Genetic variation estimated in three *Shorea* species by the RAPD analysis

KO HARADA¹*, AKIRA KINOSHITA¹, NOR AINI AB SHUKOR²,
HIDENORI TACHIDA¹ and TSUNEYUKI YAMAZAKI¹

¹Department of Biology, Faculty of Science, Kyushu University 33,
Fukuoka, 812, Japan

²Department of Forestry, Pertanian University,
Serangor, Malaysia

(Received 1 September 1994)

ABSTRACT

Three species of *Shorea* (*S. leprosula*, *S. acuminata* and *S. cursitii*) were collected from a natural forest reserve of Malaysia and analyzed for genetic variation using the technique of random amplification of polymorphic DNA (RAPD) by the polymerase chain reaction (PCR). The average number of nucleotide substitutions was estimated. The nucleotide diversities within species were very similar and larger than those found in *Drosophila melanogaster*. The nucleotide divergences between these species are about 1.5 times the nucleotide diversities within the species, indicating that these species diverged from a common ancestor relatively recently.

1. INTRODUCTION

Tropical rain forests are important not only economically in that they produce timber, but also ecologically because they preserve innumerable species of plants and animals. Today the rapid destruction of forests globally gives cause for much apprehension. We must consider how to conserve the genetic diversity of the tropical rain forests. For this purpose, we started a study of genetic variation in Dipterocarpaceae which is one of the most typical constituents of the tropical rain forests and is distributed over a large area of tropical Africa and Southeast Asia.

The genus *Shorea* is the largest and economically most important genus of Dipterocarpaceae and is distributed widely in Southeast Asia, especially in Borneo, Sumatra and the Malay Peninsula. At least 167 species are known in the genus. *Shorea* is subdivided into four groups: the Balau group, the Meranti Pa'ang group, the Meranti Damar Hitam group and the Red Meranti group (Symington, 1943). We collected samples from three species of the Red Meranti group (*Shorea acuminata*, *S. cursitii* and *S. leprosula*) to be used for the study of genetic variation.

* Corresponding author.

For this study we applied a new genetic marker, random amplified polymorphic DNA (RAPD) (Williams et al., 1990) which is produced by the polymerase chain reaction (PCR) with a single short primer of arbitrary sequence. RAPD analysis does not need specific sequence information and a large amount of DNA samples. Polymorphisms detected by RAPD is widely used in gene mapping and identification of species strains (Wilkie et al., 1993). It is also faster and less expensive than any other methods such as RFLP analysis by Southern blotting. Thus, the method is particularly attractive in application to the survey of DNA sequence difference between and within populations.

Recently, Clark and Lanigan (1993) proposed a method for estimating the degree of sequence divergence between populations using RAPD data from diploid species. Their method depend on several assumptions. Amongst them the most important one is that single nucleotide substitution in the sequence matching to the primer results in the loss of amplification of the band. It has been shown that single nucleotide substitution in the primer results in completely different set of bands except in the cases in which changes occurred at the 5'-most position (Williams et al., 1990). This implies that amplification of imperfectly matching site is rare. Furthermore, Clark and Lanigan (1993) pointed out a number of criteria that must be satisfied before application of the methods to estimating the genetic divergence. One of the main criteria is that segregating bands are assumed to be allelic and the bands appeared in the same size among different species must be demonstrated to be homologous by the method, e. g., by Southern blotting. This criteria is, however, difficult to verify for a large number of bands which appeared in population samples and we assumed this condition is met in our experiment.

2. MATERIALS AND METHODS

Twenty eight samples of *Shorea leprosula*, 30 samples of *S. acuminata* and 22 samples of *S. cuspitii* were collected from Pair Panjang Forest Reserve, Negeri Sembilan in Malaysia. DNA was isolated by the modified CTAB method (Milligan, 1992). 0.2 to 2 g of leaves were ground in liquid nitrogen and suspended in 10 ml CTAB buffer (100 mM Tris-HCl, pH 8.0, 20 mM EDTA, 1.4 M NaCl, 2% cetyltrimethylammonium bromide (CTAB) and 0.5% 2-mercaptoethanol). The mixtures were incubated at 65°C for 20 min. The lysate was extracted with an equal volume of chloroform/isoamyl alcohol (24:1) and centrifuged at 3500 rpm for 10 min. This process was repeated several times until a clear aqueous phase was obtained. DNA was precipitated with 2/3 volumes of isopropanol. Precipitated DNA was suspended in 2 ml of TE buffer, then 2 µl of 10 µg/µl RNase A was added and incubated for 30 min at 37°C. The DNA in TE buffer was extracted with an equal volume of phenol several times. The aqueous phase was precipitated with two volumes of ethanol and 1/10 volumes of 3 M sodium acetate. The

DNA was washed in 70% ethanol and resuspended in TE buffer to be adjusted to 50 ng/ μ l.

Ten DNA samples were randomly chosen for each species for RAPD. PCR reactions were performed in 25 μ l reaction mixtures containing 50 ng DNA, 0.1 mM each of dATP, dTTP, dGTP and dCTP, 10 mM Tris-HCl, pH 9.0, 1.5 mM $MgCl_2$, 50 mM KCl, 0.1% Triton X-100, 1 unit of Taq DNA polymerase (Promega) and 5 pmol of single decamer primer purchased from Operon technologies (OPA01 to OPA20). The nucleotide sequences of these random primers are shown in Table 1. The mixtures were overlaid with one drop of mineral oil. Amplification was performed by first denaturing at 94°C for 3 min followed by 40 cycles of 1 minute at 94°C, 1 min at 36°C, 2 min at 72°C, and ending with 10 min at 72°C. 10 μ l of mixtures were electrophoresed in 1.5% agarose gel in TAE buffer using Mupid electrophoresis apparatus (Advance Inc.). Gels were stained with ethidium bromide for 10 min and photographed with UV light.

Table 1. Nucleotide sequences of random primers

Primer	Sequence	Primer	Sequenece
A01	5'-CAGGCCCTTC-3'	A11	5'-CAATCGCCGT-3'
A02	5'-TGCCGAGCTG-3'	A12	5'-TCGGCGATAG-3'
A03	5'-AGTCAGCCAC-3'	A13	5'-CAGCACCCAC-3'
A04	5'-AATCGGGCTG-3'	A14	5'-TCTGTGCTGG-3'
A05	5'-AGGGGTCTTG-3'	A15	5'-TTCCGAACCC-3'
A06	5'-CGTCCCTGAC-3'	A16	5'-AGCCAGCGAA-3'
A07	5'-GAAACGGGTG-3'	A17	5'-GACCGCTTGT-3'
A08	5'-GTGACGTAGG-3'	A18	5'-AGGTGACCGT-3'
A09	5'-GGGTAACGCC-3'	A19	5'-CAAACGTCGG-3'
A10	5'-GTGATCGCAG-3'	A20	5'-GTTGCGATCC-3'

We estimated the average number of nucleotide substitutions between these *Shorea* species following the methods described by Clark and Lanigan (1993). The average number of nucleotide substitutions within species was estimated similarly. Instead of tallying scores for a pair of individuals each from different populations, we tally scores for a pair of different individuals from the same population. The corresponding formulas to estimate n_x and n_{xy} are

$$n_x = \sum_i [4(1 - H_{x(i)}) + 2H_{x(i)}]$$

$$n_{xy} = \sum_i [4(1 - H_{x(i)})^2 + 4H_{x(i)}(1 - H_{x(i)}) + H_{x(i)}^2].$$

The sums are taken for all pairs of individuals in the population and for all bands.

3. RESULTS AND DISCUSSION

For each primer employed, a multiple band profile appeared. Apparent poly-

morphism occurred in band profiles within species for all twenty primers. When we duplicated the PCR using one primer, OPA12, we obtained exactly the same band profile, showing that the banding pattern is reproducible if PCR is done under the same condition. There are several major bands that are fixed or exist in high frequency within a species (Fig. 1). These bands often appeared to be common to all three species. There also exist varying number of minor bands (Fig. 1). These bands comprised rare polymorphic bands and unique bands. It is often difficult to determine whether rare polymorphic bands are identical if they occur in different species. We considered bands to be the same if they occurred at exactly the same location but evaluated as different if there are slight differences in location or density. The number of bands identified for each primer ranged from 5 to 26 and the average was 16.7 ± 4.9 (standard deviation). A total of 333 bands were identified. Bands on RAPD gels were scored as present (1) or absent (0) for each DNA sample and the data were analyzed by a computer program developed following the method described in the materials and methods section.

Fig. 1. Ethidium bromide stained amplified DNA of *Shorea* using random primer OPA01 (lane 1 to lane 15). Lanes 1-5 represent DNA from *S. cursitii*, lanes 6-10 represent DNA from *S. acuminata* and lanes 11-15 represent DNA from *S. leprosula*. Lanes on both ends are 0.2 μ g of *Hind*III digested λ DNA. There is a monomorphic band that is common to all species (a) or appeared only in one species (b). There are also polymorphic bands (c) and a unique band (d).

The average number of nucleotide substitutions within and between these species were summarized in Table 2. It is apparent that there is considerable variations within the species. These values are similar to each other and there are no significant differences among them. These values are much larger than those obtained in *Drosophila* genes (Began and Aquadro, 1992). Relatively high

Table 2. Estimates of the average number of nucleotide substitutions within and between three species of *Shorea*

	<i>S. cursitii</i>	<i>S. leprosula</i>	<i>S. acuminata</i>
<i>S. cursitii</i>	0.0207 ± 0.0021		
<i>S. leprosula</i>	0.0286 ± 0.0024	0.0183 ± 0.0017	
<i>S. acuminata</i>	0.0303 ± 0.0032	0.0308 ± 0.0033	0.0177 ± 0.0024

The figures on the diagonal refer to the average number of nucleotide substitutions within species and those below the diagonal represent between species. The numbers following the plus-minus sign are standard deviations.

degrees of polymorphism revealed within the species suggests that these species have large effective population sizes and/or large mutation rate due to longer generation time since nucleotide diversity is proportional to the product of the two quantities (Nei, 1987). Another possibility is interspecific gene flow among these species (see below).

Interspecific nucleotide divergences are about 1.5 times that of the diversities within species. There is a significant difference between the divergences and intraspecific diversities. However, the difference between the two is very small. Combined with the large amount of intraspecific variation, this may suggest that there is some interspecific gene flow among these species. It is suggested from the comparison of interspecific divergences that *S. cursitii* and *S. leprosula* are more closely related to each other than to *S. acuminata* although the difference is not statistically significant. Speciation may have occurred first between *S. acuminata* and the *cursitii-leprosula* group and speciation then occurred between *S. cursitii* and *S. leprosula*.

Interspecific genetic divergence includes intraspecific genetic diversities, and the net genetic divergence after speciation is obtained by subtracting the means of the two intraspecific genetic diversities from the divergence between species (Nei, 1987). Thus, the divergence between *S. acuminata* and *S. leprosula*, *S. cursitii* and *S. acuminata* and *S. cursitii* and *S. leprosula* are 0.0128, 0.0110 and 0.0091, respectively. These values are relatively small suggesting that speciation among these species occurred relatively recently. Assuming the nucleotide substitution rate to be constant, the time of divergence between these species can be estimated. The average rate of synonymous nucleotide substitution in nuclear genes of higher plants is estimated to be 6×10^{-9} per site per year (Wolfe et al., 1989). Using this value, the divergence time between *S. cursitii* and *S. leprosula* is estimated to be 0.76×10^6 ($=0.0091/(2 \times 6 \times 10^{-9})$) years. Similarly, the divergence times between *S. cursitii* and *S. acuminata* and between *S. acuminata* and *S. leprosula* are estimated to be 0.92×10^6 and 1.06×10^6 years, respectively. These values are, however, maximum estimates. As pointed by Clark and Lanigan (1993), if a substantial portion of polymorphism in plants is caused by

insertion/deletion variation as in *Drosophila*, the phylogenetic relations estimated by RAPD will be overestimated. Actual sequence data showed this is true in various *Pinus* species (Kinoshita, unpublished results).

In conclusion, we found that these three species belonging to *Shorea* have fairly large amount of intraspecific genetic variation and are closely related each other. Although the RAPD method is useful as a first step in estimating genetic variations within and between populations especially in closely related species as was demonstrated here, there are several limitations in using the method for population genetic analyses. Some problems are mentioned in the introduction section. Also the assumption of random mating in estimating allele frequency in diploid organisms must be substantiated by other methods. Thus, the validity of the present result should be verified and refined by the accumulation of sequence data and we are now sequencing nuclear genes in these species.

We thank Dr. D. A. Johnson, Seinan Gakuin University for reading and improving the manuscript. We thank Mr. Ab Latib and other members of the Department of Forestry, Pertanian University for collecting *Shorea* samples. This work was supported by Heiwa-Nakajima Foundation and the Grant-in-Aid from the Ministry of Education, Science and Culture of Japan.

REFERENCES

- Began, D. J. and Aquadro, C. F. (1992). Levels of naturally occurring polymorphism correlate with recombination rates in *D. melanogaster*. *Nature* **356**, 519–520.
- Clark, A. G. and Lanigan, M. S. (1993). Prospects for estimating nucleotide divergence with RAPDs. *Mol. Biol. Evol.* **10**, 1096–1111.
- Milligan, B. G. (1992). Plant DNA isolation. In: *Molecular Genetic Analysis of Populations* (ed.: A. R. Hoelzel), pp. 59–86. IRL Press, Oxford.
- Nei, M. (1987). *Molecular Evolutionary Genetics*. Columbia University Press, New York.
- Symington, C. R. (1943). *Foresters' Manual of Dipterocarps*. Penerbit University, Kuala Lumpur.
- Wilkie, S. E., Isaac, P. G. and Slater, R. J. (1993). Random amplified polymorphic DNA (RAPD) markers for genetic analysis in *Allium*. *Theor. Appl. Genet.* **86**, 497–504.
- Williams, J. G. K., Kubelik, A. R., Livak, K. J., Rafalski, J. A. and Tingey, S. V. (1990). DNA polymorphisms amplified by arbitrary primers are useful as genetic markers. *Nucl. Acids Res.* **18**, 6531–6535.
- Wolfe K. H., Sharp, P. M. and Li, W.-H. (1989). Rates of synonymous substitution in plant nuclear genes. *J. Mol. Evol.* **29**, 208–211.