 bp. PCR amplification using allele specific markers produced a total of 22 bands, out of which 14 were found to be polymorphic. The total number of bands amplified per marker varied from of 1 to 3 with an average of 1.4. Genetic divergence in terms of percent polymorphism ranged from 50 to 100% with a mean of 63.3% per marker.

Keywords: Terpene, allele specific primers, polymorphism, *Pinus roxburghii* Sarg.

Introduction

Pinus roxburghii commonly known as Chir pine or long needle pine is one of the most important conifers of North-Western Himalayas and is an important timber and resin yielding species. In Chir pine, natural variation in resin is associated with the extremely diverse geographic regions in which it grows. To exploit the natural variability, selection of desirable parents and crossing them in a desired way is a key step for tree improvement programme. By revealing differences in the DNA sequence among individual trees, DNA markers provide the potential to increase genetic gain from tree improvement programmes through DNA fingerprinting of elite genotypes, parentage testing of superior seed and through the identification of DNA markers associated with traits of economic value in an integrated marker assisted breeding programme. There is an urgent need to identify high resin pine yielders at nursery stage using marker assisted selection to minimize ecological threat. Considering the potential applications of allele specific primers in pine forest conservation, the present study was conducted in Forest Research Institute (FRI) Dehradun, to identify the potential microsatellites responsible for high resin production in Chir pine in nursery stage.

With the help of bioinformatics tools, sequences coding for these terpene synthase gene regions were downloaded from National Centre for Biotechnology Information (NCBI) database and were tabulated in two groups based on genus. Terpene enzyme coding sequences from *Pinus* genus comprised first group while the other sequences comprised the second group. Total 337 sequences were downloaded from the NCBI, out of which 102 sequences were from *Pinus* genus. In previous study conducted in Genetics and Tree Propagation division, fifty-three resin yielders were characterized using SSR, ISSR, RAPD and AFLP markers[1]. In the present study, these genotypes were tested using allele specific primers.

Primers were designed using conserved and variable domains of selected sequences using Primer 3 software. Out of 45 allele specific primers, 10 primers showed polymorphism. DNA was extracted from young needles and sapwood samples available in the division. The protocol[2] with modifications successfully amplified the allele specific loci in *P. roxburghii.* Concentrations of 2.00 mM, 2.5 mM and 3.00 mM of $MgCl_2$, 0.2 mM of dNTPs, 0.2 µM of forward and reverse primer, 0.1 unit of *Taq* DNA polymerase and 15 ng/µl of template DNA was found to be optimal for the amplification. Using gradient PCR, the optimum annealing temperature for each primer was standardized.

PCR amplification using allele specific markers produced a total of 22 bands, out of which 14 were found to be polymorphic. The total number of bands

*Author for correspondence:
allenswati@gmail.com

Table 1 — Analysis of molecular variance (AMOVA) of *Pinus roxburghii* genotypes based on collection site and resin yield using allele specific markers

Basis	Source of variation	df	Sum of squares	Mean sum of squares	Variance components	Percentage of variation	F-statistic
Collection site	Among populations	2	0.53	0.265	0.02Va	3.91	0.005
	Within populations	51	25.92	0.50	0.49Vb	96.09	
	Total	53			0.51		
Resin yield	Among populations	4	11.44	3.8	0.72 Va	22.3	0.01**
	Within populations	49	124.5	2.5	2.5 Vb	77.7	
	Total	53			3.22		

**denotes significance at 1% probability level

amplified per marker varied from a minimum of 1 to a maximum of 3, with an average of 1.4. Genetic divergence in terms of per cent polymorphism ranged from 50 to 100% with a mean of 63.3% per marker. This was in accordance with the findings[1,3] where percentage of polymorphic loci over the 55 populations ranged from 50 to 100%. Similar results were quoted in *Jatropha curcas* where the number of allelels per SSR marker ranged from 20 to 4 and all the amplified primer pairs showed 100% polymorphism among the various accessions[4]. Other reports where the level of polymorphism was found to be 100% using SSR markers include *Dalbergia sissoo*[5-6].

The polymorphism information content (PIC) ranged from 0.171 to 0.491 with an average of 0.281 per marker. The average PIC 0.281 per primer pair observed in the present investigation was comparable to that reported earlier in *P. roxburghii* using SSR markers. The marker index (MI) ranged from 0.357 to 0.885 with an average of 0.758 per marker which was in agreement with the observations where the average MI was 0.695 (Ref. 1). The utility of a given marker system is a balance between the level of polymorphism detected (information content) and the extent to which an assay can identify multiple polymorphisms. There was a strong correlation between resolving power (RP) and PIC (r^2 = 0.86). Analysis of molecular variance was carried out in two ways i.e. genotypes grouped by collection site and by resin yield (Table 1). Analysis of molecular variance (AMOVA) by collection site revealed only 3.91% of the total variation to be among the groups with an F-statistic (F_{ST}) value of 0.005. On the other hand, in AMOVA by resin yield, the percentage of variation among the groups was 22.3% with an F_{ST} value of 0.01 which was highly significant ($p < 0.001$). This was in accordance where the resin composition of *P. pinaster* has been linked with a biosynthetic process and is controlled by genetic factors[7].

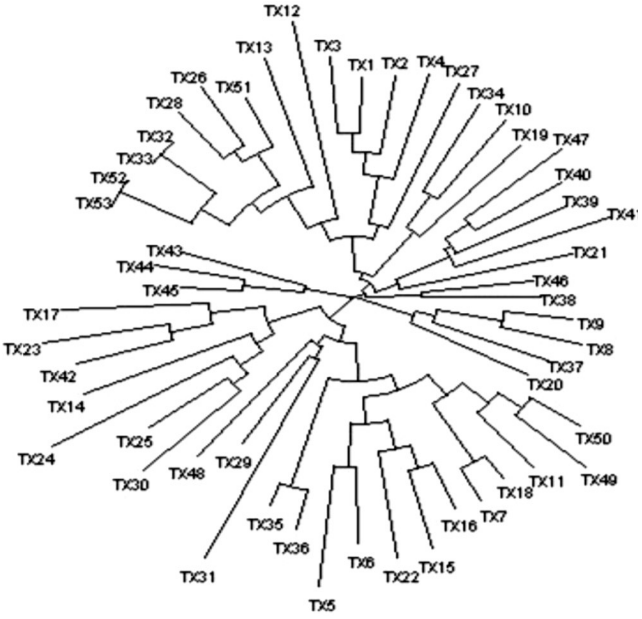

Fig. 1 — Cladogram generated by Tassel.

The relative proportions of constitutive monoterpenes are strongly inherited, little influenced by environmental traits[8] and have been used for many years as biochemical markers in forest genetics[9-10]. This study shows that resin yield of the genotypes was not attributed to their ancestry but it was because of their genetic constitution. This confirms that terpenes can be used as trait specific molecular marker in resin yield studies using allele mining. The cladogram generated by TASSEL is based on neighbor-joining trees using only simple parsimony substitution models. This cladogram generated different clusters of genotypes of *Pinus roxburghii* irrespective of resin yield (Fig. 1)

In order to validate the specificity of allele specific primers, PCR product of two primers coding for pinene synthase were sent for sequencing. These sequences were then BLASTed online to check for sequences with similar results. Both of these

sequences showed more than 90% similarity with other pinene sequences from *Pinus* genus, thus confirming the specific product amplification using allele specific primers. In order to identify low and high resin pine genotypes at an early stage, we can design allele specific primers which can help us in marker assisted selection (MAS) of genotypes. MAS is a relatively new tool for phenotypic selection of superior individuals among the genotypes. When a marker is found that co-segregates with a major gene for an important trait, it may be easier and cheaper to screen for the presence of the marker allele linked to the gene, than to evaluate the trait. Thus, the above results reveal the potential of allele specific markers to be used in MAS programmes for the selection of superior genotypes in wild domesticated forest genetic resources.

References

1 Rawat A, *Identification of molecular markers for high resin yield in Pinus roxburghii* Sarg, Ph.D thesis, Forest Research Institute University, Dehradun 2013.

2 Vendramin G G, Lellilr Rossi P & Morgante M, A set of primers for the amplification of 20 chloroplast microsatellites in Pinaceae, *Mol Ecol,* 5 (1996) 595-598.

3 Chauhan P, *Molecular genetic analysis of Chir pine* (*Pinus roxburghii* Sarg.) *through microsatellite markers*, Ph.D thesis, Forest Research Institute University, Dehradun 2011.

4 Mittal N & Dubey A K, A novel set of highly polymorphic chloroplast microsatellite and ISSR markers for the biofuel crop *Jatropha curcas*, *Eurasian J* Bio Sci 4 (2010) 119-131

5 Maurya S S, *Molecular characterization of different clones* of *Dalbergia sissoo* Roxb. *through DNA based markers*, Ph.D Thesis, Forest Research Institute University, Dehradun 2010.

6 Russell J R, Fuller J D, Macaulay M, Hatz B G, Jahoor A *et al,* Direct comparison of levels of genetic variation among barley accessions detected by RFLPs, AFLPs, SSRs and RAPDs, *Theor Appl Gene,* 95 (1993) 714-722.

7 Tadesse W, Nanos N, Aunon F J, Alia R & Gil L, Evaluation of high resin yielders of *Pinus pinaster* Ait, *For Gene,* 8 (2001) 271-278.

8 Mihcelozzi M, Radicati M & Properzi M, Resin monoterpene composition of Mediterranean pines of group "Halepensis". ler simposio de aprovechamiento de resinas naturales *Actas cientificas, Segovia,* (1998), 62-71.

9 Hanover, J W Applications of terpene analysis in forest genetics, *New For Gene*, 6 (1992) 159-178.

10 Plomion C, Pionneau C & Baille` Res H, Analysis of protein expression along the normal to tension wood gradient in *Eucalyptus gunnii, Holzforschung*, 57 (2003) 353-358.