Demographic history and interspecific hybridization of four *Shorea* species (Dipterocarpaceae) from Peninsular Malaysia inferred from nucleotide polymorphism in nuclear gene regions

Hiroko Ishiyama, Nobuyuki Inomata, Tsuneyuki Yamazaki, Nor Aini Ab Shukor, and Alfred E. Szmidt

Abstract: *Shorea acuminata* Dyer, *Shorea curtisii* Dyer ex King, *Shorea leprosula* Miq., and *Shorea parvifolia* Dyer are dominant tree species in the tropical rainforest of Peninsular Malaysia, which experienced several climatic changes during Pleistocene. To investigate the current population structure and demographic history of these species, we analyzed levels and patterns of nucleotide polymorphism of the nuclear gene region *PgiC*. We also used sequence data of the *GapC* gene region obtained in our previous study. Negative Tajima's *D* values observed in both investigated gene regions for *S. curtisii*, *S. leprosula*, and *S. parvifolia* implied that all three species have experienced population expansion events. Little to moderate levels of population differentiation in *S. acuminata* and *S. curtisii* suggested recent divergence of the investigated populations after postglacial colonization of the Peninsular Malaysia. In addition, some haplotypes were similar or identical to haplotypes of the other species. The existence of such haplotypes could be partially explained by interspecific hybridization. Indeed, we found some putative hybrid individuals. Interspecific hybridization among closely related species might have contributed to the polymorphism of the investigated species.

Résumé : *Shorea acuminata* Dyer, *Shorea curtisii* Dyer ex King, *Shorea leprosula* Miq. et *Shorea parvifolia* Dyer constituent des espèces d'arbres dominantes dans la forêt humide tropicale de la péninsule malaisienne, qui a connu plusieurs changements climatiques durant le Pléistocène. Afin d'étudier la structure de populations actuelle et l'histoire démographique de ces espèces, les auteurs ont analysé les niveaux et patrons de polymorphisme nucléotidique de la région génique nucléaire *PgiC*. Ils ont également utilisé les données de séquences de la région génique *GapC* provenant d'une étude précédente. Les valeurs *D* de Tajima négatives, observées dans les deux régions géniques étudiées chez *S. curtisii*, *S. leprosula* et *S. parvifolia*, impliquent que ces trois espèces ont connu des épisodes d'expansion de population. Des degrés de différenciation de populations allant de faibles à modérés ont été observés chez *S. acuminata* et *S. curtisii*, indiquant qu'une divergence récente des populations étudiées est survenue après la colonisation postglaciaire de la péninsule malaisienne. De plus, certains haplotypes étaient similaires ou identiques aux haplotypes d'autres espèces. L'existence de ces haplotypes pourrait s'expliquer en partie par l'hybridation interspécifique, puisque des individus vraisemblablement hybrides ont été observés. L'hybridation interspécifique entre espèces rapprochées pourrait avoir contribué au polymorphisme chez les espèces étudiées.

[Traduit par la Rédaction]

Introduction

Evaluating the levels and patterns of nucleotide polymorphism in coding regions of tree species is currently a powerful tool to detect natural selection on some traits (Savolainen and Pyhäjärvi 2007), as well as to infer the demographic history and population structure (Kado et al. 2003; Ingvarsson 2005). Furthermore, the information about nucleotide polymorphism in natural populations of forest trees is very useful for the development of strategies for conservation of genetic resources and sustainable utilization of timber trees. Yet, in spite of considerable ecological and economic importance and specific characteristics of the tropical tree species, previous studies of genetic variation in their populations were based on molecular markers (Harada et al. 1994; Lee et al. 2000; Cao et al. 2006), while nucleotide polymorphism in nuclear coding regions was seldom investigated (Ishiyama et al. 2003).

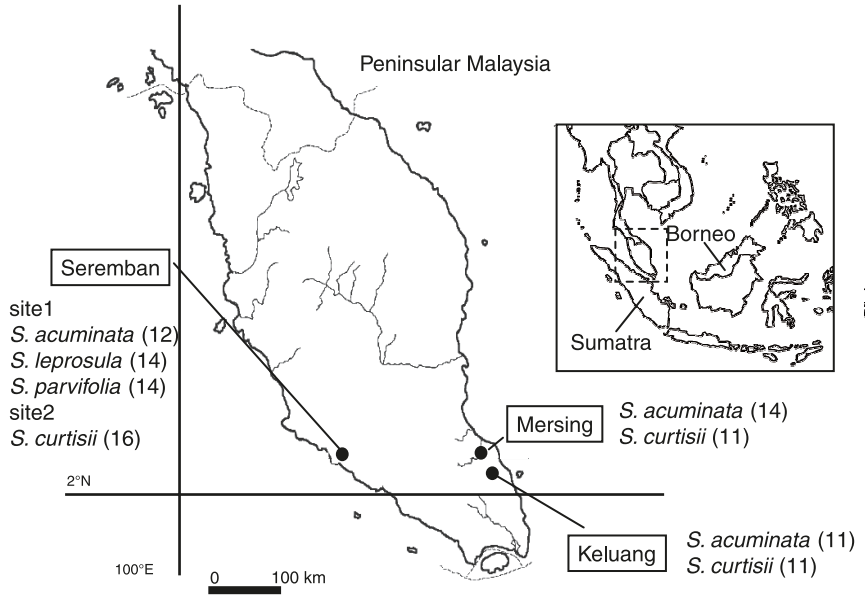
Received 24 June 2007. Accepted 10 November 2007. Published on the NRC Research Press Web site at cjfr.nrc.ca on 25 April 2008.

H. Ishiyama, N. Inomata, and A.E. Szmidt.[1] Department of Biology, Faculty of Sciences, Kyushu University, 6-10-1 Hakozaki, Higashi-ku, Fukuoka 812-8581, Japan.
T. Yamazaki. Research Institute of Evolutionary Biology, 2-4-28 Kamiyoga, Setagaya-ku, Tokyo 158-0098, Japan.
N.A.A. Shukor. Faculty of Forestry, University Putra Malaysia, 43400 UPM Serdang, Selangor Darul Ehsan, Malaysia.

[1]Corresponding author (e-mail: aszmiscb@mbox.nc.kyushu-u.ac.jp).

 doi:10.1139/X07-218

Fig. 1. Map of Southeastern Asia (inset) and sampling locations in Peninsular Malaysia. Number of individuals sampled from each location is indicated in parentheses.



Tropical tree species in Southeastern Asia are characterized by low density, exceptionally high number and synchronized flowering at irregular intervals (Ashton et al. 1988; Appanah 1993). The genus *Shorea* (family Dipterocarpaceae) represents a major component of canopy trees in Southeastern Asian rainforest. It contains approximately 190 species (Ashton 1982) and is found only in Southern and Southeastern Asia (Symington 1943). The four species investigated in this study (*Shorea acuminata* Dyer, *Shorea curtisii* Dyer ex King, *Shorea leprosula* Miq., and *Shorea parvifolia* Dyer) are typical representatives of canopy trees of the lowland rainforests in Peninsular Malaysia, Sumatra, and, except for *S. acuminata*, Borneo islands. They belong to the same section; *Mutica* (Ashton 1982) and timber group; Red Meranti (Symington 1943).

The climatic history of Peninsular Malaysia and of the neighboring islands is relatively complex and unique. At the Last Glacial Maximum (approximately 20 000 years ago), Peninsular Malaysia was covered by savanna, and rainforest refugia were present only on Borneo, Sumatra, and the Mentawai islands (Thomas 2000; Gathorne-Hardy et al. 2002). At that time, Peninsular Malaysia, Sumatra, Java, and Borneo islands were connected by the exposed Sunda shelf (Morley 2000; Voris 2000). Therefore, the extant tropical tree species in Peninsular Malaysia have expanded into this region only recently, after the last glacial period. The eastern and the western coast of Peninsular Malaysia are facing Borneo and Sumatra islands, respectively, which could represent independent and isolated sources of recolonization. Independent expansions from these sources could have led to population differentiation between eastern and western areas of Peninsular Malaysia.

In the previous study, we investigated nucleotide polymorphism of the *GapC* gene region in the four *Shorea* species included in the present study (Ishiyama et al. 2003). The results revealed that polymorphism of the investigated species was lower than that reported for herbaceous plants, and similar or lower than that of other woody species. In addition, we found that some individuals had unusual haplotypes consisting of sequences from different species, suggesting the possibility of interspecific hybridization (Ishiyama et al. 2003). Allozyme and (or) DNA polymorphism of the *Pgi* (glucose-6-phosphate isomerase, EC.5.3.1.9) in animals and the *PgiC* (cytosolic glucose-6-phosphate isomerase, EC.5.3.1.9) in plants have been well studied. Natural selection was suggested to act on this locus in some species: butterfly *Colias*, Watt (1977) and Wheat et al. (2006); *Leavenworthia*, Filatov and Charlesworth (1999); and *Arabidopsis thaliana* (L.) Heynh., (Kawabe et al. 2000). To substantiate our previous findings and to infer evolutionary features, including demographic history, population structure, and natural selection in *Shorea* species in Peninsular Malaysia, we studied levels and patterns of nucleotide polymorphism of a partial region of the *PgiC* gene. We analyzed *PgiC* sequence data together with those of the *GapC* gene region investigated in our previous study (Ishiyama et al. 2003).

Materials and methods

Plant materials

The locations of the investigated populations and the number of individuals for each population are presented in Fig. 1. Leaf samples were collected from three populations of *S. acuminata* (Mersing, Keluang, and Seremban), three populations of *S. curtisii* (Mersing, Keluang, and Seremban), one population of *S. leprosula* (Seremban), and one population of *S. parvifolia* (Seremban). Mersing and Keluang are located in the eastern, while Seremban is located in the western areas of Peninsular Malaysia, respectively (herein-

after referred to as eastern and western areas). Eleven to 16 individuals were randomly sampled from each population (Fig. 1). Most of these individuals were the same as those used in the study of the *GapC* gene (Ishiyama et al. 2003). Additionally, one leaf sample of *Shorea maxwelliana* King belonging to the Balau timber group (Symington 1943) was collected from the arboretum of the Forest Research Institute Malaysia. This species was used as an outgroup in some analyses. Species identification in the field was done by a taxonomist, and then herbarium specimens were examined again by an independent taxonomist.

DNA isolation, amplification, and sequencing

Genomic DNA was isolated from ~300 mg of leaves with a modified cetyltrimethylammonium bromide (CTAB) method (Murray and Thompson 1980) and used as template for amplification. A partial region (1248–1260 bp in length) of the putative *PgiC* gene, homologous to exons 13 through 19 of the *PgiC* gene of *A. thaliana* was amplified. The polymerase chain reaction (PCR) primers were designed for the conserved region of the *PgiC* gene of *A. thaliana* and *Dioscorea tokoro* Makino and used for amplification and sequencing of the corresponding region of *Shorea*. Based on the obtained sequences, the following specific forward and reverse primers were designed: F3, 5′-CATTTCTATTCAGCACCTTT-3′ (in the exon 13) and R4, 5′-ATGAGATGCTGTGGAACATTCTC-3′ (in the exon 19), respectively (T. Kado, personal communication). Amplification was performed using 35 cycles of denaturation at 94 °C for 30 s, annealing at 55 °C for 30 s, and extension at 72 °C for 2 min. Amplification products were purified using Wizard SV gel and PCR Clean-Up System kit (Promega). The purified products were directly sequenced using ABI Prism 3100 automatic sequencer (Applied Biosystems). Both strands were sequenced using the following six primers: F3, 5′-CATTTCTATTCAGCACCTTT-3′; F5, 5′- TTTGCTCCACACATTCAACAGG-3′; F7, 5′-GACGTGTAATTCCATGTGATTT-3′; R4, 5′-ATGAGATGCTGTGGAACATTCTC-3′; R6, 5′-CTTTAAGGTGGATAGGTTGCTG-3′; and R8, 5′-CCATTACTCTCCATGCTTACCT-3′.

We determined sequences of both haplotypes for each individual. When the sequences obtained by direct sequencing had no or only one heterozygous site, we could infer sequences of both haplotypes of an individual. On the other hand, when two or more heterozygous sites or indels were detected by direct sequencing, purified amplification products were cloned into the pGEM T-easy vector (Promega). Individual clones were sequenced using primers designed for the promoter sites of the vector, T7, 5′-TAATACGACTCACTATAGGG-3′ and SP6, 5′- TATTTAGGTGACACTATAG-3′. To obtain the sequence of the first haplotype of a heterozygous individual, sequencing of individual clones was carried out until two identical sequences were obtained. This procedure was repeated using additional clones to obtain the sequence of the second haplotype. Thus, at least four clones were sequenced for each individual with two or more heterozygous sites or indels.

To identify the maternal parent of the putative interspecific hybrids, we determined sequences of their *trnL*-intron and *trnL-trnF* spacer regions of the chloroplast genome. Furthermore, the sequences of *trnL*-intron and *trnL-trnF* spacer regions of at least two additional individuals from each of the four investigated *Shorea* species were determined. We assumed that, similar to most other angiosperms (Corriveau and Coleman 1988), chloroplast genome is maternally inherited in *Shorea* species. Amplification of the *trnL*-intron and *trnL-trnF* spacer regions was performed using universal primers developed by Taberlet et al. (1991) in the same conditions as those used for the *PgiC* gene region. Sequences of both strands were obtained by direct sequencing using the same primer sets. Sequences of the *GapC* gene region for the outgroup species *S. maxwelliana* were obtained as described by Ishiyama et al. (2003). Sequences of the *PgiC* gene region of *S. maxwelliana* were obtained from the GenBank (accession numbers: AB189522, AB1895223) (Kamiya et al. 2005). All sequences obtained in the present study have been deposited in the DDBJ/EMBL/GenBank databases under accession Nos. AB212414-AB212621 and AB368855-AB368898.

Data analyses

DNA sequences were verified, and contigs were assembled using the ATGC program ver. 4 (Genetyx Corporation). Multiple sequence alignment was done using the ClustalX program (Thomson et al. 1997) and corrected manually. We also used sequences of the *GapC* gene region obtained in our previous study (Ishiyama et al. 2003).

To assess the level of nucleotide polymorphism, the nucleotide diversity (π; Nei 1987) was calculated. The minimum number of recombination events (R_M; Hudson and Kaplan 1985) was calculated using a four-gamete test. The extent of nucleotide divergence between species was estimated as the average number of nucleotide substitutions per site between pairs of haplotypes from two species (K; Nei 1987) with Jukes and Cantor correction (Jukes and Cantor 1969). Sequence data for all populations of individual species were pooled when K was calculated. To investigate the degree of population differentiation, fixation index (F_{ST}; Hudson et al. 1992) was estimated. The significance of population differentiation (F_{ST} statistic) was tested using permutation test with 10 000 replicates. The permutation test for the F_{ST} statistic was performed using ProSeq version 2.91 program (Filatov 2002). Since both demographic changes and selection affect neutrality test statistics, several neutrality tests were performed to check for departures from neutrality: Tajima's D test (Tajima 1989) for individual loci, Tajima's D test for multiple loci (multilocus test of neutrality), MK test (McDonald and Kreitman 1991), and HKA test (Hudson et al. 1987). All calculations (except for the permutation tests for F_{ST} statistic, multilocus test of neutrality for Tajima's D statistic, and HKA test) were performed using the DnaSP program version 4.00 (Rozas et al. 2003). The multilocus test of neutrality and HKA test were performed using the HKA program obtained from Jody Hey's web site (lifesci.rutgers.edu/~heylab/). In the multilocus test, P values for average Tajima's D statistic over two loci were obtained.

Haplotype networks were constructed for both gene regions to visualize the frequency of haplotypes and to infer relationships among haplotypes. The networks were constructed using median-joining algorithm (Bandelt et al. 1999) implemented in the program NETWORK, Version 4.2.0.1, obtained from the following website: www.fluxus-

Table 1. DNA polymorphism within *Shorea* species.

Locus	Species	Population	*n*	*S*	Hd	Total ($\pi_{\_total}$)	Intron ($\pi_{\_intron}$)	Silent ($\pi_{\_sil}$)	Synonymous ($\pi_{\_syn}$)	Replacement ($\pi_{\_rep}$)	R_M
PgiC	*S. acuminata*	Total	74	45	0.87	0.0087	0.0123	0.0112	0.0009	0.0016	1
(1253 bp)[a]		Mersing	28	34	0.81	0.0076	0.0108	0.0097	0.0000	0.0016	1
		Keluang	22	32	0.91	0.0100	0.0143	0.0132	0.0029	0.0008	0
		Seremban	24	34	0.86	0.0076	0.0105	0.0095	0.0000	0.0023	0
	S. curtisii	Total	76	42	0.86	0.0042	0.0046	0.0055	0.0142	0.0006	1
		Mersing	22	20	0.87	0.0041	0.0042	0.0055	0.0182	0.0000	0
		Keluang	22	18	0.82	0.0041	0.0045	0.0053	0.0133	0.0006	0
		Seremban	32	29	0.80	0.0044	0.0049	0.0056	0.0124	0.0009	1
	S. leprosula		28	27	0.84	0.0022	0.0021	0.0026	0.0060	0.0013	0
	S. parvifolia		28	32	0.85	0.0049	0.0064	0.0064	0.0065	0.0004	1
GapC[b,c]	*S. acuminata*	Total	62	23	na	0.0026	0.0025	0.0034	0.0101	0.0002	2
(806 bp)[a]		Mersing	28	14	na	0.0020	0.0021	0.0026	0.0062	0.0004	0
		Keluang	14	18	na	0.0037	0.0048	0.0049	0.0063	0.0000	0
		Seremban	20	12	na	0.0037	0.0034	0.0049	0.0169	0.0000	2
	S. curtisii	Total	70	29	na	0.0034	0.0037	0.0045	0.0114	0.0000	1
		Mersing	24	8	na	0.0027	0.0022	0.0035	0.0143	0.0000	0
		Keluang	16	8	na	0.0026	0.0031	0.0035	0.0066	0.0000	0
		Seremban	30	27	na	0.0041	0.0048	0.0055	0.0113	0.0000	0
	S. leprosula		28	23	na	0.0037	0.0040	0.0050	0.0119	0.0004	0
	S. parvifolia		28	22	na	0.0035	0.0041	0.0047	0.0093	0.0000	1

Note: *n*, number of sequences; *S*, total number of segregating sites; Hd: haplotype diversity (Nei 1987); π, nucleotide diversity (Nei 1987); R_M, minimum number of recombination events (Hudson and Kaplan 1985); na, not analyzed.

[a]Mean length of each locus.

[b]Ishiyama et al. 2003.

[c]The values for the *GapC* gene region are different from those reported by Ishiyama et al. (2003), because the previous analysis did not include putative hybrid individuals.

technology.com/index.htm. For the *GapC* gene region, direct sequence data were available for 31, 35, 14, and 14 individuals of *S. acuminata*, *S. curtisii*, *S. leprosula*, and *S. parvifolia*, respectively. However, haplotype sequences were available for only 19, 15, 13, and 13 individuals, respectively (Ishiyama et al. 2003). Therefore, for estimations of R_M and for the construction of the haplotype network, haplotypes of only these individuals were used.

To further check haplotype relationships among the haplotypes observed in this study and those of other *Shorea* species, the Neighbor–Joining (NJ) tree (Saitou and Nei 1987) was constructed. The NJ tree for the *PgiC* gene region was constructed including our sequences and those of 28 *Shorea* species from the Red Meranti timber group obtained by Kamiya et al. (2005). The NJ tree for the *GapC* region was not constructed, because the sequences for species other than those included in our study were not available. The NJ tree was constructed using the program MEGA4 (Tamura et al. 2007).

To infer the maternal parent of putative hybrids, sequences of the *trnL*-intron and *trnL-trnF* spacer regions of the chloroplast genome were compared with individuals of the four *Shorea* species included in this study and additional 50 *Shorea* species, which are currently available in the DDBJ/EMBL/GenBank databases.

Shorea maxwelliana was used as an outgroup species in the MK test, HKA test, and for constructing the NJ tree for the *PgiC* gene region. Alignment gaps were excluded in all analyses. Putative hybrids were included in all analyses. They were assigned to individual species based on morphological identification of fresh and herbarium samples.

Results

The levels of DNA polymorphism and divergence

A partial region of the *PgiC* gene, homologous to exons 13 through 19 of the *PgiC* gene of *A. thaliana*, was sequenced. Its total length ranged from 1248 to 1260 bp, including 414 bp coding regions and 834 through 846 bp introns depending on the occurrence of indels. In total, we obtained 74, 76, 28, and 28 sequences for *S. acuminata*, *S. curtisii*, *S. leprosula*, and *S. parvifolia*, respectively. The haplotype diversity (Hd), number of segregating sites (*S*), π, and R_M are summarized in Table 1. The numbers of distinct *PgiC* haplotypes found in *S. acuminata*, *S. curtisii*, *S. leprosula*, and *S. parvifolia* were 17, 22, 14, and 11, respectively. Individual haplotypes of *S. acuminata*, *S. curtisii*, *S. leprosula*, and *S. parvifolia* were designated as sapgi-1 through sapgi-17, scpgi-1 through scpgi-22, slpgi-1 through slpgi-14, and sppgi-1 through sppgi-11, respectively. The positions of segregating sites and the frequencies of haplotypes are available as supplementary material (Supplementary Fig. 1).[2] The numbers of distinct *GapC* haplotypes found

[2] Supplementary data for this article are available on the journal Web site (cjfr.nrc.ca) or may be purchased from the Depository of Unpublished Data, Document Delivery, CISTI, National Research Council Canada, Building M-55, 1200 Montreal Road, Ottawa, ON K1A 0R6, Canada. DUD 3738. For more information on obtaining material refer to cisti-icist.nrc-cnrc.gc.ca/cms/unpub_e.html.

Table 2. The average number of nucleotide substitutions per site between *Shorea* species at silent sites (K_{sil}) with the Jukes and Cantor correction.

Species comparisons		Locus	
Species 1	Species 2	*PgiC*	*GapC*[a]
S. acuminata	*S. curtsii*	0.0236	0.0142
S. acuminata	*S. leprosula*	0.0168	0.0215
S. acuminata	*S. parvifolia*	0.0118	0.0173
S. curtisii	*S. leprosula*	0.0245	0.0221
S. curtisii	*S. parvifolia*	0.0193	0.0216
S. leprosula	*S. parvifolia*	0.0123	0.0257

[a]Ishiyama et al. 2003.

in *S. acuminata*, *S. curtisii*, *S. leprosula*, and *S. parvifolia* were 10, 10, 8, and 11, respectively. Individual haplotypes for the *GapC* gene region in *S. acuminata*, *S. curtisii*, *S. leprosula*, and *S. parvifolia* were designated as sagap-1 through sagap-10, scgap-1 through scgap-10, slgap-1 through slgap-8, and spgap-1 through spgap-11, respectively.

In the *PgiC* gene region, nucleotide diversity within species for silent sites (π_{sil}) was highest in *S. acuminata* (0.0112) and lowest in *S. leprosula* (0.0026). The nucleotide diversity within species for replacement sites (π_{rep}) was highest in *S. acuminata* (0.0016) and lowest in *S. parvifolia* (0.0004). The values of both π_{sil} and π_{rep} in the *PgiC* gene region were slightly higher than those in the *GapC* gene region, except for *S. leprosula* (Table 1). The haplotype diversity in the *PgiC* gene region was relatively high (Hd > 0.8) in all four species (Table 1).

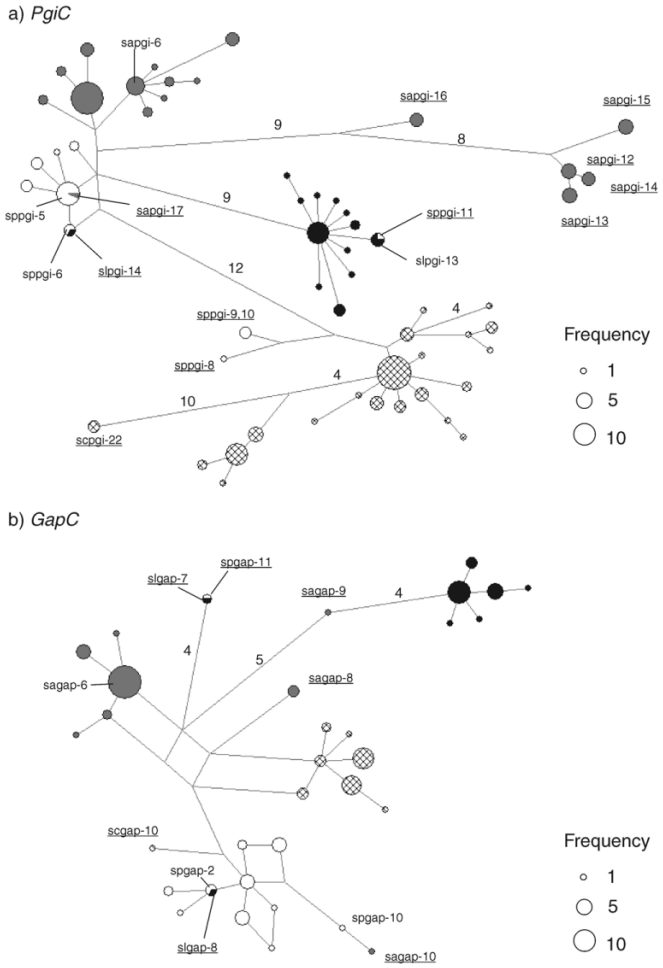
Estimates of the divergence between species at silent sites (K_{sil}) are presented in Table 2. The K_{sil} values were similar for the *PgiC* and *GapC* gene regions and ranged between 0.0118 and 0.0257.

Relationships among haplotypes

Haplotype networks for the *PgiC* and *GapC* gene regions are presented in Figs. 2*a* and 2*b*, respectively. In all investigated species, the number of distinct haplotypes was rather high, but the number of nucleotide differences between most of them was small. Furthermore, nearly half of the haplotypes were rare, that is, observed only once or twice. This was particularly prominent in the *PgiC* gene region of *S. curtisii* and *S. leprosula*. At both loci, most haplotypes belonging to individual species formed distinct, separate groups (Figs. 2*a* and 2*b*).

However, some haplotypes were diverged from typical haplotypes of corresponding species (Figs. 2*a* and 2*b*). Here we designate haplotypes that were either frequent (>10%) in a given species or differed from such haplotypes by fewer than four nucleotides as "typical" haplotypes of that species. In the *PgiC* gene region, the haplotypes sapgi-12 through sapgi-17 of *S. acuminata*, scpgi-22 of *S. curtisii*, slpgi-14 of *S. leprosula,* and sppgi-8 through sppgi-11 of *S. parvifolia* were diverged from the typical haplotypes of the corresponding species (Fig. 2*a*). In the *GapC* gene region, the haplotypes sagap-8 through sagap-10 of *S. acuminata*, scgap-10 of *S. curtisii*, slgap-7 and slgap-8 of *S. leprosula*, and spgap-11 of *S. parvifolia* were also diverged from the

Fig. 2. Haplotype networks for the four *Shorea* species. (*a*) haplotype network for the *PgiC* gene region; (*b*) haplotype network for the *GapC* gene region. The size of circles is proportional to the observed number of each haplotype. The branch lengths are proportional to the number of mutations. The number of mutations is indicated on branches when it was more than three. The shaded, crossed, solid, and open circles represent the haplotypes observed in *S. acuminata*, *S. curtisii*, *S. leprosula*, and *S. parvifolia*, respectively. The names of diverged haplotypes are underlined.



typical haplotypes (Fig. 2*b*). The K_{sil} between the typical haplotypes and the divergent haplotypes were comparable to the divergence between the four species investigated in this study. For instance, K_{sil} between the typical haplotype sapgi-6 and the diverged haplotype sapgi-15 was 0.0220. The diverged *PgiC* haplotypes sapgi-12 through sapgi-15 of *S. acuminata* formed a separate group (Fig. 2*a*). Haplotypes from this group accounted for 26% of all 74 haplotypes scored in this species. In addition, the diverged *PgiC* haplotypes sppgi-8 through sppgi-10 of *S. parvifolia* also formed a separate group (Fig. 2*a*). The remaining diverged haplotypes had low frequency (Figs. 2*a* and 2*b*).

Some of the diverged haplotypes were identical or similar to haplotypes from different species. For instance, the haplotype sapgi-17 of *S. acuminata* was identical to a typical haplotype of *S. parvifolia* (sppgi-5) (Fig. 2*a*). The slpgi-14 haplotype of *S. leprosula* was identical to a typical haplotype of *S. parvifolia* (sppgi-6) (Fig. 2*a*). The sppgi-11 haplo-

type of *S. parvifolia* was identical to a typical haplotype of *S. leprosula* (slpgi-13) (Fig. 2*a*). The NJ tree for the *PgiC* gene region revealed that the diverged haplotype group from *S. acuminata* (sapgi-12 through sapgi-15) clustered with the haplotypes of *S. quadrinervis* with high bootstrap value (92%) (Supplementary Fig. 2). They differed from the haplotypes of *S. quadrinervis* by only one to two nucleotides. In addition, the diverged haplotype sapgi-16 of *S. acuminata* was partially (positions 640–1236) similar to haplotypes of *S. quadrinervis* with one to three nucleotide differences. In contrast, its remaining part (positions 1–632) was similar to the typical haplotypes of *S. acuminata* with two to four nucleotide differences (Supplementary Fig. 3*a*). Similarly, the diverged haplotype scpgi-22 of *S. curtisii* showed partial (positions 90–699) similarity to the corresponding part of haplotypes of *S. pauciflora* with three or four nucleotide differences. On the other hand, the remaining part of this haplotype (positions 797–1261) was similar to the typical haplotypes of *S. curtisii* (scpgi-7 through scpgi-10) with only one to two nucleotide differences (Supplementary Fig. 3*b*). In the NJ tree, the scpgi-22 haplotype from *S. curtisii* was clustered with the haplotypes of *S. kunstleri* and *S. pauciflora* with high bootstrap value (92%) (Supplementary Fig. 2). In the *GapC* gene region, the 3′ part of the diverged haplotype sagap-9 of *S. acuminata* was identical to the 3′ part of a typical haplotype of *S. leprosula*. However, the 5′ part of this haplotype was identical to the 5′ part of a typical haplotype of *S. acuminata* (Ishiyama et al. 2003). The three haplotypes (sapgi-16, scpgi-22, and sagap-9) appeared to be recombinants between typical haplotypes and haplotypes of different species.

Putative hybrids

In total, we found 32 individuals that harbored the diverged haplotypes, which were different from the typical haplotypes of the corresponding species identified based on morphology. Their genotypes are summarized in Table 3. Hybrids are expected to have genotypes consisting of sequences from two different species in different regions of the nuclear genome. Three of the 32 individuals harbored diverged haplotypes at both loci. The individual sa399, identified as *S. acuminata* both in the field and in the herbarium, had one typical *PgiC* haplotype (sapgi-6) of *S. acuminata*. The other haplotype of this individual (sapgi-17) was identical to a typical haplotype of *S. parvifolia* (sppgi-5). Similarly, in the *GapC* gene region, the individual sa399 had one typical haplotype of *S. acuminata* (sagap-6), while the other haplotype (sagap-10) was similar to a typical haplotype of *S. parvifolia* (spgap-10) with one nucleotide difference. The chloroplast DNA haplotype of the individual sa399 was identical to *S. parvifolia*. The second individual (sl326), identified in the field as *S. leprosula*, had one typical *PgiC* haplotype (slpgi-13) of *S. leprosula*, while the other haplotype (slpgi-14) was identical to a typical haplotype of *S. parvifolia* (sppgi-6). In the *GapC* region, the individual sl326 had one haplotype (slgap-8), which was identical to a typical haplotype of *S. parvifolia* (spgap-2), while the other haplotype (slgap-7) was not similar to any of the typical haplotypes found in the four investigated species. The cpDNA haplotype of the individual sl326 was identical to *S. leprosula*. The third individual (sp296) identified in the field as *S. parvifolia* had one typical *PgiC* haplotype of *S. parvifolia* (sppgi-6), while the other haplotype (sppgi-11) was identical to a typical hapotype of *S. leprosula* (slpgi-13). In the *GapC* region, the individual sp296 had one typical haplotype of *S. parvifolia* (spgap-2) and one haplotype (spgap-11) that was not similar to any of the typical haplotypes found in the four investigated species. The cpDNA haplotype of the individual sp296 was identical to *S. leprosula*.

The remaining 29 individuals had diverged haplotypes in only one gene region (*PgiC* or *GapC*), and their cpDNA haplotypes were consistent with the morphological identification (Table 3).

Table 3. Genotypes of individuals with diverged haplotypes.

Individual name	Morphological identification	Genotype		
		PgiC	*GapC*[a]	cpDNA
sa399	*Shorea acuminata*	A/P	A/P	P
sa88		A/d1	A/A	A
sa91		A/d1	A/A	A
sa341		A/d1	A/A	A
sa345		A/d1	—	A
sa346		A/d1	A/A	A
sa393		A/d1	—	A
sa394		A/d1	A/A	A
sa395		A/d1	—	A
sa396		A/d1	—	A
sa397		A/d1	—	A
sa398		A/d1	A/A	A
sa402		A/d1	A/A	A
sa403		A/d1	A/A	A
sa342		d1/d1	A/A	A
sa401		d1/d1	A/A	A
sa83		rec1/d1	A/A	A
sa90		rec1/d1	A/A	A
sa85		A/A	A/rec2	A
sa86		A/A	A/d2	A
sa347		A/A	A/d2	A
sa350		A/A	A/d2	A
sc261	*S. curtisii*	C/rec3	C/C	C
sc271		C/rec3	C/C	C
sc274		C/rec3	-	C
sc275		C/C	C/d3	C
sl326	*S. leprosula*	P/L	P/d4	L
sp282	*S. parvifolia*	P/d5	P/P	P
sp283		P/d5	P/P	P
sp287		P/d5	P/P	P
sp291		P/d5	P/P	P
sp296		P/L	P/d4	L

Note: A, typical haplotypes of *S. acuminata*; C, typical haplotypes of *S. curtisii*; L, typical haplotypes of *S. leprosula*; P, typical haplotypes of *S. parvifolia*; d1, diverged haplotypes, similar to *S. quadrinervis* (sapgi-12 to 15); d2, diverged haplotype that differed from typical haplotypes of the four *Shorea* species included in this study (sagap-8); d3, diverged haplotype that differed from typical haplotypes of the four *Shorea* species included in this study (scgap-10); d4, diverged haplotype that differed from typical haplotypes of the four *Shorea* species included in this study (slgap-7 or spgap-11); d5, diverged haplotype that differed from typical haplotypes of the four *Shorea* species included in this study (sppgi-8 through sppgi-10); rec1, recombinant-like haplotype between *S. acuminata* and *S. quadrinervis* (sapgi-16); rec2, recombinant-like haplotype between *S. acuminata* and *S. leprosula* (sagap-9); rec3, recombinant-like haplotype between *S. curtisii* and *S. pauciflora* (scpgi-22); —, not studied.

[a]Ishiyama et al. 2003.

Table 4. Summary of F_{ST} values for population differentiation.

Locus	Mersing–Keluang (E–E)	Mersing–Seremban (E–W)	Keluang–Seremban (E–W)
Shorea acuminata			
PgiC	0.095	–0.015	0.091
GapC[a]	–0.030	0.030	0.013
S. curtisii			
PgiC	–0.010	0.015	0.011
GapC[a]	0.073	0.083*	0.005

Note: E, popluation from the eastern area; W, population from the western area. *, significance after the Bonferroni correction for multiple comparisons at $P < 0.05$.

[a]Ishiyama et al. 2003

Table 5. Tajima's *D* statistic for the *PgiC* and *GapC* genes and multilocus test of neutrality.

Species	Population	*PgiC*	*GapC*[a]	Average over two loci[b]	
Shorea acuminata	Total	0.560	–1.925*	–0.683	(0.180)
	Mersing	0.322	–1.997*	–0.837	(0.123)
	Keluang	1.616	–1.940*	–0.162	(0.456)
	Seremban	0.156	–0.417	–0.131	(0.484)
S. curtisii	Total	–1.241	–1.706	–1.474	(0.008)
	Mersing	–0.233	0.020	–0.106	(0.498)
	Keluang	0.153	–0.415	–0.131	(0.475)
	Seremban	–0.852	–1.811*	–1.331	(0.019)
S. leprosula		–2.220**	–1.904*	–2.062***	(0.000)
S. parvifolia		–1.038	–1.768	–1.403	(0.013)

Note: Significant before the Bonferroni correction for multiple comparisons, *$0.01 \le P < 0.05$, **$0.001 \le P < 0.01$; significant after the Bonferroni correction for multiple comparisons, ***$0.001 \le P < 0.01$.

[a]Ishiyama et al. 2003.

[b]The values in parentheses indicate *P* values for multilocus test of neutrality before the Bonferroni correction.

Population differentiation within species

The obtained F_{ST} values for all pairs of populations of *S. acuminata* and *S. curtisii* are presented in Table 4. Based on the qualitative guidelines for interpretation of F_{ST} suggested by Wright (1978), moderate genetic differentiation (F_{ST} = 0.05–0.15) was found in the *PgiC* gene region for the following two pairs of populations of *S. acuminata*: Mersing–Keluang and Keluang–Seremban. In the *GapC* gene region, moderate genetic differentiation was found for the following two pairs of populations of *S. curtisii*: Mersing–Keluang and Mersing–Seremban. For other population comparisons, little genetic differentiation (F_{ST} = 0–0.05) or negative F_{ST} values were obtained. Negative F_{ST} values were probably due to small number of samples and (or) low levels of population differentiation. After the Bonferroni correction, a significant F_{ST} value ($P < 0.05$) was obtained for only one pair of populations of *S. curtisii*: Mersing–Seremban (for the *GapC* gene region). Moderate population differentiation was observed both within and between eastern and western areas of the Malaysian Peninsula. Namely, there was no clear relationship between the location of population (eastern or western) and the level of population differentiation. We did not find any pattern in the geographic distribution of the haplotypes. Same or similar haplotypes were shared among populations within *S. acuminata* and *S. curtisii* in both *PgiC* and *GapC* gene regions (data not shown). In summary, our results suggest relatively low levels of population differentiation in both *S. acuminata* and *S. curtisii* from Peninsular Malaysia.

Tests of neutrality

The values of Tajima's *D* statistic (hereinafter referred to as *D*) for the *PgiC* and *GapC* gene regions are summarized in Table 5. Significant negative *D* values were observed in the *PgiC* gene region of *S. leprosula* and in the *GapC* gene region of *S. acuminata* total and populations Mersing and Keluang of *S. acuminata*, population Seremban of *S. curtisii* and *S. leprosula*. None of these deviations were significant after the Bonferroni correction. In the multilocus test, average *D* value over two loci was significantly negative in *S. curtisii* total, population Seremban of *S. curtisii*, *S. leprosula*, and *S. parvifolia* (Table 5). After the Bonferroni correction, only deviation in *S. leprosula* was significant ($0.001 \le P < 0.01$).

In both gene regions and in all species, there was no significant difference in the ratio of the number of synonymous to replacement substitutions between polymorphism and divergence (MK test) (data not shown). Similarly, there was no significant difference in the ratio of polymorphism to divergence between the *PgiC* and *GapC* gene regions (HKA test) (data not shown).

Discussion

Polymorphism and divergence

The average nucleotide diversities (π) over the two loci in *S. acuminata*, *S. curtisii*, *S. leprosula*, and *S. parvifolia* were 0.0083, 0.0051, 0.0035, and 0.0058 at silent sites, and 0.0011, 0.0003, 0.0009, and 0.0003 at replacement sites, re-

Table 6. Comparison of the levels of nucleotide polymorphism (π) in various organisms.

Species	Family	No. of loci	$\pi_{\_sil}$	$\pi_{\_rep}$	References
Plants					
Woody species					
Leaf trees					
Shorea acuminata	Dipterocarpaceae	2	0.0083	0.0011	This study
Shorea curtisii	Dipterocarpaceae	2	0.0051	0.0003	This study
Shorea leprosula	Dipterocarpaceae	2	0.0035	0.0009	This study
Shorea parvifolia	Dipterocarpaceae	2	0.0058	0.0003	This study
Betula pendula	Betulaceae	1	0.0047	—[a]	Järvinen et al. 2003
Populus tremula	Salicaceae	5	0.0160	0.0059	Ingvarsson 2005
Populus trichocarpa	Salicaceae	Genome wide	0.0032	0.0010	Tuskan et al. 2006
Conifers					
Cryptomeria japonica	Taxodiaceae	7	0.0038	0.0009	Kado et al. 2003
Pinus densata	Pinaceae	7	0.0122	0.0028	Ma et al. 2006
Pinus tabuliformis	Pinaceae	7	0.0119	0.0030	Ma et al. 2006
Pinus yunnanensis	Pinaceae	7	0.0095	0.0023	Ma et al. 2006
Pinus taeda	Pinaceae	19	0.0064	0.0011	Brown et al. 2004
Pseudotsuga menziesii	Pinaceae	18	0.0106	0.0021	Krutovsky and Neale 2005
Herbaceous species					
Dicots					
Arabidopsis thaliana	Brassicaceae	876	0.0060[b]	0.0015[c]	Nordborg et al. 2005
Helianthus annuus (wild)	Asteraceae	9	0.0234	0.0034	Liu and Burke 2006
Monocots					
Oryza rufipogon	Poaceae	10	0.0072	—[a]	Zhu et al. 2007
Animals					
Homo sapiens	Hominidae	10 kb noncoding region	0.0007	—[a]	Zhao et al. 2006
Drosophila melanogaster	Drosophilidae	188	0.0106	0.0018	Andolfatto 2005

Note: π, nucleotide diversity (Nei 1987) based on multilocus sequences.

[a]Data not available.

[b]Approximate value for intergenic regions.

[c]Approximate value for nondegenerate sites.

spectively (Table 6). The levels of polymorphism at silent sites were in the range observed in a wide range of organisms, and moderate in comparison with other woody species. On the other hand, the observed levels of polymorphism at replacement sites were slightly lower than in other organisms (Table 6). The two enzymes encoded by the *PgiC* and *GapC* genes are crucial in glycolysis pathway, thus selective constraint against amino acid change in these genes may be strong. This could be one of the reasons for low levels of polymorphism at replacement sites observed in our study.

The levels of divergence between the four *Shorea* species included in our study were quite low. When K_{sil} values between the four *Shorea* species were compared with the π_{sil} values, K_{sil} values were on average only 3.6 times and 4.8 times higher than π_{sil} of individual species for the *PgiC* and *GapC* gene region, respectively. Therefore, our present results suggest that the species investigated in our study diverged recently. Alternatively, interspecific hybridization could be an additional cause for the observed low divergence (see below).

Interspecific hybridization

We found that 32 individuals harbored diverged haplotypes (Table 3). Some of them were similar to haplotypes of another species. The presence of such diverged haplotypes can be due to either ancestral polymorphism or interspecific hybridization, and it is usually very difficult to distinguish between these two alternatives. However, we suspect that at least some of the diverged haplotypes detected in our study are due to interspecific hybridization rather than ancestral polymorphism. For instance, some individuals had diverged haplotypes in both *GapC* and *PgiC* regions. The individual sa399 could be a hybrid between *S. acuminata* and *S. parvifolia*, and its maternal parent could be *S. parvifolia*, as inferred from cpDNA sequences. The other two individuals (sl326 and sp296) could be hybrids between *S. parvifolia* and *S. leprosula*, and their maternal parent could be *S. leprosula*, as inferred from cpDNA sequences. However, in the *GapC* gene region, these two individuals had an unknown haplotype, which differed from both *S. leprosula* and *S. parvifolia* haplotypes. Since in the areas where our samples were collected there are many other *Shorea* species, we could not rule out the possibility of hybridization between *S. parvifolia* and another species.

The remaining 29 individuals had diverged haplotypes in only one gene region (*PgiC* or *GapC*), and some of them appeared to represent recombinants between haplotypes of different species. For instance, sequences of haplotypes scpgic-22 and sagap-9 were partially similar or identical to haplotypes of different species. The presence of the recombinant-like haplotype sagap-9 appears to support the occurrence of interspecific hybridization. In the recombi-

nant-like haplotype sagap-9, the 5′ part was exactly the same as a typical haplotype of *S. acuminata*, while the 3′ part was exactly the same as a typical haplotype of *S. leprosula* (Ishiyama et al. 2003). Therefore, the recent recombination in a hybrid between *S. acuminata* and *S. leprosula* is a more plausible explanation than ancestral polymorphism for the structure of this haplotype. The minimum number of recombination events (R_M) was one or more for both gene regions except for *S. leprosula* (Table 1), which indicates the possibility of recombination at least within species.

Shorea species from the section *Mutica* are pollinated by the same insect groups (Appanah and Chan 1981; Momose et al. 1998; Sakai et al. 1999), and their flowering periods (Ashton et al. 1988) and habitats often overlap. Moreover, most species from the genus *Shorea* have the same chromosome number (n = 7) (Jong and Lethbridge 1967; Somego 1978). Thus, the occurrence of interspecific hybridization is possible. In fact, Chan (1981) showed successful fruit formation in a cross between *S. splendida* and *S. stenoptera*, further indicating the potential for natural hybridization. In this study, species identification was carried independently by two different taxonomists, and yet the putative hybrids between *S. leprosula* and *S. parvifolia*, sl326 and sp296, were repeatedly identified as *S. leprosula* and *S. parvifolia*, respectively. This indicates that based on morphology, the hybrids of *Shorea* species may be difficult to distinguish from their parental species.

Despite the evolutionary importance of hybridization (Stebbins 1966; Rieseberg et al. 1996) and recombination in hybrids, their role in the evolution of *Shorea* species have not been studied experimentally. Both hybridization and recombination between genomes of different species are far more efficient than mutation to introduce new alleles. Further studies are necessary to determine the importance of interspecific hybridization and recombination in the evolution of *Shorea* species.

Population differentiation and demographic history of Peninsular Malaysia populations

During the last glacial period (approximately 20 000 years ago), Peninsular Malaysia, where our samples were collected, is supposed to have been covered by savanna, and rainforest refugia were present only in the northern and eastern Borneo, northern and western Sumatra and the Mentawai Islands (Thomas 2000; Gathorne-Hardy et al. 2002; Slik et al. 2003). Therefore, the extant tropical tree species have colonized this region only recently after the last glacial period (no more than 20 000 years ago). Considering such climatic history, we expected that independent expansions from isolated refugia on Borneo and Sumatra could have led to population differentiation between populations from the eastern and western parts of Peninsular Malaysia. In addition, taking into account their pollinators (thrips (Appanah and Chan 1981) and small beetles (Momose et al. 1998; Sakai et al. 1999)) and the limited seed dispersal (20–50 m (Chan 1980)), we expected high population differentiation as a result of limited gene flow. However, the levels of population differentiation in *S. acuminata* and *S. curtisii* revealed in our study (F_{ST} = –0.030–0.095) were similar to those reported for wind-pollinated trees such as *Cryptomeria japonica* D. Don (F_{ST} = 0.011–0.040) (Kado et al. 2003), *Pinus densata* Mast., *Pinus tabuliformis* Carrière, and *Pinus yunnanensis* Franch. (F_{ST} = 0.031–0.105) (Ma et al. 2006), *Betula pendula* Roth (F_{ST} = –0.023–0.093) (Järvinen et al. 2003), and *Populus tremula* L. (F_{ST} = 0.117) (Ingvarsson 2005). The relatively low level of population differentiation observed in our study is likely to be due to a short time that passed since divergence of these populations, which is consistent with the recent colonization of Peninsular Malaysia. Furthermore, we did not find a clear relationship between the location of population (eastern or western) and the level of population differentiation. Thus, extant populations in Peninsular Malaysia might have originated from a single refugium. Alternatively, populations expanding from Borneo and Sumatra refugia could have fused throughout Peninsular Malaysia.

Both change in population size and natural selection will cause a departure from neutrality. We observed significantly negative D values in both loci and in the multilocus test (Table 5). Negative D values can be caused by recent population expansion (Tajima 1993) or by a selective sweep (Braverman et al. 1995). The presence of rare diverged haplotypes due to interspecific hybridization will also skew the D statistic towards a negative value, since they produce many segregating sites with low frequency. The pattern of reduced polymorphism compared with the divergence, which is expected under selective sweep, was not detected by HKA test. In addition, changes in population size are expected to affect all loci. Actually, negative D values were found at both loci in *S. curtisii*, *S. leprosula*, and *S. parvifolia*, and their average D values over two loci were significantly negative (Table 5). This result suggests that all three species experienced past population expansion, which is consistent with climatic history of Southeastern Asia. The presence of diverged haplotypes was also likely to contribute to the observed deviations from neutrality. On the other hand, we could not find significant evidence for a population expansion event in *S. acuminata*. The power of Tajima's D test depends on the elapsed time and strength of the expansion event (Simonsen et al. 1995; Ramos-Onsins and Rozas 2002), and the signature of population expansion cannot always be detected by this test. Thus, our present results do not exclude the possibility of such an event in this species.

Conclusions

Significant negative Tajima's D in the multilocus test for *S. curtisii*, *S. leprosula*, and *S. parvifolia* supported recent population expansion events in these species. In spite of the limited pollen and seed dispersal, differentiation among populations within *S. acuminata* and *S. curtisii* was low in most cases. This suggested recent diversification of populations in the Peninsular Malaysia, which is consistent with recent expansion. Furthermore, the observed genetic homogeneity of these populations suggested expansion from single refugial population or fusion of multiple refugial populations. Additional samples from Sumatra and Borneo islands, where refugia could have existed during the last glacial period, will help clarify the origins of populations in Peninsular Malaysia.

From the conservation viewpoint, maintaining genetic

variation is important. In the four *Shorea* species investigated in our study, many rare haplotypes were found and some of them carried substitutions at replacement sites. Rare haplotypes are easily lost following reduction of population size caused by logging.

Our present results confirmed our previous findings suggesting the occurrence of interspecific hybridization in *Shorea* species. Interspecific hybridization can introduce new alleles to a species and thus affect its evolution. Additional studies are necessary to evaluate the evolutionary importance of natural interspecific hybridization in this group of tropical trees.

Acknowledgement

We are grateful to the members of our department for technical support and useful advice during our study. We thank Tadashi Kajita for providing taxonomic and sample information and Tomoyuki Kado for providing primer sequences of the *PgiC* gene region. This study was supported by grants 11490026 and 12575002 to T.Y. and 13490022, 13575002, and 17405032 to A.E.S. from the Ministry of Education, Culture, Sports, Science and Technology of Japan.

References

Andolfatto, P. 2005. Adaptive evolution of non-coding DNA in Drosophila. Nature (London), **437**: 1149–1152. doi:10.1038/nature04107. PMID:16237443.

Appanah, S. 1993. Mass flowering of Dipterocarp forests in the aseasonal tropics. J. Biosci. **18**: 457–474. doi:10.1007/BF02703079.

Appanah, S., and Chan, H.T. 1981. Thrips: the pollinators of some Dipterocarps. Malay. For. **44**: 234–252.

Ashton, P.S. 1982. Dipterocarpaceae. *In* Flora Malesiana. series I, Spermatophyta. *Edited by* C.G.G.J. van Steenis. Noordhoff, Leiden. pp. 237–552.

Ashton, P.S., Givnish, T.J., and Appanah, S. 1988. Staggered flowering in the Dipterocarpaceae: New insights into floral induction and the evolution of mast fruiting in the aseasonal tropics. Am. Nat. **132**: 44–66. doi:10.1086/284837.

Bandelt, H.J., Forster, P., and Rohl, A. 1999. Median-joining networks for inferring intraspecific phylogenies. Mol. Biol. Evol. **16**: 37–48. PMID:10331250.

Braverman, J.M., Hudson, R.R., Kaplan, N.L., Langley, C.H., and Stephan, W. 1995. The hitchhiking effect on the site frequency spectrum of DNA polymorphisms. Genetics, **140**: 783–796. PMID:7498754.

Brown, G.R., Gill, G.P., Kuntz, R.J., Langley, C.H., and Neale, D.B. 2004. Nucleotide diversity and linkage disequilibrium in loblolly pine. Proc. Natl. Acad. Sci. U.S.A. **101**: 15255–15260. doi:10.1073/pnas.0404231101. PMID:15477602.

Cao, C.-P., Finkeldey, R., Siregar, I., Siregar, U., and Gailing, O. 2006. Genetic diversity within and among populations of *Shorea leprosula* Miq., and *Shorea parvifolia* Dyer (Dipterocarpaceae) in Indonesia detected by AFLPs. Tree Genet. Genomes, **2**: 225–239. doi:10.1007/s11295-006-0046-0.

Chan, H.T. 1980. Reproductive biology of some Malaysian Dipterocarps II. Fruiting biology and seedling studies. Malay. For. **43**: 438–451.

Chan, H.T. 1981. Reproductive biology of some Malaysian Dipterocarps III. Breeding systems. Malay. For. **44**: 28–36.

Corriveau, J., and Coleman, A. 1988. Rapid screening method to detect potential biparental inheritance of plastid DNA and results for over 200 angiosperm species. Am. J. Bot. **75**: 1443–1458. doi:10.2307/2444695.

Filatov, D.A. 2002. PROSEQ: A software for preparation and evolutionary analysis of DNA sequence data sets. Mol. Ecol. Notes, **2**: 621–624. doi:10.1046/j.1471-8286.2002.00313.x.

Filatov, D.A., and Charlesworth, D. 1999. DNA polymorphism, haplotype structure and balancing selection in the Leavenworthia *PgiC* locus. Genetics, **153**: 1423–1434. PMID:10545470.

Gathorne-Hardy, F.J., Syaukani, D.R.G., Eggleton, P., and Jones, D.T. 2002. Quaternary rainforest refugia in south-east Asia: using termites (Isoptera) as indicators. Biol. J. Linn. Soc. **75**: 453–466. doi:10.1046/j.1095-8312.2002.00031.x.

Harada, K., Kinoshita, A., Shukor, N.A.A., Tachida, H., and Yamazaki, T. 1994. Genetic variation estimated in three *Shorea* species by the RAPD analysis. Jpn. J. Genet. **69**: 713–718. doi:10.1266/jjg.69.713. PMID:7857675.

Hudson, R., and Kaplan, N. 1985. Statistical properties of the number of recombination events in the history of a sample of DNA sequences. Genetics, **111**: 147–164. PMID:4029609.

Hudson, R., Kreitman, M., and Aguade, M. 1987. A test of neutral molecular evolution based on nucleotide data. Genetics, **116**: 153–159. PMID:3110004.

Hudson, R., Slatkin, M., and Maddison, W. 1992. Estimation of levels of gene flow from DNA sequence data. Genetics, **132**: 583–589. PMID:1427045.

Ingvarsson, P.K. 2005. Nucleotide polymorphism and linkage disequilibrium within and among natural populations of European Aspen (*Populus tremula* L., Salicaceae). Genetics, **169**: 945–953. doi:10.1534/genetics.104.034959. PMID:15489521.

Ishiyama, H., Kado, T., Iwasaki, M., Matsuoka, M., Shukor, N.A., Szmidt, A.E., and Yamazaki, T. 2003. Nucleotide variation in the *GapC* region of four species of *Shorea* and their putative hybrids. Tropics, **32**: 89–99.

Järvinen, P., Lemmetyinen, J., Savolainen, O., and Sopanen, T. 2003. DNA sequence variation in *BpMADS2* gene in two populations of *Betula pendula.* Mol. Ecol. **12**: 369–384. doi:10.1046/j.1365-294X.2003.01740.x. PMID:12535088.

Jong, K., and Lethbridge, A. 1967. Cytological studies in the Dipterocarpaceae. I. Chromosome numbers of certain Malaysian genera. Notes R. Bot. Gard. Edinb. **27**: 175–184.

Jukes, T.H., and Cantor, C.R. 1969. Evolution of protein molecules. *In* Mammalian protein metabolism. *Edited by* H.W. Munro. Academic Press, New York. pp. 21–120.

Kado, T., Yoshimaru, H., Tsumura, Y., and Tachida, H. 2003. DNA Variation in a Conifer, *Cryptomeria japonica* (Cupressaceae sensu lato). Genetics, **164**: 1547–1559. PMID:12930759.

Kamiya, K., Harada, K., Tachida, H., and Ashton, P.S. 2005. Phylogeny of *PgiC* gene in *Shorea* and its closely related genera (Dipterocarpaceae), the dominant trees in Southeast Asian tropical rain forests. Am. J. Bot. **92**: 775–788.

Kawabe, A., Yamane, K., and Miyashita, N. 2000. DNA polymorphism at the cytosolic phosphoglucose isomerase (*PgiC*) locus of the wild plant *Arabidopsis thaliana.* Genetics, **156**: 1339–1347. PMID:11063706.

Krutovsky, K.V., and Neale, D.B. 2005. Nucleotide diversity and linkage disequilibrium in cold hardiness and wood quality related candidate genes in Douglas-fir. Genetics, **171**: 2029–2041. doi:10.1534/genetics.105.044420. PMID:16157674.

Lee, S.L., Wickneswari, R., Mahani, M.C., and 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.

Liu, A., and Burke, J.M. 2006. Patterns of nucleotide diversity in wild and cultivated sunflower. Genetics, **173**: 321–330. doi:10.1534/genetics.105.051110. PMID:16322511.

Ma, X.F., Szmidt, A.E., and Wang, X.R. 2006. Genetic structure and evolutionary history of a diploid hybrid pine *Pinus densata* inferred from the nucleotide variation at seven gene loci. Mol. Biol. Evol. **23**: 807–816. doi:10.1093/molbev/msj100. PMID:16446291.

McDonald, J., and Kreitman, M. 1991. Adaptive protein evolution at the *Adh* locus in *Drosophila.* Nature (London), **351**: 652–654. doi:10.1038/351652a0. PMID:1904993.

Momose, K., Yumoto, T., Nagamitsu, T., Kato, M., Nagamasu, H., Sakai, S., Harrison, R., Itioka, T., Hamid, A., and Inoue, T. 1998. Pollination biology in a lowland dipterocarp forest in Sarawak, Malaysia. I. Characteristics of the plant-pollinator community in a lowland dipterocarp forest. Am. J. Bot. **85**: 1477–1501. doi:10.2307/2446404.

Morley, R.J. 2000. Origin and evolution of tropical rain forest. John Wiley Sons Ltd., Chichester, England.

Murray, M.G., and Thompson, W.F. 1980. Rapid isolation of high molecular weight plant DNA. Nucleic Acids Res. **8**: 4321–4325. doi:10.1093/nar/8.19.4321. PMID:7433111.

Nei, M. 1987. Molecular evolutionary genetics. Columbia University Press, New York.

Nordborg, M., Hu, T.T., Ishino, Y., Jhaveri, J., Toomajian, C., Zheng, H., Bakker, E., Calabrese, P., Gladstone, J., Goyal, R., Jakobsson, M., Kim, S., Morozov, Y., Padhukasahasram, B., Plagnol, V., Rosenberg, N.A., Shah, C., Wall, J.D., Wang, J., Zhao, K., Kalbfleisch, T., Schulz, V., Kreitman, M., and Bergelson, J. 2005. The pattern of polymorphism in *Arabidopsis thaliana.* PLoS Biol. **3**: 1289–1299. doi:10.1371/journal.pbio.0030196. PMID:15907155.

Ramos-Onsins, S.E., and Rozas, J. 2002. Statistical properties of new neutrality tests against population growth. Mol. Biol. Evol. **19**: 2092–2100. PMID:12446801.

Rieseberg, L.H., Sinervo, B., Linder, C.R., Ungerer, M.C., and Arias, D.M. 1996. Role of gene interactions in hybrid speciation: Evidence from ancient and experimental hybrids. Science (Washington, D.C.), **272**: 741–745. doi:10.1126/science.272.5262.741. PMID:8662570.

Rozas, J., Sanchez-DelBarrio, J.C., Messeguer, X., and Rozas, R. 2003. DnaSP, DNA polymorphism analyses by the coalescent and other methods. Bioinformatics, **19**: 2496–2497. doi:10.1093/bioinformatics/btg359. PMID:14668244.

Saitou, N., and Nei, M. 1987. The neighbor-joining method: A new method for reconstructing phylogenetic trees. Mol. Biol. Evol. **4**: 406–425. PMID:3447015.

Sakai, S., Momose, K., Yumoto, T., Kato, M., and Inoue, T. 1999. Beetle pollination of *Shorea parvifolia* (section *Mutica*, Dipterocarpaceae) in a general flowering period in Sarawak, Malaysia. Am. J. Bot. **86**: 62–69. doi:10.2307/2656955.

Savolainen, O., and Pyhäjärvi, T. 2007. Genomic diversity in forest trees. Curr. Opin. Plant Biol. **10**: 162–167. doi:10.1016/j.pbi.2007.01.011. PMID:17292660.

Simonsen, K.L., Churchill, G.A., and Aquadro, C.F. 1995. Properties of statistical tests of neutrality for DNA polymorphism data. Genetics, **141**: 413–429. PMID:8536987.

Slik, J.W.F., Poulsen, A.D., Ashton, P.S., Cannon, C.H., Eichhorn, K.A.O., Kartawinata, K., Lanniari, I., Nagamasu, H., Nakagawa, M., vanNieuwstadt, M.G.L., Payne, J., Purwaningsih, Saridan, A., Sidiyasa, K., Verburg, R.W., Webb, C.O., and Wilkie, P. 2003. A floristic analysis of the lowland dipterocarp forests of Borneo. J. Biogeogr. **30**: 1517–1531. doi:10.1046/j.1365-2699.2003.00967.x.

Somego, M. 1978. Cytogenetical study of Dipterocarpaceae. Malay. For. **41**: 358–365.

Stebbins, G.L. 1966. Processes of organic evolution. Prentice-Hall Inc., Englewood Cliffs, N.J.

Symington, C.F. 1943. Foresters' manual of dipterocarps: Malayan forest records No. 16. Penerbit Universiti Malaya, Kuala Lumpur.

Taberlet, P., Gielly, L., Pautou, G., and Bouvet, J. 1991. Universal primers for amplification of three non-coding regions of chloroplast DNA. Plant Mol. Biol. **17**: 1105–1109. doi:10.1007/BF00037152. PMID:1932684.

Tajima, F. 1989. Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. Genetics, **123**: 585–595. PMID:2513255.

Tajima, F. 1993. Measurement of DNA polymorphism. *In* Mechanisms of molecular evolution. *Edited by* N. Takahata and A.G. Clark. Sinauer Associates. Inc., Sunderland, Massachusetts. pp. 37–59.

Tamura, K., Dudley, J., Nei, M., and Kumar, S. 2007. MEGA4: Molecular evolutionary genetics analysis (MEGA) software version 4.0. Mol. Biol. Evol. **24**: 1596–1599. doi:10.1093/molbev/msm092. PMID:17488738.

Thomas, M.F. 2000. Late Quaternary environmental changes and the alluvial record in humid tropical environments. Quat. Int. **72**: 23–36. doi:10.1016/S1040-6182(00)00018-5.

Thompson, J.D., Gibson, T.J., Plewniak, F., Jeanmougin, F., and Higgins, D.G. 1997. The CLUSTAL_X windows interface: flexible strategies for multiple sequences alignment aided by quality analysis tools. Nucleic Acids Res. **25**: 4876–4882. doi:10.1093/nar/25.24.4876. PMID:9396791.

Tuskan, G.A., DiFazio, S., Jansson, S., Bohlmann, J., Grigoriev, I., Hellsten, U., Putnam, N., Ralph, S., Rombauts, S., Salamov, A., Schein, J., Sterck, L., Aerts, A., Bhalerao, R.R., Bhalerao, R.P., Blaudez, D., Boerjan, W., Brun, A., Brunner, A., Busov, V., Campbell, M., Carlson, J., Chalot, M., Chapman, J., Chen, G.L., Cooper, D., Coutinho, P.M., Couturier, J., Covert, S., Cronk, Q., Cunningham, R., Davis, J., Degroeve, S., Dejardin, A., dePamphilis, C., Detter, J., Dirks, B., Dubchak, I., Duplessis, S., Ehlting, J., Ellis, B., Gendler, K., Goodstein, D., Gribskov, M., Grimwood, J., Groover, A., Gunter, L., Hamberger, B., Heinze, B., Helariutta, Y., Henrissat, B., Holligan, D., Holt, R., Huang, W., Islam-Faridi, N., Jones, S., Jones-Rhoades, M., Jorgensen, R., Joshi, C., Kangasjarvi, J., Karlsson, J., Kelleher, C., Kirkpatrick, R., Kirst, M., Kohler, A., Kalluri, U., Larimer, F., Leebens-Mack, J., Leple, J.C., Locascio, P., Lou, Y., Lucas, S., Martin, F., Montanini, B., Napoli, C., Nelson, D.R., Nelson, C., Nieminen, K., Nilsson, O., Pereda, V., Peter, G., Philippe, R., Pilate, G., Poliakov, A., Razumovskaya, J., Richardson, P., Rinaldi, C., Ritland, K., Rouze, P., Ryaboy, D., Schmutz, J., Schrader, J., Segerman, B., Shin, H., Siddiqui, A., Sterky, F., Terry, A., Tsai, C.J., Uberbacher, E., Unneberg, P. 2006. The genome of black cottonwood, *Populus trichocarpa* (Torr. Gray). Science (Washington, D.C.), **313**: 1596–1604. doi:10.1126/science.1128691. PMID:16973872.

Voris, H.K. 2000. Maps of Pleistocene sea levels in Southeast Asia: shorelines, river systems and time durations. J. Biogeogr. **27**: 1153–1167. doi:10.1046/j.1365-2699.2000.00489.x.

Watt, W.B. 1977. Adaptation at specific loci. I. Natural selection on phosphoglucose isomerase of *Colias* butterflies: biochemical and population aspects. Genetics, **87**: 177–194. PMID:914029.

Wheat, C.W., Watt, W.B., Pollock, D.D., and Schulte, P.M. 2006. From DNA to fitness differences: Sequences and structures of adaptive variants of *Colias* phosphoglucose isomerase (PGI). Mol. Biol. Evol. **23**: 499–512. doi:10.1093/molbev/msj062. PMID:16292000.

Wright, S. 1978. Evolution and the genetics of populations. University of Chicago Press, Chicago, Ill.

Zhao, Z., Yu, N., Fu, Y.-X., and Li, W.-H. 2006. Nucleotide variation and haplotype diversity in a 10-kb noncoding region in three continental human populations. Genetics, **174**: 399–409. doi:10.1534/genetics.106.060301. PMID:16783003.

Zhu, Q., Zheng, X., Luo, J., Gaut, B.S., and Ge, S. 2007. Multilocus analysis of nucleotide variation of *Oryza sativa* and its wild relatives: Severe bottleneck during domestication of rice. Mol. Biol. Evol. **24**: 875–888. doi:10.1093/molbev/msm005. PMID: 17218640.