COMPARATIVE GENETIC DIVERSITY STUDIES OF *SHOREA CURTISII* (DIPTEROCARPACEAE): AN ASSESSMENT USING SSR AND DAMD MARKERS

W. S. Ho[1,*], R. Wickneswari[2], M. C. Mahani[2] & M. N. Shukor[2]

[1]Faculty of Resource Science and Technology, Universiti Malaysia Sarawak, 94300 Kota Samarahan, Sarawak, Malaysia
[2]School of Environmental and Natural Resource, Faculty of Science and Technology, Universiti Kebangsaan Malaysia, 43600 Bangi, Selangor Darul Ehsan, Malaysia

Received April 2004

HO, W. S., WICKNESWARI, R., MAHANI, M. C. & SHUKOR, M. N. 2006. Comparative genetic diversity studies of *Shorea curtisii* (Dipterocarpaceae): an assessment using SSR and DAMD markers. Genetic diversity of *Shorea curtisii* from different age cohorts, namely, seedlings, saplings and adult trees were determined using six SSR loci and 33 DAMD loci. To quantify genetic diversity in *S. curtisii* we used standard genetic diversity measures for SSR data, and both phenotypic and genotypic methods with null-allele frequency corrected for deviation from Hardy-Weinberg equilibrium (HWE) with SSR markers for DAMD data. Results showed that the genetic diversity measured using DAMD genotypic method was lower than those derived from SSR data based on the same set of samples. This suggests that DAMD allele frequencies corrected from HWE deviation using fixation index derived from SSR data may be underestimated. The genetic distance matrix generated from SSR data was significantly correlated with DAMD genotype data ($r = 0.990$, $p < 0.05$), indicating a similar genetic structure of *S. curtisii* being depicted by both marker types among the age cohorts tested. The relationship between sample size and genetic diversity measures demonstrated a threshold level, i.e. $n = 20$ and $n = 30$ for seedlings and saplings respectively, and $n = 15$ and $n = 20$ for adult trees revealed by SSR and DAMD markers respectively. Genetic diversity measures dropped drastically below these levels. These results further imply that a highly heterogeneous population was observed in *S. curtisii* from each age cohort. Collectively, both SSR and DAMD markers have good genome coverage in the *S. curtisii* genome.

Keywords: DNA analyses, polymorphisms, genetic distance, tropical timber species, hill dipterocarp forest

HO, W. S., WICKNESWARI, R., MAHANI, M. C. & SHUKOR, M. N. 2006. Kajian perbandingan kepelbagaian genetik *Shorea curtisii* (Dipterocarpaceae): satu penilaian menggunakan penanda SSR dan DAMD. Kepelbagaian genetik bagi *Shorea curtisii* daripada kohort umur berbeza, iaitu anak benih, anak pokok dan pokok matang telah ditentukan daripada enam lokus SSR dan 33 lokus DAMD. Bagi menganggar kepelbagaian genetik dalam *S. curtisii* kami menggunakan parameter kepelbagaian genetik piawai untuk data SSR. Bagi data DAMD pula kami menggunakan kedua-dua kaedah fenotip dan genotip dengan frekuensi alel nol dibetulkan penyisihannya daripada keseimbangan Hardy-Weinberg (HWE) dengan penanda SSR. Hasil kajian menunjukkan bahawa anggaran kepelbagaian genetik dengan kaedah genotip DAMD adalah lebih rendah berbanding dengan data SSR menggunakan set sampel yang sama. Ini mencadangkan frekuensi alel nol DAMD yang dibetulkan daripada sisihan HWE menggunakan indeks penetapan daripada data SSR mungkin di bawah jangkaan. Matriks jarak genetik yang dijana daripada data SSR berkorelasi secara signifikan dengan data genotip DAMD ($r = 0.990$, $p < 0.05$). Ini menunjukkan bahawa struktur genetik yang serupa bagi *S. curtisii* telah digambarkan oleh kedua-dua jenis penanda antara kohort umur yang diuji. Hubungan antara saiz sampel dengan kepelbagaian genetik memaparkan satu aras ambang, iaitu $n = 20$ dan $n = 30$ masing-masing bagi anak benih dan anak pokok, serta $n = 15$ dan $n = 20$ bagi pokok matang yang masing-masing dicerap oleh penanda SSR dan DAMD. Nilai kepelbagaian genetik menurun dengan mendadak di bawah nilai ambang. Hasil ini turut menunjukkan bahawa *S. curtisii* daripada setiap kohort umur mempunyai tahap keheterogenan yang tinggi. Secara keseluruhan, kedua-dua penanda SSR dan DAMD meliputi taburan yang luas dalam genom *S. curtisii*.

Introduction

Genetic diversity provides the template for adaptation and evolution of populations and species (Thomas *et al.* 1999). Therefore, preservation and maintenance of genetic diversity of all species are important for both short-term adaptations to environmental change and long-term impact on

**E-mail: wsho@frst.unimas.my*

species and communities (Templeton 1995). Loss of genetic diversity immediately after harvesting has been recognized as a potentially serious problem in commercially managed forest tree species (Buchert *et al.* 1997, Wickneswari *et al.* 1999, Rajora *et al.* 2000, Wickneswari & Lee 2001, Lee *et al.* 2002). Hence, assessment of genetic diversity has a high priority in developing successful management guidelines and effective conservation strategies or programmes at the species level and above. Indeed, conservation of genetic diversity may be one of the most important issues influencing future forestry practices (Namkoong 1992).

In recent years, there has been increasing interest in the use of DNA-based markers for a variety of applications in population genetics, conservation and tree improvement. Both SSRs (simple sequence repeats) and DAMD (direct amplification of minisatellite-region DNA) markers show much promise in this regard. DAMD has recently been used to identify wheat (Somers *et al.* 1996, Bebeli *et al.* 1997) and rice (Zhou *et al.* 1997), and quantify genetic variation in *Phaseolus vulgaris* (Métais *et al.* 2001), *Asimina triloba* (Rogstad *et al.* 1991) and *Plantago major* (Schaal *et al.* 1991). Meanwhile SSR markers have been used to quantify genetic diversity and examine population differentiation in agricultural crops (Morgante *et al.* 1994, Maughan *et al.* 1995) and trees, including radiata pine (Smith & Devey 1994), *Pithecellobium elegans* (Chase *et al.* 1996), *Quercus robur* (Lefort *et al.* 1998) and *Pinus strobus* (Rajora *et al.* 2000).

Since DAMD is a dominant marker type, the presence and absence of a band are defined as representing two alleles at a locus and, therefore, estimation of genetic diversity values must assume Hardy-Weinberg equilibrium(HWE). Consequently, these data do not allow for the estimation of allelic richness, effective number of alleles or fixation indices. However, with the advancement of PCR technology and statistical analysis methods, this problem can be minimized. For instance, Chong *et al.* (1994) analysed RAPD data genotypically by incorporating codominant allozyme data for the estimation of null-allele frequencies. SSRs, on the other hand, are codominant and tend to have multiple alleles per locus so that individuals can be identified as homozygotes or heterozygotes. The data can be used to compare observed and expected heterozygosities and other genetic diversity parameters. However, there are several shortcomings with SSR markers, i.e. the development costs of SSR markers are high if primers are not yet available, heterozygotes may be misclassified as homozygotes when null-alleles occur due to mutation in the primer annealing sites, underlying mutation model (IAM or SMM) is largely unknown, and homoplasy due to different forward and backward mutations may underestimate genetic divergence.

This study was undertaken to quantify genetic diversity of a tropical timber species, namely, *Shorea curtisii* (Dipterocarpaceae) using SSR and DAMD markers. This species is locally known as meranti seraya and is an emergent, abundant canopy tree species occurring in the ridges of hill dipterocarp forests (300–850 m) in Peninsular Malaysia. It is found in southern Thailand, Peninsular Malaysia, Sumatra and Borneo. *Shorea curtisii* flowers heavily at irregular intervals (3–5 years) after severe drought, in synchrony with mass flowering of other emergent tree species. The flowers are hermaphroditic, pollinated by insects such as thrips and meliponid bees, and produce single-seeded fruits; more than half of the mature seeds fall within 20 m of the parent tree. It is one of the main sources of the light hardwood of dark red meranti timber, which has already established a market both locally and overseas, and are used for furniture, high-class interior finishing, flooring, panelling, moulding and veneers. In this study we also examined the correlation coefficient of SSRs and DAMD based on genetic distance on the same set of samples of *S. curtisii* used to quantify genetic diversity. Both SSR and DAMD markers have advantages and limitations as genetic markers for assessing genetic diversity, genetic erosion resulting from logging practices and population structure. Therefore, combination of these two markers would yield complementary information to quantity genetic diversity.

Materials and methods

Plant materials and DNA isolation

Samples of *S. curtisii* at different age cohorts, namely, seedlings, saplings and adult trees were collected from Compartment 118 before logging in Ulu Sedili Forest Reserve, Johor, Malaysia (1° 55' N latitude and 103° 45' E longitude). Details of the sample size are listed in Table 1. The size of the compartment is about 164.31 ha with elevation ranging from 120 to 1000 m above sea level. The commonly occurring tree species in the compartment belong to the families Dipterocarpaceae, Myrtaceae, Annonaceae and Lauraceae. Total genomic DNA was isolated from fresh leaves or inner bark tissues (for mature trees) using a CTAB method modified from Murray and Thompson (1980) and Doyle and Doyle (1990).

SSR analysis

Six SSR loci specific for *S. curtisii* were used in this study for estimation of genetic diversity measures (Table 2). These SSR loci were developed by Ujino *et al.* (1998) using two different methods, i.e. colony hybridization and vectorette PCR methods. PCR amplification was carried out for 5 min at 95 °C, followed by 35 cycles of 45 s at 9 °C, 30 s at 51.6–56.8 °C, as appropriate for each pair of primers, and 45 s at 72 °C, with a final 5 min extension at 72 °C. The amplifications were done in a

Table 1 Number of samples collected and used for DNA analysis of *S. curtisii* from different age cohorts in Compartment 118, Ulu Sedili Forest Reserve

Diameter class	Sample size	
	Collected*	Analysed
Seedling	125 (17)	108
Sapling	66 (6)	60
Adult	32 (2)	30
Total	223 (25)	198

* Number in parentheses indicated wrongly identified samples and weak amplification which were not included in SSR and DAMD markers analysis.

Table 2 Selected SSR loci in *S. curtisii*, core sequence, allele number, length of PCR products and optimum primer annealing temperatures obtained in this study

Locus	Core sequence*	Number of alleles	PCR product length (bp)	Annealing temperature (°C)	Number of alleles#
Shc02	$(CT)_2CA(CT)_nGC(AT)_2$	4	137–153	55.2	2
Shc03	$(CT)_n$	4	124–140	56.8	3
Shc07	$(CT)_nCA(CT)_mCAC$ $CC(CTCA)_lCT(CA)_k$	12	137–217	56.2	11
Shc08	$(CT)_n$	14	239–353	52.5	-
Shc09	$(CT)_n$	11	189–241	51.6	9
Shc11	$(CT)_m(A/T)T(CT)_n$	5	142–174	52.5	4
	Total	50			29
	Average	8.33			4.83

number of alleles (Ujino *et al.* 1998)
* n, m, l and k show the repeated nucleotides (Ujino *et al.* 1998)

GeneAmp PCR System Model 9700 (Perkin-Elmer Co. Ltd.). The 25 µl reaction mixtures contained 1× PCR buffer (200 mM Tris-HCl at pH 8.4, 500 mM KC1), 1.5 mM $MgCl_2$, 0.2 mM of each dNTPs, 5 pmol of each primer, 20 ng of template DNA and 0.5 unit of Platinum Taq DNA polymerase (GIBCO BRL.). PCR amplification reaction was stopped with 5 µl loading dye (15% Ficoll, 0.25% Bromophenol blue and 0.25% Xylene cyanol FF). 10 µl of the PCR product was separated using 3.5% (w/v) Metaphor™ (FMC Bioproducts, Rockland, Maine) agarose gel in 1× TBE (Tris-Borate-EDTA). It should be noted that the 3% of Metaphor™ agarose gels run in 1× TBE have a degree of resolution equal to that of a 6% polyacrylamide gel for the targeted PCR product size range (FMC Bioproducts manual). The gels were running at approximately 80 V for three to four hours. After staining with Gelstar® nucleic acid gel stain (FMC Bioproducts, Rockland, Maine) for one hour, the gel was visualized under ultraviolet light. Polaroid MP-4 land camera with Polaroid 667 film was used to document the gels. The gels were scored for SSR genotypes of individual trees. The alleles at a SSR locus were identified by their molecular size and were designated alphabetically from larger to smaller molecular size.

DAMD analysis

PCR amplification reaction was performed in volumes of 25 µl containing 1× PCR buffer (200 mM Tris-HCl at pH 8.4, 500 mM KCl), 3.0 mM $MgCl_2$, 0.2 mM each of dNTP, 0.2 µM of M13 universal primer (5'-TTATGAAACGACGGCCAGT–3') (Chong *et al.* 1995), 1 unit of Taq DNA polymerase (GIBCO BRL.) and 15 ng of *S. curtisii* genomic DNA. The amplifications were done in a GeneAmp PCR System Model 9700 (Perkin-Elmer Co. Ltd.) thermocycler, programmed for 35 cycles of 1 min of denaturing at 94 °C, 1 min of annealing at 48 °C and 2 min of extension at 72 °C after an initial 2 min denaturing at 94 °C. The program was then followed by a final extension step at 72 °C for 10 min. After the final cycle, samples were held at 4 °C until the tubes were removed from the thermocycler. The DAMD products were mixed with 5 µl loading dye (15% Ficoll, 0.25% Bromophenol blue and 0.25% Xylene cyanol FF). The PCR products (10 µl) were then separated on a 1.2% (w/v) agarose gel for three to four hours at 80 V in 1× TBE buffer. The gel was stained with Gelstar® nucleic acid gel stain (FMC Bioproducts, Rockland, Maine) for an hour and then visualized under ultraviolet light. Polaroid 667 film was used to document the gel. The DAMD bands were named after the primer and a hyphenated number corresponding to the order of their migrations. Starting from the slowest to the fastest migrating fragment, the DAMD bands were designated as M13-01, M13-02, M13-03 and so on. The presence of a band was scored as 1 and absence was scored as 0, based on the following criteria: (1) each DAMD locus was assumed as independent or non-allelic, (2) there was no bias in scoring monomorphic fragments versus polymorphic fragments, and (3) only fragments in the range of 450 bp to 2.1 kb were considered in order to increase the data reliability.

SSR data analysis

Standard genetic diversity parameters (allele frequencies, allelic complements, percentage of polymorphic loci, average number of alleles per locus) and Nei's (1978) unbiased estimates of mean expected and observed heterozygosity from the SSR data were calculated using POPGENE Version 1.32 software. For each polymorphic locus in each population, deviation from HWE was examined by calculating Wright's (1978) fixation indices (F_{is}) using POPGENE Version 1.32 software. The population F_{is} values were estimated as the mean of F_{is} values over all loci. To test whether F_{is} values differed significantly from zero, two-tailed chi-square values were calculated using the formula of Li and Horvitz (1953) as

$$\chi^2 = F_{is}^2 \, N \, (n-1), \text{ for } \frac{[n(n+1)]}{2} \text{ degrees of freedom}$$

where

n = number of alleles and

N = number of samples examined.

To assess the dissimilarity among the age cohorts investigated, Nei's (1978) unbiased genetic distance (D) was calculated for each pairwise combination of age cohorts. Nei's (1978) genetic distance is a

parameter which is based on the normalized genetic identity (I) between two populations. The genetic identity indicates the probability that two alleles either sampled within a population or one taken from each population are identical.

DAMD data analysis

Genetic diversity parameters were calculated using two approaches: (1) phenotypic method using Shannon's index of diversity (Lewontin 1973), and (2) genotypic method with null-allele frequency corrected for deviation from HWE with SSR markers (Chong *et al.* 1994). For phenotypic approach, the presence (1) and absence (0) of fragments were treated as phenotypic characteristics. Shannon's index of diversity was used to quantify the level of polymorphism detected in *S. curtisii* at different age cohorts. Shannon's index of diversity was estimated using POPGENE Version 1.32 software with the following formula (Lewontin 1973):

$S = - \sum p_i \text{ in } p_i$,

where

p_i is the frequency of phenotype i.

For the genotypic approach, the primary assumption is that each polymorphic fragment will be considered as a locus, so that every locus has two alleles, identified by the presence or absence of the fragment. The generally neutral character of DAMD and SSR markers suggests deviation from HWE affecting both types of markers in a similar way. Hence, the null-allele frequency was estimated by using the previous estimate of the departure from HWE (F_{is}) with SSR data for the same populations. Then, these corrected allele frequencies were used for the estimation of the following genetic diversity parameters: percentage of polymorphic loci (P), expected heterozygosity (H_e) (Nei 1978), and unbiased genetic identity (D) (Nei 1978). Calculation of the corrected allele frequencies and genetic distance were made using POPGENE Version 1.32. The corrected allele frequency data set was then converted into BIOSYS-1 (Swofford & Selander 1981) type 3 format for the estimation of P and H_e.

Analysis of relationship between sample size and genetic diversity parameters, and correlation coefficient of SSR vs DAMD

The relationship between sample size and genetic diversity parameters was analysed for both SSR and DAMD markers based on the same set of samples collected from Compartment 118 before logging in Ulu Sedili Forest Reserve. Mean of genetic diversity parameters (H_e and n_e) were obtained from different sample size classes (5, 10, 15, 20, and so on). In order to minimize sampling error, analysis was performed based on two sets of random data for each sample size class, namely, samples number 1 to 5, 1 to 10, 1 to 15, … and so on in ascending and descending order respectively. Mean values of the genetic diversity parameters derived from two data sets for each sample size class was calculated. Besides these, the correlation coefficient of SSR and DAMD markers was performed using the Mantel matrix-correspondence test (Mantel 1967) based on the genetic distant-matrix generated from SSR data and DAMD genotypic data.

Results and discussion

SSR assay

All six SSR loci surveyed were polymorphic in all the age cohorts studied. Examples of the SSR profiles of the respective loci for *S. curtisii* are presented in Figure 1. The number of alleles per locus ranged from 4 to 14 and the average was 8.3 (Table 2). The mean expected heterozygosity (H_e) detected at different age cohorts ranged from 0.687 to 0.740 (Table 3). Obayashi *et al.* (2002) reported the H_e of *S. curtisii* adult trees from undisturbed and logged forests in Semangkok Forest Reserve, Peninsular Malaysia ranged from 0.769 to 0.794 using 3 SSR loci. In general, a high level of genetic diversity was observed in seedlings, saplings and adult trees for *S. curtisii* in Compartment 118. In fact, the levels of genetic diversity detected by SSR loci in this study were higher than those recorded using allozyme markers (Buchert *et al.* 1997, Wickneswari *et al.* 1999, Lee *et al.* 2001a) or by SSR markers as reported in Rajora *et al.* (2000).

Two SSR loci, Shc08 and Shc09, were highly polymorphic ($H_e > 0.80$), while the other two SSR loci, i.e. Shc02 and Shc03, each detecting four alleles showed moderate polymorphism in this study. Earlier studies have suggested that the longer dinucleotide loci are generally more variable and detect more alleles than those with a short repeat length (van de Ven & McNicol 1996). Thus, the results obtained in this study are concordant with these suggestions. In fact, locus Shc08 exhibited the highest genetic diversity measures amongst all the SSR loci examined for different age cohorts in this study (Table 3).

SSR loci such as Shc03, Shc08 and Shc09, allele or length polymorphism was found to depend mainly on differences in the number of CT repeats, e.g. in Shc08, $(CT)_{16}$ meant CT tandemly repeated 16 times. However, for Shc02 and Shc11, short imperfect repeats were found in the SSR core sequence. The polymorphism at these loci was found to depend on additions or deletions of one or more repeat units in the flanking region (Ujino *et al.* 1998). Locus Shc07 comprises short compound and imperfect repeats found in the SSR core sequence. Thus, the length polymorphism depended not only on differences in the repeat number of each repeat unit, but also on the combination of different repeat units. Such complexity in the core sequence may lead to errors in designating genotypes (Ujino *et al.* 1998) or in other words, a substantial amount of size homoplasy occurred (Estoup *et al.* 1995, Grimaldi & Crouau-Roy 1997). Two alleles are the same in size but they may not always be having the same sequences. However, SSR containing complex compound and interrupted repeats can be used for analysis of homoplasy (Estoup *et al.* 1995).

The amplified DAMD products with high fragment size may be not well resolved using 1.2% agarose gel. Therefore, to be conservative, only DAMD bands in the range of 450 bp to 2.1 kb were considered in this study for all age cohorts. Examples of DAMD profiles amplified by M13 universal primer for *S. curtisii* at different age cohorts are given in Figure 2. Heath *et al.* (1993) reported that minisatellite core sequences are larger than RAPD primers. Therefore, DAMD can be effectively carried out at relatively high stringencies and subsequently yield highly reproducible results. In addition, Skroch and Nienhuis (1995)

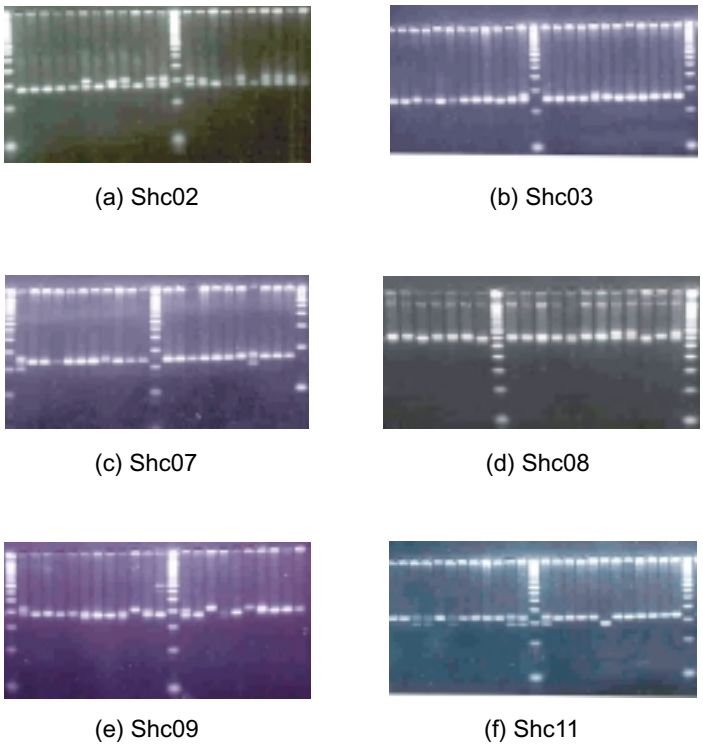


Figure 1 SSR profiles for *S. curtisii* (samples from Compartment 118, Ulu Sedili Forest Reserve) generated by six SSR loci. 50 bp DNA ladder was used as DNA size markers.

Table 3 Estimates of genetic diversity measures in *S. curtisii* from each age cohort in Compartment 118, Ulu Sedili Forest Reserve

Locus	Age cohorts					
	Seedling		Sapling		Adult	
	H_e	n_e	H_e	n_e	H_e	n_e
Shc02	0.615	2.577	0.626	2.638	0.606	2.472
Shc03	0.585	2.392	0.522	2.072	0.633	2.651
Shc07	0.669	2.988	0.595	2.441	0.814	5.014
Shc08	0.853	6.617	0.897	9.079	0.831	5.455
Shc09	0.819	5.399	0.813	5.147	0.801	4.670
Shc11	0.677	3.070	0.665	2.938	0.754	3.871
Mean	0.703	3.841	0.687	4.052	0.740	4.024
	(0.045)	(0.711)	(0.058)	(1.099)	(0.040)	(0.509)
F_{is}	0.422*		0.456*		0.459*	

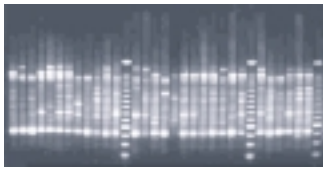
Mean = mean expected heterozygosity (H_e, Nei 1978) and number of effective alleles (n_e), per locus
F_{is} = Wright's (1978) fixation indices. * = Significantly different from 0.00 ($p < 0.05$)
Figures in parentheses are standard errors.

demonstrated that reproducibility is strongly dependent on uniformity of amplification conditions used between experiments. Hence, throughout the whole study, the same brand of chemical reagents as well as thermocycler was used for all PCR amplifications to increase data reliability.

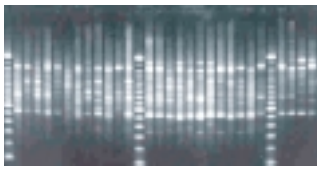
Estimation of genetic diversity at different age cohorts using DAMD markers based on 33 reproducible fragments is summarized in Table 4. For the phenotypic method, the number of polymorphic fragments scored at different age cohorts ranged from 25 (75.8%) for the adult trees to 30 (90.9%) for the seedlings. Similarly, adult trees exhibited relatively lower levels of Shannon's diversity (S = 0.325) whereas seedlings exhibited the highest (S = 0.390). Levels of genetic diversity obtained using genotypic method showed similar trends as estimated by phenotypic method. The mean expected heterozygosity (unbiased estimate, Nei 1978) obtained using DAMD genotypic method ranged from 0.223 to 0.285. These values were lower than those derived from SSR data (Table 3). It is argued that genetic diversity measures derived from DAMD allele frequencies corrected from HWE deviation using fixation index derived from SSR analysis may be underestimated. The genotypic method may not be as accurate as inferred from direct DAMD allele frequencies (e.g. derived from haploid tissues) or SSR loci. In fact, Szmidt *et al.* (1996) reported the deviation from HWE for RAPD loci to be generally higher than that for allozyme loci. Therefore, genetic diversity measured from RAPD allele frequencies corrected for HWE deviation using fixation index derived from allozyme analysis may be underestimated.

Detection of wrongly identified samples

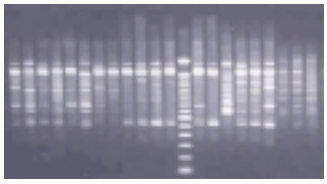
Several recent studies have demonstrated that DAMD yield individual or species-specific banding patterns and could be used for species or cultivar identification (Somers *et al.* 1996, Bebeli *et al.* 1997, Zhou *et al.* 1997). In this study, DAMD profiles generated by M13 universal primer consistently revealed distinctively different banding patterns for certain *S. curtisii* seedling samples. These samples were inferred as misidentified species during sample collection. Closely related dipterocarp tree species can share the same characters at seedling stage (Symington 1943). Such misidentified samples were excluded in the statistical analyses (Table 1). Figure 3 shows some examples of the banding patterns of the ambiguous samples of *S. curtisii*. Thus, in this regard, the M13 universal primer was indeed valuable.



(a) Seedling



(b) Sapling



(c) Adult trees

Figure 2 DAMD profiles in *S. curtisii* from different age cohorts generated by M13 universal primer. 100 bp DNA ladder was used as DNA size markers.

Table 4 Estimates of genetic diversity measures in *S. curtisii* from each age cohort in Compartment 118, Ulu Sedili Forest Reserve

Genetic diversity parameter	Age cohorts		
	Seedling	Sapling	Adult
Based on genotypic approach:			
Mean observed number of alleles, n_a	1.909	1.849	1.758
	(0.028)	(0.047)	(0.080)
Mean effective number of alleles, n_e	1.478	1.418	1.364
	(0.034)	(0.044)	(0.066)
Expected heterozygosity, H_e (Nei 1978)	0.285	0.257	0.223
	(0.017)	(0.022)	(0.035)
Percentage of polymorphic loci, P	90.9	84.9	75.8
Based on phenotypic approach:			
Shannon's index of diversity, S	0.390	0.354	0.325
	(0.022)	(0.029)	(0.059)

Figures in parentheses are standard errors.

Analysis of relationship between sample size and genetic diversity parameters

The effects of sample size on genetic diversity measures at different age cohorts of *S. curtisii* are presented in Figures 4–9. There were threshold levels where genetic diversity measures dropped drastically. For example, a threshold level of n = 20 for seedlings and saplings, and n = 15 for adult trees as revealed by SSR markers, and n = 30 for seedlings and saplings and n = 20 for adult trees, revealed by DAMD markers. These results suggest that a highly heterogeneous population was observed for each age cohort of *S. curtisii* in Compartment 118, Ulu Sedili Forest Reserve. The high heterogeneities of *S. curtisii* may have contributed to buffering genetic erosion later resulting from logging practices, as this species is one of the targets of selective logging in Malaysia. In comparison, Lee *et al.* (2002) estimated a threshold level of n = 15 in Serting Tambahan Forest Reserve and n = 10 in Pasoh Forest Reserve using RAPD markers for *Scaphium macropodum*. The lower threshold level for Pasoh Forest Reserve may indicate higher heterogeneity compared with the Serting Tambahan Forest Reserve.

Correlation coefficient of SSR vs DAMD

Coefficient of genetic distance matrix computed from SSR data and DAMD genotypic data (Nei 1978) are presented in Table 5. Mean genetic distance derived from SSR data (0.076) was slightly higher than that based on DAMD genotypic data (0.062). The genetic distance matrices obtained were compared with the Mantel statistics test for matrix comparisons. This test gave the product moment correlation (r) and a statistic test, z, to measure the degree of relationship between the two matrices. The distance-matrix generated from SSR data was significantly correlated with DAMD genotypic data (r = 0.990, p = 0.04). The high correlation coefficient obtained between the SSR and DAMD markers implied that a similar genetic structure of *S. curtisii* was depicted by both marker types among the age cohorts tested. In one study, Huang *et al.* (2000) reported that there was significant correlation in the distance-matrix generated by SSR and RAPD markers (r = 0.514, p = 0.012). Roa *et al.* (2000) also demonstrated high correlation between AFLP (dominant markers) and SSR markers (r = 0.81, probability of random z < obs. z: p = 1.00). However, Lee *et al.* (2001b) showed no significant correlation between similarity

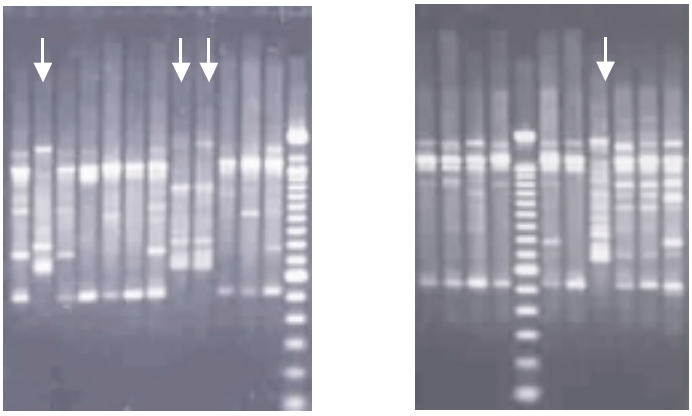


Figure 3 DAMD profile amplified by M13 universal primer of some *S. curtisii* samples from Compartment 118 in Ulu Sedili Forest Reserve. Distinctively different banding patterns of misidentified samples are indicated by arrow(s).

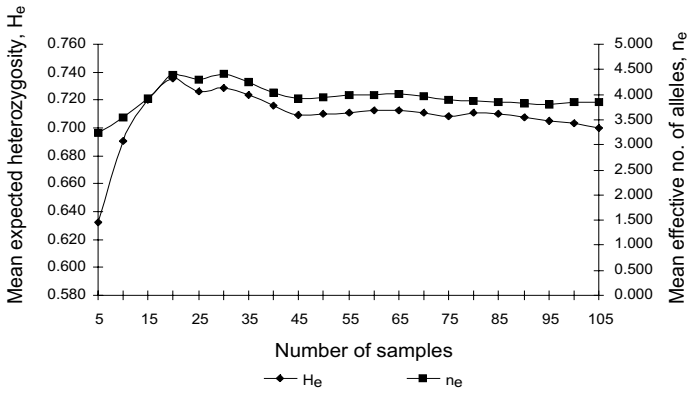


Figure 4 Relationship between measures of genetic diversity and number of samples of *S. curtisii* seedlings in Compartment 118, Ulu Sedili Forest Reserve revealed by SSR markers

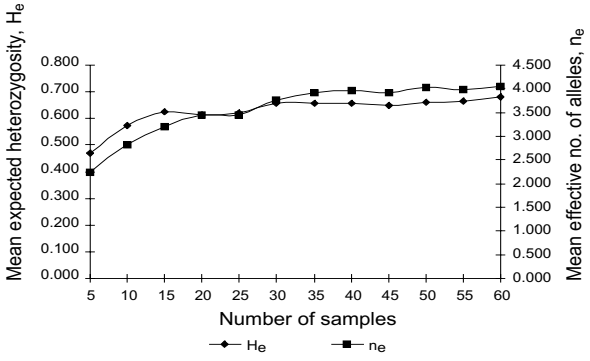


Figure 5 Relationship between measures of genetic diversity and number of samples of *S. curtisii* saplings in Compartment 118, Ulu Sedili Forest Reserve revealed by SSR markers

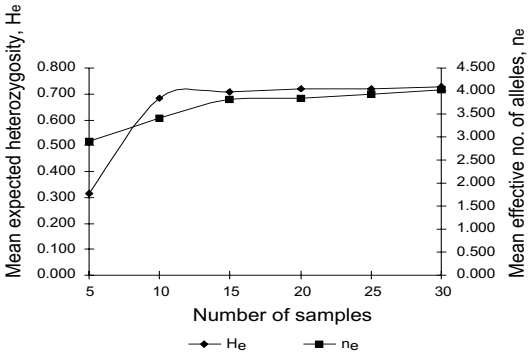


Figure 6 Relationship between measures of genetic diversity and number of samples of *S. curtisii* adult trees in Compartment 118, Ulu Sedili Forest Reserve revealed by SSR markers

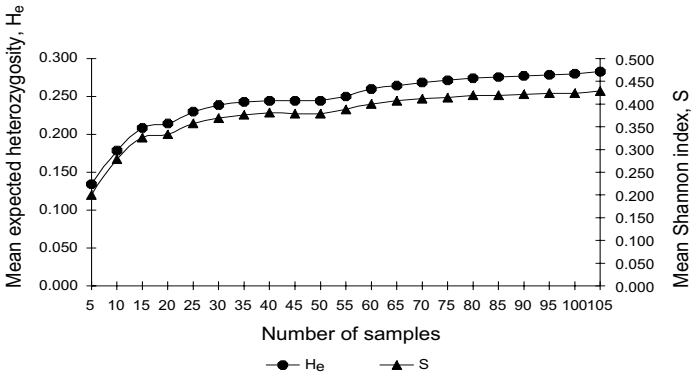


Figure 7 Relationship between measures of genetic diversity and number of samples of *S. curtisii* seedlings in Compartment 118, Ulu Sedili Forest Reserve revealed by DAMD markers

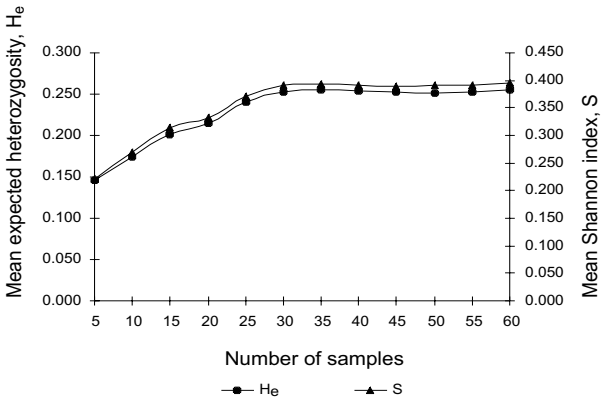


Figure 8 Relationship between measures of genetic diversity and number of samples of *S. curtisii* saplings in Compartment 118, Ulu Sedili Forest Reserve revealed by DAMD markers

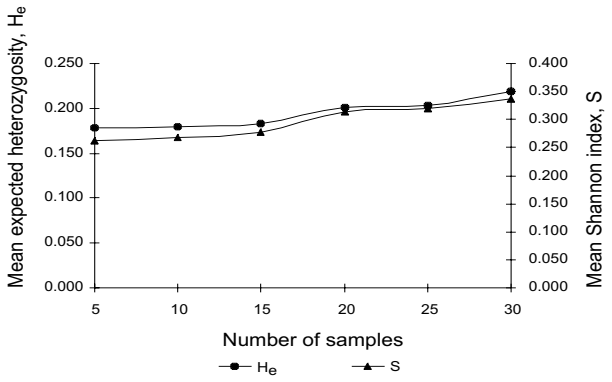


Figure 9 Relationship between measures of genetic diversity and number of samples of *S. curtisii* adult trees in Compartment 118, Ulu Sedili Forest Reserve revealed by DAMD markers

Table 5 Genetic distance between pairs of age cohorts of *S. curtisii* using SSR and DAMD data

Age cohort	Seedling	Sapling
Sapling	*0.042	
Adult	** 0.014	
	*0.081	*0.103
	** 0.076	**0.096

*SSR data (Nei's [1978] genotypic coefficient), mean pairwise genetic distance is 0.076 ± 0.018
**DAMD genotypic method (Nei's [1978] genotypic coefficient), mean pairwise genetic distance is 0.062 ± 0.025

matrices based on RAPD and allozyme indicating that both markers provide different estimates of genetic relationship among the populations tested. The greater discriminatory power of RAPD compared with allozyme may be due to its greater coverage of the genome. Ujino *et al.* (1998) found that the frequency of dinucleotide repeats in the *Shorea* genome occurred approximately once in every 1200 kb. Condit and Hubbel (1991) screened five tropical tree species and reported that one dinucleotide repeat was found in every 64 to 1105 kb. Thus, dinucleotide repeats in *S. curtisii* are apparently abundant and evenly dispersed throughout the genome. In other words, both codominant (SSR) and dominant (DAMD) markers have sufficient coverage in the *S. curtisii* genome. Therefore, combination of these two markers would yield complementary information for assessing genetic diversity, and genetic erosion resulting from logging practices and population structure.

Acknowledgements

We gratefully acknowledge the foresters from Renjer Office, Kota Tinggi as part of Johor Selatan Forestry Department for establishing the demographic plots in Compartment 118, Ulu Sedili Forest Reserve. The funding for this study was provided by Ministry of Science, Technology and Innovation (MOSTI) through IRPA 01-02-02-0022 and Malaysia Toray Science Foundation (MTSF) grants.

References

Bebeli, P. J., Zhou, Z., Somers, D. J. & Gustafson, J. P. 1997. PCR primed with minisatellite core sequences yield DNA fingerprint probes in wheat. *Theoretical and Applied Genetics* 95: 276–283.

Buchert, G. P., Rajora, O. P., Hood, J. V. & Dancik, B. P. 1997. Effects of harvesting on genetic diversity in old-growth Eastern White Pine in Ontario, Canada. *Conservation Biology* 11(3): 747–758.

Chase, M. R., Kesseli, R. & Bawa, K. S. 1996. Microsatellite markers for population and conservation genetics of tropical trees. *American Journal of Botany* 83: 51–57.

Chong, D. K. X., Yang, R. C. & Yeh, F. C., 1994. Nucleotide divergence between population of trembling aspen (*Populus tremuloides*) estimated with RAPDs. *Current Genetics* 26: 374–376.

Chong, D. K. X., Yeh, F. C., Aravanopoulos, F. A. & Zsuffa, L. 1995. DNA fingerprinting willows (*Salix* L.) using polymerase chain reaction with the M13 universal primer. *Scandinavian Journal of Forest Research* 10: 27–31.

Condit, R. & Hubbell, S. P. 1991. Abundance and DNA sequence of 2-base repeat regions in tropical tree genomes. *Genome* 34: 66–71.

Doyle, J. & Doyle, L. 1990. Isolation of plant DNA from fresh tissue. *Focus* 12: 13–15.

Estoup, A., Tailliez, C., Cornuet, J. M. & Solignac, M. 1995. Size homoplasy and mutational processes of interrupted microsatellites in two bee species, *Apis mellifera* and *Bombus terrestris* (Apidae). *Molecular Biology and Evolution* 12: 1074–1084.

Grimaldi, M. C. & Crouau-Roy, B. 1997. Microsatellite allelic homoplasy due to variable flanking sequences. *Journal of Molecular Evolution* 44: 336–340.

Heath, D. D., Iwana, G. K. & Delvin, R. H. 1993. PCR primed with VNTR core sequences yield species specific patterns and hypervariable probes. *Nucleic Acids Research* 21: 5782–5785.

Huang, B. X., Peakall, R. & Hanna, P. J. 2000. Analysis of genetic structure of blacklip abalone (*Haliotis rubra*) populations using RAPD, minisatellite and microsatellite markers. *Marine Biology* 136: 207–216.

Lee, C. T., Wickneswari, R., Mahani, M. C. & Zakri, A. H. 2002. Effect of selective logging on the genetic diversity of *Scaphium macropodum*. *Biological Conservation* 104(1): 107–118.

Lee, S. L., Wickneswari, R., Mahani, M. C. & Zakri, A. H. 2001a. Genetic diversity of a tropical tree species, *Shorea leprosula* Miq. (Dipterocarpaceae), in Malaysia: implications for conservation of genetic resources and tree improvement. *Biotropica* 32: 213–224.

Lee, S. L., Wickneswari, R., Mahani, M. C. & Zakri, A. H. 2001b. Comparative genetic diversity studies of *Shorea leprosula* Miq. (Dipterocarpaceae) using RAPD and allozyme markers. *Journal of Tropical Forest Science* 13(1): 202–215.

Lefort, F., Lally, M., Thompson, D. & Douglas, G. C. 1998. Morphological traits, microsatellite fingerprinting and genetic relatedness of a stand of elite oaks, *Quercus robur* L. at Tullynally, Ireland. *Silvae Genetica* 47: 257–262.

Lewontin, R. C. 1973. The apportionment of human diversity. *Evolutionary Biology* 6: 381–398.

Li, C. C. & Horvitz, D. G. 1953. Some methods of estimating the inbreeding coefficient. *American Journal of Human Genetics* 90: 27–32.

Mantel, N. 1967. The detection of disease clustering and a generalized regression approach. *Cancer Research* 27: 209–220.

Maughan, P. J., Saghai Maroof, M. A. & Buss, G. R. 1995. Microsatellite and amplified sequence length polymorphisms in cultivated and wild soybean. *Genome* 38: 715–723.

Métais, I., Aubry, C., Hamon, B., Jalouzot, R. & Peltier, D. 2001. Assessing common bean genetic diversity using RFLP, DAMD-PCR, ISSR, RAPD and AFLP markers. *Acta Horticulturae* 546: 459–461.

Morgante, M., Rafalski, A., Biddle, P., Tingey, S. & Olivieri, A. M. 1994. Genetic mapping and variability of seven soybean simple sequence repeat loci. *Genome* 37: 763–769.

Murray, M. G. & Thompson, W. F. 1980. Rapid isolation of high molecular weight plant DNA. *Nucleic Acids Research* 8: 4321–4325.

Namkoong, G. 1992. Biodiversity—issues in genetics, forestry and ethics. *Forestry Chronical* 68: 438–443.

Nei, M. 1978. Estimation of average heterozygosity and genetic distance from a small number of individuals. *Genetics* 89: 583–590.

Obayashi, K., Tsumura, Y., Ihara-Ujino, T., Niiyama, K., Tanouchi, H., Suyama, Y., Washitani, I., Lee, C. T., Lee, S. L. & Norwati, M. 2002. Genetic diversity and outcrossing rate between undisturbed and selectively logged forests of *Shorea curtisii* (Dipterocarpaceae) using microsatellite DNA analysis. *International Journal of Plant Science* 163(1): 151–158.

Rajora, O. P., Rahman, M. H., Buchert, G. P. & Dancik, B. P. 2000. Microsatellite DNA analysis of genetic effects of harvesting in old-growth eastern white pine (*Pinus strobus*) in Ontario, Canada. *Molecular Ecology* 9: 339–348.

Roa, A. C, Chavarriaga-Aguirre, P., Duque, C. M., Maya, M. M., Bonierbale, M. W., Iglesias, C. & Tohme, J. 2000. Cross-species amplification of cassava (*Manihot esculenta*) (Euphorbiaceae) microsatellites: allelic polymorphism and degree of relationship. *American Journal of Botany* 87(11): 1647–1655.

Rogstad, S. H., Wolff, K. & Schaal, B. A. 1991. Geographical variation in *Asimina triloba* Dunal. (Annonaceae) revealed by the M13 DNA fingerprinting probe. *American Journal of Botany* 78(10): 1391–1396.

Schaal, B. A., O'Kane, S. L. & Rogstad, S. H. 1991. DNA variation in plant populations. *Trends in Ecology and Evolution* 6: 329–333.

Skroch, P. & Nienhuis, J. 1995. Impact of scoring error and reproducibility of RAPD data on RAPD based estimates of genetic distance. *Theoretical and Applied Genetics* 91: 1086–1091.

Smith, D. N. & Devay, M. E. 1994. Occurrence and inheritance of microsatellites in *Pinus radiata*. *Genome* 37: 977–983.

Somers, D. J., Zhou, Z., Bebeli, P. J. & Gustafson, J. P. 1996. Repetitive, genome-specific probes in wheat (*T. aestivum* L. em Thell) amplified with minisatellite core sequences. *Theoretical and Applied Genetics* 93: 982–989.

Swofford, D. L. & Selander, R. B. 1981. BIOSYS-1: A FORTRAN program for the comprehensive analysis of electrophoretic data in population genetics and systematics. *Journal of Heredity* 72: 281–283.

Symington, C.F. 1943. *Foresters' Manual of Dipterocarps*. Malayan Forester Records 16. University of Malaya Press, Kuala Lumpur.

Szmidt, A. E., Wang, X. R. & Lu, M. Z. 1996. Empirical assessment of allozyme and RAPD variation in *Pinus sylvestris* (L.) using haploid tissue analysis. *Heredity* 76: 412–420.

Templeton, A. R. 1995. Biodiversity at the molecular genetic level: experiences from disparate macroorganisms. Pp. 59–64 in Hawksworth, D. L. (Ed.) *Biodiversity. Measurement and Estimation*, Chapmen & Hall, London.

Thomas, B. R., Macdonald, S. E., Hicks, M., Adams, D. L. & Hodgetts, R. B. 1999. Effects of reforestation methods on genetic diversity of lodgepole pine: an assessment using microsatellite and randomly amplified polymorphic DNA markers. *Theoretical and Applied Genetics* 98: 793–801.

Ujino, T., Kawahara, T., Tsumura, Y., Nagamitsu, T., Yoshimaru, H. & Wickneswari, R. 1998. Development and polymorphism of simple sequence repeat DNA markers for *Shorea curtisii* and other Dipterocarpaceae species. *Heredity* 81: 422–428.

Van de Ven, W. T. G. & McNicol, R. J. 1996. Microsatellites as DNA markers in Sitka spruce. *Theoretical and Applied Genetics* 93: 613–617.

Wickneswari, R. & Lee, C. T. 2001. Assessment of genetic diversity of *Scaphium macropodum* in logged-over forests using isozyme and RAPD markers. Pp. 104–109 in Ganeshaiah, K.N., Shaanker, U. & Bawa, K. S. (Eds.) *Tropical Ecosystems: Structure, Diversity and Human Welfare*, Oxford-IBH Press, New Delhi.

Wickneswari, R., Lee, C. T., Norwati, M. & Boyle, T. J. 1999. Impact of logging on genetic diversity in humid tropical forests. Pp. 171–181 in Matyas, C. (Ed.) *Forest Genetics and Sustainability,* Kluwer Academic Publishers, Netherlanda.

Wright, S. 1978. *Variability Within and Among Natural Populations.* Volume 4. The University of Chicago Press, Chicago.

Zhou, Z., Bebeli. P., Somers, D. J. & Gustafson, J. P. 1997. Direct amplification of minisatellite-region DNA with VNTR core sequences in the genus *Oryza. Theoretical and Applied Genetics* 95: 942–949.