Conservation genetics of exploited Amazonian forest tree species and the impact of selective logging on inbreeding and gene dispersal in a population of *Carapa guianensis*

Dominic Cloutier
Department of Biology
McGill University
Montréal, Québec, Canada

A thesis submitted to McGill University in partial fulfillment
of the requirements of the degree of Doctor of Philosophy

Your file Votre référence
ISBN: 978-0-494-32166-9
Our file Notre référence
ISBN: 978-0-494-32166-9

Table of contents

Résumé

L'Amazonie est une région convoitée pour sa richesse en biodiversité et en ressources naturelles. La coupe sélective d'arbres en Amazonie est une activité lucrative et de plus en plus répandue qui pourrait entraîner une dégradation des ressources génétiques des populations d'arbres exploités. Afin d'examiner cette hypothèse, j'ai mesuré les niveaux de consanguinité, de flux génique, et de diversité génétique dans des populations d'arbres de la forêt amazonienne soumises à l'exploitation forestière. Des marqueurs génétiques de type microsatellite ont été choisis pour l'étude en raison de leur grande variabilité dans les populations. J'ai d'abord vérifié que l'utilisation de ce type de marqueurs était envisageable chez les arbres, même si leur grande taille et leur longévité pourraient accroître l'incidence de mutations, puisqu'ils se sont avérés suffisamment stables dans les lignées germinales des arbres. En échantillonnant des arbres adultes et leur descendance maternelle à plusieurs loci microsatellites, j'ai découvert une dispersion à grande distance des gènes et une consanguinité faible dans des populations de *Sextonia rubra* et de *Carapa guianensis*, deux importantes espèces d'arbres de la forêt amazonienne pollinisées par des insectes. En comparant des graines récoltées avant et après une coupe sélective, aucun impact mesurable n'a pu être détecté sur la dynamique de la génétique des populations de *Carapa guianensis*, puisque les niveaux de consanguinité, de flux de gènes et de structure génétique se sont révélés similaires avant et après la coupe. En comparant différentes populations distribuées dans le bassin de l'Amazone, j'ai pu mettre en évidence une structure phylogéographique de l'ADN de chloroplaste de *Carapa guianensis* qui semble correspondre aux tributaires majeurs de l'Amazone, suggérant que l'hydrochorie pourrait contribuer à des échanges génétiques entre différentes populations. Les résultats des études menées dans le cadre de cette thèse suggèrent que les grandes tailles efficaces des populations, les importants flux de gènes, et les faibles niveaux de consanguinité observés dans les populations d'arbres exploitées en Amazonie pourrait signifier qu'elles sont résistantes à d'éventuelles conséquences

génétiques de la coupe sélective, du moins pour la population de *Carapa guianensis* étudiée, et pour le type de récolte pratiqué.

Abstract

The Amazon region is one of the richest areas on the planet in terms of its biodiversity and natural resources. The large scale harvesting of trees in this region is a relatively new activity, and it is uncertain whether the exploitation of timber species will result in depletion of forest genetic resources. To examine this, I have assessed levels of inbreeding, gene flow, and genetic diversity in populations of Amazonian forest trees undergoing logging. Because of their high variability within populations, microsatellite genetic markers were chosen for the study, and it was verified through an initial sampling experiment that this class of markers is sufficiently stable within somatic tissue of large and long-lived trees such that population studies could be undertaken with them. By sampling adult trees and seed progenies at several microsatellite loci, high levels of gene flow and low levels of inbreeding were found within populations of *Sextonia rubra* and *Carapa guianensis*, two important insect-pollinated Amazonian forest tree species. Comparing seed progeny collected before versus after selective logging of a population of *Carapa guianensis*, no measurable evidence was found that that the population genetic dynamics is impacted by logging. In particular, levels of inbreeding, gene flow, and population substructure were the same before and after logging. Comparing different populations distributed over the Amazon basin, a phylogeographical structure in the chloroplast DNA of *Carapa guianensis* that corresponds to major tributaries of the Amazon river was discovered, suggesting that seed dispersal through rivers may contribute to genetic connectivity among populations. Overall, the results of this thesis suggest that the large effective population sizes, the high levels of gene flow, and the low levels of inbreeding in exploited Amazonian tree populations may allow them to counteract potential negative genetic impacts of selective logging, at least at the levels of harvesting carried out during this study, and for the *Carapa guianensis* population investigated.

Resumo

A Amazônia é uma região rica em biodiversidade e recursos naturais. O corte seletivo de árvores na Amazônia é uma atividade lucrativa e cada vez mais importante, que podera provocar uma degradação dos recursos genéticos das populações de árvores exploradas. A fim de examinar esta hipótese, foram medidos os níveis de consangüinidade, fluxos gênicos, e diversidade genética em populações de árvores da floresta amazônica sujeitas à exploração florestal. Marcadores genéticos de tipo microsatellite foram utilizados para o estudo, devido à sua grande variabilidade nas populações. Inicialmente verificou-se que a utilização deste tipo de marcador era possível nas árvores, ainda que sua grande dimensão e a sua longevidade pudesse aumentar a incidência de mutações, dados que revelaram-se suficientemente estáveis nas linhagens germinais das árvores. Amostrando árvores adultas e sua descendência materna a vários locos microsatellites, descobriu-se uma dispersão à grande distância dos genes e uma consangüinidade fraca em populações de *Sextonia rubra* (Louro Vermelho) e de *Carapa guianensis* (Andiroba), duas importantes espécies de árvores da floresta amazônica polinizadas por insetos. Comparando sementes colhidas antes e após um corte seletivo, nenhum impacto mensurável pôde ser detectado sobre a dinâmica da genética das populações de *Carapa guianensis*, dado que os níveis de consangüinidade, fluxos de genes e de estrutura genética revelaram-se similares antes e após o corte. Comparando diferentes populações distribuídas na Amazônia, foi possivel destacar uma estrutura filogeográfica do DNA de cloroplasto de *Carapa guianensis* que parece corresponder com os tributários principais do Amazonas, sugerindo que a hidrocoria pudesse contribuir para as trocas genéticas entre diferentes populações. Os resultados dos estudos efetuados no âmbito desta tese sugerem que os grandes tamanhos efetivos das populações, os importantes níveis de fluxos de genes, e os níveis fracos de consangüinidade observados nas populações de árvores exploradas da Amazônia, possam significar que tais populações são resistentes a eventuais conseqüências genéticas de corte seletivo, pelo menos para a população de

Carapa guianensis investigada, e para o tipo de corte praticado.

Acknowledgements

I wish to express my first gratitude to my supervisor Daniel Schoen, from whom I learned about population genetics, for the insightful comments he constantly provides, for being accessible and helpful, and for allowing me to conduct all those studies I was interested in. Thanks also to my supervisory committee, Selvadurai Dayanandan, Jean Beaulieu and to my qualifying examiners Frédéric Guichard and Andrew Hendry for their assistance.

Thanks to all members of the W6-17 lab for discussions and mutual help, in reverse chronological order, Mattieu Bégin, Jeremiah Busch, Rod Docking, Julia Sharma, Anne-Marie l'Heureux, Fabienne Saade, Pierre-Olivier Cheptou, Aura Navarro-Quezada, Stephen Wright, Joelle Pérusse.

Merci également aux autres qui ont momentanément contribué à cette thèse; du cinquième étage, Hien Le, Nabil Elrouby, Thomas Bureau and his lab; du Service Canadien des Forêts, Centre de Foresterie des Laurentides, Jean Beaulieu, Danny Rioux, Marie Deslauriers, Manuel Lamothe; de l'Université Laval, Louis; de l'Université de Montréal, Anne Bruneau, Simon Joly, Mario Cappadocia; de l'INRA, Bernd Degen, Antoine Kremer, Henri Caron, Olivier Hardy, Saint-Omer Cazal; et du CIRAD, Valérie Troispoux et Laurent Maggia. Merci aussi au CRDI, qui a financé une partie de la recherche présentée dans cette thèse, d'avoir contribué à développer mon intérêt pour la question du développement durable, et à NATEQ et au CRSNG pour m'avoir permis de me loger et de me nourrir durant toutes ces années.

Gostaria agradecer todos brasileiros, amigos e parceiros, que estiveram comigo em momentos desta tese. Primeiro o projeto Dendrogene, Milton Kanashiro, Ian Thompson e Ana Ciampi para acolher-me na vossa equipa na Embrapa. Obrigado para o pessoal do IBAMA em Santarém. Em Belém, Francimary, Licurgo, Mike, André. Em Brasília, Erich, Vânia, Tályta, Valci, Joema, Christina, Janaína. No campo, Campolina, as duas

Ana Claudia, Marivana, Jaqueline, Geane, Sr. Olinto, Sr. Parente, e toda communidade Branco-1 km83 Br163, e Belterra, Laércio(Jaca), Buru, Giovanni, Preto, Isaac, Xaropinho, Benilson e todos outros.

Thanks to Nature Publishing Group, Blackwell Publishing and J.D. Sauerländer's Verlag for giving me permission to reproduce here the articles I have published during my PhD thesis.

Et finalement merci à mes parents et à ma famille pour m'avoir donné tout ce dont j'avais besoin pour faire ce que j'avais envie de faire, même si je devais m'expatrier à Montréal, en France ou au Brésil pour le faire, Jeannot Cloutier, Pauline Groleau, Geneviève, Etienne, Joey, Myriam, Mélanie.

Contribution of authors

Each of the four main chapters of this thesis was written with the purpose of producing an original research account for publication. Chapter 1 (*Heredity*) and Chapter 4 (*Silvae Genetica*) are already published, Chapter 2 is accepted for publication into *Biotropica*, and Chapter 3 is accepted for publication in *Molecular Ecology*. I am the first author on each article, reflecting the fact that I initiated and supervised each project, conducted the experiments, analyzed the data, and wrote up the research results independently. Daniel Schoen is last author on each article, as he provided advice and assistance during my thesis, and contributed to the editing of all manuscripts. The conception of the study in Chapter 1 was shared between myself and my supervisor, but the ideas behind Chapters 2, 3 and 4 were entirely my own. Other authors have made only minor contributions to the writing of the articles, but have been important for the realization in helping in the lab (J. Povoa, L. Procopio, D. Rioux), helping to get samples (N. Leao, L. Wadt, H. Caron, O. Hardy, J. Beaulieu), and providing logistical support and access to research facilities (A. Ciampi, M. Kanashiro, B. Degen, H. Caron, J. Beaulieu, D. Rioux).

General introduction and literature review

Biodiversity is the sum of all different kinds of life. While the most general aspect of biodiversity consists of all the varieties of existing biological species, communities, and ecosystems, the different populations composing a given species also harbor another, and perhaps, cryptic, component of biodiversity, *i.e.* the genetic diversity existing within and among populations. In some cases, this within-species genetic diversity can be seen (or tasted) as, for example, when one contemplate the tens or hundreds of different varieties of potatoes sold in the markets of Bolivia, an extreme contrast to the four or five potatoes varieties found in a typical Montréal supermarket. But in most cases, genetic diversity is hidden, and requires the use of special tools to assess it (*e.g.*, marker genes such as microsatellites).

Most people see biodiversity as a "good thing", worth preserving. This is why international agencies have worked to create legislative tools such as the "Convention on Biological Diversity", an agreement committing 150 signatory countries, including Canada, to " develop national strategies, plans or programmes for the conservation and sustainable use of biological diversity (...)"[1]. Conservation biology is the name of an applied branch of biology that aims to study the dynamics of biodiversity, with the goal of suggesting practical measures that can help to conserve it. When conservation biology involves the application of genetic theory to conserve genetic diversity found within or among populations, it is called "conservation genetics".

Trees are central to the study of biodiversity because they make up a large portion of the Earth's biomass and sustain biodiversity in many places of the world by their role as dominant members of ecosystems. Tree populations are also exploited for a wide range of culturally and economically important products, including edible fruits, medicines, timber, and fiber. Documenting the impacts of exploitation and other anthropogenic disturbances on tree populations has been proven difficult for a number of reasons,

---

[1] http://www.biodiv.org/convention/default.shtml

including the fact that the size and life span of trees is much larger than that of humans. As a result, few empirical studies have been conducted on the subject (Ledig 1992). Moreover, a historically repeated pattern of conflict of interests, between the short-term profit generated through extraction of the resource, and the long-term objectives of genetic diversity conservation in exploited tree populations, has impeded the development of practical solutions to the conflict over the use of forests (Namkoong 2005). New approaches that reconcile the use and the conservation of forest resources are needed to optimize the conservation of genetic diversity in trees.

One widely reported human disturbance to forests is deforestation. This is a special threat to the forests of Amazonia, one of the last, large, and relatively undisturbed reservoirs of biodiversity on the planet. While the impact of large scale, clear-cutting of forests has been studied, less is known about the extent of selective logging (*i.e.*, logging of a few high-value tree stems, leaving some forest cover). We know relatively little about the scale of this mode of forest disturbance, partially because the satellite-based methodology employed to generate deforestation figures has not captured areas that have been selectively logged. However, a recent study has shown that the total area selectively logged each year in the Brazilian Amazon is about the same as the total area deforested, thus doubling previous estimates of the total area recently disturbed by humans in Amazonia (Asner *et al* 2005). Under such a scenario, it is essential to document whether large-scale activities, such as selective logging, have consequences on the genetic diversity of harvested tree species.

The types of knowledge required to address the conservation genetics of a given species can be broken down into two components: demographic and genetic (Nunney 2000). Demographic approaches are designed to minimize the risk of extinction due to demographic or environmental stochasticity (Shaffer 1981), and are of primary importance to the short-term survival of the population (Lande 1988). However, because a genetic response to changing conditions (or a lack of them) may also have an effect on demographic parameters, the short-term importance of demography does not negate the importance of genetics (Nunney 2000).

In conservation genetics, a central parameter is the size of a population, as this will influence patterns of gene sampling (*i.e.* genetic drift) that occurs during mating of the individuals (Savolainen and Kuittinen 2000). In large populations, the gene sampling effect is reduced and large amounts of genetic variability may be maintained; while in small populations some genetic diversity is inevitably lost by chance as a result of the small number of individuals mating. However, other factors, such as unequal male and female numbers and variation in fecundity among individuals, complicate this and normally result in an increase of this sampling effect, so that even populations with large census size may be "genetically small" (Savolainen and Kuittinen 2000). The degree to which gene sampling can change gene diversity levels can be approximated under these complications by calculating an "effective population size" (Wright 1931), *i.e* as it is most commonly used, the number of individuals in a theoretically ideal population having the same magnitude of random genetic drift or inbreeding as the actual population.

Logging could have genetic consequences for the harvested population, as it potentially contributes to decreased levels of genetic exchange among individuals, and more generally to the reduction of effective population sizes. Two genetic consequences of a reduction in effective population size are increased genetic drift and inbreeding (Ellstrand and Elam 1993, Lowe *et al.* 2005). These two consequences may, in turn, influence small populations by contributing to changing patterns of genetic diversity and fitness (Ledig 1992, Ellstrand and Elam 1993, Lowe *et al.* 2005). Moreover, logging may modify physical and biological features of the forest habitat in a way that can further contribute to the reductions of reproductive population size. This is especially the case for trees that rely on animal pollinators for mating, *e.g.*, if pollinator communities are impacted by logging, this could have a direct result on the mating system or levels of gene flow (Bawa 1990, Didham *et al.* 1996, Dick *et al.* 2003). Gene flow among populations, through seed dispersal for instance, is yet another aspect of conservation genetics that may impact levels of genetic diversity within species, and again, is vulnerable to disturbance.

The potential for logging to bring about changes in genetic composition of populations arising from inbreeding and loss of variation may pose significant threats to

long-term viability of populations. This is because such changes may decrease the ability of populations to adapt to abiotic and biotic environmental change, including the ability of the population to cope with short-term challenges such as pathogens and herbivores, e.g., due to loss of resistance genes (Ledig 1992, Lowe *et al.* 2005).

Two different and conflicting views have been articulated regarding potential human impacts on the genetic make-up of exploited tree populations: (1) trees are vulnerable because of demographic and reproductive characteristics, such as low population density, complex breeding systems, high rates of outcrossing, and complex interactions with animal pollinators and seed dispersers (Finkeldey & Ziehe 2004, Lowe *et al.* 2005); (2) trees are resistant because they have overlapping generations (Degen *et al.* 2006), and the high levels of gene flow are sufficient to counteract the effects of genetic drift; moreover, their genetic variation is distributed within rather than among populations (Hamrick 2004). It has also been suggested that favorable conditions for growth may reoccur several times during the long lifetime of the tree, and that past glacial contractions in geographic ranges of old widespread tree species have not resulted in the loss of genetic variation (Hamrick 2004). Depending on the genetics, ecology and reproductive biology, a given tree population may be expected to be more vulnerable, or more resistant, to genetic changes.

The main goal of this PhD thesis was to investigate whether selective logging has neutral genetic consequences for harvested Amazonian tree populations. In carrying out this work, I benefited from a joint collaborative effort with Brazilian partners involved in Amazonian forest resources management. Within the framework of the Dendrogene Project [2], an Amazonian undisturbed forest was studied before and after commercial selective logging of nearly 40 tree species. Several tree species, chosen for their different reproductive biology and ecological characteristics, were monitored during the logging. The study species assigned to me by the Dendrogene Project was *Carapa guianensis*, a cousin of the mahogany. The studies on *Carapa guianensis* within the Project aimed to provide utilizable genetic information necessary for future studies outside the framework

[2] http://www.cpatu.embrapa.br/dendro/ingles/all.htm

of this thesis, where the impact of different forest management scenarios on exploited tree populations could be simulated for the next tens or hundreds of years (*e.g.*, Degen *et al.* 2006).

Literature cited

Asner GP, Knapp DE, Broadbent EN, Oliveira PJC, Keller M, Silva JN. (2005) Selective logging in the Brazilian Amazon. Science 310: 480-482.

Bawa KS. (1990) Plant-pollinator interactions in tropical rain forests. Annual Review of ecology and Systematics 21: 399-422.

Degen B, Blanc L, Caron H, Maggia L, Kremer A, Gourlet-Fleury. (2006) Impact of selective logging on genetic composition and demographic structure of four tropical tree species. Biological Conservation 131: 386-401.

Dick CW, Etchelecu G, Austerlitz F. (2003) Pollen dispersal of tropical trees (*Dinizia excelsa*: Fabaceae) by native insects and African honeybees in pristine and fragmented Amnazonian rainforest. Molecular Ecology 12: 753-764.

Didham RK, Ghazoul J, Stork NE, Davis AJ. (1996) Insects in fragmented forests: a functional approach. Trends in Ecology and Evolution 11: 255-261.

Ellstrand NC, Elam DR. (1993) Population genetic consequences of small population size: Implications for plant conservation. Annual Review of Ecology and Systematics 24: 217-242.

Hamrick JL. (2004) Response of forest trees to global environmental changes. Forest Ecology and Management 197: 323-335.

Lande R (1988) Genetics and demography in biological conservation. Science 241: 1455-1460.

Ledig FT. (1992) Human impacts on genetic diversity in forest ecosystems. Oikos 63:87-108.

Lowe AJ, Boshier D, Ward M, Bacles CFE, Navarro C. (2005) Genetic resource impacts of habitat loss and degradation; reconciling empirical evidence and predicted theory for neotropical trees. Heredity 95: 255-273.

Namkoong G. (2005) The misunderstood forest. (http://genetics.forestry.ubc.ca/book.html)

Nunney L. (2000) The limits to knowledge in conservation genetics – the value of effective population size. Evolutionary Biology 32: 179-194.

Savolainen O, Kuittinen H. (2000) Small population processes. In: Forest conservation genetics: Principles and practice. Eds.: A Young, D Boshier and T Boyle. CSIRO publishishing, Australia. pp. 91-99.

Shaffer ML (1981) Minimum population sizes for species conservation. Bioscience 31: 131-134.

Wright S (1931) Evolution in Mendelian populations. Genetics 16: 97-159.

Linking statement for Chapter 1

Before studies on population genetic variation are undertaken, one must consider the tools used to reveal genetic diversity and ask whether they are adequate for the questions one wishes to tackle. In particular, microsatellite gene markers, increasingly used in population studies, are well-known for their "hyper-variability", the corollary of which is that the mutation rate at those markers is high. Plants, as opposed to animals, lack a true germline that acts to protect cell lineages from mutations. This may be especially problematic for trees that are large and long-lived, and therefore have higher potential for mutation accumulation in their somatic tissues. Therefore, when genetic markers such as microsatellite loci are assayed from somatic tissue or from seed on different positions of the plant, the question of their stability must arise. The first chapter of this thesis addresses this question by assessing the somatic stability of microsatellite markers in the long-lived species, *Pinus strobus*.

Chapter 1:

Somatic stability of microsatellite loci in eastern white pine *Pinus strobus* L.

Cloutier D, Rioux D, Beaulieu J, Schoen DJ. 2003. Somatic stability of microsatellite loci in eastern white pine *Pinus strobus* L. ***Heredity*** 90: 247-252.

Abstract

Variation at nuclear- and chloroplast-encoded microsatellite loci was studied among and within clonally propagated individuals of Eastern white pine. Total DNA was extracted and assayed from gamete-bearing tissue (megagametophytes) located on six different branch positions on each of 12 individual genets. No within-individual variation was observed among 12 loci studied. Estimates of numbers of mitotic cell divisions required to produce the tissue used as the source of genomic DNA were obtained by combining tree growth and anatomical data. This allowed for the calculation of upper bound estimates of numbers of mutations per locus per somatic cell division. The estimated somatic mutation rate was found to be substantially lower than those published for genomic microsatellite mutation rates in other plant species.

Introduction

Unlike animals, the germ line in plants is not sequestered early in development, and thus gametes may arise from cell lineages that have undergone many mitotic divisions after zygotes have formed (Klekowski, 1988). If errors of DNA replication accompany mitosis and are not selectively removed during vegetative growth of the plant (eg, by cell lineage sorting), they can be passed on to the gametes (Klekowski, 1988). It has been argued that somatic mutation may play an important role in producing adaptive variability among plant shoots (Whitham and Slobodchikoff, 1981; Antolin and Strobeck, 1985; Walbot, 1996), in allele frequency change within populations (Orive, 2001), and in life-history patterns of inbreeding depression and mating system evolution (Morgan, 2001). If somatic mutations are passed on to gametes this may result in genomic mutation rates that are higher in plants than in animals, especially in long-lived plants where there

may be many somatic cell divisions prior to gamete production (Klekowski and Godfrey, 1989).

There is evidence for the occurrence of somatic mutations in plants (Whitham and Slobodchikoff, 1981; Antolin and Strobeck, 1985; Schaal, 1987; Klekowski, 1988; Gill et al, 1995; Kovalchuk et al, 2000b). Within individual variation for polygenic and cytogenetic traits has been documented by several authors (Lewis et al, 1971; Klekowski, 1988; Klekowski and Godfrey, 1989). In addition, molecular genetic variation has been detected among naturally occurring clones in several plant species (Schaal and Learn, 1988; Capossela et al, 1992; Tuskan et al, 1996). Such somatic variation might be expected to be most common at loci with high mutation rates, eg, as reported in some cases for microsatellite DNA sequences (Weber and Wong, 1993; Schlötterer et al, 1998; Udupa and Baum, 2001). There is, however, little information about the stability of microsatellite loci during somatic development of plants (Klekowski, 1988; Smith and Devey, 1994; Gill et al, 1995; Elsik et al, 2000). Given the increasing use of microsatellites markers in genetic analyses of long-lived plants, it is important to characterize their stability within individuals. In this study, we report the results of an investigation of somatic mutation in the Eastern white pine, *Pinus strobus* L. Information useful for gauging the rate of somatic mutation was obtained by combining genetic, growth, and anatomical data. Rather than restricting the analysis solely to variation expressed within the vegetative body of the plant, this study examines tissue that has differentiated into gamete-bearing structures (megagametophytes), thereby allowing us to determine the potential for chimeric transmission of somatically occurring mutations at microsatellite loci.

Materials and methods

Tree and tissue sampling

Tree sampling was conducted in a *P. strobus* breeding orchard set up by the Canadian Forest Service near Cap- Tourmente, Quebec, Canada. Propagation of the trees located in this orchard was originally carried out by selecting a number of adjacent shoots (ramets) from selected parent trees (genets) in natural populations and grafting them to separate rootstocks in the orchard. Two ramets each from 12 different genets were sampled for a total of 24 trees. From each ramet, female cones were collected at three branch positions that were marked for measurements made later on in the study (see below). Seeds were extracted from the cones in September 2000, kept separately in bags, and stored in a dry container at 41C. All genetic analyses were carried out using haploid megagametophyte tissue. Depending upon the analysis, DNA was extracted either from individual megagametophytes and analysed separately, or from the pooled tissue of 10 megagametophytes collected from cones at a single branch position. Figure 1 outlines the sampling scheme.

Microsatellite analyses

Genomic DNA was extracted from megagametophytes using the DNeasy plant mini-kit (Quiagen, Valencia, CA, USA). DNA samples were screened for variation at eight nuclear and four chloroplast microsatellite loci (Table 1). The choice of loci has been made on the basis of the level of polymorphism detected in previous studies (Echt et al, 1996). The more highly variables of these loci were selected since they are expected to have the highest mutation rate. DNA sequences of the PCR primers for the loci examined in this study have been published elsewhere (Cato and Richardson, 1996; Echt et al, 1996; Vendramin et al, 1996). The PCR reaction and amplification conditions were as described by Echt et al (1996, 1999). Fluorescently labeled amplified PCR products were separated

by electrophoresis on a 41 cm polyacrylamide gel (Li-Cor model 4000L automatic DNA sequencer) (Li-Cor, Lincoln, NE, USA).

Genotypes of individual genets at each microsatellite locus were determined by examining gel-banding patterns of PCR amplification products obtained using template DNA extracted from a single branch position (ie, DNA from a set of 10 pooled megagametophytes taken from a single cone). As this procedure is based on inferring the parent genotype from the combined haploid genotypes of meiotically derived products, it is possible that one of the two alleles at a locus for which the parent is heterozygous was not detected. However, with 10 megagametophytes per genet, the probability of such an event is small ($P=0.5^{10}<0.001$, if we consider that the allele of only one megagametophyte could be detected). To assay the parental genotype at other branch positions, and thereby establish the presence of somatically occurring mutations passed on to seeds, this procedure was applied to pooled megagametophytes collected from cones at each of the remaining five positions per genet.

The genetic interpretation of the microsatellite banding patterns was confirmed by genotypic segregation analyses conducted with each locus, using sets of individual megagametophytes taken from three separate parent trees. If no amplification was observed from the DNA at one locus, but amplification of the expected size product was observed at a control locus, the tree was assumed to have a segregating nonamplifying 'null' allele at the tested locus. When only one allele was observed from the three genets (Table 3), segregation analyses allowed us to discriminate between homozygotes and heterozygotes for null alleles.

Estimates of number of cell divisions

For each ramet, the length of vegetative growth separating the branch positions where cones were sampled was estimated. This was done by measuring the length from each branch position to a common position on the tree trunk, and then summing the distances. The distance from this common position to the position where the ramet (scion)

was grafted onto the rootstock was added to this measure. The same set of measurements was made for the paired ramet of this genet, and finally, the two ramet lengths were summed to obtain the total vegetative growth per genet (Figure 1).

To convert vegetative growth into a number of mitotic divisions, estimates of mean cell length are required. Branch tips were collected in May 2001, and radial sections of the shoot extending back through the past 2 years of vegetative growth were used. Tissue sections (1 cm long) were cut in half and fixed in FAA (4% formaldehyde, 5% acetic acid, and 47.5% ethanol) for 21 h. They were dehydrated by increasing the ethanol concentration over several days. Ethanol was gradually replaced by n-butanol, and the samples were put in nbutanol: Paraplast+ (Oxford, St Louis, MO, USA) (1:1) overnight at 60^0C, followed by transfer to three baths of pure Paraplast+ over 10 days. Longitudinal sections (7 mm) were obtained with a Leica Biocut rotary microtome (Heidelberg, Germany), deparaffined and stained with 0.05% toluidine blue O and 0.25% safranin O, and examined with a Leitz Orthoplan microscope. Cell lengths were determined from shoot sections. As selected in other studies (Brown and Sommer, 1992), the lengths of the pith cells along the median longitudinal axis of the shoot were measured with a micrometer.

Since cell enlargement accompanies cell division, measures of cell length were taken within the apical meristem (the region of active cell division), directly below the apical meristem (where cells begin to differentiate into pith tissue), and within mature pith tissue (where cell enlargement is completed). Cells within each of 10 separate longitudinal files of cells per section were counted. These files of cells were randomly selected throughout each location. Cell divisions that do not give rise to vegetative growth (eg, cell divisions within cones) were not considered in this study, but as these account for a small fraction of overall cell divisions, the introduced error is considered to be negligible.

Results

Microsatellite variation among and within genets

Microsatellite variation among separate genets was detected at all but one of the loci surveyed (Table 2). Variation in the number of alleles per locus and number of genotypes observed was higher for nuclear as compared with chloroplast loci, a finding that accords with the results of Provan et al (1999). The average observed heterozygosity per tree was 0.57 (SD=0.23). Four of the nuclear microsatellite loci appear to produce null alleles (Table 2). Analysis of microsatellite genotypes from individual megagametophytes produced by single genets confirmed the presence of segregation in families from trees judged to be heterozygous on the basis of the analysis of pooled megagametophytes (Table 3). No microsatellite mutation events were detected within genets; that is, alleles detected in megagametophytes at one branch position were identical to those detected at other positions within the genet (Figure 2). This was true of both nuclear and chloroplast loci. However, it is possible that mutations to null alleles could have occurred and remained undetected.

Estimates of number of cell divisions

The total vegetative growth unit separating the sampled branch positions within each genet varied between 15 and 21m (Table 4). Microscopic investigation of shoot sections revealed that cell lengths increased from within the apical meristem (mean length of 236 cells=21.2 um, SD=1.4 um) through to 1-year-old shoot tissue (mean length of 143 cells=70.4 um, SD=6.6 um), where they attained their maximum size. These cell length values are similar to those observed in *P. taeda* (Brown and Sommer, 1992).

Rather than basing estimates of number of cell divisions per unit of vegetative growth on cell lengths obtained from only a single region of the shoot, estimates were

based on both the smallest and largest mean cell lengths observed within the shoot tips (those from the apical meristem and 1-year-old region). Such a procedure takes into account the fact that cell enlargement and cell division may occur simultaneously within meristems (Brown and Sommer, 1992); that is, the possibility that there is no one characteristic size for an actively dividing meristematic cell. This yields two different estimates of cell divisions per unit vegetative growth, which should include the true value: 47 170 and 14 205 divisions per metre of vegetative growth.

Discussion

Our data suggest that microsatellites are stable genetic markers within Eastern white pine, as no evidence of mutation was observed during somatic growth and gamete formation. These findings are in accord with other analyses of microsatellite loci inheritance. For instance, in two separate studies of three-generation pedigrees in pine, no mutations were observed (Elsik et al, 2000; Smith and Devey, 1994). One possible exception is the study conducted by Fisher et al (1998), who reported the appearance of nonparental alleles at a single nuclear microsatellite locus in 6% of the progeny and megagametophytes from *P. radiata* crosses. Unfortunately, the sampling design of the study prevents any inference about the origin of mutations (somatic or germline) as well as their rate of occurrence. Moreover, each of the novel alleles observed by Fisher et al (1998) was at least four repeat units different from the parental allele, and thus contradicts the expectation that mutations at microsatellite loci should involve single repeat unit changes following slippage during DNA replication (Ellegren, 2000; Schlotterer, 2000). Such hypermutable microsatellite loci seem to be the exception rather than the rule, since mutations have never been reported in any other study conducted in *P. strobus* (eg Echt et al, 1996), in other pine species (eg Smith and Devey, 1994; Hicks et al, 1998; Elsik et al, 2000), and in conifers in general (eg Ven et al, 1996; Pfeiffer et al, 1997; Khasa et al, 2000; Hodgetts et al, 2001; Rajora et al, 2001). Overall, these results suggest that the use

of microsatellites as markers in long-lived organisms (eg, in marker-assisted selection and population studies) should, in general, provide reliable results.

There is little quantitative information available on rates of somatic mutation in plants. Somatic base substitution frequency in plants has been estimated to be two to three orders of magnitude higher than in other organisms (Kovalchuk et al, 2000b). Genomic mutation rates for microsatellite loci in plants have been estimated to fall within the range of 10^{-2}–10^{-3} for nuclear-encoded loci with tri- and dinucleotide repeats (Kovalchuk et al, 2000a; Udupa and Baum, 2001) to 10^{-5} for chloroplast encoded loci with mononucleotide repeats (Provan et al, 1999). Comparison of estimated mutation rates of microsatellite loci in related shorter- and longer-lived annual plant species of the genus *Cicer* led Udupa and Baum (2001) to argue that mutations occurring during somatic growth contribute to enhanced levels of genomic mutation in plants; a two- to three-fold increase in genomic mutation rates was found for the longer-lived species of *Cicer*. The results from the present investigation do not support the interpretation that somatic mutations can contribute significantly to the rate of genomic mutation.

As no mutations were observed in this study, direct calculation of the rate of somatic mutation was not possible. One can, however, use probabilistic methods to set upper boundaries on the mutation rate per cell division for each locus. For instance, in the case of a single tree and single diploid locus, if the haploid mutation rate per cell division per locus is u, and r cell divisions have occurred, the probability of observing no mutations at this locus can be taken from the zero term of the Poisson distribution with parameter $2ru$. This probability is $P=e^{-2ru}$. To find a boundary at the P=5% level for u, one can solve for u in the previous equation, setting P=0.05. Using this procedure with the data in Table 4 and the higher and lower estimates of cell division per metre of stem growth (see above) yields estimates of maximum somatic mutation for the nuclear microsatellite loci of 1.38×10^{-7} and 4.59×10^{-7} mutations per cell division, respectively. For the haploid, chloroplast loci, the procedure is similar, except we use the equation $P=e^{-ru}$. This gives estimates of 2.77×10^{-7} and 9.19×10^{-7} mutations per cell division.

Since the apical region of the meristem may divide less often than adjacent regions, the above procedure may underestimate the true somatic mutation rate. In a

detailed cytological study, Owens and Molder (1973) quantified mitotic activity of different regions of the shoot apical meristem in the related pine family member, *Pseudotsuga menziesii*. The regions studied included the apical and rib meristems. Cell divisions within the apical region give rise to the rib meristems, which contribute to vegetative growth but ultimately cease to divide within the first year. Only the apical region contains permanently dividing cells (apical initials). Thus, when tracing tissue located in different regions of the plant body back to a common 'ancestral' meristematic cell, it is the number of cell divisions within the apical meristematic region that should be counted. To quantify the relative contribution of cell division activity in the apical versus rib zones, Owens and Molder (1973) calculated mitotic frequencies using Feulgen staining methods. In the most active growth stage, that is, late bud-scale initiation, the mitotic frequency of the apical zone was 0.8, while in the rib zone it was 5.6. The proportion of cell divisions that rise from the mitotic activity in the apical zone can, therefore, be approximated as 0.8/5.6=0.14. If these relative rates are also characteristic of the apical meristem in *P. strobus*, then the estimated numbers of cell divisions separating different positions along the shoot axis (Table 4) should be reduced (ie, by multiplying by 0.14). Doing this would increase the estimated somatic mutation rates at nuclear locus, roughly 10-fold, to 9.9×10^{-7} and 3.3×10^{-6} mutations per cell division per locus, respectively, for cell division estimates based on cell lengths of 21.2 and 70.4 um. Given the likelihood that most microsatellite mutations are selectively neutral, cell lineage selection probably plays little or no role in distorting these estimates of somatic mutation (Otto and Hastings, 1998; Otto and Orive, 1995).

In theory, such estimates of mutation could also be obtained by pooling information across loci, but because it has been shown that microsatellite mutation rates may differ widely among loci (Di Rienzo et al, 1998; Harr et al, 1998), this approach is probably not warranted. Several factors may contribute to differences in mutation rates among microsatellite loci. They include repeat number, sequence of the repeat motif, length of the repeat unit, flanking sequence, and interruptions in the microsatellite (Ellegren, 2000; Schlotterer, 2000).

The above calculation of the somatic mutation rate also does not take into account mutations to null alleles. Null alleles, however, are thought to occur by base substitution, deletion or insertion within the priming site, and these events are not expected to be as frequent as mutations by slippage at DNA replication.

Estimates of the number of cell divisions separating different positions on an individual tree are difficult to obtain. One possible source of error is that cell length may differ within and along the stem. Our estimates of maximum cell length were based on pith cells, but it has been observed that cortex cells were of similar size and would therefore give the same results. Measurements of maximum length of cells were made at 7cm from the branch tip (ie in the previous growing year) and hence these cells had completed their elongation. Furthermore, measurements made with two additional white pine branches that had completed their elongation gave similar results to the values reported in this study and in the study of Brown and Sommer (1992) (data not shown). The observed level of variation in maximum cell length is negligible and the estimated somatic mutation rate remains low even when we use higher cell length estimations.

The meristem is perhaps better viewed as a heterogeneous pool of initials, some of them dividing rapidly (ie, the rib meristem) and producing the greater portion of the somatic tissue, while other meristematic initials may act as permanent stem cells (ie, apical meristem). In this study, we proposed a correction for this feature of meristems based on studies conducted with *P. menziesii*, but it is possible that the extension to *P. strobus* meristems is not accurate. Meristem zones of *P. menziesii* varied in size and in mitotic activity at different times of the growing season (Owens and Molder, 1973). It is therefore problematic to circumscribe precisely the different regions of the meristem – instead there may be a gradual transition of mitotic activity from the apical to the rib meristem. Additional studies are needed to quantify the role of apical initials as stem cells for meristematic tissues in plants. Perhaps this pattern of cell division activity within meristems may have evolved as a means to minimize errors during somatic DNA replication from being passed on from one generation to the next.

Acknowledgements

We thank Pierre-Olivier Cheptou, Marie Deslauriers, Craig Echt, Frank Ewers, Jim Grob, Manuel Lamothe, Aura Navarro-Quezada, Marie Simard, John Owens, and GG Vendramin for their advice and assistance with this study. The Natural Sciences and Engineering Research Council of Canada provided financial support through its scholarship program to DC and an Operating Grant to DJS. DC also acknowledges financial support from the Canadian Forest Service.

References

Antolin MF, Strobeck C (1985). The population genetics of somatic mutation in plants. Am Nat 126: 52–62.

Brown CL, Sommer HE (1992). Shoot growth and histogenesis of trees possessing diverse patterns of shoot development. Am J Bot 79: 335–346.

Capossela A, Silander Jr JA, Jansen RK, Bergen B, Talbot DR (1992). Nuclear ribosomal DNA variation among ramets and genets of white clover. Evolution 46: 1240–1247.

Cato SA, Richardson TE (1996). Inter- and intraspecific polymorphism at chloroplast SSR loci and the inheritance of plastids in Pinus radiata D. Don. Theor Appl Genet 93: 587– 592.

Di Rienzo A, Donnelly P, Toomajian C, Sisk B, Hill A, Petzl-Erler M et al (1998). Heterogeneity of microsatellite mutations within and between loci and implications for human demographic histories. Genetics 148: 1269–1284.

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

Echt CS, Vendramin GG, Nelson CD, Marquardt P (1999). Microsatellite DNA as shared genetic markers among conifer species. Can J For Res 29: 365–371.

Ellegren H (2000). Microsatellite mutations in the germline: implications for evolutionary inference. Trends Genet 16: 551– 558.

Elsik CG, Minihan VT, Hall SE, Scarpa AM, Williams CG (2000). Low-copy microsatellite markers for Pinus taeda L. Genome 43: 550–555.

Fisher PJ, Richardson TE, Gardner RC (1998). Characteristics of single- and multi-copy microsatellites from Pinus radiata. Theor Appl Genet 96: 969–979.

Gill DE, Chao L, Perkins SL, Wolf JB (1995). Genetic mosaicism in plants and clonal animals. Annu Rev Ecol Syst 26: 423–444.

Harr B, Zangerl B, Brem G, Schlo¨ tterer C (1998). Conservation of locus specific microsatellite variability across species: a comparison of two Drosophila sibling species D. melanogaster and D. simulans. Mol Biol Evol 15: 176–184.

Hicks M, Adams D, O'Keefe S, Macdonald E, Hodgetts R (1998). The development of RAPD and microsatellite markers in lodgepole pine (Pinus contorta var. latifolia). Genome 41: 797– 805.

Hodgetts RB, Aleksiuk MA, Brown A, Clarke C, Macdonald E, Nadeem S et al (2001). Development of microsatellite markers for white spruce (Picea glauca) and related species. Theor Appl Genet 102: 1252–1258.

Khasa PD, Newton CH, Rahman MH, Jaquish B, Dancik BP (2000) Isolation, characterization, and inheritance of microsatellite loci in Alpine larch and Western larch. Genome 43: 439–448.

Klekowski EJ (1988). Mutation, Developmental Selection, and Plant Evolution. Columbia University Press: New York.

Klekowski EJ, Godfrey PJ (1989). Aging and mutation in plants. Nature 340: 389–391.

Kovalchuk O, Dubrova YE, Arkhipov A, Hohn B, Kovalchuk I (2000a). Wheat mutation after Chernobyl. Nature 407: 583.

Kovalchuk I, Kovalchuk O, Hohn B (2000b). Genome-wide variation of the somatic mutation frequency in transgenic plants. EMBO J 19: 4431–4438.

Lewis WH, Oliver RL, Luikart TK (1971). Multiple genotypes in individuals of Claytonia virginica. Science 172: 564–565.

Morgan MT (2001). Consequences of life history evolution for inbreeding depression and mating system evolution in plants. Proc R Soc London B 268: 1817–1824.

Orive ME (2001). Somatic mutations in organisms with complex life histories. Theor Popul Biol 59: 235–249.

Otto SP, Hastings IM (1998). Mutation and selection within the individual. Genetica 103: 507–524.

Otto SP, Orive ME (1995). Evolutionary consequences of mutation and selection within an individual. Genetics 141: 1173–1187.

Owens JN, Molder M (1973). A study of DNA and mitotic activity in the vegetative apex of Douglas fir during the annual growth cycle. Can J Bot 51: 1395–1409.

Pfeiffer A, Olivieri AM, Morgante M (1997). Identification and characterization of microsatellites in Norway spruce (Picea abies K.). Genome 40: 411–419.

Provan J, Soranzo N,Wilson NJ, Goldstein DB, PowellW(1999). A low mutation rate for chloroplast microsatellites. Genetics 153: 943–947.

Rajora OP, Rahman MH, Dayanandan S, Mosseler A (2001). Isolation, characterization, inheritance and linkage of microsatellite DNA markers in white spruce (Picea glauca) and their usefulness in other spruce species. Mol Gen Genet 264: 871–882.

Schaal BA (1987). Somatic variation and genetic structure in plant populations. In: Davy AJ, Hutchings MJ, Watkinson AR (eds) Plant Population Ecology, Blackwell Scientific Publications: Boston. pp 47–58.

Schaal BA, Learn GH (1988). Ribosomal DNA variation within and among plant populations. Ann Mol Bot Gard 75: 1207– 1216.

Schlotterer CR (2000). Evolutionary dynamics of microsatellite DNA. Chromosoma 109: 365–371.

Schlotterer CR, Ritter R, Harr B, Brem G (1998). High mutation rate of a long microsatellite allele in Drosophila melanogaster provides evidence for allele-specific mutation rates. Mol Biol Evol 15: 1269–1274.

Smith DN, Devey ME (1994). Occurrence and inheritance of microsatellites in Pinus radiata. Genome 37: 977–983.

Tuskan GA, Francis KE, Russ SL, Romme WH, Turner MG (1996). RAPD markers reveal diversity within and among clonal and seedling stands of aspen in Yellowstone National Park, USA. Can J For Res 26: 2088–2098.

Udupa SM, Baum M (2001). High mutation rate and mutational bias at (TAA)n loci in chickpea (Cicer arietinum L). Mol Gen Genet 265: 1097–1103.

Van de Ven WTG, McNicol RJ (1996). Microsatellites as DNA markers in Sitka spruce. Theor Appl Genet 93: 613–617.

Vendramin GG, Lelli L, Rossi P, Morgante M (1996). A set of primers for the amplification of 20 chloroplast microsatellites in Pinaceae. Mol Ecol 5: 595–598.

Walbot V (1996). Sources and consequences of phenotypic and genotypic plasticity in flowering plants. Trends Plant Sci 1: 27–32.

Weber JL, Wong C (1993). Mutation of human short tandem repeats. Hum Mol Genet 2: 1123–1128.

Whitham TG, Slobodchikoff CN (1981). Evolution by individuals, plant–herbivore interactions, and mosaics of genetic variability: the adaptive significance of somatic mutations in plants. Oecologia 49: 287–292.

Tables

Table 1

Microsatellite loci examined in 12 genets of *Pinus strobus*

Microsatellite loci	Location	Repeat	Number of alleles observed
rps-6[a]	Nucleus	$(T)_6(AC)_n(T)_6$	5
rps-12[a]	Nucleus	$(AC)_n$	7
rps-20[a]	Nucleus	$(AC)_n(AT)_6$	8
rps-25b[a]	Nucleus	$(AC)_nAG(AT)_9$	2
rps-39[a]	Nucleus	$(AC)_n$	3
rps-50[a]	Nucleus	$(AC)_n$	9
rps-60[a]	Nucleus	$(AC)_n(AT)_7$	7
rps-118b[a]	Nucleus	$(AC)_n$	6
cpSSR-1[b]	Chloroplast	$(A)_n(G)_n$	2
cpSSR-5[b]	Chloroplast	$(T)_n$	1
Pt63718[c]	Chloroplast	$(T)_n$	2
Pt71936[c]	Chloroplast	$(T)_n$	2

[a]:Echt *et al.* (1996)

[b]:Cato and Richardson (1996)

[c]:Vendramin *et al.* (1996)

Table 2

Genotypes of ramets tested at nuclear- and chloroplast-encoded microsatellite loci

	Nuclear loci								Chloroplast loci
Genet	rps-6	rps-12	rps-20	rps-25b	rps-39	rps-50	rps-60	rps-118b	cpSSR 1/cpSSR 5/ Pt63718/Pt71936
77116	170[a]/170	169/189	138/140	113/113	170/170	170/176	251/267	138/138	239/102/98/146
77122	170/170	189/189	140/148	null/null	170/180	168/170	259/259	152/152	239/102/98/146
77130	169/170	185/185	138/142	111/111	170/170	170/172	251/271	152/152	240/102/98/146
78508	170/170	187/187	140/154	null/null	170/172	168/174	null/null	160/160	239/102/98/146
82531	169/170	185/185	140/152	111/111	172/172	170/172	null/null	152/152	240/102/98/146
84541	172/188	null/null	148/148	111/111	172/172	168/168	268/268	160/160	240/102/98/146
84544	170/196	187/187	138/140	null/null	170/172	160/170	264/264	142/176	239/102/98/146
84547	169/170	151/151	138/146	null/null	170/170	168/170	null/null	138/152	240/102/98/146
84548	170/170	185/185	138/154	null/113	172/172	168/180	269/271	152/152	239/102/99/146
84550	169/170	185/185	144/154	null/null	170/170	170/182	null/null	162/162	239/102/98/146
84554	169/170	157/185	152/152	111/111	170/172	170/174	259/259	162/162	240/102/98/146
845102	169/172	null/195	138/140	null/111	null/172	168/178	null/null	138/152	240/102/98/147

[a]:Allele sizes given as lengths (bp) of PCR products.

Table 3

Allele segregation in megagametophytes of the genets tested

Genet	Nuclear loci							
	rps-6	*rps-12*	*rps-20*	*rps-25b*	*rps-39*	*rps-50*	*rps-60*	*rps-118b*
77122	170,170	189,189	140,148	null,null	170,180	168,170	259,259	152,152
84548	170,170	185,185	138,154	null,113	172,172	168,180	269,271	152,152
845102	169,172	null,195	138,140	null,111	null,172	168,178	null,null	138,152

Table 4

Vegetative growth separating sampled cone positions

Genet	Vegetative growth (m)	Genet	Vegetative growth (m)
77116	21.13	84544	19.74
77122	20.65	84547	15.60
77130	19.73	84548	16.70
78508	21.22	84550	17.82
82531	19.51	84554	17.01
84541	20.41	845102	19.86

Figures

Figure 1

Scheme for sampling megagametophytes from individual ramets. Cones were collected from three positions per ramet, and megagametophytes were extracted from seeds within these cones. The total vegetative growth separating all sampled cones was determined by summing tree segments *a* through *h*.

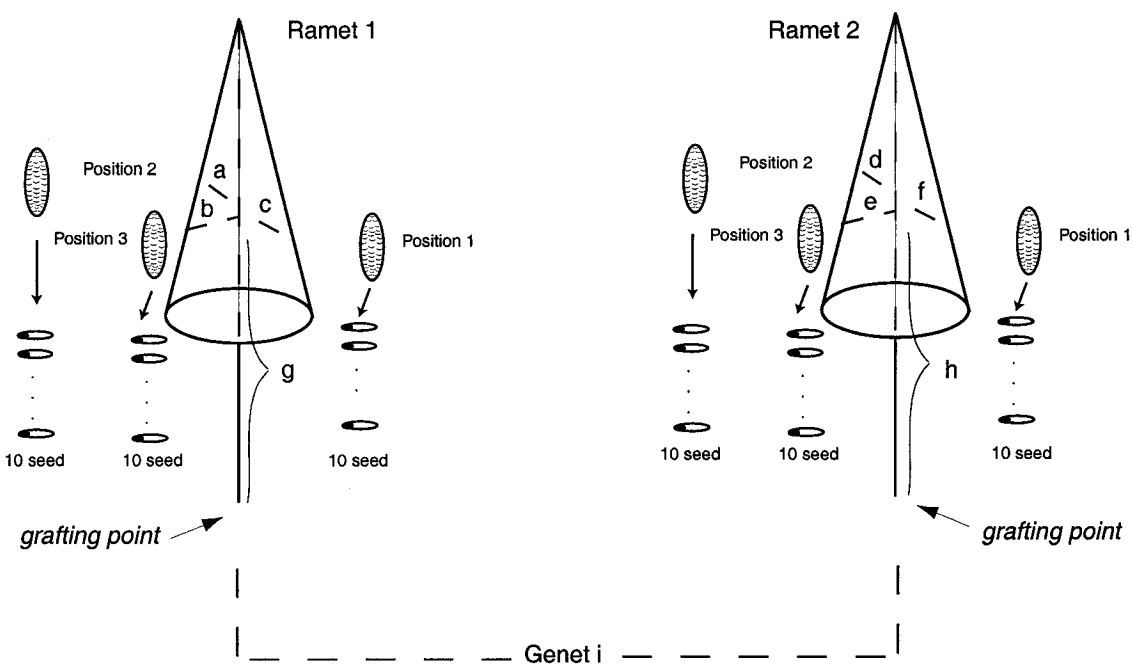
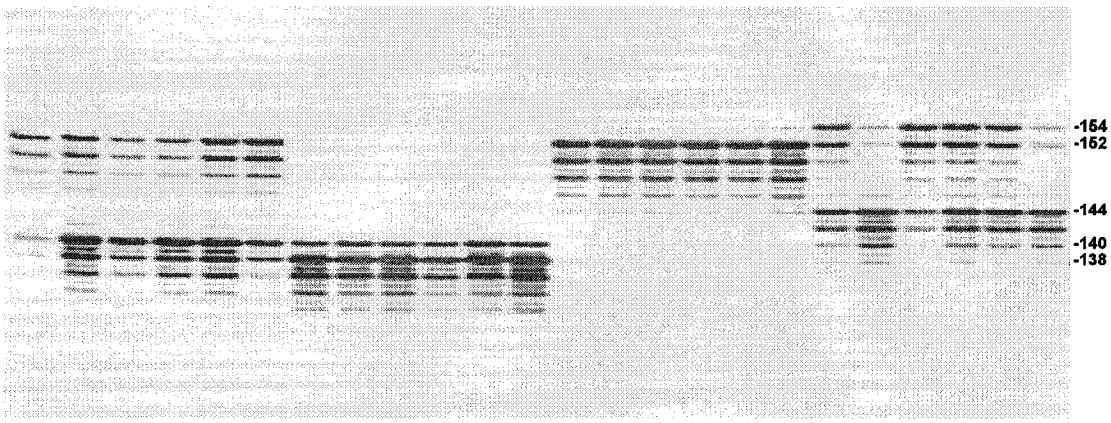


Figure 2

Gel image showing lack of evidence for microsatellite mutation at the *rps-20* locus within genets of *Pinus strobus*. From left to the right, the four genets are 82531 (lanes 1-6), 845102 (lanes 7-12), 84554 (lanes 13-18) and 84550 (lanes 19-24): Similar lack of variation was observed for all loci and genets tested. Numbers in margin represent length of the PCR product (bp).



Linking statement for Chapter 2

Having verified in Chapter 1 that microsatellite genetic markers are adequate and stable tools to measure inbreeding and gene flow within tree populations, it is of interest to apply those tools to study the population biology of a natural, relatively undisturbed system. Several studies have documented high levels of gene flow and low levels of inbreeding in tree populations, but the bulk of those studies have been conducted in temperate or boreal areas. Therefore, because of the relative paucity of information about gene flow and inbreeding in tropical plants, and because adequate description of the pre-disturbance situation is required to understand human impact on levels of genetic diversity in trees, I undertook a study to document those genetic parameters within populations of *Sextonia rubra* in French Guiana and *Carapa guianensis* in Brazil.

Chapter 2:

Low inbreeding and high pollen dispersal distances in populations of two Amazonian forest tree species

Cloutier D, Hardy OJ, Caron H, Ciampi AY, Degen B, Kanashiro M, Schoen DJ. Low inbreeding and high pollen dispersal distances in populations of two Amazonian forest tree species. ***Accepted in Biotropica***

Abstract

Recent studies suggest that tropical tree species exhibit low inbreeding and high gene dispersal despite the typically low density of conspecifics in tropical forests. To examine this, we have undertaken a study of pollen gene dispersal and mating system of two Amazonian tree species. We analyzed 341 seeds from 33 trees at four microsatellite loci in a *Carapa guianensis* population from Brazil, and 212 seeds from 22 trees at four microsatellite loci in a *Sextonia rubra* population from French Guiana. Differentiation of allele frequencies among the pollen pool of individual trees was $\Phi_{FT} = 0.053$ (95%CI:0.027-0.074) for *C. guianensis* and $\Phi_{FT} = 0.064$ (95%CI:0.017-0.088) for *S. rubra.* The mean pollen dispersal distances were estimated at between 69 and 355 meters for *C. guianensis*, and between 86 and 303 meters for *S. rubra*, depending on the pollen dispersal model and the estimate of reproductive tree density. The multi-locus outcrossing rate was estimated at 0.918 and 0.945, and the correlation of paternity at 0.089 and 0.096, for *C. guianensis* and *S. rubra* respectively, while no significant levels of biparental inbreeding were detected. Comparing trees with high and low local density of conspecifics, we found no evidence for differences in inbreeding levels. The results are discussed within the framework of the emerging picture of the reproductive biology of tropical forest trees.

Résumé

De récentes études suggèrent que les populations d'arbres tropicaux sont caractérisées par une faible consanguinité et une dispersion des gènes à grande distance, malgré la faible densité des conspécifiques particulière aux forêts tropicales. Afin de vérifier ceci, nous avons étudié la dispersion des gènes par le pollen et le régime de reproduction de deux espèces d'arbres en Amazonie. Nous avons analysé 341 graines de 33 arbres à quatre loci microsatellites d'une population de *Carapa guianensis* du Brésil,

et 212 graines de 22 arbres à quatre loci microsatellites d'une population de *Sextonia rubra* de Guyane Française. La différentiation de fréquence des allèles entre les nuages de pollen captés par les arbres était de Φ_{FT} = 0.053 (95%IC:0.027-0.074) pour *C. guianensis* et de Φ_{FT} = 0.064 (95%IC:0.017-0.088) pour *S. rubra.* La distance moyenne de pollinisation a été estimée à des valeurs entre 69 et 355 mètres pour *C. guianensis*, et entre 86 et 303 mètres pour *S. rubra*, selon le modèle de dispersion du pollen et la valeur de densité efficace d'arbres utilisés. Le taux d'allo-pollinisation a été estimé à 0.918 et à 0.945, et la corrélation de paternité à 0.089 et à 0.096, pour *C. guianensis* et *S. rubra* respectivement, tandis qu'aucune consanguinité biparentale n'a pu être détectée. En comparant des arbres à forte et à faible densité locale de conspécifiques, nous n'avons pas constaté de différences dans les niveaux de consanguinité de leur descendance. Les résultats sont discutés au regard de l'état actuel des connaissances de la biologie de la reproduction des arbres de la forêt tropicale.

Introduction

The very first predictions about mating and pollen dispersal in tropical tree species were that the relatively low population densities would lead to high levels of inbreeding and limited pollen dispersal due to restricted among-tree movement of pollinators (*e.g.* Baker 1959, Federov 1966). However, direct investigations of tropical tree breeding systems soon revealed that obligate outcrossing is the main mode of reproduction in tropical trees (*e.g.,* Bawa 1974). Moreover, early studies based on mark-recapture experiments with pollinators and direct observations of pollinator flight patterns revealed that some pollinators of tropical trees forage over large distances and thereby may mediate pollen flow among widely spaced conspecific trees (Bawa 1990).

Genetic markers such as allozymes or microsatellites have been used recently to infer realized mating and gene flow patterns in tropical tree species. Such investigations have revealed that most tropical tree species are predominantly outcrossing (reviewed in Doligez & Joly 1997, Nason & Hamrick 1997, Loveless 2002, Ward *et al.* 2005),

although a number of species showed a significant proportion of progeny produced by selfing. Other genetic marker-based studies suggest that pollen movement is the principal contributor to gene flow and can occur over large distances (reviewed in Nason & Hamrick 1997). For instance, a paternity analysis unambiguously showed that animal-mediated movement of pollen can occur at a distance of 3.2 km (Dick 2001), but estimates of average pollination distances are usually in the range of a few hundreds of meters (reviewed in Nason & Hamrick 1997, Loveless 2002). Other studies based on comparisons among populations suggest that pollen dispersal (Stacy *et al.* 1996, Degen *et al.* 2004) and selfing rate (Murawski & Hamrick 1991, 1992) are negatively correlated with tree density. As well, comparisons of pollen flow in undisturbed and disturbed habitats suggest that larger among-tree distances resulting from habitat fragmentation lead to increases in pollination distances and enhance pollen flow (Chase *et al.* 1996, Dick 2001, White *et al.* 2002, Dick *et al.* 2003), though in some cases, artificial reduction of tree density by selective logging increases levels of self-fertilization (Murawski *et al.* 1994, Doliguez & Joly 1997). Indirect estimates of historical gene dispersal distances also found a negative relationship with local population densities (Hardy *et al.* 2006).

Diverse factors such as the prevalence of animal pollination and the interactions between pollinator and plant populations, the diversity of breeding systems, the lack of seasonality of pollination at the community level, and the typically shorter longevity of the flowers may impact mating and gene flow patterns in tropical tree populations (Bawa 1990). Several authors have stressed the importance and possible application of understanding mating and gene flow patterns in the context of sustainable management of tropical forests (Bawa 1990, Ledig 1992, Kanashiro *et al.* 2002, Loveless 2002). Moreover, recently developed and powerful analytical tools are available (Austerlitz & Smouse 2001, Smouse *et al.* 2001, Ritland 2002), but relatively few attempts have been made to apply these to studies of the mating system and gene dispersal in high biodiversity Amazonian forests.

Our aim was to examine the predictions that Amazonian tropical trees in continuous forests are highly outcrossing and effectively disperse their pollen over large distances. As such, we selected two canopy forest tree species, *Carapa guianensis* from Brazil and *Sextonia rubra* from French Guiana, sharing two ecologically relevant

features: (1) an unspecialized pollination system based on small diverse insects, and (2) a high local population density (>1tree/ha) relative to other Amazonian tree species. In addition to these features, both species are commercially exploited for timber in the local and international markets, and as such, they are representative of many exploited tree populations. Our results are discussed within the framework of the emerging picture of the reproductive biology of tropical forest trees. Specifically, this paper examines the following predictionss regarding these species: (1) tropical insect pollinators disperse pollen to other conspecifics over hundreds of meters, even though the small size of the pollinators suggests restricted between-tree movement; (2) tropical tree populations are largely outcrossing, and have limited opportunities for mating among relatives in natural settings, even though tropical tree species have sexual systems (*i.e.,* hermaphrodite or monoecious) that enable self-fertilization; (3) a large number of fathers pollinate maternal parent trees in a single reproductive episode, even though factors such as population density and reproductive phenology suggest that near-neighbor mating may prevail; and (4) maternal parent trees with reduced local density of conspecific neighbors produce seed crops with higher levels of inbreeding, depending on the reproductive biology of the tree species.

Materials and methods

Study species and study sites

Carapa guianensis Aublet (Meliaceae) is a canopy tree species widely distributed in the Caribbean islands, Central America, and northern South America. It can reach one meter in diameter at breast height (DBH) and 30-40 meters in height. In the Brazilian Amazon, it is found in various habitats, from seasonally flooded river edges to dry uplands, and is occasionally planted by local indigenous and riverine communities where

it grows. In Brazil, it is an important timber tree species, and the oil extracted from the seed is extensively traded and used for its medicinal properties.

Carapa guianensis is a monoecious species harboring both male and female flowers on the same tree. The flowers are small, white, and are found on conspicuous inflorescences. During several hours of observation on a few branches of a tree, small stingless bees, beetles and moths have been seen in small numbers visiting the flowers in the morning (D. Cloutier, pers. obs.; M. Maues, pers. comm.), suggesting that it is pollinated by small insects. The fruit is a four-valved capsule containing 4-20 seeds weighting 20-40 g each. The seed may float on water and be dispersed over large distances in flood-prone forests (Scarano *et al.* 2003, Cloutier *et al.* 2005). In the upland *terra firme* populations, the seeds are dispersed by gravity and secondarily by medium-sized scatterhoarding rodents for distances less than 25 meters (Guariguata *et al.* 2002). This species flowers intensively from August through May in the Tapajós river area, and fruits are released from January through July.

The study site for *Carapa guianensis* is Flona Tapajós (03°01'S 54°58'W), Para state, Brazil. Under the auspices of the Dendrogene project (Kanashiro *et al.* 2002), an intensive study plot is currently being monitored within the Flona Tapajós. This is part of a larger, ongoing study focusing on genetic conservation in managed forests of Amazonia. Prior to this study, a commercial pre-logging inventory was conducted in the study plot (500 ha), and every tree larger than 20 cm DBH was identified and mapped. A mean density of 2.5 *C. guianensis* trees/ha ≥ 30 cm DBH was found in the 400 ha of the central plot area (Figure 1), indicating that *C. guianensis* has the second highest average tree density at this site.

Sextonia rubra (Mez) van der Werff (Lauraceae) is a canopy tree species reaching one meter DBH and 40 meters in height. It is found in the Guyanas and the Eastern Amazon basin. It is one of the most intensely harvested timber tree species in French Guiana.

Plants of *Sextonia rubra* produce small hermaphroditic, scented, white flowers in August and September in French Guiana. The pollinators of *S. rubra* have not been observed, but the size and the color of the flowers suggest that pollination is mediated by small insects (*sensu* Bawa *et al.* 1985). The fruits contain a single seed and are produced

in January in French Guiana. It has been reported that the seeds are dispersed by birds (Hardy *et al.* 2006), but the bulk of seeds are released on the ground beneath the trees.

The study site for *Sextonia rubra* is Paracou (5°18'N, 52°53'W), French Guiana, France. Paracou is composed of 16 non-contiguous permanent study square plots (15 of 6.25 ha plus 1 of 25 ha; Figure 1) administered since 1984 by CIRAD-Forêt (French Agricultural Research Centre for International Development-Forestry Department) to study forest regeneration under natural and logging conditions (Schmitt & Bariteau 1988). All trees larger than 10 cm DBH have been identified and mapped in the plots. A total of 154 *S. rubra* trees above 30 cm DBH are found in the 119 ha of the Paracou plots, for an estimated average density of 1.3 trees/ha.

Field sampling

For *Carapa guianensis*, fruits were collected directly beneath 33 maternal parent (mother) trees between April and July 2003 at the Tapajós site. These trees were selected to provide a range of inter-individual distances, from trees that are close enough to receive pollen from the same source trees, to those far apart and likely to receive pollen from different source trees. The inter-tree distances varied between 6 m and 1883 m, averaging at 722 m. The number of fruits collected from each tree varied from 8 to 13, for a grand total of 341 progeny analyzed. A separate genetic analysis conducted on *C. guianensis* multi-seeded fruits collected in 2004 at Tapajós revealed that seeds within individual fruits are significantly more likely to share the same father than are seeds taken from different fruits of the same mother tree (D. Cloutier, unpublished data), suggesting correlated pollen dispersal events (*i.e.,* pollen co-dispersal from the same male parent(s) to single flowers). Thus, to sample as many independent pollination events as possible, only one seed per fruit was genotyped. In order to get information on the *C. guianensis* adult trees, a piece of cambium (~150mg) was extracted from 199 trees with known spatial location, including or distributed around the mother trees from which seed were collected. The cambium samples were stored in a buffer of 70% ethanol and 0.3% β-mercaptoethanol before being sent to the lab.

For *Sextonia rubra*, single-seeded fruits were collected directly beneath mother trees in January 2000 at Paracou. Preliminary trials suggested that the DNA extracted from roots of germinated seedlings gave the best PCR results when used at low concentrations. Therefore, seeds were planted and the roots of the seedlings were collected and kept in the freezer until DNA extraction in 2002. Fewer seed genotypes were available than in the *C. guianensis* population — the final data set was composed of 212 seeds taken from 22 mother trees, with each of the 22 seed families containing between 2 and 15 seeds. For population level analyses of mating system, the 22 families were used, while only the 15 families with at least 8 seeds genotyped were used for pollen dispersal analyses. The inter-tree distances of the 15 families with at least 8 seeds genotyped varied between 33 m and 2128 m, with a mean of 833 m.

The maternal parent trees of the fruits collected on the ground were identified with certainty in most cases—at both sites, there was no evidence that fruits were previously handled by animals, and because large distances separate conspecific trees, the probability that seed shadows overlap is low. In a few cases, however, fruits could have been incorrectly assigned to a mother tree because of the proximity of a conspecific, or because of incorrect labeling of the fruits sampled. To rule out incorrect assignment of progeny to the mother trees, and to check the inheritance and segregation of alleles at the microsatellite loci used, we compared mother tree and seed genotypes to detect and remove mismatched individuals from the data set. Maternal parent genotypes were available from the adult tree sample of *C. guianensis*, while most maternal parent genotypes of *S. rubra* were known from another study (Hardy *et al.* 2006). The genotypes of the few trees missing were inferred from seed family genotypes (Brown & Allard 1970).

Laboratory analyses

Carapa guianensis cambium and seed samples were brought into the plant genetic laboratory of EMBRAPA-CENARGEN (Brazilian Agricultural Research Corporation - Center for Genetic Resources and Biotechnology) in Brasilia, Federal District, Brazil.

The samples were ground in liquid nitrogen and total genomic DNA was extracted from cambium samples, fresh embryos and dried leaves of germinated seed using a standard CTAB-chloroform-isoamyl alcohol protocol. Genotypes were scored at four microsatellite loci (*i.e.*, Cg1, Cg6, Cg16 and Cg17), described by Vinson *et al.* (2004). The PCR amplification cocktail (total volume of 13 μL) contained: PCR reaction buffer (Invitrogen), forward (labeled with fluorescence) and reverse primer (0.28 μM), $MgCl_2$ (1.5 mM), BSA (0.25 mg/mL), dNTP (0.25 mM), Taq DNA polymerase (1.3 U, Invitrogen), and DNA (~3 ng). PCR conditions were: denaturation at 94°C for 5 min; 35 cycles of denaturation at 94°C for 1 min, annealing at 53-59°C for 1 min, and extension at 72°C for 1 min, followed by a final extension at 72°C for 30 min. Optimal annealing temperature and concentrations of primers, DNA template and $MgCl_2$ were adjusted empirically for each locus. Reaction products were separated by electrophoresis on an ABI 377 automatic sequencer (Applied Biosystems, USA) jointly with a custom-made size standard. Genotypes were determined using the software Genotyper (Applied Biosystems, USA).

Sextonia rubra seed samples were brought to the laboratory of INRA (French National Institute for Agricultural Research) at Kourou, French Guiana, France. DNA was extracted using a standard CTAB-chloroform-isoamyl alcohol protocol from frozen roots of germinated seed grounded in liquid nitrogen. Genotypes were scored at four microsatellite loci (*i.e.*, Sr3, Sr9, Sr10 and Sr51) developed by H. Caron (Veron *et al.* 2005). The PCR cocktail (10.0 μl total) contained reaction buffer (Invitrogen), 0.04 μM of each dNTP, 0.01-0.04 μM of $MgCl_2$, 0.5 U of Taq DNA polymerase (Invitrogen), 0.001-0.01 μM of each primer (forward primer labeled with fluorescence) and ~0.2 ng of DNA. The PCR protocol was: 5 min at 94°C; 35 cycles of 45 sec at 94°C, 45 sec at 55-60°C and 45 sec at 72°C; ending with 45 min at 72°C. Optimal annealing temperature and concentrations of primers, DNA template and $MgCl_2$ were adjusted empirically for each primer pair. The amplification products were determined by electrophoresis with a ROX-500 size standard on an ABI 310 automated sequencer (Applied Biosystems, USA). Allele sizes were scored using the program Genotyper (Applied Biosystems, USA).

Standard genetic diversity parameters were estimated from the seed samples for the microsatellite loci used with the program GENEPOP (Raymond & Rousset 1995).

Using the same program, no linkage disequilibrium among the *C. guianensis* loci was detected in the adult sample, while no sample was available to test this for *S. rubra*.

Adult genetic structure analyses

The mating system and pollen dispersal analyses presented in this study require some knowledge of the genetic structure of the adult trees of each population. The adult tree sample of *Carapa guianensis* was analyzed, therefore, using the program SPAGEDI (Hardy & Vekemans 2002) to estimate adult fixation index (F_{IS}) and the statistic "*Sp*". This statistic, based on the regression of genetic similarity between pairs of individuals over the natural logarithm of the physical distance separating them, allows quantitative comparison among species of the intensity of spatial genetic structure (SGS), and is described in detail by Vekemans & Hardy (2004). The standard error for *Sp* was obtained by jackknifing genetic loci and the *Sp* value was tested against the null hypothesis of random distribution of genotypes (*i.e.* no SGS) by permuting the spatial positions of the adults 1000 times. The genetic structure estimates of the *Sextonia rubra* adult trees from Paracou are published elsewhere (Hardy *et al.* 2006) and gave a F_{IS} of 0.076, and a *Sp* value of 0.005 (0.001), indicating very weak (but statistically significant) SGS among the adult trees.

"TwoGener" pollen dispersal analyses

Pollen dispersal was analyzed with the "TwoGener" algorithm developed by Smouse *et al.* (2001) using a program obtained from F. Austerlitz (Université Paris-Sud, France). The statistic ϕ_{FT}, which measures the differentiation in allelic frequencies among the pollen clouds sampled by different seed parents whose spatial coordinates are known, was computed to estimate pollen dispersal parameters. The formal relationship between ϕ_{FT} and pollen dispersal parameters has been derived for different pollen flow models (Austerlitz & Smouse 2001). A global ϕ_{FT}, based on average pollen pool differentiation

and average physical distance over all pairs of sampled mothers, or a pairwise ϕ_{FT}, based on pollen pool differentiation and physical distance between each pair of sampled mother, can be used to estimate pollen dispersal parameters.

First, we estimated the effective number of pollen donors contributing to the progeny of the average mother tree (N_{ep})—this is derived directly from the global estimate of ϕ_{FT} as $N_{ep} = 1/2\phi_{FT}$ (Smouse *et al.* 2001, Austerlitz & Smouse 2001). Second, we estimated δ (*i.e.*, the average distance of realized pollen dispersal) from the ϕ_{FT} estimates (Austerlitz & Smouse 2001, Austerlitz & Smouse 2002, Austerlitz *et al.* 2004). The global estimate of ϕ_{FT} from the whole set of sampled mother trees (Austerlitz & Smouse 2001), or the pairwise estimates of ϕ_{FT} between each pair of sampled mother tree (Austerlitz & Smouse 2002, Austerlitz *et al.* 2004) are used to estimate δ by assuming an underlying density of reproducing trees (d) and a pollen dispersal model (*i.e.*, normal, exponential or power exponential). Because N_e/N ratios in plant populations vary from values close to one, to values as low as 1/10 (*e.g.*, due to varying fertility and non-synchronous flowering) (Frankham 1995), we estimated δ from global ϕ_{FT} assuming an upper bound estimate of the density of reproducing trees (d_{max}) (*i.e.*, the global density of conspecific trees above 30 cm DBH in the field), and a lower bound estimate of d ($d_{min} = d_{max}/10$). The pairwise ϕ_{FT} can also be used to jointly infer δ, d, and the shape of the pollen dispersal curve (*i.e.*, whether the curve is exponential, thin-tailed and decreasing faster than the exponential, or fat-tailed and decreasing slower than the exponential). A least squares criterion is used to optimize the fit between observed pairwise ϕ_{FT} values and those expected from the pollen dispersal model. The residual error for the different dispersal models tested can be compared to assess which provides the better fit to the data (Austerlitz *et al.* 2004). To assess the 95 percent confidence interval of the global ϕ_{FT} dispersal estimates and the pairwise ϕ_{FT} dispersal estimates based on the normal model, 100 bootstrap estimates were calculated using the seed families as the resampling unit. Because of the prohibitive amount of computer processing time required for the calculations, the bootstraps for the pairwise ϕ_{FT} were not performed for the exponential and the power exponential model of pollen flow.

Mating system analyses

The mating system analyses were done using the mixed-mating model with the MLTR program (Version 3.1, Ritland 2002). This program calculates maximum likelihood estimates of: (1) the rate of outcrossing; (2) the level of biparental inbreeding—this is estimated as the difference between multi- and single-locus estimates of outcrossing rate, using the fact that the likelihood of confusing selfed and biparentally inbred progeny is reduced when the number of loci used in the estimation increases; (3) the correlation of paternity, *i.e.* the extent to which siblings share the same male parent. We assumed no apomixes and no seed mortality due to inbreeding prior to the assay. The MLTR program options were set to their default values, except that the parental fixation indices were constrained to values estimated from the adult tree samples (*i.e.*, 0.029 for *Carapa guianensis* and 0.076 for *Sextonia rubra*). The variance of the MLTR estimates was estimated by bootstrapping 1000 times at the level of seed family.

Two types of mating system analyses were conducted. In the first type, we estimated the mating system at the level of the whole population. In the second type, we assessed the impact of within-population variation in local density of conspecific neighbors on the mating system. To do this, we classified each mother tree into one of two groups: (1) trees with higher neighbor density than global density; and (2) trees with lower density of neighbors relative to the global density. "Neighbors" were defined as conspecific trees found within the average pollination distance of the mother trees, as estimated by TwoGener when effective density and pollen dispersal are jointly estimated. Since the study plots at the Paracou field site were small and not contiguous, estimates of neighbor density for *S. rubra* were based on individual plot tree density.

Results

Microsatellite loci

Table 1 shows the allele frequencies, gene diversities (H_e), observed heterozygosity (H_o), and fixation indexes (F_{IS}) obtained with the four microsatellite loci used in each species. We detected high levels of gene diversity, varying between 0.42 and 0.83 for *Carapa guianensis*, and between 0.80 and 0.86 for *Sextonia rubra*, indicating that the microsatellite loci diversity is adequate for mating system and pollen dispersal estimation (Ritland 2002, Austerlitz & Smouse 2002).

Adult genetic structure

The fixation index (F_{IS}) of *Carapa guianensis* adult trees was 0.029 and the value of the *Sp* statistic was 0.0045 (SE:0.0024). This very weak level of spatial genetic structure was, nevertheless, statistically significant with the permutation test at the 0.05 level.

Pollen dispersal

The global estimate of pollen pool differentiation (ϕ_{FT}) was 0.053 for *Carapa guianensis* with a 95 percent confidence interval of 0.027-0.074, and 0.064 (95%CI: 0.017-0.088) for *Sextonia rubra*. Both values are significantly greater than 0, indicating significant substructure of pollen pool allele frequencies. These global estimates of ϕ_{FT} yield estimates of effective number of fathers per seed parent (N_{ep}) of 9.4 (6.8-18.5) for *C. guianensis,* and of 7.8 (5.7-29.4) for *S. rubra.*

The results for the mean pollen dispersal distance (δ) based on global estimates of ϕ_{FT} are presented in Table 2. Using the estimated density of trees above 30 cm DBH in size (d_{max}), and assuming the normal model of pollen dispersal, we estimated δ at 69 (58-96) m for *C. guianensis* and 86 (74-167) m for *S. rubra.* Using the estimated minimum effective density (d_{min}), δ is estimated at 218 (184-288) m for *C. guianensis* and 271 (233-

448) m for *S. rubra.* Assuming the exponential model, we obtained slightly larger values of 78 (66-109) m for *C. guianensis* and 97 (83-188) m for *S. rubra* using d_{max}, compared with 244 (207-317) m for *C. guianensis* and 303 (262-491) m for *S. rubra* using d_{min}. Overall, both species have similar levels of pollen dispersal that are influenced more by the estimate of density than by the dispersal model assumed.

The results of the pairwise estimates of ϕ_{FT} are presented in Table 3. Overall, these results are similar to those derived from the global ϕ_{FT} estimates, with estimates of δ varying between 71 and 355 m for *C. guianensis,* and between 81 and 287 m for *S. rubra.* The joint estimation of density and dispersal distance suggests that the effective density of reproducing individuals is lower than the estimated field density for both species—approximately an order of magnitude lower than the true density of reproductive trees in the population. For *C. guianensis*, the exponential power model of pollen dispersal provides the best fit to the observed data. The joint estimation of the three parameters gave a shape parameter of 0.77 (indicating that the dispersal curve is more fat-tailed than an exponential distribution), an estimated effective density of 0.14 trees/ha, and a mean pollen dispersal distance of 338 m. For *S. rubra*, the best fit is the normal model, which gave an estimated effective density of 0.14 trees/ha, and a mean pollen dispersal distance of 231 m. The program failed to converge in the case of the exponential power model in the case of *S. rubra,* possibly because of the reduced data set for this species. The 95 percent confidence intervals on the normal model parameters jointly estimated by the pairwise analysis were large, with δ (43-514) m and d (8.2–0.018) tree/ha for *C. guianensis*, and δ (49-460) m and d (3.1-0.024) tree/ha for *S. rubra.*

Mating system

The results of the population level mating system analyses are presented in Table 4. Similar results were obtained for both species. The multi-locus outcrossing rate was estimated at 0.918 for *Carapa guianensis* and at 0.945 for *Sextonia rubra*. Both values are significantly less than 1.0 suggesting that these species are able to self-fertilize. We

did not detect any significant levels of biparental inbreeding in either species, as measured by the difference between multi- and single-locus estimates. The estimated multi-locus correlation of paternity was low for both species, but significantly different than 0.

The impact of local variation in density of neighbors on the mating system is presented in Table 5. For both species, we found modest variation in the density of conspecific neighbors within pollination neighborhood of individual mother trees, with values between 1.8-3.2 neighbors/ha for *C. guianensis,* and between 0.5-2.6 neighbors/ha for *S. rubra.* When the mother trees were classified into two groups of lower and higher neighbor density, *C. guianensis* had a mean of 2.2 (16 mother trees) and 2.8 neighbors/ha (17 mother trees), respectively, while *S. rubra* had a mean of 1.0 (8 mother trees) and 2.6 neighbors/ha (10 mother trees). When estimating mating system parameters separately for those two groups, we found no evidence for an impact of neighbor density on inbreeding levels. The estimates of outcrossing rate, level of biparental inbreeding, and correlation of paternity for both species were not significantly different between the groups (0.05 level). For *S. rubra*, the correlation of paternity was only marginally higher for mother trees with lower local density of neighbors compared to trees with higher local density of neighbors.

Discussion

Pollen dispersal

We found low pollen pool differentiation and high levels of pollen dispersal for both *Carapa guianensis* and *Sextonia rubra*, with estimates of average pollination distances (δ) varying between ~70 and ~350 meters. This suggests that trees in both species mate with a large number of conspecifics, distributed over a large area in their vicinity, despite (or owing to) the large distances separating conspecifics in tropical

forests. The pollen dispersal distance estimates from this study appear larger than other high-density (*i.e.*, > 10 trees/ha) tropical trees found in similar habitats, such as *Symphonia globulifera* (Degen *et al.* 2004), and also larger than high-density tropical species with different ecological requirements, such as *Gliricidia sepium* (Dawson *et al.* 1997) and *Cordia alliodora* (Boshier *et al.* 1995). On the other hand, the pollen dispersal distances in *C. guianensis* and *S. rubra* are similar to other tropical tree species in continuous undisturbed rainforest, commonly found at low-density (*i.e.*, < 1 tree/ha), such as *Dinizia excelsa* (Dick *et al.* 2003), *Swietenia humilis* (White *et al.* 2002) or *Neobalanocarpus heimii* (Konuma *et al.* 2000). This comparison of our results to previous studies adds support to the hypothesized negative relationship between population density and gene dispersal, and suggests that density variation below ~5 trees/ha may not have a strong impact on the mean distance of pollen dispersal, and that other factors related to the reproductive biology of the tree species may be more important.

It is worth noting that our estimates of pollen dispersal distances, which represent "ongoing" dispersal, are substantially lower than "historical" estimates of gene dispersal distances derived from the observed level of spatial genetic structure (SGS) among adult trees. For *S. rubra*, the historical estimate of gene dispersal was between 728 m and 848 m (Hardy *et al.* 2006), while even higher estimates may be expected for *C. guianensis* since its SGS is weaker than *S. rubra*. This suggests that the contribution of seed dispersal to global gene flow may be significant, at least for *S. rubra*, whose seeds are potentially dispersed by birds. Moreover, at least for *C. guianensis*, whose seed are unlikely to be dispersed at large distances at the study site, this suggests that gene dispersal estimated from a single episode of reproduction (*e.g.*, this study) is more restricted than the cumulative effect of gene flow over hundreds of years, as measured by indirect estimates of historical gene dispersal based on spatial genetic structure. In other words, ecological time scale estimates of gene dispersal, such as obtained in our study, may be lower simply because they measure a different parameter, with different effective density perhaps, than historical estimates of gene dispersal based on levels of SGS among adults.

Among the assumptions required to estimate pollen dispersal distances using the TwoGener algorithm (see also Hardy *et al.* (2004) for a discussion of additional aspects), the two that have potential to bias estimates are the pollen dispersal model and the tree density used. For instance, models with a thin-tailed dispersal curve (*e.g.*, the normal model) may result in the underestimation of the true dispersal distances in cases when the actual pollen dispersal curve is fat-tailed. However, estimates of δ that assume different dispersal models are actually quite similar (Tables 2 and 3), suggesting that any bias introduced by these assumptions may be low. In comparison, the assumed tree density has a significant impact on δ estimates. Since some trees in the population may be more successful than others at pollinating their neighbor flowers, simply using trees above 30 cm DBH to estimate density may lead to an overestimate of the effective tree density, and an underestimate of mean pollen dispersal distances (Austerlitz & Smouse 2002). For that reason, we have estimated δ: (1) with global ϕ_{FT} using maximal (d_{max}) and minimal (d_{min}) estimates of the effective population density, providing, respectively, lower and upper bound estimates of pollen dispersal distance; and (2) with pairwise ϕ_{FT} to jointly infer density and dispersal distance.

The joint estimation revealed that the effective density of reproducing trees was about one tenth that of the estimated field density of trees. This is among the smallest N_e/N ratios observed in plant populations (Frankham 1995). However, the 95 percent confidence intervals estimated for the effective density estimates were very large (*i.e.*, between about 8 and 0.02 trees/ha for both species), reflecting the fact that data sets larger than ours are needed to estimate *d* precisely (Austerlitz & Smouse 2002). Nonetheless, the low estimated *d* likely indicates that there is variation in male reproductive success in our populations. Tree crown observations to determine which individuals were flowering in a sample of *C. guianensis* trees in 2003, and *S. rubra* trees in 2002, revealed that about 75 percent of trees above 30 cm DBH at the flowering peak had the potential to exchange pollen (Dendrogene project for *C. guianensis*, unpublished data; D. Cloutier and O. Hardy for *S. rubra*, unpublished data), which suggests that some sampled trees were not perfectly synchronized with the population. Paternity analyses of animal-pollinated tropical trees (*e.g.*, Konuma *et al.* 2000) also revealed some asymmetry among pollen contributors to a single mother, which could not be explained by "isolation by distance"

alone. Other factors, such as variation in the number of male gametes produced by each pollen donor, may play an additional role in fertility variation if there are differences in tree size and/or access to light.

Mating system

The high outcrossing rates found in this study ($t_m > 90\%$) fall within the range of typical values estimated for tropical tree populations (Loveless 2002). Some significant, but low, selfing levels were detected in this study, as in other studies of tropical tree populations (Loveless 2002), suggesting that *Carapa guianensis* and *Sextonia rubra* are able to self-fertilize. It is worth noting that because of the potential presence of inbreeding depression, the observed level of selfing may depend on the developmental stage at which the progenies are analyzed. Thus, we may have underestimated the level of self-fertilization for both species because we measured inbreeding levels in the seeds (for *C. guianensis*) or in the seedlings (for *S. rubra*), *i.e.* after early-acting inbreeding depression may have selected out some inbred progeny. Our results nevertheless reinforce the perception of a predominance of outcrossing in tropical tree mating systems that has developed over the past three decades (Loveless 2002, Ward *et al.* 2005).

When inferring levels of biparental inbreeding from differences between single and multi-locus estimate of outcrossing rate (Ritland 2002), we do not detect any significant levels of biparental inbreeding in our populations. In the case of *S. rubra*, this result may be due to the small data set we used for the estimation, as another study reported significant levels of biparental inbreeding for this species (Veron *et al.* 2005). Weak levels of biparental inbreeding could be interpreted as a consequence of extensive pollen flow in these populations, and weak genetic structure among reproducing trees, since opportunities for mating among near-neighbor relatives in such a situation may be infrequent.

We estimated the correlation of paternity at less than 0.1 in both *C. guianensis* and *S. rubra*. This indicates that the seed crop of individual trees is composed mostly of half-sibs. For both species, the values of correlation of paternity are approximately two times

the values of ϕ_{FT}, reflecting the fact that these statistics estimate the same quantity (Hardy *et al.* 2004). In the case of tropical trees, which are mostly animal-pollinated, we might expect varying levels of correlated paternity depending on the reproductive biology of the species, but most tropical rainforest trees studied to date show a pattern of low correlated paternity (*e.g.,* Doligez & Joly 1997). The large pool of potential pollen donors inferred for *C. guianensis* and *S. rubra* likely explains the low level of correlated paternity observed in these species.

Assuming pollinator behavior follows an optimal cost/benefit foraging strategy, reduced density of flowering trees has been hypothesized to increase levels of inbreeding (Murawski & Hamrick 1991, 1992; Murawski *et al.* 1994). We found no evidence, however, for an impact of local density of neighbors on mating system parameters within our populations. Moreover, applying a different approach (Hardy *et al.* 2004) which estimates correlated paternity for each mother tree and regresses this value on local density, again we found no significant impact of local density on the correlation of paternity (D. Cloutier, unpublished data). This could be partly explained by the scale of effective pollen dispersal, and by the relative abundance of conspecifics for *C. guianensis* and *S. rubra* compared to other tree species found in the same sites. For both species, we found extensive pollen dispersal, *i.e.* beyond the distances separating the mother trees to their immediate near-neighbors. Since the mother trees are pollinated by a large pool of potential pollen donors distributed over a large area, local variation in neighbor density within the pollination neighborhood may be smoothed out, and thus may be insufficient to have significant consequences on variation in the outcrossing rate. On the other hand, the correlation of paternity seemed to be more influenced by variation in conspecific density, but consequences on inbreeding levels appear to be marginal at best. In the case of *S. rubra*, the small data set and the lack of information about the density of conspecifics in between the plots at Paracou also limited our ability to detect any impact of neighbor density on inbreeding levels in the first place.

Comparisons of our results with other studies on *C. guianensis* and *S. rubra* may also help to detect potential mating system consequences associated with population density. An allozyme study in Costa Rican populations of *C. guianensis* with more than 10 trees/ha revealed high outcrossing rates (0.97-0.99) (Hall *et al.* 1994). Another

microsatellite study in the same area, based on adult and sapling genotypes, found that the population is completely outcrossed (Dayanandan *et al.* 1999). Comparison of our Amazonian low-density population to the Costa Rican high-density populations suggests that higher outcrossing rates may be favored by an increase in population density. In other studies, there was no evidence for reduced inbreeding in high density populations of the neotropical *Symphonia globulifera* in French Guiana (Degen *et al.* 2004) compared to lower density populations in Panama (Stacy *et al.* 1996). Furthermore, a genetic study of *S. rubra* seeds at Paracou, collected two years after those analyzed in this study, revealed similar levels of pollen dispersal, but slightly higher outcrossing rate (0.99) with significant levels of biparental inbreeding (Veron *et al.* 2005), despite having the same field population density between the two years. The limited precision of estimates obtained from genetic data may explain some of the differences between studies, but also these differences suggest that tree density is not the only factor influencing the mating system, and that non-overlapping phenologies, variation in pollen production among trees, and variation in the composition and the behavior of the animal pollinators might play an important role. It appears that only large changes in population density, such as a tenfold increase or decrease, could have significant mating system consequences. This also suggests that for tree species that have relatively dense populations and are pollinated by small insects, minor changes in population density (*e.g.*, as when a small proportion of trees are exploited for timber) may not result in significant population genetic impacts.

To conclude, the two Amazonian tree species investigated here produce genetically diverse seed crops in their natural settings, have high levels of pollen flow, and minimal levels of inbreeding. However, variation in local density of neighbors within pollination neighborhood did not have a significant impact on inbreeding levels in the populations investigated here. The mating system and pollen flow parameters estimated in this study are strikingly similar for these two timber species with high population density and pollinated by small insects, but having different sexual systems that do not completely prevent selfing. These results, while adding support to the emerging paradigm that tropical trees have evolved in a context of regular outcrossing with distant individuals, also suggest that tropical tree populations pollinated by small

insects can bypass near-neighbor mating, not only in low density populations (Stacy *et al.* 1996), but also for medium to high density populations such as those examined here.

Acknowledgements

We thank IBAMA and CIRAD-Forêt for access to the experimental sites. We thank the subject editor and the two reviewers for their suggestions to improve the manuscript, as well as I. S. Thompson, J. Campolina, inhabitants of comunidade Branco-1 km 83 and Belterra, V. C. R. Azevedo, T. N. Almeida, V. P. da Silva, M. S. Borges, F. Carneiro, E. Bandou, S.-O. Cazal, V. Troispoux for their help during field and lab activities. This study is part of Dendrogene Project: Genetic Conservation in Managed Forests in Amazônia (2000-2004, www.cpatu.embrapa.br/dendro/index.htm) carried out by Embrapa Amazônia Oriental and its partners, within bilateral cooperation between Brazil and United Kingdom, through Brazilian Cooperation Agency (ABC) and Department for International Development (DFID). This work was carried out with the aid of a grant from the International Development Research Center, Ottawa, Canada (www.idrc.ca). D. Schoen acknowledges support from an NSERC Discovery Grant. The Fonds Nature et Technologies, Québec (FQRNT), Natural Sciences and Engineering Council, Canada (NSERC) and the Organization of American States (OAS) provided financial support to D. Cloutier.

References

Austerlitz, F., C. W. Dick, C. Dutech, E. K. Klein, S. Oddou-Muratorio, P. E. Smouse, and V. L. Sork. 2004. Using genetic markers to estimate the pollen dispersal curve. Molecular Ecology 13: 937-954.

Austerlitz, F., and P. E. Smouse 2002. Two-generation analysis of pollen flow across a landscape. IV. Estimating the dispersal parameter. Genetics 161: 355-363.

Austerlitz, F., and P. E. Smouse 2001. Two-generation analysis of pollen flow across a landscape. II. Relation between Phi(ft), pollen dispersal and interfemale distance. Genetics 157: 851-857.

Baker, H. G. 1959. Reproductive methods as factors in speciation in flowering plants. Cold Spring Harbor Symposia on Quantitative Biology 24: 177-191.

Bawa, K. S. 1990. Plant-pollinator interactions in tropical rain-forests. Annual Review of Ecology and Systematics 21: 399-422.

Bawa, K. S. 1974. Breeding systems of tree species of a lowland tropical community. Evolution 28: 85-92.

Bawa, K. S., S. H. Bullock, D. R. Perry, R. E. Coville, and M. H. Grayum. 1985. Reproductive biology of tropical lowland rainforest trees. II. Pollination systems. American Journal of Botany 72: 346-356.

Boshier, D. H., M. R. Chase, and K. S. Bawa. 1995. Population genetics of *Cordia alliodora* (Boraginaceae), a neotropical tree. 3. Gene flow, neighborhood, and population substructure. American Journal of Botany 82: 484-490.

Brown, A. H. D., and R. W. Allard. 1970. Estimation of the mating system in open-pollinated maize populations using isozyme polymorphisms. Genetics 66: 133-145.

Chase, M. R., C. Moller, R. Kesseli, and K. S. Bawa. 1996. Distant gene flow in tropical trees. Nature 383: 398-399.

Cloutier, D., J. S. R. Povoa, L. C. Procopio, N. V. M. Leao, L. H. O. Wadt, A. Y. Ciampi, and D. J. Schoen. 2005. Chloroplast DNA variation of *Carapa guianensis* in the Amazon basin. Silvae Genetica 54: 270-274.

Dayanandan, S., J. Dole, K. Bawa, and R. Kesseli. 1999. Population structure delineated with microsatellite markers in fragmented populations of a tropical tree, *Carapa guianensis* (Meliaceae). Molecular Ecology 8: 1585-1592.

Dawson, I. K., R. Waugh, A. J. Simons, and W. Powell. 1997. Simple sequence repeats provide a direct estimate of pollen-mediated gene dispersal in the tropical tree *Gliricidia sepium*. Molecular Ecology 6: 179-183.

Degen, B., E. Bandou, and H. Caron. 2004. Limited pollen dispersal and biparental inbreeding in *Symphonia globulifera* in French Guiana. Heredity 93: 585-591.

Dick, C. W. 2001. Genetic rescue of remnant tropical trees by an alien pollinator. Proc. Roy. Soc. Lond. B. 268: 2391-2396.

Dick, C. W., G. Etchelecu, and F. Austerlitz. 2003. Pollen dispersal of tropical trees (*Dinizia excelsa*: Fabaceae) by native insects and African honeybees in pristine and fragmented Amazonian rainforest. Molecular Ecology 12: 753-764.

Doligez, A., and H. I. Joly. 1997. Mating system of *Carapa procera* (Meliaceae) in the French Guiana tropical forest. American Journal of Botany 84: 461-470.

Federov, A. A. 1966. Structure of tropical rain forest and speciation in humid tropics. Journal of Ecology 54: 1-11.

Frankham, R. 1995. Effective population size / adult population size ratios in wildlife – a review. Genetical Research 66: 95-107.

Guariguata, M. R., H. Arias-LeClaire, and G. Jones. 2002. Tree seed fate in a logged and fragmented forest landscape, northeastern Costa Rica. Biotropica 34: 405-415.

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

Hardy, O. J., L. Maggia, E. Bandou, P. Breyne, H. Caron, M-H. Chevallier, A. Doligez, C. Dutech, A. Kremer, C. Latouche-Halle, V. Troispoux, V. Veron, and B. Degen. 2006. Fine-scale genetic structure and gene dispersal inferences in 10 neotropical tree species. Molecular Ecology 15: 559-571.

Hardy, O. J., S. C. Gonzalez-Martinez, B. Colas, H. Freville, A. Mignot, and I. Olivieri. 2004. Fine-scale genetic structure and gene dispersal in *Centaurea corymbosa* (Asteraceae). II. Correlated paternity within and among sibships. Genetics 168: 1601-1614.

Hardy, O. J., and X. Vekemans. 2002. SPAGEDI: a versatile computer program to analyse spatial genetic structure at the individual or population levels. Molecular Ecology Notes 2: 618-620.

Kanashiro, M., I. S. Thompson, J. A. G. Yared, M. D. Loveless, P. Coventry, R. C. V. Martins-Da-Silva, B. Degen, and W. Amaral. 2002. Improving conservation values

of managed forests: the Dendrogene Project in the Brazilian Amazon. Unasylva 53: 25-33.

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

Ledig, F. T. 1992. Human impacts on genetic diversity in forest ecosystems. Oikos 63: 87-108.

Loveless, M. D. 2002. Genetic diversity and differentiation in tropical trees. In B. Degen, M. D. Loveless, and A. Kremer (Eds.). Modelling and experimental research on genetic processes in tropical and temperate forests (2000 : Kourou, French Guiana), pp. 3-30. Embrapa Amazonia Oriental, Belem, Brazil.

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

Murawski, D. A., and J. L. Hamrick. 1992. The mating system of *Cavanillesia platanifolia* under extremes of flowering-tree density – A test of predictions. Biotropica 24: 99-101.

Murawski, D. A., and J. L. Hamrick. 1991. The effect of the density of flowering individuals on the mating systems of 9 tropical tree species. Heredity 67: 167-174.

Nason, J. D., and J. L. Hamrick. 1997. Reproductive and genetic consequences of forest fragmentation: Two case studies of neotropical canopy trees. Journal of Heredity 88: 264-276.

Raymond, M., and F. Rousset. 1995. GENEPOP (version 1.2) Population genetics software for exact tests and ecumenicism. Journal of Heredity 86: 248-249.

Ritland, K. 2002. Extensions of models for the estimation of mating systems using n independent loci. Heredity 88: 221-228.

Scarano, F. R., T. S. Pereira, and G Rocas. 2003. Seed germination during floatation and seedling growth of *Carapa guianensis*, a tree from flood-prone forests of the Amazon. Plant Ecology 168: 291-296.

Schmitt, L., and M. Bariteau. 1988. Gestion de l'ecosysteme forestier guyanais: Etude de la croissance et de la regeneration naturelle: Dispositif de Paracou. Bois et Forets des Tropiques 220: 3-24.

Smouse, P. E., R. J. Dyer, R. D. Westfall, and V. L. Sork. 2001. Two-generation analysis of pollen flow across a landscape. I. Male gamete heterogeneity among females. Evolution 55: 260-271.

Stacy, E. A., J. L. Hamrick, J. D. Nason, S. P. Hubbell, R. B. Foster, and R. Condit. 1996. Pollen dispersal in low-density populations of three neotropical tree species. American Naturalist 148: 275-298.

Vekemans, X., and O. J. Hardy. 2004. New insights from fine-scale spatial genetic structure analyses in plant populations. Molecular Ecology 13: 921-935.

Veron, V., H. Caron, and B. Degen. 2005. Gene flow and mating system of the tropical tree *Sextonia rubra*. Silvae Genetica 54: 275-280.

Vinson, C. C., V. C. R. Azevedo, I. Sampaio, A. Y. Ciampi. 2005. Development of microsatellite markers for *Carapa guianensis* (Aublet), a tree species from the Amazon forest. Molecular Ecology Notes. 5: 33-34.

Ward, M., C. W. Dick, R. Gribel, and A. J. Lowe. 2005. To self, or not to self... A review of outcrossing and pollen-mediated gene flow in neotropical trees. Heredity 95: 246-254.

White, G. M., D. H. Boshier, and W. Powell. 2002. Increased pollen flow counteracts fragmentation in a tropical dry forest: An example from *Swietenia humilis* Zuccarini. Proc. Natl. Acad. Sci. 99: 2038-2042.

Tables

Table 1

Allele frequencies, number of seed genotyped (N), gene diversity (H_e), observed frequency of heterozygotes (H_o), and fixation index (F_{IS}) in microsatellite loci of *Carapa guianensis* and *Sextonia rubra*.

	Carapa guianensis microsatellite loci				*Sextonia rubra* microsatellite loci			
Allele	Cg1	Cg6	Cg16	Cg17	Sr3	Sr9	Sr10	Sr51
1	0.729	0.285	0.015	0.031	0.300	0.193	0.227	0.003
2	0.215	0.054	0.011	0.034	0.008	0.050	0.002	0.065
3	0.055	0.006	0.034	0.275	0.029	0.024	0.024	0.013
4		0.001	0.105	0.056	0.037	0.147	0.268	0.182
5		0.039	0.414	0.201	0.086	0.009	0.085	0.240
6		0.010	0.233	0.029	0.026	0.110	0.080	0.138
7		0.271	0.023	0.141	0.081	0.088	0.002	0.214
8		0.188	0.027	0.099	0.117	0.211	0.251	0.005
9		0.095	0.011	0.133	0.144	0.066	0.051	0.016
10		0.017	0.001		0.003	0.086	0.002	0.055
11		0.013	0.020		0.016	0.018	0.007	0.070
12		0.020	0.012		0.003			
13		0.003	0.071		0.140			
14			0.013		0.008			
15			0.007		0.003			
16			0.004					
N	320	317	331	341	176	212	191	176
H_e	0.42	0.80	0.76	0.83	0.84	0.86	0.80	0.83
H_o	0.37	0.83	0.73	0.80	0.81	0.85	0.80	0.79
F_{IS}	0.12	-0.04	0.03	0.04	0.03	0.02	-0.01	0.05

Table 2

Mean pollen dispersal distance (δ) under the normal and exponential dispersal models using maximum and minimum estimates of the effective density of reproductive trees (d_{max} and d_{min}) (trees/ha) based on global estimates of pollen pool differentiation in *Carapa guianensis* and *Sextonia rubra.*

Dispersal model	*Carapa guianensis* δ (m) (95% CI)[a]		*Sextonia rubra* δ (m) (95% CI)	
	$d_{max} = 2.50$	$d_{min} = 0.25$	$d_{max} = 1.30$	$d_{min} = 0.13$
Normal	69 (58-96)	218 (184-288)	86 (74-167)	271 (233-448)
Exponential	78 (66-109)	244 (207-317)	97 (83-188)	303 (262-491)

[a]: Bootstrap 95% confidence interval in parentheses

Table 3

Mean pollen dispersal distance (δ) based on pairwise pollen pool differentiation estimates for *Carapa guianensis* and *Sextonia rubra*. The density of reproductive trees (*d*) was constrained or estimated jointly with δ for different dispersal models. The error indicates the level of departure of the observed pairwise estimates to the dispersal model fitted.

Carapa guianensis

Dispersal model	Error	*d* (trees/ha)		δ (m)
		Constrained	Estimated	Estimated
Normal	2.129	2.50		71
Normal	2.063		0.23	218
Exponential	2.127	2.50		80
Exponential	2.062		0.18	276
Exponential power[a]	2.127	2.50		91
Exponential power[b]	2.062		0.13	355
Exponential power[c]	2.062		0.14	338

Sextonia rubra

Dispersal model[d]	Error	*d* (trees/ha)		δ (m)
		Constrained	Estimated	Estimated
Normal	0.263	1.30		81
Normal	0.228		0.14	231
Exponential	0.264	1.30		92
Exponential	0.229		0.11	287

[a]:Estimated shape parameter = 0.73

[b]:Shape parameter constrained to 0.73

[c]:Estimated shape parameter = 0.77

[d]:The exponential power model was not applicable (see results for details)

Table 4

Mating system estimates for *Carapa guianensis* and *Sextonia rubra*, where t_m represents the multi-locus outcrossing rate, t_s the single-locus outcrossing rate, and r_p the multi-locus correlation of outcrossed paternity.

Mating system parameters	*Carapa guianensis*	*Sextonia rubra*
t_m (SE)	0.918[a] (0.033)	0.945[a] (0.022)
t_s (SE)	0.902[a] (0.038)	0.941 (0.042)
$t_m - t_s$ (SE)	0.015 (0.015)	0.004 (0.026)
r_p (SE)	0.089[b] (0.024)	0.096[b] (0.035)

[a]:Less than 1.000 at 5 percent level

[b]:Greater than 0.000 at 5 percent level

Table 5

Impact of variation in local density of conspecific neighbors on the mating system of *Carapa guianensis* and *Sextonia rubra*, where t_m represents the multi-locus outcrossing rate, t_s the single-locus outcrossing rate and, r_p the multi-locus correlation of outcrossed paternity.

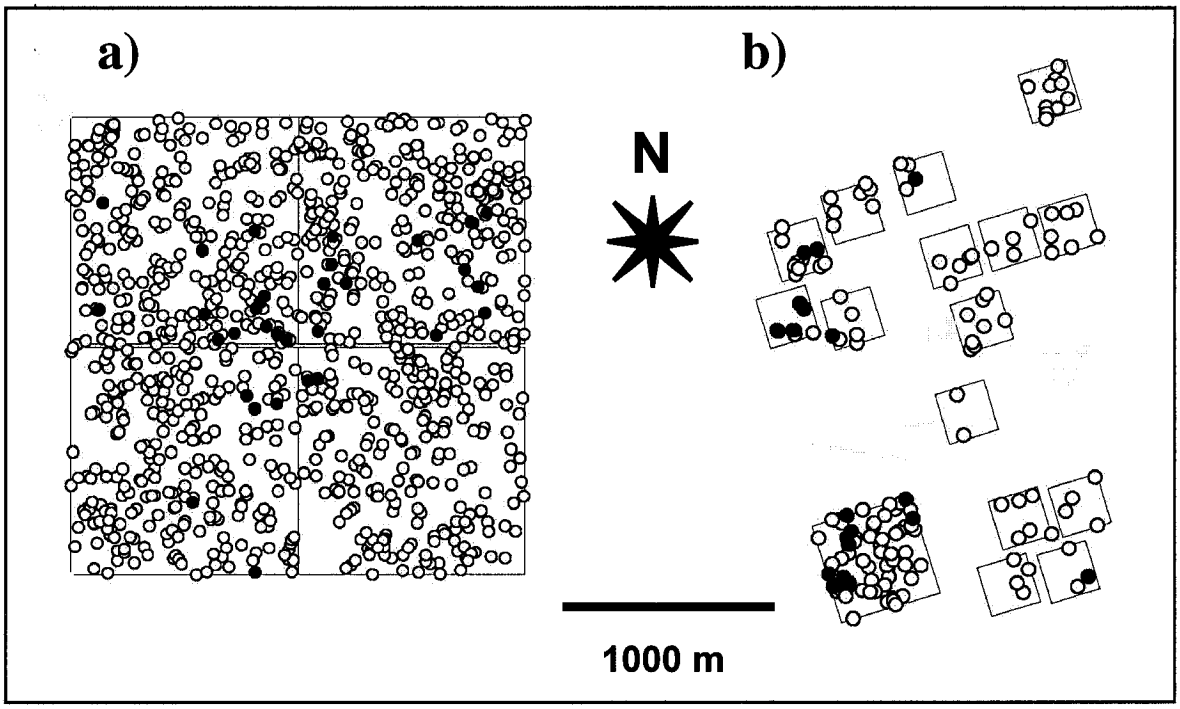
	Carapa guianensis		*Sextonia rubra*	
Local density[a]	Low density	High density	Low density	High density
t_m (SE)	0.92 (0.06)	0.93 (0.04)	0.95 (0.09)	0.90 (0.02)
t_s (SE)	0.93 (0.07)	0.90 (0.05)	0.88 (0.09)	0.92 (0.03)
t_m - t_s (SE)	-0.02 (0.02)	0.02 (0.03)	0.07 (0.08)	-0.02 (0.03)
r_p (SE)	0.11 (0.08)	0.05 (0.02)	0.19 (0.18)	0.03 (0.05)

[a]:Mother trees are classified into two groups of low and high mean number of neighbors/ha (see text for details).

Figures

Figure 1

Distribution of *Carapa guianensis* at Tapajos (a) and *Sextonia rubra* at Paracou (b) with sizes above 30 cm DBH. For both sites the plots are embedded within continuous forest for which the distribution of *C. guianensis* and *S. rubra* trees is not shown here. Filled circles are the trees from which seed was sampled.



Linking statement for Chapter 3

Having documented high levels of gene flow and low levels of inbreeding within two Amazonian forest tree populations in Chapter 2, I come to the central question of this thesis, that is, whether selective logging impacts the genetic diversity of tropical tree populations. In the Brazilian Amazon, *Carapa guianensis* is a tree species used by the timber industry, and relative to other tree species harvested in the area, is considered as locally abundant and fast-growing climax species for the purposes of forest management. Moreover, the reproductive biology of *Carapa guianensis* is characterized, among other things, by a pollination system based on small insects, the presence of monoecious flowers, and a relative lack of phenological synchrony at the population level. In Chapter 2, I showed that those features seem to translate into high gene dispersal and low inbreeding for *Carapa guianensis*. In Chapter 3, seed progeny samples from a *Carapa guianensis* population collected before and after selective logging are used to assess whether levels of genetic diversity, inbreeding and gene flow are impacted by selective harvesting. This is the first such study to investigate the population genetic characteristics of the same tropical forest before and after logging (real time series). Other studies on the impact of logging rely on artificial "time series", *i.e.* the comparison of different populations.

Chapter 3:

Impact of selective logging on inbreeding and gene dispersal in an Amazonian population of *Carapa guianensis* Aubl.

Cloutier D, Kanashiro M, Ciampi AY, Schoen DJ. Impact of selective logging on inbreeding and gene dispersal in an Amazonian population of *Carapa guianensis* Aubl. ***Accepted in Molecular Ecology***

Abstract

Selective logging may impact patterns of genetic diversity within populations of harvested forest tree species by increasing distances separating conspecific trees, and by modifying physical and biotic features of the forest habitat. We measured levels of gene diversity, inbreeding, pollen dispersal and spatial genetic structure (SGS) of an Amazonian insect-pollinated *Carapa guianensis* population before and after commercial selective logging. Similar levels of gene diversity and allelic richness were found before and after logging in both the adult and the seed generations. Pre- and post-harvest outcrossing rates were high, and not significantly different from one another. We found no significant levels of biparental inbreeding either before or after logging. Low levels of pollen pool differentiation were found, and the pre- versus post-harvest difference was not significant. Pollen dispersal distance estimates averaged between 75 and 265 m before logging, and between 76 and 268 m after logging, depending on the value of tree density and the dispersal model used. There were weak and similar levels of differentiation of allele frequencies in the adults and in the pollen pool, before and after logging occurred, as well as weak and similar pre- and post-harvest levels of SGS among adult trees. The large neighbourhood sizes estimated suggest high historical levels of gene flow. Overall our results indicate that there is no clear short-term genetic impact of selective logging on this population of *Carapa guianensis*.

Introduction

The exploitation of a proportion of forest trees chosen for the economic value of their wood, *i.e.* selective logging, has a direct impact on the demography of harvested tree populations, as it: (1) removes large individuals that likely contribute more to reproduction than smaller individuals, and (2) reduces the overall population density of reproductive individuals. Additionally, tree harvesting may result in "collateral damage"

to the forest habitat (*e.g.*, a shift in the abundance, diversity and behaviour of animal pollinators) that in turn can have an impact on the reproductive biology of remaining tree populations. From a population genetics perspective, such impacts may induce changes in patterns of genetic diversity, inbreeding, gene flow, and the effective size of populations. Two expected genetic consequences of small population size include increases in the level of genetic drift and inbreeding (Ellstrand & Elam 1993). The potential losses of genetic variation associated with logging may decrease the ability of populations to adapt to directional abiotic and biotic environmental changes as well as reduce their flexibility to respond to short-term challenges such as pathogens and herbivores (Lowe *et al.* 2005). Thus reductions in population size and increase in inbreeding resulting from selective logging may pose significant threats to the long-term viability of harvested tree populations.

The directly measurable population genetic effects of selective logging could fall into several categories. First, selective harvesting could lead to an immediate loss of genetic diversity as a consequence of removal of adult trees; or in the seed progeny generation, as a consequence of reduced availability of pollen donors. Such changes, if they occur, should be detectable with marker genes. A study of *Carapa guianensis* in Central America found that marker allelic richness in saplings was lower than in the adult population in logged forests, but not in unlogged forests (Dayanandan *et al.* 1999; but see Hall *et al.* 1994). Second, coupled with the reduction in density following tree harvest, there may be a reduction in the number of individuals available to donate pollen. This may restrict pollen movement and increase inbreeding, two factors that could lead to loss of genetic diversity in the long-term. In Lowe *et al.*'s (2005) review, six out of eight experimental studies that assessed progeny arrays for levels of inbreeding highlighted significant anthropogenic impacts, including a study of *Carapa procera* in French Guiana where lower outcrossing rates were found in logged plots compared to undisturbed plots (Doligez & Joly 1997). Alternatively, one may hypothesize that pollinators are able to cope with the increased inter-tree distance following selective logging, and that gene dispersal could be maintained (or even increased). This alternate hypothesis is supported by studies where high levels of pollen dispersal were discovered in an undisturbed population of *Carapa guianensis* (Cloutier *et al.* in press) while increased pollen flow has

been reported following habitat fragmentation in *Swietenia humilis* (White *et al.* 2002). Third, population thinning due to logging could modify the spatial genetic structure of reproductive adults, potentially leading to an increased proportion of matings among unrelated individuals in the population. Support for this hypothesis comes from studies where self-thinning in tree populations resulted in a reduction of spatial genetic structure from younger to older cohorts (Hamrick *et al.* 1993).

The goals of the present study were to assess the level of genetic diversity, inbreeding, gene flow and spatial genetic structure in a natural population of the "small-insect"-pollinated tree species *Carapa guianensis* managed for timber harvest. For the purposes of forest management in the Brazilian Amazon, *C. guianensis* is classified as a high-density and fast-growing climax species. Seed progenies were collected over a 22 month period, before and after selective logging occurred, allowing us to test the following broad genetic hypotheses: (1) selective logging may change levels of genetic diversity; (2) selective logging may change gene flow through pollen and change levels of inbreeding; and (3) selective logging may change levels of spatial genetic structure among adult trees. The results are discussed with respect to demographical and ecological aspects of *C. guianensis*, the historical context of the study population, and the short-term consequences of logging for the genetic resource base of *C. guianensis*.

Materials and methods

Study species

Carapa guianensis Aubl. (Meliaceae) is a widespread Neotropical tree found in the Caribbean islands, Central America and northern South America. The species inhabits upland *terra firme* forests as well as floodplain *varzea* forests in the Amazon basin. In Brazil, it is an important timber tree species, and the oil extracted from the seed is extensively traded and used for its medicinal properties. *C. guianensis* is a monoecious

species, and the small white flowers (3-5 mm diameter) are visited by small insects (*e.g.*, stingless bees, beetles and moths) (D. Cloutier, pers. obs.). Individual trees below a size of ~30 cm diameter at breast height (DBH) typically do not produce fruits in closed undisturbed forests, while trees above the size of 30 cm DBH may produce no fruits or produce fruits once or several times per year. Both flowers and fruits may be observed simultaneously on the same tree (D. Cloutier, pers. obs.). At the population level, flowers and fruits are produced year round, with several intra-year peaks of flowering intensity and fruit release (Dendrogene project, unpublished data). During the study period, most fruits were released around the onset (January-February) and the end (May-June) of the wet season, allowing us to sample fruits at each of those times. The fruit of *C. guianensis* contains 4-20 seeds, weighing 20-40 g each, which may be dispersed by water in flood-prone forests. However, in the upland *terra firme* populations, the seeds are dispersed by gravity, and secondarily by medium-sized scatter-hoarding rodents over distances less than 25 meters (Guariguata *et al.* 2002, Cloutier *et al.* 2005). As a result, seeds and seedlings are typically found directly beneath the maternal parent tree at upland sites.

Study site

The study site lies at the Tapajós National Forest (Flona Tapajós), Pará state, Brazil, at 2°51'S 54°57'W and 175 m above sea level. The area is a flat plateau covered by a dense *terra firme* forest. The climate is tropical with a mean annual rainfall of 2000 mm (peaks from February to May) and an annual dry season of 2-3 months (August to October). Before the creation of the Flona Tapajós in 1974, the only known recent disturbances were sporadic low-impact activities such as hunting and selective logging of a few tree species, but not of *Carapa guianensis* (Jennings 1997). Under the auspices of the Dendrogene project (Kanashiro *et al.* 2002), an intensive study plot (500 ha) is being monitored at Tapajós for the effects of selective logging (Reduced Impact Logging – RIL) on the ecology and genetics of this and other species. The study species were chosen to have different characteristics such as growth response to light, reproductive ecology, and population demography. This was done in order to place species into groups

for management purposes, and to predict (through genetic modelling) impacts on future generations with respect to genetic diversity. Among the species included in the study, *C. guianensis* was classified as a fast-growing climax species, locally abundant relatively to other species, and with reproduction through monoecy and non-synchronous flowering. Prior to this study, a commercial pre-logging inventory was conducted in the study plot, and every tree larger than 20 cm DBH was identified and mapped.

A mean density of 2.5 *C. guianensis* trees/ha ≥ 30 cm DBH was found in the 400 ha of the central plot area (Figure 1), indicating that *C. guianensis* has the second highest average tree density at this site. At the end of 2003, the study plot was selectively logged and at least 40 tree species were harvested, *C. guianensis* being one of them. The size below which *C. guianensis* trees were not harvested (cutting size limit) was about 53 cm DBH. A total of 257 *C. guianensis* trees were harvested and sent to the sawmill (from the 474 between 53 and 111 cm DBH present prior to logging), *i.e.* more than 50% of *C. guianensis* trees of harvestable size were removed. The distribution of the harvested and non-harvested *C. guianensis* trees at the study site is shown in Figure 1. Trees left standing often had a hollow or poorly formed trunk. It is worth noting that a substantial number of additional trees were likely damaged or killed directly during the logging operations, or indirectly (trees that fell following increased exposure to wind after the logging), but no estimate of this additional damage is available, and therefore the harvested trees in Figure 1 should be considered as a conservative estimate of the loss of *C. guianensis* conspecifics.

Field sampling

An important feature of this study is that seed progeny were collected from the same *Carapa guianensis* trees monitored during 22 months from 2003 to 2005, within which time interval controlled commercial selective logging was conducted at the end of 2003. The samples taken in the field during that time interval can be divided into two types: (1) seed sampled from 21 reproductive trees monitored before and after logging,

and (2) leaf and cambium sampled before logging from three cohorts of the existing population (*e.g.*, seedling, juvenile and adult).

Seeds that were not previously handled by animals were sampled directly beneath selected seed parents both before and after logging. The seed parents can be described as biological "pollen traps", and are referred to hereafter as "trap trees". We began by regularly visiting a number of potential trap trees, separated by a range of different inter-tree distances, ranging from trap trees that are close enough to receive pollen from the same source trees, to those far apart and likely to receive pollen from different source trees. Trees that produced at least 10 fruits in both the pre-logging and the post-logging period were selected as trap trees (Figure 1). The distance among the 21 trap trees varied between 16 m and 1883 m, averaging at 832 m. We sampled seeds from the same 21 unlogged trap trees (before and after logging) aiming to minimize the effects of local spatial variation so as to better examine temporal variation in the population. A separate genetic analysis conducted on *C. guianensis* multi-seeded fruits collected in 2004 at the study site revealed that seeds within individual fruits are significantly more likely to share the same father than are seeds taken from different fruits of the same mother tree (D. Cloutier, unpublished data), suggesting pollen co-dispersal from the same male parent(s) to single flowers. Thus, to sample as many independent pollination events as possible (a requirement for obtaining good estimates of pollen dispersal; Hardy *et al.* 2004), we analyzed a single seed per fruit sampled. An average of 37 fruits were sampled from each of the 21 trap trees, varying between 12 and 37 (total of 391) for the pre-harvest period, and between 10 and 33 (total of 390) for the post-harvest period. The seed were sent to the lab in closed plastic bags were the small living embryos were extracted for DNA analyses.

Leaf and cambium tissue were sampled in the population to assess levels of genetic diversity and spatial genetic structure in three different cohorts (seedling, juvenile and adult), defined as follows. The adults (n=199) were reproductive individuals, defined here as trees of size above 30 cm DBH. The juveniles (n=82) were trees between 10 and 30 cm DBH, mostly non-reproductive. The seedlings (n=84) were plants less than 2 years old and less than 1 m in height, and were entirely non-reproductive. The distribution of sampled adults, juveniles and seedlings is shown in Figure 1. When leaves were

accessible, a piece of leaf was taken in the field and dried in silica. Otherwise, a piece of cambium (~150mg) was extracted and stored in a buffer of 70% ethanol and 0.3% β-mercaptoethanol before being sent to the lab. The leaf and cambium samples analyzed were randomly selected in the lab from the set of field samples available, but since some areas were more intensively sampled than others in the field, the distribution of analyzed samples is not uniform (Figure 1) (*e.g.,* seedlings were sampled along three transects).

Laboratory analyses

Samples were brought into the plant genetic laboratory of EMBRAPA-CENARGEN (Brazilian Agricultural Research Corporation - Centre for Genetic Resources and Biotechnology) in Brasilia, Federal District, Brazil. The six microsatellite (SSR) loci used in this study are described by Dayanandan *et al* (1999) for loci cg05 and cg07, and by Vinson *et al.* (2005) for loci cg01, cg06, cg16 and cg17. New reverse primers were designed for cg05 (5'GAGGATCTTGTACGTTGGC3') and cg07 (5'CTGTTCGTTGAAGAACTTGG3') in order to obtain shorter PCR products. DNA extraction, PCR conditions and electrophoresis were carried out as described in Cloutier *et al.* (in press). To rule out incorrect assignment of progeny to the mother trees, we compared trap tree and seed genotypes to detect and remove mismatched individuals from the data set (thus 0.9% of seed genotypes were removed).

Genetic diversity analyses

Standard genetic diversity parameters were estimated for the seeds collected before and after logging; for the seedling, juvenile and adult cohort; and for the harvested and non-harvested adult trees larger than the size harvest limit. The number of alleles per locus (N_{all}), the allelic richness (R) (*i.e.,* the number of alleles independent of sample size, allowing comparisons between samples of different sample sizes; El Mousadik & Petit 1996), the unbiased gene diversity (H_e), and the observed heterozygosity (H_o) were

calculated with the program FSTAT version 2.9.3 (Goudet 1995). Exact tests for fixation index estimates (F_{is}) (Weir & Cockerham 1984) and for linkage disequilibrium among loci were conducted with the program GENEPOP (Raymond & Rousset 1995).

Mating system analyses

Maximum likelihood estimates of single-locus (t_s) and multi-locus (t_m) outcrossing rates were calculated following the mixed-mating model (Ritland 2002) using the seed together with their known maternal genotypes. If mating among relatives (biparental inbreeding) occurs, t_m should be higher than t_s, the difference giving a minimum estimate of apparent selfing due to biparental inbreeding. Multi-locus correlation of paternity (r_p) was calculated to estimate the probability that two outcrossed progenies drawn at random from the same seed family are full-sibs. Mating system analyses were conducted using the program MTLR version 3.1 (Ritland 2002). The default settings were used, except that the parental fixation index was constrained to 0.029 (*i.e.*, the value observed in the adult cohort). We assumed no apomixes and no seed mortality due to inbreeding prior to the assay. Standard errors for t_s, t_m and r_p were estimated from 1000 bootstraps at the level of the seed families.

Two different mating system analyses were performed. First, we contrasted the seed progenies sampled before logging (n=391) versus those sampled after logging (n=390), to assess the impact of logging on mating system. Second, we divided the seed progenies into four temporal groups in order to assess whether there was seasonal variation in mating system. The four temporal groups were composed of seed collected in April-June 2003 (n=243) and January-February 2004 (n=148), thus resulting from pre-logging pollination; and seed collected in May-July 2004 (n=158) and January-February 2005 (n=232), thus resulting from post-logging pollination.

Pollen dispersal distance analyses and effective number of pollen donors

Genotyping of every potential pollen donor, as in a paternity analysis, was not practical in this study, since trees were continuously distributed over a very large area. Instead, we analyzed pollen dispersal with the TWOGENER algorithm developed by Smouse *et al.* (2001), using an implementation of the algorithm programmed by F. Austerlitz (Université Paris-Sud, France). The statistic ϕ_{FT}, which measures the differentiation in allelic frequencies among the pollen clouds sampled by different seed parents (*i.e.*, trap trees, whose spatial coordinates are known), was computed to estimate the pollen dispersal parameters. A formal relationship between ϕ_{FT} and pollen dispersal parameters has been derived for different pollen flow models (Austerlitz & Smouse 2001). To estimate pollen dispersal parameters either a global ϕ_{FT}, based on average pollen pool differentiation and average physical distance over all sampled mothers, or alternatively, a pairwise ϕ_{FT}, based on pollen pool differentiation and physical distance between each pair of sampled mother can be used—both are employed below, and referred to respectively as "global" and "pairwise" estimators.

We estimated the effective number of pollen donors contributing to the progeny of the average mother tree (N_{ep}), which is derived directly from the global estimate of ϕ_{FT} as $N_{ep} = 1/2\phi_{FT}$ (Smouse *et al.* 2001), as well as the parameter δ (*i.e.*, the average distance of realized pollen dispersal) (Austerlitz & Smouse 2001, Austerlitz & Smouse 2002). The global ϕ_{FT} (Austerlitz & Smouse 2001), and the pairwise ϕ_{FT} (Austerlitz & Smouse 2002) were used to estimate δ by assuming an underlying density of reproducing trees (d) and a pollen dispersal model (*i.e.*, normal, exponential or power exponential). Because N_e/N ratios in plant populations may vary from values close to one, to as low as 1/10 (*e.g.*, due to varying fertility and non-synchronous flowering) (Frankham 1995), we estimated δ from global ϕ_{FT} assuming a upper bound estimate of the density of reproducing trees (d_{max}) (*i.e.*, the global density of conspecific trees above 30 cm DBH in the field), and a lower bound estimate of d ($d_{min} = d_{max}/10$). The pairwise ϕ_{FT} was used also to jointly infer δ and d. To assess the 95 percent confidence intervals of the pollen dispersal parameters,

we created bootstrap data sets by resampling the 21 trap tree families. Because of the prohibitive amount of computer processing time required for the calculations, we used 200 bootstrap samples (this required 52 Apple Macintosh G4 CPUs over two months). While the "power exponential" model has been suggested to approximate the shape of the dispersal curve (Austerlitz *et al.* 2004), this model repeatedly failed to converge to realistic values using the original and bootstrapped data sets and, therefore, only results based on the "normal" and "exponential" dispersal model are reported here.

Detection of bottlenecks in pollen pool allele frequencies

To assess the magnitude of logging-mediated population bottlenecks, we estimated the extent of differentiation (or divergence) between the estimates of pollen allele frequencies and of adult tree allele frequencies, using the estimator F_k (Pollack 1983, Waples 1989). This statistic is related to the 'variance effective population size' of the pollen donor pool. We estimated, separately, pollen pool allele frequencies before and after logging using MLTR version 3.1 (Ritland 2002), while the actual allele frequencies of the 199 adults sampled were used for the adult allele frequencies. A 95% confidence interval on F_k was derived from the chi-square distribution following Waples (1989).

Spatial genetic structure analyses

To take advantage of different approaches for assessing the consequences of historical gene dispersal, spatial genetic structure (SGS) was estimated using two measures of genetic similarity. First, we used a relative kinship coefficient (F_{ij}) described in Loiselle *et al.* (1995). $F_{ij}= (Q_{ij}-Q_m)/(1-Q_m)$, where Q_{ij} is the probability of identity in state for random genes from individuals i and j, and Q_m is the probability of identity in state for genes from random individuals in the population. Second, we used the relationship coefficient R_{ij}, computed as the correlation between individual allele frequencies (*i.e.,* Moran's I autocorrelation statistic), as described in Hardy & Vekemans

(1999). We used the program SPAGEDI version 1.2 (Hardy & Vekemans 2002) for all SGS analyses.

To visualize SGS, F_{ij} values were plotted between individuals according to the physical distance separating them, but since using all pairwise F_{ij} values on a graph would obscure the overall trend of the SGS, we presented average pairwise F_{ij} for pairs of individuals falling into a set of specific distance classes. The upper limit of each distance class was 3.6, 7.6, 14.4, 24, 38, 67, 109, 191, 362, 612, 1014, 1536 m for the seedlings; 30, 50, 70, 91, 118, 156, 211, 303, 450, 722, 1213, 2745 m for the juveniles; and 30, 46, 61, 80, 102, 135, 175, 222, 273, 332, 397, 464, 542, 635, 740, 864, 1020, 1220, 1579, 2945 m for the adults. The distance classes were chosen to facilitate the visualization of fine-scale SGS on a logarithmic graph, while ensuring a minimum of 30 pairs in each distance class. To test for significant deviations from random SGS, observed values for each distance class were compared to the 95% confidence interval derived from 1000 permutations of individuals among locations.

To estimate historical gene dispersal from SGS in terms that are equivalent to Wright's neighbourhood size (Nb), we used two approaches. For the first approach, we followed Rousset (2000), who showed that under isolation by distance in a two dimensional space, F_{ij} is expected to be approximately linearly related to the logarithm of distance between individuals, and therefore the slope of the regression (b_k) of F_{ij} values over ln(distance) estimates $1/4\pi D\sigma^2$, where σ^2 is the second moment of the parent-progeny dispersal distance and D is the effective density of individuals. The product $4\pi D\sigma^2$ is related to Wright's neighbourhood size, allowing us to estimate $Nb_1 = -(1-F_{ij}^{[1]})/b_k$ (applying the correction proposed by Vekemans & Hardy 2004), where $F_{ij}^{[1]}$ is the average kinship coefficient between neighbouring individuals. For the second approach, we followed Epperson (2005), who described a nearly linear relationship between the natural logarithm of Nb and the finest scale spatial autocorrelation of individual allele frequencies, as measured by Moran's I relationship coefficient (R_{ij}) between neighbours. This allows one to estimate $Nb_2 = \exp[(0.544-R_{ij}^{[1]})/0.102]$, where $R_{ij}^{[1]}$ is the average Moran's I coefficient between neighbour individuals. Both approaches require one to define what "neighbour" implies; *i.e.* what is the maximal distance between two "neighbouring" individuals. Following Epperson (2003), we

defined this distance relative to the average distance between contiguous individuals, as if they were uniformly distributed on a lattice model, including pairs of individuals separated by rook and bishop moves, and using 1.5 times the square root of the inverse of adult density, which for this *Carapa guianensis* population gave as neighbours, individuals separated by less than 95 m. $F_{ij}^{[1]}$, $R_{ij}^{[1]}$, and b_k values were tested against the null hypothesis of random distribution of genotypes by randomizing the spatial positions 1000 times. Standard errors for $F_{ij}^{[1]}$, $R_{ij}^{[1]}$ and b_k were obtained by jackknifing genetic loci. The 95% confidence interval for Nb_1 was estimated as $-(1-F_{ij}^{[1]})/(b_k \pm 2SE_{bk})$, and for Nb_2 as $\exp[(0.544-(R_{ij}^{[1]} \pm 2SE_{Rij[1]}))/0.102]$.

To estimate historical gene dispersal distance (σ) from SGS, we first estimated $\sigma=(b_k/4\pi D)^{(0.5)}$, using the census density of adults as an estimate of D. But the linear relationship of F_{ij} over ln(distance) holds best within a restricted range between 1-20 times σ (Rousset 2000), therefore we applied an iterative procedure to estimate σ using restricted b_k implemented in SPAGEDI (Hardy & Vekemans 2002).

Results

Genetic diversity and heterozygosity

Allele frequencies for both the adults and the seed sampled at the six microsatellite loci are shown in Table 1. The total number of alleles for all samples was 4, 12, 13, 9, 17 and 12, respectively, at the loci cg01, cg06, cg05, cg07, cg16 and cg17. Gene diversity (H_e) was high at ~0.8 for cg06, cg05, cg16 and cg17, while lower H_e was found for cg07 (~0.6) and cg01 (~0.4). No linkage disequilibrium was detected among the loci.

Genetic diversity statistics averaged over loci for the different samples are shown in Table 2. There were minor differences in the mean number of alleles per locus (N_{all}) among samples, apparently depending on the sample size. Using the measure of allelic

richness (R), there is a modest but non-significant trend toward higher R values in oldest cohorts (adults and juveniles), and lower R values in the youngest cohorts (seeds and seedlings). Trees above harvest size limit that were left unlogged had a slightly lower R than logged trees, and consequently, the post-harvest seeds appeared to have a lower R than pre-harvest seeds, however, in all cases the differences were not significant. Gene diversity (H_e) was similar in all samples indicating that common alleles remain at about the same frequencies regardless of logging or age effects. The observed heterozygosities (H_o) were slightly lower than H_e in all samples, and fixation indexes (F_{is}) were all low, but significantly greater than zero at the 5% level, except for the harvested trees. There were no significant differences in F_{is} values between the samples.

Mating system

Mating system estimates are shown in Table 3. Outcrossing rates (t_m) were high both before (0.94) and after logging (0.93), but were significantly less than 1.00, suggesting the presence of a small amount of selfing in the population. The level of mating between relatives (biparental inbreeding) was not significant, as measured by the lack of difference between single and multilocus outcrossing rates (t_m-t_s). The correlation of outcrossed paternity (r_p) was low, but significantly different than 0, suggesting that most seed within families had a different father (*i.e.* half-sibs). Overall, no significant differences in mating system parameters were detected between pre-harvest and post-harvest seed progenies.

Seasonal variation in mating system is summarized in Figure 2. All parameters t_m, t_m-t_s and r_p showed little variation over the 22 month period. However, slightly higher, though non-significant, levels of inbreeding were measured in seeds collected around January 2004 (r_p: 0.112 (0.139)) and May-June 2004 (t_m: 0.859 (0.122)); both samples corresponding to periods where fruit production in the population was relatively low. Overall, no significant temporal variation was detected over the time period investigated despite potential for variation in population phenology, pollen production, and pollinator composition.

Pollen dispersal

Global pollen pool differentiation (ϕ_{FT}) was estimated at 0.044 before logging, with a 95 percent confidence interval (95% CI) of 0.024-0.063, and at 0.054 after logging, with a 95% CI of 0.033-0.070. These values translate into an effective number of pollen donors (N_{ep}) of 11.3 (95%CI: 7.9-20.8) before logging, and of 9.3 (95%CI: 7.1-15.2) after logging. These differences before versus after logging are not significant.

Pollen dispersal distance analyses are summarized in Table 4. The minimum (δ_{min}) and maximum (δ_{max}) estimates of average distance of pollen dispersal using global ϕ_{FT} were 75 and 265 m before logging, when both the normal and the exponential model were considered, while the estimates were similar after logging, with a minimum of 76 m and a maximum of 268 m. Before logging, the joint estimation of *d* and δ using pairwise ϕ_{FT} gave an effective density of pollen donors (*d*) of 0.30 and 0.25 trees/ha, and an average pollen dispersal distance (δ) of 204 and 252 m, for the normal and exponential model respectively. After logging, the estimates were slightly different, as *d* increased to 0.75 and 0.55 trees/ha, and δ decreased to 123 and 159 m, for the normal and exponential model, respectively. The bootstrap 95% confidence intervals on the parameters were large, however, and no statistically significant differences were detected between pre-harvest and post-harvest periods. Minimum and maximum estimates derived from the pairwise ϕ_{FT} are not shown since they were very similar to the estimates from the global ϕ_{FT} reported in Table 4.

Detection of bottlenecks in pollen pool allele frequencies

The differentiation in allele frequencies (F_k) between the adult and the pre-harvest seeds was 0.0084, with a 95% confidence interval of (0.0056-0.0118), while the differentiation between adult and post-harvest seeds was 0.0066 with a 95% confidence

interval of (0.0044-0.0093). Therefore, no significant differences between before and after logging were detected.

Spatial genetic structure (SGS)

The SGS in the different cohorts is shown in Figure 3. The estimates of average F_{ij} decreased with the logarithm of distance in all cohorts, as expected under a model of isolation by distance. For the seedling cohort, average F_{ij} values at distances from 2 m to 60 m departed significantly from the hypothesis of absence of SGS, while few or no values exhibited significant SGS for the adult and juvenile cohorts. Moreover, the average F_{ij} values for the seedling cohort at short distances were higher than observed in the juvenile and adult cohorts. The overall pattern of SGS for pre-harvest and post-harvest adult population was nearly indistinguishable, indicating that adult tree SGS is not significantly impacted by selective logging.

Estimates of SGS and neighbourhood sizes (Nb) are presented in Table 5. The global regression slope (b_k) of the adult and the seedling cohorts is significantly different from a random spatial structure, confirming the presence of isolation by distance. Significant positive genetic similarity between neighbours separated by less than 95 m was detected in the adult and seedling cohorts, but not in the juveniles. The seedling cohort gave estimates of Nb that are significantly smaller than the adult and juvenile cohorts (Table 5). The estimates of neighbourhood size are slightly different according to the approach used, but agree in indicating large values for Nb.

The historical gene dispersal distance estimate was $\sigma = 238$ meters, assuming $D = 2.5$ trees/ha and unrestricted b_k. Using the iterative procedure that estimates σ within a restricted range, we obtained $\sigma = 273$ meters. This latter value should be considered as the lower bound estimate for historical gene dispersal distance because D is likely to be lower than the census density of 2.5 individuals/ha (*i.e.* the iterative procedure applied with $D < 0.8$ led to an estimate of gene dispersal that is infinite).

Discussion

Impact of tree harvesting on genetic diversity and heterozygosity

Following logging in this population, no evidence is seen for immediate losses of genetic variation in either the adult trees or their progeny, or for increases in fixation indexes. This may be partially explained by the fact that *Carapa guianensis* is relatively common in the forest, that its reproductive phenology is asynchronous, that the population has never been previously logged, and that the proportion of trees harvested during the logging operation was not sufficient to cause a detectable decline in genetic diversity or increase in inbreeding. Perhaps more importantly, we have measured genetic variation using a small number of neutral SSR markers. It is possible that selective logging, because it targets the largest and healthiest trees in the population also reduced the diversity of genes underlying these characteristics, and is therefore dysgenic, leading to erosion of the species' genetic resources (*e.g.*, quantitative trait genes for rapid growth, or qualitative trait genes for disease- or drought-resistance, etc.). As has been pointed out elsewhere, even when surveys of neutral genetic diversity indicate the presence of substantial polymorphism, genetic diversity for loci important to long-term survival and reproduction may still be in decline (Lowe *et al.* 2005).

From earlier studies, we might have expected to detect a decrease in F_{is} with age, indicating the presence of inbreeding depression. Doligez & Joly (1997) found significant excess homozygosity in seeds and excess heterozygosity in adults in *Carapa procera,* which they explained as due to selection in favour of heterozygotes. There is no such trend in our study, however, as the estimated multilocus F_{is} was 0.025 for the pre-logging seed and 0.049 for the post-logging seed, values that are similar to the seedling (0.079), juvenile (0.048) and adult (0.029) F_{is}. Lack of evidence for significant inbreeding depression in this population could be explained by a low level of inbreeding, potentially due to the presence of a pre-zygotic incompatibility system (M. Maues, pers. comm.).

Impact of tree harvesting on inbreeding, pollen dispersal distance and effective number of pollen donors

A reduction in density of adult trees following logging could, in theory, change levels of inbreeding and gene flow, but we found no statistically significant differences in the genetic parameters assessed before versus after the tree harvest. Estimates of the mating system show that there are low levels of inbreeding in this population, and that inbreeding levels remained unchanged after the harvest. We found no biparental inbreeding in this population (Table 1), a result that is in accord with the weak spatial genetic structure observed among adult trees (Table 5). Estimates of global pollen pool differentiation revealed that pollen dispersal is extensive, that there are a large effective number of pollen donors, and that pollen dispersal parameters did not appear to be strongly influenced by logging. For instance, the global ϕ_{FT} calculated with TWOGENER using seeds obtained before logging was 0.044 (0.010), while after it, increased (non-significantly) to 0.054 (0.010) (Table 4). Multilocus r_p estimates, when t_m is constrained to 1.0 in MLTR, were 0.067 (0.094) before logging and increased to 0.104 (0.161) after logging. Both approaches yield values that point toward a slight, though non-significant decrease in the effective number of pollen donors (N_{ep}) following logging. On the other hand, using pairwise ϕ_{FT} with TWOGENER to jointly estimate d and δ, we obtained (for the normal model) d = 0.3 trees/ha and δ = 204 m before logging, and d = 0.75 trees/ha and δ = 123 m after logging, a difference that suggests an (non-significant) increase in N_{ep} following logging. The reduced differentiation in allele frequencies between adult and progeny (F_k) observed after logging also points to an increase in N_{ep}. Overall it appears reasonable to conclude that these results do not provide evidence for strong genetic consequences of selective logging on *C. guianensis*.

There are a number of potential sources of error when estimating pollen dispersal parameters. It is possible that we could have slightly underestimated average pollen dispersal distances, due to the limited number of trap trees available, and the resulting restricted number of pollen dispersal models that could be applied to the data (Austerlitz *et al.* 2004). Another potential source of error is the value of tree density assumed. In this study, we defined reproductive individuals as trees larger than 30 cm DBH.

However, it is likely that some trees above this size were also not reproductive, and that some trees below this size did, in fact, contribute to reproduction. Moreover some reproductive trees are likely to be more important contributors than others due to interplant variation in flower production. The reproductive phenology of *Carapa guianensis* further complicates the matter as trees can be seen flowering intermittently. In this study, we tried to take into account these factors by estimating minimum and maximum levels of pollen dispersal, based on maximum and minimum estimates of effective tree density, but it is nevertheless possible that the actual pollen dispersal distance is higher than the estimated intervals.

At least two ecological factors may help account for the stability of gene dispersal and inbreeding in the face of selective logging. First, *Carapa guianensis* trees are relatively abundant at the study site and they are reproductively mature at a relatively small DBH compared to other tree species. Thus, as long as the size harvest limit (53 cm DBH in this logging experiment) is substantially higher than the size at which trees become reproductive, the impact on mating patterns and gene flow may be negligible in the short term. For instance, logging resulted in a modest decrease of tree density from 2.5 trees/ha to 2.0 trees/ha when all trees above 30 cm DBH are considered. A second factor could lie in features of the reproductive biology of *C. guianensis* that potentially buffer the effect of logging. For instance, despite likely intra-year fluctuations in the number of flowering trees and the efficiency of the animal pollinators, *C. guianensis* can successfully produce seeds with high levels of genetic diversity at different times within a year. Its reproductive system may be more sensitive to seasonal factors, and relatively less sensitive to logging. This is suggested in Figure 2, where seasonal variation in mating system, albeit low, appears to be higher than the variation detected as a result of the logging.

Spatial genetic structure and historical gene flow

The level of SGS observed in *Carapa guianensis* is weak, and the different estimates of neighbourhood sizes (Nb) (Table 5) are large, which suggests that gene

dispersal was high and that the impact of genetic drift has been historically limited in this population. Moreover, the pattern of SGS for the juvenile and the adult cohorts is similar (Figure 3), which suggests that gene flow has remained stable in the past in this population. In fact, the SGS in *C. guianensis* is amongst the weakest observed to date in plants (Vekemans & Hardy 2004), and is comparable to tropical tree species such as *Symphonia globulifera* (Degen *et al.* 2004), *Dicorynia guianensis* and *Sextonia rubra* (Hardy *et al.* 2006). The existence of high levels of genetic diversity and large Nb values in this undisturbed (and never logged) population at time of logging may also have helped to buffer negative impacts (Hamrick 2004); we may have found different results had the study been conducted in a forest that had previously been subjected to one or more cycles of selective logging, as it is often the case for Amazonian forests following road construction in their vicinity. The fact that Nb estimates indicate apparently higher historical gene flow levels than the pollen dispersal δ estimates also suggests that monitoring of our tree population for a two year period—a short period of time relative to the life span of the tree—may be insufficient to capture all the genetic consequences of selective logging for harvested tree populations.

In other studies, natural self-thinning of tree populations was associated with the loss of SGS over time (*e.g.*, Epperson & Alvarez-Buylla 1997). This seems to have occurred in our population as well, since the neighbourhood size (Nb) estimates for the seedlings are significantly lower than for the juvenile and adult cohorts (Table 5). This supports our observation of apparently higher levels of historical than contemporary gene flow (Table 4), and could be interpreted as evidence that gene flow measured over one or two bouts of reproduction in populations of trees is more restricted than the cumulative gene flow that occurs over several tens (or hundreds) of years (*i.e.*, within the lifetime of a typical tree). On the other hand, weak SGS levels are observed overall in the adults, and minor changes in adult SGS caused by artificial thinning are unlikely to have any consequences for mating patterns in this population. Overall, these results suggests that self-thinning has potential to erase SGS during the aging of the different cohorts, but that selective logging thinning only has a marginal effect on adult SGS in our population.

While we have not attempted to directly estimate seed dispersal in this population, several lines of evidence suggest that the contribution of seed dispersal to overall gene

flow is negligible. At short distances, the estimated average kinship coefficient (F_{ij}) between seedlings is close to 0.125 (Table 5), which is the expected value for maternal half-sibs groups. Moreover, the estimates of average F_{ij} between adult-seedling pairs separated by less than 20 m were 0.148 (0.018) (results not shown), which is around the expected genetic similarity for mother-offspring pairs. Of course we cannot rule out that rare long-distance seed dispersal occurs, but this seems unlikely because: (1) the seeds are heavy; (2) there is no flooding in this area and so seeds cannot be dispersed by water; and (3) many fruits are seen intact on the forest floor for months, suggesting a lack of animal dispersers. However, the absence of animal dispersers at time of seed collection (2003-2005) may also suggest a recent increase in hunting (facilitated by the construction of the roads to extract the trees a few years before) that may have decimated a significant part of the local animal dispersers population at the study site.

In conclusion, when several indicators of genetic diversity are taken into account, this previously undisturbed *Carapa guianensis* population does not seem to have been strongly impacted by the single episode of selective logging that occurred during the course of this investigation. We have verified that neutral genetic diversity, inbreeding level and gene flow do not measurably change after selective logging of trees in this study, which suggests that pollination of *C. guianensis* was adequate immediately after the harvest. Over the long run however, it remains unknown whether selective logging is dysgenic or will cause genetic damage to the population. Future monitoring of this population should address demographic concerns regarding the establishment and the survival of seedling cohorts in the post-logging forest habitat, as well as possible long-term changes to the mating system that could be a consequence of a more gradual erosion of pollinator density and diversity.

Acknowledgements

We thank IBAMA (Brazilian Institute for Environment and Natural Renewable Resources) for access to study site, and staff at Embrapa in Santarem and Belterra, B

Degen and IS Thompson for logistical help. We thank J Campolina and many others from Belterra and Communidade Branco-1 km 83 for help in sampling fruits, and VCR Azevedo, TN Almeida, VP daSilva, MS Borges and F Saade for help in the lab. This study is part of Dendrogene Project: Genetic Conservation in Managed Forests in Amazônia (www.cpatu.embrapa.br/dendro/index.htm) carried out by Embrapa Amazônia Oriental and its partners, within bilateral cooperation between Brazil and United Kingdom, through Brazilian Cooperation Agency (ABC) and Department for International Development (DFID). This work was carried out with the aid of a grant from the International Development Research Center, Ottawa, Canada (www.idrc.ca). D Schoen acknowledges support from an NSERC Discovery Grant. The Fonds Nature et Technologies, Québec (FQRNT), Natural Sciences and Engineering Council, Canada (NSERC) and the Organization of American States (OAS) provided financial support to D Cloutier.

References

Austerlitz F, Dick CW, Dutech C, *et al.* (2004) Using genetic markers to estimate the pollen dispersal curve. Molecular Ecology, 13, 937-954.

Austerlitz F, Smouse PE (2002) Two-generation analysis of pollen flow across a landscape. IV. Estimating the dispersal parameter. Genetics, 161, 355-363.

Austerlitz F, Smouse PE (2001) Two-generation analysis of pollen flow across a landscape. II. Relation between Phi(ft), pollen dispersal and interfemale distance. Genetics, 157, 851-857.

Cloutier D, Povoa JSR, Procopio LC, *et al.* (2005) Chloroplast DNA variation of *Carapa guianensis* in the Amazon basin. Silvae Genetica, 54, 270-274.

Dayanandan S, Dole J, Bawa K, Kesseli R (1999) Population structure delineated with microsatellite markers in fragmented populations of a tropical tree, *Carapa guianensis* (Meliaceae). Molecular Ecology, 8, 1585-1592.

Degen B, Bandou E, Caron H (2004) Limited pollen dispersal and biparental inbreeding in *Symphonia globulifera* in French Guiana. Heredity, 93, 585-591.

Doligez A, Joly HI (1997) Mating system of *Carapa procera* (Meliaceae) in the French Guiana tropical forest. American Journal of Botany, 84, 461-470.

Ellstrand NC, Elam DR (1993) Population genetic consequences of small population size – Implications for plant conservation. Annual Review of Ecology and Systematics, 24, 217-242.

El Mousadik A, Petit RJ (1996) High level of genetic differentiation for allelic richness among populations of the argan tree (*Argania spinosa* (L) Skeels) endemic to Morocco. Theoretical and Applied Genetics, 92, 832-839.

Epperson BK (2003) Geographical genetics. Princeton University Press, Princeton, USA.

Epperson BK (2005) Estimating dispersal from short distance spatial autocorrelation. Heredity, 95, 7-15.

Frankham R (1995) Effective population size / adult population size ratios in wildlife – a review. Genetical Research, 66, 95-107.

Goudet J (1995) FSTAT (version 1.2): A computer program to calculate F-statistics. Journal of Heredity, 86, 485-486.

Guariguata MR, Arias-LeClaire H, Jones G (2002) Tree seed fate in a logged and fragmented forest landscape, northeastern Costa Rica. Biotropica, 34, 405-415.

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

Hamrick JL (2004) Response of forest trees to global environmental changes. Forest Ecology and Management, 197, 323-335.

Hamrick JL, Murawski DA, Nason JD (1993) The influence of seed dispersal mechanisms on the genetic structure of tropical tree populations. Vegetatio, 108, 281-297.

Hardy OJ, Gonzalez-Martinez SC, Colas B, *et al.* (2004) Fine-Scale Genetic Structure and Gene Dispersal in *Centaurea corymbosa* (Asteraceae). II. Correlated Paternity Within and Among Sibships. Genetics, 168, 1601-1614.

Hardy OJ, Maggia L, Bandou E, *et al.* (2006) Fine-scale genetic structure and gene dispersal inferences in 10 neotropical tree species. Molecular Ecology, 15, 559-571.

Hardy OJ, Vekemans X (1999) Isolation by distance in a continuous population: reconciliation between spatial autocorrelation analysis and population genetics models. Heredity, 83, 145-154.

Hardy OJ, Vekemans X (2002) SPAGEDi: a versatile computer program to analyse spatial genetic structure at the individual or population levels. Molecular Ecology Notes, 2, 618-620.

Jennings SB (1997) The response of tree seedlings to canopy disturbances in an Amazonian rain forest. DPhil thesis, University of Oxford.

Kanashiro M, Thompson IS, Yared JAG, *et al.* (2002) Improving conservation values of managed forests: the Dendrogene Project in the Brazilian Amazon. Unasylva, 53, 25-33.

Loiselle BA, Sork VL, Nason JD, Graham C (1995) Spatial genetic structure of a tropical understory shrub, *Psychotria officinalis* (Rubiaceae). American Journal of Botany, 82, 1420-1425.

Lowe AJ, Boshier D, Ward M, *et al.* (2005) Genetic resource impacts of habitat loss and degradation; reconciling empirical evidence and predicted theory for neotropical trees. Heredity, 95, 255-273.

Pollak E (1983) A new method for estimating the effective population size from allele frequency changes. Genetics, 104, 531-548.

Raymond M, Rousset F (1995) GENEPOP (version 1.2) Population genetics software for exact tests and ecumenicism. Journal of Heredity, 86, 248-249.

Ritland K (2002) Extensions of models for the estimation of mating systems using n independent loci. Heredity, 88, 221-228.

Rousset F (2000) Genetic differentiation between individuals. Journal of Evolutionary Biology, 13, 58-62.

Smouse PE, Dyer RJ, Westfall RD, Sork VL (2001) Two-generation analysis of pollen flow across a landscape. I. Male gamete heterogeneity among females. Evolution, 55, 260-271.

Vekemans X, Hardy OJ (2004) New insights from fine-scale spatial genetic structure analyses in plant populations. Molecular Ecology, 13, 921-935.

Vinson CC, Azevedo VCR, Sampaio I, Ciampi AY (2005) Development of microsatellite markers for *Carapa guianensis* (Aublet), a tree species from the Amazon forest. Molecular Ecology Notes, 5, 33-34.

Waples RS (1989) A generalized approach for estimating effective population size from temporal changes in allele frequency. Genetics, 121, 379-391.

Weir BS, Cockerham CC (1984) Estimating F-statistics for the analysis of population structure. Evolution, 38, 1358-1370.

White GM, Boshier DH, Powell W (2002) Increased pollen flow counteracts fragmentation in a tropical dry forest: An example from *Swietenia humilis* Zuccarini. Proceedings of the National Academy of Sciences of the United States of America, 99, 2038-2042.

Tables

Table 1

Allele frequencies and gene diversities at the six microsatellite loci used in *Carapa guianensis*.

	Cg01			Cg06			Cg05			Cg07			Cg16			Cg17		
	Allele	Adult	Seed	Allele	Adult	Seed	Allele	Adult	Seed	Allele	Adult	Seed	Allele	Adult	Seed	Allele	Adult	Seed
1	196[1]	0.769	0.769	139	0.238	0.295	153	0.003	0.001	177	0.022	0.021	108	0.013	0.008	100	0.000	0.001
2	199	0.194	0.191	143	0.107	0.064	157	0.053	0.036	183	0.024	0.085	110	0.008	0.014	104	0.032	0.055
3	202	0.038	0.039	149	0.008	0.001	161	0.021	0.015	185	0.146	0.093	112	0.044	0.054	106	0.048	0.053
4	205	0.000	0.001	151	0.021	0.008	163	0.024	0.006	187	0.100	0.071	114	0.154	0.137	108	0.278	0.299
5				153	0.029	0.011	165	0.021	0.050	189	0.611	0.604	116	0.370	0.404	110	0.067	0.094
6				155	0.235	0.207	167	0.259	0.267	191	0.068	0.099	118	0.188	0.163	112	0.163	0.141
7				157	0.209	0.217	169	0.249	0.201	193	0.005	0.006	120	0.018	0.028	114	0.040	0.017
8				159	0.094	0.148	171	0.217	0.242	195	0.019	0.019	122	0.031	0.021	116	0.126	0.112
9				161	0.029	0.032	173	0.136	0.167	197	0.005	0.002	124	0.010	0.032	118	0.102	0.129
10				163	0.013	0.003	175	0.000	0.001				126	0.003	0.005	120	0.139	0.097
11				165	0.008	0.011	177	0.013	0.002				128	0.013	0.004	122	0.003	0.001
12				167	0.008	0.003	179	0.000	0.001				130	0.023	0.023	124	0.003	0.001
13							183	0.003	0.014				132	0.000	0.001			
14													134	0.068	0.050			
15													136	0.036	0.030			
16													138	0.003	0.001			
17													140	0.018	0.027			
N		337	1488		369	1500		368	1424		362	1380		381	1541		370	1554
H_o		0.38	0.31		0.82	0.80		0.75	0.80		0.53	0.59		0.76	0.76		0.86	0.79
H_e		0.37	0.37		0.82	0.80		0.80	0.80		0.59	0.60		0.80	0.78		0.84	0.84

N, number of alleles sampled; H_o, observed heterozygosity; H_e, gene diversity

[1]: Allele named according to the length of the PCR product in base pairs

Table 2

Genetic diversity averaged over six microsatellite loci in cohorts of *Carapa guianensis*.

Cohort	N	N_{all}	R (SE)	H_o (SE)	H_e (SE)	F_{is}
Pre-harvest seeds	390	10.0	7.7 (2.8)[3]	0.68 (0.21)	0.69 (0.19)	0.025 *
Post-harvest seeds	391	10.6	7.6 (2.8)	0.67 (0.18)	0.70 (0.17)	0.049 *
Seedlings	84	8.8	7.6 (2.9)	0.65 (0.19)	0.71 (0.19)	0.079 *
Juveniles	82	9.0	7.9 (3.3)	0.68 (0.17)	0.72 (0.15)	0.048 *
Adults	199	10.3	8.1 (2.9)	0.69 (0.19)	0.71 (0.18)	0.029 *
Harvested[1]	46	8.3	8.0 (3.2)	0.67 (0.24)	0.71 (0.18)	0.056 ns
Unharvested[2]	49	8.3	7.7 (2.9)	0.65 (0.18)	0.70 (0.18)	0.059 *

N, sample size; N_{all}, number of alleles; R, allelic richness based on a minimum sample size of 33 diploid individuals; H_o, observed heterozygosity; H_e, gene diversity; F_{is}, fixation index (test of significance of the fixation index *:$P<0.05$ and ns: $P>0.05$)

[1]: Harvested adults >53cm DBH

[2]: Adults >53cm DBH left unharvested

[3]: Standard error over loci in parentheses

Table 3

Mating system of *Carapa guianensis* as estimated from seeds collected before and after selective logging.

Mating system parameter	Before logging		After logging	
t_m (SE)	0.939	(0.027)	0.927	(0.033)
t_s (SE)	0.924	(0.028)	0.899	(0.041)
$t_m - t_s$ (SE)	0.015	(0.015)	0.028	(0.018)
r_p (SE)	0.050	(0.014)	0.054	(0.013)

t_m, multilocus outcrossing rate; t_s, singlelocus outcrossing rate; r_p, multilocus correlation of outcrossed paternity

Table 4

Pollen dispersal analyses based on estimates of pollen pool differentiation before and after selective logging in *Carapa guianensis*.

Type of estimate	Dispersal model	Parameter estimated	Before logging[1] (95% CI)	After logging[2] (95% CI)
Estimation of δ using global ϕ_{FT}	Normal	δ_{min}[4]	75 (63-102)[3]	76 (67-98)
		δ_{max}[5]	237 (206-329)	240 (214-311)
	Exponential	δ_{min}	85 (71-115)	86 (75-110)
		δ_{max}	265 (230-369)	268 (240-348)
Joint estimation of *d* and δ using pairwise ϕ_{FT}	Normal	*d*	0.30 (0.06-3.01)	0.75 (0.19-2.08)
		δ	204 (2-358)	123 (5-230)
	Exponential	*d*	0.25 (0.04-9.36)	0.55 (0.10-581)
		δ	252 (18-534)	159 (5-306)

ϕ_{FT}, pollen pool differentiation; δ, average distance of pollen dispersal (m); *d*, adult effective density (trees/ha)

[1]: Adult field density >30 cm DBH before logging, 2.5 trees/ha

[2]: Adult field density >30 cm DBH after logging, 2.0 trees/ha

[3]: Estimated 95 percent confidence intervals based on bootstrapping the seed families

[4]: Assuming adult field density (d_{max})

[5]: Assuming adult field density x 1/10 (d_{min}) (see methods for details)

Table 5

Spatial genetic structure averaged over loci and Wright's neighbourhood size in cohorts of *Carapa guianensis*.

Cohort	N	b_k (SE)	$F_{ij}^{[1]}$ (SE)	Nb_1 (95% CI)	$R_{ij}^{[1]}$ (SE)	Nb_2 (95% CI)
Seedling	3486	-0.0179*	0.0480*	53	0.0698*	104
		(0.0025)	(0.0078)	(42-74)	(0.0110)	(84-130)
Juvenile	3321	-0.0047	0.0156	209	0.0258*	161
		(0.0031)	(0.0104)	(90-∞)	(0.0160)	(118-∞)
Adult	19701	-0.0045*	0.0070	221	0.0175	174
Pre-harvest		(0.0024)	(0.0080)	(107-∞)	(0.0134)	(134-∞)
Adult	9591	-0.0063*	0.0041	-[1]	0.0154	-
Post-harvest		(0.0037)	(0.0128)		(0.0243)	

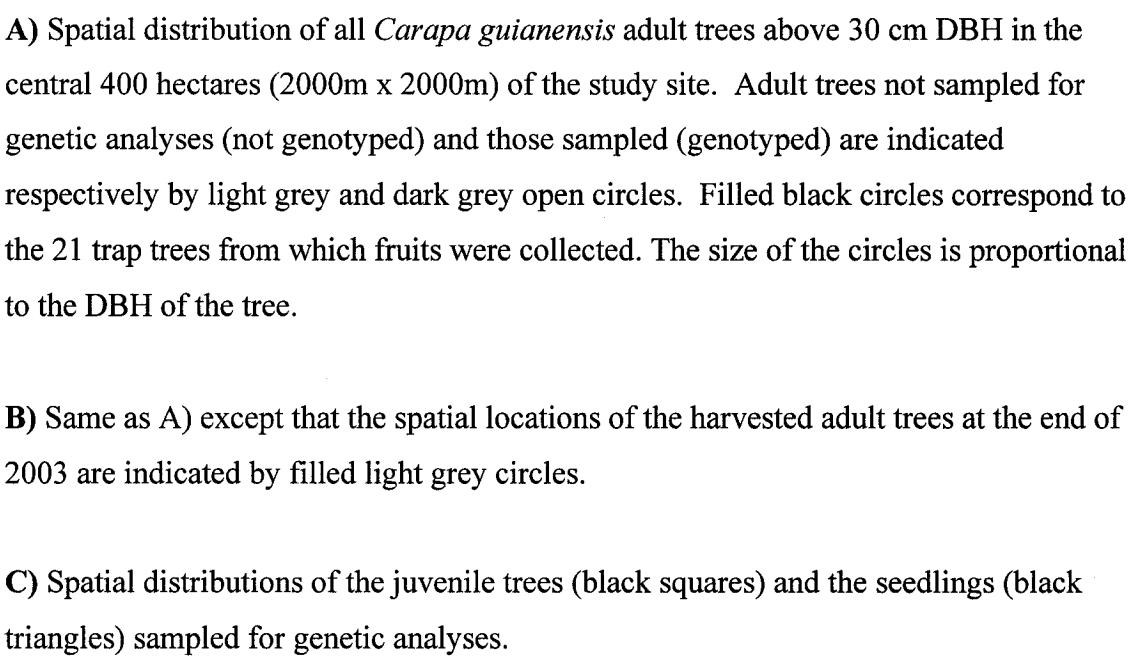
N, total number of pairs; b_k, slope of the regression of F_{ij} values over the natural logarithm of the spatial distance; $F_{ij}^{[1]}$, average kinship coefficient between neighbour individuals; $R_{ij}^{[1]}$, average Moran's I index between neighbour individuals; Nb_1, neighbourhood size estimated as $-(1-F_{ij}^{[1]})/b_k$; Nb_2, neighbourhood size estimated as $\exp[(0.544-R^{[1]})/0.102]$

*, type I error rate < 0.05 associated with the hypothesis of no isolation by distance (see methods for details)

[1]: Not estimated since the pre-harvest estimates are accurate regarding historical dispersal-drift balance

Figures

Figure 1



A) Spatial distribution of all *Carapa guianensis* adult trees above 30 cm DBH in the central 400 hectares (2000m x 2000m) of the study site. Adult trees not sampled for genetic analyses (not genotyped) and those sampled (genotyped) are indicated respectively by light grey and dark grey open circles. Filled black circles correspond to the 21 trap trees from which fruits were collected. The size of the circles is proportional to the DBH of the tree.

B) Same as A) except that the spatial locations of the harvested adult trees at the end of 2003 are indicated by filled light grey circles.

C) Spatial distributions of the juvenile trees (black squares) and the seedlings (black triangles) sampled for genetic analyses.

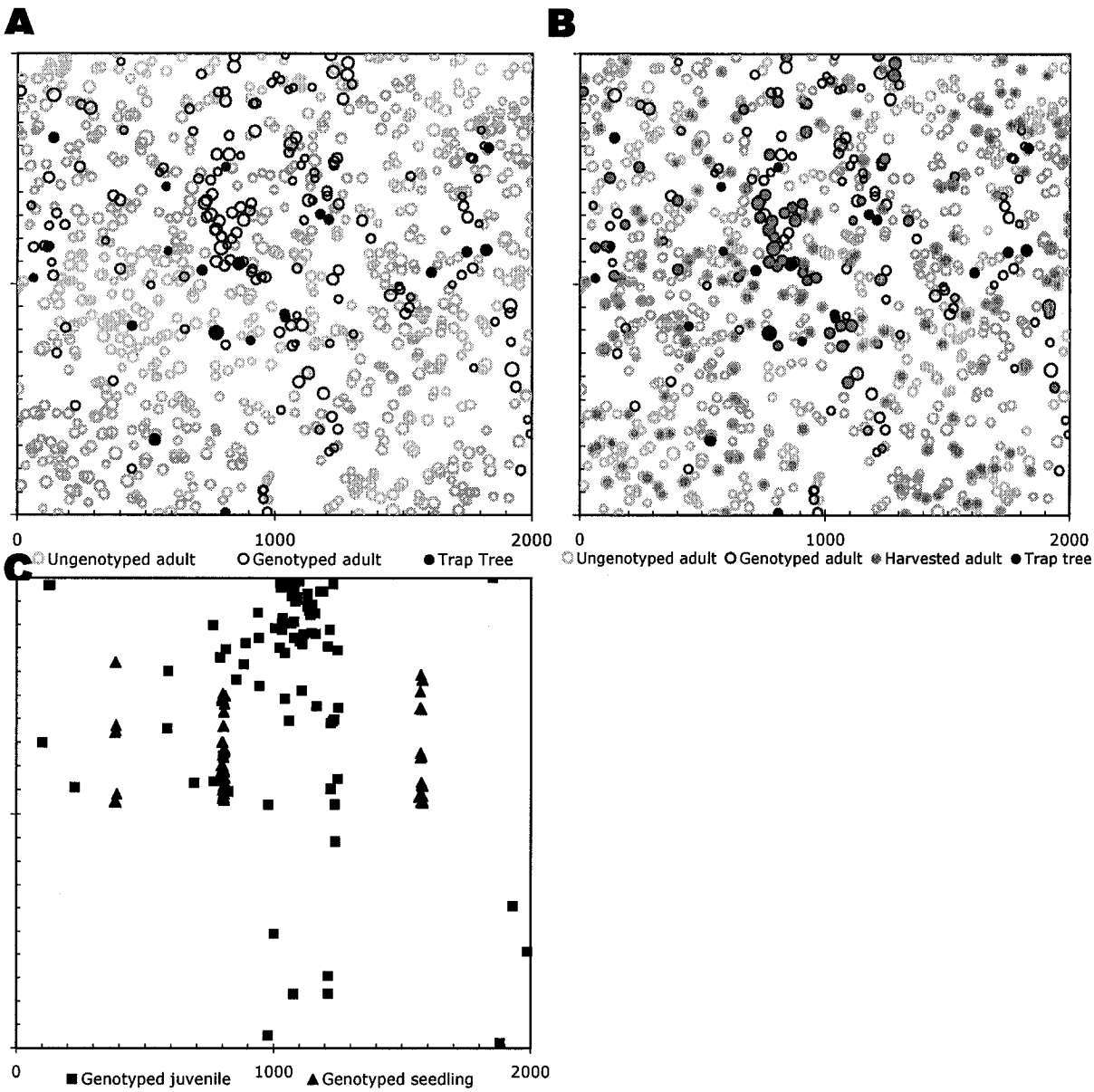

A
B
0
1000
2000
Ungenotyped adult
Genotyped adult
Trap Tree
Harvested adult
Trap tree
C
Genotyped juvenile
Genotyped seedling

Figure 2

Seasonal variation in the mating system of *Carapa guianensis*. Circles, squares and lozenges symbols represent respectively the outcrossing rate (t_m), the level of biparental inbreeding (t_m-t_s) and the correlation of paternity (r_p). Filled symbols indicate values estimated from trees that flowered after harvesting. Bars correspond to the standard error of the estimates.

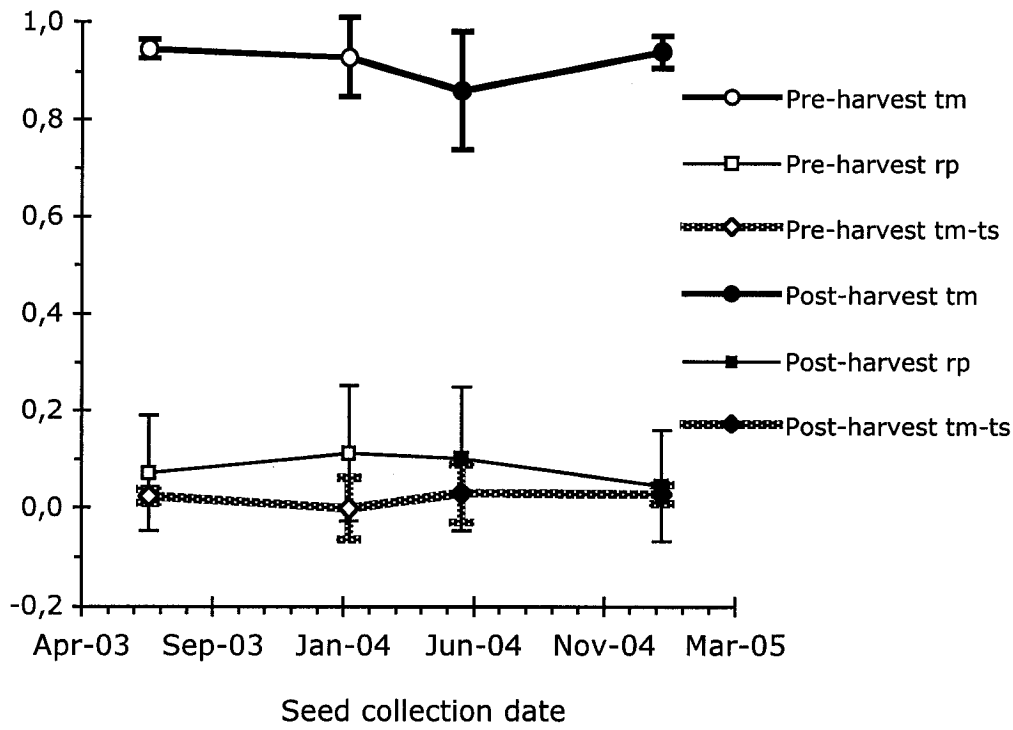
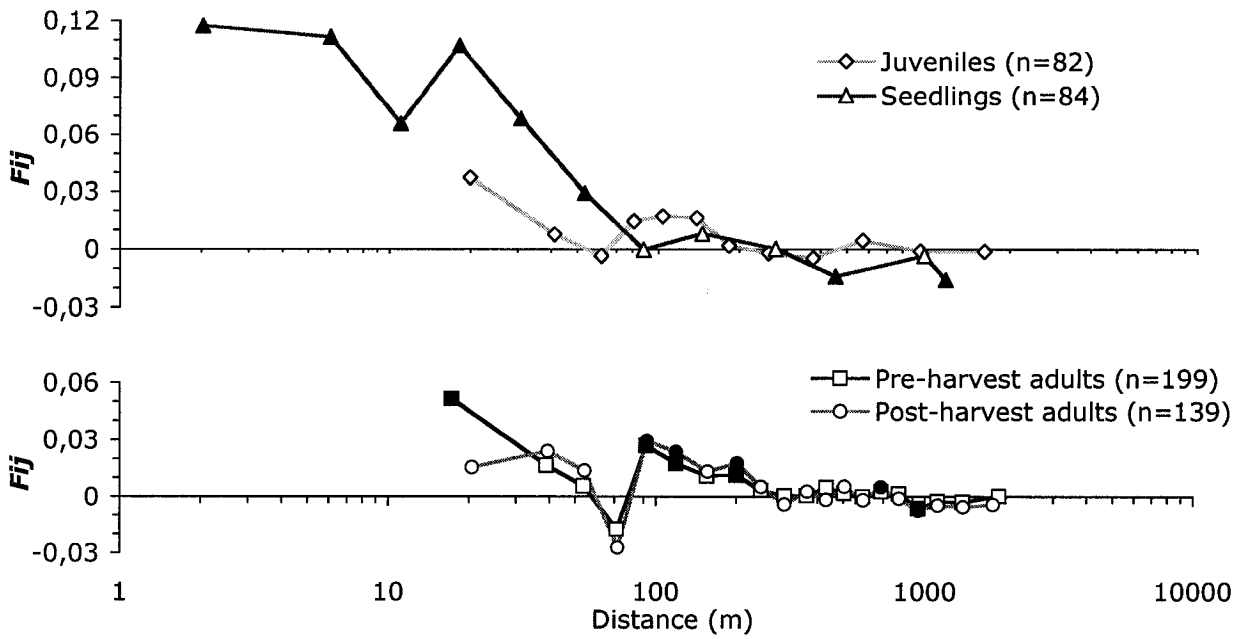


Figure 3

Mean pairwise kinship coefficient (F_{ij}) between individuals versus inter-individual distance for the seedling and juvenile cohorts (upper), and the pre-harvest and post-harvest adult population (lower). Filled symbols show values departing significantly (α=5%) from a random spatial distribution of genotypes.



Linking statement for Chapter 4

Having documented the dynamics of within-population genetic diversity in pre- and post disturbance environment in Chapters 2 and 3, and having shown that at the population level, trees are highly connected through long-distance gene dispersal, it is of interest to know whether this pattern of high genetic connectivity is also present at larger spatial scales. *Carapa guianensis* is a widespread species occupying a large part of the Neotropics, and the genus *Carapa* is found both in the American and African continents, suggesting that *Carapa guianensis* is an ancient species. Phylogeographic studies using variation at the maternally inherited chloroplast DNA among different populations of the same species can give some insight into current levels of gene flow among populations and the historical dynamics of tree migration. To date, there are no published phylogeographical studies of tree populations focusing on the Amazon basin, partially because of limited physical access, and legal restrictions to the access of the material. The forth chapter of this thesis is a preliminary study of phylogeographical chloroplast DNA variation in *Carapa guianensis*, a forest tree species capable of water dispersal. It is one of the first such studies to be conducted with a tropical tree species.

Chapter 4:

Chloroplast DNA variation of *Carapa guianensis* in the Amazon basin

Cloutier D, Povoa JSR, Procopio LC, Leao NVM, Wadt LHO, Ciampi AY, Schoen DJ. 2005. Chloroplast DNA variation of *Carapa guianensis* in the Amazon basin. ***Silvae Genetica*** 54: 270-274.

Abstract

Carapa guianensis is a widespread Neotropical tree species that produces a seed adapted for water dispersal. We conducted a pilot study of chloroplast DNA (cpDNA) variation in order to investigate the consequences of hydrochory on genetic diversity and geographic population structure in the lower Amazon basin. A survey of cpDNA haplotype variation reveals strong regional structure, which suggests limited gene flow by seeds. Within site variation was detected only in one floodplain forest (*varzea*), suggesting that seed dispersal by water in these forests has the potential to mix maternal lineages. Several phylogeographic hypotheses are discussed with respect to these data.

Introduction

An understanding of gene dispersal can help account for patterns and levels of biodiversity. In plant populations, genes disperse through the movement of pollen and seeds. Seed dispersal, and in particular, long-distance seed dispersal, allows for migration and colonization of new areas, and contributes to the genetic structure of populations.

Chloroplast DNA (cpDNA) is usually maternally inherited in angiosperms, and is thereby dispersed by seed. This allows one to use patterns of variation in cpDNA to help infer historical processes of seed migration and colonization. Since cpDNA evolves slowly, the observed patterns of cpDNA variation are likely to arise from processes occurring over long time periods and large spatial scales. CpDNA has emerged as the marker of choice for examining patterns of gene flow through seed dispersal. An especially interesting set of examples is the study of Pleistocene tree colonization in Europe (Petit *et al.* 2003).

The few studies using cpDNA that have been conducted in the tropics suggest that spatiotemporal patterns of cpDNA variation are the product of both biogeographical history and contemporary seed dispersal, but the imprint of Pleistocene recolonization

events in genetic diversity may be more diffuse and difficult to detect in the tropics compared with temperate zones (Caron *et al* 2000, Dutech *et al.* 2000, Cavers *et al.* 2003). Despite long-standing interest in the origin and history of biodiversity in the Amazon basin, and the potential role of the Amazon River system as a conduit of dispersal, to date there have been no published studies using cpDNA to infer phylogeographic patterns there.

Hydrochory (seed dispersal by water) is an important means of seed dispersal in several types of rainforests that flood seasonally or tidally. Near the ocean, seed from forests subjected to tidal flooding may be dispersed both upstream or downstream. Fish could also play a role in dispersing seed, and this could further contribute to upstream dispersal in seasonally inundated forests (Gottsberger 1978). In areas where seed are dispersed by water, there is the potential for extensive genetic exchange between populations in the Amazon basin. Therefore we expect cpDNA variation to be weakly structured at this scale. We would also expect to find a higher level of variation at downstream flooded sites located near the Amazon River mouth compared with upstream sites.

This study was initiated in an attempt to assess the consequences of hydrochory in shaping the genetic diversity of the Neotropical forest tree, *Carapa guianensis* (Meliaceae). We describe cpDNA variation for this species as revealed by polymerase chain reaction (PCR) of chloroplast intergenic sequences, followed by analysis of restriction fragment length polymorphisms (RFLPs), and we use the variation observed to examine hypotheses pertaining to seed dispersal in this species. Specifically, we examine: (1) whether genetic differentiation among sites is consistent with the presence of hydrochory in the Amazon basin; and (2) whether there is evidence that flooded forests receive seed migrants from distant sites.

Materials and Methods

Study species

Carapa guianensis Aubl. (Meliaceae) is a widespread Neotropical tree found in the Caribbean islands and Central America to northern South America, including Amazonia. The species inhabits upland *terra firme* forests as well as floodplain *varzea* forests in the Amazon basin. A second species in the genus, *C. procera*, is also found in the New World. This species is restricted to the Guyanas and Central Amazonia, but is also found in Africa. The two species can be differentiated from one another by reproductive morphology (Pennigton 1981, Dendrogene 2004), but it is difficult to distinguish trees and saplings in the field when flowers or fruits are absent. At the seedling stage, *C. guianensis* has composite leaves, while *C. procera* has simple leaves (Fisch *et al.* 1995).

The seed of *Carapa guianensis* is composed of a large kernel, variable in weight (10-50 g), and containing a high proportion of lipids. The seeds and seedlings are typically found directly beneath the parent tree, and occasionally are dispersed a few meters by scatter-hoarding rodents (Forget *et al.* 1999). Along streams and in areas subjected to flooding, however, the seed may float and be carried by water currents. They are often found in large numbers on the riverbanks at time of seed release. The seeds of this species are adapted for water dispersal and can easily germinate and establish after a floating phase of several months (Scarano *et al.* 2003). *C. guianensis* is also used by the local human population as source of timber and medicinal oil, which is extracted from the seed.

Sampling

Due to legal restrictions on sampling *Carapa guianensis* material for genetic analyses, we used material held by the EMBRAPA (Empresa Brasileira de Pesquisa Agropecuaria) research network, together with seedlings sold by local merchants. Twenty-two dried samples from the Ian herbarium of EMBRAPA-Bélem and one sample from the Cen herbarium of EMBRAPA-Brasilia were collected. These samples represent

diverse locations from Brazil, Venezuela and Suriname, and collection dates ranging from 1942 to 2002. The fresh samples were collected from EMBRAPA research sites, or bought from people involved in the trade of seed and seedling of tree species (Table 1). These samples were dried in silica gel before being sent to the lab.

Laboratory analyses

For the herbarium specimens, DNA was extracted using the DNeasy plant mini-kit (Quiagen) following the standard protocol with slight modifications as described in Drabkova *et al.* (2002). For the fresh samples, DNA was extracted from about 200 mg of leaf tissue using a standard CTAB protocol. Chloroplast DNA intergenic sequences were amplified using a set of 16 universal primer pairs developed by Taberlet *et al.* (1991), Demesure *et al.* (1995) and Dumolin-Lapegue *et al.* (1997). The PCR reaction mixture was in a final volume of 13μL containing: DNA (0.3ng/μL), reaction buffer (10mM Tris-HCl, pH 8.3, 50mM KCl, 1.5mM $MgCl_2$), primers (0.5μM each), $MgCl_2$ (1.5mM), BSA (0.25mg/mL), dNTP (0.25mM) and *Taq* polymerase (1.3 U, Invitrogen). PCR amplifications were carried out using 1 cycle of 4 min 94°C, 30 cycles of 45 s at 92°C, 45 s at 54°C, 3 min at 72°C and one cycle of 10 min at 72°C. PCR products were digested using various restriction enzymes, following manufacturer's specifications. Restriction fragments were then separated in 3% agarose gels stained with ethidium bromide and visualized by UV fluorescence. Several digital pictures were taken during migration to maximize the chances of detecting variants.

To allow an initial screen of geographical variation a subset of chloroplast amplification products from Acre, Belterra, Porto Trombetas and Belem were digested with four different restriction enzymes (*Hinf*I, *Taq*1, *Mse*I, *Alu*I). Those enzyme-primer pair combinations that revealed variation among locations were used for analysis of the remaining samples. When several different enzymes revealed variation for a given primer pair, the enzyme-primer pair combination that gave the clearest variable bands under migration conditions used was chosen. Eight primer pairs gave no or only weak amplification products, and these were not used in subsequent analyses. Of the eight

chloroplast PCR products tested, three were apparently monomorphic with the enzymes tested, and therefore were excluded in subsequent analyses, leaving a total of five polymorphic PCR products.

Data analysis

The Neighbour-Joining (NJ) tree of haplotypes was constructed using PAUP* (Swofford 2003), with the characters (restriction sites, indels) coded as present or absent, and distance expressed as mean character distance. Scoring haplotypes as ordered and unordered alleles respectively, we estimated levels of genetic differentiation (G_{st} and N_{st}), average within-location diversity (H_S and V_S), and total diversity (H_T and V_T) using the software PERMUT (Pons and Petit 1995; 1996). We also calculated the mean number of differences between haplotypes (D_m) and the weighted mean number of differences between haplotypes (D_{wm}). For these analyses, locations with a single individual assessed for cpDNA variation were omitted, however Belem and Santa Barbara, and the 2 Curua-Una locations, were treated as two synthetic locations because they were separated by less than 100 km. Since we can not rule out anthropogenic seed transfer in the plantation samples analyzed, we conducted two separate sets of analyses, one including all the 11 locations, and another including only the 7 locations of "natural" origin.

Results

The amplification of cpDNA regions from dried herbarium specimens failed, likely because of the low quality of the DNA extracted from these samples. Only short sequences < 500 bp could be amplified (results not shown).

A total of 53 samples from 19 sites yielded DNA amenable to PCR. Chloroplast variation is described in Tables 1, 2 and 3. Six polymorphisms were observed--four indels and two restriction site variants. The geographic distribution and phylogenetic

relationship of these haplotypes is shown in Figure 1 and Figure 2. A single haplotype was found in all locations except Gurupa-Island, which is an island on the Amazon River with floodplain *varzea* forest. The analysis of genetic differentiation is presented in Table 3.

Discussion

Contrary to our hypothesis that hydrochory has played a major role in determining phylogeographic patterns in this species, the general pattern of cpDNA variation (Figure 1) is consistent with limited seed dispersal at the scale of the Amazon basin. The global genetic differentiation measures (Table 4) are higher than the average reported in plant species (Petit *et al.* 2005), which suggests that gene flow is limited. This may be explained by the fact that hydrochory does not result in isotropic gene flow on a large scale, and that most individuals in upland sites do not have opportunity to exchange seed over long distances. There is, however, limited evidence that hydrochory allows some exchange between populations, since the only sampled location exhibiting within-site variation in cpDNA is a flooded *varzea* population situated on an island in the Amazon River ("Gurupa-Island") (Table 1). But while the observed mixing of haplotypes is evidence that hydrochory may influence the genetic diversity of *Carapa guianensis* populations, this influence appears to be localized and does not seem to have promoted extensive genetic mixing. Although we analyzed only a limited number of individuals per site (for 10 sites a single individual was analyzed) our results suggest that most locations are fixed for a single haplotype, as reported for other species (*e.g.,* Cavers *et al.* 2003). More locations will need to be analyzed to test whether downstream locations could be locations of genetic mixing in this hydrochorous tree species.

Several historical hypotheses have been put forward to explain patterns of biodiversity in the Amazon basin, and may help to explain the pattern of cpDNA variation found in this study. These include the presence of ecological gradients (Endler 1982), the dynamics of Amazonian rivers (Salo *et al* 1986), past climatic changes with retreat of forest habitat to forest refuges (Haffer 1969), and the glacial cycles with sea

level changes causing marine incursion into Amazonia (Vuilleumier 1971, Nores 1999). The last two hypotheses deserve some consideration here.

In Eastern Amazonia, there is evidence from studies of lake pollen that forest regression and savanna expansion occurred during the Pleistocene (Absy *et al.* 1991, Van der Hammen and Absy 1994), suggesting that reduction of rainfall in the past converted areas presently receiving less than 2000 mm of rain per year into savanna. If this climatic reconstruction is correct, recently recolonized zones could harbor a subset of cpDNA diversity that were present in the neighboring refugial areas. From our results, it can be seen that the total chloroplast diversity in the dry zones putatively "recolonized" does not appear lower (H_T= 0.60; V_T= 0.60; D_{wm}=6.0) (8 locations) than the diversity in the combined putative refugial areas (H_T= 0.62; V_T= 0.22; D_{wm}=1.3) (11 locations). However our sample size is small, and additional samples would be required to rigorously test this hypothesis. But our preliminary results do not support the hypothesis that a Pleistocene "refuge" Amazonian forest along the Atlantic coast has acted as the main source of recolonization for the interior by *Carapa guianensis*.

The phylogeny and geographic distribution of the haplotypes suggest that this species has colonized the Amazon basin from multiple sources (Figures 1 and 2). Large areas with low topography were flooded during past marine introgressions into Amazonia that occurred during the Tertiary and the Quaternary (Nores 1999). Interestingly the distribution of haplotypes "C" and "E", differentiated by at least six independent mutations, corresponds roughly to high elevations areas of the Amazon basin, respectively the Guyanan shield in the North, and the Brazilian shield in the South. Therefore these areas may have been historical "refuges" for *Carapa guianensis* or may represent two independent colonization events. Conservation and management efforts should take into account the presence of such cryptic genetic differences in the Amazon basin in *Carapa guianensis*.

Acknowledgements

We would like to thank the IBAMA office in Santarem, L. Anchieta, I. Thompson, M. Kanashiro and RCVM da Silva for the logistical help. Thanks to Geane Alvarenga, Valdecy de Santos, Isaac Monteiro, staff at RESEX Verde para Sempre and at Flona Tapajos for help in sampling. This study is part of Dendrogene Project: Genetic Conservation in Managed Forests in Amazônia (2000-2004, www.cpatu.embrapa.br/dendro/index.htm) carried out by Embrapa Amazônia Oriental and its partners, within bilateral cooperation between Brazil and United Kingdom, through Brazilian Cooperation Agency (ABC) and Department for International Development (DFID). D. Schoen acknowledges support from an NSERC Discovery Grant. The International Development Research Center, Canada (IDRC), Fonds Nature et Technologies, Québec (NATEQ), Natural Sciences and Engineering Council, Canada (NSERC) and the Organization of American States (OAS) provided financial support to D. Cloutier.

References

Absy, M.L., A. Cleef, M. Fournier, L. Martin, M. Servant, A. Sifeddine, M.F. DaSilva, F. Soubies, K. Suguio, B. Turcq, T. VanDerHammen, 1991: Occurrence of 4 episodes of rain-forest regression in Southeastern Amazonia during the last 60,000 YRS – 1st comparison with other tropical regions. Comptes Rendus de l'Académie des Sciences Série II 312: 673-678.

Caron, H., S. Dumas, G. Marque, C. Messier, E. Bandou, R.J. Petit, A. Kremer, 2000: Spatial and temporal distribution of chloroplast DNA polymorphism in a tropical tree species. Molecular Ecology 9: 1089-1098.

Cavers, S, C. Navarro, A.J. Lowe, 2003: Chloroplast DNA phylogeography reveals colonization history of a Neotropical tree, Cedrela odorata L., in Mesoamerica. Molecular Ecology 12: 1451-1460.

Demesure, B, N. Sodzi, R.J. Petit, 1995: A set of universal primers for amplification of polymorphic noncoding regions of mitochondrial and chloroplast DNA in plants. Molecular Ecology 4: 129-131.

Dendrogene, 2004: ANDIROBA, /Carapa guianensis/. Belém,PA: Embrapa Amazônia Oriental : Projeto Dendrogene, 2004. (Projeto Dendrogene. Espécies Arbóreas da Amazônia, 2).

Drabkova, L, J. Kirschner, C. Vlcek, 2002: Comparison of seven DNA extraction and amplification protocols in historical herbarium specimens of Juncaceae. Plant Molecular Biology Reporter 20: 161-175.

Dumolin-Lapegue, S, M.H. Pemonge, R.J. Petit, 1997: An enlarged set of consensus primers for the study of organelle DNA in plants. Molecular Ecology 6: 393-397.

Dutech, C, L. Maggia, H.I. Joly, 2000: Chloroplast diversity in Vouacapoua americana (Caesalpiniaceae), a Neotropical forest tree. Molecular Ecology 9: 1427-1432.

Endler, J.A., 1982: Problems in distinguishing historical from ecological factors in biogeography. American Zoologist 22: 441-452.

Fisch, S.T., I.D.K. Ferraz and W.A. Rodrigues, 1995: Distinguishing Carapa guianensis Aubl. from Carapa procera DC (Meliaceae) by morphology of young seedlings. Acta Amazonica 25: 193-200.

Forget, P.-M., F. Mercier, F. Collinet, 1999: Spatial patterns of two rodent-dispersed rain forest trees Carapa procera (Meliaceae) and Vouacapoua americana (Caesalpiniaceae) at Paracou, French Guiana. Journal of Tropical Ecology 15: 301-313.

Gottsberger, G., 1978: Seed dispersal by fish in inundated regions of Humaita, Amazonia, Biotropica 10: 170-183.

Haffer, J., 1969: Speciation in Amazonian forest birds. Science 165: 131.

Nores M., 1999: An alternative hypothesis for the origin of Amazonian bird diversity, Journal of Biogeography 26: 475-485.

Pennington T.D. (1981) Meliaceae, Flora Neotropica, Monograph 28, Bronx: The New York Botanical Garden.

Petit, R.J., J. Duminil, S. Fineschi, A. Hampe, D. Salvini, G.G. Vendramin, 2005: Comparative organization of chloroplast, mitochondrial and nuclear diversity in plant populations. Molecular Ecology 14: 689-701.

Petit, R.J., I. Aguinagalde, J.L. de Beaulieu, C. Bittkau, S. Brewer, R. Cheddadi, R. Ennos, S. Fineschi, D. Grivet, M. Lascoux, A. Mohanty, G.M. Muller-Starck, B. Demesure-Musch, A. Palme, J.P. Martin, S. Rendell, G.G. Vendramin, 2003: Glacial refugia: Hotspots but not melting pots of genetic diversity. Science 300: 1563-1565.

Pons, O., R.J. Petit, 1996: Measuring and testing genetic differentiation with ordered versus unordered alleles. Genetics 144: 1237-1245.

Pons, O., R.J. Petit, 1995: Estimation, variance and optimal sampling of gene diversity. 1. Haploid locus. Theoretical and Applied Genetics 90: 462-470.

Salo, J., R. Kalliola, I. Hakkinen, Y. Makinen, P. Niemela, M. Puhakka, P.D. Coley, 1986: River dynamics and the diversity of Amazon lowland forest. Nature 322: 254-258.

Scarano, F.R., T.S. Pereira, G. Rocas, 2003: Seed germination during floatation and seedling growth of Carapa guianensis, a tree from flood-prone forests of the Amazon. Plant Ecology 168: 291-296.

Swofford, D. L. 2003. PAUP*. Phylogenetic Analysis Using Parsimony (*and Other Methods). Version 4. Sinauer Associates, Sunderland, Massachusetts.

Taberlet, P., L. Gielly, G. Pautou, J. Bouvet, 1991: Universal primers for amplification of 3 noncoding regions of chloroplast DNA. Plant Molecular Biology 17: 1105-1109.

VanDerHammen, T, M.L. Absy, 1994: Amazonia during the last glacial. Palaeogeography Palaeoclimatology 109: 247-261.

Vuilleumier, B.S, 1971: Pleistocene changes in fauna and flora of South-America. Science 173: 771.

Tables

Table 1

Carapa guianensis populations analyzed for chloroplast DNA variation.

Location	Number* and name of location	Source of material	Species	Sample size	Haplotype
00°05'N 51°08'W	1.Macapa AP	Terra firme forest	*C. guianensis/procera*	1	A
01°00'S 51°30'W	2.Gurupa Island PA	Varzea forest	C. guianensis	9	A(8), B(1)
01°10'S 48°20'W	3.Santa Barbara PA	Terra firme forest	*C. guianensis*	1	A
01°15'S 56°38'W	4.Porto Trombetas PA	Tree nursery	*C. guianensis*†	3	C
01°20'S 51°38'W	5.Gurupa Town PA	Seedlings on riverbanks	*C. guianensis*†	6	A
01°25'S 53°48'W	6.Prainha PA - Jatuarana	Terra firme forest	*C. guianensis*	1	A
01°30'S 48°30'W	7.Belem PA - Embrapa	Plantation	*C. guianensis*	1	A
01°34'S 56°15'W	8.Oriximina PA - Trombetas river	Plantation	*C. guianensis/procera*	1	D
01°40'S 55°52'W	9.Oriximina PA - Town	Tree nursery	*C. procera*†	3	C
01°54'S 52°27'W	10.Porto de Moz PA - Arurua	Plantation	*C. guianensis*	1	A
02°31'S 54°39'W	11.Santarem PA - Curua-Una road	Terra firme forest	*C. guianensis*	1	E
02°32'S 54°07'W	12.Santarem PA - Curua-Una	Plantation	*C. guianensis*	1	E
02°50'S 52°01'W	13.Vitoria do Xingu PA - Town	Plantation	*C. guianensis*	1	A
02°50'S 55°01'W	14.Belterra PA - Flona km64	Terra firme forest	*C. guianensis*	2	E
03°01'S 54°58'W	15.Belterra PA - Flona km83	Terra firme forest	*C. guianensis*	3	E
06°13'S 57°55'W	16.Jacareacanga PA	Terra firme forest	*C. guianensis/procera*	1	E
09°45'S 67°38'W	17.Porto Acre AC	Terra firme forest	*C. guianensis*	4	F
10°01'S 67°41'W	18.Rio Branco AC	Varzea forest	*C. guianensis*	6	F
10°02'S 67°40'W	19.Rio Branco AC	Terra firme forest	*C. guianensis*	7	F

*: Numbers used in the distribution map (Fig. 1b)

†: Species identification based on presence of composite or simple leaves at seedling stage

Table 2

Chloroplast DNA haplotypes in *Carapa guianensis* based on presence (1) or absence (0) of the least common variable band for the primer pairs and restriction enzymes used.

Haplotype	Primer pairs and restriction enzymes used							
	trnH-trnK*	trnK-trnK*	trnK-trnK*	trnK-trnK*		trnC-trnD *	trnD-trnT *	trnF-trnVr †
	*Mse*I	*Hinf*I	*Mse*I	*Taq*1#		*Hinf*I	*Alu*I	*Hinf*I
A	0	0	0	0	0	0	0	0
B	1	1	1	0	0	1	0	0
C	0	0	0	0	1	0	1	1
D	0	0	0	0	1	0	0	0
E	1	0	1	0	0	1	0	0
F	0	0	0	1	0	0	0	0

*: Primer sequence from Demesure *et al.* (1995)

†: Primer sequence from Dumolin-Lapegue *et al.* (1997)

#: The digestion of trnK-trnK with *Taq*1 gives two markers

Table 3

Global genetic differentiation and diversity levels at chloroplast DNA among populations of *Carapa guianensis* in the Amazon basin.

Type of measures	Parameters	Values (SD) for natural populations (7 locations)	Values (SD) for all populations (11 locations)
Unordered alleles	G_{st}	0.96 (0.04)	0.98 (0.02)
	H_S	0.03 (0.03)	0.02 (0.02)
	H_T	0.77 (0.06)	0.82 (0.02)
Ordered alleles	N_{st}	0.94 (0.07)	0.97 (0.03)
	V_S	0.04 (0.04)	0.02 (0.02)
	V_T	0.69 (0.13)	0.73 (0.11)
	D_m	3.0	3.8
	D_{wm}	2.7	3.4

Figures

Figure 1

(**A**) Geographic range of *Carapa guianensis* (light gray) and areas where both *C. guianensis* and *C. procera* are found (dark gray).

(**B**) Map of the 19 locations sampled for chloroplast DNA variation in *Carapa guianensis* in the Amazon basin. Each location is represented by a numbered circle of size proportional to the number of individuals sampled, and each haplotype is represented by a different color (see Table 1 for details).

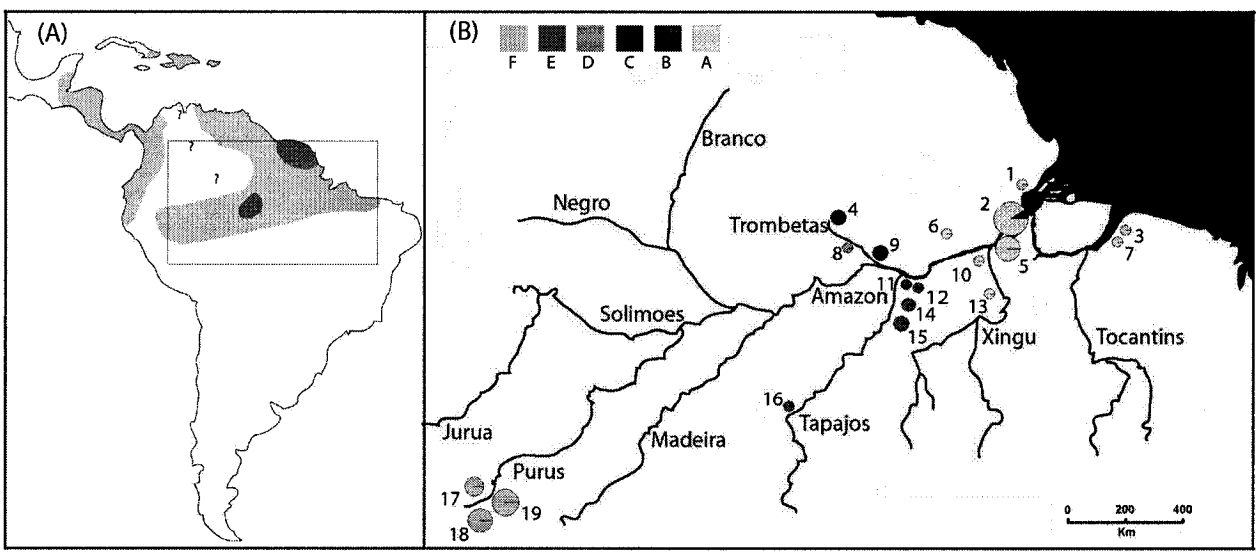
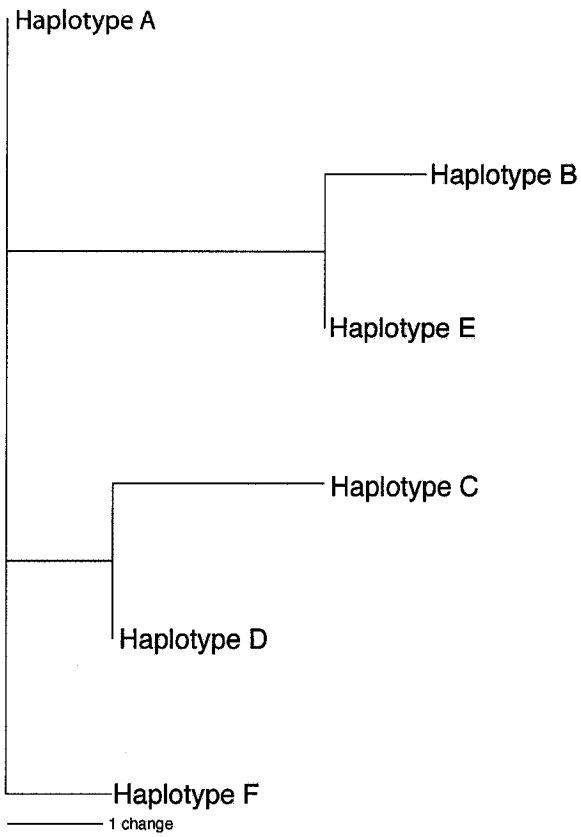


Figure 2

Neighbor-joining tree of *Carapa guianensis* chloroplast haplotypes.



General conclusions

This PhD thesis investigated whether selective logging would impact the levels of genetic diversity found in unexploited Amazonian forest tree populations. In Chapter 1, I have checked the mutability of microsatellite genetic markers, the main genetical tool used to infer patterns of mating in the populations. It was found that while such markers provide the genetic variability necessary to conduct within-population level studies, mutability within the lifetime of a tree is not an obstacle to their use in studies of inbreeding and gene flow. In Chapter 2, I have documented low levels of inbreeding and high levels of within-population gene flow in two Amazonian forest tree species pollinated by small insects. This study has addressed the paucity of information on gene flow and inbreeding among tropical plants, relative to their temperate counterparts. In Chapter 3, I reported that the structure and the dynamics of genetic diversity within an Amazonian tree population were not measurably altered by selective logging. In Chapter 4, I demonstrated the presence of among-population genetic diversity that may have been shaped by historical (tree seed migration) or current (seed dispersal by water) events.

The entirety of the population genetic information collected in these studies from Amazonian harvested forest tree species suggests that the genetic makeup of these populations may be resistant to the exploitation of their resources, such as when selective logging is practiced. This is basically because of the high level of the genetic connectivity among and within populations, gene exchange due to outcrossing among distant individuals, and because effective population sizes are large enough to counteract the effects of genetic drift, at least for the harvesting levels practiced during this study. This population genetic dynamics may be linked to specific features of the *Carapa guianensis* population we investigated, such as the small size and the large number of reproductively mature trees, and to non-synchronous phenology at the population level, but additional work is needed to clarify this potential cause and effect relationship.

The results of this PhD research project are good news for "conservationists" and "exploitationists" alike, because they provide some evidence that the use and the

conservation of Amazonian forest resources can co-exist. However, far more work is needed, at the scientific, legal and political level, to ensure that genetic diversity, as a component of biodiversity, is preserved during exploitation, and perhaps even enhanced, in a way that allows humans to enjoy its benefits in the future. It will be important to extend studies such as this one to larger time and spatial scales, and to additional species, as our information about the population genetic consequences of tree harvesting is based on very few studies.