Molecular Ecology (1999) **8**, 1585–1592

# Population structure delineated with microsatellite markers in fragmented populations of a tropical tree, *Carapa guianensis* (Meliaceae)

S. DAYANANDAN,* J. DOLE, K. BAWA and R. KESSELI
*Biology Department, University of Massachusetts, Boston, Massachusetts 02125 USA*

**Abstract**

**Deforestation and selective logging in the tropics may have serious consequences on genetic processes in tropical tree populations, affecting long-term survival of a given species as well as tropical forest communities. Because understanding the effects of human-induced changes on genetic processes is of utmost importance in formulating sound conservation and management plans for tropical forest communities, we developed microsatellite or simple sequence repeat (SSR) markers for the tropical tree *Carapa guianensis* (Meliaceae) and assessed the polymorphism of SSRs in adult and sapling populations in a large contiguous forest and in selectively logged and fragmented forests. The number of alleles in polymorphic loci ranged between 4 and 28. No inbreeding was detected in saplings or adult cohorts, but the allelic richness was lower in the sapling cohort of the isolated fragment. Genetic distances, Nei's *D* and $(\delta\mu)^2$, and $R_{ST}$ values among saplings were greater than among adult cohorts, suggesting restriction of gene flow due to deforestation and habitat fragmentation. These SSR loci may be used to address many related questions regarding the population and conservation genetics of tropical trees.**

*Keywords*: conservation genetics, deforestation, gene flow, La Selva, Royal Mahogany, simple sequence repeats (SSR)

*Received 9 January 1999; revision received 17 May 1999; accepted 17 May 1999*

## Introduction

Deforestation, forest fragmentation, and extraction of timber in the form of selective logging could have serious consequences on the long-term maintenance of genetic diversity and fitness in plants (Young *et al.* 1996; Laikre & Ryman 1996). Forest fragmentation is particularly acute in tropical forests, which harbour disproportionately high levels of biodiversity. Forest fragmentation can modify gene flow (Bawa 1993; Hamrick 1993; Young *et al.* 1996). Selective logging can reduce the density of trees, thereby altering mating patterns and increasing the probability of inbreeding (Murawski *et al.* 1994). Although fragmentation and disturbance such as selective logging are common in tropical forests, there is a poor understanding of the effect of both processes on mating systems, gene flow and genetic diversity (Bawa 1993; Hamrick 1993; Namkoong *et al*. 1996; Nason *et al.* 1997).

Because of their high allelic diversity, microsatellites or simple sequence repeats (SSRs) have strong discriminatory power, and are becoming a popular tool in population and conservation genetics (Chase *et al*. 1996b). The application of SSRs to the population and conservation genetics of tropical forests is limited due to a paucity of DNA sequence information for many tropical plant species. However, SSR primers developed for one species can be used to detect polymorphism in related species, minimizing laborious cloning and screening procedures (Schlotterer *et al*. 1991; Kijas *et al*. 1995; Levine *et al*. 1995; Dayanandan *et al.* 1997). Therefore, characterizing SSR loci from selected species representative of diverse groups would be valuable for population and conservation genetic studies of tropical forest communities, particularly for tropical forests possessing a series of closely related species groups. SSRs have now been identified for many plant species (Akkaya *et al*. 1992; Terauchi &

Correspondence: R. Kesseli. Fax: +01-617-287-6650; E-mail: rick.kesseli@umb.edu
*Present address: Department of Renewable Resources, University of Alberta, Edmonton, Alberta T6G 2H1 Canada.

Konuma 1994; Dow & Ashley 1996; Chase *et al*. 1996a; Echt *et al*. 1996; Van de Ven & McNicol 1996; Steinkellner *et al*. 1997). However, only a limited number of studies have characterized SSRs for wild species, particularly tropical trees (Chase *et al*. 1996a; Aldrich & Hamrick 1998).

Here, we report three SSR loci for *Carapa guianensis* (Meliaceae), a valuable timber species in forests of Central and South America. In Costa Rica this species, commercially known as Royal Mahogany, is common in Atlantic lowlands, where the forests are highly fragmented (Sanchez-Azofeifa *et al*. 1998). We characterized the SSRs by estimating their genomic frequency and by screening open pollinated progeny arrays to assess inheritance. We assayed genetic diversity by scoring polymorphisms among adults and saplings in three populations. We also tested the hypothesis that forest fragmentation reduces gene flow among populations. If this is correct, we expect greater genetic differentiation among populations for the sapling cohorts that were established after fragmentation than for the adult cohorts that pre-date fragmentation. Thus, apart from identifying SSR loci and describing allelic diversity at these loci, we addressed the following questions. (i) What is the level of differentiation among populations of *Carapa guianensis*? (ii) Is gene flow lower and differentiation greater among sapling cohorts than among adults cohorts? (iii) Is inbreeding in seedling cohorts higher than that in adults? We also compared results obtained from SSR loci with those obtained from isozyme loci in a previous study (Hall *et al.* 1994).

## Materials and methods

### *Plant material and populations*

We sampled mature individuals and saplings in La Selva biological reserve and in two managed forests, Gerardo Fallas and Mata Banano. The La Selva population is located within the boundaries of the 1510 ha biological reserve in which there is no evidence of past logging. The density of reproductively mature *Carapa* trees in a plot close to the study site at La Selva was less than 10 individuals per ha (Lieberman & Lieberman 1987). Gerardo Fallas and Mata Banano sites are owned by a private company, Portico SA, and managed for timber production through selective logging. The Gerardo Fallas site was selectively logged in 1988, and approximately four *Carapa* trees per ha, greater than 70 cm diameter at breast height (d.b.h.), were removed. The Mata Banano site was selectively logged in 1989, and three trees per ha were removed. The density of *Carapa* trees (35 cm < d.b.h.), after logging, in Mata Banano and Gerardo Fallas sites was 11 and 15 trees per ha, respectively (Hall *et al.* 1994). The area surrounding the Gerardo Fallas site is composed of pasture and fragmented logged forests, whereas the Mata Banano and La Selva sites are part of larger contiguous blocks of forest. The geographical distances from La Selva to Gerardo Fallas and Mata Banano sites were approximately 41 and 44 km, respectively. The distance between Mata Banano and Gerardo Falls was about 5 km.

Leaf samples were collected from 28, 26 and 20 adult canopy trees from La Selva, Gerardo Fallas and Mata Banano, respectively, and from 32 randomly distributed young saplings in each population. Given the size of the adults and known growth rate analyses, these adults are probably older than 100 years (MacHargue & Hartshorn 1983; Lieberman & Lieberman 1987; Clark & Clark 1992). Based on size and growth form, the selected saplings are probably under 10 and most may be under 5 years of age, although the precise age is unknown. Leaf samples were frozen in liquid nitrogen and transported to the laboratory at the University of Massachusetts for analysis.

### *DNA extraction*

DNA extraction followed a CTAB protocol modified from Bernatzky & Tanksley (1986). About 200–300 mg of fresh leaves was placed in a plastic bag with 2 mL of ice-cold extraction buffer (0.35 M sorbitol, 100 mM Tris pH 7.5, 5 mM EDTA, and 0.02 M sodium bisulphite), and crushed with a wooden roller. After several standard steps, the final pellet was washed with 70% ethanol, dried and dissolved in 150 μL of TE (10 mM Tris pH 7.5, 1 mM EDTA).

### *Library construction*

The extracted DNA (100 μL) was digested to completion with *Sau*3a (New England Biolabs, Beverly, MA, USA). The digested DNA was concentrated (ethanol precipitation) and dissolved in 30 μL of TE (pH 8.0). The M13mp19 vector (Life Technologies, Gaithersberg, MD, USA) DNA was digested with *Bam*HI and dephosporylated with 5 units of calf intestinal phosphatase (New England Biolabs, Beverly, MA, USA). The phosphatase was deactivated by adding 0.25 μL of 0.5 M EDTA and heating to 75 °C for 10 min. The digested DNA samples were electrophoresed in 1% agarose gel with 1× TBE buffer. DNA fragments between 350 and 750 bp and M13 DNA were excised from the gel and purified with QIAquick (Qiagen, Santa Clarita, CA, USA) gel extraction kit.

Ligation reactions were performed with 4 μL each of gel-purified plant DNA and M13mp19 DNA (80 ng of each DNA) with 1 μL of T4 DNA ligase (approximately 2000 units) in a total volume of 15 μL at 16 °C for 16 h. Ligated M13 DNA was transfected with XL1BlueMRF′ competent bacterial cells (Stratagene, CA, USA) following the manufacturer's recommendations and plated on 150 mm culture plates with Luria–Bertani (LB)/tetracycline

agar with Xgal and IPTG. A total of 18 plates, each with approximately 200 plaques was prepared.

### *Screening for microsatellite repeat sequences*

Plates were blotted with nylon membranes (Hybond N+, Amersham) for 1 min. Membranes were baked at 80 °C for 2 h to fix DNA onto the membranes and then soaked in 2× SSC for 10 min and washed with 5× SSC, 0.5% SDS for 1 h and prehybridized overnight in plastic bags at 60 °C. Each bag with nine membranes contained 150 mL of hybridization medium (20× SSC, 150 mL; 100× Denhardt's, 25 mL; BSA, 5 g; and SDS, 2 g dissolved in a total volume of 500 mL with distilled water). Following prehybridization, hybridization medium in each bag was replaced with 75 mL of fresh hybridization medium and 30 µL of the denatured probe (prepared as described below), and incubated at 50 °C in a shaking incubator for 2 h. Hybridized blots were washed twice with 150 mL of 0.1% SDS, 6× SSC at room temperature, washed with 2000 mL of 0.1% SDS and 6× SSC for 10 min at 45 °C and exposed to autoradiographic film after blot drying. The probe (an oligonucleotide of known repeats such as $(CA)_{15}$) was prepared by end-labelling 20 pmol of 30-mer oligo of desired SSR, with 30 µCi of $\gamma^{32}P$ using T4 polynucleotide kinase (Epicentre technology) at 37 °C for 30 min, in a total volume of 15 µL. The labelled probe was denatured at 94 °C for 2 min and purified using Biospin SP6 columns (Bio-Rad), after adding 35 µL of distilled water.

### *Sequencing positive clones*

Positive plaques were picked and placed into 15-mL culture tubes containing 2 mL of LB with 100 µL of XL1BlueMRF′ bacterial cells grown overnight. Tubes were incubated at 37 °C with shaking for 4 h. The culture (1.5 mL) was transferred to a microcentrifuge tube and centrifuged at maximum speed (14 000 ***g***) for 5 min. Supernatant (1000 µL) was transferred to a sterile microfuge tube and 200 µL of PEG (20%) and 2.5 M NaCl was added. The solution was vortexed and incubated at room temperature for 15 min. After centrifuging at 12 000 g for 5 min at room temperature, the pellet was resuspended in 100 µL of TE (pH 8.0). Phenol (50 µL) was added, and the solution was vortexed and centrifuged at 12 000 g for 1 min at room temperature. Supernatant was transferred to a microfuge tube with 100 µL of chloroform:isoamyl alcohol (24:1), mixed and centrifuged at 12 000 g for 2 min. Supernatant was transferred to a microfuge tube with 250 µL of 100% ethanol sodium acetate (25:1), mixed, incubated at −20 °C for 1 h and centrifuged at 14 000 ***g*** for 10 min. The pellet was washed with 70% ethanol, dried and dissolved in 50 µL of TE (pH 7.5). This DNA (4 µL) was sequenced with fmol sequencing system (Promega Corporation, Madison, WI, USA) and M13 forward primer with [$\alpha^{33}P$]-dATP radionucleotides, following the manufacturer's recommendations. The thermal cycling parameters used for sequencing were 60 s denaturation at 94 °C, 30 s annealing at 55 °C, and 60 s extension at 70 °C for 30 cycles. Electrophoresis of the sequencing reaction products was performed on 6% denaturing acrylamide gels with 6 M urea.

### *Designing primers and resolving SSR polymorphism*

Oligonucleotide primers complementary to flanking regions of identified repeats were synthesized. The PCR conditions were as follows. After the initial denaturation step at 94 °C, five cycles were performed with 55 °C annealing, 72 °C extension and 94 °C denaturation temperatures, 1 min at each step, and then for the remaining 30 cycles each step was set to 30 s. Each amplification was performed in a volume of 15 µL, which consisted of 0.2 mM dNTP, 2.0 mM $MgCl_2$, 50 mM KCl, 10 mM Tris-HCl (pH 9.0), 0.1% Triton X-100, 2.5 pmol of reverse primer, 2 pmol of forward primer, 0.5 pmol (0.5 µCi) of $^{33}P$-end-labelled forward primer and 1 unit *Taq* DNA polymerase. For primers *Cg6* and *Cg7*, a 60 °C annealing temperature was found to be optimal. The primer end-labelling mix consisted of 3 µL of 10× PNK buffer, 12.5 µL of $H_2O$, 1 µL of T4 polynucleotide kinase (Epicentre technology), 1.5 µL of Gamma $^{33}P$-labelled ATP and 12 µL of primer (30 pmol). The reaction mix was incubated at 37 °C for 30 min and then the enzyme was inactivated at 70 °C for 5 min. Amplified products were electrophoresed on 6% denaturing polyacrylamide gel with 6 M urea and 1× TBE. Electrophoresis was performed on a S2 electrophoretic apparatus (Owl Scientific, Woburn, MA, USA) for about 4 h at a setting of 1750 V, 55 W or until xylene cyanol marker migrated about 35–40 cm from the origin. A radioactively ($^{33}P$) labelled 10 bp ladder DNA (Research Genetics) was used as a size standard.

### *Data analysis*

Inbreeding coefficients ($F_{IS}$) and gametic disequilbria were estimated using GENEPOP (Raymond & Rousset 1995), and significance levels were tested using Fisher's exact tests. Allelic richness, $r_{(n)}$, the expected number of alleles in a standard sample size of $n$, was characterized by adapting the rarefaction methods of species diversity models as suggested by Petit *et al*. (1998). The MATLAB-based programs of Dr Eugene Gallagher (HTTP://www.es.umb.edu/edgwebp.htm) were used for these analyses with the standard sample size set at 40 gametes (the smallest population size). The genetic distances among populations for both the adult and sapling cohorts were estimated by the Nei's $D$ using MICROSAT (Minch *et al*. 1997) and the $(\delta\mu)^2$ statistic of Goldstein *et al*. (1995) provided by the computer program RST-CALC (Goodman 1997). This latter measure

provides an estimate of distance which is linear with time under the stepwise mutation model, which is believed to most accurately describe mutation at microsatellite loci (e.g. Valdes *et al*. 1993; Goldstein *et al*. 1995), however, see Paetkau *et al*. (1997). Population differentiation was estimated by Slatkin's (1995) $R_{ST}$ based upon the distribution of variance in SSR length within and between populations. $R_{ST}$ is a better estimate of population differentiation with higher stepwise mutation rates than Wright's $F_{ST}$ (Slatkin 1995). We first standardized allele length to eliminate bias and used RST-CALC (Goodman 1997) to compute $R_{ST}$ and migration rate statistics based upon 5000 bootstraps for both adults and seedlings separately. Migration rate, *Nm*, is a simple function of $R_{ST}$, because $R_{ST}$ can be treated as the SSR analogue of Wright's $F_{ST}$; migration rate may be estimated as $Nm = 1/4\ (1/R_{ST} - 1)$ (cf. Slatkin & Barton 1989; Slatkin 1995). The paternity exclusion probability of each locus was calculated as given in Weir (1990).

## Results

### *Frequency and informativeness of SSRs*

The partial genomic library of *Carapa guianensis*, consisting of 18 plates with about 3600 M13 plaques, was screened with five di- and tetranucleotide repeat oligonucleotide probes [$(CA)_{15}$ $(AG)_{15}$ $(AAAG)_7$ $(AAAT)_7$ $(AGAT)_7$]. Fifty-one positive plaques with potential microsatellite repeats were identified. DNA isolated from 19 plaques was sequenced and 10 clones containing 7–26 repeats were identified. Oligonucleotide primers were designed for seven sequences and PCR conditions were optimized for four loci (Table 1). Assuming a mean insert size of 550 bp in 3600 clones, and the presence of $(CA)_n$ repeats in four, and $(AG)_n$ repeats in six clones in a total of 19 sequenced positive clones, we estimated that $(CA)_n$ and $(AG)_n$ repeats are present at least once in every 184 kb and 123 kb, respectively, in the *Carapa* genome.

Of the four optimized loci, three were polymorphic. These primers were used to assess the level of polymorphism in populations of *C. guianensis* in natural forests. We examined the segregation of these alleles in open-pollinated half-sib arrays (data not shown). All individuals of a given half-sib family shared at least one allele with the maternal parent. Also, each maternal allele had a frequency of about 0.5, supporting a Mendelian pattern of segregation.

From 170 individuals assayed, 28, 4 and 12 alleles were detected in the loci *Cg5*, *Cg6* and *Cg7*, respectively (Table 1). The sizes of alleles in the most polymorphic locus, *Cg5*, ranged from 164 to 222 bp with consecutive alleles differing by 1–6 bp. Alleles at *Cg7* ranged from 215 to 241 bp with consecutive alleles differing by 2 or 4 bp and the alleles at *Cg6* ranged from 106 to 110 bp differing by 1 or 2 bp. Only for one allele ($Cg6^{108}$) did the allele frequency of the most common allele greatly exceed, by fourfold or more, that of any other allele and only one other allele ($Cg7^{229}$) in one site exceeded a frequency of 0.50 (data not shown). Generally alleles showed a relatively even distribution of frequencies (Table 2).

The paternity exclusion probability for SSR loci ranged from 0.14 for the less polymorphic *Cg6* locus to 0.77 for the highly polymorphic *Cg5* locus. There was no evidence for gametic disequilibrium between *Cg5* and *Cg7* ($\chi^2 = 3.8$, d.f. = 2, $P > 0.10$); calculations involving *Cg6* were inconclusive because of the low level of polymorphism at this locus. The overall joint paternity exclusion probability of three SSR loci was 0.93. Every adult had a unique multilocus genotype. From another study of *C. guianensis*, paternity exclusion probabilities for six isozyme loci ranged from 0.04 to 0.46 (Hall *et al.* 1994). The joint paternity exclusion probability achieved with the isozyme loci was only 0.73.

### *Gene diversity, inbreeding and allelic richness*

Allelic diversity (expected heterozygosity, $H_E$) ranged from a low of 0.12 for *Cg6* in La Selva to nearly 0.90 for *Cg5* in all populations. Allele frequencies in each population were in agreement with Hardy–Weinberg expectations

**Table 1** Oligonucleotide primer sequences, repeat pattern, number of alleles and product length of four microsatellite loci

| Locus | Primer (5′–3′) | Repeat | No. of alleles | Allele size |
|---|---|---|---|---|
| Cg4 | ACACATATGAATAGATATGCAGAGA<br>CACATAAAGCAACAGATACGCCAGC | $(AG)_5GG(AG)_5$ | 1 | 202 |
| Cg5 | TGCATTTCTTCTCCTTGCTTCTGG<br>TCTTGTACGTTGGCTCTTGCACATTC | $(AG)_{13}(AAT)_{13}$ | 28 | 164–222 |
| Cg6 | ACATTCTCTCTCTTTTTCTCTCAGC<br>TAGTGAAAAAGACTAGACCTCACAGT | $(AAG)_4(AG)_{11}$ | 4 | 106–110 |
| Cg7 | AAGACTTCCCCCAGATTTGTTTTT<br>TGTACCAGTTCAGTTGATGCTCA | $(AG)_{12}$ | 12 | 215–241 |

**Table 2** Combined allelic frequencies for the three SSR markers in all populations of *Carapa guianensis*

| Locus | Allele* | Freq | Locus | Allele | Freq | Locus | Allele | Freq |
|---|---|---|---|---|---|---|---|---|
| *Cg5* | 164 | 0.038 | *Cg6* | 106 | 0.003 | *Cg7* | 215 | 0.111 |
| | 166 | 0.243 | | 107 | 0.003 | | 219 | 0.015 |
| | 168 | 0.003 | | 108 | 0.799 | | 221 | 0.039 |
| | 172 | 0.003 | | 110 | 0.195 | | 223 | 0.096 |
| | 177 | 0.015 | | | | | 225 | 0.081 |
| | 183 | 0.027 | | | | | 229 | 0.401 |
| | 184 | 0.009 | | | | | 231 | 0.129 |
| | 186 | 0.080 | | | | | 233 | 0.024 |
| | 189 | 0.006 | | | | | 235 | 0.006 |
| | 190 | 0.012 | | | | | 237 | 0.033 |
| | 191 | 0.012 | | | | | 239 | 0.024 |
| | 192 | 0.186 | | | | | 241 | 0.042 |
| | 195 | 0.109 | | | | | | |
| | 197 | 0.018 | | | | | | |
| | 198 | 0.059 | | | | | | |
| | 200 | 0.018 | | | | | | |
| | 201 | 0.006 | | | | | | |
| | 202 | 0.047 | | | | | | |
| | 204 | 0.012 | | | | | | |
| | 206 | 0.006 | | | | | | |
| | 208 | 0.012 | | | | | | |
| | 210 | 0.018 | | | | | | |
| | 212 | 0.015 | | | | | | |
| | 214 | 0.006 | | | | | | |
| | 216 | 0.009 | | | | | | |
| | 218 | 0.012 | | | | | | |
| | 220 | 0.018 | | | | | | |
| | 222 | 0.003 | | | | | | |

*Alleles are fragment sizes in base pairs.

(Fisher's Exact test $P > 0.05$) and $F_{IS}$ values (Table 3) showed no evidence of inbreeding in saplings or adult cohorts. In all sapling cohorts, observed heterozygosity values at all loci were higher than expected, although not significantly. In adults, only *Cg7* showed a higher level of heterozygosity and the remaining two loci showed a lower than expected heterozygosity value, although again not significant.

In general, the sapling cohorts of a population resembled their adult cohorts. However, the genetic distances, $D$ and $(\delta\mu)^2$, between adults and saplings of the Gerardo Fallas were greater than that found at the La Selva and Mata Banano sites ($D = 0.10$ vs. $D = 0.01$ and 0.02 and $(\delta\mu)^2 = 0.11$ vs. $(\delta\mu)^2 = 0.05$ and 0.06, respectively). This distinction between the adult and sapling cohorts of Gerardo Fallas is consistent across all loci (Fisher's Exact Test $P < 0.05$ for *Cg5* and *Cg7*). In the Gerardo Fallas population, which is surrounded by pastures and degraded forests, the sapling cohort had a lower average allelic richness index than the adult population (the expected number of alleles in a standard sample of 40 gametes, $r_{(40)} = 7.3$ vs. 7.9). For the other two sites the allelic richness index was greater in the sapling cohorts; $r = 7.4$ vs. 7.1 at the La Selva site and $r = 8.0$ vs. 6.3 at the Mata Banano site.

### *Genetic differentiation and gene flow among similar cohorts of different populations*

The genetic distances, $D$ and $(\delta\mu)^2$, among sapling cohorts averaged twice that among adult cohorts; $D = 0.147$ vs. 0.075 and $(\delta\mu)^2 = 0.153$ vs. 0.076, respectively. A pairwise comparison showed that this was due to large increases in the distance among sapling cohorts of La Selva and the other two populations, the most separated pairs (Table 4). Because of relatively large standard errors, the pairwise distances were not generally significantly different from each other. Saplings showed higher levels of $R_{ST}$ than adults, also suggesting greater genetic differentiation; observed values were 0.060 for saplings and 0.017 for adults. Mean $R_{ST}$ values ± SE based on 5000 bootstraps were 0.074 ± 0.003 and 0.041 ± 0.003, respectively. The pairwise comparison of mean $R_{ST}$ among populations again shows the greatest increase between generations for the distant population pairs and a comparatively lower increase

**Table 3** Number of alleles ($A_O$), Effective number of alleles ($A_E$), observed ($H_O$) and expected ($H_E$) heterozygosity, and fixation index ($F_{IS}$) for three loci in adult and sapling populations of La Selva, Gerardo Fallas and Mata Banano forests. Number of individuals sampled in each cohort is given in parenthesis

| Locus | $A_O$ | $A_E$ | $H_E$ | $H_O$ | $F$ |
|---|---|---|---|---|---|
| Selva Adults (28) | | | | | |
| Cg5 | 13 | 7.70 | 0.87 | 0.86 | 0.01 |
| Cg6 | 2 | 1.34 | 0.26 | 0.22 | 0.14 |
| Cg7 | 9 | 3.17 | 0.70 | 0.79 | −0.13 |
| La Selva saplings (32) | | | | | |
| Cg5 | 14 | 5.43 | 0.82 | 0.84 | −0.02 |
| Cg6 | 2 | 1.13 | 0.12 | 0.13 | −0.05 |
| Cg7 | 10 | 2.87 | 0.66 | 0.72 | −0.09 |
| Gerardo Fallas Adults (26) | | | | | |
| Cg5 | 14 | 5.83 | 0.85 | 0.81 | 0.05 |
| Cg6 | 4 | 2.15 | 0.54 | 0.50 | 0.08 |
| Cg7 | 7 | 3.82 | 0.75 | 0.80 | −0.06 |
| Gerardo Fallas saplings (32) | | | | | |
| Cg5 | 14 | 7.10 | 0.86 | 0.88 | −0.02 |
| Cg6 | 2 | 1.44 | 0.31 | 0.38 | −0.22 |
| Cg7 | 9 | 5.22 | 0.82 | 0.91 | −0.11 |
| Mata Banano Adults (20) | | | | | |
| Cg5 | 11 | 5.81 | 0.86 | 0.80 | 0.07 |
| Cg6 | 2 | 1.51 | 0.36 | 0.35 | 0.02 |
| Cg7 | 6 | 4.60 | 0.80 | 0.83 | −0.04 |
| Mata Banano Saplings (32) | | | | | |
| Cg5 | 15 | 8.29 | 0.89 | 0.93 | −0.05 |
| Cg6 | 2 | 1.32 | 0.27 | 0.31 | −0.17 |
| Cg7 | 11 | 6.49 | 0.86 | 0.93 | −0.10 |

**Table 4** Genetic distances, Nei's $D$ and $(\delta\mu)^2$ (Goldstein *et al.* (1995)) among adult and among sapling cohorts of *Carapa guianensis* in La Selva (LS), Gerardo Fallas (GF) and Mata Banano (MB) forests

| | Adults | | | Saplings | | |
|---|---|---|---|---|---|---|
| | LS | GF | MB | LS | GF | MB |
| Nei's $D$ | | | | | | |
| La Selva (LS) | — | | | — | | |
| Gerardo Fallas (GF) | 0.08 | — | | 0.18 | — | |
| Mata Banano (MB) | 0.09 | 0.05 | — | 0.19 | 0.07 | — |
| $(\delta\mu)^2$ | | | | | | |
| La Selva (LS) | — | | | — | | |
| Gerardo Fallas (GF) | 0.11 | — | | 0.17 | — | |
| Mata Banano (MB) | 0.03 | 0.08 | — | 0.17 | 0.12 | — |

in the neighbouring Mata Banano–Gerardo Fallas populations (from $R_{ST}$ = 0.043 in adults to $R_{ST}$ = 0.051 in saplings; Table 5). Correspondingly lower values of *Nm* (calculated from the observed $R_{ST}$ values not the bootstrap values) were found for the saplings (*Nm* = 3.9) verses the adults (*Nm* = 14.1) and the overall results suggest a reduction in gene flow prior to the establishment of the sapling cohort. The sapling *Nm* values may be in error due to a lack of population equilibrium if the recent deforestation has indeed reduced gene flow.

**Table 5** Mean $R_{ST}$ values among adult and among sapling cohorts of *Carapa guianensis* in La Selva (LS), Gerardo Fallas (GF) and Mata Banano (MB) forests

| | Adults | | | Saplings | | |
|---|---|---|---|---|---|---|
| | LS | GF | MB | LS | GF | MB |
| La Selva (LS) | — | | | — | | |
| Gerardo Fallas (GF) | 0.001 | — | | 0.087 | — | |
| Mata Banano (MB) | 0.022 | 0.043 | — | 0.076 | 0.051 | — |

## Discussion

The dinucleotide repeats $(CA)_n$ and $(AG)_n$, in *Carapa guianensis* occur approximately once every 184 kb and 123 kb, respectively; a frequency that fits within the range of values estimated for other plant taxa (Condit & Hubbell 1991; Lagercrantz *et al.* 1993; Terauchi & Konuma 1994; Chase *et al.* 1996b).

Several factors dictate the level of polymorphism to be expected from different types of SSRs. In general, SSRs with many repeats have been shown to be more polymorphic than ones with few repeats (Weber 1990; Chase *et al.* 1996a; Van de Ven & McNicol 1996). The locus *Cg5* is a long compound repeat $(AG)_{13}$ $(AAT)_{13}$, and showed the highest polymorphism. Also as expected, *Cg4*, which had short interrupted repeats, was monomorphic. A second factor that may affect the levels of polymorphism detected with SSRs is the type of repeat. Repeats of AT (or possibly AAT as in *Cg5* of our study) have been noted to be more polymorphic than other classes (Witsenboer *et al.* 1998). Consecutive alleles in the highly polymorphic locus *Cg5* often vary by one and three base steps, suggesting that both length and class of repeat may be important and should be considered when choosing markers for a study. The gene diversity (expected heterozygosity) values estimated from SSR loci were over five-times greater than those estimated from isozyme loci in *Carapa* (Hall *et al.* 1994), suggesting that SSR loci would better detect fine-scale genetic differentiation.

While one effect of forest fragmentation may be to increase local inbreeding levels in progeny cohorts, our study detected no evidence of inbreeding in either the sapling or adult cohorts of any of the sites. However, we did detect some interesting trends that suggest an effect of fragmentation. The Gerardo Fallas site is immediately surrounded by logged pastures and degraded forests, unlike the two other sites which are part of larger forests.

The genetic analyses showed Gerardo Fallas to be different from the other forests in two ways. First, the adult and sapling cohorts of this site were much more distinct. The genetic distance was two- to threefold greater than that found between the cohorts of the other two sites. Second, the allelic richness of this sapling cohort was slightly lower than its adult counterpart again unlike the other two sites which showed slightly higher allelic richness in sapling cohorts. These differences could be caused by changes in local mating patterns similar to that described recently by Aldrich & Hamrick (1998). They showed that remnant populations of the tropical tree *Symphonia globulifera* received a substantial influx of pollen and seed from neighbouring pasture trees. A few off-site pasture trees contributing disproportionally to the sapling cohort could both reduce its allelic richness and increase the genetic distance between the on-site cohorts of the Gerardo Fallas population. As the differences are small in our populations, larger sample sizes would be needed to detect widespread statistical significance.

Data from this study also suggest that deforestation may alter the regional-scale population structure and dynamics of fragmented remnant forests. As forest trees are long-lived, the large adults sampled probably predate the destruction of surrounding forests while saplings do not. The deforestation has been extensive during the last 50 years. The annual rate of loss in the region of this study since 1976 has been 2.15% and the number of forest fragments has more than doubled while the average size of these fragments has decreased from 0.95 $km^2$ to 0.25 $km^2$ (Sanchez-Azofeifa *et al*. 1998). Adult cohorts are probably remnants of a large contiguous population that once occupied the lowlands. The saplings are part of a cohort that has largely arisen from a fragmented landscape. We have shown that the genetic distances $D$ and $(\delta\mu)^2$ among the sapling cohorts are higher than the corresponding distances among adult cohorts. Similarly, the substructure ($R_{ST}$) is higher and the subsequent gene flow estimates are lower among the sapling cohorts. These data suggest that the gross landscape changes occurring in this region may decrease gene flow and increase allelic diversity among sites, although other possible explanations such as selection during various life-cycle stages cannot be discounted and the long-term significance of these small differences cannot be determined. Recently, others have begun to characterize the effects of deforestation and fragmentation on the structure of remnant populations. Young *et al*. (1996) discussed several studies that suggested that increased or decreased gene flow may result. Aldrich & Hamrick (1998) found that gene flow from outside the sampling area was three-times greater in contiguous than remnant forests. Future studies comparing cohorts from different subpopulations within contiguous and remnant forests and directly characterizing paternity, mating patterns and gene flow events (Chase *et al*. 1996b) will be needed to develop a full picture of the complex changes induced by landscape fragmentation and degradation.

## Acknowledgements

The authors thank Dr Mike Chase, Jennifer Forman, Joel Gerwein, Kristen Long and Cameron Moller for technical help and comments on the manuscript. This research was supported by NSF grant DEB 94–00147 and a grant from the International Plant Genetic Resources Institute, Rome, Italy.

## References

Akkaya MS, Bhagwat AA, Cregan PB (1992) Length polymorphisms of simple sequence repeat DNA in soybean. *Genetics*, **132**, 1131–1139.

Aldrich P, Hamrick J (1998) Reproductive dominance of pasture trees in a fragmented tropical forest mosaic. *Science*, **281**, 103–105.

Bawa K (1993) Effects of deforestation and forest fragmentation on genetic diversity in tropical tree populations. In: *Genetic Conservation and Production of Tropical Forest Seed* (eds Drysdale R, John S, Yapa A), 10–16. ASEAN–Canada Forest Tree Seed Centre, Chiang Mai, Thailand.

Bernatzky R, Tanksley S (1986) Genetics of actin related sequences in tomato. *Theoretical and Applied Genetics*, **72**, 314–321.

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

Chase M, Moller C, Kesseli R, Bawa K (1996b) Long distance gene flow patterns precisely characterized by simple sequence repeats in a fragmented tropical tree population. *Nature*, **383**, 398–399.

Clark DA, Clark DB (1992) Life history of canopy and emergent trees in a neotropical rain forest. *Ecological Monographs*, **62**, 315–344.

Condit R, Hubbell S (1991) Abundance and DNA sequence of two-base repeat regions in tropical tree genomes. *Genome*, **34**, 66–71.

Dayanandan S, Bawa K, Kesseli R (1997) Conservation of microsatellites among tropical trees (Leguminosae). *American Journal of Botany*, **84**, 1658–1663.

Dow B, Ashley M (1996) Microsatellite analysis of seed dispersal and parentage of saplings in bur oak, *Quercus macrocarpa*. *Molecular Ecology*, **5**, 615–627.

Echt CS, May-Marquardt P, Hseih M, Zahorchak R (1996) Characterization of microsatellite markers in eastern white pine. *Genome*, **39**, 1102–1108.

Goldstein D, Linares AR, Cavalli-Sforza LL, Feldman M (1995) Genetic absolute dating based on microsatellites and the origin of modern humans. *Proceedings of the National Academy of Science of the USA*, **92**, 6723–6727.

Goodman S (1997) RST CALC: a collection of computer programs for calculating unbiased estimates of genetic differentiation and determining their significance for microsatellite data. *Molecular Ecology*, **6**, 881–885.

Hall P, Orrell L, Bawa K (1994) Genetic diversity and mating system in a tropical tree, *Carapa guianensis*. *American Journal of Botany*, **81**, 1104–1111.

Hamrick J (1993) Genetic diversity and conservation in tropical

forests. In: *Proceedings, International Symposium on Genetic Conservation and Production of Tropical Forest Seed* (eds Drysdale RM, John SET, Yapa AC), pp. 1–9. ASEAN–Canada Forest Tree Seed Center, Chiang Mai, Thailand.

Kijas J, Fowler JCS, Thomas MR (1995) An evaluation of sequence tagged microsatellite site markers for genetic analysis within citrus and related species. *Genome*, **38**, 349–355.

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

Laikre L, Ryman N (1996) Effects on intraspecific biodiversity from harvesting and enhancing natural populations. *Ambio*, **25**, 504–509.

Levine I, Cheng H, Bacter J, Hillel J (1995) Turkey microsatellite DNA loci amplified by chicken specific primers. *Animal Genetics*, **26**, 107–110.

Lieberman D, Lieberman M (1987) Forest tree growth and dynamics at La Selva, Costa Rica (1969–82). *Ecology*, **3**, 347–358.

MacHargue LA, Hartshorn GS (1983) *Carapa guianensis*. In: *Costa Rican Natural History* (ed. Janzen DH), 206–207. University of Chicago Press, Chicago, IL.

Minch E, Ruiz-Linares A, Goldstein D, Feldman M (1997) Microsat (Version 1.5b): a computer program for calculating various statistics on microsatellite allele data. WWW: http://lotka.stanford.edu/microsat.html.

Murawski DA, Gunatilleke IAUN, Bawa KS (1994) The effects of selective logging in *Shorea megistophylla* (Dipterocarpaceae) from Sri Lanka. *Conservation Biology*, **8**, 99–1002.

Namkoong G, Boyle T, Gregorius H, Joly H, Savikaubebm O, Ratnum W, Young A (1996) Testing criteria and indicators for assessing the sustainability of forest management: Genetic criteria and indicators. In: Center for International Forestry Research, Indonesia: Working Paper 10.

Nason JD, Aldrich PR, Hamrick JL (1997) Dispersal and dynamics of genetic structure in fragmented tropical tree populations. In: *Tropical Forest Remnants: Ecology, Management, and Conservation of Fragmented Communities* (eds Laurance WF, Bierregard RO), pp. 304–320. University of Chicago Press, Chicago, IL.

Paetkau D, Waits L, Clarkson P, Craighead L, Strobeck C (1997) An empirical evaluation of genetic distance statistics using microsatellite data from bear (Ursidae) populations. *Genetics*, **147**, 1943–1957.

Petit R, El-Mousadik A, Pons O (1998) Identifying populations for conservation on the basis of genetic markers. *Conservation Biology*, **12**, 844–855.

Raymond M, Rousset F (1995) GENEPOP: a population genetics software for exact test and ecumenicism. *Journal of Heredity*, **86**, 248–249.

Sanchez-Azofeifa GA, Quesada-Mateo C, Dayanandan S, Bawa K (1998) Protected areas and conservation of biodiversity in the tropics. *Conservation Biology*, **13**, 407–411.

Schlotterer C, Amos B, Tautz D (1991) Conservation of polymorphic simple sequence loci in cetacean species. *Nature*, **354**, 63–65.

Slatkin M (1995) A measure of population subdivision based on microsatellite allele frequencies. *Genetics*, **139**, 457–462.

Slatkin M, Barton N (1989) A comparison of three indirect methods for estimating average levels of gene flow. *Evolution*, **43**, 1349–1368.

Steinkellner J, Fluch S, Turetschek E, Streiff C, Kremer A, Burg K, Gloessl J (1997) Identification and characterization of (GA/CT)n microsatellite loci in *Quercus petraea*. *Plant Molecular Biology*, **33**, 1093–1096.

Terauchi R, Konuma A (1994) Microsatellite polymorphism in *Dioscorea tokoro*, a wild yam species. *Genome*, **37**, 794–801.

Valdes A, Slatkin M, Freimer N (1993) Allele frequencies at microsatellite loci: a stepwise mutation model revisited. *Genetics*, **84**, 639–659.

Van de Ven WTG, McNicol RJ (1996) Microsatellites as DNA markers in Sitka spruce. *Theoretical and Applied Genetics*, **88**, 1–6.

Weber JL (1990) Human DNA polymorphisms based on length variations in simple-sequence tandem repeats. In: *Genome Analysis*, Vol. 1: *Genetic and Physical Mapping* (ed. Davis KE), pp. 159–181. Cold Spring Harbor Laboratory Press, Cold Spring Harbor.

Weir BS (1990) *Genetic Data Analysis*. Sinauer Associates, Sunderland, MA.

Witsenboer H, Vogel J, Michelmore RW (1998) Identification, genetic localization and allelic diversity of selectively amplified microsatellite polymorphic loci (SAMPL) in lettuce and wild relatives (*Lactuca* spp.). *Genome*, **40**, 923–936.

Young A, Boyle T, Brown T (1996) The population genetic consequences of habitat fragmentation for plants. *Trends in Ecology and Evolution*, **11**, 413–418.

---

S. Dayanandan has interests in plant systematics and tropical and temporate forest ecology. Jeffrey Dole studies plant mating systems and is interested in the genetic basis of herterosis. Kamal Bawa has been studying the reproductive biology of tropical forest trees for many years. Recently, he has been evaluating land use and its effects on biodiversity. Rick Kesseli is interested in the molecular bases of disease resistance and breeding systems in plants. He also studies the population genetics of rare species.

---