

A preliminary investigation on AFLP marker-wood density trait association in teak (*Tectona grandis* L. f.)

V. Vaishnav, S. A. Wali, S. B. Tripathi, M. S. Negi, S. A. Ansari

Vaishnav V., Wali S.A., Tripathi S.B., Negi M.S., Ansari S.A., 2018. A preliminary investigation on AFLP marker-wood density trait association in teak (*Tectona grandis* L. f.). Ann. For. Res. 61(1): _-_.

Abstract. Association between 276 AFLP loci and wood density of 46 teak (*Tectona grandis* L. f.) genotypes was evaluated, confirming the genetic structure among the genotypes and significant ($p < 0.01$) linkage disequilibrium between 9.4% loci-pair. AFLP markers with Bayesian correction for inbreeding coefficient detected a low genetic structure vis-à-vis high genetic diversity (0.23) and high polymorphism ($57.41 \pm 9.62\%$). AMOVA allocated 26.34% variation among the populations and 73.65% variation among the genotypes with $F_{ST} = 0.16$. The wood density with 8.71% variation displayed significant normal distribution. The careful control of statistical estimates incorporating Q and K to avoid the false discovery resulted in four AFLP loci significantly associated with the wood density trait. This is the first report dealing with marker-trait association in teak against the scarcity of background genomic information in this species. The AFLP markers associated with the wood density trait may be developed into STS markers for marker-assisted selection and breeding for genetic improvement of the species.

Keywords association mapping, Bayesian statistics, linkage disequilibrium, general linear model, mixed linear model

Authors. Vivek Vaishnav - Institute of Forest Productivity, Lalgotwa, NH-23, Ranchi 835303, India; Syed Arif Wali, Shashi Bhushan Tripathi, Mahender Singh Negi - The Energy Resource Institute, Darbari Seth Block, IHC Complex, Lodhi Road, New Delhi 110 003, India; Shamim Akhtar Ansari (shamimansari_1@yahoo.co.uk) - Flat #206, Gemini Apartment, Khurram Nagar, Lucknow 226022, India.

Manuscript received March 8, 2018; revised May 29, 2018; accepted June 10, 2018; online first July 3, 2018.

Introduction

Teak (*Tectona grandis* L. f.) is the foremost valuable tropical hardwood of the world,

valued for strength and natural resistance to termites. The timber finds a wide range of applications starting from making the most exquisite furniture through interior joinery

to cultural uses. It naturally occurs in India, Myanmar, Laos, and Thailand and later introduced to other tropical countries (Kaosa-ard 1981). With increased overexploitation of the species combined with agricultural expansion, shifting cultivation, population pressure and grazing, the natural teak forest area has substantially been declined. The situation necessitates conservation and multiplication of available genetic resource of the species together with identification of elite genotypes, applying marker-assisted selection (MAS). The MAS has been found to considerably reduce breeding cycles and increase selection accuracy (Rao 2004). It relies on the identification of DNA markers explaining the high proportion of the variation in the desired trait. For the approach, the linkage maps locating many quantitative trait loci (QTL) have been developed for forest species but with the limitation of low resolution to capture the trait variation and low success rate to validate the QTL at the different environment and pedigree lines (Mackay 2001). This limitation has led to change the research focus towards association mapping to identify the DNA marker associated with the trait that offers higher resolution and transferability of the marker to different family and species than QTL mapping (Wilcox et al. 2007, Khan & Korban 2012). The association mapping approach has extensively been used to detect loci related to quantitative traits in forest trees like microfibril angle (Thumma et al. 2005), early wood relative density and proportion of latewood (Gonzalez-Martinez et al. 2007), flowering traits (Ingvarsson et al. 2008), carbon isotope discrimination (Gonzalez-Martinez et al. 2008), cellulose content (Thumma et al. 2009), cold-hardiness-related traits (Eckert et al. 2009, Holliday et al. 2010), density (Dillon et al. 2010), growth cessation (Ma et al. 2010), disease resistance (Quesada et al. 2010), composition of early wood, tracheid wall thickness and annual ring width (Beaulieu et al. 2011), and growth traits such as height and girth (Lepoittevin et al. 2012).

The association mapping in different spe-

cies has led to the advancement for robust experimental design to eliminate confounding factors causing false discovery through type-I or type-II error (Jansen 1994). The association mapping conceptually captures the history of 'non-random association of alleles at different loci' or linkage disequilibrium (LD) significantly correlated to the phenotypic polymorphism that is mainly influenced by the genetic structure and mating behavior of the species (Slatkin 2008). A small population is prone to lose the rare variants leading to high LD vis-à-vis false association discovery (McCarthy et al. 2008). Therefore, the reliable association mapping prefers a large random mating population with highly admixed genotypes that extends the LD due to the introduction of different ancestry in species genome, on the one hand, and facilitates increased equal distribution of allele frequency, on the other (Falush et al. 2003). In case of a petite population comprising *ex situ* conserved germplasm bank or a small set of genotypes, one should cautiously confirm the true representation of the available genetic resources of the species in the wild. The higher LD decay is observed in out-crossing species than the self-crossing species. On the other hand, the stratified population with strong genetic structure, and with unequal distribution of alleles and limited gene flow, may lead to the spurious marker-trait association based on significant LD. Due to such troublesomeness, it has been essential to confirm the extent of LD among the applied markers and the genetic structure of the sampled set of genotypes (Myles et al. 2009). The advanced statistical modeling has made a provision to incorporate the information of genetic structure and genetic relatedness of the sampled genotypes to avoid the false-discovery due to the structured population (Pritchard et al. 2000, Evanno et al. 2005, Falush et al. 2007, Hubisz et al. 2009). There are two approaches for association mapping based on the species genome coverage through the selected DNA marker; first, genome-wide association study (GWAS) and the second, candidate gene-based association

study (CGAS). The GWAS does not require genomic information related to the targeted trait and hence suits to the random markers. The genetic differentiation and structure of teak distributed in India (Ansari et al. 2012, Vaishnav et al. 2015) and other native countries (Verhaegan et al. 2010, Hansen et al. 2015) has been extensively reported. However, no investigation has ever been performed to understand the genetic architecture of its wood quality or to develop the genetic probe for its marker-assisted selection. The highly heritable wood density of teak (Moya et al. 2013) is expected to have a direct influence on its various mechanical properties (Kokutse et al. 2004). Incidentally, the insignificant association between wood density and recorded growth descriptors i.e. height or girth has eluded to develop morpho-metric marker. Nevertheless for the association mapping and MAS, the wood density of teak is an implicit and suitable trait because

of its high commercial value and heritability. Therefore, we sampled a teak germplasm bank to find out the association between the genetic polymorphism exhibited on the amplified fragment length polymorphism (AFLP) system and wood density trait to target its causal QTL.

Materials and methods

Plant materials

The plant material for the present investigation comprised clones of teak plus trees selected from natural populations distributed to the various states of India and assembled in the national teak germplasm bank (NTGB), Chandrapur, India (Table 1) since 1978 (Kumar et al. 1998). The wood radial core samples were non-destructively extracted at the breast height (1.37 m) of the trees using an increment borer

Table 1 Estimates of genetic diversity parameters in 46 genotypes of teak (*Tectona grandis* L.f.) sampled from national teak germplasm bank, Chandrapur, India obtained by AFLP markers

States (acronyms)	N	Accessions	P% ^a	H _s ^b	H _e ^c
Andhrapradesh (AP)	10	APNPL-8, APJNB-1, APKEA-23, APKEA-24, APKEL-2, APT-7, APT-11, APT-22, APT-60, SBL-1	69.90	0.27	0.21
Arunanchal Pradesh (AR)	2	AC-I, AC-II	56.50	0.27	0.27
Gujrat (GJ)	3	G+4, G+10, G+13	51.40	0.26	0.20
Karnataka (KR)	7	MYHD-3, MYHD-4, MYSA-1, ST-8, ST-17, ST-19, ST-20	62.30	0.25	0.17
Kerala (KE)	1	KLS-2	42.80	0.27	0.15
MadhyaPradesh (MP)	1	PT-41	46.00	0.27	0.16
Maharashtra (MH)	5	MHALA-7, MHAL-8, MHALP-9, MHALA-4, MHALA-9	58.70	0.26	0.19
Orissa (OR)	5	ORPUB-18, ORPUB-27, ORJEK-1, ORANR-7, ORP-8	66.70	0.28	0.24
Tamilnadu (TN)	5	TNT-5, TNT-6, TNT-7, TNT-15, TNT-17	61.60	0.27	0.21
Uttar Pradesh (UP)	5	UP-A, UP-C, UP-E, UP-H, UP-O	69.20	0.28	0.26
West Bengal (WB)	2	WB, WB+4	46.40	0.28	0.15
Total	46	Average (\pm std. deviation)	57.41 ± 9.62	0.27 ± 0.01	0.20 ± 0.04

Note. Abbreviations: ^a - percentage of polymorphism, ^b - panmictic heterozygosity resulted by HICKORY, ^c - heterozygosity resulted by AFLP-SERV.

and delivered to the laboratory for the wood density measurement. The collected branch cuttings of teak plus tree clones were treated with IAA (5 mM) at the root end for 24h and sealed at shoot end through the paraffin wax and planted in the nursery. The sprouted young leaves were plucked and preserved in a cryo freezer for genomic DNA isolation.

Wood Density Measurement

The density of wood samples was measured applying geometrical volume based formula of density (oven-dried weight/green volume). The green weight of each sample was measured following previously described method (Chave 2005) using weighing balance (0.01 mg resolution) just after bringing it from field to the laboratory. Oven-dry weight of the same sample was measured by drying it in an oven (at 75 °C) until it achieved constant weight (this took three weeks). The samples were weighed immediately after being taken out of the drying oven. The green volume of samples was measured assuming regular cylindrical shape. Total length and diameter at middle and both ends of the sample was measured using a digital caliper. The green volume of samples was measured using formula $(\pi*d^2*l)/4$ (Chave 2005), where, 'd' is the mean diameter and 'l' is the length of the wood radial core sample.

DNA Isolation, restriction digestion and amplification

The DNA was isolated from the leaf sam-

ples following a modified CTAB method (Narayanan et al. 2007). After quality and quantity estimation described in the method (Narayanan et al. 2007), a total of 46 genomic DNA samples were selected for final amplification. The extracted genomic DNA was subjected to AFLP analysis following the procedure of Vos et al. (1995). We used AFLP kit I (Invitrogen Inc.) along with its instructions. 200-250ng genomic DNA was restricted for 2h at 37^o C with 2.5 units of *Mse* I and 5 unit of *Eco* RI in a 25 µl reaction mixture that additionally contained 5 µl of 5X restriction ligation buffer enriched with ATP, 2.5 µl of 0.5 M NaCl, 1.25 µl of BSA (1 mg/ml). The reaction was terminated by enzyme inactivation at 70 °C for 10 min. Subsequently, 1 unit of T4 DNA ligase along with 5 pmole of *Eco* RI adaptor, 50 pmole of *Mse* I adaptor was added to the digestion reaction to perform ligation for 2h at 20 °C. The ligation mix obtained so was diluted 1:10 ratio with TE buffer (10 mM Tris, 0.1 mM EDTA). A 2 µl of the diluted ligation mix was taken as a template for pre-amplification with adaptor specific primers *Eco* RI + A and *Mse* I + C in a total of 20 µl volume (Table 2). Pre-amplification was performed on Gene Amp PCR 9700 thermal cycler that was programmed as: 20 cycles of 30 s at 94 °C, 60 s at 56 °C and 60 s at 72 °C. The resultant mix was diluted to 50 fold, of which 10 µl volume was taken for selective amplification, employing *Eco* RI (labeled with ³²P) and *Mse* I primers with 3 selective nucleotides in four combinations (Table 3) and subjected to PCR conditions as: one cycle of the 30s at 94 °C, 30s at

Table 2 Sequence information of AFLP primers applied for the assay *Tectona grandis* L.f. and *Tectona hamiltoniana* Wall.

EcoRI		MseI	
Primers	Sequences (5'-3')	Primers	Sequences (5'-3')
E-A	GACTGCGTACCAATTC-A	M-C	GATGAGTCCTGAGTAA-C
E-AAG	GACTGCGTACCAATTC-AAG	M-CTC	TGAGTCCTGAGTAA-CTC
E-ACT	GACTGCGTACCAATTC-ACT	M-CAT	TGAGTCCTGAGTAA-CAT
E-AAC	GACTGCGTACCAATTC-AAC	M-CAA	TGAGTCCTGAGTAA-CAA
		M-CAC	TGAGTCCTGAGTAA-CAC

Table 3 Performance of AFLP primer combinations in 46 genotypes of teak (*Tectona grandis* L.f.)

Primers	AF ^a	PIC ^b	RP ^c
E _{AAG} -M _{CTC}	0.86	13.93	33.34
E _{AAG} -M _{CAC}	0.80	10.48	38.14
E _{AAC} -M _{CAT}	0.83	19.97	43.25
E _{ACT} -M _{CAA}	0.83	14.13	35.60
Average	0.83 ± 0.02	14.62 ± 3.93	37.58 ± 4.25

Note. Abbreviations: ^a - average allele frequency, ^b - polymorphic information content, ^c - resolving power.

65 °C and 72 °C for 60s, decreasing annealing temperature by 1 °C per cycle during the first 11 cycles, followed by 23 cycles each at 94 °C for the 30s, 56 °C for 30s and 72 °C for 60s. The amplified product was resolved on 6% polyacrylamide gel followed by autoradiography for fragment detection. The AFLP assay was repeated thrice. Only reproducible well resolved bands ranging 80-400 bp across teak genotypes were enumerated for binary matrix.

Data analysis

The measures of central tendency and the coefficient of variation were calculated for height, girth at the breast height (GBH), and the wood density and their normal distribution was also confirmed using program PAST v2.10 (Hammer et al. 2010). A binary genetic profile of 46 genotypes on amplified AFLP dominant markers was generated and analyzed following the both band- and allele frequency-based approaches as suggested in Bonin et al. (2007). Only polymorphic loci were used in all the estimates. The amplification performance of the markers was evaluated through the generation of the mean allelic frequency (AF) of the primers on sampled genotypes through program AFLP-SURV (Vekemans 2002). The polymorphic information content (PIC) of the primers (Botstein et al. 1980) was calculated applying program POWERMARKER v3.1 (Liu & Muse 2005). The resolving power (RP) of the primers was calculated following the formula given by Prevost & Wilkinson (1999). The level of LD between markers and distance (in

base pair or 'bp') of LD decay were estimated using the pair-wise recombination coefficient (R^2 values) calculated by software TASSEL v3.0 (Bradbury et al. 2007). The percentage of the marker combinations at significant ($p < 0.01$) LD was calculated by estimating the number of marker combinations at different R^2 values (>0.2).

The genetic diversity of the germplasm collection was estimated applying Bayesian statistics appropriate for dominant markers and 46 sampled genotypes. Consequently, we applied DIC based models implemented in program HICKORY v1.1 (Holsinger & Lewis 2007) for detecting Hardy-Weinberg Equilibrium (HWE) and total panmictic heterozygosity (Ht) of the germplasm as well as panmictic heterozygosity (Hs) of the sampled states. The program was run with 50,000 steps of burn-in, 500,000 replicates, 'thinning' = 20 following the best-suited model for the data amongst the all four models (full, $f = 0$, $\theta = 0$ and f-free models) on the basis of the least DIC (deviance information criterion) value as suggested by Spiegelhalter et al. (2002). For the genetic differentiation estimates, we preferred the θ -statistics (*viz.* θ -I, θ -II, θ -III and G_{ST} -B) implemented in the program that allows direct F_{ST} estimates from dominant markers without the assumption of prior knowledge of the extent of inbreeding. Incorporating the prior information of inbreeding (F_{IS}) obtained from the model in HICKORY, we estimated polymorphism percentage ($P\%$), Nei's gene diversity or expected heterozygosity (He), among and within-population genetic diversity through

Bayesian approach based program AFLP-SURV with non-uniform prior distribution of allele frequencies at 10,000 permutations (Zhivotovsky 1999). The hierarchical genetic structure of the germplasm bank was determined through among and within population heterozygosity by the program AFLP-SURV and through analysis of molecular variance (AMOVA) by program ARLEQUIN v3.11 (Excoffier & Lischer 2010).

For the confirmation of the distinct genetic resource of *Tectona grandis* L. f. (Indian teak) amplified by the dominant markers, we included two genotypes of *Tectona hamiltoniana* Wall. (Myanmar teak) in our analysis as an outlier. *T. hamiltoniana* Wall is endemic to Myanmar and the number of private and common alleles for the accession was evaluated through program GENALEX v6.0 (Peakall & Smouse 2012). Jaccard genetic distance coefficients with 100 bootstrap replications was used to generate a rooted dendrogram for phylogenetic relationship among the genotypes, employing program DARWIN v5.0 (Perrier & Jacquemoud-Collet 2006). The principal coordinate analysis (PCoA) was performed to cluster the genotypes in different axes applying program GENALEX v6.0 (Peakall & Smouse 2012). The Bayesian algorithm based program STRUCTURE v2.3.1 (Pritchard et al. 2000) was applied to determine the appropriate number of cryptic populations (K) assigned by the genotypes. The program was run applying admixture model with correlated allele frequencies among the samples on 100,000 burn-in and 1000,000 MCMC repeats with 10 runs for each number of K ($2 \leq K \leq 12$). The best-suited 'K' was determined based on the highest delta-K value extracted from online program STRUCTURE HARVESTER (Evanno et al. 2005).

We performed the marker-trait association analysis through software TASSEL v3.0 (Bradbury et al. 2007) applying general linear model (GLM) and mixed linear model (MLM) as suggested by Yu et al. (2006) applying 10,000 simulations. The GLM is a least square

fixed linear model to find out an association between segregating markers and trait that disregards the population structure of the sampled genotypes. The MLM considers the markers applied to the study as a fixed-effect factor and the population structure information of the sampled genotypes are considered as random effect factors. The required K was calculated through the software TASSEL v3.0 and Q-value was obtained from software STRUCTURE following the most suitable number of K to implement the MLM+K+Q method. The significant marker-trait association was determined based on marker at $p \leq 0.01$.

Results

The average GBH of the sampled 46 teak genotypes was 80.94 cm (CV = 37.25%), the average height, 18.20 m (CV = 42.71%) and the average wood density, 781.52 kg/m³ (CV = 8.71%). The wood density was found in normal distribution at significant level $p < 0.01$. A significant correlation ($p < 0.001$) between the GBH and height existed.

With 63.04% polymorphic loci, the five AFLP primer combinations amplified 276 loci. An average of 121.7 loci was amplified for each teak genotype. The average allele frequency was 0.83 ± 0.02 . PIC and RP values were obtained to be the highest in the primer E-AAC/M-CAT (Table 3). Among the 8525 loci-pairs, 886 loci (10.39%) were found to be in significant LD ($p < 0.01$). Overall 9.04% loci-pairs were found with $R^2 > 0.2$ in distance up to 80 bp (Figure 1).

The sampled teak genotypes representing 11 states exhibited 57.41 ± 9.62 % of polymorphism (Table 1). The Bayesian algorithm based program resulted in the lowest DIC value (3785.99) for full-model, supporting its suitability for the data and its deviation from HWE. For the sampled collection of the germplasm bank, the average panmictic heterozygosity (Hs) was found to be 0.27 (Table 1) and the total panmictic heterozygosity (Ht),

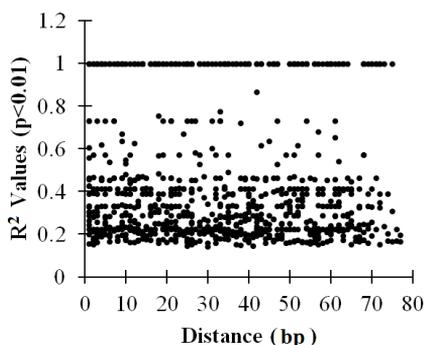


Figure 1 An LD plot with distribution of loci found in significant LD ($p < 0.01$) at different distance (in bp) for different level of recombination coefficient (R^2)

0.29. The measures of the θ -statistics were found to be 0.36 for θ -I, 0.09 for θ -II, and 0.08 for θ -III. The G_{ST-B} was 0.085. With the confirmation of deviation from HWE, the f -value (0.43) obtained from program HICKORY was incorporated as F_{IS} value in program AFLP-SURV that produced the corrected estimate of average heterozygosity (H_e) to be 0.20 ± 0.04 (Table 1) and the H_t to be 0.23. No significant correlation was obtained between the H_s and H_e . The estimates for H_e among and within populations (states) were 0.04 and 0.19, respectively. The F_{ST} was 0.16. AMOVA allocated 26.34% genetic variation to the among-population and 73.65% variation to within-population.

The genotypes of *T. hemiltoniana* Wall. included as outlier exhibited 52 (18.84%) private and 19 (6.88%) common loci, amplifying only 38.76% of 276 loci. In consensus rooted dendrograph, *T. grandis* L. f. genotypes, except G+4 genotype were distinctly and widely separated from *T. hamiltoniana* Wall. genotypes, supported by an absolute bootstrap value, i.e. 100%. *T. grandis* L. f. genotype G+4 was included in *T. hamiltoniana* Wall. cluster, but belonged to a different sub-cluster within the cluster, supported by an absolute bootstrap value, i.e.100%. Fur-

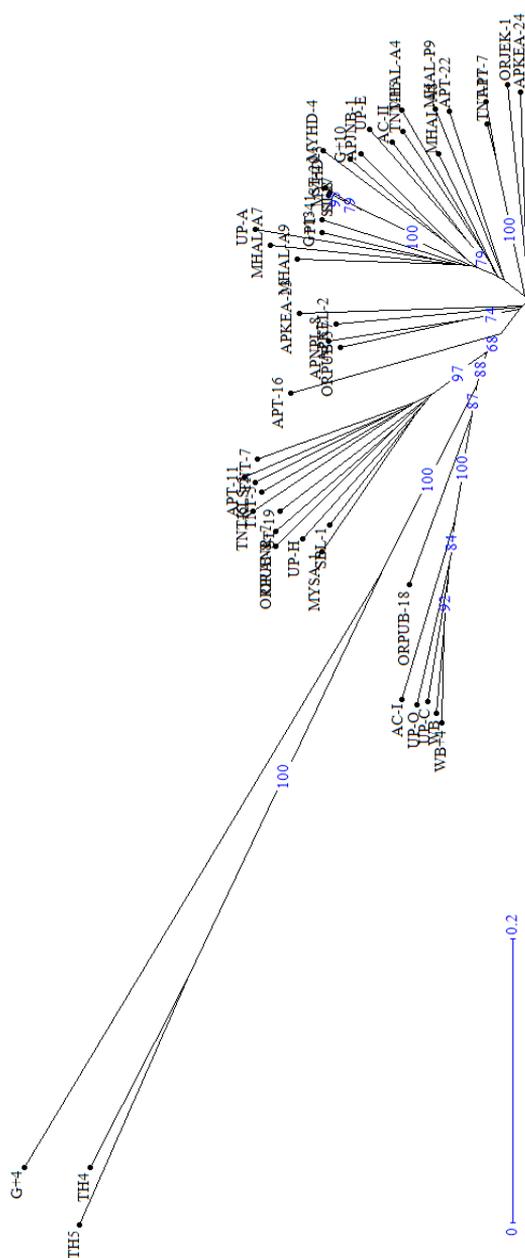


Figure 2 Jaccard genetic distance coefficients based rooted phylogenetic (radial) tree for 46 *Tec-tona grandis* L. f. genotypes along with two genotypes of *T. hamiltoniana* Wall. (TH). The bootstrap values $>60\%$ are shown on the nodes.

ther, *T. grandis* L. f. genotypes made several hierarchical stable clusters (supported by high bootstrap values, i.e. > 60%) that intermixed

genotypes from different locations, i.e. states (Figure 2). The PCoA representing 96.25% variation also separated both genotypes of *T.*

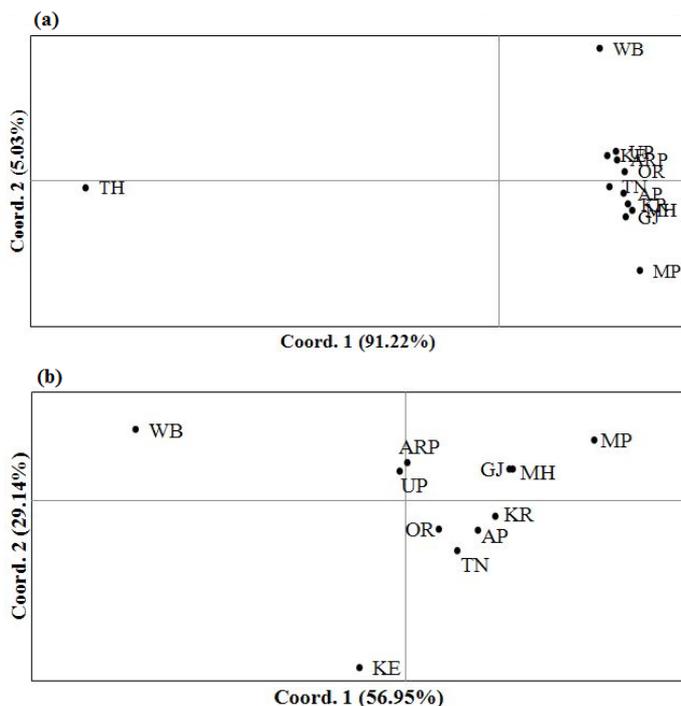


Figure 3 Principal coordinate analysis among sampled states of *T. grandis* L. f. genotypes: (a) Including genotypes of *T. hamiltoniana* Wall., (b) Excluding *T. hamiltoniana* Wall. The name of states for the acronyms are given in Table 1.

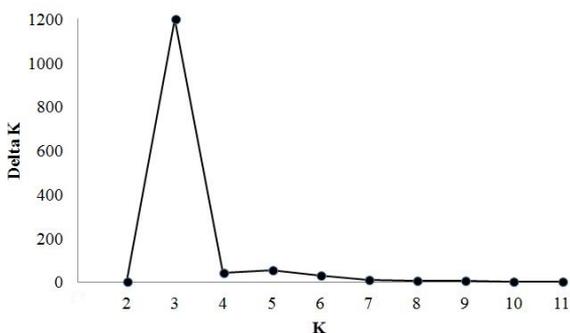


Figure 4 The delta K values for cryptic populations detected among 48 genotypes by program STRUCTURE HARVESTER: The best value for K = 3 cryptic populations.

hamiltoniana Wall. from *T. grandis* L. f. genotypes (Figure 3a). On exclusion of *T. hamiltoniana* Wall. genotypes, the PCoA representing 86.09% variation among the genotypes of *T. grandis* L. f. showed the KE, MP, and WB distinctly and far from other three groups of ARP-UP, GJ-MH, and AP-OR-KR-TN (Figure 3b). STRUCTURE analysis implementing admixture with correlated allele frequencies identified three cryptic populations to which the sampled genotypes subscribed their memberships (Figure 4). *T. grandis* genotypes exhibited admixed subscription to two cryptic populations. However, *T. grandis* genotype G+4 subscribed to all three detected cryptic populations. In contrast, *T. hamiltoniana* genotypes did not subscribe to two

cryptic populations to which *T. grandis* genotypes belonged; but they subscribed exclusively to the third cryptic population only (Figure 5). In both GLM and MLM models, three AFLP loci amplified by the primer E-AAC/M-CAT were found to be significantly ($p < 0.01$) associated with wood density (Table 4). An AFLP locus E-AAG/M-CTC_41 was found to be significantly ($p < 0.01$) associated with the trait in MLM model only and was found in LD block with maximum 19 loci (Table 4).

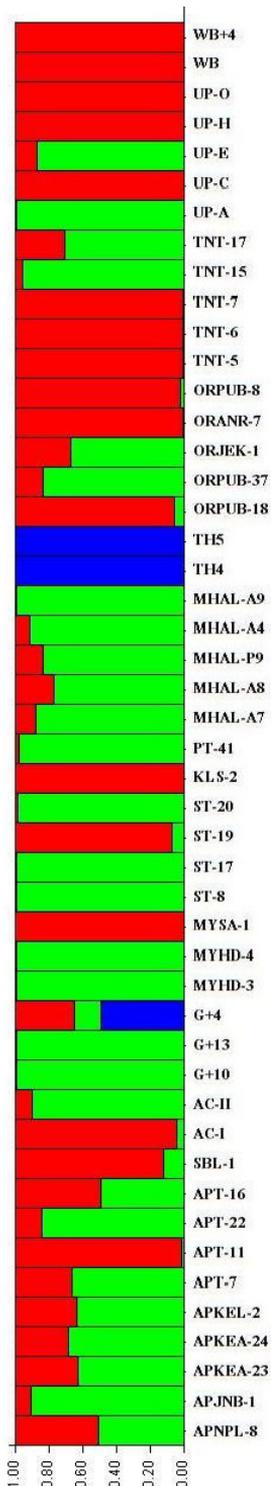


Figure 5

Bar plot for 46 genotypes of *T. grandis* L. f. and two genotypes of *T. hamiltoniana* wall. (TH) generated by program STRUCTURE applying admixing with correlated allele frequencies model. Each color bar represents single genotypes of corresponding states. The acronyms of the related sampling state are given in Table 1.

Discussion

A genotype-phenotype association study gets enhanced power with a sampling strategy that covers the most of the heterogeneity in the species genome for the targeted trait making DNA markers under investigation as effective as those of single nucleotide polymorphism (Amos et al. 2011). Moreover, the GWAS or CGAS is advantageous only if the number of employed markers includes the functional mutation itself. Thus, this highlights that the robustness of marker trait association depends more on the number of markers than the sample size (Nazareno et al. 2017). To find out the QTL associated with the

wood density in teak, we sampled the 46 genotypes from a germplasm bank, assuming it to be true representation for phenotypic and genetic heterogeneity exhibited by the metapopulation of the species in India. The efficacy and the specificity of the applied AFLP markers to cover the most of the heterogeneity of *T. grandis* L. f. is verified including two genotypes of species *T. hamiltoniana* Wall. as an outlier.

Variation in wood density

The wood density of the sampled teak genotypes representing 8.71% of variation among the genotypes is higher than the within-species variation in wood density (2.9-7.2 %) exhibited by 14 hardwood species sampled from Canada (Alemdag et al. 1984) and the variation among teak genotypes (7.54 ± 2.90 %) from clonal orchards of two southern Indian locations (Indira & Bhat 1998). In order to investigate the genotype-phenotype association, the wood quality traits with lower variation (<5%) than what we obtained for teak has been successfully employed in *Pinus taeda* (González-Martínez et al. 2007), *Pinus radiata* (Dillon et al. 2010) and *Picea glauca* (Beaulieu et al. 2011). On the other hand, the normal distribution of the wood density strengthens the statistical power to detect the QTL (Li et al. 2015), avoiding the demographic influence in the teak genotypes of the germplasm bank. Thus, we find the variation and normal distribution of teak wood density exceptionally fit for our investigation.

The genetic information of AFLP markers and LD

The genetic polymorphism (63.04%) of the

Table 4 Statistics of AFLP markers for associations with wood density in 46 genotypes of teak (*Tectona grandis* L.f.) applying GLM and MLM models.

Marker	Model				
	GLM ^a		MLM ^b		LD block ^c
	F ^d	p ^f	F ^d	p ^f	
E-AAC/M-CAT_38	13.348	0.0009	12.223	0.0014	5
E-AAC/M-CAT_11	9.525	0.0042	9.525	0.0042	6
E-AAC/M-CAT_90	9.342	0.0046	14.489	0.0006	1
E-AAG/M-CTC_41*	-	-	9.405	0.0044	19

Note. Abbreviations: ^a - general linear model, ^b - mixed linear model, ^c - number of loci making LD block, ^{d,f} - value for the association, ^p - value for the association, * - no significant association in GLM model.

AFLP markers has been found comparable to the earlier findings reporting the polymorphism of AFLP on teak populations sampled from Costa Rica (Araya et al. 2005), Southern India (Sreekanth et al. 2012) and Central and Southern India (Vaishnav et al. 2015). The allelic frequency (0.83 ± 0.02) is also higher than those reported by nuclear SSR markers on teak (Fofana et al. 2008). The high polymorphism of markers is synonymous with the high mutation rate of the sequence covered, and their frequent presence on teak genome makes it suitable to cover the functional variant useful for the investigation. Unlike the other dominant markers, the AFLP markers are reproducible (Vos et al. 1995). Thus, the primer E-AAC/M-CAT with the highest genetic information can be further used to develop a diagnostic marker for different ecotypes of teak (Griffiths & Orr 1999).

The LD among the loci indicates their historical association for functional variants; the magnitude of LD and its decay with genetic distance determine the resolution of association mapping (Vos et al. 2017). Teak is an entomophilous species exhibiting $78.40 \pm 7.78\%$ of the variation within population consisted of highly admixed genotypes (Verhaegan et al. 2010, Hansen et al. 2015). We have also found 73.65% variation among the sampled genotypes; so the low number of loci-pair in LD (10.39%) is as per the expectations. On the other hand, the R^2 value as the measure of LD strength is not influenced by small size popu-

lation. Further, the loci-pair in significant LD (9.04%) with $R^2 > 0.2$ confirms the large size of LD blocks present in the species genome for high resolution mapping. The LD blocks may have a little influence of the low genetic structure of the highly admixed teak genotypes or of the limited number of historical recombination events caused by long breeding cycle in teak. As a result, no LD decay has been observed in the loci distributed up to 80 bp of distance. This is advantageous as the absence of LD decay may enhance the possibility of GWAS in teak even applying low-density markers (Neale & Savolainen 2004).

Genetic diversity and structure of the germplasm bank

A germplasm collection may have diverse genetic structure causing false-positive association (Furlotte et al. 2012). This makes essential to confirm genetic structure and relative kinship among the sampled genotypes prior to statistical modeling based association mapping. In our investigation, Hickory resulted in a lowest value of DIC for the full-model indicating the effect of inbreeding depression in the germplasm bank and its deviation from HWE. In this case, we incorporated the level of inbreeding ($F_{IS} = 0.43$) to estimate the genetic diversity measures through AFLP markers, resulting in low He (0.20 ± 0.04) value by AFLP-SURV compared to Hs (0.27 ± 0.01) value by Hickory. The polymorphism and the

heterozygosity (H_s and H_e) of the sampled teak germplasm bank is lower than the polymorphism (80.55%) and the heterozygosity (0.63 ± 0.09) of Indian teak metapopulation estimated by nuclear SSR markers (Fofana et al. 2008, Fofana et al. 2009, Hansen et al. 2015). This is expected due to genotype selection for superior growth and quality traits from the wild that leads to heterogeneity loss vis-à-vis genetic gain in germplasm collection. It may, however, be reiterated that the heterogeneity that could not be conserved in the germplasm has nothing to do with the functional variants linked with the wood density trait as we have observed enough variation in the wood density with a significant normal distribution.

The outgrouping of *T. hamiltoniana* Wall. from the sampled genotypes of *T. grandis* L. f. in phylogenetic trees and the PCoA has substantiated its inclusion as an outlier and confirmed the specificity of AFLP markers to detect true kinship among the teak genotypes. The 6.88% loci common to both species of the genus attest their common evolutionary link that also gains strength from *T. grandis* L. f. genotype G+4 exhibiting admixing of all three cryptic populations. Further, the 18.84% private loci of *T. hamiltoniana* Wall. also indicate its genetic divergence from *T. grandis* L. f. during the course of speciation on a long evolutionary scale. The differentiation of teak genotypes representing natural teak metapopulation in India actually confirms the earlier finding on teak depicting two centers of genetic origin of teak in India; central and peninsular India (Verhaegan et al. 2010, Vaishnav et al. 2015). For the first time, the present investigation incorporates the samples from North-Eastern states that have found to be slightly distinct from the reported centers of genetic origin of the species. The Bayesian algorithm of delta-K value has resulted in the very low genetic structure and very high admixing for the sampled genotypes. The low value for genetic differentiation measure ($F_{ST} = 0.16$) also validates the low differentiable set

of sampled genotypes. The possible explanation for the low genetic differentiation and structure in spite of higher polymorphism may be the selective sweeps associated with a local adaptation that has lowered the variation only over the local conditions (Rydholm et al. 2006). Interestingly, the low genetic structure with high admixing among the genotypes is a prior to significant association mapping avoiding false discovery; hence strengthens our statistical design.

The QTL associated with wood density

Applying a GLM model that does not consider the genetic structure and kinship among the genotypes, we have found three AFLP loci of marker E-AAC/M-CAT significantly associated with the wood density in teak. Consideration of the genetic structure and kinship among the genotypes as priori has helped us to avoid the false-negative error that has detected an additional locus, E-AAG/M-CTC_41 associated with the wood density trait. AFLP marker system is semi-conservative due to its selective assay of the genome and further amplification, and makes it reproducible and reliable. Hence, the markers found significantly associated with the wood density are robust and reproducible to perform the marker-assisted selection in teak.

In order to detect the marker associated with wood density, we sampled a heterogeneous set of teak genotypes from a germplasm bank that is known to have a high resolution of association mapping (Brescaglio & Sorrels 2006). Only a few number of microsatellite markers characterized for the species (Verhaegan et al. 2005) made us use reproducible and semi-conserved AFLP markers for the initial investigation of GWAS. To avoid the false discovery we carefully controlled the statistical analysis and found the genetically diverse genotypes with low genetic structure and exceptional LD of loci. The significant LD without decay within the distance of 80 bp representing large size

of LD blocks in diploid genome of the species has supported the efficacy of LD mapping of complex traits with fewer numbers of markers. With the MLM model incorporating genetic structure and relative kinship of the genotypes, we have found four AFLP markers associated with the wood density trait. These markers can further be developed to the sequence tagged site (STS) markers to apply in the marker-assisted selection of the teak with higher wood density. For a better confirmation, the investigation can be carried out with a larger number of genotypes employing next-generation sequencing technology.

Conclusions

We presented a preliminary investigation on AFLP markers-wood density association through sampling 46 teak genotypes representing the heterogeneity of teak metapopulation in India. Being a non-model species, no genomic information has been available for teak, which makes it difficult to explore the genetic causal factor controlling the wood trait. Since there is scarcity of the genomic information of the species, we believe that the present investigation will help enhance the marker database of the species. The markers in association with the wood density can be utilized for the recurrent selection of the improvement breeding program. It is also expected that the association of wood density with the AFLP loci would be verified in the future through genome sequencing studies, further promoting the application of marker assisted selection to teak genetic improvement and possibly other tree improvement and breeding programs.

Acknowledgements

Authors are grateful to Department of Biotechnology, Ministry of Science and Technology, Government of India, New Delhi, for the sanction of the research grant (BT/PR/3000/12

AGR/16/236/ 2002) to carry out this investigation.

References

- Alemdag, I. S. 1984. Wood density variation of 28 tree species from ontario. Canada: Petawawa National Forestry Institute, Canadian Forest Service.
- Amos, W., Driscoll, E., Hoffman, J. I., 2011. Candidate genes versus genome-wide associations: which are better for detecting genetic susceptibility to infectious disease? *Proceedings of the Royal Society of London Series B* 278: 1183-1188. DOI: 10.1098/rspb.2010.1920
- Ansari, S. A., Narayanan, C., Wali, S. A., Kumar, R., Shukla, N., Kumar, S. R. 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.
- Araya, E., Murillo, O., Agullar, G., Rocha, O., Woolbright, S., Kelm, P. 2005. Possibilities of breeding teak (*Tectona grandis* L.f.) in Costa Rica assisted by AFLP markers. *Kurú: Revista Forest (Costa Rica)* 2(5): 1-8.
- Beaulieu, J., Doerksen, T., Boyle, B., Clément, S., Deslauriers, M., Beauseigle, S., Poulin, P., Lenz, P., Caron, S., Rigault, P., Bicho, P., Bousquet, J., Mackay, J. 2011. Association genetics of wood physical traits in the conifer white spruce and relationships with gene expression. *Genetics* 110: 197-214. DOI: 10.1534/genetics.110.125781
- Bonin, A., Ehrlich, D., Manel, S. 2007. Statistical analysis of amplified fragment length polymorphisms data: a tool box for molecular ecologists and evolutionists. *Molecular Ecology* 16: 3737-3758. DOI: 10.1111/j.1365-294X.2007.03435.x
- Botstein, D., White, R. L., Skolnick, M., Davis, R. W. 1980. Construction of a genetic linkage map in man using restriction fragment length polymorphism. *American Journal of Human Genetics* 32: 314-331.
- Bradbury, P. J., Zhang, Z., Kroon, D. E., Casstevens, T. M., Ramdoss, Y., Buckler, E. S. 2007. TASSEL: Software for association mapping of complex traits in diverse samples. *Bioinformatics* 23 (19): 2633-2635. DOI: 10.1093/bioinformatics/btm308
- Breseghele, F., Sorrells, M. S. 2006. Association mapping of kernel size and milling quality in wheat (*Triticum aestivum* L.) cultivars. *Genetics* 172: 1165-1177. DOI: 10.1534/genetics.105.044586
- Chave, J. 2005. Measuring wood density for tropical forest trees: A field manual for the CTFS sites. Web: [http://www.rainfor.org/upload/ManualsEnglish/wood_density_english\[1\].pdf](http://www.rainfor.org/upload/ManualsEnglish/wood_density_english[1].pdf). Accessed 02.2018.
- Dillon, S.K., Nolan, M., Li, W., Bell, C., Wu, H. X., Southerton, S. G. 2010. Allelic variation in cell wall candidate genes affecting solid wood properties in natural populations and land races of *Pinus radiata*. *Genetics* 185 (4): 1477-1487. DOI: 10.1534/genetics.110.116582

- Eckert, A. J., Pande, B., Ersoz, E. S., Wright, M. H., Rashbrook, V. K., Neale, D. B. 2009. High-throughput genotyping and mapping of single nucleotide polymorphisms in loblolly pine (*Pinus taeda* L.). *Tree Genetics & Genomes* 5: 225-234. DOI: 10.1007/s11295-008-0183-8
- Evanno, G., Regnaut, S., Goudet, J. 2005. Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Molecular Ecology* 14: 2611-2620. DOI: 10.1111/j.1365-294X.2005.02553.x
- Excoffier, L., Lischer, H. E. L. 2010. ARLEQUIN suite ver 3.5: A new series of programs to perform population genetics analyses under LINUX and WINDOWS. *Molecular Ecology Notes* 10: 564-567. DOI: 10.1111/j.1755-0998.2010.02847.x
- Falush, D., Stephens, M., Pritchard, J. K. 2003. Inference of population structure using multilocus genotype data: linked loci and correlated allele frequencies. *Genetics* 164: 1567-1587.
- Falush, D., Stephens, M., Pritchard, J. K. 2007. Inference of population structure using multilocus genotype data: dominant markers and null alleles. *Molecular Ecology Notes* 7: 574-578. DOI: 10.1111/j.1471-8286.2007.01758.x
- Fofana, I., Ofori, D., Poitel, M., Verhaegen, D. 2009. Diversity and genetic structure of teak (*Tectona grandis* Linn f) in its natural range using DNA microsatellite markers. *New Forest* 37: 175-195. DOI: 10.1007/s11056-008-9116-5
- Fofana, I. J., Lidah, Y. J., Diarrassouba, N., N'guetta, S. P. A. 2008. Genetic structure and conservation of teak (*Tectona grandis*) plantations in Cote d'Ivoire, revealed by site specific recombinase (SSR). *Tropical Conservation Science* 1(3): 279-292. DOI: 10.1177/194008290800100308
- Furlotte, N. A., Kang, E. Y., Van Nas, A., Farber, C. R., Lusi, A. J., Eskin, E. 2012. Increasing association mapping power and resolution in mouse genetic studies through the use of meta-analysis for structured populations. *Genetics* 191: 959-967. DOI: 10.1534/genetics.112.140277
- Gonza'lez-Marti'nez, S.C., Huber, D., Ersoz, E., Davis, J. M., Neale, D. B. 2008. Association genetics in *Pinus taeda* L. II. Carbon isotope discrimination. *Heredity* 101: 19-28. DOI: 10.1038/hdy.2008.21
- Gonza'lez-Marti'nez, S.C., Wheeler, N. C., Ersoz, E., Nelson, C. D., Neale, D. B. 2007. Association genetics in *Pinus taeda* L. I. Wood property traits. *Genetics* 175: 399-409. DOI: 10.1534/genetics.106.061127
- Griffiths, R., Orr, K. 1999. The use of amplified fragment length polymorphism (AFLP) in the isolation of sex-specific markers. *Molecular Ecology* 8: 671-674. DOI: 10.1046/j.1365-294x.1999.00578.x
- Hammer, Ø., Harper, D.A.T., Ryan, P.D. 2001. Past: paleontological statistics software package for education and data analysis. *Palaeontology Electronica* 4: 9.
- Hansen, O. K., Changtragoon, S., Ponoy, B., Kjør, E. D., Minn, Y., Finkeldey, R., Nielsen, K. B., Graudal, L. 2015. Genetic resources of teak (*Tectona grandis* Linn f) strong genetic structure among natural populations. *Tree Genetics & Genomes* 11 DOI:10.1007/s11295-014-0802-5
- Holliday, J.A., Ritland, K., Aitken, S. N. 2010. Wide-spread, ecologically relevant genetic markers developed from association mapping of climate-related traits in Sitka spruce (*Picea sitchensis*). *New Phytology* 188: 501-514. DOI: 10.1111/j.1469-8137.2010.03380.x
- Holsinger, K. E., Lewis, P. O. 2007. Hickory, software for analysis of geographic structure in genetic data. Web: <http://darwin.eeb.uconn.edu/hickory/hickory.html>. Accessed: 02.2018.
- Hubisz, M. J., Falush, D., Stephens, M., Pritchard, J. K. 2009. Inferring weak population structure with the assistance of sample group information. *Molecular Ecology Research* 9: 1322-1333. DOI: 10.1111/j.1755-0998.2009.02591.x
- Indira, E. P., Bhat, K. M. 1998. Effects of site and place of origin on wood density of teak. *Journal of Tropical Forest Science* 10(4): 537-541.
- Ingvarsson, P. K., Garcia, M. V., Luquez, V. Hall, D., Jansson, S. 2008. Nucleotide polymorphism and phenotypic associations within and around the phytochrome B2 locus in European Aspen (*Populus tremula*, Salicaceae). *Genetics* 178: 2217-2226. DOI: 10.1534/genetics.107.082354
- Jansen, R. C. 1994. Controlling the type I and type II errors in mapping quantitative trait loci. *Genetics* 138: 871-881.
- Kaosa-ard, A. 1981. Teak (*Tectona grandis* L.f.) - its natural distribution and related factors. *Natural History Bulletin of Siam Society* 29: 55-74.
- Khan, M.A., Korban, S. S. 2012. Association mapping in forest trees and fruit crops. *Journal of Experimental Botany* 63: 4045-4060. DOI: 10.1093/jxb/ers105
- Kokutse, A. D., Bailleres, H., Stokes, A., Kokou, K. 2004. Proportion and quality of heartwood in Togolese teak (*Tectona grandis* L. f.). *Forest Ecology Management* 189: 37-48. DOI: 10.1016/j.foreco.2003.07.041
- Kumar, V., Kotrange, H. R., Dhotekar, U. P. 1998. Genetic improvement of teak. *Indian Forester* 124 (9): 687-695.
- Lepoittevin, C., Harvengt, L., Plomion, C., Garnier-Gere, P. 2012. Association mapping for growth, straightness and wood chemistry traits in the *Pinus pinaster* Aquitaine breeding population. *Tree Genetics & Genomes* 8: 113-126. DOI: 10.1007/s11295-011-0426-y
- Li, Z., Mottonen, J., Sillanpaa, M. J. 2015. A robust multiple locus method for quantitative trait locus analysis of non-normally distributed multiple traits. *Heredity* 115: 556-564. DOI: 10.1038/hdy.2015.61
- Liu, K., Muse, S. V. 2005. Integrated analysis environment for genetic marker data. *Bioinformatics* 21(9): 2128-2129. DOI: 10.1093/bioinformatics/bti282
- Ma, X., Hall, D., St.Onge, K. R., Jansson, S., Ingvarsson, P. K. 2010. Genetic differentiation, clinal variation and phenotypic associations with growth cessation across

- the *Populus tremula* photoperiodic pathway. *Genetics* 186: 1033-1044. DOI: 10.1534/genetics.110.120873
- Mackay, T. F. C. 2001. Quantitative trait loci in *Drosophila*. *Nature Review of Genetics* 2: 11-20. DOI: 10.1038/35047544
- McCarthy, M.I., Abecasis, G. R., Cardon, L. R. et al. 2008. Genome-wide association studies for complex traits: consensus, uncertainty and challenges. *Nature Review* 9: 356–369. DOI: 10.1038/nrg2344
- Moya, R., Marin, J. D., Murillo, O., Leandro, L. 2013. Wood physical properties, color, decay resistance and stiffness in *Tectona grandis* clones with evidence of genetic control. *Silvae Genetica* 62 (3). 142 -152. DOI: 10.1515/sg-2013-0019
- Myles, S., Peiffer, J., Brown, P. J., Ersoz, E. S., Zhang, Z., Costich, D. E., Buckler, E. S. 2009. Association mapping: critical considerations shift from genotyping experimental design. *Plant Cell* 21: 2194 – 2202. DOI: 10.1105/tpc.109.068437
- Narayanan, C., Wali, S. A., Shukla, N., Kumar, R., Mandal, A. K., Ansari, S. A. 2007. RAPD and ISSR markers for molecular characterization of teak (*Tectona grandis*) plus trees. *Journal of Tropical Forest Science* 19 (4): 218 -225.
- Nazareno, A. G., Bemmels, J. B., Dick, C. W., Lohmann, L. G. 2017. Minimum sample sizes for population genomics: an empirical study from an Amazonian plant species. *Molecular Ecology Resources* 17(6): 1136 – 1147. DOI: 10.1111/1755-0998.12654
- Neale, D. B., Savolainen, O. 2004. Association genetics of complex traits in conifers. *Trends in Plant Science* 9: 325–330. DOI: 10.1016/j.tplants.2004.05.006
- Oraguzie, N.C., Rikkerink, E.H.A., Gardiner, S.E. 2007. Association mapping in plants. Springer Science + Business Media, LLC, New York. DOI: 10.1007/978-0-387-36011-9
- Peakall, R., Smouse, P. E. 2012. GenAlEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research - an update. *Bioinformatics* 28: 2537-2539. DOI: 10.1093/bioinformatics/bts460
- Perrier, X., Jacquemond-Collet, J. P. 2006. DARwin software. Web: <http://darwin.cirad.fr/darwin/>. Accessed: 02.2018.
- Prevost, A., Wilkinson, M.J. 1999: A new system of comparing PCR primers applied to ISSR fingerprinting of potato cultivars. *Theoretical & Applied Genetics* 98: 107–112. DOI: 10.1007/s001220051046
- Pritchard, J. K., Stephens, J. K., Rosenberg, N. A., Donnelly, P. 2000. Association mapping in structured populations. *American Journal of Human Genetics* 67: 170–181. DOI: 10.1086/302959
- Quesada, T., Gopal, V., Cumbie, W. P., Echert, J. A. J., Wegrzyn, L., Neale, D. B. 2010. Association mapping of quantitative disease resistance in a natural population of loblolly pine (*Pinus taeda* L.). *Genetics* 186: 677–686. DOI: 10.1534/genetics.110.117549
- Rao, N. K. 2004: Plant genetic resources: Advancing conservation and use through biotechnology. *African Journal of Biotechnology* 3(2): 136- 145.
- Rydhholm, C., Szakacs, G., Lutzoni, F. 2006. Low genetic variation and no detectable population structure in *Aspergillus fumigatus* compared to closely related Neosartorya species. *Eukaryotic Cell* 5: 650–657. DOI: 10.1128/EC.5.4.650-657.2006
- Slatkin, M. 2008. Linkage disequilibrium—understanding the evolutionary past and mapping the medical future. *Nature Reviews of Genetics* 9: 477–485. DOI: 10.1038/nrg2361
- Spiegelhalter, D. J., Best, N. G., Carlin, B. P., van der Linde, A. 2002. Bayesian measures of model complexity and fit. *Journal of the Royal Statistical Society Series B Statistical Methodology* 64: 583– 689. DOI: 10.1111/1467-9868.00353
- Sreekanth, P. M., Balasundaran, M., Nazeem, P. A., Suma, T. B. 2012. Genetic diversity of nine natural *Tectona grandis* L. f. populations of the Western Ghats in Southern India. *Conservation Genetics* 13: 1409-1419. DOI: 10.1007/s10592-012-0383-5
- Thumma, B. R., Nolan, M. F., Evans, R., Moran, G. F. 2005. Polymorphisms in cinnamoyl CoA reductase (CCR) are associated with variation in microfibril angle in *Eucalyptus* spp. *Genetics* 171: 1257–1265. DOI: 10.1534/genetics.105.042028
- Thumma, B. R., Matheson, B. A., Zhang, D., Meeske, C., Meder, R., Downes, G. M., Southerton, S. G. 2009. Identification of a Cis-acting regulatory polymorphism in a Eucalypt COBRA-Like gene affecting cellulose content. *Genetics* 183 (3): 1155-1164. DOI: 10.1534/genetics.109.106591
- Vaishnav, V., Mohammad, N., Wali, S. A., Kumar, R., Tripathi, S. B., Negi, M. S., Ansari, S. A. 2015. AFLP markers for analysis of genetic diversity and structure of teak (*Tectona grandis* L f) in India. *Canadian Journal of Forest Research* 44: 297- 306. DOI: 10.1139/cjfr-2014-0279
- Vekemans, X. 2002. AFLP-SURV version 1.0. Laboratoire de Génétique et Ecologie Végétale. In. Université Libre de Bruxelles, Belgium
- Verhaegan, D., Ofori, D., Fofana, I. J., Poitel, M., Vaillant, A. 2005. Development and characterization of microsatellite markers in *Tectona grandis* (Linn. f). *Molecular Ecology Note*. 5: 945– 947. DOI: 10.1111/j.1471-8286.2005.01124.x
- Verhaegen, D., Fofana, I. J., Logossa, Z. A., Ofori, D. 2010. What is the genetic origin of teak (*Tectona grandis* L) introduced in Africa and Indonesia? *Tree Genetics & Genomes* 6: 717-733. DOI: 10.1007/s11295-010-0286-x
- Vos, P., Hogers, R., Bleeker, M., Reijmans, M., Vandelee, T., Hornes, M., Frijters, A., Pot, J., Peleman, J., Kuiper, M., Zabeau, M. 1995. AFLP - a new technique for DNA fingerprinting. *Nucleic Acids Research* 23: 4407– 4414. DOI: 10.1093/nar/23.21.4407
- Vos, P. G., Paulo, M. J., Voorrips, R. E., Visser, R. G., van Eck, H. J., van Eeuwijk, F. A. 2017. Evaluation of LD decay and various LD-decay estimators in simulated

- and SNP-array data of tetraploid potato. *Theoretical & Applied Genetics* 130(1): 123–135. DOI: 10.1007/s00122-016-2798-8
- Wilcox, P. L., Echt, C. E., Burdon, R. D. 2007. Application of association genetics for forest tree breeding. In: Oraguzie N.C., Rikkerink E.H.A., Gardiner D.E., De Silva H.N. (eds), *Association mapping in plants*. Springer. New York, USA, pp. 213–249. DOI: 10.1007/978-0-387-36011-9_10
- Yu, J., Pressoir, G., Briggs, W. H., Bi, I. V., Yamasaki, M., Doebley, J. F., McMullen, M. D., Gaut, B. S., Nielsen, D. M., Holland, J. B. 2006. A unified mixed-model method for association mapping that accounts for multiple levels of relatedness. *Nature Genetics* 38: 203–208. DOI: 10.1038/ng1702
- Zhivotovsky, L.A. 1999. Estimating population structure in diploids with multilocus dominant DNA markers. *Molecular Ecology* 8: 907–913. DOI: 10.1046/j.1365-294x.1999.00620.x