ASSESSMENT OF GENETIC STRUCTURE OF THE ENDANGERED FOREST SPECIES *BOSWELLIA SERRATA* ROXB. POPULATION IN CENTRAL INDIA

Vivek Vaishnav[1], Shashank Mahesh[2], Pramod Kumar[3]*

[1]Institute of Forest Productivity, Ranchi, Jharkhand, India, 835303
[1]CSIR- National Botanical Research Institute, Lucknow, India, 226601
[2]National Institute for Research in Tribal Health, Jabalpur, Madhya Pradesh, 482001
[3]Tropical Forest Research Institute, Jabalpur, Madhya Pradesh, 482021
** Corresponding author: pramod_kt@rediffmail.com*

ABSTRACT

Boswellia serrata Roxb., a commercially important species for its pulp and pharmaceutical properties was sampled from three locations representing its natural distribution in central India for genetic characterization through 56 RAPD + 42 ISSR loci. The wood fiber dimensions measured for morphometric characterization confirmed 11.36% of the variation in the length and 8.75% of the variation in the width indicating its fitness for local adaptation. Bayesian and non-Bayesian approach based diversity measures resulted moderate within population gene diversity (0.26±0.17), Shannon's information index (0.40±0.22) and panmictic heterozygosity (0.28±0.01). A high estimate for genetic differentiation measures i.e. G_{ST} (0.31), G_{ST}-B (0.33±0.02) and θ-II (0.45) led to the distinct clusters of the sampled genotypes representing their regional variability due to limited gene flow and total absence of natural regeneration. We report the first investigation of the species for its molecular characterization emphasizing the urgent need for the genetic improvement program for the *In-situ*/*Ex-situ* conservation and sustainable commercialization.

Keywords: Burseraceae, θ-statistics, Shannon's information index, genetic improvement

INTRODUCTION

Boswellia serrata Roxb. (family- Burseraceae), commonly known as salai guggul or Indian frankincense (olibanum indicum) is a commercially important deciduous tree of India (Shah et al. 2008). It is found in the region having rainfall between 500mm -2000mm and temperature up to 45 °C. The species has the ability to thrive in the poorest and the shallowest soils (Bhat et al. 1952). In India, it is distributed in states Rajasthan, Maharashtra, Madhya Pradesh, Karnataka and Chhattisgarh (Pawar et al. 2012). The oleo-gum-resin of the tree contains boswellic acid, which is effective for the treatment of inflammatory disorders, arthritis, cardiovascular diseases (Ammon 2006), diarrhea, dysentery and other skin diseases (Khare 2004). Apart from these medicinal values, the resin of the species was found a more effective sizing agent in papermaking than the rosin obtained from *Pinus* species (Sharma et al. 1985). The convincing wood quality for pulp making led to the establishment of the first paper mill in India in Nepanagar of Madhya Pradesh (Khan 1972) as the region had been occupied by wide range of natural patch of the species.

The species is highly out-crossing supported by the self-incompatibility to selfing (Sunnichan et al. 2005). But, a poor fruit setting (2.6% - 10%) under open pollination condition, inadequate production of viable seeds and scanty seed germination (10% - 20%) limit the distribution of the species in nature (Ghorpade et al. 2010). The species population has been harvested for the frankincence also (Sharma 1983). The scarcity of protocols to regenerate it through seeds and clones makes the mass multiplication difficult (Purohit et al. 1995). This situation resulted in declined abundance of the species and the International Union for Conservation of Nature (IUCN) has enlisted it under the status of endangered. Therefore, the available natural patches of the species require keen attention for both *In-situ*

48 and *Ex-situ* conservation with actual knowledge of available genetic resources of the species
49 for management and breeding objectives.

50 Information on actual genetic variability and a cryptic number of the genetically
51 differentiated genetic resource of any species is an important aid for its conservation and
52 genetic improvement. For such purposes, DNA based markers are of value because unlike the
53 morphometric traits the molecular markers are independent of the influence of environment
54 and can be assessed at any growth stage. Among the various molecular markers employed to
55 assess diversity studies, PCR-based dominant markers such as RAPD (random amplified
56 polymorphic DNA, Williams et al. 1990) and ISSR (inter-simple sequence repeat,
57 Zietkiewiez et al. 1994) have become popular due to its polymorphism and discrimination
58 power, as their application does not need any prior sequence information. These primer
59 systems have been successfully applied for the genetic characterization of populations of
60 tropical tree species (Ansari et al. 2012, Abuduli et al. 2016, Vaishnav & Ansari 2018).
61 Bayesian statistics has been extended to dominant markers for precise estimates of population
62 genetic hierarchies equivalent to those obtained with codominant markers, circumventing
63 inbreeding estimate within the population (Holsinger et al. 2002). Therefore, the present
64 investigation was conducted to differentiate the morphometric and the genetic variability
65 exhibited by the natural population of *B. serrata* Roxb. in three agro-climatic regions of
66 central India applying dominant marker system.

67 **MATERIAL AND METHODS**

68 **Population sampling**

69 The information of the distribution of the patches or population of the species was obtained
70 from the old forest working plans of the concerned forest departments of the Indian state

Chhattisgarh, which is classified in three agro-climatic zones *viz.* Northern Hills, Chhattisgarh plain and Bastar plateau (Table 1).

Table 1. Geo-climatic conditions of the three regions of natural distribution of *B. serrata* Roxb. populations sampled for the investigation

Regions	Agro-climatic zones	GPS coordinate ranges		Altitude (m)	Mean annual temperature (^{0}C)		Annual Precipitation (mm)
		Latitude (N)	Longitude (E)		Min	Max	
Dhamtari (DT)	Chhattisgarh Plain	20.37861-20.40511	81.96636-81.98122	445.80 ±5.44	21.60 ±5.22	33.02 ±3.56	1488±0
Narayanpur (NP)	Bastar Plateau	19.71365-19.71586	81.19447-81.19836	595.60 ±21.39	20.80 ±5.22	32.40 ±3.36	1628±0
Sarguja (SG)	Northern Hills	23.39300-23.39446	83.46277-83.46783	619.75 ±7.15	19.02 ±6.35	31.02 ±4.07	1670±0

± - Standard deviation (SD)

Forests of these agro-climatic regions were visited during July 2014 to September 2014 to collect samples from 20 trees of each region in natural distribution of *B. serrata* respecting inter-tree distance of 100 m along the latitude (Figure 1).

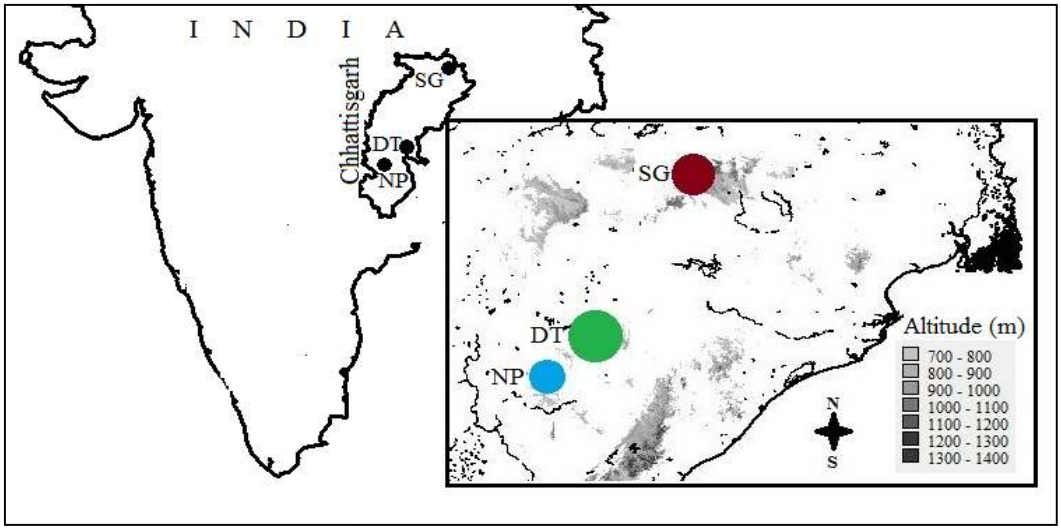


Figure 1 Three regions i.e. Sarguja (SG), Dhamatri (DT) and Narayanpur (NP) of natural distribution of *B. serrata* Roxb from central Indian region (Chhattisgarh) were sampled for the investigation. Every location represents distinct agro-climatic zone namely Northern hills (SG), Chhattisgarh plain (DT) and Bastar plateau (NP). An inset view of altitudinal variation is given in the box and size differences among the location's legend dot depicts differences in gene diversity of the species estimated employing the dominant markers

83 The girth at breast height (GBH) was measured by measuring tape. A wood radial core
84 sample was extracted from each tree at breast height (1.34 m) and was stored in a tube with
85 40% formaldehyde. A leaf sample was also collected from each tree. The samples were
86 brought to the laboratory in a cryo-box.

87 **Measurement of wood fiber dimensions**

88 In the laboratory, the wood samples were macerated following the method described in
89 Mahesh et al. (2015). Five slides were prepared for each tree and the wood fiber length
90 (WFL) and width (WFW) of five fibers per slide were measured through the compatible
91 program integrated with the compound microscope (Leica, EC3, Switzerland).

92 **DNA isolation and amplification**

93 DNA from the leaf samples was isolated following modified CTAB method (Deshmukh et al.
94 2007). To select the reproducible markers, 10 RAPD and 10 ISSR primers were cross-
95 amplified on set of 12 genotypes of the species (four from each region) and finally, only five
96 from each of RAPD and ISSR markers were selected based on their polymorphism and
97 reproducibility for the final amplification of all sixty genotypes (Table 3). For RAPD/ISSR
98 amplification, the final reaction mixture for PCR assay was consisted of 15 µl/10 µl
99 containing 50 ng/ 40ng genomic DNA, 0.66 µM/0.8 µM of primers, 0.2 mM/ 0.1mM of
100 dNTP mix, 1.5 mM/2.5 mM $MgCl_2$, 1X buffer with KCl and 1 unit of Taq polymerase. PCR
101 parameters included an initial 4 minutes denaturation step at 94 °C followed by 35 cycles of
102 30s at 94 °C, 30s at the annealing temperature of 35 °C/50 °C, 45s extension at 72 °C
103 followed by a final extension of 5 minutes at 72 °C. The amplified products were
104 electrophoresed on 1.5% agarose gel containing 0.5µg/ml ethidium bromide (EtBr) in 0.5 X
105 TBE (pH 8.0). The separation was carried out by applying constant 100V for 3 hours. The
106 fractionated amplified products on agarose gel were visualized on gel documentation system

107 under UV light. To avoid the homology, the amplification profile of all sixty genotypes was
108 evaluated through the molecular weight of the bands for each primer and the bands were
109 scored in a Microsoft Excel sheet in binary data format.

110 **Data analysis**

111 The measures of central tendency and the coefficient of variation (CV) were calculated for
112 GBH, WFL, and WFW. A genetic profile of 60 genotypes on 10 markers was generated and
113 analyzed following both band-based and allele frequency-based approaches as suggested in
114 Bonin et al. (2007). The information from the markers was evaluated estimating mean allelic
115 frequency (AF), gene diversity (GD), and polymorphic information content (PIC) by program
116 POWERMARKER v3.1 (Liu & Muse 2005). Resolving power (RP) of the markers was
117 calculated following the formula given by Prevost and Wilkinson (1999).

118 The genetic variability and the differentiation of the sampled population were evaluated
119 applying both non-Bayesian and Bayesian approaches. The program POPGENE v1.32 (Yeh
120 et al. 1999) was applied for the non-Bayesian estimates of genetic diversity measures i.e.
121 Nei's gene diversity (GD), Shannon's information index (I) and the percentage of
122 polymorphism (P%) for each location and G_{ST} was calculated to evaluate the level of genetic
123 differentiation. The genetic relationship among the genotypes was calculated based on
124 Jaccard's genetic similarity coefficient by the program DARWIN v5.0 (Perrier &
125 Jacquemoud-Collet 2006). The principal coordinate analysis (PCoA) was performed to
126 cluster the genotypes in different axes applying program GENALEX v6.5 (Peakall & Smouse
127 2012). Analysis of molecular variance (AMOVA) was performed applying program
128 ARLEQUIN v3.11 (Excoffier & Lischer 2010) to estimate the hierarchical variation. The
129 isolation by distance (IBD) was confirmed calculating the correlation coefficient between two
130 pairwise matrices of genetic and geographical distances among genotypes through Mantel's

131 test performed by program Alleles in Space (Miller 2005) in independent runs (with and
132 without logarithmic transformation) on 10,000 permutations.

133 For the Bayesian approach based genetic estimates we preferred the θ-statistics (*viz.* θ-I,
134 θ-II, θ-III and G_{ST}-B) implemented in program HICKORY v1.1 (Holsinger & Lewis 2007)
135 that allows direct estimation of genetic differentiation measure (F_{ST}) from dominant markers
136 without assumption of prior knowledge of the extent of inbreeding and Hardy –Weinberg
137 Equilibrium (HWE) within population even with small size and number. Region-wise
138 panmictic heterozygosity (Hs) and total panmictic heterozygosity (Ht) were also estimated.
139 All these estimations were performed with 50,000 steps of burn-in, 500,000 replicates,
140 'thinning' = 20 following the best-suited model for the data amongst the all four models (*viz.*
141 full, f = 0, θ = 0 and f-free models) implemented in the program based on the minimum DIC
142 (deviance information criterion) value as suggested by Spiegelhalter et al. (2002). To
143 determine the most suitable number of cryptic populations (K) for the dataset, the program
144 STRUCTURE v2.3.1 (Pritchard et al. 2000) was run applying model combinations of
145 admixture/ no-admixture model with correlated/ independent allele frequencies among the
146 samples on 100000 burn-in and 1000000 MCMC repeats with three-run from K=1 to K=9,
147 for number of K $2 \leq K \leq 8$. The most suitable model and the best-suited 'K' was determined
148 based on the highest Delta-K value resulted by online program STRUCTURE HARVESTER
149 (Earl and VonHoldt 2012).

150

151 **RESULTS**

152 **Variation in wood fiber dimension**

153 The WFL of the species ranged between 0.803 mm to 1.397 mm with average of 0.968±0.11
154 mm (11.36% CV) and the WFW of the species ranged between 0.019 mm to 0.030 mm with
155 average of 0.025±0.002 mm (8.75% CV). Both traits followed normal distribution and no
156 significant correlation was observed between them. WFL and WFW both were evenly
157 distributed on the standard deviation (σ) based classes and no significant variation was
158 observed among the locations based on WFL and WFW (Table 2).

Table 2. Variability in wood trait parameters of three sub-populations of *B. serrata* Roxb.

Regions	GBH (m)	WFL (mm)	WFW (mm)
Dhamtari (DT)	0.99 ±0.25	0.919 ±0.097	0.025 ±0.001
Narayanpur (NP)	1.07 ±0.22	0.975 ±0.095	0.026 ±0.002
Sarguja (SG)	1.40 ±0.25	1.009 ±0.120	0.024 ±0.003
Average	1.15 ±0.30	0.968 ±0.11	0.025 ±0.002
CV (%)	26.01	11.36	8.75

GBH- girth at breast height, WFL- wood fiber length, WFW- wood fiber width, CV- coefficient of variation, ±- standard deviation (SD)

159 **Genetic informativeness of markers**

160 The five RAPD primers amplified 826 bands and 56 loci (14.75 bands/locus). The diversity
161 measures AF and GD were 0.81±0.065 and 0.27±0.075 respectively. PIC and RP were
162 0.22±0.054 and 5.51±3.17 subsequently (Table 3).

163

Table 3. Genetic informativeness of RAPD and ISSR markers

Marker System	Primers	Sequence (5' – 3')	N	AF	GD	PIC	RP
RAPD	OPA1	CAGGCCCTTC	12	0.77	0.32	0.26	6.10
	OPA2	TGCCGAGCTG	13	0.72	0.37	0.29	10.80
	OPY1	GTGGCATCTC	10	0.80	0.27	0.22	3.97
	OPY2	CATCGCCGCA	12	0.87	0.19	0.16	3.70
	OPP3	CTGATACGCC	9	0.87	0.21	0.18	2.97
	Average		11.20	0.81	0.27	0.22	5.51
ISSR	UBC840	GAGAGAGAGAGAGAGAAT	9	0.67	0.42	0.33	9.47
	UBC830	TGTGTGTGTGTGTGTGG	6	0.74	0.36	0.29	6.50
	UBC822	TCTCTCTCTCTCTCTCA	8	0.74	0.37	0.30	6.50
	UBC845	CTCTCTCTCTCTCTCTGG	8	0.84	0.24	0.20	6.17
	UBC836	AGAGAGAGAGAGAGAGTA	11	0.80	0.28	0.23	4.83
	Average		8.40	0.76	0.33	0.27	6.69

N- number of amplified loci, AF- allele frequency, GD- gene diversity, PIC- polymorphic information content, RP - resolving power

For RAPD markers, the diversity measures and discriminatory power measures were found in significant ($p<0.05$) linear correlation. The five ISSR primers amplified 1004 bands for 42 loci (23.90 bands/locus). In a significant linear correlation ($p<0.01$) to each other, the AF was 0.76±0.068 and the GD was 0.33 ±0.073. The PIC and the RP were 0.27 ±0.052 and 6.69±1.69 subsequently (Table 3).

Genetic diversity and differentiation

Among the non-Bayesian estimates of genetic diversity measure for the species population, 100% polymorphism, 0.26±0.17 GD, and 0.40±0.22 I were observed among the regional locations (Table 4). For all measures, no significant difference was found among the three regions. Dhamtari was found slightly more diverse and polymorphic than the other two regions (Table 4). The G_{ST} was 0.31. The Bayesian model implemented in Hickory estimated the lowest DIC value (896.96) for the full model and the panmictic heterozygosity was 0.28±0.01. No significant variation in panmictic heterozygosity was found among the three regional sub-populations (Table 4). The θ-I, θ-II, and θ-III values were 0.50±0.04, 0.45±0.04 and 0.25±0.01 subsequently and the G_{ST}-B was 0.33±0.02.

Table 4. Population genetics parameters for 60 genotypes sampled from three sub-populations of *B. serrata* Roxb.

Regions	P%	GD	I	Hs
Dhamtari (DT)	63.27	0.20 ±0.20	0.30 ±0.28	0.20 ±0.01
Narayanpur (NP)	57.14	0.19 ±0.19	0.28 ±0.28	0.19 ±0.01
Sarguja (SG)	55.10	0.14 ±0.16	0.22 ±0.24	0.17 ±0.01
Overall*	100	0.26 ±0.17	0.40 ±0.22	0.28 ±0.01

±- SD, P% - polymorphism%, GD - Nei (1973) gene diversity, I - Shannon's information index, Hs - panmictic heterozygosity resulted by Hickory, * considering all 60 genotypes as a group

The dendrograph based on Jaccard's genetic similarity coefficient (average 0.67) grouped the genotypes following their regions (Figure 2). The genotypes sampled from Sarguja further bifurcated into two different but conjoint clusters supported by the bootstrap values >60 (Figure 2).

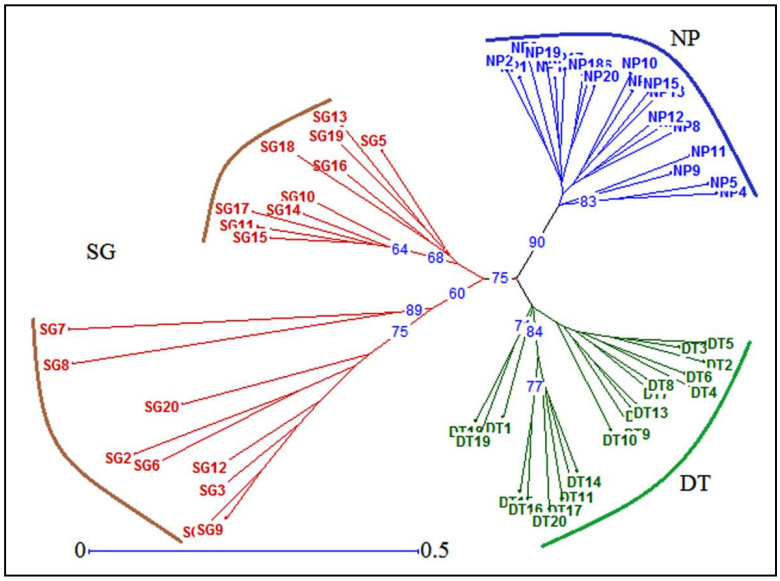
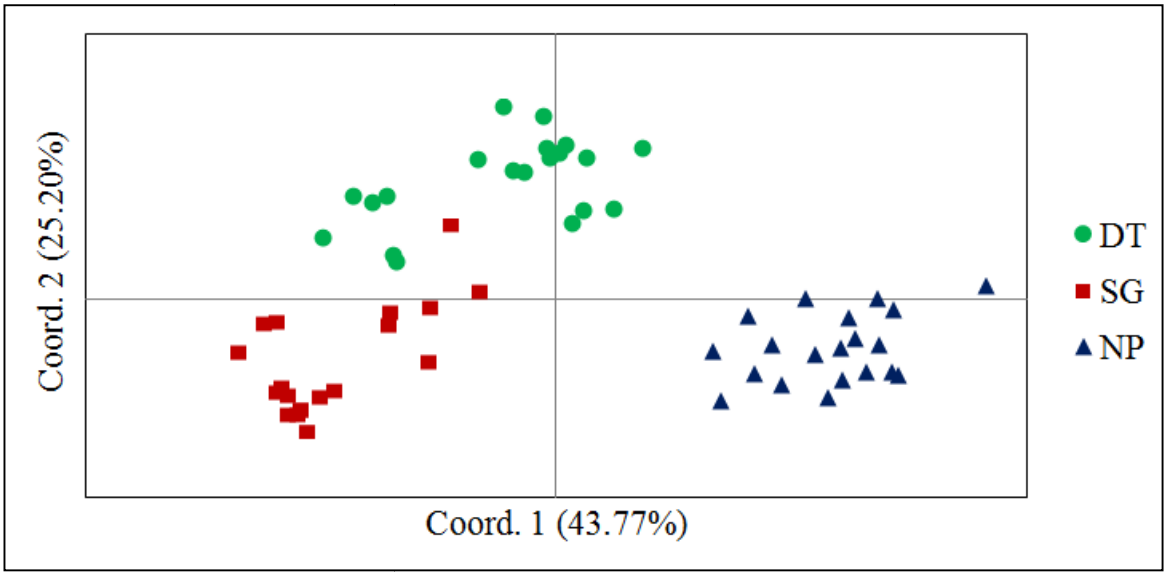


Figure 2 A Jaccard's similarity coefficient based radial dendrograph of 60 accessions from three natural populations i.e. Sarguja (SG), Dhamtari (DT), and Narayanpur (NP) of *B. serrata* Roxb. Bootstrap values (>60) are shown on nodes of accessions

188 The PCoA accounting for 68.97% cumulative separation based on the same genetic
189 distance separated the genotypes of Narayanpur from rest of the sampled genotypes and the
190 equal proportion of the genotypes from Dhamatari and Sarguja was in the separate group
191 from their sampled locations making overall four small clusters (Figure 3).



192

193 **Figure 3** The principal coordinate analysis, showing closeness of genotypes of *B. serrata*
194 Roxb sampled from Dhamtari (DT) and Sarguja (SG) and the genotypes from Narayanpur
195 (NP) cluster distinctly, in two coordinates covering 68.97% of variation

196 The Bayesian program STRUCTURE assigned K=2 as the most appropriate number of
197 cryptic population for the sampled genotypes based on the highest deltaK value in
198 STRUCTURE HARVESTER. The bar plot generated (Figure 4) shows admixing of
199 genotypes assigned to the two cryptic populations (K=2) with >80% of ancestry coefficient.
200 The genotypes from Dhamtari were found admixed with both of the cryptic populations, the
201 genotypes from Narayanpur and Dhamatri were assigned to be in the same cluster. On the
202 other hand, genotypes from Sarguja clustered distinctly (Figure 4). AMOVA resulted
203 46.03±0.65% variation among groupings suggested by dendrograph, PCoA, and STRUTURE
204 (K=2) and 53.38±0.15% variation within these groups. Mantel's test found significant

205 (p<0.001) correlation between pair-wise genetic and geographical distances among the
206 genotypes.

207 **DISCUSSION**

208 *B. serrata* Roxb. a member of family Burseraceae, bifurcated from "Terebinthaceae" along
209 with Anacardiaceae during the Miocene epoch due to lesser intrinsic ability to adopt the
210 changing conditions across the climatic niches (Weeks et al. 2014). Therefore, the
211 distribution of genus *Boswellia* limited within the mediterranian and tropical regions. *B.*
212 *serrata* Roxb. is found only in central India and the species is renowned for its olio-gum-resin
213 and pulp content. Due to uncontrolled harvesting, it has become scarce in nature and the
214 literature regarding its distribution is also not adequate to locate it. During the investigation,
215 the species was found in the temperature range of 19.02±6.35 °C to 33.02±4.07 °C and the
216 annual rainfall of 1595.33±95.30 mm confirming its suitability for moist conditions along
217 with the arid regions. For molecular characterization of the species, the RAPD and ISSR
218 markers are preferred due to the absence of background genomic information of the species
219 to design microsatellites and low-cost applicability. The problems with the dominant markers
220 such as low reproducibility and uncertain homology of alleles have been handled during
221 amplification and data profiling. Bayesian model-based θ-statistics has been applied based on
222 its successful estimation of the genetic variability and the differentiation of population
223 applying dominant markers system (Holsinger & Wallace 2004) in forest species as well (Ci
224 et al. 2008,Vaishnaw et al. 2015).

225 **Variation in wood fiber dimension**

226 The wood fiber length along with the thickness is considered to be the most important
227 characteristics for the mechanical strength of wood (Migneault et al. 2008) and for the
228 making of high-quality groundwood in paper industries (Zobel & Jett 1995). The comparison

229 with the other tropical trees employed in pulp making process establishes that the wood fiber
230 dimensions of *B. serrata* Roxb. obtained in our investigation is much higher than those of
231 *Ailanthus altissima* (WFL= 0.747±0.035 mm and WFW= 0.023±0.0004 mm) and *Eucalyptus*
232 *globulus* (WFL = 0.785±0.012 mm and WFW= 0.018±0.0005) reported in Baptista et al.
233 (2014). The WFL is lower than of *Eucalyptus grandis* (1.01±0.14 mm) reported in Bhat et al.
234 (1990). The coefficient of variation of wood fiber dimension in *B. serrata* is higher than *A.*
235 *altissima* (WFL= 4.76% and WFW= 1.66%) and *E. globulus* (WFL= 1.5% and WFW=
236 2.72%) and lower than *E. grandis* (WFL= 14.45%) in above studies. In terms of adaptability,
237 such high variation confirms the adaptive fitness of the population to the local ecological
238 conditions (MacPherson et al. 2015).

239 **Genetic informativeness of the primers**

240 With an approach to handle the limitation of the non-reproducibility, the dominant markers
241 (RAPD/ISSR) have exhibited moderate values for AF and GD. It indicates that the approach
242 avoiding null alleles have controlled the informativeness of the primers helping genetic
243 differentiation (Lynch & Milligan 1994). The AF exhibited by the RAPD/ISSR markers are
244 slightly higher to the information resulted in populations of *Tectona grandis* L. f. sampled
245 from India on ISSR markers (0.73±0.02, Vaishnav & Ansari 2018). On the other hand,
246 negligible differences in AF among the markers of RAPD and ISSR systems ensure the
247 qualitative support for the diversity assessment (Lynch & Milligan 1994, Nybom & Bartish
248 2000). The discriminatory power measures exhibit a strong and linear relationship between
249 the ability of a primer to distinguish genotypes (Prevost & Wilkinson 1999). The values of
250 the PIC and RP are comparable to those obtained with dominant markers in other species, e.g.
251 *Podophyllum hexandrum* (Naik et al. 2010), *Jatropha curcas* (Grativol et al. 2011) and
252 *Pongamia pinnata* (Sharma et al. 2014).

253 **Genetic diversity and differentiation**

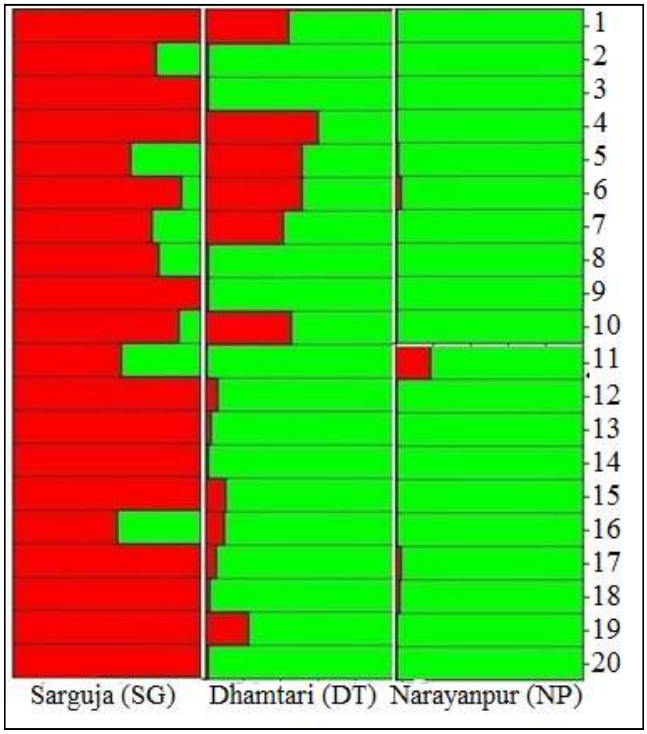
254 In order to estimate the genetic diversity of *B. serrata* Roxb. population, we applied both;
255 band-based and allele frequency-based approach with non-Bayesian and Bayesian statistics
256 so that the under-estimation or over-estimation of the diversity measures caused by the
257 dominant marker systems could be avoided. A significant correlation between diversity
258 measures i.e. P%, GD, I and Hs; with slight differences among them for different regions,
259 confirms almost equal heterogeneity of the species population. The gene diversity of the
260 sampled population has been found slightly higher than that reported by Nybom (2004) for
261 out-crossing species (0.18). Nevertheless, Holsinger & Wallace (2007) suggested avoiding
262 the comparison among the diversity measures resulted by the neutral markers (AFLP/RAPD)
263 as these markers only show the differences in rates of mutations at their loci rather than
264 reflecting different migration rates. However, there are few studies based on the tree species
265 of the central Indian regions *viz.* Wang et al. (2011) investigated the genetic diversity of *D.*
266 *sissoo* sampled from Indian regions and found 89.11% of polymorphism, 0.27±0.16 GD and
267 0.41±0.23 I. Ansari et al. (2012) sampled *Tectona grandis* L. f. from all along the central and
268 peninsular Indian regions and found 80.30% polymorphism, 0.32 GD and 0.45 I. The
269 comparison of the results from present investigation with above studies indicates that the
270 species natural distribution and the level of genetic diversity exhibited are somehow equally
271 influenced by the geographical conditions and the landscapes. The significant correlation
272 found between genetic and geographical distances of the genotypes in our result also supports
273 the fact.

274 The genetic structure of a population reflects the long-term evolutionary history of the
275 species along with its mating system (Duminil et al. 2007), reproductive biology and gene
276 flow influenced by the adaptive selection and fragmentation (Slatkin 1987, Nybom & Bartish

277 2000). In our investigation, the results of Bayesian model-based investigation through
278 HICKORY found full-model as the most suitable for our data determining the influence of
279 inbreeding within the populations. The low level of panmictic heterozygosity (Hs) for
280 different regions and for overall population also supports it. Like other out-crossing species
281 (Hamrick & Godt 1989; Nybom 2004) the *B. serrata* Roxb also exhibited higher within-
282 population variability than among populations but the genetic variation within the population
283 resulted by AMOVA was found comparatively lower than those reported in other out-
284 crossing forest species (Nybom 2004, Wang et al. 2011, Ansari et al. 2012). A limited gene
285 flow in *B. serrata* Roxb. is one of the important factors that have led to the high
286 differentiation among sampled populations and the genotypes. The species has an
287 entomophilous mode of pollination and the self-incompatible flowers support only cross-
288 pollinated pollen to grow the pollen tubes. Moreover, only up to 10% fruit set has been
289 observed in open-pollination condition (Sunnichan et al. 2005). In our field observation also,
290 we noticed earlier abscission of unripe fruits and ripen fruit with empty seeds led to the total
291 absence of natural regeneration in all sites of sampling. These bottlenecks contribute to
292 genetic drift and signatory inbreeding depression to the species population reflected through
293 different estimates.

294 The genetic differentiation measures G_{ST} and G_{ST}-B have resulted in comparatively equal
295 values indicating the cautions taken to employ the dominant markers for the investigation
296 could be equivalent to the Bayesian model correction. The θ-statistics measures have
297 exhibited differences in their values due to a low number of populations sampled for the
298 investigation (Holsinger & Wallace 2004). The θ-I is supposed to be equivalent to Wright's
299 F_{ST} (Wright 1969) and its higher value (0.50±0.04) is expected because it results in the scaled
300 allele frequency measured across the evolutionary time. The θ-II is a more appropriate

301 measure for a small number of admixed population and its higher value (0.45±0.004) reflects
302 the problematic gene exchange among the genotypes leading to high differentiation.



303

304 **Figure 4** The Bayesian clustering resulted by program STRUCTURE depicted two cryptic
305 populations (K) in 60 genotypes of *B. serrata* Roxb sampled from three locations of its
306 natural distribution. Each genotype is presented by a separate colored bar with corresponding
307 number. Green and Red color represents two different cryptic populations.

308 Among the Bayesian clusterings resulted by STRUCTURE, K=2 has been found the most
309 suitable for our data on the basis of deltaK value. Nevertheless, the presence of inbreeding in
310 sampled population and deviation from the HWE suggest preferring the dendrograph and the
311 PCoA over the two admixed cryptic populations resulted by the program STRUCTURE
312 because of its prior assumption of HWE. In PCoA, We have found Sarguja population closely
313 related to Dhamtari population and the Narayanpur population has represented itself
314 distinctly even after geographical closeness to Dhamatri. It might be due to the gene flow
315 among these populations follows geographical limits along with their own biological
316 constraints. On the other hand, the Narayanpur population might represent a different source
317 of genetic origin of southern Indian region that could not be sampled. An almost equal

318 polymorphism, genetic diversity, and heterozygosity but with highly differentiable genetic
319 resource resulting to the distinctly different clusters indicate that the sampled populations
320 may belong to a large metapopulation of the species that has been fragmented with due
321 course of time due to its own intrinsic inability to cope up with natural and anthropogenic
322 pressures later leading to geographical isolation.

323 *B. serrata* Roxb. is a highly out-crossing species with self-incompatibility and therefore,
324 its own constraint to exchange the genes and the fragmentation of the population due to
325 geographical isolation and commercial harvesting might be the reason for detected inbreeding
326 pressure. The estimates of genetic diversity measures, genetic structure, and level of
327 differentiation of a species population depend on the molecular marker system employed for
328 the investigation. In our investigation, the dominant markers might be a reason behind low to
329 moderate estimates of gene diversity, panmictic heterozygosity, and within-population
330 variability. On the other hand, the Bayesian model-based analysis has given a higher value of
331 θ-II because the θ-statistics has been developed specifically for the dominant markers
332 (Spiegelhalter et al. 2002) and has been found more appropriate than G_{ST} and G_{ST}-B to
333 estimate the inbreeding and population differentiation (Holsinger et al. 2002, Holsinger &
334 Wallace 2004). The comparatively higher variation in morphometric traits also indicates the
335 existence of higher variability in the population. The northernmost region Sarguja and the
336 southernmost region Narayanpur might represent different centers of genetic origin of a large
337 metapopulation of the species and the genes from Sarguja are immigrating to the central
338 region Dhamtari. For a better understanding of the influence of landscape and other
339 geographical features on distribution and gene diversity of *B. serrata* Roxb. an investigation
340 can be conducted with a larger number of locations representing its entire range of natural
341 distribution applying co-dominant multiallelic marker system to cross verify the results of our
342 investigation.

CONCLUSION

B. serrata Roxb. has been over-exploited due to its significant quality of wood for pulp and paper industries. The technology in paper industries to use the species as a raw material has become obsolete and nowadays the oleo-gum-resin obtained from its wood is of significant medicinal value to heal arthritis and asthma. The species is endemic to India and unfortunately, there is not a single source of sustainable supply for the raw material required in pharmaceutical industries. Apart from all, the species is understudied and has not received much scientific attention to report on population genetic. Through the present investigation, we report that the available genetic resource of the species in central Indian region maintains high variation in wood fiber dimension and exhibits higher genetic variability within and among its sub-populations. In order to maintain the valuable forest genetic resource, the species need prior attention for the conservation. The locations of its natural distribution can be developed as strict nature reserves. On the other hand, a genetic improvement program should be initiated for selection and multiplication of elite genotypes. There is need to develop protocols for the artificial generation of the species through clonal and tissue culture techniques.

ACKNOWLEDGEMENT

Authors are grateful to Director, Tropical Forest Research Institute, Jabalpur for Institutional support and to Director General, Indian Council of Forestry Research and Education, Dehradun for the financial support received under the project ID-175/TFRI/2011/Gen-4 (23).

REFERENCES

ABUDULI A, AYDIN Y, SAKIROGLU M. 2016. Molecular evaluation of genetic diversity in wild-type mastic tree (*Pistacia lentiscus* L.). *Biochemical Genetics* 54 (5): 619-635. doi:10.1007/s10528-016-9742-0

AMMON HP. 2006. Boswellic acid in chronic inflammatory disease. *Planta Medica* 72 (12): 1100-1116. doi:10.1055/s-2006-947227

ANSARI SA, NARAYANAN C, WALI SA, KUMAR R, SHUKLA N, KUMAR SR. 2012. ISSR markers for analysis of molecular diversity and genetic structure of Indian teak (*Tectona grandis* L f) locations. *Annals of Forest Research* 55 (1): 1-13.

BAPTISTA P, COSTA AP, SIMOES R, AMARAL ME. 2014. *Ailanthus altissima*: An alternative fiber source for papermaking. *Industrial Crops & Products* 52: 32-37. doi: 10.1016/j.indcrop.2013.10.008

BHAT KM, BHAT KV, DHAMODARAN. 1990. Wood density and fiber length of *Eucalyptus grandis* grown in Kerala, India. *Wood Fiber Science* 22 (1): 54-61.

BHAT RV, GUHA SRD, PANDA SP. 1952. Indigenous cellulosic raw materials for the production of pulp, paper and board, part VIII- writing and printing papers from *Boswellia serrata* Roxb. (Salai). *Indian Forester* 78 (4): 169 – 175.

BONIN A, EHRICH D, MANEL S. 2007. Statistical analysis of amplified fragment length polymorphism data: a toolbox for molecular ecologists and economists. *Molecular Ecology* 16: 3737–3758. Doi:10.1111/j.1365-294X.2007.03435.x

CI X, CHEN J, LI Q. 2008. AFLP and ISSR analysis reveals high genetic variation and inert-population differentiation in fragmented populations of the endangered *Litsea szemaois* (Lauraceae) from Southwest China. *Plant Systematics & Evolution* 273: 237-346. https://doi.org/10.1007/s00606-008-0012-4

387 DESHMUKH VP, THAKARE PV, CHAUDHARI US, GAWANDE PA. 2007. A simple
388 method for isolation of genomic DNA from fresh and dry leaves of *Terminalia arjuna*
389 (Roxb.) Wight and Argot. *Electronic Journal of Biotechnology* 10: 468-472.

390 DUMINIL J, DAINOU K, KAVIRIRI DK, GILLET P, LOO J, DOUCET JL, HARDY OJ.
391 2016. Relationships between population density, fine-scale genetic structure, mating
392 system and pollen dispersal in a timber tree from African rainforests. *Heredity* 166: 295-
393 303. doi: 10.1038/hdy.2015.101

394 EARL DA,VONHOLDT BM. 2012. STRUCTURE HARVESTER: a website and program
395 for visualizing STRUCTURE output and implementing the Evanno method. *Conservation*
396 *Genetics Resources* 4: 359-361.

397 EXCOFFIER L, LISCHER HEL. 2010. ARLEQUIN Suite ver 3.5: A new series of programs
398 to perform population genetics analyses under LINUX and WINDOWS. *Molecular*
399 *Ecology Notes* 10: 564–567. doi: 10.1111/j.1755-0998.2010.02847.x

400 GHORPADE RP, CHOPRA A, NIKAM TD. 2010. *In vitro* zygotic embryo germination and
401 propagation of an edangered Boswellia serrata Roxb., a source of boswellic acid.
402 *Physiology & Molecular Biotechnology of Plants* 16 (2): 159-165. doi: 10.1007/s12298-
403 010-0017-7

404 GRATIVOL C, DA FONSECA LIRA-MEDEIROS C, HEMERLY AS, FERREIRA PC.
405 2011. High efficiency and reliability of inter-simple sequence repeats (ISSR) markers for
406 evaluation of genetic diversity in Brazilian cultivated *Jatropha curcas* L. accessions.
407 *Molecular Biology Reporter* 38 (7): 4245-4256. doi: 10.1007/s11033-010-0547-7

408 HAMRICK JL, GODT MJW. 1989. Allozyme diversity in plant species. In: Brown AHD,
409 Clegg MT, Kahler AL, Weir BS. (ed), Plant Population Genetics: Breeding and Genetic
410 Resources. Sinauer, Sunderland; Massachusetts, USA, pp 43–63

411 HOLSINGER KE, LEWIS PO, DEY DK. 2002. A Bayesian approach to inferring population
412 structure from dominant markers. *Molecular Ecology* 11: 1157 – 1164.

413 HOLSINGER KE, LEWIS PO. 2007. Hickory: A package for analysis of population genetic
414 data, version 1.1. Computer program and documentation, Department of Ecology and
415 Evolutionary Biology, University of Connecticut, Storrs, Connecticut, USA

416 HOLSINGER KE, WALLACE LE. 2004. Bayesian approach for analysis of population
417 structure: an example from *Planthanthera leucophaea* (Orchidaceae). *Molecular Ecology*
418 13: 887–894.

419 KHAN MAW. 1972. Propagation of *Boswellia papyrifera* through branch-cuttings. *Indian*
420 *Forester* 98 (7): 437-440.

421 KHARE CP. 2004. Encyclopedia of India, Rational Western Therapy, Ayurvedic and other
422 Traditional Usage. Botany. Springer, Berlin

423 LIU K, MUSE SV. 2005. POWERMARKER: Integrated analysis environment for genetic
424 marker data. *Bioinformatics* 21 (9): 2128-2129. doi:10.1093/bioinformatics/bti282

425 LYNCH M, MILLIGAN BG. 1994. Analysis of population genetic structure with RAPD
426 markers. *Molecular Ecology* 3 (2) : 91-99. doi:10.1111/j.1365-294x.1994.tb00109.x.

427 MACPHERSON A, HOHENLOHE PA, NUISMER SL. 2015. Trait dimensionality explains
428 widespread variation in local adaptation. *Proceedings of the Royal Society of London B*
429 282: 20141570. http://dx.doi.org/10.1098/rspb.2 014.1570

430 MAHESH S, KUMAR P, ANSARI SA. 2015. A rapid and economical method for the
431 maceration of wood fibers in *Boswellia serrata* Roxb. *Tropical Plant Research* 2 (2):
432 108-111.

433 MIGNEAULT S, KOUBAA A, ERCHIQUI F, CHAALA A, ENGLUND K, KRAUSE C,
434 WOLCOTT M. 2008. Effect of fiber length on processing and properties of extruded wood-
435 fiber/HDPE composites. *Journal of Applied Polymer Science* 110 (2): 1085-1092.
436 https://doi.org/10.1002/app.28720

437 MILLER MP. 2005. Alleles in space: computer software for the joint analysis of inter-
438 individual spatial and genetic information. *Journal of Heredity* 96: 722–724. doi:
439 10.1093/jhered/esi119

440 NAIK PK, ALAM A, SINGH H, GOYAL V, PARIDA S. 2010. Assessment of genetic
441 diversity through RAPD, ISSR and AFLP markers in *Podophyllum hexandrum*: a
442 medicinal herb from the Northwestern Himalayan region. *Physiology & Molecular*
443 *Biology of Plants* 16:135–148. doi: 10.1007/s12298-010-0015-9.

444 NYBOM H, BARTISH IV. 2000. Effects of life history traits and sampling strategies on
445 genetic diversity estimates obtained with RAPD markers in plants. *Perspectives in Plant*
446 *Ecology Evolution & Systematics* 3: 93-114. *https://doi.org/10.1078/1433-8319-00006*

447 NYBOM HD. 2004. Comparison of different nuclear DNA markers for estimating
448 intraspecific genetic diversity in plants. *Molecular Ecology* 13: 1143-1155.
449 https://doi.org/10.1111/j.1365-294X.2004.02141.x

450 PAWAR GV, SINGH L, JHARIYA MK, SAHU KP. 2012. Regeneration status in relation to
451 antropogenic disturbance in tropical deciduous forest of Chhattisgarh. *Ecoscan* 1: 281-
452 286.

453 PEAKALL R, SMOUSE PE. 2012. GenAlEx 6.5: genetic analysis in Excel. Population
454 genetic software for teaching and research—an update. *Bioinformatics* 28: 2537-2539.
455 doi: 10.1093/bioinformatics/bts460

456 PERRIER X, JACQUEMOUD-COLLET JP. 2006. DARwin software http://darwin.cirad.fr/

457 PREVOST A, WILKINSON MJ. 1999. A new system of comparing PCR primers applied to
458 ISSR fingerprinting of potato cultivars. *Theoretical and Applied Genetics* 98: 107–112.
459 https://doi.org/10.1007/s001220051046

460 PRITCHARD JK, STEPHENS M, DONNELLY P. 2000. Inference of population structure
461 using multilocus genotype data. *Genetics* 155: 945–959.

462 PUROHIT SD, SHARMA P, NAGORI R. 1995. Assessment of molecular diversity and
463 *Invitro* clonal fidelity in some Indian medicinal plants. *ISHS Acta Horticulturae* 1098: 47-
464 60. doi: 10.17660/ActaHortic.2015.1098.4

465 SHAH SA, RATHOD IS, SUHAGIA BN, PANDYA SS, PARMAR VK. 2008. A simple
466 high performance liquid chromatography for the estimation of boswellic acids from the
467 market formulations containing *Boswellia serrata* extract. *Journal of Chromatography*
468 *Science* 46(8): 735-738.

469 SHARMA S. 1983. A census of rare and endemic flora of South East Rajasthan. In: Jain,
470 S.K. and Rao, R.R. (eds.), Threatened Plants of India. Botanical Survey of India ,
471 Howrah, pp. 630-670.

472 SHARMA SS, AADIL K, NEGI MS, TRIPATHI SB. 2014. Efficacy of two dominant
473 marker systems, ISSR and TE-AFLP for assessment of genetic diversity in biodiesel
474 species *Pongamia pinnata*. *Current Science* 106: 1576-1580.

475 SHARMA YK, BHANDARI KS, SRIVASTAVA A. 1985 *Boswellia serrata* resin gum: a
476 potential internal sizing agent for paper. *Indian Forester* 111(3): 149-157.

477 SLATKIN M. 1987. Gene flow and the geographic structure of populations. *Science* 236:
478 787–792.

479 SPIEGELHALTER DJ, BEST NG, CARLIN BP, VAN DER LINDE A 2002. Bayesian
480 measures of model complexity and fit. *Journal of Royal Statistical Society Series B*
481 *Statistical Methodology* 64: 583-689. https://doi.org/10.1111/1467-9868.00353

482 SUNNICHAN VG, MOHANRAM HY, SHIVANNA KR. 2005. Reproductive biology
483 of *Boswellia serrata*, the source of salai guggul, an important gum-resin. *Botanical*
484 *Journal of Linean Society* 147 (1): 73-82. https://doi.org/10.1111/j.1095-
485 8339.2005.00349.x

486 VAISHNAV V, ANSARI SA. 2018. Genetic differentiation and adaptation in teak (*Tectona*
487 *grandis* L.f.) metapopulations. *Plant Molecular Biology Reporter*. doi: 10.1007/s11105-
488 018-1101-3

489 VAISHNAW V, MOHAMMAD N, WALI SA, KUMAR R, TRIPATHI SB, NEGI MS,
490 ANSARI SA. 2015. AFLP markers for analysis of genetic diversity and structure of teak
491 (*Tectona grandis* L f) in India. *Canadian Journal of Forest Research* 44: 297-
492 306. https://doi.org/10.1139/cjfr-2014-0279

493 WANG BY, SHI L, RUAN ZY, DENG J. 2011. Genetic diversity and differentiation in
494 *Dalbergia sissoo* (Fabaceae) as revealed by RAPD. *Genetics & Molecular Research* 10
495 (1): 114-120.

496 WEEKS A, ZAPATA F, PELL SK, DALY DC, MITCHELL JD, FINE PVA. 2014. To move
497 or to evolve: contrasting patterns of intercontinental connectivity and climatic niche
498 evolution in "Terbinthaceae" (Anacardeacea and Burseraceae). *Frontier Genetics*. doi:
499 10.3389/fgene.2014.0 0409

500 WRIGHT S. 1969. Evolution and the Genetics of Populations, vol. 2: The Theory of Gene
501 Frequencies. University of Chicago.

502 YEH FC, YANG RC, BOYLE T. 1999. POPGENE 32-version 1.32. Software.
503 http://www.ualberta.ca/~fyeh/fyeh/.

504 ZIETKIEWIEZ E, RAFALSKI A, LABUDA D. 1994. Genome fingerprinting by simple
505 sequence repeat (SSR)-anchored polymerase chain reaction amplification. *Genomics* 20
506 :176–183.

507 ZOBEL B, JETT JB. 1995. Genetics of wood production. Springer-Verlag, New York, pp
508 128.