

Genetic diversity and spatial genetic structure in populations of *Orbignya phalerata* Mart. under different exploitation intensities in the Brazilian savanna

By B. IBANES^{*,1)}, A. M. SEBBENN²⁾, V. C. R. AZEVEDO³⁾, M. A. MORENO¹⁾, F. B. GANDARA¹⁾, E. V. TAMBARUSSI⁴⁾, E. M. FERRAZ¹⁾, K. J. DAMASCENO-SILVA⁵⁾, P. S. C. LIMA⁵⁾ and M. A. CARVALHAES⁵⁾

(Received 6th August 2015)

Abstract

Genetic studies in tropical tree species have found signs of decreased genetic diversity and increased levels of inbreeding and spatial genetic structure (SGS) in fragmented and exploited populations. The aim of this paper was to investigate genetic diversity, structure, and intrapopulation SGS using eight microsatellite loci for three *Orbignya phalerata* populations that have undergone different intensities of seed harvesting. From each population, we georeferenced and sampled 30 seedlings, 30 juveniles, and 30 adult trees. The total number of alleles over all loci (k), and observed (H_o) and expected heterozygosity (H_e) presented lower values for the population experiencing more intense fruit harvesting than less heavily exploited populations, suggesting that fruit harvesting may decrease genetic diversity. Null alleles were detected in practically all loci among seedlings, juveniles, and adults in all populations, indicating that the estimates of H_o , H_e , and fixation index (F) are biased. When corrected for null alleles (F_{Null}), the fixation index decreased for all samples, resulting in signifi-

cantly higher than zero results for seedlings of all populations, but not for juveniles and adults of all populations. The comparison of F_{Null} values between cohorts in the most heavily exploited population (ESP) suggests that inbred individuals are eliminated between seedling and adult stages. Significant SGS was detected up to 60 m in all populations, which indicates short distance seed dispersal. Genetic differentiation (G'_{ST}) between pairwise populations was related to spatial distance between populations, with the greatest difference between more distant populations.

Key words: babassu; *Orbignya phalerata*; forest management; exploitation; genetic diversity.

Introduction

Genetic diversity is necessary to ensure species can maintain high evolutionary and adaptive potential to withstand environmental stochasticity (GEBUREK, 2000). Loss of genetic diversity may have significant consequences for a population as it affects the mating system, as well as pollen and seed flow, and such processes determine levels of genetic diversity within populations (YOUNG and BOYLE, 2000). Studies of tropical tree species have found signs of decreased genetic diversity and increased inbreeding in fragmented forest populations (AGUILAR et al., 2008). However, such changes may take several generations to manifest because the remaining adults still represent the pre-disturbance population (HAMRICK, 2004; AGUILAR et al., 2008; MANOEL et al., 2012).

The main consequence of inbreeding is the occurrence of homozygosity for deleterious recessive alleles, which may lead to reduced yield, fertility, and seed viability, thus increasing the risk of extinction (FRANKHAN, 2012). In this context, analysis of spatial genetic structure (SGS) is useful as it characterizes the levels and spa-

¹⁾ Escola Superior de Agricultura "Luiz de Queiroz", Universidade de São Paulo", Av. Pádua Dias, 11, CP 9, Piracicaba, SP, 13418-900, Brazil.

²⁾ Instituto Florestal de São Paulo, CP 1322, São Paulo, SP, 01059-970, Brazil.

³⁾ Empresa Brasileira de Pesquisa Agropecuária – Recursos Genéticos e Biotecnologia, Av. W5 Norte (final), CP 02372, Brasília, DF, 70770-917, Brazil.

⁴⁾ Programa de Pós Graduação em Ciência Florestal – FCA/UNESP – Botucatu, Rua José Barbosa de Barros, nº 1780, Portaria II: Rodovia Alcides Soares, Km 3, Botucatu, SP, 18610-307, Brazil.

⁵⁾ Empresa Brasileira de Pesquisa Agropecuária – Meio Norte, Av. Duque de Caxias, 5650, Teresina, PI, 64006-220, Brazil.

^{*} Corresponding author: BRUNA IBANES. Escola Superior de Agricultura "Luiz de Queiroz", Universidade de São Paulo", CP 9, Piracicaba, SP, CEP: 13418-900, Brazil, Phone: +55 19 2105-8675.
E-Mail: brunaibanes@yahoo.com.br

tial distribution of genetic diversity within populations (HARDY et al., 2006). SGS is defined as a non-random distribution of genetic variation across space and time and may be influenced by mating system, pollen and seed dispersal, pollinator community, and population density (EPPERSON, 1990; HARDY et al., 2006; DICK et al., 2008), generating a greater degree of relatedness between individuals within populations than would be expected by chance. Understanding SGS is essential for seed harvesting strategies used in conservation and forest restoration programs, to avoid collecting seeds from related mother trees, thus decreasing the probability of sampling related seeds from different mothers (SEBBENN et al., 2011).

Babassu is the common name given to several palm trees from *Attalea* and *Orbignya* genera. The taxonomic identification of babassu species is difficult due to the occurrence of natural hybrids resulting from interspecific crosses. The palm species *Orbignya phalerata* Mart. (Areaceae) occurs widely across Maranhão, Piauí, Tocantins, Pará, Amazonas, Goiás, and Mato Grosso States in Brazil, ranging over an area of about 18 million hectares (LORENZI et al., 2010). The species reaches up to 30 m in height and 60 cm in diameter at breast height (DBH). Its leaves are pinnate and erect, and inflorescence may be either staminate or hermaphroditic (LORENZI et al., 2010). Protogyny occurs in the species and there is a low frequency of self-pollination (ANDERSON and BALICK, 1988). Although the primary pollinator is the small beetle *Mystrops mexicana*, wind also plays an important role in pollination across open areas (ANDERSON and BALICK, 1988), and fruits are dispersed in bunches (LORENZI et al., 2010). The palm is an important source of income for regional communities and local economies as it is a source of renewable energy (LORENZI et al., 2010) and is widely used in the cosmetics and food industries (i.e., a source of fat for ice cream, margarine, and cocoa butter production, HAUMANN, 1992). To date, the effects of seed harvesting on genetic diversity of the species have not been assessed.

This study investigated the genetic diversity and structure, and intrapopulation spatial genetic structure, of three *O. phalerata* populations experiencing different intensities of seed harvesting. To conduct our analyses, we developed eight microsatellite markers specifically for *O. phalerata*.

Material and Methods

Study sites and sampling

We sampled three populations of *O. phalerata* that experience different intensities of fruit harvesting: José de Freitas (JF), Teresina (TER), and Esperantina (ESP), all located in Piauí State, Brazil (Fig. 1). Site JF (04° 45' 23" S and 42° 34' 32" W) has an area of 9.7 ha and is located on a privately owned farm in which the owner, Mr. Freitas Filho, does not allow fruit collection (personal communication). The land was acquired in the 20th century by the owner's great-grandfather, and the family has managed and preserved the area for more than 40 years. As there has been no previous harvesting activity in the area, we classified this population as modestly exploited. The region's vegetation includes patches of Caatinga and savanna, as well as dense *O. phalerata* forests. It has a semiarid tropical climate with a hot and dry season of six months and temperatures ranging from 18 to 38°C. The TER sample area (5° 5' 21' S and 42° 48' 07" W) belongs to

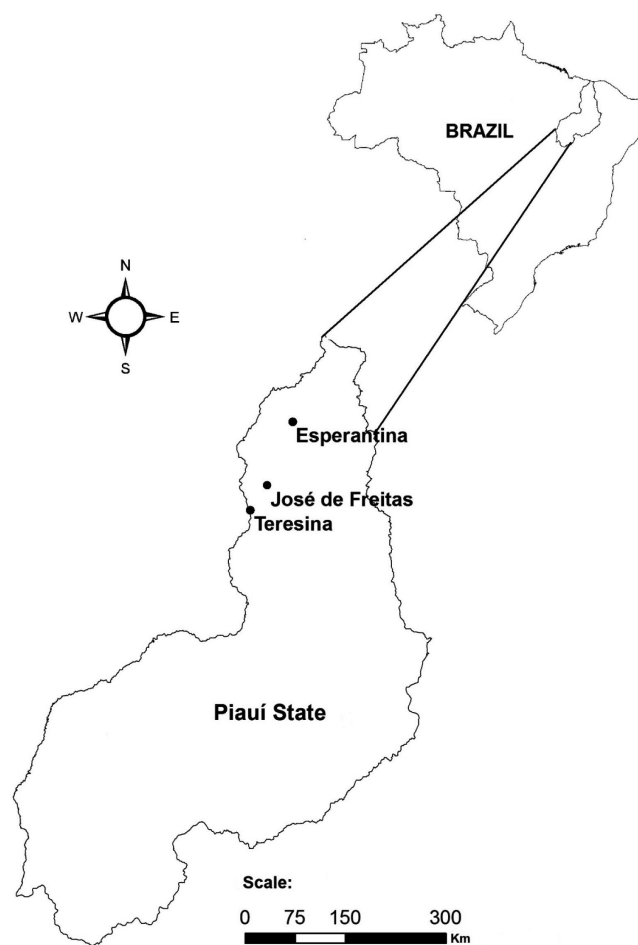


Figure 1. – Location of the study populations in Piauí State, Brazil.

Embrapa Meio-Norte (a branch of the Brazilian government agricultural research agency) and it is made up of 5.2 ha of dense forest with native tree species, low species diversity, and a predominance of *O. phalerata*. In this area, fruit harvesting of medium intensity occurs, local access is legally controlled, and around 10 rural families within 400 ha are allowed to exploit the area. We estimate that less than 10% of the *O. phalerata* population is being exploited; therefore, seed harvesting is not intensive, and dropped fruit are left on the ground to be dispersed by fauna. The area is characterized by a hot, sub-humid tropical climate with a dry season of six months, temperatures ranging from 22 to 38°C, and average annual rainfall of 1,370 mm. Site ESP (03° 54' 06" S and 42° 14' 01" W) consists of 22.4 ha of typical mixed deciduous forest with patches of Caatinga and savanna. This site has been used by rural populations for centuries for *O. phalerata* exploitation. There are several agrarian reform settlements, aggregate rural workers, and land tenants living in the region. The intensity of the harvest depends on the demand for the oil and charcoal; during some harvests, all fruit on the ground is collected. The babassu forests are freely accessible to all communities wishing to exploit them (access is not controlled at this site) and the fruit of *O. phalerata* is a major source of income for the local population. We consider this population as heavily exploited. The area has a sub-humid tropical climate with a dry season of six months, temperatures range from 25 to 34°C, and average annual rainfall is 1,400 mm.

From each population, 30 seedlings (individuals with 3 true leaves), 30 juveniles (before reproductive stage), and 30 adults (reproductive individuals) were sampled and georeferenced using GPS (Garmin: GPSMAP 76CXS). For genetic analyses, leaf samples were collected, stored in containers with silica gel, and transported to the laboratory.

Microsatellite analysis

Genomic DNA extraction was performed using CTAB following the method described by DOYLE and DOYLE (1990), where 150-mg of leaf tissue was macerated in liquid nitrogen. After extraction, samples underwent electrophoresis on 1% agarose gel, and then stained with ethidium bromide, and visualized under ultraviolet light (UV) to quantify the DNA by visually compar-

ing to a standard (λ DNA). The samples were further diluted in ultrapure water to a required concentration of 2.5 ng. μ L⁻¹.

DNA amplifications were performed in solutions containing 2.5 ng. μ L⁻¹ genomic DNA, 250 mM each of deoxyribonucleotides (dATP, dCTP, dGTP, dTTP), 10 mM Tris-HCl, 50 mM KCl, 1.5 mM MgCl₂, 2.5 μ g.mL⁻¹ BSA, 0.2 M of each primer, 1 U. μ L⁻¹ Taq DNA polymerase, and ultrapure water. Amplifications were conducted in a thermocycler (MJ Research PTC-100) programmed as follows: denaturation at 96°C for 2 min; 30 cycles of denaturation at 94°C for 1 min, specific annealing temperature per primer pair for 1 min, an extension step at 72°C for 1 min, and a final extension at 72°C for 10 min. After amplification, DNA fragments were separated on a 5% denaturing polyacrylamide gel in 1X TBE buffer (90 mM Tris-HCl, 2.0 mM EDTA, and 90 mM boric acid at pH 8.3) for 1.5 h in a vertical tank. Amplification products were stained with silver nitrate and analysed based on the method of CRESTE et al. (2001). Allele size in base pairs (bp) was determined by comparison with a standard molecular weight (MW) marker DNA ladder of 10 bp (Invitrogen®). Fragments of different sizes were considered different alleles.

Primer development

Specific SSR markers were developed for genetic analyses of the species. Total genomic DNA was extracted from leaves of a single *O. phalerata* plant and digested with restriction enzyme MSE I (New England BioLabs, Ipswich, MA). Fragments from 200 to 800 bp were recovered and used to construct a library enriched for SSR loci sequences as described by RAFALSKI et al. (1996) and further optimized for tropical species (BUSO et al., 2003). These fragments were ligated into pGEM-T Easy Vector (Promega, Madison, WI) and transformed to *E. coli* XL1-Blue. Positive clones were selected and sequenced using an automatic ABI 3700 DNA analyser (Applied Biosystems). The sequences obtained were processed by Genescan version 3.7 and Genotyper version 2.0 NT software (Applied Biosystems) and transformed into „FASTA“ format using the BioEdit Sequence Alignment Editor Program. Primer design was performed using the Primer3 software with fragments between 100 bp to 350 bp and GC percentage from 40 to 60%. The primer pairs were then tested in PCR reactions and 24

individuals were randomly selected for polymorphism detection (Table 1). Finally, eight identified loci were selected for *O. phalerata* genotyping and characterization.

Analyses of genotypic disequilibrium and genetic diversity/structure

The presence of null alleles within seedling, juvenile, and adult samples of each population was estimated using INEST (CHYBICKI and BURCZYK, 2009). Genotypic disequilibrium between pairs of loci was tested to verify the allelic association of different loci. We assessed statistical significance using Monte Carlo permutations of alleles among individuals, followed by Bonferroni correction ($\alpha=0.05$: 5% probability). Genetic diversity was assessed for each population for seedlings, juveniles, and adults separately and the whole population together, to determine total number of alleles per locus (k), average allelic richness (R), observed heterozygosity (H_o), and expected heterozygosity (H_e). Inbreeding within the samples was inferred from the fixation index [$F = 1 - (H_o/H_e)$] and the significance of mean values was tested using Monte Carlo permutation of alleles among individuals, followed by Bonferroni correction ($\alpha=0.05$). As the presence of null alleles results in bias in the estimate of H_o , consequently resulting in bias in the estimates of F , we corrected the fixation index for null alleles (F_{Null}) using a Bayesian approach (IIM) imple-

mented in INEST (CHYBICKI and BURCZYK, 2009). Genetic differentiation (G'_{ST}) among populations and generations was estimated according to HEDRICK (2005). All analyses were performed using FSTAT (GOUDET, 1995). An unpaired t-test was used to determine whether the estimated genetic indices were significantly different between development phases within the same population.

Spatial genetic structure analysis

The coancestry coefficient (θ_{xy}) was used to evaluate the spatial genetic structure (SGS) as described in LOISELLE et al. (1995), and implemented in SPAGeDi 1.3 (HARDY and VEKEMANS, 2002). Because robust estimates of SGS require at least 100 individuals (CAVERS et al., 2005), we grouped seedlings, juveniles, and adults in each population for this analysis because each cohort has a maximum of 30 individuals. However, in conducting such grouping we must assume that there are no differences in SGS between cohorts. The distance classes were determined using the same number of pairwise individuals within each class. One thousand permutations of individuals were used to estimate the 95% confidence interval among the different distance classes. The extent of SGS was compared using the S_p statistic (VEKEMANS and HARDY, 2004), as: $S_p = b_k / (1 - \theta_l)$; where θ_l is the average coancestry coefficient calculated between individuals in the first distance class and b_k is the

Table 1. – Nuclear microsatellite loci from *O. phalerata* with motifs employed in the library enrichment, primer sequences, allele ranges in base pairs (amplitude), annealing temperature in degrees Celsius (T_a), and Genbank accession numbers.

Locus	Motifs	Primer sequences (5'-3')	Amplitude (pb)	T_a	Access no.
OPH03	(CT) ₁₀ (CA) ₇	F: TGCTGAAGAAGAGAAACATGGGA R: TATCCAACGGTATCAGATGGC	130/150	60	KJ803841
OPH12	(GA) ₁₇	F: ACAGAGGATCACTATTTGCCGT R: ATAAATGCTTGCGTCCGTTG	120/140	60	KJ803842
OPH17	(GA) ₁₅	F: GTAAACTTCAGCCAAATCGAC R: CTTATGGTGTTCATGGGACATTG	138/156	60	KJ803843
OPH20	(GA) ₁₉	F: CGAACCCAACACACATACAGAC R: TTCATGTTCACCTTTGCAAGTCC	116/122	60	KJ803844
OPH22	(AG) ₁₀	F: AAAACCAAGGTTCACTGGACTG R: CTCCTCTCCATAAGCCCTCTCT	250/270	60	KJ803845
OPH30	(GAA) ₁₅	F: GATGGAAAAGAAAACATCAGGG R: CTTATCTCACGCATGCACTCTCT	140/176	60	KJ803846
OPH39	(CT) ₁₇	F: CTCTCTATGCTTTGCAGATT R: AAGGAGGAAGAGTGAAGGTA	158/192	60	KJ803847
OPH40	(GA) ₁₈	F: GGAAGAAAGGAGAGGGAAAGG R: CCACCATCTTCTCCTCAGTCTT	116/140	60	KJ803848

regression slope of the coancestry coefficient on the logarithm of distance between individuals (0-100 m). Statistical significance of SGS was

tested by permuting individual positions 1,000 times to obtain the b_k frequency distribution, assuming θ_l and $\ln(d_{xy})$ are not correlated.

Table 2. – Description of genetic diversity and inbreeding for populations and generations of *Orbignya phalerata* (k is the total number of alleles for all loci; R is allelic richness estimated for 14 individuals; H_o and H_e are the observed and expected heterozygosity, respectively; F is the fixation index likely biased due the presence of null alleles; F_{Null} is the corrected fixation index for null alleles; \pm SE is the 95% standard error; * $P < 0.05$).

Sample	n	k	$R \pm SE$	$H_o \pm SE$	$H_e \pm SE$	$F \pm SE$	$F_{Null} \pm SE$
José de Freitas: JF							
Seedlings	30	71	7.07 \pm 2.15	0.43 \pm 0.24	0.70 \pm 0.21	0.39 \pm 0.34*	0.11 \pm 0.01*
Juveniles	30	86	10.75 \pm 4.03	0.53 \pm 0.20	0.78 \pm 0.24	0.32 \pm 0.26*	0.03 \pm 0.01
Adults	30	78	8.06 \pm 2.39	0.56 \pm 0.22	0.77 \pm 0.20	0.27 \pm 0.26*	0.09 \pm 0.02*
Total	90	99	8.57 \pm 2.52	0.51 \pm 0.21	0.77 \pm 0.22	-	-
Teresina: TER							
Seedlings	20	51	6.64 \pm 2.37	0.49 \pm 0.25	0.68 \pm 0.21	0.29 \pm 0.25*	0.09 \pm 0.03*
Juveniles	26	81	8.36 \pm 2.24	0.60 \pm 0.21	0.79 \pm 0.13	0.24 \pm 0.20*	0.02 \pm 0.01
Adults	30	89	8.93 \pm 2.25	0.57 \pm 0.21	0.79 \pm 0.14	0.28 \pm 0.24*	0.09 \pm 0.01*
Total	76	110	8.70 \pm 2.11	0.56 \pm 0.20	0.78 \pm 0.16	-	-
Esperantina: ESP							
Seedlings	30	55	6.03 \pm 2.65	0.36 \pm 0.20	0.64 \pm 0.28	0.44 \pm 0.31*	0.12 \pm 0.03*
Juveniles	30	63	6.56 \pm 2.68	0.45 \pm 0.26	0.66 \pm 0.26	0.32 \pm 0.28*	0.15 \pm 0.02*
Adults	30	62	6.60 \pm 2.41	0.43 \pm 0.19	0.67 \pm 0.26	0.37 \pm 0.25*	0.01 \pm 0.00
Total	90	77	6.98 \pm 2.71	0.41 \pm 0.18	0.67 \pm 0.28	-	-

Table 3. – Expected null allele frequencies for seedlings (Se), juveniles (Ju), and adults (Ad) for the studied *O. phalerata* populations. (\pm SD is the standard deviation).

Locus	José de Freitas: JF			Teresina: TER			Esperantina: ESP		
	Se	Ju	Ad	Se	Ju	Ad	Se	Ju	Ad
OPH03	0.042	0.070	0.001	0.013	0.043	0.055	0.139	0.109	0.050
OPH12	0.191	0.086	0.114	0.001	0.206	0.134	0.223	0.040	0.173
OPH17	0.045	0.128	0.083	0	0.148	0.012	0.107	0.100	0.114
OPH20	0.018	0.042	0.054	0.004	0.190	0.019	0.110	0.035	0.297
OPH22	0.161	0.060	0.022	0.005	0.036	0.001	0.277	0.085	0.161
OPH30	0.302	0.274	0.148	0.108	0.152	0.371	0.398	0.192	0.290
OPH39	0.351	0.308	0.230	0.264	0.208	0.248	0.327	0.232	0.346
OPH40	0.036	0.040	0.044	0	0.017	0.006	0.028	0.099	0.144
Mean	0.143	0.126	0.087	0.049	0.125	0.105	0.201	0.111	0.197
\pm SD	\pm 0.130	\pm 0.106	\pm 0.075	\pm 0.094	\pm 0.080	\pm 0.137	\pm 0.126	\pm 0.069	\pm 0.103

Results

Genetic diversity, inbreeding, and null alleles

Ten seedlings and six juveniles sampled from TER were excluded due to the poor quality of DNA (Table 2). We found no significant genotypic disequilibrium between any pairs of loci, suggesting random allelic association between

pairwise loci within the three populations. Furthermore, this result shows that the developed microsatellite loci can be used in population genetic studies where no genetic association among loci needs to be assumed. The total number of alleles per locus (k) in all cohorts combined was highest in TER (110), with the lowest found in ESP (77), the most exploited popula-

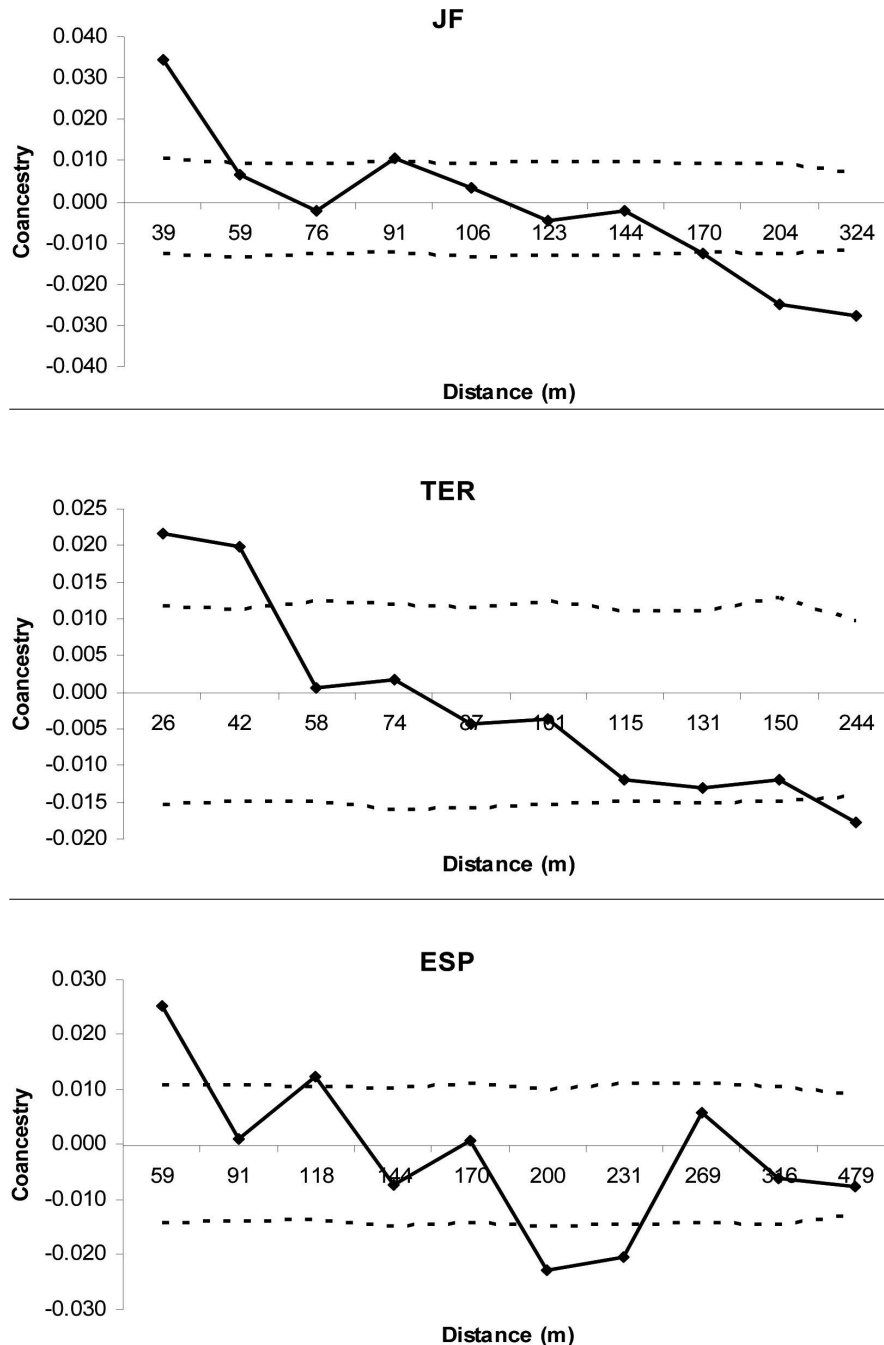


Figure 2. – Correlogram of the coancestry coefficient in eight distance classes for seedlings, juveniles, and adults of *O. phalerata* populations in Teresina (TER), Esperantina (ESP), and José Freitas (JF). The dotted lines indicate a 95% confidence interval (CI) around the null hypothesis that coancestry coefficient is not related to distance.

Table 4. – Estimated index of spatial genetic structure (seedlings, juveniles, and adults) from the three populations of *O. phalerata* (θ_1 is the coancestry coefficient in the first distance class; b_k is the slope of regression curve of coancestry coefficient on the logarithm of the distance between individuals (0–100 m); and S_p is the *Sp*-statistic; \pm SE is the 95% standard error; ** $P < 0.01$).

Sample	n	θ_1	$b_k \pm$ SE	S_p
José de Freitas: JF	90	0.0345**	-0.0202 \pm 0,0098	0.0209
Teresina: TER	76	0.0217**	-0.0164 \pm 0,0053	0.0168
Esperantina: ESP	90	0.0251**	-0.0197 \pm 0,0019	0.0202

tion (Table 2). We also detected 10, 11, and 13 private alleles in ESP, JF, and TER populations, respectively. The paired t-test detected significantly higher mean allelic richness (R) in JF than ESP ($P=0.040$), and in TER compared to ESP ($P=0.039$), as well as higher observed heterozygosity (H_o) and fixation index (F) for ESP than TER ($P=0.008$).

Among cohorts within the JF population, the index R was significantly greater for juveniles than seedlings ($P=0.046$), H_o was greater for adults than seedlings ($P=0.029$), and H_e was greater for adults and juveniles than seedlings ($P=0.047$, $P=0.041$, respectively). In TER, R was significantly higher in juveniles than seedlings ($P=0.046$), H_o was greater for adults than seedlings ($P=0.029$), and H_e was higher for adults and juveniles than seedlings ($P=0.047$, $P=0.041$, respectively). In ESP, no significant differences were observed for all indices. Null alleles seem to be present in practically all loci of seedlings, juveniles, and adults of the three populations, with the exception of locus OPH17 and OPH40 in seedlings of the TER population (Table 3). In JF, the expected null allele frequencies ranged among loci and samples from 0.001 to 0.351, in TER from 0 to 0.371, and in ESP from 0.028 to 0.398, indicating that the estimates of H_o , H_e , and F are likely biased. The corrected fixation index for null alleles (F_{Null}) decreased the fixation indexes for all samples of the populations (Table 2). Seedlings of all populations, juveniles of ESP, and adults of JF and TER presented significantly positive F_{Null} values, indicating inbreeding. The comparison of F_{Null} values between cohorts in ESP suggest that inbred individuals are eliminated between seedling and adult stages.

Spatial genetic structure (SGS)

The three populations showed significant differences for SGS (Fig. 2), reaching approximately 40, 42, and 60 m, in JF, TER, and ESP populations, respectively. These results suggest that near neighbours within each population are genetically related. The mean θ_1 values in the first distance class were significantly greater than zero, ranging from 0.0217 to 0.0345 (Table 4). The slope of regression curve (b_k) on the logarithm of the distance between pairs of individuals (0–100 m) and their average coancestry coefficient was significant, suggesting a dispersal pattern of isolation by distance. The extent of SGS, measured by the S_p statistic, was greater for ESP than JF and TER.

Genetic differentiation among populations

The genetic differentiation (G'_{ST}) between pairwise populations were significantly greater than zero ($P > 0.05$), but lower between the nearest populations (JF x TER: $G'_{ST} = 0.157$, 52 km; JF x ESP, $G'_{ST} = 0.379$; 128 km; TER x ESP: $G'_{ST} = 0.392$; 178 km), suggesting genetic isolation by distance.

Discussion

Genetic diversity

In comparison with other studies on palm and tree species in savanna regions, the eight microsatellite loci used in this study revealed a high level of genetic diversity within the three populations (H_e , ranging from 0.64 to 0.78). For example, from 25 *Dypterix alata* populations in the savanna, COLLEVATTI et al. (2013) found an

average H_e across eight microsatellite loci of only 0.43; yet for *Himatanthus drasticus*, a tree species, BALDAUF et al. (2013) found an average H_e across six microsatellite loci of 0.73 in adults and 0.68 in juveniles. In contrast, studying two populations of the Atlantic Forest palm *Euterpe edulis*, GAIOTTO et al. (2003) found an average H_e over 18 microsatellite loci of 0.86, a result higher than that found in the current study. *O. phalerata* has high genetic diversity which is likely due to a high outcrossing rate combined with the occurrence of protogyny, extensive pollen flow among populations via beetles and wind (ANDERSON and BALICK, 1988), as well as the high density of palms per hectare (on average 453.3 trees ha⁻¹).

However, the mean allelic richness (R) and observed heterozygosity (H_o) were significantly lower for the ESP population ($R=6.98$, $H_o=0.41$), which experiences more intense seed harvesting than TER, which is moderately exploited ($R=8.70$, $H_o=0.56$), and JF, which is the least exploited ($R=8.57$, $H_o=0.51$). In fact, all ontogenic stages (adults, juveniles, and seedlings) of ESP presented lower R and H_o values than those found for TER and JF. These results suggest that long term, intensive seed harvesting is decreasing genetic diversity in the ESP population. The intensity of fruit harvesting in ESP may create a genetic bottleneck for seedlings, resulting in regeneration from fewer parents, thus decreasing the genetic diversity. Alternatively, the low level of genetic diversity may be associated with population history, such as founder effect, genetic bottleneck, and drift.

Seedlings of all populations exhibited lower levels of genetic diversity than adults and juveniles (Table 2). This pattern of limited genetic diversity in seedlings is likely the result of genetic drift, as well as correlated mating and mating among relatives, considering that, due to protogyny, selfing is unlikely. Based on the corrected fixation index for null alleles (F_{Null}), seedlings from the three populations were inbred, although adults of JF and TER also showed evidence of inbreeding. In ESP, the absence of inbreeding for adults and significant inbreeding in seedlings and juveniles indicates selection against inbred individuals between seedling and adult stages. As all populations present SGS, the occurrence of inbreeding could be explained by the behaviour of pollinator vectors (the small beetle *M. mexicana*) that tend to forage pollen at greater frequency among near

neighbour and related plants. Selection between seed, seedling, and juvenile to adult stage in tree species is well documented. For example, in *Astrocaryum mexicanum* (EGUIARTE et al., 1992), *E. edulis* (GAIOTTO et al., 2003), and *Copaifera langsdorffii* (SEBBENN et al., 2011), higher levels of inbreeding were found in seeds and juveniles than in adults.

The expected occurrence of null alleles in the loci likely biased our estimates of heterozygosity and fixation index in all populations and cohorts. The F values decreased substantially after correction for null alleles (F_{Null}). Null alleles result in the absence of amplification of some alleles, leading to incorrect determination of some genotypes, increasing the estimates of homozygosity in populations and, consequently, the estimates of fixation index (CHYBICKI and BURCZYK, 2009). As inbreeding in *O. phalerata* is likely the result of mating among relatives, which is equal to the coancestry among parents, our results indicate that the inbreeding detected in some of the samples (Table 2, maximum F_{Null} of 0.15) is mainly the result of mating among half-sib parents (0.125).

Spatial genetic structure (SGS)

The significant SGS detected for the three studied populations is due to restricted seed dispersal. *O. phalerata* seed dispersal is primarily barochorous, although some secondary dispersal by rodent species is likely to occur (BONJORNE and GALETTI, 2007). Previous studies suggest that the habits of rodents can restrict the dispersal of seeds near to mother trees, resulting in a non-random pattern of seed dispersal. Our results support the observations of CHOO et al. (2012), who found that 99% of *O. phalerata* seeds and fruit are dispersed at distances shorter than 30 m from the dispersal point and 73% of seedlings are established up to 30 m away from the mother tree. These results are consistent with the SGS found in our current study which showed a seed dispersal range from 40 to 60 m (Fig. 2).

The extent of SGS (S_p) was greatest for JF and lowest for the TER population (Table 4). The limited extent of SGS in TER may be attributed to less frequent human intervention as compared to ESP. The result for JF is similar to ESP, the most exploited population, and could be due to the scattering behaviour of seed dispersers. It appears that less human inter-

vention may result in less significant impacts on populations of dispersing animals. All populations showed the presence of spatial genetic structure (SGS) and mating between related individuals resulted in inbreeding. However, studies on mating systems using open-pollinated progenies should be undertaken to confirm our conclusions.

Genetic differentiation among populations

Genetic differentiation between pairwise populations (G'_{ST} ranged from 0.157 to 0.391) showed that genetic diversity is mainly distributed within populations, as expected for outcrossing tropical tree species (LOVELESS and HAMRICK, 1984). The G'_{ST} was lowest between populations located in close proximity (JFxTER: $G'_{ST}=0.157$, 52 km) in comparison to more distant populations ($G'_{ST}=0.392$; 178 km). This suggests a pattern of isolation by distance in which greater spatial distances between populations result in greater genetic differentiation. This isolation can be attributed to decreasing levels of pollen and seed flow with an increase in the spatial distance. *O. phalerata* fruit and seeds are dispersed by barochory and rodents, such as the species *Dasyprocta agouti* (GALVEZ and KRANSTAUBER, 2009). *D. agouti* lives within a limited territory and moves between the same places; it is therefore an easy target for hunters, who have exploited the species both for food (TRINCA and FERRARI, 2006) and its hide (KASPER et al., 2007). Gene flow via seed dispersal is therefore limited to short distances. On the other hand, pollen is dispersed by wind and the beetle *M. mexicana* which enable pollen dispersal over longer distances than seeds. Even though *O. phalerata* pollen has the potential of being dispersed over great distances, the results indicate that long distances among populations could result in increased genetic differentiation.

Conclusions

The three *O. phalerata* populations studied herein maintained high levels of genetic diversity, which may be attributed to protogyny and high population density. However, the population that experiences the most intense fruit harvesting showed the lowest levels of genetic diversity, indicating that fruit harvesting may decrease genetic diversity in the species. Null alleles were present in practically all loci of

seedlings, juveniles, and adults, indicating bias in the estimates of H_o , H_e , and F . The corrected fixation index (F_{Null}) for null alleles decreased the fixation indices for all samples, resulting in significantly greater than zero fixation index for seedlings of all populations, but not for juveniles and adults of all populations. The comparison of F_{Null} values between cohorts in ESP suggests that inbred individuals are eliminated between seedling and adult stages. Significant SGS was detected up to 60 m in the populations, which indicates short distance seed dispersal. Genetic differentiation is greater for the most spatially distant populations, demonstrating a typical isolation by distance pattern of dispersal or low historical gene flow between distant populations. In practical terms, as all populations present high levels of genetic diversity and SGS, all can be used for seed collection for conservation and breeding purposes. Seeds for *ex situ* conservation, forest restoration, and breeding programs must be sampled from seed trees located at least 40 m apart to avoid seed collection from related mother trees.

Acknowledgments

This study was supported financially by the Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq, contract number: 564670/2010-3). We would like to thank REGINALDO PINHEIRO for his assistance in sample collection, GIULLIA FORTI and ELZA MARTINS FERRAZ for laboratory work, and JEFFERSON LORDELLO POLIZEL for map building. BRUNA IBANES was supported by a CNPq scholarship (382890/2012-4). ALEXANDRE MAGNO SEBBENN was supported by a research scholarship granted by the Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq). Finally, we thank Dr. EVELYN R. NIMMO for editing the English of the manuscript.

References

- AGUILAR, R., M. QUESADA, L. ASHWORTH, Y. HERRERIAS-DIEGO and J. LOCO (2008): Genetic consequences of habitat fragmentation in plant populations: susceptible signals in plant traits and methodological approaches. *Molecular Ecology* **17**: 5177–5188.
- ANDERSON, A. B. and M. J. BALICK (1988): Taxonomy of the babassu palm complex. *Systematic Botany* **13**: 32–50.

- BALDAUF, C., M. CIAMPI-GUILLARDI, F. A. M. SANTOS, A. P. SOUZA and A. M. SEBBENN (2013): Tapping latex and alleles? The impacts of latex and bark harvesting on the genetic diversity of *Himatanthus drasticus* (Apocynaceae). *Forest Ecology and Management* **310**: 434–441.
- BONJORNE, L. A. and M. GALETTI (2007): Seed dispersal and spatial distribution of *Attalea geraensis* (Arecaceae) in two remnants of Cerrado in South-eastern Brazil. *Acta Oecologica* **32**: 180–187.
- BUSO, G. S. C., A. Y. CIAMPI, M. C. MORETZSOHN, Z. P. S. AMARAL and R.V. BRONDANI (2003): Marcadores microssatélites em espécies vegetais. *Biocientífica Ciência e Desenvolvimento* **7**: 46–50.
- CAVERS, S., B. DEGEN, H. CARON, M. R. LEMES, R. MARGIS, F. SALGUEIRO and A. J. LOWE (2005): Optimal sampling strategy for estimation of spatial genetic structure in tree populations. *Heredity* **95**: 281–289.
- CHYBICKI, I. J. and J. BURCZYK (2009): Simultaneous estimation of null alleles and inbreeding coefficients. *Journal of Heredity* **100**: 106–113.
- CHOO, J., T. E. JUENGER and B. B. SIMPSON (2012): Consequences of frugivore-mediated seed dispersal for the spatial and genetic structure of a neotropical palm. *Molecular Ecology* **21**: 1019–1031.
- COLLEVATTI, R. C., M. P. C. TELLES, J. C. NABOUT, L. J. CHAVES and T. N. SOARES (2013): Demographic history and the low genetic diversity in *Dipteryx alata* (Fabaceae) from Brazilian Neotropical savannas. *Heredity* **111**: 97–105.
- CRESTE, S., A. TULMANN NETO and A. FIGUEIRA (2001): Detection of single sequence repeat polymorphism in denaturing polyacrylamide sequencing gels by silver staining. *Plant Molecular Biology Reporter* **19**: 299–306.
- DICK, C. W., O. J. HARDY, F. A. JONES and R. J. PETIT (2008): Spatial scales of pollen and seed-mediated gene flow in tropical rain forest trees. *Tropical Plant Biology* **1**: 20–33.
- DOYLE, J. J. and J. L. DOYLE (1990): Isolation of plant DNA from fresh tissue. *Focus* **12**: 13–15.
- EGUIARTE, L. E., N. PEREZ-NASSER and D. PIÑERO (1992): Genetic structure, outcrossing rate and heterosis in *Astrocaryum mexicanum* (tropical palm): implications for evolution and conservation. *Heredity* **69**: 217–228.
- EPPELSON, B. K. (1990): Spatial patterns of genetic variation within plant populations, pp. 229–253. *In: Plant Population genetics, breeding and genetic resources*, edited by A. H. D. BROWN, M. T. CLEGG, A. L. KAHLER and B. S. WEIR, Sinauer Associates Inc., Massachusetts.
- FRANKHAN, R. (2012): How closely does genetic diversity in finite populations conform to predictions of neutral theory? Large deficits in regions of low recombination. *Heredity* **168**: 167–178.
- GAJOTTO, F. A., D. GRATTAPAGLIA and R. VENCOVSKY (2003): Genetic structure, mating system, and long-distance gene flow in heart of palm (*Euterpe edulis* Mart.). *Journal of Heredity* **94**: 399–406.
- GALVEZ, D. and B. KRANSTAUBER (2009): Scatterhoarding by the Central American agouti: a test of optimal cache spacing theory. *Animal Behaviour* **78**: 1327–1333.
- GEBUREK, T. (2000): Effects of environmental population on the genetics of Forest trees, pp. 135–155. *In: Forest conservation genetics: principles and practice*, edited by A. YOUNG, D. BOSHIER and T. BOYLE, Csiro Publishing, Canberra.
- GOUDET, J. (1995): Fstat Version 1.2. A computer program to calculate F-statistics. *Heredity* **86**: 485–486.
- HAMRICK, J. L. (2004): Response of forest trees to global environmental changes. *Forest Ecology and Management* **197**: 323–335.
- HARDY, O. J. and X. VEKEMANS (2002): SPAGeDI: a versatile computer program to analyses spatial genetic structure at the individual or population levels. *Molecular Ecology Notes* **2**: 618–620.
- HARDY, O. J., L. MAGGIA, E. BANDO, P. BREYNE, H. CARON, M. E. CHEVALLIER, A. DOLIGEZ, C. DUTECH, A. KREMER, C. LATOCHE-HALLE, V. TROISPOUX, V. Veron and B. DEGEN (2006): Fine-scale genetic structure and gene dispersal influences in 10 Neotropical tree species. *Molecular Ecology* **15**: 559–571.
- HAUMANN, B. F. (1992): Here's a list of who's producing what. *Food Technology* **12**: 1284–1287.
- KASPER, C. B., F. D. MAZIM, J. B. G. SOARES, T. G. OLIVEIRA and M. E. FABIÁN (2007): Composição e abundância relativa dos mamíferos de médio e grande porte no Parque Estadual do Turvo, Rio Grande do Sul, Brasil. *Revista Brasileira de Zoologia* **24**: 1087–1100.
- LOISELLE, B. A., V. L. SORK, J. NASON and C. GRAHAM (1995): Spatial genetic structure of a tropical understory shrub, *Psychotria officinalis* (Rubiaceae). *American Journal of Botany* **82**: 1420–1425.
- LORENZI, H., L. NOBLICK, F. KAHN and E. FERREIRA (2010): *Flora Brasileira: Arecaeae – Palmeiras*. Plantarum, Nova Odessa.
- LOVELESS, M. D. and J. L. HAMRICK (1984): Ecological determinants of genetic structure in plant populations. *Annual Review of Ecology and Systematics* **15**: 65–95.
- MANOEL, R. O., P. F. ALVES, C. L. DOURADO, A. P. S. C. GAINO, M. L. M. FREITAS, M. L. T. MORAES and A. M. SEBBENN (2012): Contemporary pollen flow, mating patterns and effective population size inferred from paternity analysis in a small fragmented population of the Neotropical tree *Copaiifera langsdorffii* Desf. (Leguminosae-Caesalpinioideae). *Conservation Genetics* **13**: 613–623.
- RAFALSKI, J. A., J. M. VOGEL, M. MORGANTE, W. POWELL, C. ANDRE, and S. V. TINGEY (1996): Generating and using DNA markers in plants pp. 75–134. *In: Non mammalian Genomic Analysis: A Practical Guide*, edited by B. BIRREN and E. LAI, Academic Press, San Diego.

- SEBBENN, A. M., A. C. M. CARVALHO, M. L. M. FREITAS, S. M. B. MORAES, A. P. S. C. GAINO, J. M. SILVA, C. JOLIVET and M. L. T. MORAES (2011): Low levels of realized seed and pollen gene flow and strong spatial genetic structure in a small, isolated and fragmented population of the tropical tree *Copaifera langsdorffii* Desf. *Heredity* **106**: 134–145.
- TRINCA, C. T. and S. F. FERRARI (2006): Caça em assentamento rural na amazônia matogrossense, pp. 155–167. *In*: Diálogos em ambiente e sociedade no Brasil, edited by P. JACOBI and L. C. FERREIRA, Annablume, Pinheiros.
- VEKEMANS, X. and O. J. HARDY (2004): New insights from fine-scale spatial genetic structure analyses in plant populations. *Molecular Ecology* **13**: 921–935.
- YOUNG, A. G. and T. J. BOYLE (2000): Forest fragmentation, pp. 123–134. *In*: Forest conservation genetics: Principles and practice, edited by A. YOUNG, D. BOSHIER and T. BOYLE, Csiro Publishing, Canberra.

Development of PCR based markers in Terpene synthase genes for marker assisted selection of high resin yielders in *Pinus roxburghii* Sarg.

S. ALLEN^{*}, H. S. GINWAL and S. BARTHWAL

Division of Genetics and Tree propagation, Forest Research Institute, Dehradun 248006 India

(Received 27th August 2015)

Abstract

Pine oleoresin is a complex mixture of volatile and nonvolatile terpenes and is exploited for commercial production in India. Pines have long reproductive cycle vegetative phase extending upto many years. Therefore, there is an urgent need to identify high resin pine yielders at an early stage. Due to the strong influence of genetic factors on resin yield in pine species, marker assisted selection may serve as a potential tool for early identification of the genotype of interest. This study was planned to identify high resin pine yielders at early stage to avoid exploitation of natural germplasm of chir pine. Protein sequences of different terpene synthases were downloaded from the NCBI database and were multiply aligned to identify conserved and variable regions across the sequences. Primers were designed based on the related coding sequences to target these regions using Primer 3 software to amplify and isolate related genomic loci in chir pine genotypes. These primers were characterized for specificity using Primer Blast and the presence of hairpin loop formation using OligoCalc software. DNA

was extracted from different resin yielding pine genotypes and primers were run on them for molecular studies. With the help of bioinformatics tools, we were able to associate some markers with resin yield. This study holds a key promise for the conservation of pine germplasm in natural habitat.

Key words: *Pinus roxburghii* Sarg, Terpene synthase, PCR-marker, Bioinformatics, association analysis.

Introduction

Pinus roxburghii Sarg., commonly known as Chir pine, belongs to family Pinaceae, and is one of the most extensively distributed pine species in India which is mainly exploited for resin. The forests of *P. roxburghii* are found ranging from longitudes 70°E to 93°E and latitudes 26°N to 36°N in subtropical and warm temperate monsoon belts, between 450 to 2300 m altitudes in Siwaliks and Himalayan main river valleys, from Kashmir to Bhutan. 'Naval Stores' is a term used to denote the products obtained from the oleoresin or resin of pine trees (COPPEN and HONE, 1995). Oleoresin is a complex mixture of terpenoids (KEELING and

^{*}) Corresponding author: allenswati@gmail.com