HAL
open science

What is the genetic origin of teak (Tectona grandis L.) introduced in Africa and in Indonesia?

Daniel Verhaegen, Jesus Fofana Inza, A. Zénor Logossa, Daniel Ofori

To cite this version:

Daniel Verhaegen, Jesus Fofana Inza, A. Zénor Logossa, Daniel Ofori. What is the genetic origin of teak (Tectona grandis L.) introduced in Africa and in Indonesia?. Tree Genetics and Genomes, 2010, 6, pp.717-733. cirad-00846130

HAL Id: cirad-00846130

https://hal.science/cirad-00846130

Submitted on 18 Jul 2013

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Tree Genetics & Genomes (2010) 6:717–733
DOI 10.1007/s11295-010-0286-x

ORIGINAL PAPER

What is the genetic origin of teak (*Tectona grandis* L.) introduced in Africa and in Indonesia?

Daniel Verhaegen · Inza Jesus Fofana · Zénor A. Logossa · Daniel Ofori

Received: 10 September 2009 / Revised: 5 March 2010 / Accepted: 19 March 2010 / Published online: 22 April 2010
© Springer-Verlag 2010

Abstract The first introduction of *Tectona grandis* in Indonesia took place between the fourteenth and sixteenth centuries and in Africa in the nineteenth century. A total of 1.1 and 0.3 million ha, respectively were planted in the two areas. The extension of teak plantations often started from these first introductions. Unfortunately, the documentation concerning dates, planting stocks and the sources of origin of the teak imported into the different countries was very inaccurate. In this study, the use of 15 microsatellite molecular markers enabled us to compare 22 exotic teak provenances with 17 provenances of the natural range. Results of the analysis showed that the provenances from South India were not related to the provenances that were first introduced in either Africa or Indonesia. Nearly 95% of teak landraces in Benin, Cameroon, Côte d'Ivoire, Tanzania, Togo and Senegal came from North India, and 96% of Indonesian and Ghanaian teak appeared to be very closely linked to Central Laos. The genetic origin of introduced teak was confirmed by the main traits of interest of provenances observed in international trials. Thus, trees from North India had very bad stem forms compared to Laotian and Thai provenances, which generally had good stem forms but low vigour. This genetic knowledge is essential for programmes to develop varieties and to improve the quality of plantations, particularly in Africa.

Keywords Genetic diversity · Genetic structure · Phylogenetic origin · Microsatellite · *Tectona grandis*

Communicated by: A. Kremer

D. Verhaegen (✉) · I. J. Fofana · Z. A. Logossa · D. Ofori
CIRAD Biological System Department Research Unit 39: "Genetic Diversity and Breeding of Forest Tree Species", Campus International de Baillarguet TA A-39/C, 34398 Montpellier Cedex 5, France
e-mail: daniel.verhaegen@cirad.fr

I. J. Fofana
Laboratoire de Génétique, Université de Cocody, Abidjan 22, BP 582, Abidjan 22, Côte d'Ivoire

Z. A. Logossa
Université de Lomé Faculté des Sciences, BP 1515, Lomé, Togo

D. Ofori
Council for Scientific and Industrial Research, Forestry Research Institute of Ghana, Biodiversity and Land Use Division, Biotechnology Laboratory, UST, P.O. Box 63, Kumasi, Ghana

Introduction

Teak (*Tectona grandis* L.f.) is one of the most important tropical hardwood species in the international market of high-quality timber extracted from both natural forest and plantations. Teak forests occur naturally in India, Myanmar, Thailand and Laos and cover an area of about 23 million ha. Teak is also grown in plantations in at least 36 countries throughout the tropics and about 5.7 million ha of teak plantations are recorded (Bhat and Ma 2004). Teak constitutes about 75% of the world's high-quality tropical hardwood plantations (FAO 2001). About 43% of all teak plantations are located in India, 31% in Indonesia, 7% in Thailand, 6% in Myanmar and 5% in tropical Africa. In Africa, teak plantations are mainly concentrated in Côte d'Ivoire, Nigeria, Ghana, Sudan, Togo, Benin and Tanzania.

Export of teak genetic material outside its natural range is very recent and is hardly more than a century old in many countries. Unfortunately, the exact sources of the genetic materials were not well documented, sometimes forgotten, or were hypothetical, and it is thus difficult to trace the

origins of the materials in existing plantations. Today it is also impossible to say in which form (seeds or stumps) and how much teak was exported outside its natural range for the first time, which is believed to have been to the island of Java where it was used as planting material. The scientific community agrees that teak was naturalised in Java for several centuries, has been acclimatised to ecological conditions there, and has regenerated naturally throughout the area (Altona 1922; Begué 1957; Behaghel 1999; Kaosa-ard 1999; Pandey and Brown 2000; White 1991). The probable establishment of these plantations was during the period of Hindunization, from the beginning of the fourteenth to the sixteenth century. The local presence of a sizeable Hindu population at that time justifies the assumption that Hindus planted the teak. It was assumed that teak seedlings were imported from South India by Brahman monks (Altona 1922). At almost the same time, Buddhism was being introduced in Indonesia by missionaries. Teak seedlings were probably planted in rows around temples as teak was considered to be a religious tree, and the incarnation of the souls of the ancestors was believed to reside in the teak tree. Today teak is still planted around temples in North Thailand.

In Africa, the first teak introduced in trial gardens can be attributed to the German colonial administration at the end of the nineteenth century, and the extension of these species trials appears to have occurred at the beginning of the twentieth century. In Tanzania, teak is assumed to have been introduced in 1898 at sites in Dar Es Salam and Mhoro using seeds from Calcutta (Madoffe and Maghembe 1988) or from several genetically distant populations (moist, south western part of India), Myanmar and/or Java (Kjaer and Siegismund 1996; Wood 1967). Trial plots were established between 1905 and 1936 with provenances from Myanmar (Tennasserim), Indonesia, South India (Travancore) and Thailand (Rance and Monteuuis 2004). Provenances in Mtibwa and Bigwa are first and second generations of the Kihuwi population (Pedersen et al. 2007). In 1902, teak was planted for the first time in Nigeria using seeds from India and later using seeds from Myanmar or from Thailand (Chollet 1958). In Togo and Ghana, teak was introduced in 1905 using seeds from Nigeria (Chollet 1958; Kokutse 2002). In 1916, the first teak plantations were created in Benin. In Côte d'Ivoire, teak was planted for the first time in 1926 in the Banco National Park near Abidjan using seeds collected in Togo where natural regeneration was very abundant, then in 1929 in Bouaké (Tariel 1966). In Senegal, teak was introduced for the first time in 1932 in Casamance.

The history of the seed distribution pathway should make it possible to trace the introduction of teak in Indonesia and Africa, but in fact it is quite difficult to identify the exact origin of a specific introduction since African landraces are assumed to have originated from three different sources (India, Myanmar or Thailand). Many plantations have since been established using offspring of local trees (i.e., Togo, Ghana or Côte d'Ivoire, or Tanzanian landraces), whose sources are unknown. Our ignorance of the genetic origin of the teak that was introduced in different parts of the world raises a number of problems for researchers. First, the natural genetic variability of the provenances and the genetic relationship between the origins are poorly described; yet knowledge of the genetic pool and of natural genetic variability is indispensable for any genetic improvement programme for teak since it is required for the estimation of heritability and genetic gain, and also for genetic resources management. Provenance trials have shown that plantations based on seeds imported from South India can grow 30% faster and have a better stem form than plantations created with seeds from local sources (Kjaer et al. 1995). Investing in teak plantations with seeds from a narrow unknown range of genetic sources is very risky because of poor growth, bad form and high susceptibility to pests and diseases.

Several studies on the genetic diversity of teak have been conducted (Kertadikara and Prat 1995; Nicodemus et al. 2003; Shrestha et al. 2005), but none offers decisive information on the origin of the African landraces (Koskela et al. 2010). Given the existence of the microsatellite teak bank (Verhaegen et al. 2005) and a database containing SSR markers for teak in its natural range (Fofana et al. 2009), microsatellite marker analysis and Bayesian analysis could help determine the origin of African and Indonesian teak.

Materials and methods

Sampling for molecular DNA analysis

Between 1971 and 1973, seeds were collected from 17 natural populations originating from India, Thailand and Laos and distributed for international provenance trial by the Danida Forest Seed Centre (Keiding et al. 1986). These trials were used as a reference for samples of natural origin. Microsatellite DNA analysis revealed four main clusters, two of Indian origin (one on either side of latitude 19°25′N), one cluster consisted mainly of populations from Thailand and Laos, and the fourth cluster was made up of provenances from the centre of Laos (Fofana et al. 2009). This sample was used as the reference for natural genetic clusters.

In 2003, leaf samples were collected from 165 unknown trees in 22 African and Indonesian populations for three comparative provenance trials in Ghana and Côte d'Ivoire. A total of ten African and four Indonesian landraces

selected by Danida were completed by eight local African landraces (Table 1). The aim of the collections was to obtain as broad a representation as possible of the whole range of distribution, covering both the most typical and distinctly different types of environments (Keiding et al. 1986). In order to represent the maximum variability within each provenance, the samples were collected randomly, trees could be crooked, forked or buttressed as well as skewed or have many protuberant buds. This sample was used to establish the origins of African and Indonesian teak.

SSR genotyping and polymorphism

DNA extraction, PCR conditions and electrophoresis conditions were as described in (Verhaegen et al. 2005). Automated infrared fluorescence DNA sequencing was used to identify allele variability according to (Steffens et al. 1993). Fifteen microsatellite loci (the same loci as those used by Fofana et al. 2009) were amplified by PCR in a 15-µl reaction volume containing: 25 ng of genomic DNA in a 0.5× reaction buffer (10 mM Tris–HCl, 50 mM KCl, and 2 mM $MgCl_2$), 0.2 mM dNTPs, 0.10 µM of forward primer, 0.06 µM of reverse primer, 0.10 µM of IRdye M13/700 or M13/800 and 0.13 U/µl Taq DNA polymerase (Invitrogen™). Amplifications were carried out with a thermal-cycler Stratagene® Robocycler gradient 96 under the following conditions: denaturation at 94°C for 4 min; 30 denaturation cycles at 94°C for 30 s, an annealing cycle at 51°C for 45 s and an extension at 72°C for 45 s, and a final extension at 72°C for 5 min. The reverse PCR primers were probed with a 19-base extension at its 5′ tail end with the sequence 5′-CACGACGTTG TAAAACGAC-3′. This sequence is complementary to an IR-labelled universal M13 forward sequencing primer included in the PCR reaction. During PCR, the tailed primer generates a complementary sequence, which is subsequently used for priming in the amplification reaction thereby generating IR-labelled PCR products. The samples were electrophoresed on an IR DNA analyzer (LI-COR, Inc.), which detects primer-labelled extension products at two different wavelengths (IRDye 700 nm and IRDye 800 nm), this enabled the loading of a multiplex of four PCR products in one well. The individual trees were genotyped using Saga^{GT} software (LI-COR, Inc).

Table 1 The 22 unknown tested provenances cover a wide area in Africa and Indonesia

Harvest number	Country	Provenance name or local name	Latitude	Longitude	Annual rainfall (mm)	Trials	*N*
5	Benin	Djigbé	6°52′	2°20′	1,100	Séguié	5
6		Toffo Lama	6°52′	2°07′	1,100	Séguié	6
7	Cameroon	Bambuku	4°15′	9°15′	1,780	Séguié	5
3067[a]		Bambuku	4°26′	9°16′	1,900	Téné	5
4	Côte d'Ivoire	Bamoro A20	7°48′	5°05′	1,200	Séguié	5
3		Bamoro A29	7°48′	5°05′	1,200	Séguié	5
3037[a]		Bouaké	7°48′	5°07′	1,200	Téné	7
TB 73[a]		Landrace	7°48′	5°07′	1,200	Téné	6
3044[a]	Ghana	Jema	7°50′	1°50′	1,100–1,600	Tain II	9
SG 01[a]		Landrace	7°50′	1°50′	1,100–1,600	Tain II	9
SG 03[a]		Landrace	7°50′	1°50′	1,100–1,600	Tain II	9
SG 04[a]		Landrace	7°50′	1°50′	1,100–1,600	Tain II	8
3065[a]	Tanzania	Bigwa	6°50′	38°39′	900	Téné	14
3066[a]		Kihuwi	5°12′	38°39′	1,880	Séguié	10
17		Mtibwa	6°	37°39′	1,160	Séguié	10
3063[a]	Togo	Tové	6°40′	0°40′	1,300	Téné	5
1	Sénégal	Djibelor	12°35′	16°06′	1,650	Séguié	6
2		Kalounayes	12°45′	16°05′	1,650	Séguié	5
3047[a]	Indonesia	Bangsri, Pati	6°30′	110°48′	3,900	Tain II	8
3048[a]		Nanas, Blora	6°57′	111°30′	1,700	Tain II	10
3049[a]		Ngliron, Ngliron	7°12′	111°22′	1,200	Tain II	9
3050[a]		Temandsang	7°12′	111°22	1,200	Tain II	9

The provenance names of eight countries with their main geographic traits are given. The trials of Séguié and Téné were planted in Côte d'Ivoire and Tain II in Ghana

N number of trees used

[a] Number designated by Danida Forest Seed Centre

Analysis of molecular genetic data

Genotyping errors were estimated according to the recommendations of Pompanon et al. (2005). For each SSR locus, the DNA of three heterozygote teak samples of contrasted natural origins whose alleles were known, was mixed and used at the same time as the internal control and ladder of size reference to estimate the size of the unknown alleles. Ultimately, genotyping errors were identified on eight loci (2TB07; 1TF05; 1TA06; 3TF01; 2TC03; 3TB02; 4TD12; and 4TF02), 22 repeated known genotypes with 42 alleles repeated 32 times.

The origin of the exotic provenances was determined by two complementary analyses. The first used the teak from its natural range as the reference cluster (Fofana et al. 2009). Geneclass2 software (Piry et al. 2004) was used to assign or exclude natural genetic clusters as the origin of 165 diploid individuals from 22 exotic provenances on the basis of multi-locus genotype data. The 165 teak trees of unknown origin are grouped together in only one data file. The probability for each individual to be assigned to one or more of the defined genetic clusters was estimated using a Bayesian criterion (Rannala and Mountain 1997) and a frequency criterion (Paetkau et al. 2004). The log likelihood of each individual multi-locus genotype in each cluster was calculated assuming that the individual came from the genetic cluster in question. The log–log plots of genotypes for pairs of population likelihood were drawn to identify genotypes that appeared to be better explained by belonging to another cluster than the one in which they were sampled (Excoffier et al. 2007; Paetkau et al. 1997).

The second method consisted of adding the 165 individuals from the natural range to the 166 individuals from the exotic provenances. With the complete genotyped data, the 331 individuals were subdivided into genetic clusters using a model-based clustering method to infer population structure and assign individuals to populations with the software package Structure (Pritchard et al. 2000). This programme estimates the number of genetically homogeneous populations (K) and does not require prior information on the number of locations or on the location where each individual was sampled. At least ten clustering runs were carried out by setting the number of populations (K) from 1 to 8. For each run, burnin time and replication number were respectively 40,000 and 400,000. Two models for the ancestry of individuals included in the software were used. Individuals can have mixed ancestry (admixture model) or come from one of the K populations (no admixture model). The true number of populations (K) is often identified using the maximal value of Ln Pr($X|K$) returned by the software. However, the authors warn that Pr($X|K$) is really only an indication of the number of clusters (p. 949 in Pritchard et al. 2000). Evanno et al. (2005) observed that in most simulations, once the real K is reached, Pr($X|K$) at larger K levels off or continues to increase only slightly. In order to detect the uppermost hierarchical level of structure, the statistic ΔK was calculated based on the rate of change in the log probability of data between successive K values (Evanno et al. 2005).

Pairwise genetic distances between pairs of provenances were computed with the Cavalli-Sforza chord measure (Cavalli-Sforza and Edwards 1967) of the Microsatellite Analyzer (Dieringer and Schlötterer 2003). The results included the exotic teak populations added to the natural teak populations studied by Fofana et al. (2009). The distance tree was constructed using the neighbour-joining method (Saitou and Nei 1987). The robustness of each node was evaluated by bootstrapping data over loci for 1,000 replications using the Seqboot programme of PHYLIP 3.67 (Felsenstein 2005). The consensus tree obtained using PHYLIP 3.67 was displayed with Darwin 5.0.148 software (Perrier and Jacquemoud-Collet 2006).

To investigate the hierarchical structure of genetic variation, analysis of molecular variance (AMOVA) was performed using Arlequin ver 3.11 (Excoffier et al. 2005) with 1,000 permutations, which tests the genetic structure by partitioning the total variance into covariance components due to intra-individual differences, inter-individual differences and/or inter-population differences. Components of genetic variance were computed at two hierarchical levels: among exotic teak populations added to the natural teak populations studied by Fofana et al. (2009), and among the genetic clusters found with the model-based clustering method (Pritchard et al. 2000).

Genetic diversity within populations and genetic clusters were estimated by the number of alleles per locus (Ao), the expected unbiased (H_{nb}) and the observed (H_0) heterozygosity (Nei 1978) using Genetix 4.05.2 software (Belkhir et al. 1996–2004). To check if the differences in sample size and the different spatial scales at which individuals were pooled into provenances and regions affected the diversity estimates, the allelic richness (El Mousadik and Petit 1996) was calculated per provenance and genetic cluster taking into account the dependence on sample size with an adaptation of the rarefaction index (Hurlbert 1971). The calculation was performed with Fstat 2.9.3.2 software (Goudet 2001). The principle is to estimate the expected number of alleles in a sub-sample of $2n$ genes, given that $2N$ genes have been sampled ($N>n$), with n defined as the smallest number of individuals typed for a locus sample.

Results

Genotyping errors

With the 232 PCR amplifications on eight SSR loci, the mean error rate per allele was estimated at 89×10^{-5}; the error rate per locus varied from 0 (loci: 4TD12; 3TB02; 4TF02; 1TA06; 1TF05; 3TF01; 2TC03) to 33×10^{-3} (locus: 2TB07), and the mean error rate per locus was calculated at 43×10^{-4}, and the error rate per multi-locus genotype and the error rate per reaction were 4×10^{-4}.

Population genetic structure

The individual-based assignment method with Bayesian criterion placed 86% of the 36 teak trees growing in Indonesia and 83% of the 35 teak trees growing in Ghana in the genetic cluster from Central Laos (Table 2). Four Indonesian and five Ghanaian trees were assigned to the genetic cluster from North India. One Indonesian and one Ghanaian tree were assigned to the Thai cluster, and no Indonesian or Ghanaian trees were assigned to genetic cluster from South India. Concerning the African provenances, the method assigned 82%, 70%, 78%, 82%, 60% and 73%, respectively of the trees from Benin, Cameroon, Côte d'Ivoire, Tanzania, Togo and Senegal to the genetic cluster from North India. Sixteen, two and two African trees were assigned to the clusters from Thailand, South India and Central Laos, respectively.

The Bayesian criterion assigned 85% of the trees from Indonesia and Ghana to the genetic cluster in Central Laos and 78% of the trees from Benin, Cameroon, Côte d'Ivoire, Tanzania, Togo and Senegal to the genetic cluster in North India. Assignment using the Frequency criterion gave very similar results to those obtained using the Bayesian criterion with 75% of the trees from Indonesia and Ghana assigned to the genetic cluster in Central Laos and 69% of the trees from Benin, Cameroon, Côte d'Ivoire, Tanzania, Togo and Senegal in the genetic cluster in North India (Table 2).

The log–log plots of genotypes for pairs of clusters showed a clear separation between trees from South India and trees belonging all the other provenances tested (Fig. 1a to c) and also strong opposition between the genetic cluster from Thailand and the cluster comprising Laos, Indonesia and Ghana (Fig. 1d). Ultimately the graphs confirmed a change in cluster of some African trees to the Thai cluster or the reverse, and also moved some trees from the cluster comprising Laos, Indonesia, and Ghana to the African cluster (Fig. 1e, f). The graph and the data tables made it possible to identify the genotypes that appeared to be better explained by belonging to another cluster than the one in which they were sampled. Only 12 trees (i031, i050, i153, i074, i107, i110, i218, i219, g020, g004 g031 and i065) did not fit their provenance and cluster origin.

Using a Bayesian approach, clustering was performed on the entire data set with increasing numbers of inferred clusters. With the admixture and the no admixture models and for ten runs of clustering with 40,000 burnin time and 400,000 replications, the most appropriate number of clusters for modelling the data did not appear clearly with the Ln Pr($X|K$) of the Structure software (Fig. 2a). The very strong genetic structuring related to the trees from South India (71 individuals) disturbed the analysis. On the other hand, when the 71 samples from South India were removed, all analyses using all the different models in the software for the 260 teak trees clearly revealed three genetic clusters (Fig. 2b). These clusters were analysed separately and none presented a genetic substructure (Fig. 2c–f). Ultimately, there were four genetic clusters for the 331 samples analysed (Fig. 3). Teak provenances were grouped in four clusters which corresponded to i) South India with provenances 15; 16; 3016; 3021; 20 and 3022, ii) North Thailand and South Laos with provenances 12; 10; 3038; 3040; 13; 3061 and 3054, iii) Central Laos, Indonesia and Ghana with provenances 3059; 3055; 3056; 3047; 3048; 3049; 3050; 3044; SG 01; SG 03 and SG 04, iv) North India, Benin, Cameroon, Côte d'Ivoire, Tanzania, Togo and Senegal with provenances 3034, 5, 6, 7, 3067, 4, 3, 3037, TB 73, 3065, 3066 and 17.

A total of 41 individuals in the sample presented a genome composed of two different origins (Fig. 3). These were:

- Thirty trees (i023, i022, i003, i008, i034, i042, i051, i053, i062, i199, i073, i065, i068, i069, i160, i044, i128, i133, i226, g018, g074, g082, g006, g060, g043, g065, g002, g014, g053 and g089) of 20 provenances (1, 2, 3, 6, 7, 17, 3066, 3034, 3063, 3065, 3067, 12, 3054, 3061, SG01, SG03 SG04, 3044, 3047 and 3048) had a small fraction of their genome belonging to another cluster, but their classification corresponded well with the cluster containing their provenance.
- Eleven trees (i031, i050, i153, i074, i107, i110, i218, i219, g020, g004 and g031) did not group with their provenance or cluster of origin. The level of confidence for a particular sample belonging to a particular origin varied from 65% to 100%. Two trees (i031 from Côte d'Ivoire and i050 from Cameroon) mainly clustered with samples from North Thailand and South Laos. Two trees (i153 from Tanzania and i074 from Togo) were grouped with the cluster comprising Central Laos, Indonesia and

Table 2 Assignment of 165 teak trees from 22 exotic provenances genotyped with 15 microsatellite markers to four pre-defined genetic clusters (Fofana et al. 2009) using the algorithm of GeneClass2 with criterion Rannala and Mountain (1997) and criterion of Paetkau et al. (2004)

			Cluster 1: South India		Cluster 2: North India		Cluster 3: Thailand		Cluster 4: Central Laos	
			P value		*P* value		*P* value		*P* value	
	Country	*N*	0.5–0.95	>0.95	0.5–0.95	>0.95	0.5–0.95	>0.95	0.5–0.95	>0.95
Rannala and Mountain (1997)	Indonesia	36			2	2	1		9	22
	Ghana	35			3	2	1		3	26
	Total	71			5	4	2		12	48
	Benin	11		1	3	6	1			
	Cameroun	10				7	2	1		
	Côte d'Ivoire	23	1		1	17	1	2	1	
	Tanzania	34	1		1	27	1	4		
	Togo	5			1	2	1		1	
	Senegal	11				8	1	2		
	Total	94	2	1	6	67	7	9	2	
Paetkau et al. (2004)	Indonesia	36			5	4	1		4	22
	Ghana	35			2	5	1		3	24
	Total	71			7	9	2		7	46
	Benin	11			3	4	2	2		
	Cameroun	10			1	6	1	2		
	Côte d'Ivoire	23			1	14	4		3	1
	Tanzania	34			1	26	3	4		
	Togo	5			1	1	1		2	
	Senegal	11			1	6	2	2		
	Total	94			8	57	13	10	5	1

N number of trees studied

Ghana. Four trees native of the Ban Pha Lai provenance of Thailand (i107, i110, i218, i219), two trees from Ghana (g020, g004) and one from Indonesia (g031) were assigned to the cluster of North India, Benin, Cameroon, Côte d'Ivoire, Tanzania, Togo and Senegal.

Analysis of population differentiation

All F_{ST} values were significant, except in the cluster containing North Indian and African populations (Table 3). F_{ST} values were 0.16 and 0.14 among 39 teak populations and four genetic clusters respectively. The F_{ST} values were 0.03, 0.08 and 0.12, respectively within six populations from South India, 11 populations from Central Laos, Indonesia and Ghana and seven populations from Thailand. The F_{IS} values were all non-significant except for the cluster comprising Laos, Indonesia and Ghana.

With the genetic distances of Cavalli-Sforza and Edwards, the 39 provenances were clearly separated and the robustness of nodes varied from 591/1,000 to 1,000/1,000 (Fig. 4). The phylogram allowed separation of four groups: i) the seven populations from South India (Nellicutha 15, Nellicutha 16, Virnoli Range, Masale Valley, Nilambur and Bairluty), which formed a single highly significant group (robustness of node 1,000/1,000). ii) The three central populations from Laos (Vientiane, Savannakhet I and Savannakhet II) together with four provenances from Indonesia (Ngliron Ngliron, Bangsri Pati, Temandsang and Nanas Blora) and four provenances from Ghana (Jema and Landraces G1, G3, G4), which were clearly separated from other provenances (robustness of node 1,000/1,000). iii) The provenance from North India (Purunakote) was grouped with the African populations (Bambuku 3067, Mtibwa, Bigwa, Toffo Lama, Tové, Djibelor, Bamoro A29, Kalounayes, Djigbé, Kihuwi, Bamoro A20, Bouaké), which were very clearly separated from all other origins or clusters (robustness of node 671/1,000). iiii) The group of provenances from Thailand and two Laotian populations (Ban Cham Pui, Mae Huat, Pong Salee, Pak Lai, Houi Na Soon, Ban Pha Lai and Pakse) formed the last group (robustness of node 604/1,000). The two provenances from Bambuku (7) and Landrace TB73 were not clearly associated with any of the clusters.

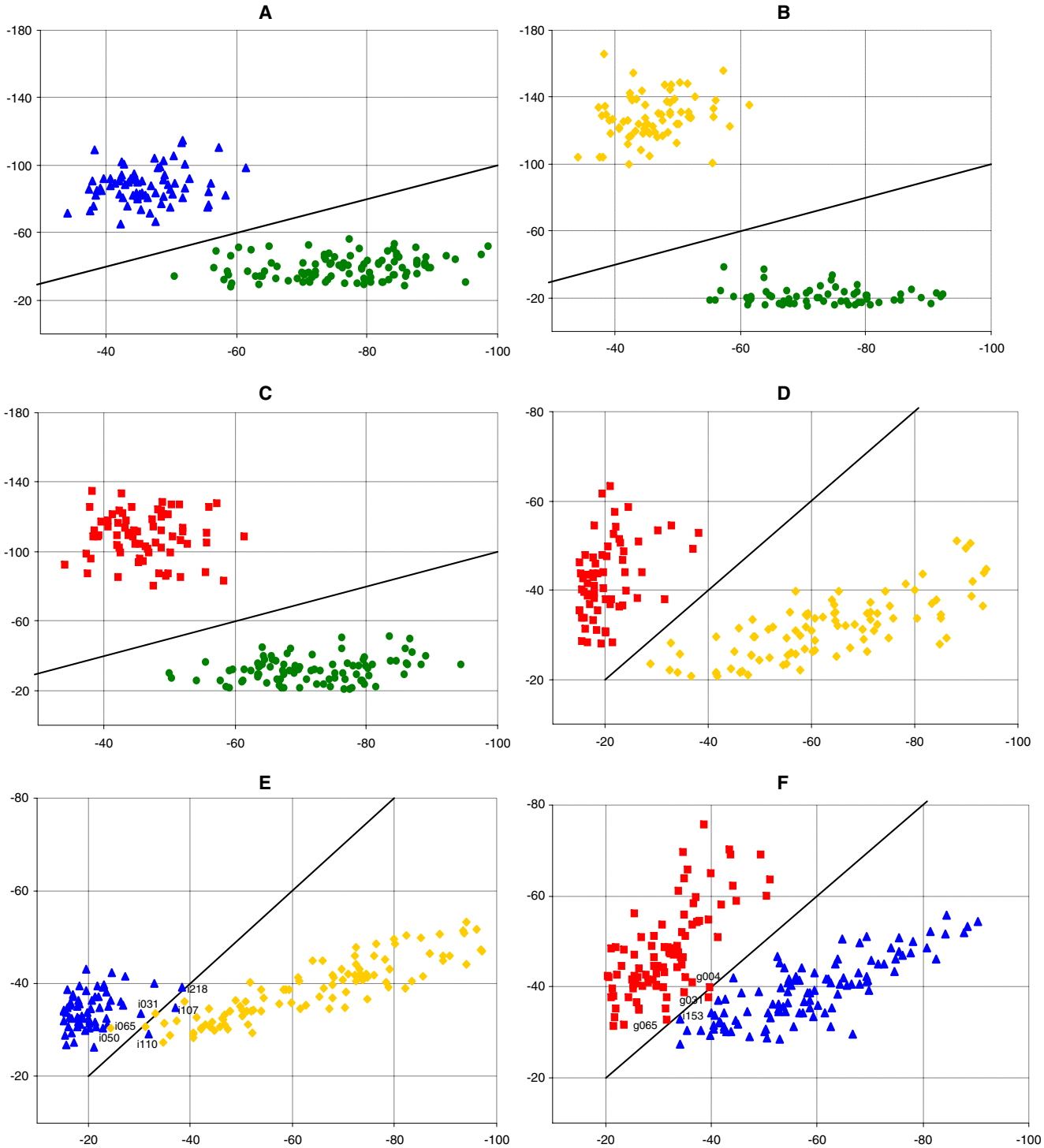

Fig. 1 Log likelihood of individuals sampled in one cluster versus those of another cluster for 15 SSR loci. **a** South India (*green circles*) vs. North India and Africa (*blue triangles*). **b** South India (*green circles*) vs. Thailand (*yellow diamonds*). **c** South India (*green circles*) vs. Laos, Indonesia and Ghana (*red squares*). **d** Thailand (*yellow diamonds*) vs. Laos, Indonesia and Ghana (*red squares*). **e** Thailand (*yellow diamonds*) vs. North India and Africa (*blue triangles*). **f** Laos, Indonesia and Ghana (*red squares*) vs. North India and Africa (*blue triangles*)

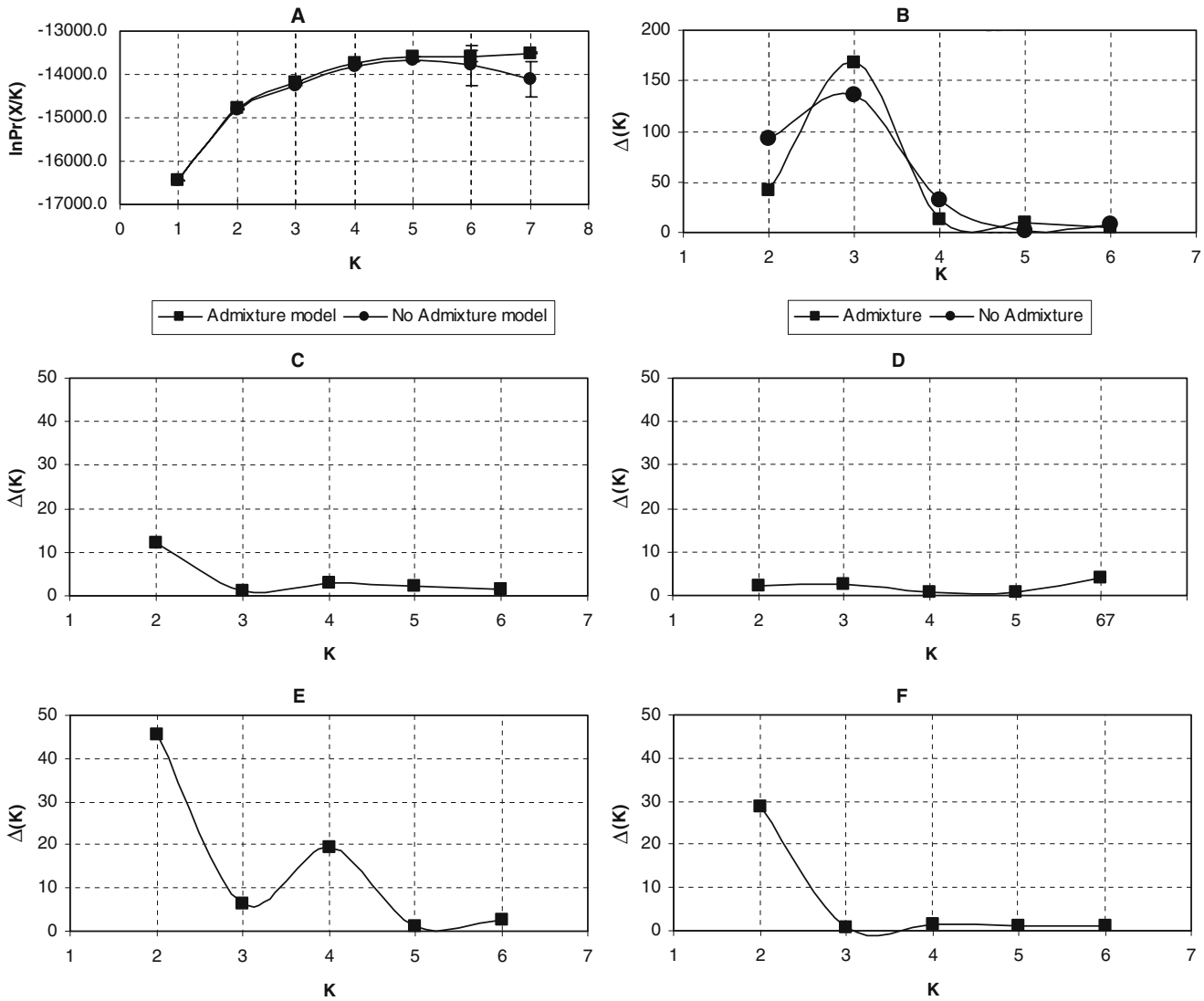

Fig. 2 Bayesian estimates of population structure based on microsatellite variation among 17 teak natural populations and 22 African and Indonesian populations. Description of the three steps allowing the detection of the true number of clusters K. **a** Distribution of Pr($X|K$) (+/−SD) (Pritchard et al. 2000) with 331 individuals and 15 SSR loci for admixture and no admixture models. **b** Magnitude of ΔK as a function of K (Evanno et al. 2005) calculated on 260 teak trees excluding 71 individuals from South India. **c** Magnitude of ΔK calculated on the South Indian cluster (71 individual trees). **d** Magnitude of ΔK calculated on the Thailand cluster (64 individual trees). **e** Magnitude of ΔK calculated on the cluster including Laos, Indonesia and Ghana (92 individual trees). **f** Magnitude of ΔK calculated on the North India and Africa cluster (104 individual trees)

Within-population and within-cluster genetic diversity

The 15 microsatellite loci were polymorphic across all 331 genotypes, and the number of alleles per locus ranged from 3 for 1TG02 to 21 for 1TH10 (Table 4). Considering each locus, the distribution of allele frequencies was not unbalanced for ten loci, and one allele presented a frequency higher than 50% for the following loci: 4TF02 (size of 227 bp), 4TH09 (size of 157 bp), 3TE06 (size of 218 bp), 3TD09 (size of 208 bp) and 1TG02 (size of 166 bp).

The number of rare alleles ranged from 1 to 13 for locus 1TG02 and 4TD12, respectively. No SSR locus presented rare alleles specific to only one provenance or only one genetic cluster. With fifteen SSR loci, the provenances of Bairluty, Mtibwa, Masale Valley and Purunakote showed 7.8%, 7.7%, 7.0% and 6.7% of rare alleles. The poorest provenances with rare alleles were Pong Salee, Huoi Na Soon, Pak Lai and Savannakhet I from 0.6% to 0%. Out of a total of 224 alleles, 91 alleles showed a frequency of under 1% and the first genetic cluster of provenances from South India (Nellicutha 15;

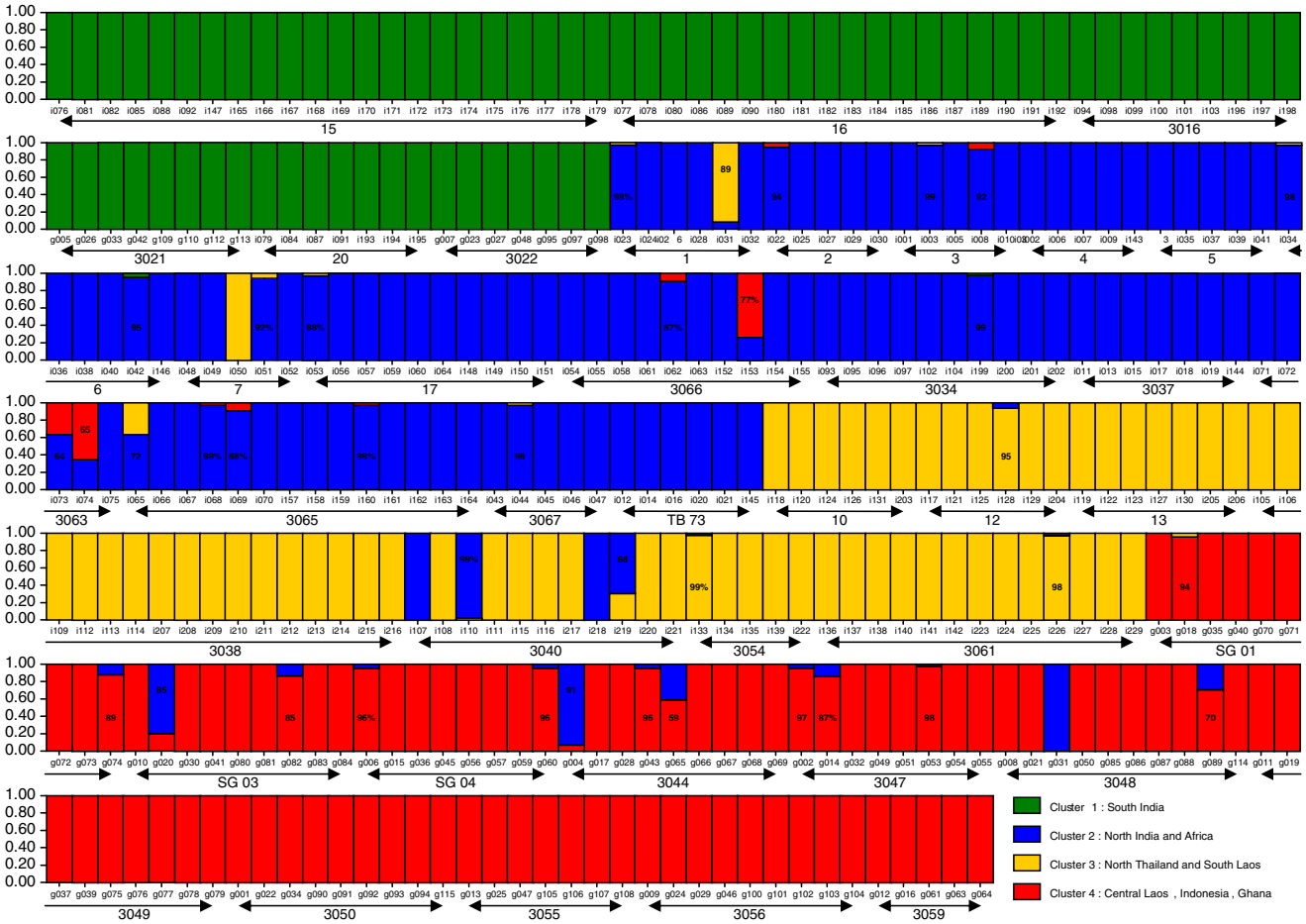

Fig. 3 Estimated population structure of 331 individuals from 17 provenances from the natural range and 22 African and Indonesian provenances of *Tectona grandis*. Each individual is represented by a *vertical bar*, which is assigned to one of the four colours that represent the estimated affiliation of the individual to one of the four clusters: South India; North India and Africa; North Thailand and South Laos; Central Laos, Indonesia and Ghana

Nellicutha 16; Masale Valley; Nilambur; Virnoli; Bairluty) and the second genetic cluster of provenances from Nord India added to African provenances (Purunakote; Djibelor; Kalounayes; Bamoro A29; Bamoro A20; Bouaké: Landrace TB73; Tové; Djigbé; Toffo Lama; Bambuku 7; Bambuku 3067; Kihuwi; Mtibwa; Bigwa) presented respectively 32.9% and 25.3% of low-frequency alleles. The two clusters of provenances from Thailand (Pong Salee; Mae Huat; Ban Cham Pui; Ban Pha Lai; Huoi Na Soon; Pak Lai; Pakse) and Laos added to provenances from Ghana and Indonesia (Vientiane; Savannakhet I; Savannakhet II; Bangsri Pati; Temandsang; Ngliron Ngliron; Nanas Blora; Jema; Landraces G1, G3, G4) presented only 3.3% of low-frequency alleles.

The observed (H_0) heterozygosity and the expected unbiased heterozygosity (H_{nb}) values ranged from 0.17 to 0.84 and from 0.20 to 0.87 for locus 1TG02 and 1TH10, respectively. All the loci showed a heterozygote deficit. The F_{IS} values were highly significant except for one locus 1TH10 (Table 4). In the samples studied, more than 12 alleles were identified per locus, with the exception of locus 3TD09 (9) and 1TG02 (3). Allelic richness varied between 2.66 and 15.95 for all the loci.

In the 22 teak from exotic provenances, the mean number of alleles per locus per population (A) ranged from 3.00 in Bambuku (7) to 5.93 in Bigwa, while allelic richness (R) ranged between 3.00 and 4.47 for Bambuku (7) and Djigbé, respectively (Table 5). The Pearson correlation coefficient between the number of alleles per locus and the allelic richness corrected with a rarefaction index was 0.41 and was highly significant, demonstrating a strong relationship between these two parameters. The allelic richness of the cluster from North Thailand (R= 5.07) was approximately half that of the cluster from South India (R=11.47) and of that from North India and Africa (R=10.26). With an allelic richness of 6.93, the cluster comprising Central Laos, Indonesia and Ghana was close to that from Thailand (Table 5).

Table 3 Hierarchical analysis of molecular variance based on 15 SSR analysed in 22 exotic provenances of *Tectona grandis* and in 39 populations regrouped in four genetic clusters defined using a model-based approach and pairwise genetic distances

Source of variation	*df*	SS	Variance components	Percent	F_{ST}	F_{IS}
Among 39 populations	38	687.93	0.8157***	16.0	0.1604***	0.0135^{NS}
Among individuals within populations	292	1,263.96	0.0578 ^{NS}	1.1		
Within individuals	331	1,394.50	4.2130***	82.8		
Total	661	3,346.38				
Among four clusters	3	391.91	0.7402***	14.1	0.1407***	0.0135 ^{NS}
Among populations within clusters	35	296.02	0.2502***	4.8		
Among individuals within populations	292	1,263.96	0.0578 ^{NS}	1.1		
Within individuals	331	1,394.50	4.2130***	80.1		
Total	661	3,346.38				
Among 6 populations of South India	5	46.88	0.1706*	3.0	0.0299*	−0.0001^{NS}
Among individuals within populations	65	360.21	0.0000 ^{NS}	0.0		
Within individuals	71	393.50	5.5423 ^{NS}	97.0		
Total	141	800.59				
Among 7 populations of Thailand	6	55.02	0.3638***	11.5	0.1151***	−0.0278^{NS}
Among individuals within populations	57	155.00	−0.0779 ^{NS}	−2.5		
Within individuals	64	184.00	2.8750**	91.0		
Total	127	394.02				
Among 11 populations of Laos Indonesia and Ghana	10	87.07	0.2910***	7.7	0.0769***	0.1014***
Among individuals within populations	81	311.84	0.3543***	9.4		
Within individuals	92	289.00	3.1413***	82.9		
Total	183	687.90				
Among 15 populations of North India and Africa	14	101.34	0.1897^{NS}	3.88	0.0388^{NS}	−0.0131^{NS}
Among individuals within populations	89	412.65	−0.0616^{NS}	−1.26		
Within individuals	104	495.00	4.7596*	97.4		
Total	207	1,008.99				

NS p value non significant; **p* value significant <0.05; ***p* value significant <0.01; ****p* value significant <0.001

Observed (H_0) heterozygosity values ranged from 0.43 in the Ngliron Ngliron population to 0.75 in the Bamoro A29 population and the expected unbiased heterozygosity (H_{nb}) values ranged from 0.49 to 0.70 in the Bambuku (7) and Bamoro A20 provenances, respectively. At the regional level, teak heterozygosity (H_0) was clearly higher in India and Africa (0.74 and 0.63, respectively) than in Thailand and Laos (0.38 and 0.47, respectively).

Discussion

Genotyping errors and missing data

In our study on teak, the 43×10^{-4} error rate per locus can be regarded as very low since the review of the literature by Pompanon et al. (2005) reported error rates ranging from 0.2% to more than 15%. Causes of genotyping errors can be: i) variation in the DNA sequence, ii) too small quantity or low quality of DNA, iii) biochemical artefacts and iv) human error (Pompanon et al. 2005). With regard to teak, genotyping errors were mainly caused by variation in the DNA sequence. Indeed, in our study, it was primarily the disappearance of one fragment that was the source of error in genotyping. This phenomenon is well known and is due to competition between fragments during PCR amplification (Wattier et al. 1998). Non-amplification of SSR alleles has been shown to be due to either mutations in flanking sequences or a low target DNA copy number (Taberlet et al. 1996). Genotyping errors in this study were probably due to either mutations in flanking sequences or a low target DNA copy number in only one heterozygote sample. The disappearance of one fragment can lead to overestimation of heterozygote deficiency and to erroneous conclusions regarding genetic populations, but in this study on teak, the overestimation of heterozygote deficiency was very low and the final results were strengthened by the low number of genotyping errors identified. Nevertheless, a recent study showed that the human factors represented nearly 93% of genotyping errors (Hoffman and Amos 2005). For teak, these human errors could have taken place either in nurseries, during the establishment of trial plots, during

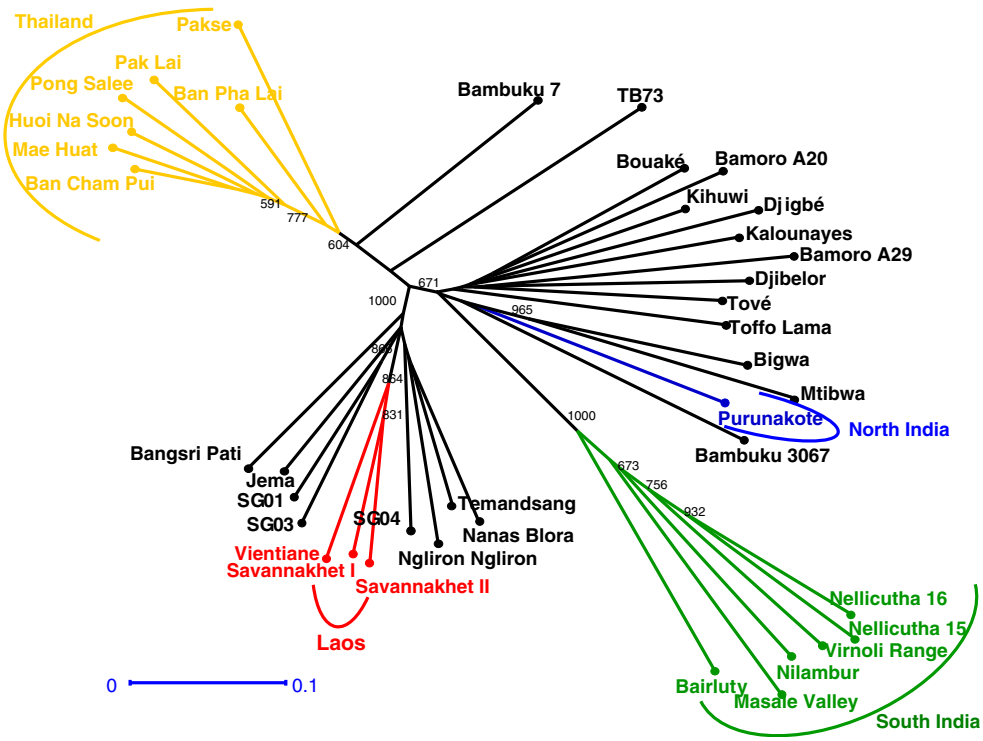

Fig. 4 Neighbour-joining phylogram based on the Cavalli-Sforza and Edwards chord method for 17 natural and 22 exotic provenances of *Tectona grandis*. The robustness of each node was evaluated by bootstrapping data over loci for 1,000 replications, only data >500 were noted

sampling in Côte d'Ivoire and Ghana, or while handling DNA in the laboratory. These human errors could explain the results of the trees i050, i107, i218 and g031 of which 100% of the genome corresponded to another genetic cluster rather than that of their origin.

Finally, for studies of the genetic structure of teak and the assignment of individuals to genetic clusters, the lack of a precise analytical correction for missing alleles was not a significant problem because results were relatively insensitive to the frequency assumed for the alleles concerned (Paetkau et al. 2004). However, the power to identify immigrants was improved by using a large sample (up to about 50 individuals) and by sampling all populations from which migrants may have originated. The genetic analyses of teak carried out in this study can thus be regarded as robust and reliable.

Genetic structure and assignment test

Concerning the assignment tests computed with Geneclass2, it is worth mentioning that criterion values were not comparable among individuals when based on a different

Table 4 Genetic diversity among 331 trees of *Tectona grandis* revealed by 15 SSR loci

No total number of observed alleles, *N*<1% number of alleles with a frequency <1%, H_0 observed heterozygosity, H_{nb} the expected unbiased heterozygosity corrected for small sample size (Nei 1978), F_{IS} the inbreeding coefficient (fixation index, Fisher) with *NS p* value adjusted using sequential Bonferroni (Rice 1989) procedure not significant, F_{ST} represents the differentiation among the four clusters within the total population, *R* is the corrected allelic richness

p* value significant <0.05;*p* value significant <0.01; ****p* value significant <0.001

Locus name	Accession number	Number	*N*<1%	H_0	H_{nb}	F_{IS}	F_{ST}	*R*
1TA06	AJ968929	20	6	0.76	0.85	0.106***	0.145	15.95
1TB03	AJ968930	15	6	0.70	0.81	0.129***	0.132	11.97
1TF05	AJ968931	15	6	0.59	0.71	0.172***	0.133	11.59
1TG02	AJ968932	3	1	0.17	0.20	0.156**	0.161	2.66
1TH10	AJ968933	21	9	0.84	0.87	0.035^{NS}	0.065	15.05
2TB07	AJ968934	13	4	0.66	0.77	0.136***	0.208	10.77
2TC03	AJ968935	17	8	0.66	0.85	0.225***	0.179	11.66
3TA11	AJ968936	15	5	0.60	0.79	0.241***	0.232	12.86
3TB02	AJ968937	19	8	0.69	0.81	0.146***	0.107	14.36
3TD09	AJ968938	9	3	0.30	0.40	0.256***	0.091	6.23
3TE06	AJ968939	12	2	0.31	0.48	0.353***	0.286	10.25
3TF01	AJ968940	18	7	0.72	0.81	0.108***	0.130	13.69
4TD12	AJ968941	20	13	0.55	0.74	0.258***	0.185	11.38
4TF02	AJ968942	14	5	0.54	0.60	0.100***	0.163	11.13
4TH09	AJ968943	13	8	0.37	0.53	0.305***	0.248	8.37

Table 5 Summary of intra-population genetic diversity at 15 microsatellite loci for 22 exotic populations and four genetic clusters of *Tectona grandis*

Population	Sample size	A	R	H_{nb}	H_0
Bangsri Pati	8	3.80	3.34	0.56	0.51
Temandsang	9	3.93	3.25	0.52	0.44
Ngliron Ngliron	9	3.80	3.17	0.51	0.43
Nanas Blora	10	4.27	3.45	0.59	0.55
Jema	9	4.20	3.51	0.58	0.53
Landrace G1	9	4.20	3.49	0.59	0.53
Landrace G3	9	4.27	3.52	0.58	0.57
Landrace G4	8	4.00	3.40	0.54	0.53
Djibelor	6	4.20	3.95	0.65	0.58
Kalounayes	5	3.60	3.60	0.58	0.59
Bamoro A29	5	4.33	4.33	0.68	0.75
Bamoro A20	5	4.13	4.13	0.70	0.67
Bouaké	7	4.60	4.05	0.65	0.72
Landrace TB 73	6	3.80	3.57	0.61	0.67
Tové	5	3.93	3.93	0.65	0.64
Djigbé	5	4.47	4.47	0.71	0.72
Toffo Lama	6	4.33	4.01	0.62	0.66
Bambuku 7	5	3.00	3.00	0.49	0.52
Bambuku 3067	5	4.07	4.07	0.63	0.72
Kihuwi	10	5.13	3.85	0.63	0.61
Mtibwa	10	5.40	3.96	0.61	0.57
Bigwa	14	5.93	3.95	0.60	0.60
South India	71	11.47	11.26	0.76	0.74
North Thailand	64	5.07	5.07	0.41	0.38
Central Laos Indonesia, Ghana	92	7.40	6.93	0.56	0.47
North India, Africa	104	11.20	10.26	0.65	0.63

Results are given for each exotic population and each cluster with the natural range defined in Fofana et al. (2009)

A mean number of alleles per population or cluster, R corrected allelic richness, H_{nb} the expected unbiased heterozygosity i.e., expected heterozygosity corrected for the small sample size (Nei 1978), H_0 the observed heterozygosity

number of loci (Piry et al. 2004). For teak, the calculations were mainly made with all the 15 loci studied and ten individuals were compared with only 14 loci in our study. At least three potentially misleading results may arise when applying the method used in this study (Rannala and Mountain 1997). Firstly, the appropriate reference populations were not included in the analysis. This is the case of the Myanmar populations which could not be included in our reference populations. If provenances from Myanmar had been widely introduced in Africa and/or if these provenances were genetically very different from those in our study, then we should have seen a new genetic cluster appear. In addition, it is difficult to accept that provenances that are located on each side of the border between Myanmar and Thailand are very different and could form a new genetic cluster. As a new genetic cluster did not appear, we can conclude either that introductions from Myanmar into Africa remained rare or that these Myanmar provenances were not different from the four genetic clusters found by Fofana et al. (2009). Secondly, an individual might incorrectly appear to have originated from a particular population other than its actual population of origin. This might be due to similarities in allele frequencies due to long-term gene flow, or to genotyping errors, or to missing data. The allele frequencies of the natural populations of teak, which were very different among the four clusters and the small number of genotyping errors in this study could not explain these discrepancies. Gene flows between African or Indonesian populations, which were widely distributed for more than one century and which could recombine could also be a source of clustering error. Thirdly, the fact that many pairwise comparisons between populations were performed for each of a large number of individuals means that some individuals will appear to be immigrants by chance. This can be corrected for by using smaller values for α (Rannala and Mountain 1997).

In a simulation study comparing different methods of assignment, the Bayesian criterion of Rannala and Mountain (1997) was always the most efficient (Cornuet et al. 1999). These authors also showed that a perfect assignment can be obtained with F_{ST} values as moderate as 0.1. Whenever F_{ST} is high (e.g., F_{ST}=0.22) many combinations from (10 loci×24 individuals per population) to (48 loci× only five individuals per population) all provide a 100% correct assignment. Our study on teak is situated among

high values of F_{ST} ($0.1488<F_{ST}<0.2762$) and an average number of loci. With these conditions, it is possible to conclude that the assignment method using the Bayesian criterion of Rannala and Mountain (1997) gives the best results. The Monte Carlo re-sampling algorithms may be useful for identifying populations whose origin are poorly known. For teak, this method gave very precise information on the most probable origin of the different introductions in Africa and Indonesia.

The number of clusters in which the individuals could be grouped was estimated using simulations with the Structure software (Evanno et al. 2005). The only situation in which the software failed to detect the real K was the partial sampling of 20 individuals and five microsatellite markers in an island model. This study of different genetic models associated with reductions in sampling or genotyping appeared to confirm the robustness of the results obtained in teak.

On the other hand, Bayesian or maximum-likelihood estimations based on Markov chain Monte Carlo (MCMC) methods often require several consecutive runs to check that the chains have converged and that parameter space has been correctly explored (Excoffier and Heckel 2006). The Structure programme tried to define the number of sub-populations from which the sampled individuals were drawn. The software performed this allocation sequentially for different numbers of clusters, and then flagged the number of clusters with the highest likelihood, which might not always be optimal (Evanno et al. 2005). For the teak data, we ran a large number of MCMC repetitions (400,000), with several repetitions of the same analyses (10) after a long burnin period (40,000) and for two different methods (admixture model, no admixture model) with two options (correlated alleles, or independent alleles). By using the whole SSR dataset, the strong specificity and variability of teak trees from South India disturbed the identification of the four genetic clusters with both models used. Nevertheless, as the number of sub-populations was exactly the same in all the other analyses, in these conditions the distribution of the teak samples in four genetic clusters could be considered as accurate and robust. What is more, our results correspond perfectly with the results of Fofana et al. (2009).

With the results of these genetic structure analyses, the four genetic clusters obtained by Fofana et al. (2009) were analysed again. The results obtained with the two methods were identical (Table 6). The 22 exotic provenances tested ended up in one of the four clusters. All the African provenances: Senegal (2), Côte d'Ivoire (4), Benin (2), Togo (1), Cameroon (2) and Tanzania (3) were in the cluster from North India. The provenances of Indonesia (4) and Ghana (4) were grouped with the cluster from Central Laos (Fofana et al. 2009). It should be noted that neither provenance nor introduced teak was related to the cluster from South India and it is consequently possible to assert that the teak introduced in Africa and Indonesia did not come from South India. The African sample of teak is believed to have originated from North India, while Indonesian and Ghanaian teak were introduced from the natural range of teak in East Thailand or Central Laos.

Origin of exotic teak

Among reports in the literature on the introduction of teak in Africa, only the paper by Madoffe and Maghembe (1988) on the first introduction of teak in Tanzania by the German administration in 1898 from seeds from Calcutta was confirmed by our results. At that time, Germany occupied Tanzania, Togo, Cameroon and Namibia. Teak was then introduced in Togo, because the German administration was interested in problems arising from savannah and soil degradation. Forest species were planted at Nuatja, Sokodé-Bassari and Galangashi. After the First World War, Togo and Cameroon were integrated into French administered Africa, and the Togolese teak plantations were used as the origin of seeds to test the species in Benin, Côte d'Ivoire and subsequently in Senegal. Results obtained with SSR molecular markers on the genetic diversity structure and the individual-based assignment analyses to their clusters of origin confirmed this expansion of the teak plantations from German-speaking to French-speaking Africa with seeds originating from North India first planted in Tanzania.

The results of the analysis of the genetic structure of teak populations using the clustering approach and individual-based assignment analyses revealed a significant difference between the teak introduced in countries that were formerly under German and/then French administration and the teaks introduced in countries that were formerly under English administration. For Benin, Cameroon, Côte d'Ivoire, Tanzania, Togo and Senegal, 94.7%, 3.2%, and 2.1% of individual teak could be assigned respectively to North India, Thailand and Central Laos. For Indonesia and Ghana, respectively 95.8% and 4.2% of the teak could be assigned to Central Laos and North India.

In Ghana, four different provenances were sampled (Jema, SG1, SG3 and SG4), but using molecular markers, these four provenances appeared to correspond to only one sampled population with close genetic distances (Fig. 4). The difference observed between Ghanaian provenances and all the other African populations could be due to the specific choice of the stand where the seeds were harvested by Danida, possibly from Indonesia or Laos.

Concerning Indonesia, results using the SSR markers made it possible to reject the speculative theory of Altona (1922) which strongly presumed teak was imported from

Table 6 Comparison between the results obtained with the method of assignment of the unknown individuals to the reference genetic clusters of the natural range (GeneClass2) and the clustering based on all the teak trees studied (Structure)

Provenance names or local names	*N*	Method	Cluster 1: South India (%)	Cluster 2: North India (%)	Cluster 3: Thailand (%)	Cluster 4: Central Laos (%)
Djigbé; Toffo Lama; Bambuku 7; Bambuku 3067; Bamoro A20; Bamoro A29; Bouaké; TB 73; Bigwa; Kihuwi; Mtibwa; Tové; Djibelor; Kalounayes.	94	Assignment	3	79	17	1
Djigbé; Toffo Lama; Bambuku 7; Bambuku 3067; Bamoro A20; Bamoro A29; Bouaké; TB 73; Bigwa; Kihuwi; Mtibwa; Tové; Djibelor; Kalounayes.	94	Clustering	0	96	2	2
Bangsri Pati ; Nanas Blora ; Ngliron; Temandsang ; Jema; SG01; SG03; SG04.	71	Assignment	0	12	3	85
Bangsri Pati ; Nanas Blora ; Ngliron; Temandsang ; Jema; SG01; SG03; SG04.	71	Clustering	0	4	0	96

South India by Brahmans or Buddhism monks. Indeed the introduction of teak in Java could have taken place between the fourteenth and the sixteenth centuries during which there was intensive interaction between India and Indonesia. However, a remark on the importing of teak to Java which was primarily planted around temples (as still observed in North Thailand around Chiang Mai), could be closer to our results (Simatupang 2000). We tested a large number of provenances from this area (Ban Cham Pui, Mae Huat, Huoi Na Soon, Ban Pha Lai and Pong Salee), and our results showed that it was probably from a smaller area with a narrow genetic base that the first teak was introduced in Java. Several centuries of teak introduction in Indonesia, during which there were an increase in the number of plantations and many genetic mixings, made it possible for SSR loci to acquire a high level of variability (Table 5).

The supply of seed is a limiting factor for planting and reduces the quality of the plantations, especially in countries where teak is grown as an introduced tree species. In Africa, for the first teak plantations, it was probably easier to use seeds harvested in stands that showed satisfactory adaptation to local conditions. Two complementary analyses of teak stands in Togo confirmed this hypothesis. The technological properties of teak trees (Kokutse et al. 2004) and the genetic diversity with molecular markers (Logossa 2006) were analysed in five very contrasted ecological zones in Togo. Results showed that the percentage of heartwood differed significantly in trees depending on the ecological zone (Kokutse et al. 2004). The genetic differentiation among the six populations studied was estimated at 2.3% and the teak stands of these five ecologically contrasted zones can be considered as a single genetic population. Moreover, these Togolese provenances are close to the North Indian cluster defined by Fofana et al. (2009). The differences in technological properties of teak in Togo are thus only due to the environmental effect.

Finally it should be noted that in Africa, even 100 years after teak was originally introduced, there are still few successive generations and consequently little genetic admixture. For example, in Côte d'Ivoire, seeds harvested in the former Bamoro stands (1929) were first generation and were used for the Matiemba plantation (1964) and by Sodefor for their mechanised plantations until the 1990s. Seeds harvested in the Matiemba stand were second generation and were also used by Sodefor until the 1990s. The programme for the genetic improvement of teak has been producing seeds since 1995, but the quantities are not sufficient to ensure extensive man-made forests. The fairly low multiplication factor of seed orchard trees could be attributed to the slow growth of trees, the duration of the breeding programmes, the low seed yield per tree and the low and sporadic germination behaviour of teak seed (Koskela et al. 2010). Contrary to common belief, in Africa there was neither very diversified use of seed stands nor many generations of genetic admixture. Consequently, the genetic admixture between the various provenances—which may have been introduced—remained very limited. This low level of genetic admixture between generations also explains the results obtained with SSR markers in our study.

Using the SSR molecular markers enabled us to trace the origin of teak imported into Africa during the nineteenth century and into Indonesia during the period of Hindunization. It is possible to assert that the teak introductions in Africa and Indonesia did not come from South India. Today, the South India provenances of teak are found mainly in provenance trials or in conservation areas and were introduced in Africa and Indonesia more recently to increase the teak genetic pool for the genetic improvement of the species.

Relationship between teak straightness and genetic clusters

The general form of African plantation teak also appears to confirm the results we obtained with the SSR molecular

markers. The bad overall shape of the trees (stem form, axis persistence, branch size, buttressing) of Jema, SG01, SG03, SG04, Bouaké, Tové, Bambuku 7, Bambuku 3067, TB73, Kihuwi, Bamoro A29, Bamoro A20, Toffo Lama, Kalounayes, Djigbé provenances from Africa has been confirmed in numerous comparative trials across the world (Dupuy and Verhaegen 1993; Kaosa-ard 1999; Keiding et al. 1986; Kjaer and Lauridsen 1996; Kjaer et al. 1995; Kjaer et al. 1996; Madoffe and Maghembe 1988; Pedersen et al. 2007; Rao et al. 2001). Stem quality, which includes straightness, clear bole, persistence of stem axis, mode of branching, branch size and branch diameter and flowering, is strongly controlled by the provenance. In the teak natural range, bad stem quality was found in provenances from Maharashtra (Kaosa-ard 1999), Hoshangabad (Pedersen et al. 2007), Chanda (Madoffe and Maghembe 1988; Pedersen et al. 2007), Berbera, Purunakote, Bakbahal (Kjaer et al. 1996), Jhirpa, Murda and Bakbahal (Rao et al. 2001). It is interesting to note that all these populations are located to the north of the limit defined by Fofana et al. (2009), which separated the genetic variability of the teak in its natural range in India. In this paper, we show that the introductions of teak in Africa probably originated from North India. It is thus possible to propose the hypothesis that the bad general shape of trees in Africa could be connected to their genetic origin.

The molecular markers revealed the strong genetic relationships between the landraces studied, but markers are neutral with respect to the evolution of the introduced provenances which may have undergone mass selection. The case of the provenances from Kihuwi, Bigwa and Mtibwa in Tanzania is a good illustration of the last point. Genetic distances measured with SSR markers (Fig. 4) confirmed that Mtibwa, and Bigwa are first and second generations of Kihuwi (Pedersen et al. 2007). For the main tree shape traits, the Kihuwi provenance is comparable to the other African provenances. On the other hand, the mass selection of seed trees in the seed-collection stand of Kihuwi and Mtibwa significantly improved the stem form of the trees belonging to the Mtibwa and Bigwa provenances (Madoffe and Maghembe 1988; Pedersen et al. 2007). The best behaviour in trials of these last two provenances gave the impression that Tanzanian provenances were of different genetic origin than the other African provenances.

The genetic cluster including Laos, Ghana and Indonesia is represented by provenances from Savannakhet I, Savannakhet II, Vientiane Town, Jema, SG01, SG03, SG04, Ngliron Ngliron, Bangsri Pati, Temandsang, Nanas Blora, Beran and Lembaga. An analysis of the literature on comparative provenance trials (Kaosa-ard 1999; Kjaer and Lauridsen 1996; Kjaer et al. 1996; Pedersen et al. 2007) showed that the provenances from Ghana (Jema, SG01, SG03, SG04) had a bad stem form and axis persistence in the international trials in Ghana (Vigneron, personal communication), Puerto Rico and Brazil. The Indonesian provenances had bad stem form and axis persistence in seven international trials. The provenances from Indonesia which had the worst stem form are the provenances from Ngliron Ngliron, Temandsang and Nanas Blora. The provenance from Bangsri Pati may have an average or a good stem form. The provenances from Savannakhet I, Savannakhet II and Vientiane Town had a bad tree shape in trials in Ghana and Brazil (Kjaer et al. 1995). It is important to note that the Laotian provenances analysed in this study were divided into the two genetic clusters defined by Fofana et al. (2009). The provenances in the cluster from Central Laos (Savannakhet I, Savannakhet II, Vientiane Town) had a bad stem form, in contrast to the provenances from Laos (Pakse and Pak Lai) allocated to the Thailand cluster which had a better stem form, but which also varied. The genetic cluster including Laos, Ghana, and Indonesia is also linked with the bad general shape of teak trees. Some provenances in this cluster were able to diversify under the effect of anthropological pressure or natural selection resulting in populations with a better general tree shape.

To close this analysis of international trials, it is worth noting that the provenances from Ban Cham Pui, Pak Lay, Ban Pha Lai, Mae Huat, Ban Mae Pam, Ban Maekut and Ban Huey that were the straightest, are part of the genetic cluster in Thailand defined by Fofana et al. (2009).

The technological properties of teak should also be analysed because it has been shown that the quality of wood varies with the provenance. In Madhya Pradesh, the timber is golden yellow and the heartwood blends into the sapwood (Bedell 1989). Perfect wood from African plantations (Togo, Benin, Côte d'Ivoire and Nigeria) is golden yellow to yellow-brown (Dupuy and Verhaegen 1993).

Population genetic parameters

It should be noted that the genetic parameters of both clusters from South India and Africa associated with North India displayed the strongest diversity and their genetic parameters were similar with respect to a number of alleles of 11.5 and 11.2, a corrected allelic richness of 11.3 and 10.3 and an expected heterozygosity of 0.76 and 0.65. The genetic diversity found in Benin, Cameroon, Côte d'Ivoire, Tanzania, Togo and Senegal remained high and corresponded well to the genetic diversity of the natural range in North India. On the other hand, the cluster comprising Indonesia, Ghana and Central Laos presented a number of alleles, a corrected allelic richness, and an expected heterozygosity lower and closer to the cluster from Thailand. Fofana et al. (2009) showed that genetic diversity in

Central Laos is very low (Na=3.8, R=3.06 and He=0.22) compared with South India where genetic diversity is maximum (Na=11.5, R=6.63 and He=0.76). These characteristics are of major importance for any programme for the conservation and (or) genetic improvement of the species. The in situ conservation of teak germplasm is discussed in the paper of Fofana et al. (2009). For genetic improvement programmes, the genetic admixture of various genetic clusters should at best help investigate the genetic variability of teak in order to benefit from different teak traits: e.g., the stem form, growth, technological wood properties and pest resistances of *ex situ* germplasm.

Conclusions

In this study on teak, the methods used were found to be very effective and it was surprising to note that more than one century after the introduction and extension of teak in Africa and several regions in Indonesia, and after a lot of intercontinental seed exchanges and several generations, it was still feasible to find the origin of introduced teak populations in Africa and Indonesia. It should be noted that the African populations corresponded perfectly with the Purunakote population, which is rather close to Calcutta; they were thus combined in order to constitute a reference genetic base of the four genetic clusters. These methods could now be used to test other as yet unknown teak provenances or to compare other countries in which teak was introduced. It is however still necessary to complete our data with provenances from Myanmar and also to increase the number of populations in North India to consolidate our reference data base.

Acknowledgements We would like to thank the European Commission project "Waft" (Inco-Dev ICA4-CT-2001-10090) that supported this research in Ghana, in Côte d'Ivoire and at CIRAD Bios (French Agricultural Research Centre for International Development) UR 39 laboratory. Many thanks to the SODEFOR (Forest Development Society BP 3770 Abidjan 01 Côte d'Ivoire) and FORIG (Forestry Research Institute of Ghana UST P.O Box 63 Kumasi Ghana), which supplied the teak plant material. Finally we wish to thank Claire Billot from CIRAD Bios for having allowed and facilitated access to the sequencer of the molecular biology laboratory, as well as Ronan Rivallan for technical support.

References

Altona T (1922) Teak and Hindoos. Tectona 15:457–506

Bedell PE (1989) Preliminary observations on variability of teak in India. Indian For 115:72–81

Begué L (1957) Problèmes forestiers de l'Indonésie. Bois For Trop 56:3–16

Behaghel I (1999) The state of teak (*Tectona grandis*) plantations worldwide. Bois For Trop 262:6–18

Belkhir K, Borsa P, Chikhi L, Raufaste N, Bonhomme F (1996–2004) GENETIX 4.05, logiciel sous Windows TM pour la génétique des populations. Laboratoire Génome, Populations, Interactions, CNRS UMR 5171, Université de Montpellier II, Montpellier (France) Available at http://www.genetix.univ-montp2.fr/genetix/genetix.htm

Bhat KM, Ma HO (2004) Teak growers unite! ITTO Trop For Update 14:3–5

Cavalli-Sforza LL, Edwards AWF (1967) Phylogenetic analysis: models and estimation procedures. Evol 32:550–570

Chollet MA (1958) Le teck au Togo. Bois For Trop 49:9–18

Cornuet J-M, Piry S, Luikart G, Estoup A, Solignac M (1999) New methods employing multilocus genotypes to select or exclude populations as origins of individuals. Genet 153:1989–2000

Dieringer D, Schlötterer C (2003) Microsatellite Analyser (MSA): a platform independent analysis tool for large microsatellite data sets. Mol Ecol Notes 3:167–169

Dupuy B, Verhaegen D (1993) Plantation-grown teak (*Tectona grandis*) in Côte d'Ivoire. Bois For Trop 235:9–24

El Mousadik A, Petit RJ (1996) High level of genetic differentiation for allelic richness among populations of the argan tree [*Argania spinosa* (L.) Skeels] endemic to Morocco. Theor Appl Genet 92:832–839

Evanno G, Regnaut S, Goudet J (2005) Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. Mol Ecol 14:2611–2620

Excoffier L, Heckel G (2006) Computer programs for population genetics data analysis: a survival guide. Nat Rev Genet 7:745–758

Excoffier L, Laval G, Schneider S (2005) Arlequin ver. 3.0: an integrated software package for population genetics data analysis. Evol Bioinformatics Online 1:47–50

Excoffier L, Laval G, Schneider S (2007) Arlequin ver 3.11: an integrated software package for population genetics data analysis. Computational and Molecular Population Genetics Lab, Institute of Zoology University of Berne Switzerland, distributed by the authors at (http://lgb.unige.ch/arlequin/).

FAO (2001) Global forest resources assessment 2000 Main report. FAO For Pap 140:1–511

Felsenstein J (2005) PHYLIP (Phylogeny Inference Package) version 3.6. Department of Genome Sciences, University of Washington, Seattle, distributed by the authors at (http://evolution.gs.washington.edu/phylip.html).

Fofana IJ, Ofori D, Poitel M, Verhaegen D (2009) Diversity and genetic structure of teak (*Tectona grandis* L.f) in its natural range using DNA microsatellite markers. New For 37:175–195

Goudet J (2001) FSTAT, a program to estimate and test gene diversities and fixation indices. Version 2.9.3. Department of Ecology & Evolution, Biology Building, UNIL, CH-1015 Lausanne, Switzerland, distributed by the authors at (http://www2.unil.ch/popgen/softwares/fstat.htm).

Hoffman JI, Amos W (2005) Microsatellite genotyping errors: detection approaches, common sources and consequences for paternal exclusion. Mol Ecol 14:599–612

Hurlbert SH (1971) The nonconcept of species diversity: a critique and alternative parameters. Ecol 52:577–586

Kaosa-ard A (1999) Gains from provenance selection. Proc of intern semin on site technol and product of teak plant Chiang Mai, Thailand, 26–29 January:191–207

Keiding H, Wellendorf H, Lauridsen EB (1986) Evaluation of an international series of teak provenance trials. Danida For Seed Cent:1–81

Kertadikara AWS, Prat D (1995) Isozyme variation among teak (*Tectona grandis* L.f.) provenances. Theor Appl Genet 90:803–810

Kjaer ED, Lauridsen EB (1996) Results from a second evaluation of DFSC coordinated teak (*Tectona grandis*) provenance trials: has

new information been obtained? Proceedings of the conference: tree improvement for sustainable tropical forestry Caloundra, Queensland, Australia, 27 October–1 November, pp 154–157

Kjaer ED, Siegismund HR (1996) Allozyme diversity in two Tanzanian and two Nicaraguan landraces of teak (*Tectona grandis* L.). For Genet 3:45–52

Kjaer ED, Lauridsen EB, Wellendorf H (1995) Second evaluation of an international series of Teak provenance trials. Danida Forest Seed Center

Kjaer ED, Siegismund HR, Suangtho V (1996) A multivariate study on genetic variation in teak (*Tectona grandis* (L.)). Silvae Genet 45:361–368

Kokutse AD (2002) Analyse de la qualité du bois de teck (*Tectona grandis* L.f) en plantation au Togo: formation du bois de coeur, propriétés mécaniques et durabilité. Thèse de l'école doctorale des sciences physiques et de l'ingénieur—Bordeaux N°d'ordre 2638:1–142

Kokutse AD, Baillères H, Stokes A, Kokou K (2004) Proportion and quality of heartwood in Togolese teak (*Tectona grandis* L.f.). For Ecol Manag 189:37–48

Koskela J, Vinceti B, Dvorak W, Bush D, Dawson I, Loo J, Kjaer ED, Navarro C, Padolina C, Bordács S, Jamnadass R, Graudal L, Ramamonjisoa L (2010) The use and exchange of forest genetic resources for food and agriculture. FAO Commission on Genetic Resources for Food and Agriculture Background study paper N° 44:1–67

Logossa ZA (2006) Caractérisation génétique des tecks (*Tectona grandis* L.f.) provenant de plantations du Togo. Sciences Technologiques. Bordeaux 1, p 29

Madoffe SS, Maghembe JA (1988) Performance of teak (*Tectona grandis* L.f.) provenances seventeen years after planting at Longuza, Tanzania. Silvae Genet 37:175–178

Nei M (1978) Estimation of average heterozygosity and genetic distance from a small number of individuals. Genet 89:583–590

Nicodemus A, Nagarajan B, Narayanan C, Varghese M, Subramanian K (2003) RAPD variation in Indian teak populations and its implications for breeding and conservation. Proceedings of the international conference on quality timber products of teak from sustainable forest management Peechi, India, 2–5 December, pp 321–330

Paetkau D, Waits LP, Clarkson PL, Craigheadg L, Strobe C (1997) An empirical evaluation of genetic distance statistics using microsatellite data from bear (Ursidae) populations. Genet 147:1943–1957

Paetkau D, Slade R, Burden M, Estoup A (2004) Genetic assignment methods for the direct, real-time estimation of migration rate: a simulation-based exploration of accuracy and power. Mol Ecol 13:55–65

Pandey D, Brown C (2000) Le teck dans le monde. Unasylva 201:3–13

Pedersen AP, Hansen JK, Mitika JM, Msangi TH (2007) Growth, stem quality and age–age correlations in a Teak provenance trial in Tanzania. Silvae Genet 56:142–148

Perrier X, Jacquemoud-Collet JP (2006) DARwin software. Cirad-Bios UPR Genetic Improvement of Vegetatively Propagated Crops, distributed by the authors at (http://darwin.cirad.fr/darwin).

Piry S, Alapetite A, Cornuet J-M, Paetkau D, Baudouin L, Estoup A (2004) GeneClass2: a software for genetic assignment and first-generation migrant detection. J Hered 95:536–539

Pompanon F, Bonin A, Bellemain E, Taberlet P (2005) Genotyping errors: causes, consequences and solutions. Nat Rev Genet 6:847–858

Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. Genet 155:945–959

Rance W, Monteuuis O (2004) Teak in Tanzania I. Overview of the context. Bois For Trop 279:5–10

Rannala B, Mountain JL (1997) Detecting immigration by using multilocus genotypes. Proc Natl Acad Sci USA 94:9197–9201

Rao PS, Venkaiah K, Murali V, Murti SSN, Sattar SA (2001) Evaluation of international teak provenance trial plot in North East Andhra Pradesh. Indian For 127:415–422

Rice WR (1989) Analysing tables of statistical tests. Evol 43:223–225

Saitou N, Nei M (1987) The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol Biol Evol 4:406–425

Shrestha MK, Volkaert H, Van der Straeten D (2005) Assessment of genetic diversity in *Tectona grandis* using amplified fragment length polymorphism markers. Can J For Res 35:1017–1022

Simatupang MH (2000) Some notes on the origin and establishment of teak forest (*Tectona grandis* L.F.) in Java, Indonesia. In: Hardiyanto EB (ed) Proceedings of third regional seminar on teak Potentials and opportunities in marketing and trade of plantation teak: Challenge for the new millennium, Yogyakarta, Indonesia, July 31–August 4, pp 91–98

Steffens DL, Sutter SL, Roemer SC (1993) An alternate universal forward primer for improved automated DNA sequencing of M13. Biotech 15:580–582

Taberlet P, Griffin S, Goossens B, Questiau S, Manceau V, Escaravage N, Waits LP, Bouvet J (1996) Reliable genotyping of samples with very low DNA quantities using PCR. Nucleic Acids Res 24:3189–3194

Tariel J (1966) Le Teck en Côte-d'Ivoire. Bois For Trop 107:27–47

Verhaegen D, Ofori D, Fofana I, Poitel M, Vaillant A (2005) Development and characterization of microsatellite markers in *Tectona grandis* (Linn. f). Mol Ecol Notes 5:945–947

Wattier R, Engel CR, Saumitou-Laprade P, Valero M (1998) Short allele dominance as a source of heterozygote deficiency at microsatellite loci: experimental evidence at the dinucleotide locus Gv1CT in *Gracilaria gracilis* (Rhodophyta). Mol Ecol 7:1569–1573

White KJ (1991) Teak: some aspects of research and development. FAO Regional Office for Asia and the Pacific 1991, pp 1–53

Wood PJ (1967) Teak planting in Tanzania. World symposium on man-made forests and their industrial importance. FAO, Canberra, Australia 14–25 April 1967, pp 1631–1643