JOURNAL OF FOREST SCIENCE, *60*, 2014 (1): 18–27

Microsatellite DNA Markers in *Shorea platyclados* (Dipterocarpaceae): genetic diversity, size homoplasy and mother trees

A.M. Javed[1], Ch.H. Cannon[1], R. Wickneswari[2]

[1]*Department of Biological Sciences, Texas Tech University, Lubbock, USA*

[2]*School of Environmental and Natural Resource Sciences, Faculty of Science and Technology, Universiti Kebangsaan Malaysia, Bangi, Selangor, Malaysia*

ABSTRACT: Cross-specific amplification of microsatellite loci greatly enhances the effectiveness of this marker system. This shortcut would greatly enhance our examination of the gene flow and population structure of trees in diverse tropical rainforests. To explore the effectiveness and limitations of this approach, we examined allelic diversity at six microsatellite loci, originally developed in a congeneric species, in three populations of *Shorea platyclados* from Peninsular Malaysia. Fragment sizing was performed by an efficient and sensitive (1 bp resolution) technique using capillary electrophoresis, ethidium bromide detection, and minimal clean-up. Fragment size ranges were conserved between species and null allele frequencies were low. Higher overall levels of genetic diversity were detected in our study. Variation among populations was directly related to geographic distance. Fragment size class distributions suggest that each locus should be studied using different evolutionary models. Direct sequencing of SSR fragments revealed that size differences were due to changes in both the flanking regions and repeat motifs. Several clear examples of size homoplasy were observed, along with the disruption of perfect repeats, suggesting that cross-specific amplification of microsatellite loci requires an additional level of confirmation at the DNA sequence level before the influence of size homoplasy and changes in repeat structure can be assessed. Simulation studies demonstrate that the increasing intensity of timber harvest leads to higher variability in levels of potential heterozygosity and decreasing total number of alleles in the remnant "mother trees" The careful selection of "mother" trees can greatly enhance the future genetic diversity of populations.

Keywords: dipterocarp; timber identification; certification

Emergent trees of the family Dipterocarpaceae dominate the lowland forests of Southeast Asia, in terms of diversity and biomass (Ashton 1964; Baillie et al. 1987; Cannon et al. 1994). Despite their great ecological and economic importance, relatively little is known about the genetics of these trees. The effects of selective harvest on genetic structure have been studied in only a few species (Murawski et al. 1994; Lee et al. 2000; Wickneswari et al. 2000; Obayashi et al. 2002) and the accumulation of information has been slow in comparison with the rate of forest conversion (Sodhi et al. 2004). Historically, the majority of timber harvest has occurred in the lowland habitats but as stocks decline, the focus of the harvest has progressively moved into the uplands. Uplands areas (> 750 m elevation) present a small fraction (7%) of the total land area in peninsular Malaysia, representing an archipelago of the habitat. These upland habitats are arguably more fragile than lowland forests (Ashton et al. 2001; Bruijnzeel 2004) and represent some of the largest contiguous intact habitats remaining in Malaysia. *Shorea platyclados* is an upland species of the family Dipterocarpaceae, limited to forests above 700-meter elevation. It accounts for 14% of the stand and is the most dominant species above the legal logging limit in 600 hectares of the Perak Integrated Timber Complex (PITC) concession (unpublished data), a recently established logging unit in the Temengor Forest Reserve. This species provides a challenging test case for the rapid

development of informative molecular markers because of its future importance to the logging industry and the relatively undisturbed patterns of population structure and gene flow. In working with PITC, we can also examine before and after effects of logging on the impact of future dynamics of genetic diversity.

Moreover, there is a rising demand of legally harvested timber, resulting in a stringent regulation imposed by importing countries, that timber and timber products are harvested in a sustainable way. European Union (EU) timber regulation No. 995/2010 prohibits the export of illegally harvested timber to European markets. This regulation demands the need for a timber tracking system to verify the source and place of origin. Conventional tracking systems used paper-based documentation for identification purposes thus increasing the possibility of forgery. Alternative methods based on DNA markers could become methods of choice (Tnah et al. 2010).

Microsatellites are the genetic marker system of choice for population biologists examining questions of genetic structure and gene flow (ISI's Web of Science contains > 21,000 citations identified by the keyword). This marker system combines several attractive features, including high levels of polymorphism, co-dominance of alleles, reproducibility, wide distribution across the genome and is amenable to a simple and rapid PCR assay (Tautz 1989; Charlesworth et al. 1994; Gupta et al. 1996; Weising et al. 2005). The widespread use of microsatellites has been limited by the fact that informative novel species specific markers have to be isolated for the focal organism before a study can begin (Steinkellner et al. 1997; Ng et al. 2013), although many different approaches are currently available (Zane et al. 2002). Cross-specific amplification of already published microsatellite loci is a possible way of quickly and efficiently expanding our knowledge about population structure, gene flow, management policies and conservation genetics of these ecologically and commercially important trees (Byrne et al. 1996). Previous work indicates that microsatellite loci might be widely conserved across the family (Ujino et al. 1998; Ng et al. 2013). In this paper, we examined the feasibility and limitations of using cross-amplified microsatellite loci in *Shorea platyclados*.

MATERIAL AND METHODS

Plant material and DNA isolation. A total of 159 individuals of *Shorea platyclados* (Dipterocarpaceae) were sampled at three different locations along the Titiwangsa mountain range encompassing the states of Pahang, Perak and Kelantan in Peninsular Malaysia. The specific sample numbers at each of the three populations were: (*i*) Fraser's Hill = 51 (Pahang); (*ii*) Perak Integrated Timber Complex (PITC) = 54 (Perak); and (*iii*) Belum = 54 (Kelantan). The first population is from a small mountain at the southern end of the range and in a well-established conservation and ecotourism area. The second population is at the northern end of the range and is fully contained within the PITC concession area. The Forest Stewardship Council and Scientific Certification Systems have certified the PITC operation as a sustainable forest management unit. The third population is also in the north and adjacent to a newly established protected area.

DNA extraction. Whole genomic DNA was extracted from inner bark and leaf samples from each individual. DNA was extracted using the modified CTAB protocol (Doyle, Doyle 1987). Crude pellet containing DNA was further purified using the High Pure PCR Template Preparation Kit (Roche-Diagnostics).

PCR reactions and capillary electrophoresis detection. Six microsatellite loci (Table 1) from a set of nine SSR markers developed in *S. curtisii* (Dipterocarpaceae) by Ujino et al. (1998) were assessed for transfer in *S. platyclados*. Modifications in $MgCl_2$, DNA template, and primer concentrations and annealing temperature were made to optimize the PCR reactions for each primer pair. Annealing temperatures observed varied in *S. platyclados* (Table 1) compared to *S. curtisii* (Ujino et al. 1998). PCR reactions were performed using an Eppendorf Mastercycler Gradient thermocycler. Reaction mixtures of 25 µl volume contained 10 ng template DNA, 1 × PCR buffer [10mM Tris-HCl, 50mM KCl, (NH_4) SO_4], and 1.5mM $MgCl_2$, 0.5 $U \cdot \mu l^{-1}$ *Taq* DNA polymerase (Eppendorf, Hamburg, Germany), 0.2mM each dNTP (Eppendorf), and 5 pmol of each primer. PCR amplification was carried out for 5 min at 95°C followed by 35 cycles of 45 s at 94°C, 30 s at 51.1–55.2°C (Table 1) and 45 s at 72°C. After the final cycle, samples were incubated at 72°C for 5 min and held at 4°C prior to analysis. Amplification success was first confirmed on a 2% agarose gel by ethidium bromide staining. To examine the effect of different detection systems on the accuracy and reproducibility of fragment sizing, PCR fragment size was determined using both polyacrylamide gel electrophoresis (PAGE) and capillary electrophoresis (CE). The PAGE system consisted of a Protean® II (BioRad, Hercules, USA) vertical gel rig. Fragments were resolved on 6–8%

acrylamide/bisacrylamide (19:1), 7M urea and 1 × TBE denaturing gel. The gels were run at 40W constant power for 2–3 h and visualized using silver staining (Kaemmer et al. 1997). All individuals were also screened using a 24-capillary Reveal Mutation Discovery System (Spectrumedix Inc., State College, USA). The use of ethidium bromide detection in this capillary system had never been explored and run conditions had to be standardized. Ethidium bromide concentration was optimized at 0.4 $\mu g \cdot ml^{-1}$ of ethidium bromide to provide standard high intensity peaks. A non-fluorescent 50 bp DNA ladder (Promega, Inc., Madison, USA) was admixed to each PCR sample as an internal standard. Different proportions of the mixture were initially tested using different volumes of PCR products and various ladder dilutions (data not shown). Finally 8 µl of PCR product was mixed with 6 µl of (15:1 diluted) 50 bp DNA step ladder (340 $\mu g \cdot \mu l^{-1}$, Promega, Madison, USA). Samples were overlaid with mineral oil (Sigma). Running, washing buffers and gel matrix were purchased from Spectrumedix, Inc. Twenty-four samples were loaded per plate and six plates were run simultaneously (144 samples per run). The runs were performed at a constant temperature (35°C) for 90 min. Fragment sizing was performed by the QuickSpec™ 2.4. software (Spectrumedix Inc., State College, USA). Reproducibility of fragment sizes was checked by running selected samples several times.

Data analysis. Due to the higher accuracy and reproducibility of the results from the CE detection system (data not shown), all subsequent analyses excluded the PAGE results. The presence of alleles could easily be detected on the electrophoregrams. Only distinct and clear peaks were scored while indistinct or shorter peaks were considered stutter bands resulting from slippage of *Taq* polymerase during PCR amplification (Lagercrantz et al. 1993; Wu, Tanksley 1993). The number of alleles and degree of polymorphism were calculated for each locus and population. Allele richness, heterozygosity and effective number of alleles were calculated for each locus-population combination, using the POPGENE 32 software (Yeh et al. 1997). For each locus in each population, deviations from the Hardy-Weinberg expectation were examined by calculating Wright's (1965) fixation indices (F_{is}). Fixation indices were tested for significance by X^2 tests according to Litz and Horvitz (1953). GENEPOP 3.1b (Raymond & Rousset, Montpellier, France) was used to calculate Weir and Cockerham's (1984) estimate of F_{st}. For each population, statistical significance of F_{st} value was based on the X^2 test (Workman, Niswander 1970). Earlier SSR studies have shown that the deviation from the Hardy-Weinberg equilibrium was caused by the heterozygote deficiency (Williams et al. 2000). At microsatellite loci, species heterozygote deficiencies were observed and may be explained by the presence of null alleles. To estimate how the frequency of null alleles (r_b) affects observed heterozygosity (H_{obs}), r_b was calculated according to Brookfield (1996).

Direct microsatellite sequencing. Amplification products from two trans-specific loci Shc02 and Shc08 were sequenced. Four homozygous individuals (with identical fragment sizes) from Shc02 and four heterozygous individuals (with different fragment sizes) from Shc08 were sequenced to examine the variation in repeat number and their flanking regions. For this purpose, PCR products were separated on a PAGE gel and stained with silver nitrate (Kaemmer et al. 1997). Target bands were cut out of the gel and soaked in 100 µl of ddH_2O for 10 min initially and then boiled for 15 min in an Eppendorf Mastercycler Gradient thermocycler. Gel debris was removed by centrifuging at 10,000 rpm for 2 min. The supernatant was collected and mixed with 10 µl of 3M ammonium acetate, 5 µl of glycogen (10 $mg \cdot ml^{-1}$), and 450 µl of 100% EtOH. Samples were incubated at –40°C for 30 min. DNA was pelleted down by centrifuging at 4°C for 10 min at 10,000 rpm. The pellet was washed twice with 200 µl of ice-cold 85% EtOH and then dissolved in 10 µl of ddH_2O and 4 µl was used subsequently for the sequencing reaction. Fragments were sequenced using a CEQ dye terminator cycle sequencing kit (http://www.beckmancoulter.com) on CEQ 2000 from both directions using forward and reverse primers. Sequence files were edited and aligned using the SeqMan and CLUSTAL-W method included in the MEGALIGN software (DNASTAR, Inc., Madison, USA), followed by manual adjustments.

Logging simulation. The Malaysian silvicultural system requires a comprehensive pre-felling inventory (Potts et al. 2005), during which mother trees are chosen. These mother trees are not harvested but are left undamaged to produce the next generation of timber trees. The process for selection of these mother trees is based almost entirely on the experience of the forest officers. To examine the interactive effects of harvest intensity and selection of mother trees on the remnant and future genetic diversity, numerous replicate jackknife data sets were produced for each population, without replacement, for no logging (0% jackknife), ¾ cut (25%), and so

forth down to two individuals. The future potential in genetic diversity was estimated from the range of maximum and minimum values found in these replicate data sets for two measurements: total potential heterozygosity of offspring and total number of alleles. Total potential heterozygosity of offspring was estimated by calculating the percentage of heterozygosity among offspring resulting from all possible matings of the remnant mother trees. Total number of alleles was the percentage of alleles remaining among all mother trees and loci.

RESULTS

Allelic diversity and geographic distribution

Six primers screened were successfully amplified in all *Shorea platyclados* individuals including the Shc08 locus which produced strong amplifications as also reported in *Shorea curtisii* (Ho 2001) contradicting the earlier findings (Ujino et al. 1998). PCR fragments amplified from all 159 individuals by the six loci tested in *Shorea platyclados* fell within the expected size range. Two loci (Shc01 and Shc09) were highly polymorphic and had similar numbers of alleles (24 and 25, respectively) but the fragment size distribution for Shc01 was clearly unimodal, nearly normal, and roughly identical in all populations while Shc09 was multimodal, randomly distributed, and varied widely across the three populations (Fig. 1). The Shc11 locus was substantially less polymorphic (11 alleles) and the fragment size distributions were unimodal in each population but the mode was different among populations (Fig. 1). Most alleles were detected at frequencies less than 10%, both within and among populations. Nevertheless, the Shc02 locus exhibited allelic frequencies greater than 50%. The Shc01 locus possessed the widest range of fragment sizes (117 to 204 or 87 bp). A higher level of polymorphism is weakly correlated with the increasing fragment length. The frequency of rare alleles (i.e. observed in a single individual or population) was not positively correlated with the allelic diversity or size range (Table 1). Two loci violated the assumptions of Hardy-Weinberg equilibrium (Shc02, and Shc11) in all three populations based on F_{is} test values. Most of the significant violations involved higher than expected levels of homozygosity, while two loci (Shc02 and Shc03) exhibited heterozygosity excesses in the Fraser's Hill population. Only three within-population fixation indices (F_{is}) were significantly higher than zero (Table 1), indicating that the gene flow within each population was random and non-assortative. Two of these significant values were observed in the PITC population and the pooled F_{is} values in this population were generally higher than in the other two populations for all loci. The pooled F_{is} values were not significantly different from zero showing that the populations in general were in Hardy-Weinberg equilibrium (Table 1). One cause of heterozygote deficiency is the presence of null alleles. The r_b estimator, which estimates the frequency of null alleles in each population at each locus, varied from –0.008 to 0.204 (Table 2) where both extremes were observed in the PITC population. When overall r_b values were calculated for all six loci per populations, the lowest value (0.036) was observed in Fraser population and the highest (0.104) was found in PITC population (Table 2). F_{st} the standardized variance in

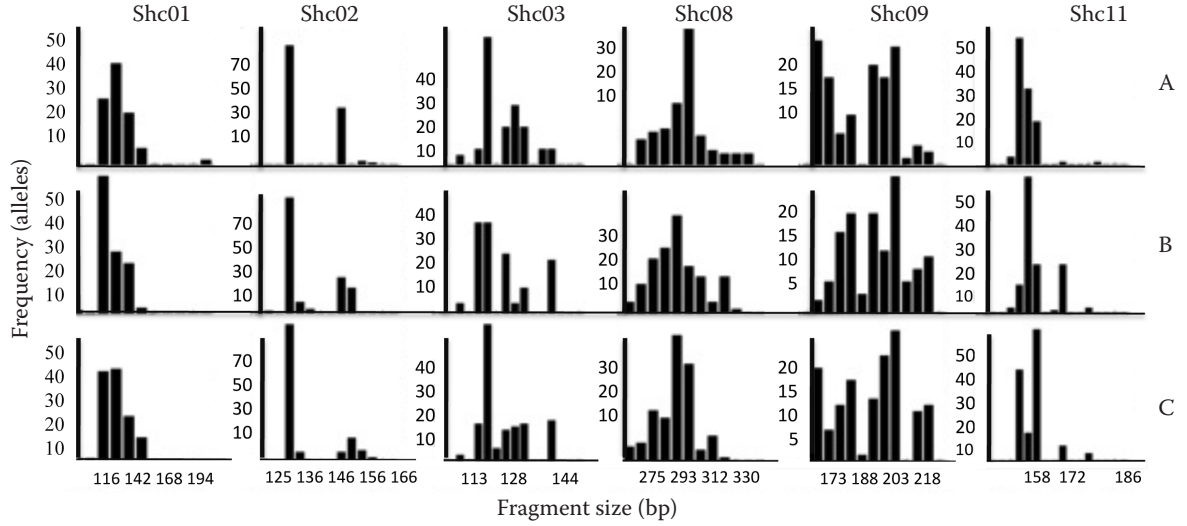

Fig. 1. Size distributions of PCR fragment sizes for six microsatellite loci in three populations of upland *Shorea platyclados* bars – scaled to relative abundance of the most frequent class in each graph, rows illustrate populations: A – Fraser's Hill; B – Belum; C – PITC

Table 1. Microsatellite loci in *Shorea platyclados*: basic information, genetic diversity, and population structure

	Sch01	Sch02	Sch03	Sch08	Shc09	Shc11
°C	54	55.2	54.2	51.1	52.2	55.2
Bp	117–204	132–166	120–146	279–340	178–228	159–182
F_{is}	0.06	0.20	–0.08	0.25	0.11	0.39
F_{it}	0.08	0.29	–0.04	0.26	0.12	0.45
F_{st}	0.02	0.11	0.03	0.01	0.01	0.10
N_m	12.7	2.7	9.9	21.1	14.4	2.9

each column refers to a different locus (Ujino et al. 1998); °C – optimized annealing temperature; bp – size range of fragments in base pairs; F_{is} – inbreeding index in total population; F_{it} – inbreeding within populations; F_{st} – inbreeding among populations; N_m – gene flow estimated from F_{st}; estimates were calculated according to standard techniques (Raymond, Rousset 1995)

Table 2. Allelic diversity and heterozygosity estimates for six microsatellite loci for three populations

	TA	HW	F_{is}	R_b
Shc01				
A	19	0.00	0.11	0.04
B	13	.77	–.02	-.01
C	19	0.00	0.07	0.05
All	24	D	0.06	
Shc02				
A	6	0.02	0.18	–.08
B	11	0.00	0.14	.01
C	12	0.00	0.27	0.20
All	14	D	0.20	
Shc03				
A	11	1.0	–.09	.08
B	10	0.79	–.04	.02
C	10	0.34	0.02	.01
All	13	E	–.08	
Shc08				
A	18	0.00	0.17	0.09
B	19	0.00	0.20	0.14
C	18	0.00	0.25	0.13
All	20	D	0.25	
Shc09				
A	20	0.01	0.10	0.09
B	23	.21	0.01	0.01
C	19	0.01	0.10	0.06
All	25	D	0.11	
Shc11				
A	8	0.00	0.25	0.11
B	10	0.00	0.34	0.20
C	8	0.00	0.24	0.11
All	11	D	0.39	

TA – total number of alleles, HW – exact probability results of the Hardy-Weinberg equilibrium test, F_{is} – Wright's fixation index; populations are indicated: A – Fraser's Hill, B – Belum, C – PITC, All – three populations together; multilocus estimates of F_{is} are shown for all populations; for each locus, significant results for excess or deficiency of heterozygotes in a global test by E or D, respectively, are shown for all populations, R_b – potential contribution of null alleles

allele frequencies between populations, is the measure of genetic distance between populations. Pairwise comparison for F_{st} revealed moderate levels of differentiation with average F_{st} (0.029 ± 0.005). The average calculated F_{st}= 0.029, which is moderately low, means approximately 2.9% of all variation found that was partitioned between populations. Fraser and PITC populations were observed to be genetically distant (F_{st} = 0.0343) followed by Kelantan & Fraser (F_{st} = 0.029) and Kelantan & PITC (F_{st} = 0.0237). These distances positively correlated with the geographical distribution of three populations.

Sequence variation

Direct sequencing of the PCR products from Shc02 and Shc08 revealed that these primer pairs amplified fragments containing similar simple sequence repeats as originally reported, although the repetitive motifs were not perfectly conserved in either locus. For Shc02, an insertion of a C was observed (Table 3) in the complex repeat sequence for all individuals at the boldfaced and underlined position: $(CT)_2CA(CT)_nCGC(AT)_2$. A transversion (T > G) has disrupted the simple dinucleotide repeat $(CT)_n$ in the Shc08 locus, isolating two copies of the motif away from the longer repetitive sequence, creating an imperfect repeat $[(CT)_2 CG(CT)_n]$ in all individuals (Table 3). Comparison of DNA sequence between the repeat and flanking regions of these two loci illustrated two different types of size homoplasy. While the repeat motif in Shc02 was generally conserved across the sequences, the flanking regions possessed several point mutations (Table 3). These fragments were of identical length and would normally be interpreted as homologs but the flanking region sequence indicates a more complex history. To uncover this history, a comprehensive study of DNA sequence variation in the flanking regions for all fragments would have

Table 3. Polymorphisms in DNA sequence detected for microsatellite loci Shc02 and Shc08

	24	**48**	**53**	**62**	**66**	***84***								
Shc02-1	A	T	A	T	T	*C*								
Shc02-2	.	.	.	.	.	.								
Shc02-3	.	G	G	G	G	.								
Shc02-4	T	.	G	.	.	.								
	12	**71**	**104**	**106**	**108**	**109**	**110**	**111**	**112**	**113**	**114**	**115**	**116**	**117**
Shc08-1	–	G	T	T	T	T	T	A	C	T	T	T	T	T
Shc08-2	A	C	C	A	.	C	C	.	T	–	–	–	–	–
Shc08-3	–	.	.	.	.	.	.	C	.	.	.	.	C	C
Shc08-4	–	.	.	.	C	A	.	T	.	.	C	.	.	.
	118	**119**	**120**	**121**	**122**	**123**	**124**	**125**	**126**	**127**	**128**	**129**	**137**	**148**
Shc08-1	T	C	A	C	–	T	A	C	A	C	C	C	–	–
Shc08-2	–	–	–	–	–	–	–	–	–	–	–	–	G	T
Shc08-3	A	.	–	–	–	–	–	–	–	.	.	.	.	.
Shc08-4	A	A	.	T	A	A	.	.	.	.	.	.	T	.

in bold – position in the reference sequence, only polymorphisms are shown; in bold italic – insertion not present in the original microsatellite sequence; dot – nucleotides which agree with the consensus; dash – gaps

to be performed (Dettman, Taylor 2004). The patterns of DNA sequence variation in Shc08 were much more complex, with variation seen in the repeat and flanking regions. Repeat number varied from 11 to 17 and both indels and point mutations were seen in the flanking regions. A large indel of 17 bp was observed at position 113, with repetitive sequences of interrupted microsatellite sequences (Table 3). Given some SSR primer-site conservation among the Dipterocarps, we performed BLAST searches for nucleotide similarity to determine whether the microsatellites were associated with conserved gene regions. The BLAST searches were based on the full-length allele sequence but only Shc02 locus showed similarity to GenBank sequences. Shc02 partial sequence showed 93.9% nucleotide similarity to a putative high mobility group (HMG) protein of mRNA (*Arabidopsis thaliana* BT005660). The microsatellite resides within an intron, with one primer anchored in the HMG coding region. HMG is one of the most abundant non-histone chromosomal proteins in eukaryotes, has multiple roles in transcription, replication and cellular differentiation. The primary sequence of HMG from various higher organisms is highly conserved (Bustin, Reeves 1996), therefore the exo-directed primer is likely to function in many related species. This is evident from Shc02 sequence similarity (97.7% identity, nucleotide similarity, AY486859) to *Hevea brasiliensis* microsatellite sequence. It contains the fully conserved one primer sequence and flanking region of *S. platyclados* and partial repeat region. While one expects less conservation of the intron, our amplification results seem to indicate that there is enough conservation to ensure cross-specific amplification across *Shorea* species and also across different genera of Dipterocarpacae (Ujino et al. 1998).

Selection of mother trees

Logging intensity has a profound impact on total allelic diversity and the overall range of outcomes becomes significantly narrower with increasing harvest intensity. Under most outcomes, allelic diversity does not begin to decline until harvest intensity goes beyond 50%. Additionally, total allelic diversity remains above 50% until harvest intensities exceed 80%. Under the best of the outcomes (the upper dashed line), allelic diversity does not fall below 50% until harvest reaches almost 90% of the stand. The potential levels of heterozygosity among offspring become more and more divergent at increasing levels of harvest and increasingly dependent on the selection of mother trees. If mother trees are selected to maximize genetic diversity, a low overall reduction of allelic diversity or potential heterozygosity could be maintained at relatively high levels of harvest (roughly 70%).

DISCUSSION

A major drawback for microsatellite analysis is the need to develop informative loci for each study

species. However, the reported frequency of cross-specific amplifications, at the generic level and sometimes even across genera, has increased as more and more attempts have been made (Konuma et al. 2000; Stacy et al. 2001; Gutierrez et al. 2005). The ability to adopt informative loci from one species to the study of a wide range of closely related species would be a major boom, particularly for tropical biologists. For example, over 200 species of *Shorea* are found in Southeast Asia (Ashton 1982) and a dozen species can easily be found in a single landscape (Cannon, Leighton 2004). Cross-specific amplification of the growing number of microsatellite loci available (Lee et al. 2004) would provide a significant shortcut to Southeast Asian foresters for this major economic tree group. In this study, we demonstrated that six microsatellite primers originally developed for *S. curtisii* could work in *S. platyclados*. The loci were highly polymorphic and a high level of genetic diversity was observed in all the sampled populations. One locus was of particular interest, i.e. Shc08 was successfully amplified in this study previously reported to have weak amplifications in preliminary studies on Dipterocarpaceae species (Ujino et al. 1998). However, it was strongly amplified and observed to be highly polymorphic in *Shorea platyclados*. Ho (2001) used the same set of loci including Shc08 and successfully amplified adult trees and seedlings of *Shorea curtisii*, a parent species from which these loci were developed (Ujino et al. 1998). In the same study Shc08 was observed with 14 alleles ranging from 239 to 353 bp and expected heterozygosity range from 0.66 to 0.74. Konuma et al. (2000) tested all nine loci (Ujino et al. 1998); however they found only four polymorphic or working well in *Neobalanocarpus heimii*. Interestingly excluded loci Shc01, Shc04 and Shc11 were earlier reported with strong amplifications in the same species by Ujino et al. (1998). Ng et al. (2004) used seven different loci developed by Ujino et al. (1998) for *Shorea leprosula* and *Shorea ovalis ssp. sericea*, and the exclusion of Shc08 and Shc09 loci was not based on poor amplification of these loci. Weak amplifications could result because of low quality of DNA (Nagamitsu et al. 2001), annealing temperature, variation in PCR reagent components used in various studies and different species affinities.

Previous studies that carefully examined the results of cross-amplification often found limitations to the approach, including loss of the internal repeat (Gutierrez et al. 2005), disruption of the repeat structure (Shao et al. 2005), increased rate of mutation in the flanking regions (González-Martínez et al. 2004), and increased levels of null alleles (Wright et al. 2004). Detecting the loss of the internal repeat requires direct DNA sequencing of the fragments, particularly when the repeat originally composed only a small fraction of the entire fragment length. We did not find evidence for the loss of the repeat motif in the two loci sequenced in our study. The disruption of repeat motif by insertions or point mutations can alter a perfect repeat into a compound repeat or further breakdown of the complex repetitive motif. These changes in the motif influence the evolution and stability of these markers (Bacon et al. 2000; Symonds, Lloyd 2003). Both of the loci sequenced here exhibited disrupted repeat motifs. Likewise, the flanking regions were highly variable among our DNA sequences, with occasional large indels and numerous point mutations. We found limited evidence for null alleles in our sample. In general, all of these aspects of cross-specific amplification of microsatellite loci will increase the levels of homoplasy and homozygosity observed in the sample. Size homoplasy can be separated into two different types. Soft homoplasy can be described as PCR fragments with identical sizes but with different DNA sequences (Byrne et al. 1996; Adams et al. 2004; Curtu et al. 2004), indicating that direct sequencing of fragments can easily overcome this problem. Hard homoplasy, on the other hand, is a much more difficult problem and occurs when two fragments with identical numbers of repeats and flanking regions have different ancestries. This type of homoplasy is a problem for all highly polymorphic markers and essentially erases more and more information with time. Some studies have found that the size of fragments is not a good indication of genealogical history (Orti et al. 1997), while other studies have found that hard homoplasy has a little impact on diversity estimates and homology of fragment sizes (Dettman, Taylor 2004; Navascues, Emerson 2005). In general, size homoplasy decreases the power of microsatellites with increasing genealogical distance and therefore will almost always provide an underestimate of gene flow and population structure.

Despite the simplicity of microsatellite DNA sequences, their evolution has proved to be difficult to model (Symonds, Lloyd 2003), even among individuals of very well known organisms like humans (Ellegren 2004). There is little evidence that a single model will explain the evolutionary dynamics of most microsatellite polymorphisms. Most evidence suggests that the evolutionary models should be locus-specific, drawn from the data itself using a null model approach. More sophisticated techniques for analysing these data are available and need to be more generally applied (Pearse, Crandall 2004). In our study, a clear geographical structure among the three populations was observed. The two northern popu-

lations obviously experience abundant gene flow and have very similar allelic frequencies at all loci, while the distant southern population appears to be quite unusual and distinct, where two loci (Shc02 and Shc03) show heterozygosity excesses while only one locus (Shc11) has a heterozygosity deficiency. The upland habitat is quite limited in Peninsular Malaysia but the genetic connection between the opposite ends of the mountain range appears to be strong.

S. platyclados showed relatively high levels of genetic variability within populations compared to that between populations and displayed little genetic differentiation among populations. Similar genetic patterns were observed for *S. leprosula* populations that showed high levels of genetic diversity and most of the diversity was partitioned within populations (Lee et al. 2000b). Weak differentiation or spatial structure observed between three populations (based on pairwise comparisons of F_{st} values) could be attributed to contiguous distribution of the populations. This close geographical affinity suggested a high rate of gene flow through pollen or seed dispersal. Konuma et al. (2000) suggested that long-distance gene flow and seed migration are responsible for the poorly developed genetic structure of *Neobalanocarpus heimii*.

The effects of selective logging on the structure and diversity of these forests have been carefully studied (Cannon et al. 1994; Ho et al. 2004) and our knowledge of the effects on genetic diversity is expanding with increasing research (Wickneswari et al. 2000; Lee et al. 2002a,b; Ng et al. 2013). From our simulation studies, the selection of mother trees clearly has a major impact on the future genetic diversity of the remnant forest. The difference between the best and worst outcomes for allelic diversity is most divergent at intermediate levels of timber harvest (at high levels, both scenarios are bad) while the range of potential heterozygosity of offspring becomes increasingly divergent with increasing harvest. Given harvest levels of roughly 60–70% currently practiced, genetic information prior to felling and selection of mother trees could provide a powerful tool to maximize the genetic flexibility and diversity of the remnant forest. Cross-specific amplification of established microsatellite loci can be an effective strategy in the exploration of previously unstudied species but it is obvious that the transfer of these loci between species is not without its complications. To verify the conservation of repeat motifs and the frequency of disruption of these motifs and mutational changes in the flanking regions which might influence the simple interpretation of fragment size, a certain level of direct sequencing of selected loci should be performed (Goldstein, Pollock 1997). In our study, relatively few sequences produced several examples of complicated and homoplasious evolution. The overall patterns of genetic diversity and population structure, in any microsatellite study and particularly for cross-amplified loci, should be viewed as an underestimate of the reality.

CONCLUSIONS

In this study, we have demonstrated the successful cross-species amplification of several *Shorea* SSR loci within the genus, with immediate potential for the study of variation within and between populations of *S. platyclados*. Sequencing results show different repeat structure between the species, and that the maintenance of variation in allele size among species is complex, reflecting mutational processes in both the repeat and flanking regions. Sequence complexity and possibility of size homoplasy will complicate the interpretation of SSR variation. Studies employing cross-species amplification must therefore be accompanied by the knowledge of the underlying DNA sequence.

Acknowledgements

Funding for this study was received from Texas Tech University, in the form of start-up and Faculty Research Enhancement grants. We would like to express our greatest gratitude to Mr. C.T. Tan, CEO of the Perak Integrated Timber Complex, Malaysia, the Perak and Pahang Forestry Departments, and the Economic Planning Unit of the Malaysian Government for assisting us in the collection of timber tree samples. Extraction of total DNA was carried out at UKM's Forest Genetics Laboratory in the School of Environmental and Natural Resource Sciences, SSR analysis was provided by the Center for Biotechnology and Genomics at Texas Tech University who acknowledge TTU's Department of Plant and Soil Sciences' investment into the Reveal Mutation Discovery System.

References

Adams R.I., Brown K.M., Hamilton M.B. (2004): The impact of microsatellite electromorph size homoplasy on multilocus population structure estimates in a tropical tree (*Corythophora alta*) and an anadromous fish (*Morone saxatilis*). Molecular Ecology, ***13***: 2579–2588.

Ashton P.S. (1964): A quantitative phytosociological technique applied to tropical mixed rain forest vegetation. Malaysian Forester, ***27***: 304–317.

Ashton P.S. (1982): Dipterocarpaceae. Flora Malesiana, ***1***: 237–552.

Ashton M.S., Gunatilleke C.V.S., Singhakumara B.M.P., Gunatilleke I. (2001): Restoration pathways for rain forest in southwest Sri Lanka: a review of concepts and models. Forest Ecology Management, ***154***: 409–430.

Avise J.C. (1994): Molecular Markers, Natural History and Evolution. New York Chapman & Hall: 511.

Bacon A.L., Farrington S.M., Dunlop M.G. (2000): Sequence interruptions confer differential stability at microsatellite alleles in mismatch repair-deficient cells. Human Molecular Genetics, ***9***: 2707–2713.

Baillie I.C., Ashton P.S., Court M.N., Anderson J.A.R., Fitzpatrick E.A., Tinsley J. (1987): Site characteristics and the distribution of tree species in Mixed Dipterocarp Forest on Tertiary sediments in central Sarawak, Malaysia. Journal of Tropical Ecology, ***3***: 201–231.

Brookfield J.F.Y. (1996): A simple new method for estimating null allele frequency from heterozygote deficiency. Molecular Ecology, ***5***: 453–455.

Bruijnzeel L.A. (2004): Hydrological functions of tropical forests: not seeing the soil for the trees? Agriculture, Ecosystem and Environment, ***104***: 185–228.

Byrne M., Marquez-Garcia M.I., Uren T., Smith D.S., Moran G.F. (1996): Conservation and genetic diversity of microsatellite loci in the genus *Eucalyptus*. Australian Journal of Botany, ***44***: 331–341.

Cannon C.H., Leighton M. (2004): Tree species distributions across five habitats in a Bornean rain forest. Journal of Vegetative Science, ***15***: 257–266.

Cannon C.H., Peart D.R., Leighton M., Kartawinata K. (1994): The structure of lowland rainforest after selective logging in West Kalimantan, Indonesia. Forest Ecology and Management, ***67***: 49–68.

Charlesworth B., Sniegowski P., Stephan W. (1994): The evolutionary dynamics of repetitive DNA in eukaryotes. Nature, ***371***: 215–220.

Curtu A.L., Finkeldey R., Gailing O. (2004): Comparative sequencing of a microsatellite locus reveals size homoplasy within and between European oak species (*Quercus* spp.). Plant Molecular Biology Reporter, ***22***: 339–346.

Dettman J.R., Taylor J.W. (2004): Mutation and evolution of microsatellite loci in neurospora. Genetics, ***168***:1231–1248.

Doyle J.J., Doyle J.L. (1987): Rapid DNA isolation procedure for small quantities of fresh leaf tissue. Phytochemical Bulletin, ***19***: 11–15.

Ellegren H. (2004): Microsatellites: Simple sequences with complex evolution. Nature Reviews in Genetics, ***5***: 435–445.

Goldstein D.B., Pollock D.D. (1997): Launching microsatellites: A review of mutation processes and methods of phylogenetic inference. Journal of Heredity, ***88***: 335–341.

González-Martínez S.C., Robledo-Arnuncio J.J., Collada C., Díaz A., Williams C.G., Alía R., Cervera M.T. (2004): Cross-amplification and sequence variation of microsatellite loci in Eurasian hard pines. Theoretical Applied Genetics, ***109***: 103–111.

Gupta P.K., Balyan I.S., Sharma P.C., Ramesh B. (1996): Microsatellites in plants: A new class of molecular markers. Current Science, ***70***: 45–54.

Gutierrez M.V., Patto M.C.V., Huguet T., Cubero J.I., Moreno M.T., Torres A.M. (2005): Cross-species amplification of *Medicago truncatula* microsatellites across three major pulse crops. Theoretical Applied Genetics, ***110***: 1210–1217.

Ho W.S., Wickneswari R., Noor S.M., Clyde M.M. (2004): The effects of the timing and method of logging on forest structure in Peninsular Malaysia. Forest Ecology and Management, ***203***: 209–228.

Kaemmer D., Fischer D., Jarret R.L., Baurens F.C., Grapin A.D., Dambier D., Noyer J.L., Lanaud C., Kahl G., Lagoda P.J.L. (1997): Molecular breeding in the genus *Musa*: A strong case for STMS marker technology. Euphytica, ***96***: 49–63.

Konuma A., Tsumura Y., Lee C.T., Lee S.L., Okuda T. (2000): Estimation of gene flow in the tropical-rainforest tree *Neobalanocarpus heimii* (Dipterocarpaceae), inferred from paternity analysis. Molecular Ecology, ***9***: 1843–1852.

Lagercrantz U., Ellegren H., Andersson L. (1993): The abundance of various polymorphic microsatellite motifs differs between plants and vertebrates. Nucleic Acids Research, ***21***: 1111–1115.

Lee S.L., Tani N., Ng K.K.S., Tsumura Y. (2004): Isolation and characterization of 20 microsatellite loci for an important tropical tree *Shorea leprosula* (Dipterocarpaceae) and their applicability to *S. parvifolia*. Molecular Ecology Notes, ***4***: 222–225.

Lee S.L., Wickneswari R., Mahani M.C., Zakri A.H. (2000): 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.

Murawski D.A., Nimal-Gunatilleke I.A.U., Bawa K.S. (1994): The effects of selective logging on inbreeding in *Shorea megistophylla* (Dipterocarpaceae) from Sri Lanka. Conservation Biology, ***8***: 997–1002.

Navascues M., Emerson B.C. (2005): Chloroplast microsatellites: Measures of genetic diversity and the effect of homoplasy. Molecular Ecology, ***14***: 1333–1341.

Ng C.H., Lee C.T., Lee S.L., Tnah L.H., Ng K.K.S. (2013): Isolation and characterization of micorsatellite markers for *Shorea platyclados* (Dipterocarpaceae). Applications in Plant Sciences, ***1***: 1–3.

Obayashi K., Tsumura Y., Ihara-Ujino T., Niiyama K., Tanouchi H., Suyama Y., Washitani I., Lee C.T., Lee S.L., Muhammad N. (2002): Genetic diversity and

outcrossing rate between undisturbed and selectively logged forests of *Shorea curtisii* (Dipterocarpaceae) using microsatellite DNA analysis. International Journal of Plant Sciences, ***163***: 151–158.

Orti G., Pearse D.E., Avise J.C. (1997): Phylogenetic assessment of length variation at a microsatellite locus. Proceedings of National Academy of Sciences, ***94***: 10745–10749.

Pearse D.E., Crandall K.A. (2004): Beyond F-ST: Analysis of population genetic data for conservation. Conservation Genetics, ***5***: 585–602.

Potts M.D., Kassim A.R., Supardi M.N.N., Tan S., Bossert W.H. (2005): Sampling tree diversity in Malaysian tropical forests: an evaluation of a pre-felling inventory. Forest Ecology Management, ***205***: 385–395.

Shao Z., Lek S., Chang J. (2005): Complex mutation at a microsatellite locus in sturgeons: *Acipenser sinensis*, *A. schrenckii*, *A. gueldenstaedtii* and *A. baerii*. Journal of Applied Ichthylogy, ***21***: 2–6.

Sodhi N.S., Koh L.P., Brook B.W., Ng P.K.L. (2004): Southeast Asian biodiversity: an impending disaster. Trends in Ecology & Evolution, ***19***: 654–660.

Stacy E.A., Dayanandan S., Dancik B.P., Khasa P.D. (2001): Microsatellite DNA markers for the Sri Lankan rainforest tree species, *Shorea cordifolia* (Dipterocarpaceae), and cross-species amplification in *S. megistophylla*. Molecular Ecology Notes, ***1***: 53–54.

Steinkellner H., Fluch S., Turetschek E., Lexer C., Streiff R., Kremer A., Burg K., Gloessl J. (1997): Identification and characterization of (GA/CT) sub (n)-icrosatellite loci from *Quercus petraea*. Plant Molecular Biology, ***33***: 1093–1096.

Symonds V.V., Lloyd A.M. (2003): An analysis of microsatellite loci in *Arabidopsis thaliana*: Mutational dynamics and application. Genetics, ***165***: 1475–1488.

Tautz D. (1989): Hypervariability of simple sequences as a general source for polymorphic DNA markers. Nucleic Acids Research, ***17***: 6463–6471.

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

Weising K., Nybom H., Wolfe K., Kahl G. (2005): DNA Fingerprinting in Plants: Principles, Methods and Applications. 2nd Ed. Boca Raton, Taylor & Francis Group: 472.

Wickneswari R., Lee C.T., Muhammad N., Boyle T.J.B. (2000): Impact of logging on genetic diversity in humid tropical forests. Forest Genetics and Sustainability, ***63***: 171–181.

Wright T.F., Johns P.M., Walters J.R., Lerner A.P., Swallow J.G., Wilkinson G.S. (2004): Microsatellite variation among divergent populations of stalk-eyed flies, genus *Cyrtodiopsis*. Genetics Research, ***84***: 27–40.

Wu K.S., Tanksley S.D. (1993): Abundance, polymorphism and genetic-mapping of microsatellites in rice. Molecular and General Genetics, ***241***: 225–235.

Yeh F.C., Yang R.C., Boyle T.B.J., Ye Z.H., Mao J.X. (1997): POPGENE, the user friendly shareware for population genetic analysis. Molecular Biology and Biotechnology Centre, University of Alberta. Available at http://www.ualberta.ca/~fyeh

Zane L., Bargelloni L., Patarnello T. (2002): Strategies for microsatellite isolation: a review. Molecular Ecology, ***11***: 1–16.

Received for publication October 18, 2013
Accepted after corrections November 25, 2013

Corresponding author:

Asif Muhammad Javed, PhD, Texas Tech University, Department of Biological Sciences, Lubbock, TX 79409, USA; e-mail: asifjaved50603@yahoo.com