Assessment of Genetic Diversity and Population Genetic Structure of *Santalum album* L. in India by Genic and Genomic SSR markers

Tanzeem Fatima^{1*}, Ashutosh Srivastava¹, Vageeshbabu S. Hanur² M. Srinivasa Rao³

¹Genetics and Tree Improvement Division, Institute of Wood Science and Technology

Bangalore-India-560003

²Department of Biotechnology, Indian Institute of Horticultural Research

Hessarghatta, Bangalore-India- 560089

³Forest Development Corporation of Maharashtra Limited Nagpur-India-440036

*Corresponding Author: tanzeem.fatima@gmail.com

Abstract

Sandalwood (*Santalum album* L.) is highly valued aromatic tropical tree. It is known for its high quality heartwood and oil. In this study, 39 genic and genomic SSR markers were used to analyze the genetic diversity and population structure of 177 *S. album* accessions from 14 populations of three states in India. High genetic diversity was observed in terms of number of alleles 127, expected heterozygosity (He) ranged from 0.63-0.87 and the average PIC was 0.85. The selected population had relatively high genetic diversity with Shannon's information index (I) >1.0. 0.02 mean coefficient of genetic differentiation (F_{ST}) and 10.55 gene flow were observed. AMOVA revealed that 92% of the variation observed within individuals. Based on cluster and Structure result, individuals were not clustered as per their geographical origin. Furthermore the clusters were clearly distinguished by principal component analysis analysis and the result revealed that PC1 reflected the moderate contribution in genetic variation (6%) followed by PC2 (5.5%). From this study, high genetic diversity and genetic differentiation was found in *S. album* populations. The genetic diversity information of *S. album* populations can be used for selection of superior genotypes and germplasm conservation to promote the tree improvement of *S. album* populations.

Keywords: AMOVA, Genic and genomic SSRs, PCA, *S. album*, Structure.

1. Introduction

Indian Sandalwood (*Santalum album* L.; Family- Santalaceae) is an economically important tropical tree species (Kuriakose et al., 2010). It is naturally distributed in India, Indonesia, Australia, Sri Lanka, Fiji, pacific islands, Juan Fernandes islands, Niue and Tonga (Harbaugh and Baldwin, 2007; Subasinghe, 2013). India has

approximate area of 9000 sq. kms of *S. album* distribution spread across eight states; Karnataka having the maximum area of 5, 245 sq. kms, Tamil Nadu (3,040 sq. kms), Andhra Pradesh (200 sq. kms), Kerala (15 sq. kms), Odisha (25 sq. kms), Madhya Pradesh (33 sq. kms) and Maharashtra (33 sq. kms) (Kumaravelu et al., 2007). It is also distributed in other states, Assam, Rajasthan, Uttar Pradesh, Bihar and Manipur (Ananthapadmanabha and Gupta 2016). The scented heartwood and its essential oil served for religious purposes, sophisticated perfumes, flavors, cosmetics, toiletries, beauty aids and medicines (Sreenivasan et al., 1992). It has a great demand in traditional Ayurveda medicines for antidepressant, anti-inflammatory, anti-fungal, astringent, sedative, as an insecticide, antipyretic and antiseptic (Kirthikar and Basu, 1987; Heubeger et al., 2006; Burdock et al., 2008; Silva et al., 2018). Due to high rate of cross-breeding and the extensive range of agro-climatic conditions, *S. album* prevailed high rate of genetic diversity and complex genetic structure (Veerendra and Ananthapadmanabha, 1996). In earlier studies it was reported that a large gap in genetic distance of *S. album* populations, belongs to different geographical regions (Angadi et al., 2003). Various of factors including illicit felling, large-scale changes in land use and poor natural regeneration, superior genetic resources of *S. album* in the country are severely threatened and natural populations have been depleted considerably (Sreenivasan et al., 1992; Rao et al., 2007; Jain et al., 2003). Due to amendment to the Karnataka Forest Act 2001, many of the private organizations and people have started raising plantations. Under this scenario it is essential to assess the morphological and genetic diversity of *S. album* natural populations and plantations. Although the morphological parameter reflects the diversity they are mostly influenced by environmental factors. Therefore molecular markers are being extensively used to assess genetic diversity of the species. To date various dominant and co-dominant marker system has been used in genetic diversity and population structure study of *S. album* (RFLP, Isozymes, DAMD, Allozymes, and RAPD) markers (Brand 1994; Suma and Balasundaran, 2003; Shashidhara et al., 2003; Rao et al., 2007; Jones et al., 2008; Azeez et al., 2009; Dani et al., 2011). Microsatellite SSR markers genic and genomic have proven to be efficient and accurate tool for the study of genetic variation in forest tree species compared with other markers (Kalia et al., 2011). However, very few genomic and genic microsatellite markers have been developed for *S. album* (Naseer et al., 2012; Patel et al. 2016; Fatima et al., 2019a) to characterize genetic diversity and genetic differentiation of *S. album*. Therefore in this

study, we aimed to evaluate the genetic diversity and reveal the population structure of the natural populations and plantations of *S. album* by using both genomic and genic microsatellite markers. These results facilitate novel insight into the genetic diversity and population genetic structure in *S. album* with more appropriateness. Evaluation of genetic diversity could useful in tree improvement, tree breeding process and conservation management of *S. album* populations.

2. Material and Methods

The basic objective of the current study was to assess genetic diversity of $n = 177$ *S. album* accessions from natural populations and plantations of three southern states in India viz, Karnataka ($n=80$), Kerala ($n=30$) and Telangana ($n=67$) (Table 1). Genetic diversity analysis of *S. album* was performed by 39 (14 genic and 25 random genomic SSR markers). The DNA isolation, PCR and gel electrophoresis details were available (Fatima et al., 2019a & Fatima et al., 2019b).

2.1. Genetic Diversity Parameters, Genetic Differentiation, Cluster analysis

Genetic diversity parameters; number of alleles (N_a), effective number of allele (N_e), observed heterozygosity (H_o), expected heterozygosity (H_e), Shannon information index (I), Nei's genetic distance and Polymorphic Information Content (PIC) were characterized by GenAlex 6.0 software and Cervus 3.0.3 software (Peakall and Smouse, 2012; Kalinowski et al., 2007). To determine the genetic differentiation, analysis of molecular variance (AMOVA) and Wright's F statistics were performed by GenAlex v.6.5 (Peakall and Smouse, 2012). Phylogenetic cluster analysis based on the UPGMA was constructed by Jaccard's similarity coefficient by using Darwin software (version 6.0) (Takezaki et al., 2014) at 0.05 significance level.

2.2. Structure and Principal Component Analysis (PCA)

For analyzing population structure, a model-based (Bayesian) cluster analysis was performed based on SSR markers. An unbiased Bayesian approach using Markov chain Monte Carlo (MCMC) clustering of samples was conducted via the STRUCTURE v.2.3.4 software (<http://www.stats.ox.ac.uk/>) (Pritchard et al. 2000). PCA was executed for selected *S. album* accessions data by using Minitab v.18 software (McKenzie, 2004). It was used to identify the similarities between variables and classify the genotypes collections from different geographical locations (Torbick and Becker, 2009).

3. Results

3.1. Population genetic diversity and variation

103 In the current study, 127 alleles were amplified by 39 SSR loci examined in the 177 *S.*
104 *album* accessions and the average N_a detected 9.39 (Table 2). Among the selected
105 populations the largest N_a (12.13), N_e (9.12), H_e (0.87) was harbored in Telangana
106 population Surya L2 TS (Table 2). The highest I 2.24 and PIC (0.96) was observed in
107 Karnataka population IWC L1 KA respectively (Table 2). However, the Karnataka
108 populations Chen L4KA showed lower than average N_e (7.90) and H_e (0.83) (Table
109 2). The lowest PIC was found in Kerala population Mara L3 KL (Table 2). In terms of
110 PIC value, 10 populations (WC L1KA, NelL2KA, KFDHaL5KA, KFDDhL7 KA,
111 ShimL6KA, GoL KA, Surya L2 TS, Kham L3 TS and HydL4TS) had highly
112 informative alleles with values higher than 0.80 while only population (ChenL4KA,
113 MaraL1 KA, MaraL2 KL, Mara L3 KL) had moderately informative alleles with a
114 PIC value between 0.72-0.79 (Table 2). The average F , F_{ST} and N_m 0.80, 0.02 and
115 10.55 respectively (Table 2 and Table 4).

116 **3.2. Population genetic structure and genetic differentiation**

117 Both AMOVA and pairwise F_{ST} analysis were conducted to investigate the genetic
118 variation among 14 populations. Hierarchical AMOVA revealed that total 3% of the
119 total genetic variation occurred among 14 populations and 92% of the total variation
120 was distributed within the individuals (Table 3). Population genetic structure of *S.*
121 *album* revealed strong population genetic structure at the species level F_{ST} (0.08,
122 $P < 0.001$) (Table 3). AMOVA indicated that a small fraction of the observed genetic
123 diversity was attributable to differences among the populations (Table 3).
124 Cluster analysis based on the neighbour-joining (NJ) dendrogram method was
125 established to demonstrate the relationship between 177 accessions of *S. album* from
126 14 populations was clearly clustered into four clusters. Cluster I comprised three
127 natural populations from Karnataka and Kerala viz, Mara L3KL, Shim L6 KA, and
128 Che L4 KA. Cluster II consisted of 11 natural populations and plantations admixtures
129 viz, SuryaL1 TS, Surya L2 TS, Mara L1 KL, Mara L2 KL, Hyd L4 TS, KFD DhL7
130 KA, IWC L1 KA, Nel L2 KA, Kham L3 TS, KFD HaL5 KA, and GoL3 KA. In
131 cluster I natural populations and plantations both were grouped together whereas,
132 cluster II only plantations were segregated. Notably many neighboring populations
133 cluster together specifically all four populations of Telangana state cluster II (Fig. 1).
134 The pairwise population matrix between any two provenances, the index values
135 varied from 0.04-0.286. The smallest value 0.04 was found in between IWCL1 KA
136 and ChenL4 KA along with the highest distant differentiation (0.286) was observed in

137 Mara L2 KL and KFD DhL7 KA, which originated from two geographical distant
138 states in southern India (Table 5).

139 Structure analysis with the proportion of cluster members of each accession from 14
140 populations is shown in (Fig. 2). Individuals with a proportion higher than 0.75 were
141 considered pure and those with lower 0.75 were considered as admixtures. The graph
142 showed that the geographical distance was relatively not clear. The first cluster (red
143 lines) had a lower proportion than second cluster (green lines) (Fig. 2). Structure
144 analysis was conducted for highest value $K = 4$ which means that all populations were
145 divided into four genetic structure groups with admixtures. At $K=4$ the estimated Ln
146 probability of all data was <0.75 , inferred cluster value 28.85% (red lines) 30.45%
147 (green lines) 20.34% blue and 22.45% (yellow lines) revealed that genetic structure of
148 all the selected populations showed admixtures of mostly in Telangana and Karnataka
149 state accessions whereas, Kerala populations were found pure showed in yellow lines
150 (Fig. 2). State wise, structure analysis and cluster analysis of all accessions ($n=80$)
151 from Karnataka revealed four clusters ($\Delta K=4$). The lines were divided into two
152 groups I, II III and IV groups with admixtures. The mean assignment probabilities
153 were <0.7 **i.** (G1 18) (red) with the dendrogram individuals group I and **ii.** (G2 31)
154 (green) **iii.** (G3 16) (blue) and **iv.** (G4 15) (yellow) individuals were observed (Fig.
155 3). Similarly in Kerala populations revealed four major groups ($\Delta K=4$) and the
156 accessions were divided into four groups I, II, III and IV with the mean assignment
157 probabilities <0.8 **i.** (G1 11) (red) **ii.** (G2 9) (green) and **iii.** (G3 4) (blue) and **iv.** (G4
158 8) (yellow) (Fig. 4). Structure analysis of four populations of Telangana with ($n=65$)
159 accessions divided into three major groups at $\Delta K= 3$. The mean assignment
160 probabilities at <0.9 inferred cluster value **i.** (G1 36) (red) **ii.** (G2 19) (green) **iii.** (G3
161 10) (blue). Around 16 accessions from Kodada, Suryapet and IFB Hyderabad was
162 found admixtures (Fig. 5). This result was consistent with that of UPGMA clustering
163 analysis (Fig. 1, 2, 3, 4 and 5). Similarly the three clusters were clearly distinguished
164 by PCA analysis (Fig. 6); the populations from Karnataka, Kerala and Telangana
165 were mixed and scattered in all three clusters (Fig. 6). The localization of individuals
166 was defined by the first principal component (PC1) and second principal component
167 (PC2). The results revealed that PC1 reflected the highest contribution in genetic
168 variation (6%) followed by PC2 (5.5%). Eigen values of 25 PC components were
169 found higher than 1 and 15 PC components were found to be lower than 1 revealed

170 high variability in selected *S. album* accessions. The result showed the combined
171 variation 87.5% and cumulative variation 98.6% (Table 6). The obtained result
172 showed that there was no significant correlation was found between genetic distance
173 and geographic distance among the investigated populations.

174 4. Discussion

175 A number of causes have been endorsed to the depletion of *S. album* populations specifically
176 due to overharvesting and illegal poaching in natural habitats (Azeez et al., 2009; Naseer et
177 al., 2012). Therefore conservation and systematic cultivation in terms of identifying the elite
178 tree/populations of *S. album*, genetic diversity study is essential as it provides the important
179 factors that direct the variability (adaptation, migration), inbreeding and gene flow (Naseer et
180 al., 2012; Patel et al., 2016).

181 4.1. Genetic Diversity and variation

182 In the current study we estimated the genetic diversity parameters of 177 *S. alum* accessions
183 at 39 loci. Overall genetic diversity among the populations was high ($N_a=9.39$ and H_e 0.85)
184 (Table 2). Similar results were obtained in previous studies of *Santalum* and other tree species
185 like *O. lanceolata* (Otieno et al., 2016), *S. spicatum* (Millar et al., 2012), *S. lanceolatum*
186 (Jones et al., 2010), *S. leptocladum* (Jones et al., 2010), *S. album* (Fatima et al., 2019b) and
187 *Phyllostachys violascens*, (Cai et al., 2019). In contrast to our study, low N_a was observed in
188 *S. album* by using genic SSR markers (Fatima et al., 2019a) and, *Ambrosia artemisiifolia*,
189 (Meyer et al., 2017). The PIC value of 14 populations ranged from 0.83-0.96 (Table 2) in
190 which 10 populations were highly polymorphic ($PIC>0.8$). The average PIC (0.85) was higher
191 than that of previous study on *S. album*, *Jatropha curcas*, (Naseer et al., 2012; Wen et al.,
192 2014). The PIC value is changed by number of genetic markers, types of SSR markers, the
193 SSR motif repeats and the analysis method (Liu et al., 2017; Chen et al., 2020). As in another
194 study high level of polymorphism was observed in *S. album*, *Populus tomentosa*, *Quercus*
195 *variabilis* and *Eucalyptus tereticornis* respectively (Patel et al., 2016; Du et al., 2012; Shi et
196 al., 2017; Zhijiao et al., 2016).

197 4.2. Population structure

198 Investigation of population structure using several approaches supported the evolutionary
199 divergence of *S. album* populations in selected states. In this study we analyzed the population
200 structure of *S. album* at different levels. At the individual level, cluster, structure and PCA
201 analysis were performed on 177 accessions which divided them into groups. The Bayesian
202 structure analysis and PCA of *S. album* defined four groups and provided scattered genetic
203 divergence among the populations from the selected three states viz, Karnataka, Kerala and
204 Telangana (Fig. 2 and Fig. 6). The AMOVA revealed these four groups were not
205 differentiated among the populations (3% $P<0.001$) (Table 3). The structure result also
206 confirmed AMOVA test and showed presence of genetic diversity within each presumed
207 population groups. Similar results were reported in *S. album* and *Corylus mandshurica*

(Fatima et al., 2019a; Fatima et al., 2019b; Zong et al., 2015;). The obtained results indicated the mix lineage of the nuclear genome and distributed randomly among the populations. Within *S. album* populations distributed in southern India, no clear geographical genetic differentiation was found by the NJ tree, Bayesian and the PCA analysis (Fig. 1, Fig. 2 and Fig. 6). In contrast to present study genotypes of similar populations/geographical region were clustered together as per their geographical locations viz. *Tectona grandis*, *Cabralea*, *Canjerana* and *Swietenia macrophylla* (Alcantara et al., 2013; Sreekanth et al., 2012; Fofana, et al., 2009; Melo et al., 2014; Alcala et al., 2014; Lemes et al., 2002). UPGMA cluster analysis showed IWC L1 KA and Nel L2 KA populations were clustered together in cluster II and confirmed the seedling source of *S. album*. Similar result was observed in another study of *S. album* by oil specific genic markers (Fatima et al., 2019a). On the other hand a previous study using random genomic markers indicated mixed clustering of *S. album* populations (Fatima et al., 2019b). This might revealed the result based on the genic markers transferability within the species, high reproducibility and their conserved nature (Powell et al., 2007; Muriira et al., 2018).

5. Conclusion

In the present study, the genetic diversity and population structure of *S. album* genotypes were analyzed. The result revealed the high genetic variability and genetic differentiation in fourteen populations of *S. album* of three selected states of India. The observed variation in selected populations of *S. album* will aid in the selection of elite genotypes for genetic improvement, conservation management and tree breeding program. Additionally, these markers would be a useful tool for investigating genetic diversity, genetic structure of other states natural populations and plantations of *S. album*. The information obtained by both genomic and genic SSR markers offer precise mark on the distribution of the genetic diversity among the selected populations.

Acknowledgement

Authors are thankful to the Director, Group Co-ordinator Research, Head- Genetics and Tree Improvement Division, Dr. D. Annapurna GTI Division (PDF), Institute of Wood Science and Technology for encouragement to carry out the present study.

Funding

I gratefully acknowledge the funding agency, the University Grant Commission (UGC) of the Government of India, for providing financial support, to complete my work, in the form of UGC Maulana Azad National Fellowship JRF/SRF (F1.17.1/2015-16/MANF-2015-17-UTT-65893/(SA-III/website) 2015-16.

242 **Author contributions**

243 First author design the experiment, completed laboratory work and written manuscript. All
244 other authors reviewed the manuscript and helped in formatting.

245 **Availability of data and materials**

246 We declare that all data generated or analyzed during this study are included in this
247 manuscript.

248 **Declaration of Competing Interest**

249 We declare that this manuscript is original, has not been published and is not under
250 consideration for publications elsewhere and there are no conflicts of interest to disclose.

251

252 **References**

253 Alcalá, R.E., Salazar, H., Granados, G.G., Snook, L.K., 2014. Genetic structure and
254 genetic diversity of *Swietenia macrophylla* (Meliaceae): implications for sustainable
255 forest management in Mexico. *Journal of Tropical Forest Science* 26(1), 142-152.
256 <https://core.ac.uk/download/pdf/132697971.pdf>.

257

258 Alcantara, B.K., Veasey, E.A., 2013. Genetic diversity of teak (*Tectona grandis*
259 L.F.) from different provenances using microsatellite markers. *Ravista arvore* 37
260 (4), 747-758. [https://www.scielo.br/scielo.php?script=sci_arttext&pid=S0100-](https://www.scielo.br/scielo.php?script=sci_arttext&pid=S0100-67622013000400018&lng=en&tlng=en)
261 [67622013000400018&lng=en&tlng=en](https://www.scielo.br/scielo.php?script=sci_arttext&pid=S0100-67622013000400018&lng=en&tlng=en).

262 Angadi, V.G., Jain, S.H., Shankaranarayana, K.H., 2003. Genetic diversity between
263 sandal populations of different provenances in India. *Sandalwood Research*
264 *Newsletter* 17:4-5.

265 Ananthapadmanabha, H.S., Gupta, S.M. 2016. Queens of sandalwood. The Indian
266 sandalwood (*Santalum album* L.) Institute of Natural resources conservation
267 education research and training (INCERT) Bangalore.

268 ArunKumar, A.N., Joshi, G., Ram, M.H.Y., 2012. Sandalwood: history, uses,
269 present status and the future. *Current Science* 103 (12), 408-415.
270 <https://www.jstor.org/stable/24089347?seq=1>.

271

272 Azeez, S.A., Nelson, R., Prasadbabu, A., Rao, M.S., 2009. Genetic diversity of
273 *Santalum album* using random amplified polymorphic DNAs. *African Journal of*
274 *Biotechnology* 8 (13), 2943-2947.
275 <https://www.ajol.info/index.php/ajb/article/view/60956>.

276 Brand, J.E. (1994) Genotypic variation in *Santalum album*. *Sandalwood Research*
277 *Newsletter* 2, 24.

278 Cai, K., Zhu, L., Zhang, K., Li, L., Zhao, Z., 2019. Development and
279 Characterization of EST-SSR Markers From RNA-Seq Data in *Phyllostachys*
280 *violascens*. *Frontiers of Plant Sciences* 10, 50.

281 <https://www.frontiersin.org/articles/10.3389/fpls.2019.00050/full>.

282

- 283 Chen, S., Dong, M., Zhang, Y., Qi, S., Liu, X., Zhang, J., Zhao, J., 2020.
284 Development and Characterization of Simple Sequence Repeat Markers for, and
285 Genetic Diversity Analysis of *Liquidambar formosana*. Forests (11), 203.
286 <https://www.mdpi.com/1999-4907/11/2/203>.
- 287 Dani, S., Ravikumar, P., Kumar, P.R, Kush, A., 2011. Genetic variation within and
288 among small isolated populations of *Santalum album*. Biologia Plantarum 55 (2),
289 323-326. [https://bp.ueb.cas.cz/artkey/bpl-201102-0016_Genetic-variation-within-](https://bp.ueb.cas.cz/artkey/bpl-201102-0016_Genetic-variation-within-and-among-small-isolated-populations-of-Santalum-album.php)
290 [and-among-small-isolated-populations-of-Santalum-album.php](https://bp.ueb.cas.cz/artkey/bpl-201102-0016_Genetic-variation-within-and-among-small-isolated-populations-of-Santalum-album.php).
- 291 Du, Q., Wang, B., Wei, Z., Zhang, D., Li, B., 2012. Genetic diversity and
292 population structure of Chinese white poplar (*Populous tomentosa*) revealed by
293 SSR markers. Journal of Heredity 103 (6), 1-10.
294 <https://academic.oup.com/jhered/article/103/6/853/891661>.
295
- 296 Fatima, T., Srivastava, A., Somashekar, P.V., Vageeshbabu, H.S., Rao, M.S., Bisht,
297 S.S., 2019a. Assessment of morphological and genetic variability through genic
298 microsatellite markers for essential oil in Sandalwood (*Santalum album* L.). 3Biotech
299 (9), 252. <https://link.springer.com/article/10.1007/s13205-019-1758-9>.
- 300 Fatima, T., Srivastava, A., Vageeshbabu, H.S., Somashekar, P.V., Rao, M.S., 2019b.
301 Genetic Diversity Estimates of *Santalum album* L. through Microsatellite Markers:
302 Implications on Conservation. American Journal of Plant Science 10 (3), 462-485.
303 <https://www.scirp.org/journal/paperinformation.aspx?paperid=91297>.
- 304 Fofana, I.J., Ofori, D., Poitel, M., Verhaegen, D., 2009. Diversity and genetic
305 structure of teak (*Tectona grandis* L. f.) in its natural range using DNA
306 microsatellite markers. New Forests 37(2), 175-195.
307 <https://link.springer.com/article/10.1007/s11056-008-9116-5>.
- 308 Harbaugh, D.T., Baldwin, B.G., 2007. Phylogeny and biogeography of the
309 sandalwoods (*Santalum*, Santalaceae); repeated dispersals throughout the pacific.
310 American Journal of Botany 94 (6), 1028-1040.
311 <https://pubmed.ncbi.nlm.nih.gov/21636472/>.
- 312 Heubeger, E., Hongratanaworakit, T., Buchbauer, G. 2006. East Indian sandalwood
313 and α -santalol odor increase physiological and self-rated arousal in humans. Planta
314 Medica 72 (9), 792–800. <https://pubmed.ncbi.nlm.nih.gov/16783696/>.
315
- 316 Jain, S.H., Angadi, V.G., Shankarnarayana, K. H., 2003. Edaphic, Environmental
317 and Genetic Factors associated with Growth and adaptability of Sandal (*Santalum*
318 *album* L) in provenances. Sandalwood Research Newsletter (17), 6-7.
319 [http://www.ausbushfoods.com/bushfoodsonline/reports/Ess%20oils/Santalwood%2](http://www.ausbushfoods.com/bushfoodsonline/reports/Ess%20oils/Santalwood%20newsletter_2003.pdf)
320 [0newsletter_2003.pdf](http://www.ausbushfoods.com/bushfoodsonline/reports/Ess%20oils/Santalwood%20newsletter_2003.pdf).
- 321 Jones, C.G., Plummer, J.A., Barbour, E.L., Byrne, M., 2008. Genetic Diversity of
322 an Australian *S. album* Collection-Implications for Tree Improvement Potential.
323 *Silvae Genetica* 58 (1-6), 5-6. <https://sciendo.com/article/10.1515/sg-2009-0036>.

324
325 Jones, B.L., Waycott, M., Robson, H.L.A., Calladine, A., Page, T., 2010. Isolation
326 and characterization of microsatellite loci in *Santalum lanceolatum* and *Santalum*
327 *leptocladum* (Santalaceae). American Journal of Botany 97 (10), E97-e98.
328 <https://researchonline.jcu.edu.au/16887/>.
329
330 Kalia, R.K., Rai, M.K., Kalia, S., Singh, R., Dhawan, A.K., 2011. Microsatellite
331 markers: an overview of the recent progress in plants. Euphytica 177, 309–33.
332 [https://www.scielo.br/scielo.php?script=sci_arttext&pid=S1415-](https://www.scielo.br/scielo.php?script=sci_arttext&pid=S1415-47572016000300312)
333 [47572016000300312](https://www.scielo.br/scielo.php?script=sci_arttext&pid=S1415-47572016000300312).
334
335 Kalinowski, S.T., Taper, M. L., Marshall, T.C., 2007. Revising how the computer
336 program CERVUS accommodates genotyping error increases success in paternity
337 assignment. Molecular Ecology 16 (5), 1099–1106.
338 <https://pubmed.ncbi.nlm.nih.gov/17305863/>.

339 Kirthikar, K.R., Basu, B.D., 1987. Indian Medicinal Plants. International Book
340 Distributors, Dehradun. [https://www.worldcat.org/title/indian-medicinal-](https://www.worldcat.org/title/indian-medicinal-plants/oclc/366247781?referer=di&ht=edition)
341 [plants/oclc/366247781?referer=di&ht=edition](https://www.worldcat.org/title/indian-medicinal-plants/oclc/366247781?referer=di&ht=edition).

342 Kumaravelu, G., Bharathi, R.K., Jainaldeen, A., Balraj, V.C., 2007. Resurrection of
343 the rare aromatic *Santalum album* population in Tamil Nadu, India, Proceedings of
344 National Seminar, Bangalore. 23-31.
345
346 Kuriakose, S., Thankappan, X., Joe, H., Venkataraman, V., 2010. Detection and
347 quantification of adulteration in Sandalwood oil through near infrared spectroscopy.
348 Analyst 135 (10), 2676-2681. <https://pubmed.ncbi.nlm.nih.gov/20820490/>.

349 Lemes, M.R., Brondani, R.P.V., Grattapaglia, D., 2002. Multiplexed systems of
350 microsatellite markers for genetic analysis of Mahogany, *Swietenia macrophylla*
351 King (Meliaceae), a threatened neotropical timber species. The Journal of Heridity
352 93(4), 287-291. [https://agris.fao.org/agris-](https://agris.fao.org/agris-search/search.do?recordID=US201900239002)
353 [search/search.do?recordID=US201900239002](https://agris.fao.org/agris-search/search.do?recordID=US201900239002).

354 Liu, S., Liu, H., Wu, A., Hou, Y., An, Y., Wei, C., 2017. Construction of
355 fingerprinting for tea plant (*Camellia sinensis*) accessions using new genomic SSR
356 markers. Molecular Breeding (37), 93.
357 <https://link.springer.com/article/10.1007%2Fs11032-017-0692-y>.
358
359 McKenzie, J.D., 2004. Minitab 18th edition student release 14: statistical software
360 for education. Bostan, Pearson Addison- Wesley Harvard.
361
362 Melo, A.T.O., Coelho, A.S.G., Pereira, M.F., Blanco, A. J.V., Franceschinelli, E. V.,
363 2014. High genetic diversity and strong spatial genetic structure in *Cabralea*
364 *canjerana* (Vell.) Mart. (Meliaceae): Implications to Brazilian Atlantic Forest tree
365 conservation. Natureza & Conservação 12(2), 129-133.
366
367 Meyer. L., Causse, R., Pernin, F., Scalone, R., Geraldine, B., Chauvel, B., Deyle,
368 C., Corre, V. L. 2017. New gSSR and EST-SSR markers reveal high genetic

diversity in the invasive plant *Ambrosia artemisiifolia* L. and can be transferred to other invasive *Ambrosia* species. PLoS One 12 (5): e0176197. <https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0176197>.

Millar, M.A., Byrne, M., Barbour, E., 2012. Characterization of eleven polymorphic microsatellite DNA markers for Australian Sandalwood (*Santalum spicatum*) (R.Br.) A. D.C. (Santalaceae). Conservation Genetic Resources 4 (1), 51-53. <https://link.springer.com/article/10.1007%2Fs12686-011-9473-9>.

Muriira, N.G., Muchugi, A., Yul, A., Xu, J., Liu, A., 2018. Genetic Diversity Analysis Reveals Genetic Differentiation and Strong Population Structure in *Calotropis* Plants. Scientific Reports 8, 7832. [https://www.nature.com/articles/s41598-018-26275x#:~:text=Moderate%20genetic%20diversity%20was%20revealed,0.528\)%20between%20the%20two%20species](https://www.nature.com/articles/s41598-018-26275x#:~:text=Moderate%20genetic%20diversity%20was%20revealed,0.528)%20between%20the%20two%20species).

Naseer. M., Shashank, M., Kumar, P., Ansari, S.A., 2012. Genotyping of *Santalum album* L. accessions through cross-species transferability of SSR markers of *Santalum austrocaledonicum* and *Santalum insulare*. Sandalwood Research Newsletter 27, 1–5. https://www.researchgate.net/publication/267211931_Genotyping_of_Santalum_album_L_accessions_through_cross

Nei, M. 1973. Analysis of genetic diversity in subdivided populations. Proceedings of National Academy of Sciences of the United State of America (PNAS) 70 (12), 3321-3323. <https://www.pnas.org/content/70/12/3321>.

Otieno, O.J., Omondi, S.F., Perry, A., Odee, W.D., Makatiani, E.T., Kiplagat, O., Cavers, S., 2016. Development and characterization of microsatellite markers for *Osyris lanceolata* Hochst. & Steud., an endangered African sandalwood tree species. Tropical Plant Research (An International Journal) 3 (3), 701-703. <https://www.tropicalplantresearch.com/archives/?year=2016&vol=3&issue=3&ArticleId=176>.

Patel, D.M., Fougat, R.S., Sakure, A.A., Kumar, S., Kumar, M., Mistry, J.G., 2016. Detection of Genetic Variation in Sandalwood Using Various DNA Markers. 3 Biotech (6), 55. [https://link.springer.com/article/10.1007/s13205-016-03910-#:~:text=Out%20of%2025%2C%20181%20\(83.13,%25\)%20ISSR%20bands%20were%20polymorphic.&text=Though%2C%20there%20was%20high%20similarity,the%20characterization%20of%20sandalwood%20genotypes](https://link.springer.com/article/10.1007/s13205-016-03910-#:~:text=Out%20of%2025%2C%20181%20(83.13,%25)%20ISSR%20bands%20were%20polymorphic.&text=Though%2C%20there%20was%20high%20similarity,the%20characterization%20of%20sandalwood%20genotypes).

Paulpandi, M., Kannan, S., Kaveri, R.T.K., Gunasekaran, P., Rejeeth, C. 2012. *In vitro* anti-viral effect of β -santalol against influenza viral replication. Phytomedicine. 19 (3), 231-235. <https://pubmed.ncbi.nlm.nih.gov/22192867/#:~:text=%CE%B2%2DSantalol%20exhibited%20anti%2Dinfluenza,concentration%20of%20100%20%CE%BCg%2Fml>.

Peakall, R., Smouse, P. E., 2012. GENALEX 6.5: genetic analysis in excel. genetic analysis in excel. Population genetic software for teaching and research- an update.

413 Bioinformatics (Oxford, England) 28 (19), 2537-2539.
414 <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3463245/>

415 Pritchard, J.K., Stephens, M., Donnelly, P., 2000. Inference of population structure
416 using multilocus genotype data. Genetics 155 (2), 945-959.
417 <https://www.genetics.org/content/155/2/945>.

418 Rai, S.N., Sarma, C.R., 1990. Depleting sandalwood production and rising prices.
419 The Indian Forester. 116: 348-355.
420 <http://www.indianforester.co.in/index.php/indianforester/article/view/8972>.

421 Rao, N.M., Ravikanth, G., Ganeshaiah, K.N., Rathore, T.S., Shaanker, R.U., 2007.
422 assessing threats and identifying the ecological niche of sandal resources to identify
423 'hot-spots' for *in situ* conservation in Southern India. Sandal proceedings of
424 National Seminar. 23-31

425 Shashidhara, G., Hema, M.V., Koshy, B., Farooqi, A.A., 2003. Assessment of
426 genetic diversity and identification of core collection in sandalwood germplasm
427 using RAPDs. Journal of Horticulture Science Biotechnology. (78): 528-536.
428 <https://www.tandfonline.com/doi/abs/10.1080/14620316.2003.11511659>.

429 Shi, X., Wen, Q., Cao, M., Guo, X., Xu, L., 2017. Genetic Diversity and Structure
430 of Natural *Quercus variabilis* Population in China as Revealed by Microsatellites
431 Markers. Forests (8), 495. <https://www.mdpi.com/1999-4907/8/12/495>.

432 Sreenivasan, V.V., Sivaramakrishnan, V.R., Rangaswamy, C.R.,
433 Ananthapadmanabha, H.S., Shankaranarayana, K.H., 1992. Sandal (*Santalum album*
434 L.): a monograph, ICFRE, Dehradun 23-24.

435 Sreekanth, P.M., Balasundaran, M., Nazeem, P.A., Suma, T.B., 2012. Genetic
436 diversity of nine natural *Tectona grandis* L.f. populations of the Western Ghats in
437 Southern India. Conservation Genetics 13(5), 1409-1419.
438 <https://link.springer.com/article/10.1007/s10592-012-0383-5>.
439

440 Subasinghe, U. 2013. Sandalwood Research: A Global Perspective. Journal of
441 Tropical Forestry and Environment 3(01), 1-8.
442 https://www.researchgate.net/publication/264239816_Sandalwood_Research_A_Global_Perspective.
443

444 Suma, T.B., Balasundaran, M., 2003. Isozyme variation in five provenances of
445 *Santalum album* in India. Australian Journal of Botany (51), 243-249.
446 <https://www.publish.csiro.au/BT/BT02094>.
447

448 Silva, J.A.T.D, Kher, M.M., Soner, D., Natraj, M., Dobranszki, J., Millar, M., 2018.
449 Santalum molecular biology: molecular markers for genetic transformation.
450 Agroforestry System (92), 1301-1315.
451 <https://link.springer.com/article/10.1007/s10457-017-0075-8>
452

453 Takezaki, N., Nei, M., Tamura, K., 2010. POPTREE2: Software for constructing
454 population trees from allele frequency and computing other population statistics
455 with Windows interface. Molecular Biology and Evolution 27 (4), 747-752.

456

457 Torbick, N., Becker, B. 2009. Evaluating Principal Components Analysis for
458 Identifying Optimal Bands Using Wetland Hyperspectral Measurements from the
459 Great Lakes, USA. Remote Sensory (1), 408-417. [https://www.mdpi.com/2072-](https://www.mdpi.com/2072-4292/1/3/408)
460 4292/1/3/408.

461

462 Veerendra, H.C.S., Ananthapadmanabha, H.S., 1996. The breeding system in
463 Sandal (*Santalum album* L.). Silvae Genetica 45(4), 188-190.
464 <https://agris.fao.org/agris-search/search.do?recordID=DE1997T12189>.

465

466 Wen, M., Wang, H., Xia, Z., Zou, M., Lu, C., Wang, W., 2010. Development of
467 EST-SSR and genomic SSR markers to assess genetic diversity in *Jatropha curcas*
468 L. BMC Research Notes (3), 42.

469 <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2844395/>

470

471 Zhang, M., Mao, W., Zhang, G., Wu, F., 2014. Development and Characterization
472 of Polymorphic EST- SSR and Genomic SSR Markers for Tibetan Annual Wild
473 Barley. PloS One 9(4), e94881.
474 <https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0094881>.

475 Zhang, Y., Chen, X., Sun, W., Li, J., 2018. Genetic diversity and structure of tea
476 plant in *Qinba area* in China by three types of molecular markers. Hereditas 9
477 (155), 22. <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5944026/>.

478 Zhijiao, S., Heyu, Y., Qijie, W., Changpin, Z., Fagen, L., Mei, L., Wanhong, L.,
479 Jianzhong, L., Siming, G., 2016. Genetic diversity and selective loci in *Eucalyptus*
480 *tereticornis* populations. Scientia Silvae Sinicae 52(9), 40-47.
481 <http://www.linyekexue.net/EN/lexeme/showArticleByLexeme.do?articleID=7431>.

482 Zong, J.W., Zhang, T.T., Ma, Q.H., Liang, L.S., Wang, G.X., 2015. Assessment of
483 Genetic Diversity and Population Genetic Structure of *Corylus mandshurica* in
484 China using SSR Markers. PloS One 10(9), e0137528.
485 <https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0137528#:~:text=Corylus%20mandshurica%2C%20also%20known%20as,among%2012%20populations%20in%20China>.

488 Captions of the figures

489 **Fig. 1.** Neighbour joining tree based UPGMA dendrogram showing the clustering
490 pattern of 14 populations of *S. album* using 39 microsatellite markers
491 (*Red:Telangana population; Blue: Karnataka populations; Green: Kerala
492 populations).

493 **Fig. 2.** Result of Bayesian structure analysis for 177 *S. album* accessions based on
494 39 microsatellite markers (A) Delta K values from the mean log-likelihood
495 probabilities from STRUCTURE runs with inferred cluster (K) ranged from 1-11
496 (B) Distribution of cluster memberships at the individual levels estimation using
497 STRUCTURE software (Red area: Group I), (Green area: Group II), (Blue area:
498 Group III), (Yellow area: Group IV).

499 **Fig. 3.** Genetic distance and structure analysis of *S. album* genotypes from 7
500 populations of Karnataka using 39 microsatellite markers (ii) The dendrogram
501 constructed based on UPGMA analysis using similarity matrix (ii) The distribution
502 of clustering of Population structure of Karnataka at estimated LnP (D) of possible
503 clusters from (1-11) (K=4) Estimated genetic clustering (K=4) for 80 genotypes of
504 Karnataka.

505 **Fig. 4.** Genetic distance and structure analysis of *S. album* genotypes from 3
506 populations of Kerala using microsatellite markers (i) The dendrogram constructed
507 based on UPGMA analysis using similarity matrix (ii) The distribution of
508 clustering of Population structure of Karnataka at estimated LnP (D) of possible
509 clusters from (1-11) (K=4) Estimated genetic clustering (K=4) for 30 genotypes of
510 Kerala.

511 **Fig. 5.** Genetic distance and structure analysis of *S. album* genotypes from 4
512 populations of Telangana using microsatellite markers (i) The dendrogram
513 constructed based on UPGMA analysis using similarity matrix (ii) The distribution
514 of clustering of Population structure of Karnataka at estimated LnP (D) of possible
515 clusters from (1-11) (K=3) Estimated genetic clustering (K=4) for 64 genotypes of
516 Telangana

517 **Fig. 6.** Principal Component Analysis of 177 accessions from 14 populations of *S.*
518 *album* based on 39 microsatellite markers. The blue represents Karnataka
519 genotypes, the green filled dots represent Kerala and the red filled dots represent
520 Telangana state individuals. The first two Principle components explain percentage
521 of total variation for PC1 6.0% and PC2 5.5% respectively
522
523
524

Highlights:

- *Santalum album* L. is important aromatic tropical tree species.
- In the present study, 39 genic and genomic SSR markers were used to analyze genetic diversity and population structure of 177 *S. album* from 14 populations of three states.
- The result revealed the high genetic variability and genetic differentiation in fourteen populations of *S. album* from the selected states.
- The genetic diversity information of *S. album* populations could be used for selection of superior genotypes to promote the genetic tree improvement of *S. album*.

(A)

(B)

5
3

8
9
1

5

1

0 ————— 0.1

Table 1. Details of *S. album* natural populations and plantations identified across Karnataka, Kerala and Telangana state

State & total number of accessions	Provenances & number of accessions	Taluka	Ann Rainfall (mm)	Agroclimatic zone	Soil type	Soil pH
Karnataka (n=80)	IWC L1 KA (n=20)	Bangalore	698	Eastern Dry zone	Red loamy	6.3-8.3
	Nel L2 KA (n=10)	Bangalore	666	North West zone	Laterite	6.5-8.5
	GoL3 KA (n=10)	Hoskote	669	North East zone	Laterite	6.5-8.5
	Che L4 KA (n= 10)	Hassan	980	North West zone	Laterite	3.72-9.71
	KFD HaL5 KA (n=10)	Hassan	985	North West zone	Red and loamy	3.72-9.71
	Shim L6 KA (n=10)	Shimmoga	813	Southern transition zone	Red and loamy	6.8-8.25
Kerala (n=30)	KFD DhL7 KA (n=10)	Dharwad	899	North west dry zone	Lateritic brown sandy	1.18-6.31
	Mara L1 KL (n=12)	Idukki	698	Eastern Dry zone	Laterite	6.6-7.0
	Mara L2 KL (n=8)	Idukki	666	North West zone	Laterite	7-7.3
	Mara L3KL (n=10)	Devikulam	669	North East zone	Alluvial	6.6-7.9
Telangana (n=67)	SuryaL1 TS (n=23)	Kodada Town	1013	North West zone	Chalkas red sandy	8.3-9.1
	Surya L2 TS (n=23)	Kodada Town	1013	North West zone	Red sandy soil	8.9-9.6
	Kham L3 TS (n=10)	Khammam	899	North west dry zone	Deep red loamy	7.9-8.7
	Hyd L4 TS (n=11)	IFB Hyderabad	821	North east dry zone	Deep black	4.5-5.8

Table 2. Number of alleles (Na), effective number of alleles (Ne), observed heterozygosity (Ho), Expected heterozygosity (He), fixation index (F), Polymorphic Information Content (PIC) of fourteen populations of *S. album* by 39 genic and genomic SSR markers

State	Population	Na	Ne	Ho	He	F	I	PIC
Karnataka	IWC L	11.38	8.90	0.18	0.87	0.86	2.24	0.96
	1 KA							
	Nel L2 KA	8.51	7.36	0.19	0.85	0.79	2.02	0.83
	KFD HaL5 KA	8.00	6.83	0.15	0.83	0.89	1.95	0.80
	Che L4 KA	7.90	6.75	0.15	0.83	0.84	1.94	0.79
	KFD DhL7 KA	8.38	6.62	0.18	0.84	0.80	1.91	0.83
	Shim L6 KA	8.23	7.01	0.18	0.84	0.79	1.98	0.88
	GoL3 KA	8.45	7.24	0.19	0.84	0.76	1.99	0.87
Kerala	Mara L1 KL	9.23	7.65	0.18	0.65	0.79	2.07	0.74
	Mara L2 KL	7.84	6.60	0.13	0.63	0.88	1.93	0.74
	Mara L3 KL	8.27	7.18	0.15	0.70	0.84	1.99	0.72
	SuryaL1 TS	12.13	9.12	0.27	0.80	0.77	2.21	0.78
Telangana	Surya L2 TS	11.40	8.58	0.19	0.87	0.79	2.22	0.88
	Kham L3 TS	8.27	7.17	0.19	0.84	0.78	1.97	0.88
	Hyd L4 TS	8.50	7.42	0.22	0.85	0.74	2.02	0.84
	Overall pop	9.39	7.50	0.18	0.85	0.80	2.04	0.85

Table 3. Analysis of Molecular Variance (AMOVA) of 177 *S. album* accessions based on 39 genic and genomic SSR markers

Source of Variation	Df	SS	Variance component	Estimated variation	% of variation	F _{ST}	P value
Overall							
Among population	13	95.808	7.370	0.000	3%	0.08	0.001
Among individuls	162	1646.883	10.166	0.000	5%		
Within individulas	176	1930.000	10.966	10.966	92%		
Karnataka							
Among population	6	45.288	7.548	0.000	2%	-0.011	0.001
Among individuls	73	740.675	10.146	0.000	7%		
Within individulas	80	876.000	10.950	10.950	91%		
Kerala							
Among population	2	15.350	7.675	0.000	4%	-0.012	0.001
Among individuls	27	273.867	10.143	0.000	6%		
Within individulas	30	328.500	10.950	10.950	90%		
Telangana							
Among population	3	22.475	7.492	0.000	2%	-0.008	0.001
Among individuls	63	605.791	10.097	0.000	6%		
Within individulas	67	703.500	10.992	10.992	92%		

Table 4. F statistics and Nm (Gene flow) of 14 populations of *S. album* by 39 genic and genomic SSR markers; Nm = [(1/Fst)- 1]/4

States	Populations	F_{ST}	Nm
Karnataka	IWC L1 KA	0.03	9.50
	Nel L2 KA	0.03	8.18
	KFD DhL7 KA	0.02	12.95
	KFD HaL5 KA	0.020	10.75
	Che L4 KA	0.02	8.75
	GoL3 KA	0.02	9.57
	Shim L6 KA	0.02	8.06
Kerala	Mara L1 KL	0.02	13.47
	Mara L2 KL	0.02	8.66
	Mara L3 KL	0.02	8.66
Telangana	SuryaL1 TS	0.02	14.57
	Surya L2 TS	0.02	12.44
	Kham L3 TS	0.03	10.51
	Hyd L4 TS	0.02	11.62
Average		0.02	10.55

Table 5. Pairwise Population Matrix of Nie's (1973) unbiased genetic distance by using 39 genic and genomic SSR markers

Surya L1 TS	Surya L2 TS	Kham L3 TS	Hyd L4 TS	Mara L1 KL	Mara L2 KL	Mara L3 KL	IWC L1 KA	Nel L2 KA	KFD HaL5 KA	Che L4 KA	KFD DhL7 KA	Shim L6 KA	GoL3 KA	
***														SuryaL1 TS
0.08	***													SuryaL2 TS
0.09	0.01	***												Kham L3 TS
0.08	0.06	0.04	***											Hyd L4 TS
0.04	0.00	0.00	0.06	***										Mara L1 KL
0.14	0.12	0.08	0.00	0.14	***									Mara L2 KL
0.19	0.09	0.11	0.05	0.09	0.193	***								Mara L3 KL
0.09	0.05	0.01	0.00	0.00	0.098	0.11	***							IWC L1 KA
0.18	0.10	0.05	0.00	0.11	0.078	0.09	0.07	***						Nel L2 KA
0.12	0.01	0.00	0.03	0.07	0.123	0.16	0.00	0.19	***					KFD HaL5 KA
0.07	0.09	0.03	0.00	0.05	0.064	0.16	0.04	0.13	0.05	***				Che L4 KA
0.22	0.18	0.07	0.12	0.17	0.286	0.17	0.10	0.17	0.05	0.14	***			KFD DhL7 KA
0.14	0.05	0.09	0.03	0.03	0.153	0.13	0.02	0.04	0.09	0.15	0.22	***		Shim L6 KA
0.14	0.10	0.12	0.10	0.13	0.171	0.08	0.04	0.16	0.08	0.08	0.13	0.13	***	GoL3 KA

Table 6. Principal Component Analysis (PCA) of selected *S. album* accessions by using 39 genic SSR and genomic SSR markers

PC	Eigen value	Variation (%)	Cumulative (%)
PC1	3.81	6.02	6.03
PC2	3.52	5.53	11.67
PC3	3.36	5.36	16.83
PC4	3.23	5.18	22.07
PC5	2.94	4.63	26.62
PC6	2.73	4.41	30.99
PC7	2.75	4.38	35.22
PC8	2.56	4.02	39.29
PC9	2.43	3.88	43.11
PC10	2.31	3.71	46.97
PC11	2.12	3.45	50.23
PC12	2.17	3.37	53.57
PC13	2.05	3.22	56.83
PC14	2.01	3.25	59.98
PC15	1.95	3.01	63.03
PC16	1.79	2.77	65.79
PC17	1.69	2.64	68.32
PC18	1.63	2.59	70.98
PC19	1.53	2.41	73.32
PC20	1.45	2.35	75.68
PC21	1.41	2.27	77.82
PC22	1.34	2.13	79.96
PC23	1.27	2.02	81.93
PC24	1.15	1.81	83.75
PC25	1.08	1.74	85.48
PC26	0.93	1.55	86.86
PC27	0.91	1.47	88.31
PC28	0.88	1.48	89.62
PC29	0.83	1.33	90.93
PC30	0.77	1.21	92.24
PC31	0.76	1.24	93.37
PC32	0.70	1.15	94.49
PC33	0.62	1.07	95.53
PC34	0.57	0.95	96.32
PC35	0.57	0.93	97.23
PC36	0.50	0.87	98.06
PC37	0.45	0.72	98.71
PC38	0.34	0.57	99.35
PC39	0.28	0.48	99.73

Table 1. Details of *S. album* natural populations and plantations identified across Karnataka, Kerala and Telangana state

State & total number of accessions	Provenances & number of accessions	Taluka	Ann Rainfall (mm)	Agroclimatic zone	Soil type	Soil pH
Karnataka (n=80)	IWC L1 KA (n=20)	Bangalore	698	Eastern Dry zone	Red loamy	6.3-8.3
	Nel L2 KA (n=10)	Bangalore	666	North West zone	Laterite	6.5-8.5
	GoL3 KA (n=10)	Hoskote	669	North East zone	Laterite	6.5-8.5
	Che L4 KA (n= 10)	Hassan	980	North West zone	Laterite	3.72-9.71
	KFD HaL5 KA (n=10)	Hassan	985	North West zone	Red and loamy	3.72-9.71
	Shim L6 KA (n=10)	Shimmoga	813	Southern transition zone	Red and loamy	6.8-8.25
Kerala (n=30)	KFD DhL7 KA (n=10)	Dharwad	899	North west dry zone	Lateritic brown sandy	1.18-6.31
	Mara L1 KL (n=12)	Idukki	698	Eastern Dry zone	Laterite	6.6-7.0
	Mara L2 KL (n=8)	Idukki	666	North West zone	Laterite	7-7.3
	Mara L3KL (n=10)	Devikullam	669	North East zone	Alluvial	6.6-7.9
Telangana (n=67)	SuryaL1 TS (n=23)	Kodada Town	1013	North West zone	Chalkas red sandy	8.3-9.1
	Surya L2 TS (n=23)	Kodada Town	1013	North West zone	Red sandy soil	8.9-9.6
	Kham L3 TS (n=10)	Khammam	899	North west dry zone	Deep red loamy	7.9-8.7
	Hyd L4 TS (n=11)	IFB Hyderabad	821	North east dry zone	Deep black	4.5-5.8

Table 2. Number of alleles (Na), effective number of alleles (Ne), observed heterozygosity (Ho), Expected heterozygosity (He), fixation index (F), Polymorphic Information Content (PIC) of fourteen populations of *S. album* by 39 genic and genomic SSR markers

State	Population	Na	Ne	Ho	He	F	I	PIC
Karnataka	IWC L	11.38	8.90	0.18	0.87	0.86	2.24	0.96
	1 KA							
	Nel L2 KA	8.51	7.36	0.19	0.85	0.79	2.02	0.83
	KFD HaL5 KA	8.00	6.83	0.15	0.83	0.89	1.95	0.80
	Che L4 KA	7.90	6.75	0.15	0.83	0.84	1.94	0.79
	KFD DhL7 KA	8.38	6.62	0.18	0.84	0.80	1.91	0.83
	Shim L6 KA	8.23	7.01	0.18	0.84	0.79	1.98	0.88
	GoL3 KA	8.45	7.24	0.19	0.84	0.76	1.99	0.87
Kerala	Mara L1 KL	9.23	7.65	0.18	0.65	0.79	2.07	0.74
	Mara L2 KL	7.84	6.60	0.13	0.63	0.88	1.93	0.74
	Mara L3 KL	8.27	7.18	0.15	0.70	0.84	1.99	0.72
	SuryaL1 TS	12.13	9.12	0.27	0.80	0.77	2.21	0.78
Telangana	Surya L2 TS	11.40	8.58	0.19	0.87	0.79	2.22	0.88
	Kham L3 TS	8.27	7.17	0.19	0.84	0.78	1.97	0.88
	Hyd L4 TS	8.50	7.42	0.22	0.85	0.74	2.02	0.84
	Overall pop	9.39	7.50	0.18	0.85	0.80	2.04	0.85

Table 3. Analysis of Molecular Variance (AMOVA) of 177 *S. album* accessions based on 39 genic and genomic SSR markers

Source of Variation	Df	SS	Variance component	Estimated variation	% of variation	F _{ST}	P value
Overall							
Among population	13	95.808	7.370	0.000	3%	0.08	0.001
Among individuls	162	1646.883	10.166	0.000	5%		
Within individulas	176	1930.000	10.966	10.966	92%		
Karnataka							
Among population	6	45.288	7.548	0.000	2%	-0.011	0.001
Among individuls	73	740.675	10.146	0.000	7%		
Within individulas	80	876.000	10.950	10.950	91%		
Kerala							
Among population	2	15.350	7.675	0.000	4%	-0.012	0.001
Among individuls	27	273.867	10.143	0.000	6%		
Within individulas	30	328.500	10.950	10.950	90%		
Telangana							
Among population	3	22.475	7.492	0.000	2%	-0.008	0.001
Among individuls	63	605.791	10.097	0.000	6%		
Within individulas	67	703.500	10.992	10.992	92%		

Table 4. F statistics and Nm (Gene flow) of 14 populations of *S. album* by 39 genic and genomic SSR markers; Nm = [(1/Fst)- 1]/4

States	Populations	F_{ST}	Nm
Karnataka	IWC L1 KA	0.03	9.50
	Nel L2 KA	0.03	8.18
	KFD DhL7 KA	0.02	12.95
	KFD HaL5 KA	0.020	10.75
	Che L4 KA	0.02	8.75
	GoL3 KA	0.02	9.57
	Shim L6 KA	0.02	8.06
Kerala	Mara L1 KL	0.02	13.47
	Mara L2 KL	0.02	8.66
	Mara L3 KL	0.02	8.66
Telangana	SuryaL1 TS	0.02	14.57
	Surya L2 TS	0.02	12.44
	Kham L3 TS	0.03	10.51
	Hyd L4 TS	0.02	11.62
Average		0.02	10.55

Table 6. Principal Component Analysis (PCA) of selected *S. album* accessions by using 39 genic SSR and genomic SSR markers

PC	Eigen value	Variation (%)	Cumulative (%)
PC1	3.81	6.02	6.03
PC2	3.52	5.53	11.67
PC3	3.36	5.36	16.83
PC4	3.23	5.18	22.07
PC5	2.94	4.63	26.62
PC6	2.73	4.41	30.99
PC7	2.75	4.38	35.22
PC8	2.56	4.02	39.29
PC9	2.43	3.88	43.11
PC10	2.31	3.71	46.97
PC11	2.12	3.45	50.23
PC12	2.17	3.37	53.57
PC13	2.05	3.22	56.83
PC14	2.01	3.25	59.98
PC15	1.95	3.01	63.03
PC16	1.79	2.77	65.79
PC17	1.69	2.64	68.32
PC18	1.63	2.59	70.98
PC19	1.53	2.41	73.32
PC20	1.45	2.35	75.68
PC21	1.41	2.27	77.82
PC22	1.34	2.13	79.96
PC23	1.27	2.02	81.93
PC24	1.15	1.81	83.75
PC25	1.08	1.74	85.48
PC26	0.93	1.55	86.86
PC27	0.91	1.47	88.31
PC28	0.88	1.48	89.62
PC29	0.83	1.33	90.93
PC30	0.77	1.21	92.24
PC31	0.76	1.24	93.37
PC32	0.70	1.15	94.49
PC33	0.62	1.07	95.53
PC34	0.57	0.95	96.32
PC35	0.57	0.93	97.23
PC36	0.50	0.87	98.06
PC37	0.45	0.72	98.71
PC38	0.34	0.57	99.35
PC39	0.28	0.48	99.73