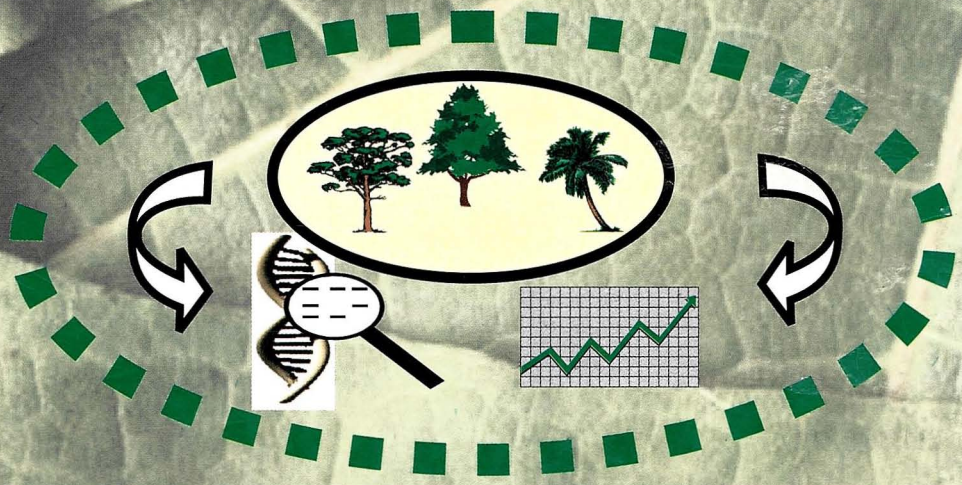
Ministério da Agricultura
Pecuária e Abastecimento

Modelling and experimental research on genetic processes in tropical and temperate forests



18-22. September 2000 - Kourou, French Guiana

Proceedings

Bernd Degen, Marilyn D. Loveless
and Antoine Kremer

editors

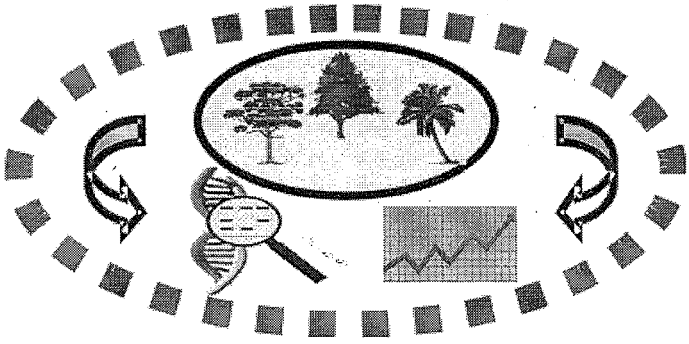
Embrapa

SilvoLab GUYANE

Ministério da Agricultura
Pecuária e Abastecimento

ISBN 85-87690-14-0

Modelling and experimental research on genetic processes in tropical and temperate forests



18-22. September 2000 - Kourou, French Guiana

Proceedings

Bernd Degen, Marilyn D. Loveless
and Antoine Kremer

editors

Embrapa

SilvoLab GUYANE

Modelling and experimental research on genetic processes in tropical and temperate forests

18-22. September 2000 - Kourou, French Guiana

Proceedings

Empresa Brasileira de Pesquisa Agropecuária
Embrapa amazônia Oriental
Ministério da Agricultura, Pecuária e Abastecimento

Modelling and experimental research on genetic processes in tropical and temperate forests

18-22. September 2000 - Kourou, French Guiana

Proceedings

Bernd Degen, Marilyn D. Loveless
and Antoine Kremer
editors

Embrapa Amazônia Oriental
Belém, PA
2002

Exemplares desta publicação podem ser adquiridos na:

Embrapa Amazônia Oriental
Trav. Dr. Enéas Pinheiro, s/n
Caixa Postal, 48 CEP: 66095-100 - Belém, PA
Fone: (91) 299-4500
Fax: (91) 276-9845
E-mail: sac@cpatu.embrapa.br

1ª edição
1ª impressão (2002): 1000 exemplares

Dados Internacionais para Catalogação da Publicação (CIP)
do Departamento Nacional do Livro

Modelling and experimental research on genetic processes in tropical and temperate forests (2000 : Kourou, French Guiana).

Proceedings, 18 a 22 setember 2000 / Editors Bernd Degen, Marilyn D. Loveless, Antoine Kremer. – Belém, PA: Embrapa amazônia Oriental, 2002.

200p.

ISBN 85-87690-14-0

1. Ecologia das florestas tropicais. I. Degen, Bernd. II. Loveless, Marilyn D. III. Kremer, Antoine. IV. Brasil, Ministério da Agricultura, Pecuária e Abastecimento Empresa Brasileira de Agropecuária..

CDD: 634.9

Edited by

Bernd Degen

UMR - CIRAD INRA ENGREF, Campus agronomique, BP 709, 97387 - Kourou cedex, French Guiana

Marilyn D. Loveless

Department of Biology, The College of Wooster, Wooster, Ohio 44691, USA

and

Antoine Kremer

INRA, Laboratoire de génétique et amélioration des arbres forestiers BP 45, 33611 Gazinet Cedex, France

Foreword

Embrapa's mission is to provide technological solutions for the sustainable development of agro-business by generating, adapting, and transferring knowledge and technologies for the good of society. It is part of this mission to develop international partnerships and networks.

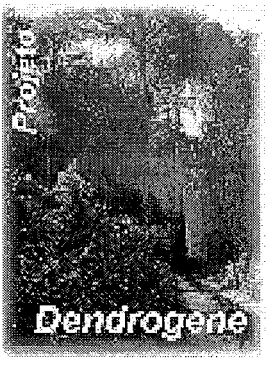
The Dendrogene project, hosted at the Embrapa Eastern Amazon research station in Belém, has created such an important scientific network in the fields of ecology, genetics, botany and modelling in order to secure the use and conservation of the tropical humid forests (see contribution of Kanashiro et al. in this issue).

Silvolab is a joint venture that coordinates scientific projects of the French institutions involved in forest research in French Guiana. The aims are both to increase our basic knowledge of the tropical rainforest ecosystem functioning, and to contribute to best practices for sustainable management. Through Silvolab activities, scientists of INRA and CIRAD in French Guiana contribute to the Dendrogene network with expertise in the fields of simulation modelling, ecology and forest genetics.

The Dendrogene project supported the Symposium "Modelling and experimental research on genetic processes in tropical and temperate forests" since this issue is of central interest to the project. Therefore the publication of the proceedings is an important contribution of Embrapa to make the topics discussed at the Symposium available to the public.

Emanuel Adilson de
Sousa Serrão
(General Director Embrapa Eastern Amazon)

Meriem Fournier
(President of Silvolab Guyane, Director of INRA Research Unit)



Introduction

Since the early fifties, forest geneticists have established provenance tests and accumulated data on population differentiation for phenotypic traits in tree populations from around the globe. This information was later complemented by numerous surveys of genetic diversity using gene markers, principally morphological traits and isozymes. A large body of experimental evidence documenting genetic organization has been accumulated for temperate trees. Over the past 20 years, similar efforts have been extended to populations of woody species in the tropics. However, the interactions between basic mechanisms that shape genetic diversity still need to be elucidated.

A central question is: what is the relative contribution of evolutionary history, genetic processes, and demographic phenomena to diversity and genetic structure in natural populations? In classical population genetics, these issues have been investigated through theoretical work using simplified scenarios. More recently, new techniques and methods have been developed which are now contributing in novel and important ways to our understanding of the evolution of genetic diversity in plant populations. First, computer simulations have been developed and are now routinely used to analyse the dynamics of genetic structure and to test hypothesis about the impact and function of specific processes. Second, new genetic markers have become available that have opened doors and made it possible to address new and complicated questions within populations. Chloroplast DNA polymorphisms offer a tool for tracking fruit dispersal and for investigating the continuity of maternal lineages in tree populations. Microsatellite markers have provided the high levels of polymorphism we need to reconstruct, in detail, mating patterns within a study area. And new techniques and tools are continuing to be developed which will offer us even more powerful methods for understanding and for documenting the ways in which elements of evolutionary history and ecological processes interact to structure natural populations.

We intended during this meeting to bring together investigators using all of these different approaches, and to provoke thoughtful discussions between practitioners about the basic mechanisms which contribute to genetic diversity. Comparisons were made between theoretical expectations, computer simulations and experimental results. Our hope was to create a fruitful setting for exchange of knowledge and for animated discussion, in order to begin to fill the gaps between theoretical and experimental methods. We hoped to initiate more cooperation between theoretical, modelling approaches and experimental studies. The two approaches are intimately linked, since empirical data are necessary to calibrate and validate simulation results in the same way as theoretical and modelling approaches can be helpful in analysing, understanding, and predicting population genetic processes.

We organized the meeting in the tropics for several reasons. Most importantly, we wanted to provide a venue to gather together representatives from the different groups currently working on forest genetics in the tropics using a diversity of approaches. This meeting was intended in part to provide an overview of current research efforts on population genetics in tropical forests and to reinforce cooperation between these groups. We also wished to compare methodologies used in temperate and tropical regions, and to examine results from these studies in a comparative framework. Finally, we wanted to illustrate to scientists and researchers from temperate countries the unique challenges that the tropics offer to forest genetics.

We received substantial help from several institutions in France, Brazil and The United Kingdom in organizing this meeting. We would like to thank CIRAD (Centre de Coopération Internationale en Recherche Agronomique pour le Développement, France), ECOFOR (Le Groupement d'Intérêt Public (GIP) Ecosystèmes forestiers, France), EMBRAPA-DFID (Empresa Brasileira de Pesquisa Agropecuaria, Brazil and Department for International Development, UK), INRA (Institut National de la Recherche Agronomique, France) and SILVOLAB Guyane for their important financial contributions to organizing and conducting this meeting. We hope that the discussions that took place during the symposium and the papers presented in this volume will contribute to ongoing efforts to better understand the dynamics of genetic structure in natural populations of plants in both temperate and tropical habitats.

Kourou, French Guiana, 26 September 2002

The editors

Contents

PART 2: MODELLING APPROACHES

PART 3: GENE FLOW AND POPULATION DIFFERENTIATION

Contributors

ALIA, R.	Departamento de Mejora Genética y Biotecnología, INIA, P.O. 8111, 28.080 Madrid, Spain
AUSTERLITZ, F.	INRA, Laboratoire de génétique et amélioration des arbres forestiers BP 45, 33611 Gazinet Cedex, France
BODÉNÈS, C.	INRA, Laboratoire de génétique et amélioration des arbres forestiers BP 45, 33611 Gazinet Cedex, France
CAVERS, S.	Centre for Ecology and Hydrology, Edinburgh Station, Bush Estate, Penicuik, EH26 0QB, UK
CERVERA, M. T.	Departamento de Mejora Genética y Biotecnología, INIA, P.O. 8111, 28.080 Madrid, Spain
DEGEN, B.	UMR CIRAD INRA ENGREF, Campus agronomique, BP 709, 97387 - Kourou cedex, French Guiana
DICK, C. W.	Smithsonian Tropical Research Institute, Unit 0948, APO AA 34002, USA
DUTECH, C.	Laboratoire de Génétique et d'Ecologie Moléculaire – Silvolab, CIRAD-Forêt, BP 701, 97387 Kourou Cedex, French Guiana
DYER, R. J.	Department of Biology, University of Missouri - St. Louis, St. Louis, Missouri 63121-4999, USA
GALLO, L.	Forest Genetics Unit, INTA Bariloche, C.C. 277, 8400 Bariloche, Argentina

GERBER, S.	INRA, Laboratoire de génétique et amélioration des arbres forestiers BP 45, 33611 Gazinet Cedex, France
GIL, L.	Unidad de Anatomía, Fisiología y Genética, ETSIM, Ciudad Universitaria s/n, 28040 Madrid, Spain
GODELLE, B.	Laboratoire Evolution et Systématique, Université Paris-Sud, F-91405 Orsay, France
GONZALEZ-MARTINEZ, S. C.	Departamento de Mejora Genética y Biotecnología, INIA, P.O. 8111, 28.080 Madrid, Spain
GOUYON, P.-H.	Laboratoire Evolution et Systématique, Université Paris-Sud, F-91405 Orsay, France
GREGORIUS, H.-R.	Institut für Forstgenetik und Forstpflanzenzüchtung, Universität Göttingen, Büsgenweg 2, 37077 Göttingen, Germany
HAGGAR, J.	International Center for Research in Agroforestry, Chetumal, Mexico
JARNE, P.	CNRS/CEFE, 1919 route de Mende, 34000 Montpellier, France
JOLY, H.	CIRAD-Forêt, Campus international de Baillarguet, 34000 Montpellier Cedex 1, France
KANASHIRO, M.	Embrapa Amazonia Oriental, C.P. 48, 66095-100 Belém-PA, Brazil
KREMER, A.	INRA, Laboratoire de génétique et amélioration des arbres forestiers BP 45, 33611 Gazinet Cedex, France

LE CORRE, V.	INRA, Unité de Malherbologie et Agronomie, B.P. 86510, 21065 Dijon Cedex, France
LOVELESS, M. D.	Department of Biology, The College of Wooster, Wooster, Ohio 44691, USA
LOWE, A.	Centre for Ecology and Hydrology, Edinburgh Station, Bush Estate, Penicuik, EH26 0QB,UK
MACHON, N.	Conservatoire Botanique du Bassin Parisien, Muséum National d'Histoire Naturelle, F-75005 Paris, France
MAGGIA, L.	Laboratoire de Génétique et d'Ecologie Moléculaire – Silvolab, CIRAD-Forêt, BP 701, 97387 Kourou Cedex, French Guiana
MARCHELLI, P.	Unidad de Genética Forestal, INTA EEA Bariloche, CC 277, 8400 Bariloche, Río Negro, Argentina
MARIETTE, S.	INRA, Laboratoire de génétique et amélioration des arbres forestiers BP 45, 33611 Gazinet Cedex, France
MARTINEZ-ZAPATER, J. M.	Departamento de Mejora Genética y Biotecnología, INIA, P.O. 8111, 28.080 Madrid, Spain
RODRIGUEZ SANTIAGO, B.	Instituto Nacional de Investigaciones Forestales y Agropecuarias A. Postal 182 Chetumal, 77000, Mexico

ROUBIK, D. W.	Smithsonian Tropical Research Institute, Apdo. 2072, Balboa, Panama and Unit 0948, APO AA 34002-0948, USA
SORK, V. L.	Department of Biology, University of Missouri - St. Louis, St. Louis, Missouri 63121-4999, USA
STREIFF, R.	INRA - URLB, Laboratoire de Modélisation et de Biologie Evolutive 488 rue de la Croix-Lavit, 34000 Montpellier, France
THOMPSON, I. S.	United Kingdom Department for International Development (DFID), C.P. 48, 66095-100 Belém-PA, Brazil
WARD, S. E.	International Institute of Tropical Forestry, PO Box 25,000, San Juan, Puerto Rico
WIGHTMAN, K.	International Center for Research in Agroforestry, Chetumal, Mexico

Part 1

Reviews and visions

Genetic diversity and differentiation in tropical trees

M. D. LOVELESS

Address: Department of Biology, The College of Wooster, Wooster, Ohio 44691 USA

Key words: genetic diversity, genetic structure, genetic variation, population differentiation, tropical trees.

Abstract: In this paper, I review the literature measuring levels of genetic diversity and patterns of between-population differentiation in tropical tree species. I summarize information on mating systems and genetic evidence for pollen and seed dispersal distances in tropical trees, and suggest areas where more study is needed. I also discuss the ecological factors which appear to influence these genetic patterns, in both isozyme and in DNA genetic markers. Finally, I consider the implications of our current understanding of tropical tree genetics for conservation and management of these genetic resources.

Introduction

Understanding the spatial pattern of genetic diversity in any species is an exercise in history as well as in ecology. The amount and the distribution of genetic variation within and between populations is a function of biogeography, isolation, gene flow, changes in population size, habitat distribution, and patterns of selection that have given rise to the genetic structure we measure and try to interpret in modern populations.

The genetic makeup of a species is closely correlated with its ecological characteristics (Hamrick et al. 1979; Loveless and Hamrick 1984; Hamrick and Godt 1990; Loveless 1992). This suggests that ecological processes, perhaps more or less stable over periods of time, are a major determinant of contemporary genetic structure (Loveless and Hamrick 1984). Such a link between ecological processes (survivorship, reproduction, and dispersal) and genetic structure is predicted from population genetic theory, in which population size and patterns of mating determine the genetics of each new generation (Lande and Barrowclough 1987; Barrett and Kohn 1991; Ellstrand and Elam 1993). At the same time, the history of any species will consist of events that have altered the numbers of individuals, their densities, mating patterns, and habitat distributions. Thus, population genetic structure will always include an elusive, historical constraint on its modern makeup.

Population genetic studies of plant species have, over the past 50 years, provided a complex and highly-textured picture of how plants sequester and partition genetic diversity. But until the mid-1980's we had virtually no information on how these genetic patterns might be manifested in tropical habitats, where species diversity reigns supreme. A large number of recent studies have begun to correct this gap. Although our understanding of tropical plant population genetics, on a global scale, is still in its infancy, patterns are beginning to emerge that provide a satisfying conceptual link between the ecology of tropical trees and their genetic makeup on a regional and a local scale.

Characteristics of tropical trees

Tropical woody species, as a group, are characterized by a variety of traits which will influence their genetic makeup (Loveless 1992; Hamrick et al. 1991). Most tropical forests exhibit very high levels of species diversity in comparison with temperate forests (Gentry 1988; 1990). The corollary of this is that individuals of any single species occur at low densities, often much less than 1 individual ha^{-1}, with large inter-individual distances among adult plants (Ashton and Bawa 1991). While dispersion is probably random for some species, others show distinct clumping (Condit et al. 1992; Gullison et al. 1996; Caron et al. 1998, Degen et al. 2001). Thus, neighborhood areas are likely to be large, and neighborhood size (N_e) will be strongly influenced by pollen and seed dispersal (Loveless and Hamrick 1984).

Pollination in tropical trees is also different, on average, than in temperate species, where wind pollination is common. Virtually all tropical tree species are woody angiosperms, mostly dicotyledons, and the vast majority of these are pollinated by animal vectors, especially by insects (Bawa et al. 1985; Prance 1985; Bawa 1990). Since populations of animal vectors are regulated by their own ecological and behavioral processes, and since most tropical woody species are perennial, pollen movement between individual trees in a tropical forest within and among years is complex. In addition, the flowering phenology of tropical trees shows more diversity than that of temperate species, and there is considerable inter-individual and inter-annual variation in flowering even within a single species (Gentry 1974; Newstrom et al. 1994). A large proportion of tropical trees, at least in some forests, have also been shown to be dioecious (Bawa and Opler 1975), further complicating pollen flow. Such variations in reproductive behavior are likely to affect both neighborhood area and neighborhood size for tropical trees.

Seed dispersal vectors of tropical trees are also much more likely to be animals than is the case for temperate tree species (Estrada and Fleming 1986; Howe 1990; Levey et al. 1994). Fruit and seed vectors are usually vertebrates,

principally birds, bats, arboreal mammals, and terrestrial seed dispersers and seed predators. In seasonal tropical forests, however, wind dispersal may be common (Augspurger 1986). In addition, the patterns of seed dispersal may affect survivorship, thus influencing spatial and genetic pattern (Hamrick et al. 1996). In either case, the physical structure and the species diversity of the tropical forest is likely to make seed dispersal by these vectors different in its outcome than in temperate forests.

The more elusive part of the genetics of tropical trees is their history. It seems clear that tropical forests globally have been dynamic habitats, both on the scale of thousands and millions of years (Vuillemier 1971; Bush 1994). In some instances, genetic structure and population isolation may have been influenced by processes of continental drift, isthmus-building, and orogenies. In others, it is likely that the dynamics of Pleistocene and Holocene environments have caused population to shrink, expand, and move about in ways that are not well-understood. While there are a few studies (Aide and Rivera 1998; Caron et al. 2000; Cavers and Lowe this volume) which attempt to include the impacts of such processes in their analyses, at this point we can only be sure that, underlying the patterns we see, there is a historical signature that we cannot yet read clearly (Utelli et al. 1999). The application of molecular methods, with their higher levels of fine-scale genetic diversity, to tropical tree populations may, however, begin to make even this element of population genetics interpretable in the future (Newton et al. 1999; Bermingham and Dick 2001).

Finally, our understanding of taxonomy and species ranges in tropical habitats continues to present a challenge to genetic studies. Some recent evidence suggests that, while many tropical tree species are uncommon on a local scale, they may be widely-distributed on a regional scale (Pitman et al. 2000). The biogeographical implications of this finding are important, and suggest that tropical taxa have either reoccupied large areas of habitat since hypothesized Pleistocene climatic changes, or that they persisted in more than one refugial site within or near their current ranges. This evidence thus suggests that many tropical tree species have a long history of broad distributions throughout varied environments. On the other hand, thorough local taxonomic studies of particular species or genera often identify as-yet undescribed taxa (M. Hopkins, pers comm). Extensive systematic studies are needed, to verify that widely separated populations are, indeed, truly conspecific, and not sibling species. Genetic information can be one source of information from which to infer such complex species histories, which are crucial to an understanding of the genetic makeup of tropical tree taxa.

As a result of these ecological and historical characteristics, tropical trees present a variety of challenges to the ecological geneticist. Speculations on the genetics of tropical tree populations by earlier authors (Corner 1954; Federov

1966) suggested that tropical trees might be highly inbred and genetically depauperate, due to their low density, hyperdispersion, phenological asynchrony, and the fragility of their small insect pollinators. Some understory species may, indeed, manifest such patterns, as was suggested for Bornean palms in the genus *Pinanga* (Shapcott 1999). But current evidence for most tropical canopy species suggests the opposite pattern. In these canopy and emergent species, outbreeding is common, some pollinators apparently move long distances, gene flow may be extensive, seeds are largely or completely outcrossed, and tropical tree populations of many species seem to be large and genetically connected. Although our understanding of genetic diversity and differentiation in tropical tree species is far from mature, recent and ongoing studies have begun to suggest a framework by which the genetic history of these species can be appreciated and evaluated.

Ecology and genetic structure

There are two basic approaches to understanding and interpreting genetic architecture in natural populations. In one approach, we can look to population genetic theory to suggest factors which are likely to have been operating within the history of a particular plant population, and which have produced the patterns we can document. At the species level, we expect high levels of genetic variability where a taxon has a long evolutionary history and a broad ecological amplitude. In this scenario, populations are widely distributed across a selectively-variable landscape, such that different alleles and multilocus genotypes are favored in different parts of the range. Historically, large population sizes maintain a high mutational input of new variation into the species (Lande 1985; Sherwin and Moritz 2000). Where gene flow between populations is high, and where populations lack geographic and genetic isolation, variation will be widely distributed among populations. In contrast, where populations are isolated, genetic differentiation will be high (Slatkin 1985). Where those isolated populations are also small, increased homozygosity due to inbreeding and loss of alleles due to genetic drift is expected to lead to low overall levels of genetic variation at the species level, along with high levels of population differentiation (Barrett and Kohn 1990; Ellstrand and Elam 1993.) If local populations in restricted aggregations or habitats have been related historically by metapopulation dynamics, then fragmentation and loss of populations may affect rates of colonization and extinction, reducing the sources from which colonists are drawn, and leading to further alterations in regional structure (Young et al. 1996).

An alternative method of identifying processes which are important in population genetic structure is to examine existing empirical studies of genetic variation and structure in plant species, and ask what identifiable factors are correlated with particular genetic patterns. Reviews of the genetic literature have identified attributes which explain significant proportions of genetic variation within species and within populations, and which are related to how variation is partitioned within and among populations of plants (Hamrick et al. 1979; Loveless and Hamrick 1984; Hamrick and Godt 1990; Hamrick et al. 1991). In these reviews, geographic range was the single best predictor of levels of within species and within population genetic variation (Table 1), while levels of potential gene flow were most strongly associated with population differentiation (Hamrick and Godt, 1990). To the degree that tropical tree species share a suite of common ecological traits, this could suggest generalizations about patterns of genetic structure within this group of plants which could lead to a more synthetic understanding of the ways in which different tropical species will respond to changes in their habitat, especially due to clearing, fragmentation, logging, and other processes of forest conversion and loss.

Patterns of genetic variation

The first generalization which has emerged from genetic studies of tropical trees is that these species maintain levels of genetic diversity at least equal to that of temperate plant species. The image of the tropical rainforest as consisting of small, isolated, ancient, inbred populations of locally-evolved species is resoundingly contradicted by emerging genetic data. In a review of allozyme studies for tropical taxa, Loveless (1992) found that native woody tropical species were polymorphic (P) at, on average, 39.0 (s.e. = $\pm$2.7)% of loci, and maintained an average heterozygosity (He) of 0.129 ($\pm$0.011).
These values are quite comparable to levels of P (34.2%) and He (0.113) in plant species overall (Hamrick and Godt 1990), but slightly less than those authors found in long-lived woody perennials (P = 50.0 % , He = 0.149), the category most ecologically comparable to tropical trees (Hamrick and Godt 1990). More recent studies have also found levels of variation in tropical trees similar to those in other plant groups (Table 2). While levels of variability reported in the literature are influenced by methodology and by the scale over which individuals and populations are sampled, genetic variation in tropical trees is widespread. This is so whether the genetic markers used are allozymes, RAPDs, AFLPs, or microsatellites. Thus, it seems clear that, whatever their history and ecology, populations of tropical trees have generated and retained substantial levels of

genetic variation. By no means are most tropical tree species "genetic fossils" which have been subjected to long-term genetic decay by drift processes operating in small, isolated populations.

Table 1. Ecological and distributional traits which account for significant differences in genetic variables among plant groups. Values are given with standard errors in parentheses. Species H_e, genetic diversity at the species level. Population H_e, genetic diversity at the population level. G_{st}, proportion of total diversity among populations. Asterisks indicate level of significance among categories within each trait. *, $p<0.05$; **, $p<0.01$, ***, $p<0.001$, NS, Not Significant. Data from Hamrick and Godt (1990), based on a survey of 653 studies of 449 plant species.

Species Characteristics	**Species H_e**	**Population H_e**	**G_{ST}**
GEOGRAPHIC RANGE	***	***	NS
Endemic	0.096 (0.010)	0.063 (0.006)	0.248 (0.037)
Narrow	0.137 (0.011)	0.105 (0.009)	0.242 (0.024)
Regional	0.150 (0.008)	0.118 (0.007)	0.216 (0.019)
Widespread	0.202 (0.015)	0.159 (0.013)	0.210 (0.025)
LIFE FORM	***	***	***
Annual	0.161 (0.009)	0.105 (0.008)	0.357 (0.024)
Short-lived Herbaceous	0.116 (0.009)	0.096 (0.008)	0.233 (0.019)
Short-lived Woody	0.097 (0.020)	0.094 (0.021)	0.088 (0.024)
Long-lived Herbaceous	0.205 (0.084)	0.084 (0.028)	0.213 (0.144)
Long-lived Woody	0.177 (0.010)	0.149 (0.009)	0.076 (0.010)
BREEDING SYSTEM	**	***	***
Selfing	0.124 (0.011)	0.074 (0.010)	0.510 (0.035)
Mixed Mating			
Animal-pollinated	0.120 (0.015)	0.090 (0.010)	0.216 (0.024)
Wind-pollinated	0.194 (0.038)	0.198 (0.041)	0.100 (0.022)
Outcrossing			
Animal-pollinated	0.167 (0.010)	0.124 (0.008)	0.197 (0.017)
Wind-pollinated	0.162 (0.009)	0.124 (0.009)	0.099 (0.012)
SEED DISPERSAL	***	**	**
Gravity-dispersed	0.136 (0.008)	0.101 (0.007)	0.277 (0.021)
Animal-attached	0.204 (0.019)	0.137 (0.012)	0.257 (0.032)
Explosive	0.092 (0.017)	0.062 (0.011)	0.243 (0.048)
Animal-ingested	0.176 (0.019)	0.129 (0.015)	0.223 (0.033)
Wind-dispersed	0.144 (0.010)	0.123 (0.010)	0.143 (0.020)

Table 2. Comparisons of genetic parameters for diversity and differentiation in various plant groups. N, number of species sampled; P, proportion of loci polymorphic; He, expected heterozygosity or gene diversity; Gst, proportion of the total variation which is between populations. Barro Colorado Island is a moist lowland tropical forest in Panama.

Category	N	P	H_e	G_{st}	Source
Long-lived woody perennials	115	0.50	0.149	-	Hamrick and Godt (1990)
Native woody tropical species	81	0.39	0.109	0.109	Loveless (1992)
Common trees on Barro Colorado Is	16	0.61	0.211	0.055	Hamrick and Loveless (1989)
Rare trees on Barro Colorado Is	16	0.42	0.142	-	Hamrick and Murawski (1991)

Hamrick and Godt (1990) emphasized the strong correlation between widespread geographic range and high levels of genetic variation. Although geographic ranges of tropical species are well-described only for a few charismatic species (such as mahogany, *Swietenia macrophylla*, Lamb 1959), the traditional view has been that most tropical species are quite local in their range. However, it may be that this generalization has arisen largely from the lack of widespread collections in many tropical localities. In an attempt to assess rarity in tropical woody taxa, Pitman et al. (1999) sampled tree species within 21 plots in the Manu National Park (Peru). These authors found that, for taxa which are well-characterized botanically, the great majority have widespread distributions, 69% (of 508 spp) being found 1500 km or more away, in Amazonian Ecuador. In addition, perhaps as few as 15% of the species were habitat specialists; the other 85% were found in more than one forest type within the sampling area. If these generalizations hold true for tropical species as a whole, they suggest that, consistent with population genetics theory, we might expect tropical trees to exhibit high levels of genetic variation at the species level (Table 3).

Table 3. Species-level measures of genetic variation for selected species of tropical trees using isozyme techniques. N, number of populations surveyed; P, proportion of loci polymorphic; H, species-level gene diversity. Superscripts indicate the geographic range of the species; n, neotropical; o, Oriental (Indian, Asian, or Australasian.) [1] values averaged only among natural populations.

Species	N	P	H	Source
Swietenia macrophylla[n]	2	0.667	0.229	Loveless and Gullison (2002)
Cordia alliodora[n]	11	0.442	0.127	Chase et al. (1995)
Carapa guianensis[n]	9	0.350	0.120	Hall et al. (1994)
Hopea odorata [o 1]	4	0.422	0.195	Wickneswari et al. (1994)
Pinus caribaea caribaea[n]	5	0.600	0.272	Zheng and Ennos (1999)
Tachigali versicolor[n]	6	0.296	0.073	Loveless et al. (1998)
Pentaclethra macroloba[n]	18	0.357	0.074	Hall et al. (1994)
Shorea leprosula[o]	8	0.889	0.369	Lee et al. (2000)
Syzygium nervosum[o]	21	0.650	0.307	Shapcott (1998c)
Pterocarpus macrophylla[o]	11	0.823	0.246	Liengisiri et al. (1995)
Tropical native species	81	0.390	0.109	Loveless (1992)

At the same time, Pitman et al. (1999) concur that the vast majority of species in tropical forests occur in very low densities throughout much of their range. The spatial scale at which genetic sampling is done is thus likely to be an important factor in measured levels of genetic variation and differentiation found in tropical genetic studies. Given the fact that species ranges of tropical taxa are poorly known, currently available estimates of genetic diversity may underestimate actual species-level variation. Such an effect of sampling scale seems to be shown in Tables 4 and 5, in which isozyme studies have been grouped according to the distances between sampled populations. It is thus important to note that our lack of understanding of the ranges of even common tropical tree taxa is a serious impediment to understanding their likely genetic structure. Biosystematics and biogeography are areas of research which are neglected, in the modern context, but which are critical to any appreciation of genetic variability in tropical trees, as well as to an understanding of tropical biotic resources in general.

Table 4. Measures of genetic variation and population differentiation from various studies of palms (Palmae). All studies are from SE Asia except that of Eguiarte et al. (1992), from Mexico. P, proportion of loci polymorphic; He, expected heterozygosity or gene diversity; Fst, population subdivision, or proportion of total variation contained between populations.

Species	P	He	Fst	Source
Pinanga aristata	0.83	0.379	0.210	Shapcott (1999)
Pinanga brevipes	0.69	0.256	0.261	Shapcott (1999)
Pinanga dumetosa	0.91	0.294	0.099	Shapcott (1999)
Pinanga tenella	0.47	0.133	0.148	Shapcott (1999)
Pinanga veitchii	0.80	0.352	0.194	Shapcott (1999)
Carpentaria aculminata	0.41	0.143	0.379	Shapcott (1998a)
Ptychosperma bleeserii	0 027	0.006	—	Shapcott (1998b)
Archontophoenix cunningmanii	0.111	0.040	0.484	Playford in Shapcott (1999)
Calamus subinermis	0 824	0.360	—	Bon et al. (1999)
Carpoxylon macrospermum	comparable to *P. bleeserii*			Dow et al. (1997)
Astrocaryum mexicanum	0.318	0.153	0.041	Eguiarte et al. (1992)

The reality seems to be that, as we expect of other plant groups, tropical trees also show different patterns of variation, depending on the details of their ecology and history. While some factors may be of particular importance, especially levels and distances of gene flow, species with different ranges and distributions will be genetically different. Highly isolated populations are expected to show lower levels of genetic variation, especially if they have likely been subject to historical reductions of population size (Wickneswari and Norwati 1993). Widespread species are more likely to be genetically variable. If we ignore historical or distributional factors, then we may be able to make additional predictions based on ecological factors. However, the unique background of each species is likely to be critical in the measured levels of variation we see today.

Table 5. Measures of population differentiation among populations of tropical trees. Populations of species shown in this table were censused on a spatial scale of 10 km distance apart or less. N, number of populations surveyed; H, species-level gene diversity; G_{st}, proportion of total variation partitioned between different populations. Superscripts indicate the geographic range of the species; n, neotropical; o, Oriental (Indian, Asian, or Australasian.)

Species	N	H	G_{st}	Source
Alseis blackiana^{n}	6	0.340	0.048	Loveless and Hamrick (1987)
Astrocaryum mexicanum^{n}	4	0.153	0.040	Eguiarte et al. (1992)
Ocotea tenera^{n}	6	0.225	0.128	Gibson and Wheelwright (1995)
Tachigali versicolor^{n}	6	0.073	0.069	Loveless et al. (1998)
Brosimum alicastrum^{n}	6	0.225	0.055	Loveless and Hamrick (1987)
Hevea brasiliensis^{n}	2	0.307	0.0003	De Paiva et al. (1994)
Pentaclethra macroloba^{n}	7	0.074	0.038	Hall et al. (1994)
Shorea leprosula^{o}	7	0.368	0.085	Lee et al. (2000)
Stemonoporus oblongifolia^{o}	4	0.282	0.162	Murawski and Bawa (1994)

One way to tease apart these variables is to look at groups of closely related species who share a similar taxonomic designation, and which occur in a common area. Few studies of this sort have been done on tropical species, but Shapcott's work (1999) on understory palms in the genus *Pinanga* in Borneo, using isozymes, is a useful example. Five different species of *Pinanga* coexist within a single region, and thus share not only a taxonomic lineage but probably historical and biogeographical processes, as well. They share a similar floral morphology, and all are monoecious, with similar pollinators. Within the forest, however, the species vary widely in their population densities. They also differ in fruit size and in vegetative growth patterns. Shapcott found a strong positive correlation between nearest neighbor distance and genetic diversity (H_c), and a strong negative correlation between plant density and genetic diversity. The most common species (*P. dumetosa*) had the highest gene flow, the lowest level of population differentiation, and the highest genetic diversity. *P.* sp. aff. *brevipes,* the least abundant species, had the lowest gene diversity, the lowest gene flow (Nm), and the most differentiation among populations (highest F_{st}). It is, however, clear that taxonomy alone is not a good predictor of genetic variation, particularly in tropical trees. Comparisons among different species of the family Aracaceae (Table 6) show wide differences in levels of genetic variation, which is not surprising, given that their life histories vary from an Indonesian climbing rattan

(*Calamus subinermis*, Bon et al. 1999) to a canopy Neotropical species (*Astrocaryum mexicanum*, Eguiarte et al. 1992). Clearly, ecology has a significant and pervasive role in shaping the genetic variation present in a species.

Table 6. Measures of population differentiation among populations of tropical trees. Populations of species shown in this table were censused on a spatial scale of more than10 km distance apart . N, number of populations surveyed; H, species-level gene diversity; G_{st}, proportion of total variation partitioned between different populations. Superscripts indicate the geographic range of the species; n, neotropical; o, Oriental (Indian, Asian, or Australasian.)

Species	N	H	Gst	Source
Swietenia macrophylla^{n}	2	0.229	0.015	Loveless and Gullison, in prep
Cordia alliodora^{n}	11	0.127	0.117	Chase et al. (1995)
Carapa guianensis^{n}	9	0.120	0.046	Hall et al. (1994b)
Pinus caribaea caribaea^{n}	5	0.272	0.020	Zheng and Ennos (1999)
Pithecellobium elegans^{n}	8	0.130	0.101	Hall et al. (1996)
Pentaclethra macroloba^{n}	12	0.207	0.219	Hall et al. (1994a)
Shorea leprosula^{o}	8	0.369	0.117	Lee et al. (2000)
Syzgium nervosum^{o}	21	0.307	0.118	Shapcott (1999)
Pterocarpus macrocarpus^{o}	11	0.246	0.121	Liengisiri et al. (1995)
Acacia auriculiformis^{o}	18	0.081	0.270	Wickneswari and Norwati (1993)
Campnosperma brevipetiolata^{o}	4	0.122	0.174	Sheely and Meagher (1996)

DNA-based markers (RAPDs, AFLPs, and microsatellites) are now being used to measure levels of genetic variation in many tropical species. While these markers are usually highly variable, and are making important contributions to our understanding of gene flow in breeding biology, it is still somewhat difficult to interpret or cross-interpret these data, in terms of levels of genetic variation. DNA based markers are valuable for their high levels of variability (Luikart and England 1999), but may not be sensitive to losses of some kinds of genetic diversity (Collevatti et al. 2001). The historical depth and comparative value of isozyme studies allows us to put tropical data into a larger framework. A similar comparative framework for considering DNA-based studies, while quickly being constructed (Newton et al. 1999), is not yet in place.

Another clear point made by surveying the tables is that existing studies are almost exclusively from Neotropics. Within this region, we are gradually building up a database that includes species from a wide variety of habitats, including montane, dry, moist, and web tropical forests. There are an increasing number of studies coming from Asia and the Indian subcontinent. But we have virtually no data from African tree species, especially from the lowland forest habitats. These regions suffered extremely severe reductions in area during the Pleistocene, and thus data from Africa could be especially helpful in understanding more about the role of geological history in population genetic structure.

Patterns of population genetic differentiation

Population differentiation is measured using hierarchical statistics which evaluate the fraction of the total variation present within and among subpopulations at various levels. The standard measures are F_{st} (Wright 1951) and G_{st} (Nei 1973). Although these measures are predicated on equilibrium dynamics of the population, and thus may not be applicable in some situations, they provide a useful way to compare different species in widely different habitats. Since population differentiation by these measures is predicated on isolation and interrupted gene flow, enhanced by local differences in selective pressures (as, for example, Eguiarte et al. 1993), the scale at which populations are sampled for genetic differentiation is crucial in understanding patterns of differentiation across the landscape. Based on this assumption, the data in Tables 5 and 6 have been organized according to the spatial scale across which populations have been sampled. A cutoff of 10 km is arbitrary, and is mostly based on the studies that are available. However, it does represent a distance across which most pollinators are unlikely to fly within a single foraging event, and thus suggests a distance across which pollen or seeds might move only between different generations.

The values shown in Tables 5 and 6 suggest that population differentiation is affected by the scale of population sampling. Values of G_{st} for populations less than 10 km distant are generally lower than those for species whose populations were sampled on a larger scale. Even so, there is overlap in absolute values, and two of the three lowest G_{st} values are found among species with wider population sampling. In the case of *Pinus caribaea caribaea* this low value for population differentiation is likely to result from the fact that it, among all the species tabled, exhibits wind pollination. Low differentiation is harder to explain in *Swietenia macrophylla* , since the two measured populations were 40 km distant from each other. While White and Boshier (2001) demonstrate pollen

movement on a scale of 4.5 km in *Swietenia humilis* in Honduras, it is unlikely that contemporary pollen movement explains this lack of population differentiation in the Bolivian *S.macrophylla* populations. It is more likely a result of the complex regeneration pattern of mahogany populations across the Beni landscape (Gullison et al. 1996).

On the other hand, relatively high levels of differentiation were found in *Pentaclethra macroloba* over a much more restricted geographic area. Hall et al. (1994) surveyed differentiation at two scales: one within the La Selva reserve (1800 ha), and one at scales of up to 70 km away. Within the reserve, adult populations showed a multilocus F_{st} of 0.038, while that value increased to 0.219 at the regional scale. The authors suggest that gene flow is restricted due to fragmentation on a regional scale and the potentially small local N_e of local populations, perhaps enhanced by phenological differences among individuals within populations.

Effects of increasing spatial separation are also demonstrated in the more detailed analyses of several studies. For instance, in *Shorea leprosula* (Lee et al. 2000), the G_{st} for the seven populations on Peninsular Malaysia was 0.085; addition of a single population from Borneo increased the overal G_{st} value to 0.117. In another example from Australasia, *Acacia auriculiformi s* (Wicknesware and Norwati 1993) was sampled in populations from two peninsulas in Australia, and from an adjacent peninsula in Papau New Guinea. Overall population differentiation gave a G_{st} of 0.270 (reported in Table 6), but this differentiation was largely the result of differences between the three different geographical regions. A tree species of southern Pacific island tropical forests, *Campnosperma brevipetiolata* (Sheely and Meagher 1996) was analyzed both within islands (at 4-5 locations) and among islands separated by up to 2175 km. An analysis of population differentiation gave mean F_{st} values within islands of 0.047. Differentiation between widely distant islands, although not high, was considerably more that that between subpopulations ($F_{st} = 0.174$).

All these results are consistent with expectations from population genetics theory. In general, increasing distances between populations, and thus increasing genetic isolation, should enhance genetic differentiation. While the baseline level of genetic diversity on which this pattern is played out may differ in different species, this geographical pattern seems to hold, not unexpectedly, for tropical species as well as for temperate ones. It is important to note, however, that in examining genetic differentiation, especially for making conservation recommendations, sampling should be done over the geographic range of the species, and not restricted to local populations, as this will not give an accurate measure of species- level genetic variation.

Mating systems

One of the surprises to come from our increased understanding of tropical tree genetics is the degree to which tropical trees are outcrossed. This is a complete reversal of the initial paradigm, but has now become the default assumption for understanding the mating patterns of tropical trees. The first indications of this outcrossed breeding structure came in seminal studies by Bawa (1974) in which he documented that a large fraction of tropical trees show genetic incompatibility systems, preventing self-pollination. Subsequent experimental studies revealed that floral morphology, especially dioecy and phenology in many tropical trees would enforce or enhance between-plant pollen movement. (Bawa and Opler 1975). With the application of genetic analyses to measure mating systems (O'Malley and Bawa 1986; Loveless and Hamrick 1987), it became clear that, even where species are capable of selfing, progeny in the field are predominantly outcrossed (Loveless 1992; Murawski 1995; Nason and Hamrick 1997; Loveless and Gullison 2002). Table 7 summarizes a variety of studies on tropical trees which demonstrate these patterns. Virtually all species examined to date have outcrossing rates of more than 0.85, indicating essentially complete outcrossing. Where a mixed mating system is present, however, patterns of outcrossing may be dependent on local flowering phenology or on plant density (Murawski et al. 1990).

The largely outcrossed mating system of tropical trees indicates that many tropical tree populations have evolved in the context of regular, often long-distance mating with distant individuals (Stacy et al. 1996; Nason and Hamrick 1997). This follows from the observation that tropical tree populations are usually low in density; outcrossing would require pollen input from often distant individuals with temporal flowering overlap. In species which characteristically outcross, however, there would have been no regular historical opportunities to purge populations of the accumulation of deleterious, recessive alleles. Thus, outcrossed species might be expected to show higher levels of inbreeding depression than species with mixed mating systems, where deleterious recessives can be purged from populations (Lande and Schemske 1985; Schemske and Lande 1985; Alvarez-Buylla et al. 1996; Dudash and Fenster 2000). As long as tropical tree populations continue to persist within relatively undisturbed, spatially extensive habitats, this should present no problems. However, as tropical habitats become disturbed by selective logging or become fragmented or lost by land conversion, the historical context within which tropical trees have evolved their mating systems becomes altered, with uncertain consequences (Collevatti et al. 2001b; see below, in the discussion of gene flow). To date, there are virtually no studies which clearly measure the potential for inbreeding depression in tropical trees or consider the consequences of inbreeding depression for long-term genetic erosion in rainforest populations.

Table 7. Estimates of outcrossing rates (mating systems) for lowland tropical forest tree species. Outcrossing rates are given with standard errors, when available. Superscripts indicate the geographic range of the species; n, Neotropical; o, Oriental (Indian, Asian, or Australasian). Where more than one value is given for t, the data represent measurements in different flowering seasons.

Species	Outcrossing Rate	Source
Astrocaryum mexicanum^{n}	1.007 ± 0.053	Eguiarte et al. (1992)
Beilschmedia pendula^{n}	0.918	Hamrick and Murawski (1990)
Bertholletia excelsa^{n}	0.850 ± 0.03	O'Malley et al. (1988)
Brosimum alicastrum^{n}	0.876	Hamrick and Murawski (1990)
Calophyllum longifolium^{n}	1.030 ± 0.085	Stacy et al. (1996)
Carapa guianensis^{n}	0.967 ± 0.022	Hall et al. (1994b)
Carapa procera^{n}	0.850	Doligez and Joly (1997)
Cavanillisia platanifolia^{n}	0.570 ± 0.02	Murawski and Hamrick (1991)
Cecropia obtusifolia^{n}	0.970 ± 0.02	Alvarez-Buylla and Garay (1994)
Ceiba pentandra^{n}	0.689 ± 0.03	Murawski and Hamrick (1992)
Cedrela odorata^{n}	0.969 ± 0.024	James et al. (1998)
Cordia alliodora^{n}	0.966 ± 0.027	Boshier et al. (1995)
Enterolobium cyclocarpum^{n}	0.990 ± 0.12	Apsit et al. in prep
Jacaranda copaia^{n}	0.943 ± 0.044	James et al. (1998)
Ocotea tenera^{n}	0.926 ± 0.13	Gibson and Wheelwright (1996)
Pithecellobium elegans	0.986 ± 0.014	Hall et al. (1996)
Pithecellobium pedicilare^{n}	0.951 ± 0.02	O'Malley and Bawa (1987)
Platypodium elegans^{n}	0.921	Hamrick and Murawski (1990)
Pterocarpus indicus^{o}	0.908 ± 0.029	Finkeldey et al. (1999)
Quararibea asterolepis^{n}	1.010 ± 0.01	Hamrick and Murawski (1990)
Samanea saman^{n}	0.990 ± .002	Cascante et al. (2002)
Sorocea affinis^{n}	1.089 ± 0.045	Murawski and Hamrick (1991)
Shorea megistophylla^{o}	0.866 ± 0.058	Murawski et al. (1994)
Spondia mombin^{n}	0.989 ± 0.163	Stacy et al. (1996)
Stemmadenia donnell-smithii	0.896 ± 0.107	James et al. (1998)
Stemonoporus oblongifolius^{o}	0.844 ± 0.021	Murawski and Bawa (1994)
Swietenia macrophylla^{n}	1.038 ± 0.024	Loveless and Gullison (2002)
Tachigali versicolor^{n}	0.998 ± 0.054	Loveless et al. (1998)
Trichilia tuberculata ^{n}	1.08	Hamrick and Murawski (1990)
Turpinia occidentalis^{n}	1.006 ± 0.090	Stacy et al. (1996)

Several papers have, however, provided preliminary evidence which suggests that, under conditions which restrict pollen movement, there could be fertility or

fecundity consequences to limited pollen movement. Nason and Hamrick (1997) showed a reduction in fruit set and a reduction in germinability of fruits of *Spondius mombin* from small, isolated populations. Aldrich and Hamrick (1998) detected reduced fruit set in *Symphonia globulifera* in residual refugial populations. And Cascante et al. (2002) measured several indicators of potential inbreeding and inbreeding depression, including low pollen tube penetration into ovules and low germinability of seeds from isolated trees, in the Costa Rican dry forest tree in *Samanea (Pithecellobium) samam*. The frequency of albinism in seedlings of tropical trees grown in greenhouse settings suggests that the potential for such deleterious recessives, under some conditions, may be significant (Loveless, per obs.).

The link between the genetics and the demography of small populations is through inbreeding depression and the possible effects of genetic erosion on recruitment and population viability (Young and Clarke 2000). Thus, we badly need studies which assess the degree to which their predominantly and historically outcrossed mating system makes tropical trees susceptible to inbreeding depression, especially in species being subjected to logging in lowland forests.

Processes of gene movement

Given that population differentiation is slight in most tropical trees on the regional level, and that trees are largely outcrossed, it follows that gene flow on the landscape level is likely to be extensive. This could be a result of either gene movement by pollen, or by seed dispersal, or a combination of both. In this area of investigation, new data derived from highly variable AFLP and microsatellite markers have been especially important in documenting distances of gene movement that appear to underlie the genetic makeup of tropical tree species, at least in the Neotropics. They have also been applied to a variety of studies asking particular questions about population differentiation and phylogeny (Newton et al. 1999; see Table 8).

The comparison of nuclear and chloroplast (maternally-inherited) genomes should enable us to distinguish between pollen and seed movement in tropical trees. While choloroplast genomes have been used in a few studies (c.f. Caron et al. 2000, and Hamilton 1999), most applications of DNA variation have been in measuring pollen movement, via paternity analyses. Isozymes can be use for such studies (Stacy et al. 1996; Nason et al. 1998), but microsatellites and AFLPs (Gerber et al. this volume) provide such variable loci and such high exclusion probabilities that they are the markers of choice for such studies, at least into the future.

Table 8: DNA marker studies which examine population differentiation in a variety of tropical tree species. Results are given qualitatively. All species listed are from the Neotropics, except *Prunus africana*. See original studies for population distributions, sample sizes, and results.

Species	Method	Scale	Results	Source
Plathymenia reticulata	RAPD	regional	85% of variation within populations	Lacerda et al. (2001)
Inga thibaudiana *Protium glabrum* *Dendropanax arboreus*	RAPD	local	>96% of variation within populations	Schierenbeck et al. (1997)
Cedrela odorata	RAPD	regional	high regional, low local differentiation	Gillies et al. (1997)
Manilkara zapota	RAPD	regional?	no genetic differentiation	Heaton et al. (1999)
Prunus africana	RAPD	regional	most variation between wide-spread populations	Dawson and Powell (1999)
Caealpinia echinata	RAPD	regional	geographic subdivisions different	Cardoso et al. (1998)
Corythophora alta	cpDNA	local	strong differences between nearby populations	Hamilton (1999)
Dicorynia guianensis	cpDNA	regional	regional differences, strong local structure	Caron et al. (2000)
Calycophyllum spruceanum	AFLP	regional	most variation found within fragments	Russell et al. (1999)
Swietenia humilis	microsat	local	most variation found within fragments	White et al. (1999)
Symphonia globulifera	microsat	local	seedlings bottlenecked	Aldrich and Hamrick (1998); Aldrich et al. (1998)
Gliricidia sepium	microsat	local	high gene flow	Dawson et al. (1997)
Caryocar brasiliense	microsat	regional	evidence of population differentiation, limited gene flow	Collevatti et al. (2001a, b)
Pithecellobium elegans	microsat	local	high gene flow among isolated trees	Chase et al. (1996)

Tracking pollen movement in plant populations can be done in a variety of ways. Pollinator observations (Frankie et al. 1976; 1983) give some indication of pollinator flight distances. In a classic study, Webb and Bawa (1983) used

fluorescent dyes to track pollen in two dry-forest herbaceous and shrub species, showing that for *Malva viscus*, pollen moved on average 37.5 m, with a maximum of 225 m, between individuals. However, most indirect measures of pollen flow have been shown to seriously underestimate actual dispersal distances of haplotypes. The use of genetic markers to measure gene flow, especially using paternity analysis, has revolutionized our understanding of realized pollen movement in tropical trees. Virtually all these studies have confirmed the fact that pollinators can move long distances between widely separated flowering individuals. These data have provided a satisfying corroboration to the studies of mating systems and population differentiation, and confirm the overall idea that effective population sizes of tropical forest tree species are far larger than our initial expectations.

The conclusion from many of these studies is that pollen, especially, can be moved regularly over long distances in tropical forests, by extremely vagile pollen vectors. Not only are vertebrate pollinators, like traplining hummingbirds (Loveless et al. unpublished data) or bats (Fleming et al. 1978; Fleming and Heithaus 1981) moving long distances, but small bees (Stacy et al. 1996) and even fig wasps, less than 2 mm in length, can travel or be moved extensive distances to pollinate distant trees. What is virtually unstudied, is the method by which these pollinators locate such distant pollen sources.

A growing number of studies, using unique alleles or paternity analysis, have given us insight into the distances over which pollen can move. Loveless et al. (1998) used a unique allozyme genotype in *Tachigali versicolor* to show that pollen moved at high frequencies (20%) over distances of at least 500 m. Hamrick and Murawski (1990) demonstrated routine pollen movement of 400-500 m in *Platypodium elegans*, using paternity analysis. In a study pioneering the use of microsatellite markers, Chase et al. (1996) showed mean pollen movement of 142 m to isolated pasture individuals of *Pithecellobium elegans* in Costa Rica. An elegant study by Kaufman et al. (1998) analyzed the pollen movement among remnant natural populations of *Cecropia obtusifolia* and populations establishing in abandoned agricultural sites at different distances from the forest populations. They showed that 37% of the seeds fathered in reference populations, in the Los Tuxtlas reserve, came from pollen derived from distant "acahaul" populations (up to 40 km away). They concluded that pollen movement over the landscape was sufficient to homogenize the genetic structure of this pioneer species. Loveless et al (unpublished data) found that the mean pollen dispersal distance for a hummingbird-pollinated, understory species (*Erythrina costaricensis*) was about 90 meters.

Other studies of isolated individuals have also demonstrated long-distance pollen flow. Nason and Hamrick (1997) showed that isolated populations of *Spondias mombin* received up to 100% of their pollen from outside the local

(insular) population (a distance of 80-1000 m). Stacy et al. (1996) showed that while nearest-neighbor mating prevailed among trees in clusters, pollen movement to the spatially isolated individuals bypassed nearest neighbours and distances of several hundred meters. They showed that, in *Calophyllum longifolium*, 62% of the effective pollen moved at least 210 m, and in *Spondius mombin*, between 2.5 and 5.2% of the effective pollen moved at least 300 m. Dick (this volume) found that isolated remnant trees of *Dinizia excelsa* in pastures could receive pollen from intact forest as much as 3.2 km away. And in the current champion for long-distance pollen movement, Nason and his coworkers used paternity analysis to demonstrate that fig wasps were regularly dispersing pollen among *Ficus* species at distances of between 6 and 14 km (Nason and Hamrick 1997; Nason et al. 1996, 1998). Based on the density of *Ficus* trees on the landscape, these data suggest that the maximum estimated breeding population size for *Ficus obtusifolia* was between 339 and 574 individuals, in an area of between 47 and 79 km^2.

It would be easy to conclude, from these examples, that animal pollinators in tropical forests have adapted to the rarity of their host plant species. This would be a premature conclusion, however, given the diversity of pollinator behaviors and population structures in tropical habitats. In fact, some pollinators can be local in their movement patterns. Territorial hummingbirds, for instance, may move pollen regularly only with a local area (Linhart 1973; Stiles 1975). In *Shorea siamensis*, Ghazoul et al. (1998) showed that, when densities of individuals in a population declined, their small *Trigona* bee pollinators showed reduced movement between trees. And although bats can be extremely strong fliers and pollen dispersers (Heithaus et al. 1975), some species are territorial, and forage locally, restricting pollen movement (Gribel and Hay 1993). Collevatti et al. (2001) suggest that such local pollen movement by small, territorial bat species has enforced genetic differentiation between populations of *Caryocar brasiliense*. One important area in which research is badly needed is in documenting the capacity for pollinators to adjust to changes in spatial distribution and density in their foraging territories. We need to understand what the thresholds for pollen movement might be, and how pollinator assemblages might change with different kinds of forest disturbance or alteration. Without good data on the behavioral responses of pollinators, we will be unable to infer the effects of changes in tree densities and distributions on the genetics of the larger population. And an intriguing suggestion is the possibility that such behavioral adaptations might even now be under selection as a result of fragmentation in plant populations (Van Dyck and Matthysen 1999).

Seed dispersal is harder to quantify, and only a few studies have attempted to assess the relative contributions of pollen vs. seed dispersal to the genetic structure of tropical trees. In most cases, the genetic data have been used to

evaluate "donor" movement, without specifying the locations of the pollen and seed parents. Aldrich and Hamrick (1998) used donor analysis to identify the pairs of *Symphonia globulifera* trees likely to have contributed to seeds and seedlings within a spatially heterogeneous population. They found extensive gene movement into remnant forest localities from genotypes of pasture trees, suggesting that bats modified their foraging movements and selectively carried these fruits into small forest tracts. The result was a genetic bottleneck, since the very high fecundity of the pasture trees essentially dominated the genetics of the recruiting seedlings.

Hamilton (1999) suggests, however, that seed dispersal in tropical trees might be much more limited than pollen movement, and could thus be important in shaping their genetic structure. In his work with *Corythophora alta* (Lecythidaceae), cpDNA haplotypes were found to be clustered, suggesting that seed dispersal is limited and reflects historical areas of continuous forest. A similar result was found by Caron et al. (2000). Within stands of *Dicorynia guianensis* in French Guiana, limited seed flow was inferred from spatial aggregation of cpDNA haplotypes. In addition, haplotype frequencies at a regional scale reflected the likely dynamics of Pleistocene refugia, and subsequent forest expansions.

Seed dispersal is even more challenging to measure than pollen movement. However, the expanded application of hypervariable DNA markers to tropical trees should permit us to use parent pair exclusion, to infer both the gamete sources of individual seedlings and samplings. The use of maternal endocarp tissue for identifying the mothers of dispersed seeds also has promise in understanding the role of dispersal in genetic structure (Godoy and Jordano 2001). This will allow us to measure the lifetime gene movement in various species, and to place that vagility (or lack of it) into a larger conservation context.

Conclusion

The data we have so far are largely drawn from studies in the neotropics, and thus may over-represent the phylogenetic and evolutionary patterns typical of Central and South American species. Nonetheless, emerging studies, especially from SE Asia, confirm several aspects of tropical tree genetics as being more than local phenomena. Species in the Old World tropics are also highly genetically variable, and that variation is distributed broadly among populations, at least at the local level. Outcrossing is high, although we have detailed paternity studies for only one Old World tropical species. Konuma et al. (2000) used paternity inference to measure pollen flow in Neobalanocarpus heimii, a Malaysian dipterocarp. They found mean pollen flow of 191.2 m (±104.9m SD), and a

maximum measured pollination event of 663.6 m. Thus, the ecological similarities between tropical forests in different biogeographic regions seem to have shaped a largely similar genetic structure. Unique historical factors or distributional patterns will still be likely to leave their marks on the genetics of particular species. But in general, tropical species studied to date exhibit a suite of genetic patterns regardless of the continent from which they come. Outcrossing is a recurrent theme, although at varying spatial levels, and tropical pollinators are adept at moving pollen long distances across highly diverse forest landscapes.

Can we then generalize about the genetics of tropical tree genetics? While there are definitely strong trends to be seen, especially among neotropical canopy emergent species, there are now enough studies to show that life form, dispersal mechanisms, and history will also play a formative role in shaping population genetic structure. The emerging data provide a fascinating picture of the dynamics of tropical tree populations and the evolutionary processes that have generated their diversity over geological time.

At the same time, these data are important in making decisions about preservation, conservation, and sustainable management strategies in tropical forests. The available genetic evidence suggests that populations within a local area are not strongly differentiated. Except in situations in which the local landscape is extremely heterogeneous or topograpically diverse, it seems likely that tree populations can resist genetic erosion if harvesting or forest conversion is planned and strongly regulated. The more important effects of landscape changes in tree density will be in altering ongoing ecological processes, especially pollen movement and seed dispersal, that govern transmission genetics in a post-disturbance environment. Where species are outcrossed and typically part genetic neighborhoods that cover tens or hundreds of hectares, the evolutionary history of that species has been played out on a much larger and more heterogeneous scales. While long-distance pollination has been documented among trees in intact forest and in mosaic landscapes, increased spatial isolation must, at some point, begin to alter the reproductive process. There may be a threshold beyond at which ecological processes are so altered that the genetic cohesion of a population simply collapses.

The difficulty lies in trying to imagine, not the next ten years in that population's genetic history, but the next 100, or the next 1000. In the final analysis, transformation of habitat has the potential to leave species without a forest to live in. While the genetic processes on the way to that end point may be resilient to rapid change, we do not yet know where the thresholds lie.

Thus, we need solid information about genetic diversity and differentiation, about processes of gene movement and the genetics of recruitment, in both intact forests and in forests under stress. We need to employ the powerful tools of genetic analysis to anticipate and, if possible, prevent, the genetic degradation of tropical forest trees.

Acknowledgments

The effort and organization of Bernd Degen made possible this symposium, and he deserves thanks for his insight and thoughtful planning, as well as for his patience during the preparation of this manuscript. Financial support for preparation of this paper was provided by The College of Wooster. Logistical support was supplied by CPATU/Embrapa Amazonia Oriental in Belém, Pará, Brazil, and CIRADForêt/INRA, in Kourou, French Guiana.

References

Aide, T. M., Rivera, E. (1998): Geographic patterns of genetic diversity in Poulsenia armata (Moraceae). Journal of Biogeography 25:695-705.

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

Aldrich, P. R., Hamrick, J. L., Chavarriaga, P., Kochert, G. (1998): Microsatellite analysis of demographic genetic structure in fragmented populations of the tropical trees Symphonia globulifera. Molecular Ecology 7:933-944.

Alvarez-Buylla, E. R., Garay, A. A. (1994): Population genetic structure of Cecropia obtusifolia, a tropical pioneer tree. Evolution 48:437-453.

Alvarez-Buylla, E. R., Garcia-Barrios, R. et al. (1996): Demographic and genetic models in conservation biology: Applications and perspectives for tropical rain forest tree species. Ann. Rev. Ecology Syst. 27: 387-421.

Augspurger, C. K. (1986): Morphology and dispersal potential of wind-dispersed diaspores of neotropical trees. Amer. J. Bot. 73:353-363.

Barrett, S. C. H., Kohn, J. R. (1991): Genetic and evolutionary consequences of small population size in plants: Implications for conservation. pp. 3-30 In: Falk, D. A. and Holsinger, K. E. (Eds.). Genetics and Conservation of Rare Plants. Oxford Univ. Press, Oxford., UK.

Bawa, K. S. (1990): Plant-pollinator interactions in tropical rain forest. Ann. Rev. Ecol. Syst. 21: 399-422.

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

Bawa, K. S., Opler, P. A. (1975): Dioecism in tropical forest trees. Evolution 29:167-179.

Bawa, K. S., Ashton, P. S. (1991): Conservation of rare trees in tropical rain forests: a genetic perspective. pp. 62-71 In: Falk, D. A. and Holsinger, K. E. (Eds.). Genetics and Conservation of Rare Plants. Oxford Univ. Press, Oxford.

Bawa, K. S., Perry, D. R., Beach, J. H. (1985): Reproductive biology of tropical lowland rain forest trees. I. Sexual systems and incompatibility mechanisms. Amer. J. Bot. 72:331-345.

Bermingham, E., Dick, C. (2001): The Inga – Newcomer or museum antiquity? Science 293:2214-2216.

Bon, M.-C., Joly, H. I., Alloysius, D. (1999): Genetic variation in remnant populations of a rattan species (Calamus subinermis) in Sabah.

Boshier, D. H., Chase, M. R., Bawa, K.. S. (1995): Population genetics of Cordia alliodora (Boraginaceae), a neotropical tree. 2. Mating system. Am. J. Bot 82:476-483.

Bush, M. B. (1994): Amazon speciation: a necessarily complex model. Journal of Biogeography 21:5-17.

Cardoso, M. A., Provan, J., Powell, W., Ferreira, P. C. G., de Oliveira, D. E. (1998): High genetic differentiation among remnant populations of the endangered Caesalpinia echinata Lam. (Leguminosae-Caesalpiniodeae). Molecular Ecology 7:601-608.

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

Cascante, A., Quesada, M., Lobo, J. J., Fuchs, E. A. (2002): Effects of dry tropical forest fragmentation on the reproductive success and genetic structure of the tree Samanea saman. Conservation Biology 16:137-147.

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

Chase, M. R., Boshier, D. H., Bawa, K. S. (1995): Population genetics of Cordia alliodora (Boraginaceae), a neotropical tree. 1. Genetic variation in natural populations. Am. J. Bot. 82:468-475.

Collevatti, R. G., Gratapaglia, D., Hay, J. D. (2001a): Population genetic structure of the endangered tropical tree species Caryocar brasiliense, based on variability at microsatellite loci. Molecular Ecology 10:449-356.

Collevatti, R. G., Gratapaglia,D., Hay, J. D. (2001b): High resolution microsatellite based analysis of the mating system allows the detection of significant biparental inbreeding in Caryocar brasiliense, an endangered tropical tree species. Heredity 86:60-67.

Condit, R., Hubbell, S. P., Foster, R. B. (1992): Recruitment near conspecific adults and the maintenance of tree and shrub diversity in a neotropical forest. Amer. Natur. 140:261-286.

Corner, E. J. H. (1954): The evolution of tropical forest. pp. 34-46 In: Huxley, J. S., Hardy, A. C. and Ford, E. B. (Eds.). Evolution as a Process. Allen and Unwin, London.

Dawson, I. K., Powell, W. (1999): Genetic variation in the Afromontane tree Prunus africana, an endangered medicinal species. Molecular Ecology 8:151-156.

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

Degen, B., Caron, H., Bandou, E., Maggia, L., Chevallier, M-H., Leveau, A., Kremer, A. (2001): Fine-scale spatial genetic structure of eight tropical tree species as analysed by RAPDs. Heredity 87:497-507.

Dick, C. W. (2001): Genetic rescue of remnant tropical trees by an alien pollinator. Proc.Royal Soc. London Series B-Biol. Sci. 268, 2391-2396.

Doligez, A., Joly, H. I. (1997): Mating system of Carapa procera (Meliaceae) in the French Guiana tropical forest. Am. J. Bot 84:461-470.

Dowe, J. L. , Benzie, J., Ballment, E. (1997): Ecology and genetics of Carpoxylon macrospermum H. Wendle. & Drude (Arecaceae), an endangered palm from Vanuatu. Biological Conservation 79: 205-216.

Dudash, M. R., Fenster, C. B. (2000): Inbreeding and outbreeding depression in fragmented populations. pp. 35-53 In: Young, A. G. and Clarke, G. M. (Eds.) Genetics, Demograph, and Viability of Fragmented Populations. Cambridge University Press, Cambridge.

Eguiarte, L. E., Búrquez, A., Rodrigues, J., Martinez-Ramos, M., Sarukan, J., Piñero, D. (1993): Direct and indirect estimates of neighborhood and effective population size in a tropical palm, Astrocaryum mexicanum. Evolution 47:75-87.

Eguiarte, L. E., Perez-Nasser, N., Piñero, D. (1992): Genetic structure, outcrossing rate and htereosis in Astrocaryum mexicanum (tropical palm): Implications for evolution and conservation. Heredity 69:217-228.

Ellstrand, N. C., Elam, D. R. (1993): Population genetic consequences of small population size: implications for plant conservation. Ann. Rev. Ecol. Syst. 24:217-242.

Estrada, A., Fleming, T. H. (Eds.) (1986): Friguvores and Seed Dispersal. Dr. W. Junk, Publishers, The Hague.

Federov, A. A. (1966): The structure of the tropical rain forest and speciation in the humid tropics. Jour. Ecol. 54:1-11.

Finkeldey, R., Guzman, N. D., Changtragoon, S. (1999): The mating system of Pterocarpus indicus Willd. at Mt. Makiling, Philippines. Biotropica 31:525-530.

Fleming, T. H., Heithaus, E. R. (1981): Frugivorous bats, seed shadows, and the structure of tropical forests. Biotropica 13 (Suppl.): 45-53.

Frankie, G. W., Opler, P. A., Bawa, K. S. (1976): Foraging behavior of solitary bees: implications for outcrossing of a neotropical forest tree species. Jour. Ecol. 64:1049-1957.

Frankie, G. W., Vinson, S. B., Newstrom, L. E., Barthell, J. F., Haber, W. A., Frankie, J. K. (1990): Plant phenology, pollination ecology, pollinator behavior and conservation of pollinators in Neotropical dry forest. pp 37-48 In: Bawa, K. S. and Hadley, M. (Eds.). Reproductive Ecology of Tropical Forest Plants. Man and the Biosphere Series, Vol. 7. Unesco, Paris.

Frankie, G. W., Haber, W. A., Opler, P. A., Bawa, K. S. (1983): Characteristics and organization of the large bee pollination system in the Costa Rican dry forest. pp. 411-447 In: Jones, E. C. and Little, R. J. (Eds.) Handbook of Experimental Pollination Ecology. Van Nostrand Reinhold, New York.

Gentry, A. H. (1974): Flowering phenology and diversity in tropical Bignoniaceae. Biotropica 6:64-68.

Gentry, A. H. (1988): Tree species richeness of upper Amazonian forests. Proc. Nat. Acad. Sciences (USA) 85: 156-159.

Gentry, A. H. (Ed.) (1990): Four Neotropical Rainforests. Yale University Press, New Haven.

Ghazoul, J., Liston, K. A., Boyle, T. J. B. (1998): Disturbance-induced density-dependent seed set in Shorea siamensis (Dipterocarpaceae), a tropical forest tree. J. Ecol. 86:462-473.

Gibson, J. P., Wheelwright, N. T. (1995): Genetic structure in a population of a tropical tree Ocotea tenera (Lauraceae): influence of avian seed dispersal. Oecologia 103:49-54.

Gibson, J. P., Wheelwright, N. T. (1996): Mating system dynamics of Ocotea tenera (Lauraceae), a gynodioecious tropical tree. Amer. J. Bot. 83:890-894.

Gillies, A. C. M., Cornelius, J. P., Newton, A. C., Navarro, C., Hernández, M. (1997): Genetic variation in Costa Rican populations of the tropical timber species Cedrela odorata L., assessed using RAPDs. Molecular Ecology 6:1133-1145.

Godoy, J. A., Jordano, P. (2001): Seed dispersal by animals: exact identification of source trees with endocarp DNA microsatellites. Molecular Ecology 10: 2275-2283.

Gullison, E. R., Panfil, S. N., Strouse, J. J., Hubbell, S. P. (1996): Ecology and management of mahogany (Swietenia macrophylla King) in the Chimanes Forest, Beni, Bolivia. Botanical Journal of the Linnean Society 122:9-34.

Hall, P. S., Orrell, L. C., Bawa, K. S. (1994b): Genetic diversity and mating system in a tropical tree, Carapa guianensis (Meliaceae). Am. J. Bot 81:1104-1111.

Hall, P. S., Chase, M. R., Bawa, K. S. (1994a): Low genetic variation but high population differentiation in a common tropical forest tree species. Conservation Biology 8:471-482.

Hall, P., Walker, S., Bawa, K. (1996): Effect of forest fragmentation on genetic diversity and mating system in a tropical trees, Pithecellobium elegans. Cons. Biology 10:757-768.

Hamilton, M. B. (1999): Tropical tree gene flow and seed dispersal. Nature 401:129-130.

Hamrick, J. L., Murawski, D. A. (1990): The breeding structure of tropical trees. Plant Species Biology 5:157-165.

Hamrick, J. L., Murawski, D. A. (1991): Levels of allozyme diversity in populations of uncommon Neotropical tree species. Jour. Trop. Ecol. 7:395-399.

Hamrick, J. L., Loveless, M. D. (1989): Genetic structure of tropical tree populations: associations with reproductive biology. pp. 129-149 In: Bock, J. H. and Linhart, Y. B. (Eds.). The Evolutionary Ecology of Plants. Westview Press, Boulder, CO.

Hamrick, J. L., Godt, M. J. W. (1990): Allozyme diversity in plant species. pp. 43-63 In: Brown, A. H. D., Clegg, M. T., Kahler, A. L. and Weir, B. S. (Eds.) Plant Population Genetics, Breeding, and Genetic Resources. Sinauer Associates, Inc., Sunderland, MA.

Hamrick, J. L., Murawski, D. A., Nason, J. D. (1993): The influence of seed dispersal on the genetic structure of tropical tree populations. Vegetatio 107:281-297.

Hamrick, J. L., Godt, M. J. W., Murawski, D. A., Loveless, M.. D. (1991): Correlations between species traits and allozyme diversity: implications for conservation biology. pp. 75-86 In: Falk, D. A. and Holsinger, K. E. (Eds.), Genetics and Conservation of Rare Plants. Oxford Univ. Press, Oxford, UK.

Hamrick, J. L., Linhart, Y. B., Mitton, J. B. (1979): Relationships between life history characteristics and electrophyoretically detetable genetic variation in plants. Ann. Rev. Ecol. Syst. 10:173-200.

Hastings, A., Harrison, S. (1994): Metapopulation dynamics and genetics. Ann. Rev. Ecol. Syst. 25:167-188.

Heaton, H. J., Whitkus, R., Gómez-Pompa, A. (1999): Extreme ecological and phenotypic differences in the tropical tree chicozapote (Manilkara zapota (l.) P. Royen) are not matched by genetic diversgence: a random amplified polymoprhic DNA (RAPD) analysis. Molecular Ecology 8:627-632.

Heithaus, E. R., Fleming, T. H., Opler, P. A. (1975): Foraging patterns and resource utilization in seven species of bats in a seasonal tropical forest. Ecology 56:841-854.

Howe, H. F. (1990): Seed dispersal by birds and mammals: Implications for seedling demography. pp. 191-218 In: Bawa, K. S. and Hadley, M. (Eds.) Reproductive Ecology of Tropical Forest Plants. Man and the Biosphere Series, Vol. 7. UNESCO, Paris.

James, T., Vege, S., Aldrich, P., Hamrick, J. L. (1998): Mating systems of three tropical dry forest tree species. Biotropica 30:587-594.

Kaufman, S. R., Smouse, P. E., Alvarez-Buylla, E. R. (1998): Pollen-mediated gene flow and differential male reproductive success in a tropical pioneer tree, Cecropia obtusifolia Bertol. (Moraceae): a paternity analysis. Heredity 81:164-173.

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

Lacerda, D. R., Acedo, M. D. P., Lemos Filho, J. P., Lovato, M. B. (2001): Genetic diversity and structure of natural populations of Plathymenia reticulata (Mimosoideae), a tropical tree from the Brazilian Cerrado. Molecular Ecology 10:1143-1152.

Lamb, F. B. (1966): Mahogany of Tropical America: Its Ecology and Management. University of Michigan Press, Ann Arbor..

Lande, R. (1995): Mutation and conservation. Conservation Biology 9:782-791.

Lande, R., Barrowclough, G. F. (1987): Effective population size, genetic variation, and their use in population management. pp. 87-124 In: Soule, M. L. (ed.) Viable Populations for Conservation. Cambridge University Press, Cambridge, UK.

Lande, R., Schemskie, D. W. (1985): The evolution of self-fertilization and inbreeding depression in plants. I. Genetic models. Evolution 39:24-40.

Lee, S. L., Wickneswari, R., Mahani, M. C., Zakri, A. H. (2000): Genetic diversity of a tropical tree species, Shorea leprosula Miq (Dipterocarpaceae), in Malaysia: Implications for conservation of genetic resources and tree improvement. Biotropica 32:213-224.

Levey, D. J., Moermond, T. C., Denslow, J. S. (1994): Frugivory: An overview. pp. 282-294 In: McDade, L. A, Bawa, D. S., Hespendeide, H. A. and Hartshorn, G. S. (Eds.) Las Selve: Ecology and natural history of a Neogropical rainforest. University of Chicago Press, Chicago. IL.

Linhart, Y. B. (1973): Ecological and behaviroral determinants of pollen dispersal in hummingbird pollinated Heliconia. Amer. Natur. 107:115-123.

Loveless, M. D. (1992): Isozyme variation in tropical trees: patterns of genetic organization. New Forests 6:67-94.

Loveless, M. D. and J. L. Hamrick, J. L. (1984): Ecological determinants of genetic structure in plant populations. Ann. Rev. Ecol. Syst. 15:65-95.

Loveless, M. D., Hamrick, J. L. (1987): Distribución de la variación genética en especies arboreas tropicáles. Revista de Biología Tropicál 35 (Supl.1):165-175.

Loveless, M. D., Gullison, R. E. (2002): Genetic variation in natural populations of mahogany in Bolivia. In: Lugo, A. (Ed.). Big-leaf Mahogany: Genetic Resources, Ecology and Management. Ecological Studies Series, Springer-Verlag. (in press).

Loveless, M. D., Hamrick, J. L., Foster, R. B. (1998): Population structure and mating system in Tachigali versicolor, a monocarpic neotropical tree. Heredity 81:134-143.

Luikart, G., England, P. R. (1999): Statistical analysis of microsatellite DNA data. Trends in Ecology and Evolution 14:253-256.

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

Murawski, D. A., Hamrick, J. L. (1991): Sorocea, Cavinallesia. Plant Ecol Syst. ? or J. Heredity 67 167-174

Murawski, D. A., Hamrick, J. L. (1992): Mating system and phenology of Ceiba pentandra (Bombacaceae) in Central Panama. Jour. Heredity 83:401-404.

Murawski, D. A., Bawa, K. S. (1994): Genetic structure and mating system of Stemonoporus oblongifolius (Dipterocarpaceae) in Sri Lanka. Am. J. Bot. 81:155-160.

Murawski, D. A., Hamrick, J. L., Hubbell, S. P., Foster, R. B. (1990): Mating systems of two Bombacaceous trees of a neotropical moist forest. Oecologia 82:501-506.

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

Nason, J. D., Herre, E. A., Hamrick, J. L. (1996): Paternity analysis of the breeding structure of strangler fig populations: evidence for substantial long-distance wasp dispersal. Journal of Biogeography 23:501-512.

Nason, J. D., Herre, E. A., Hamrick, J. L. (1998): The breeding structure of a tropical keystone plant resource. Nature 391:685-687.

Nei, M. (1973): Analysis of gene diversity in subdivided populations. Proc. Natl. Acad. Sci. USA 70:3321-3323.

Newstrom, L. E., Frankie, G. W., Baker, H. G. (1994): A new classification for plant phenology based on flowering patterns in lowland tropical rain forest trees at La Selva, Costa Rica. Biotropica 26:141-159.

Newton, A. C., Allnutt, T. R., Gillies, A. C. M., Lowe, A. J., Ennos, R. A. (1999): Molecular phylogeography, intraspecific variation and the conservation of tree species. Trends in Ecology and Evolution 14:140-145.

O'Malley, D. M., Bawa, K. S. (1987): Mating system of a tropical rain forest tree species. Am. J. Bot. 74:1143-1139.

O'Malley, D. M., Buckley, D. P., Prance, G. T., Bawa, K. S. (1988): Genetics of Brazil nut (Bertholletia excelsa Humb. & Bonpl.: Lecythidaceae). 2. Mating system. Theor. Appl. Genetics 76:929-932.

Pitman, N. C. A., Terborgh, J., Silman, M. R., Nuñez, P. V. (1999): Tree species distributions in an upper Amazonian forest. Ecology 80:2651-2661.

Prance, G. T. (1985): The pollination of Amazonian plants. pp. 166-191 In: Prance, G. T. and Lovejoy, T. E. (Eds). Key Environments: Amazonia. Pergamon Press, Oxford, UK.

Russell, J. R., Weber, J. C., Booth, A., Powell, W., Sotelo-Montes, C., Dawson, I. K. (1999): Genetic variation of Calycophyllum spruceanum in the Peruvian Amazon Basin, revealed by amplified fragment length polymorphism (AFLP) analysis. Molecular Ecology 8:199-204.

Schemske, D. W., Lande, R. (1985): The evolution of self-fertilization and inbreeding depression in plants. II. Empirical observations. Evolution 39:41-52.

Schierenbeck, K. A., Skupski, M., Lieberman, D., Lieberman, M. (1997): Population structure and genetic diversity in four tropical tree species in Costa Rica. Molecular Ecology 6:137-144.

Shapcott, A. (1998a): The patterns of genetic diversity in Carpentaria acuminata (Arecaceae), and rainforest history in northern Australia. Molecular Ecology 7:833-847.

Shapcott, A. (1998b): The genetics of Ptychosperma bleeseri, a rare palm from the Northern Territory, Australia. Biological Conservation 85:203-209.

Shapcott, A. (1998c): Vagile but inbred: patterns of inbreeding and the genetic structure within populations of the monsoon rainforest tree Syzygium nervosum (Myrtaceae). Jour. Tropical Ecology 14:595-614.

Shapcott, A. (1999): Comparison of the population genetics and densities of five Pinanga palm species at Kuala Belalong, Brunei. Molecular Ecology 8:1641-1654.

Sheely, D. L., Meagher, T. R. (1996): Genetic diversity in micronesian island populations of the tropical tree Campnosperma brevipetiolata (Anacardiaceae). Am. J. Bot. 83:1571-1579.

Sherwin, W. B., Moritz, C. (2000): Managing and monitoring genetic erosion. pp. 9-34 In: Young, A. G. and Clarke, G. M. (Eds.) Genetics, Demography, and Viability of Fragmented Populations. Cambridge University Press, Cambridge.

Slatkin, M. (1985): Gene flow and the geographic structure of populations. Science 236:787-792.

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

Stiles, F. G. (1975): Ecology, flowering phenology and hummingbird pollination of some Costa Rican Heliconia species. Ecology 56:285-310.

Utelli, A.-B., Roy, B. A., Baltisberger, M. (1999): History can be more important than "pollination syndrome" in determining the genetic structure of plant populations: the case of Aconitum lycoctonum (Ranunculaceae). Heredity 82:574-584.

Van Dyck, H., Matthysen, E. (1999): Habitat fragmentation and insect flight: a changing "design" in a changing landscape? Trends in Ecology and Evolution 14:172-174.

Vuillemier, B. S. (1971): Pleistocene changes in the fauna and flora of South America. Science 1973: 771-779.

Webb, C. J., Bawa, K. S. (1983): Pollen dispersal by hummingbirds and butterflies: a comparative study of two lowland tropical plants. Evolution 37: 1258-1270.

White, G. M., Boshier, D. H. (2000): Fragmentation in Central American dry forests: genetic impacts on Swietenia humilis (Meliaceae). pp. 293-311 In: Young, A. G. and Clarke, G. M. (Eds.) Genetics, Demography, and Viability of Fragmented Populations. Cambridge University Press, Cambridge.

White, G. M., Boshier, D. H., Powell, W. (1999): Genetic variation within a fragmented population of Swietenia humilis Zucc. Molecular Ecology 8:1899-1909.

Wickneswari, R., Norwati, M. (1993): Genetic diversity of natural populations of Acacia auriculiformis. Aust. J. Bot. 41:65-77.

Wickneswari, R., Ismail, Z., Lee, S. L., Muhammad, N. (1994): Genetic diversity of remanat and planted populations of Hopea odorata in Peninsular Malaysia. Proceedings International Workshop, BIO-REFOR, Kangar, Malaysia, 1994.

Wright, S. (1951): The genetical structure of populations. Ann. Eugen. 15: 232-254.

Young, A. G., Clarke, G. M. (Eds.) (2000): Genetics, Demography, and Viability of Fragmented Populations. Cambridge University Press, Cambridge. 438 pp.

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

Tropical bee colonies, pollen dispersal and reproductive gene flow in forest trees

D. W. ROUBIK

Address: Smithsonian Tropical Research Institute, Apdo. 2072, Balboa, Panama

Key words: honey bees stingless bees, Meliponini, tropical forests, pollination behaviour, secondary pollen dispersal, ramified pollination

Abstract: Modelling processes that control tropical tree population genetics require appraisal of pollination syndromes and systems. Tropical forests are dominated by large perennial bee colonies. Two to 6 colonies ha^{-1} of Meliponini (stingless bees) and Apini (honey bees) provide millions of potential visitors to every flower. I studied their pollen dispersal experimentally in equatorial forests of America, Asia and Africa. Bee colonies affect pollen dispersal and plant genetic neighbourhoods in three ways: 1) directly, as a result of foraging range and flower visitation patterns (including pollen removal with no pollination, by robbers and thieves), 2) indirectly (*spatially*), as a function of interaction (sometimes highly aggressive) with other flower visitors, and 3) indirectly (*temporally*), as a result of bee-to-bee pollen transfer in colonies, which can produce *secondary pollen dispersal*, and also from changing foraging patches between foraging trips, which produces *ramified pollination*. Examples and new results from pollen viability studies in primary forest, and experiments on bee recruitment in the canopy, help demonstrate the necessity of knowing details of pollinator behaviour and pollen biology.

Perennial bee colonies in natural forest

A frequent problem with studies of plant-pollinator systems is that they do not concern natural habitats and ignore dynamics in the canopy. At different heights, sun, wind, exposure, and flexibility in travel between sun-drenched gaps and shade may attain prominent importance in pollen dispersal by living animals. In forest, the distribution and abundance of the plants and pollinators, and in particular of central place foragers residing in bee colonies, take on critical importance. These factors determine generalized or specialized foraging activities of bees, best shown by pollen analysis (Nieuwstadtl et al. 1997; Kiew 1997; Momose et al. 1998; Roubik and Moreno 2000).

Whether in the upper forest canopy or walking on the ground, a visitor to any mainland tropical forest will soon be investigated and 'visited' by a bee from a large, perennial colony. While the bee's primary attractant to humans might be clothing color, fatty acids on the skin or other chemicals associated

with sweat, the lesson to be learned is that the highly eusocial bees (the only bees with perennial colonies) are ubiquitous. Surveys of nests in the neotropics (Hubbell and Johnson 1977; Johnson and Hubbell 1984; Roubik 1983) and in southeast Asian forests (Roubik 1993a; Nagamitsu and Inoue 1997) estimate there are 2 to 6 such colonies ha^{-1} in forest. These studies were made of non-*Apis* bees that form large perennial colonies. They comprise Meliponini, or stingless honey-making bees. The other forest component in much of the Old World tropics, and now also the New World tropics, after the introduction of honey bees from Africa, is *Apis*. Up to 5 species naturally living in a single forest visit a rather small proportion of local forest plant species in southeast Asia or part of the Indo-Pacific, while the exotic African honey bee may visit more species in the Neotropics (Roubik 1988, 1989, 1997; Kiew 1997). In contrast to the Meliponini, which have more abundant colonies, *Apis* usually has more foragers, considerably larger foraging ranges, or even migratory ability — best known from *Apis dorsata* in Asia (Oldroyd et al. 2000).

Conservatively, each honey bee colony might exist in forest at a density of 1 in 100 ha (see Seeley 1985). Certainly this is not true for tropical species in mature forest. *Apis dorsata*, for example, can have colonies remarkably clumped in space. A common number is 20 colonies of 40,000 bees each, in a single tree (Kiew 1997). For the stingless bees Meliponini, one might find 400 colonies in 100 ha of neotropical forest, yet only half as many populate forest having native honey bees (Roubik 1993a; Nagamitsu and Inoue 1997). The flight ranges of honey bees commonly reach 5 km from the nest, while I postulate those of Meliponini seldom exceed, on average, 0.8 km. (Maxima known for each bee group are 3 times the means given here.) Area covered per colony thus differs substantially in the two bees. It is roughly 80 km^2 and 2 km^2, respectively.

Let us assume 10 honey bee colonies live in the equivalent area, 1 km^2, of 400 stingless bee colonies. Total numbers of foragers may average 5600 $colony^{-1}$ for *Apis* (Dyer and Seeley 1991) but are near 2000 $colony^{-1}$ for Meliponini (Roubik 1993b). Taking this information together, Meliponini and *Apis* impact the forest by providing access to every flower by 2 to 5 million bees. Honey bees per flower, at the densities and colony sizes mentioned above, can be calculated as having 2.8x the impact of the Meliponini. Their potential on direct versus indirect gene dispersal is, however, not the same. For example, an 'average' honey bee colony may genetically tie together all the individual flowering trees within 80 km^2, but the stingless bee colony, on average, may only permit effective gene flow within 2 km^2.

While at present we cannot verify whether the foregoing conclusions are robust, data collected in tropical forests are suggestive. Here I examine the evidence and concepts related to highly eusocial bee colonies in the genetics and reproduction of tropical forest trees, and give results of two new studies.

Rules governing successful pollen/gene dispersal by bees

A majority of tropical trees have breeding systems that either severely restrict self-compatibility, or entirely eliminate the possibility of self-pollination in time or space (see Bawa 1990). If there are particular rules that determine whether genetic outcrossing occurs, they are that 1) pollen viability must continue over a period long enough to allow pollen transfer between trees. Yet more rules depend on little-studied variation in time and space (Fig. 1). For colonies as well as solitary bees, 2) a bee must be recruited, or make a foraging decision, to use more than one foraging area or more than one tree (Fig. 1c,d), 3), pollen left on surfaces of the nest, or transferred directly from one bee to another on body surfaces, also may join the pool of pollen donors for outcrossing (Fig. 1a,b). This can be called *secondary pollen dispersal*, originating from punctuated pollen transfer within the nest (see Hatjina et al. 1999). I suggest pollen making the trip between trees on the forager it was originally deposited upon can be called *ramified pollination* if there is an intervening return to the nest (Fig. 1a). Within a single foraging trip, it can generally be called *traplining* (Fig. 1b), if visits to multiple trees occur (see Janzen 1971; Thomson et al. 1997). Finally, when many foragers compete for foraging territory at flowering trees, there is opportunity for 4) *aggression-driven outcrossing* (Frankie et al. 1974; Johnson and Hubbell 1974; Nagamitsu and Inoue 1997; Roubik 1982, 1989). As these authors show, aerial competition can drive pollinators from patch to patch. In addition, some eusocial flower visitors simply remove pollen from the system without contacting stigmas, or discourage visitation to flowers. Aggressiveness, destructive robbing and degree of flower patch dominance and resource depletion by such groups should vary inversely with successful outcrossing by legitimate pollinators (Roubik 1989). The quantitative 'break even' point has not been worked out for this set of variables. Furthermore, to my knowledge no detailed field study of ramified pollination exists.

Some primarily botanical rules might be added to the above list, but they are at the stigmatic or carpel level and do not apply to the models discussed here. Botanical variables largely govern the realized degree of self-compatibility, which may vary among individuals. Similarly, they govern whether pollen germination on the stigma and ultimate ovule fertilization occur, which may be the result of a *mentor effect*, itself determined by whether enough grains, or sufficiently genetically varied grains, are present (Richards 1986; Kearns and Inouye 1993).

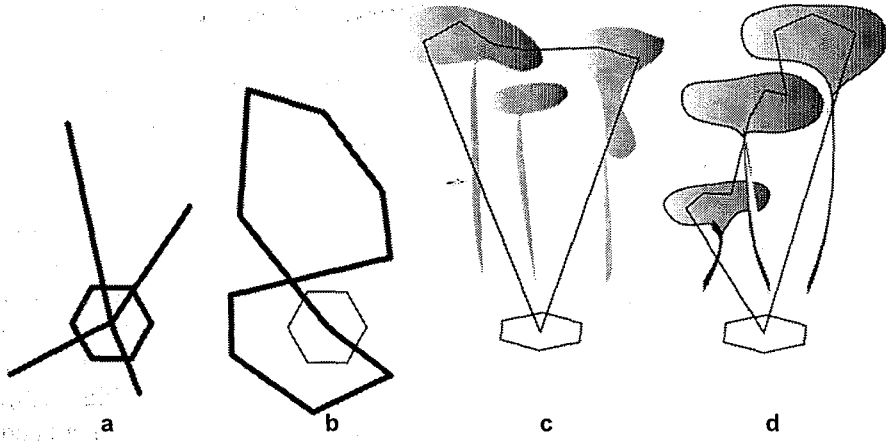


Figure 1. Schematic representation of pollen dispersal variation in space and time. The central-place forager (hexagon centre) and its foraging paths in the forest may produce: 'a' ramified pollination, 'b' traplining, 'c' a foraging bout restricted to high canopy or a single stratum, and 'd' a foraging bout incorporating multiple strata.

Experimental data on potential outcrossing ranges

The data available derive from information on African honey bees studied in lowland forest in central Gabon, Bornean honey bees in Lambir Hills field station, Sarawak (Roubik 1999; Roubik et al. 1995), and Meliponini and African honey bees in lowland neotropical forest in Panama (Roubik 1989 and present study). Only the first study measured the spatial sequences in foraging sites, while the rest determined foraging ranges and vagility of marked foragers moving between canopy and understory. In Gabon, 7,000 bees were marked over 28 days at 6 feeding stations positioned evenly 2 km along a forest trail. Most bees switched sites as food was depleted at feeders, and a negative exponential trend was detected, with the maximum distance of 1.6 km between feeders where the same bee appeared (Fig. 2). Within one day, the trend was similar. Shifts of 300-600 meters were relatively common, and the maximum was 1600 m. Distances flown by foragers between known sites could support a 177 ha gene pool among forest trees. An intriguing possibility for the behavioural mechanism was postulated but not proven. Either bees had returned to the nest and been recruited to other foraging stations, as naïve foragers, or, bees learned a number of foraging sites and shifted between them on their own *without first returning to the nest.* This subject needs further research.

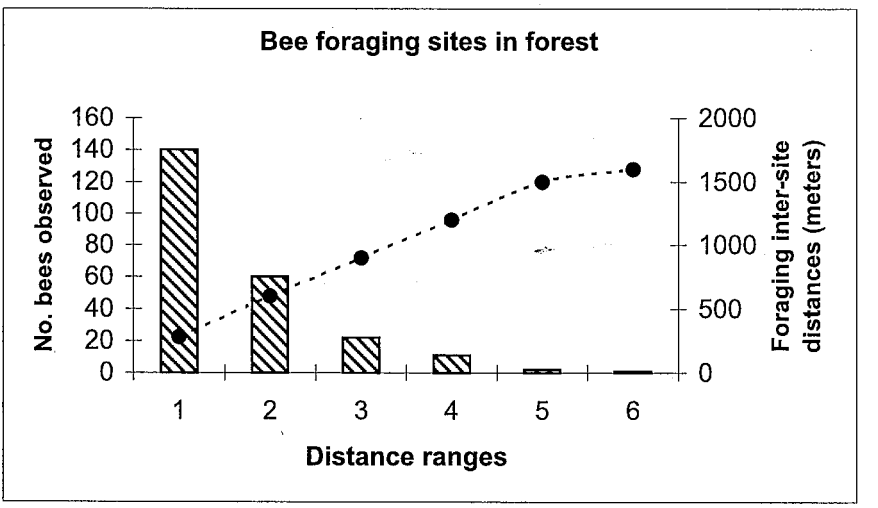


Figure 2. Frequency distribution of distance travelled between foraging sites by *Apis mellifera* in a Gabonese forest (Roubik 1999).

In the canopy walkway system at Lambir Hills, 3000 giant honey bees and Bornean honey bees (*A. dorsata* and *A. koschevnikovi)* were marked. They foraged from ground level to the highest canopy at over 60 m. Distances of 220 m between two tree towers were travelled by marked foragers within minutes, and one marked bee moved 640 m between the canopy and the ground. The most striking result was the speed with which bees discovered a second tree tower, after first feeding at a single tower. As soon as the small feeders were placed on the tree platforms, or on walkways between trees, bees arrived there. The bees quickly moved within an area of 60 ha that included canopy and understory. Some bees obviously scouted for new areas without being recruited to them. These are the individuals from eusocial colonies that presumably are responsible for a substantial portion of outcrossing effected by Meliponini or *Apis*.

Meliponini from known nests of two of the largest species were tagged with ferrous tags and 228 were released at distances up to 2.4 km from their nests in the closed forest of Barro Colorado Island, Panama (see Roubik 1989). The tags were captured from returning bees by using magnets at the nest entrances. Results showed that bees travelled from 1.5 to 2.1 km through forest producing a negative exponential distribution of bees over distance. Marked individuals from the larger species were observed at flowers 1.5 km from their nest.

In exploratory studies in another lowland Panamanian forest I utilized a canopy construction crane which provided access to forest at all levels within a 100 m diameter (Roubik 1993b). Totals of 40 canopy sites and 26 ground sites

were studied over 23 days, while marking bees at each site with paint of different colours. Feeders were placed at varied levels, including the top and up to 15 m *above* a 30-m forest canopy. The dominant bees were *Trigona*, *Partamona* or, occasionally, the African honey bee *Apis mellifera scutellata*.

Trigona corvina arrived in groups of a few hundred, as a compact cloud drifting over the top of the canopy. They attacked conspecific groups until only one colony was present, and may have indirectly discouraged visitation by most local Meliponini (22 species live there but only five arrived at feeders). Once the studies commenced, new sites were discovered by the three principal species within minutes and never took more than 30 minutes. All three foraged on the ground-level feeders and in the canopy. Contrary to all previous studies, aggressive *Trigona corvina* (summary in Roubik 1989) discovered new resources and moved to them over moderate distances of 55 to 160 m within a few hours or less. Thus even highly territorial foragers were somewhat flexible in their foraging site. Like the honey bees, flexible individuals (scout bees?) of aggressive group-foraging stingless bees searched for new resource sites by using odours, shapes or colours, or other cues. Furthermore, recruitment to feeders 15 m above the canopy demonstrated that such bees do not require deposition of odour trails, or previous experience at a particular site.

Experimental data on pollen viability

The subject of pollen dispersed by contact within a hive of honey bees has recently been brought up to the level of technically feasible pollination study in natural field conditions (Hatjina et al. 1999; Dafni 1999). The first authors reported that a surprisingly high proportion of pollen on the bodies of European *Apis*, as well as within the nest, and in the collected pollen carried on bee 'pollen-baskets' remained viable during the day. Pollen viability was slightly less than 50% in the conditions of the temperate zone summer, shown by pollen germination tests in which a sucrose medium was used.

I worked with three colonies of the stingless bee *Trigona fulviventris*, in the Cana field station area of Darién Province, Panama, and on Barro Colorado Island (BCI), Panama. Both studies took place during dry season of February and March, 2000, when seasonal flowering activity of trees was at a maximum. Sampling techniques were worked out in preliminary studies performed in 1989 using one of several available assays for pollen viability, Alexander's stain (Kearns and Inouye 1993). After generally poor results, I repeated the studies using viability stains and materials kindly supplied by A. Dafni. Sigma-fast urea hydrogen peroxidase No. U-5005 was used with 3,3 Diaminobesidine No. D-9167, in distilled water. With this method, grains that

are viable stain dark brown or purple, while the other grains are whitish.

My sampling procedure was designed to show whether bees were picking up additional pollen after they returned to the nest, and whether the pollen on the dorsum and venter of their bodies (not on the hindlegs, where pollen harvested for food is kept) was viable. The first 15 bees leaving the nest were trapped with a stationary insect net in early morning. Then at intervals of 20-30 minutes, groups of 15 incoming or outgoing foragers were sampled with an insect net. Collection vials using cyanide gas as the killing agent were used for groups of incoming and outgoing bees, and were thoroughly cleaned with tissue between collections. The insect net was swept vigorously between sampling periods, in an effort to dislodge any pollen. Replicates were made 10 to eleven times for each bee group. After removal from the collection vial, each bee was carefully washed in a droplet of approximately 10 microliters of viability stain solution. Microscope slides were then mounted with a fixative and coverslip, from which grain morphospecies and staining characteristics were scored.

Results from a single colony sampled on two successive days in the Darien forest agreed with two BCI colonies (Fig. 3). The first foragers leaving the nest were carrying multiple species of viable pollen on their bodies. Less than 10% of the pollen grains had lost viability. Furthermore, there was some indirect evidence for pollen transfer between bees. This conclusion is warranted by the dynamics of maximum pollen species diversity among groups exiting the nest, compared to those arriving. As shown in the Figure, returning bees tended to have more diversity of pollen on their bodies than did exiting bees. However, a lag period of one or two hours elapsed between peak pollen diversity on the incoming foragers, and a preceding pollen diversity peak on bees exiting the nest. Most notably in Colony 1, pollen diversity on exiting foragers increased gradually through the day. I make the inference that the gradual accumulation or tracking of peaks among incoming bees was due to pollen mixed within the next, then going out on exiting foragers. Indirectly, this is evidence that pollen transfer between bees *foraging at different individuals of the same species* could certainly occur.

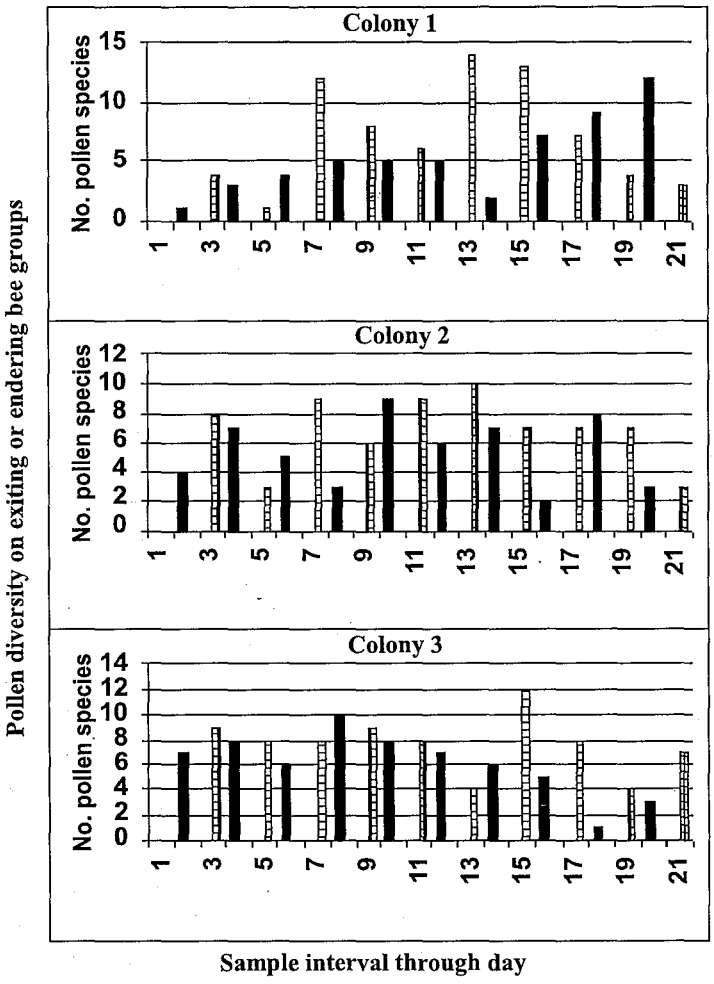


Figure 3. Pollen species carried on exiting and entering foragers in three colonies of Trigona fulviventris, sampled throughout the day in lowland primary forests in Panama. Sold bars represent exiting bees and hatched bars are the returning foragers. Pollen species were sampled from the bodies of bees, but not from the corbicula or pollen baskets.

Discussion

Individual bees in the nests of Meliponini and honey bees tend to continue their foraging activity throughout the day (Roubik 1989). At this point we know that bees not only take viable pollen to their nests, but that it is to some extent mixed

and redistributed there, among foragers. That much of this pollen is viable strongly contradicts older literature which gave the impression that pollen is rapidly killed once collected by a bee (Hatjima et al. 1999), either by glandular secretions or by anti-microbial substances such as resin within the nest (Roubik 1989).

Despite the fact that a bee may concentrate all its foraging activity on a single tree or flower patch, it may still provide some outcrossing pollen to the plant it visits. Contact with another bee or part of the nest can produce secondary pollen dispersal. The same phenomena should be possible in either small colonies or nests of solitary bees. Moreover, the results of studies reviewed here indicate that highly eusocial bees rapidly change their individual foraging sites, and that some continuously search for forage without returning to the nest. Traplining in *Melipona* has been demonstrated by their pollination of a sequentially monoecious vine, *Cayaponia macrocalyx* (Roubik in press). *Apis* also seems likely to trapline among known feeders, suggested by the Gabonese study (Roubik 1999). If the term traplining is imprecise, the outcome of foraging —outcrossing pollination— can still be embraced by the term ramified pollination. Pollen passes through the bee nest en route to its final stigmatic destination.

Because telemetry studies on a fine scale have yet to become feasible for bees in forest settings, mass-marking of highly eusocial bees recruiting to feeding stations remains the principal means of assessing the potential outcrossing ranges and gene pools supported by bees. While microsatellite studies in the tropics have begun to give solid evidence of relatively long-distance pollen dispersal (Dick, this volume; Gribel et al. 1998), we are still far from confident in assessing the qualities and parameters of plant gene flow in natural forest settings.

References

Bawa, K. S. (1990): Plant-pollinator interactions in tropical rain forest. Annu. Rev. Syst. Ecol. 21: 399-422

Dafni, A. (1999): Manual of pollination ecology: field methods. Institute of Evolution, University of Haifa, Israel

Dyer, F. C., Seeley, T. D. (1991): Nesting behavior and the evolution of worker tempo in four honey bee species. Ecology 72: 156-170

Frankie, G. W., Baker, H. G. (1974): The importance of pollinator behavior in reproductive biology of tropical trees. An. Inst. Biol. Univ. Nac. Autón. Mex. Ser. Bot. 45: 1-10

Gribel, R., Gibbs, P. e., Quieroz, A. L. (1999): Flowering phenology and pollination biology of *Ceiba pentandra* in central Amazonia. J. Trop. Ecol. 15: 247-263

Hatjina, F., Free, J. B., Paxton, R. J. (1999): Hive-entrance pollen transfer devices to increase the cross-pollination potential of honey bees. II. Examination of three materials and pollen viability. J. Apic. Res. 38: 3-9

Hubbell, S. P., Johnson, L. K. (1977): Competition and nest spacing in a tropical stingless bee community. Ecology 58: 949-963

Johnson, L. K., Hubbell, S. P. (1974): Aggression and competition among stingless bees: field studies. Ecology 55: 120-127

Johnson, L. K., Hubbell, S. P. (1984): Nest tree selectivity and density of stingless bee colonies in a Panamanian forest. pp. 147-154 in: Chadwick, A. C., Sutton, S. L. (Eds.) Tropical rain-forest: the leeds symposium. Leeds Phil. Lit. Soc., Leeds, U. K.

Kearns, K. A., Inouye, D. W. (1993): Techniques for pollination biologists. Univ. Press of Colorado, Niwot

Kiew, R. (1997): Analysis of pollen from combs of *Apis dorsata* in peninsular Malaysia. pp. 174-187 in: Mardan, M., Sipat, A., Kamaruddin, Y., Kiew, R., Abdullah, M. (Eds.) Tropical bees and the environment. Beenet Asia, Universiti Pertanian Malaysia, Serdang

Momose, K., Ishii, R., Sakai, S., Inoue, T. (1998): Plant reproductive intervals and pollinators in the aseasonal tropics: a new model. Proceed. Roy. Soc. London. B. 265: 2333-2339

Nagamitsu, T., Inoue, T. (1997): Aggressive foraging of social bees as a mechanism of floral resource partitioning in an Asian tropical rainforest. Oecologia 110: 432-439

Niuwstadtl, M. G. L. van, Sanchez, L. A., Biesmeijer, J. C., Arce, H. G., Sommeijer, M. J. (1997): pp. 100-110 in: Mardan, M., Sipat, A., Kamaruddin, Y., Kiew, R., Abdullah, M. (Eds.) Tropical bees and the environment. Beenet Asia, Universiti Pertanian Malaysia, Serdang

Oldroyd, B. P., Osborne, K. E., Mardan, M. (2000): Colony relatedness in aggregations of *Apis dorsata*. Ins. Soc. 47: 94-95

Richards, A. J. (1986): Plant breeding systems. George Allen and Unwin, London

Roubik, D. W. (1982): The ecological impact of nectar-robbing bees and pollinating hummingbirds on a tropical shrub. Ecology 63: 354-360

Roubik, D. W. (1988): An overview of Africanized honey bee populations: reproduction, diet and competition. pp. 45-54 in: Needham, C., Delfinado-Baker, M., Page, R., Bowman, C. (Eds.) Proc. Int. Conf. on Africanized honeybees and bee mites. E. Horwood, Chichester, U. K.

Roubik, D. W. (1989): Ecology and natural history of tropical bees. Cambridge Univ. Press, New York

Roubik, D. W. (1993a): Direct costs of forest reproduction, bee-cycling and the efficiency of pollination modes. J. Biosci. 18: 537-552.

Roubik, D. W. (1993b): Tropical pollinators in the canopy and understory: field data and theory for stratum "preferences". J. Ins. Behav. 6: 659-673

Roubik, D. W. (1996) Wild bees of Brunei Darussalam. Pp. 59-66 in: Edwards, D. S., Booth, W. E., Choey, S. C. (Eds.) Tropical Rainforest Research—Current Issues. Kluwer Academic, the Netherlands

Roubik, D. W. (1997): Diversity in the real world: tropical forests as pollinator reserves. pp. 25-33 in: Mardan, M., Sipat, A., Kamaruddin, Y., Kiew, R., Abdullah, M. (Eds.) Tropical bees and the environment. Beenet Asia, Universiti Pertanian Malaysia, Serdang

Roubik, D. W. (1999): The foraging and potential otucrossing pollination ranges of African honey bees in Congo forest. Journal of the Kansas Entomological Society 72: 394-401

Roubik, D. W. (in press): *Melipona* (Apidae: Meliponini) is a traplining pollinator of sequentially monoecious *Cayaponia macrocalyx* (Cucurbitaceae). Tropics.

Roubik, D. W., Inoue, T., Hamid, A. (1995): Canopy foraging by two tropical honeybees: bee height fidelity and tree genetic neighborhoods. Tropics 5: 81-93

Roubik, D. W., Moreno J. E. (2000): Pollen specialization and generalization by stingless bees. Pp. 112-118 in: Proc. Sixth Int. Conf. on Apiculture in Tropical Countries. Int. Bee Res. Assn. Cardiff, Wales, UK.

Seeley, T. D. (1985): Honeybee ecology. Princeton Univ. Press, Princeton, New Jersey

Thomson, J. D., Slatkin, M., Thomson, B. A. (1997): Trapline foraging by bumble bees: II. Definition and detection from sequence data. Behav. Ecol. 8: 199-210

An integrative approach to modelling mating systems of tree populations

H.-R.GREGORIUS

Institut für Forstgenetik und Forstpflanzenzüchtung, Universität Göttingen, Büsgenweg 2, 37077 Göttingen, Germany

Key words: mating systems, trees, modelling, estimation, functions of mating systems, adaptation, effective size, selection load, gene flow, parentage analysis, mating preferences, metapopulation structure

Abstract: The evaluation of existing approaches and the development of alternative approaches to modelling and estimation of mating system characteristics is put on a firm basis by considering the three fundamental functions of mating systems: (1) generative reproduction, (2) selection for participation in generative reproduction, and (3) combination of genetic information into genotypes. Intactness of the corresponding mating system mechanisms directly affects (1) adaptedness to current environmental conditions, (2) preservation of adaptability to future changes, and (3) preservation of adaptedness. Indicators for the intactness of these mechanisms are estimates of (1) reproductive success, (2) reproductively effective population size, and (3) selection load. The latter two are elaborated conceptually. All three indicators are integrated into the modelling and estimation of mating system characteristics by utilizing parentage distributions. A parentage distribution consists of a pool of zygotes and a group of potential parents such that each zygote has at least one parent from the group. This approach is useful for the design of models, methods of estimation and for the exploration of mating system characteristics. It is applied to measurement of the effects of mating systems on the demarcation of populations, on the reproductively effective number of maternal and paternal parents, on the amount of gene flow including its two sexual components, and on subpopulation differentiation by spatial variation in mating relations. Application of measures of mating preference as defined by the parental pairs of zygotes is shown to provide conceptually more satisfactory information on the reproductive isolation and coherence patterns that determine metapopulation structure and initiate or prevent speciation. Self-fertilization, for example, can be viewed in this context as establishing an extreme type of genealogical metapopulation structure, which is detectable with the help of estimates of "self-preference" but not of proportions of self-fertilization.

Introduction

In the present paper, an attempt will be made to bring to attention the fundamental functions of mating systems as the ultimate goals to which modelling and estimation of mating system characteristics are directed. By this, it is hoped

that the comparative evaluation of existing approaches can be put on a firmer basis and that the development of alternative approaches, if desirable, is aided.

Objectives of modelling

For a start, recall that any analysis of the consequences of mating system characteristics for population development and survival must be preceded by a description and estimation of these characteristics. Modelling is required here, either as a means of predicting this development on the basis of observed characteristics, or as a means of obtaining these characteristics if no methods of direct observation are available. In the latter case, where target characteristics are not observable, a model is designed in which the target variables depend on observable variables. Calibration of this model with respect to the observations (usually by maximization of likelihood) then yields "indirect" (or model-dependent) estimates of the desired target characteristics. These estimates are only acceptable if a test of the calibrated model does not recommend its rejection. Estimation of rates of self-fertilization in trees is a well-known example of indirect estimation.

A third common area of application of models concerns the detection of equivalencies in performance between different systems. This is usually realized in terms of "effective sizes", which result from the comparison of a complex system with an "ideal" model system. In particular, this comprises all non-testable models, which yield "estimates" of model parameters after calibration with respect to observations. Such "estimates" do not refer to characteristics of the observed system and should thus be addressed as effective parameter values. An example is to be seen in F_{ST} used to "estimate" amounts of gene flow (in the form of $N \cdot m$) on the basis of Wright's idealized model of drift and migration. Experimental tests of this model are very difficult and possibly infeasible. $N \cdot m$ may, in this case, be addressed as an effective amount of migration, but it does not estimate any realized number of migrants (for a detailed criticism of the F_{ST} method as yielding indirect estimates of gene flow, see Whitlock and McCauley 1999). This category of model is thus of limited value in the analysis of characteristics of real systems.

In summary, one can distinguish three major objectives of modelling,

- to provide *testable hypotheses on causal mechanisms* and serve the *prediction* of developments and *planning* of actions,
- to enable *indirect estimation* of system characteristics, after passing a test ("model-dependent estimation"), and
- to enable the detection of *equivalencies in performance* between different systems ("effective sizes")

Mating system functions

While the above objectives are pursued to various degrees in each analytical study of mating systems, the design of the applied models frequently does not permit clear recognition of their relations to general biological functions of mating systems. Since the inherent principles can be expected to guide any modelling effort it is useful to recall briefly the three *fundamental functions of mating systems*, which are

1. generative *reproduction*,
2. *selection* for participation in generative reproduction,
3. *combination* of genetic information (genes) into genotypes.

These functions of mating systems determine those adaptational capacities of populations which can be realized during the transition from one generation to the next. Consideration of the functions in experimental analyses is thus required to assess the significance of observable mating system characteristics for population development and survival. The same requirement applies, of course, to the design and parameterisation of mating system models. These models always contain the observable mating system characteristics as variables. They must, however, frequently employ non-observable variables (free parameters) in order to enable the desired analysis. As was mentioned above, non-observable model variables serve in the calibration of the model and by this enable indirect estimation. They can also be varied with the aim to predict effects of certain scenarios on the three fundamental functions.

Major determinants of plant mating systems

In order to simplify reference of the following reflections to their biological basis, a short list of categories of factors is compiled, which affect plant mating systems. Among the categories most frequently considered in experimental studies, the following can be distinguished:

i. *Spatial* relations: Spatial distribution of individuals, in relation to their pollen dispersal characteristics.
ii. *Temporal* relations: Temporally varying spatial distribution patterns and activities or behaviour, age of female and male sexual maturity, timing of female and male flowering, time-dependent expression of reproductively relevant phenotypes.
iii. *Phenotypic* relations: Prezygotic incompatibility or isolation mechanisms, including biochemical or physiological agents, and morphological barriers (concerning e.g. flowering phenology). With the exception of purely genetic control, expression of the relevant traits involves interactions of genotypes with environments.

iv. *Ecological* conditions: Availability and selectivity of pollinators, physical barriers to pollen dispersal, species composition.
v. *Genealogical* relations: With the exception of self-fertilization, they are rarely direct determinants of mating relations. Preferential mating among relatives in plants is mostly a consequence of limited seed and pollen dispersal, which reveals spatial rather than genealogical relations as direct determinants of mating.

With particular reference to forest tree mating systems, mixed mating in the form of selfing and random cross-fertilization (category (*v*)), preferential mating among neighbours (category (*i*)), and gametophytic and sporophytic incompatibility (category (*iii*)) have received most attention. More recently, ecological conditions (category (*iv*)) are attracting some interest because of increased concern about tropical tree species with their animal dominated pollination systems.

The adaptational context of mating systems

Intactness of a mating system

Adaptedness and adaptability of mating systems to environmental conditions can be realized only if the mechanisms performing the three basic functions of the mating system are intact. Corresponding to the three functions,

(1) the mechanisms of generative reproduction are intact if sufficient numbers of offspring are produced, in the sense that the number of offspring compensates for the number of deaths (i.e. the number of offspring of a cohort over its total life span is at least equal to the cohort's initial size);
(2) the mechanisms of selection for generative reproduction are intact if all of the adults' genes are represented in their successful gametes (gametes appearing in zygotes);
(3) the mechanisms of combination are intact if the genotypic composition of the offspring guarantees sufficient chances for survival to adulthood and reproduction in the next generation.

Translated back into adaptational terms, this states that intact performance of the three functions implies that the mating system of a population

(1) is adapted to the current environmental conditions,
(2) preserves the adaptability to future environmental changes by preserving genetic variation,
(3) preserves the adaptedness by reducing the mortality implied by the adaptational pressures on the next generation.

As opposed to mating system functions (1) and (2), adaptational forces do not directly act on function (3). The adaptational effect of the performance of function (3) is determined by the initial conditions that it provides for the performance of functions (1) and (2) in the next generation. In other words, function (3) should be performed such that the genotypes, which are adaptationally advantageous under the environmental conditions of the next generation, are produced at sufficient frequencies.

The mechanisms of the mating system are therefore impaired in the case of insufficient offspring production (function (1)), participation of only a small fraction of adults in the offspring production (function (2)), or in the case of excessive formation of adaptationally inferior genotypes (function (3)).

Example: inbreeding depression

An example of the performance of mating system function (3) is provided by the degree of self-fertilization, in combination with homozygote disadvantage, which is frequently addressed as inbreeding depression. For given allele frequencies, the share of adaptationally inferior genotypes increases with increasing degree of selfing. Over the generations, this share will gradually lessen for dominant gene action, until a selection-mutation equilibrium is reached. When starting with a high share of inferior genotypes, the following reduction in population size could be so drastic that the implied genetic drift effects could entail substantial losses of adaptationally important variation in the genetic background. Under such conditions, the intactness of the mating system becomes manifest in the effects of the degree of selfing on the average survival and reproduction of the offspring generation.

Example: mode of pollination

Performance of functions (1) and (3) can be affected simultaneously, for example, by the mode of pollination. According to Lloyd (1979), three such modes can be distinguished: prior, delayed, and competing self-pollination. Ziehe and Gregorius (1988) demonstrated that each of these modes affects the pollination efficiency (function (1)) and the degree of self-fertilization (combination of genes into genotypes, function (3)) differently. In particular, delayed self-pollination may increase pollination efficiency in a supplementary way if cross-pollination was insufficient or failed as a result of low population density or of colonization events. Note that the assumptions of the classical mixed mating model (random cross-fertilization, fixed ovule selfing proportions, all ovules have the same chance to be fertilized) postulate the performance of

the three fundamental functions rather than explain how they are affected by certain mating relations.

Example: evolutionary effects of mating systems

The adaptational pressures on populations may reinforce the evolution of reproductive separation (the "Wallace effect" after Wallace 1889) or of reproductive coherence (Steiner and Gregorius 1997) between unlike genetic types. Separation and coherence correspond to preventing and enhancing heterotypic matings. Consequently, reproductive separation initiates speciation, and reproductive coherence stabilizes the population as a reproductive community. In both cases the losses due to selection are reduced and thus the adaptedness to the respective environmental conditions improved. Again, intactness of the mating mechanisms refers to the performance of function (3).

Example: subpopulation differentiation

The concepts of reproductive separation and reproductive coherence are also relevant at the population or metapopulation level. Adaptation to spatially heterogeneous environments requires a certain degree of reproductive isolation in order to limit the formation of adaptationally disadvantageous genotypes (function (3)). Depending on the type of environmental heterogeneity (spatial, temporal, etc.), the causes for the isolation may fall into any of the above-listed categories (*i*) to (*iv*) of determinants of mating. Reproductive isolation leading to limited reproductive neighbourhoods may be a prerequisite for the evolution of local adaptations. Selfing need not contribute to the formation of reproductive neighbourhoods, but it directly affects the selection loads within these neighbourhoods.

Indication of the adaptational status of a mating system

It remains to demonstrate the practicability of the present approach of tracing back the adaptational status of mating systems to the intactness of their mechanisms performing the fundamental functions. This requires us to specify the indicator variables for quantification of the intactness of mechanisms of the mating system, such that they attest fulfilment of the criterion of population survival. Reflecting the three fundamental functions and their conditions of intact performance, the above exposition suggests as primary indicator variables

(1) the *reproductive success*, as defined by the number of successful gametes produced per member of a cohort,

(2) the *reproductively effective population size*, as defined by the effective number of population members contributing to the zygotes produced in a specified period of time, and

(3) the *cohort selection load*, as referred to the zygotes that established the cohort and the number of successful gametes of each cohort member (fitness). The cohort selection load is then defined by minimum reduction of the reproductive capacity of a cohort that is required to arrive at the actual differences in production of successful gametes between types (a generalized concept of selection load including survival and reproduction is introduced in the Appendix).

Even though the experimental verification of indicator (1) simply requires counts of cohort members and their offspring, these counts may be difficult to obtain in iteroparous organisms, since offspring cannot be unambiguously assigned to a single cohort. In such cases, model-dependent methods of estimating life table data must be applied, which are based on observations from different cohorts.

The example of inbreeding depression can again be used to demonstrate the effects of the mating system via self-fertilization on the cohort selection load (indicator (3)). In combination with low reproductively effective population sizes (indicator (2)), large selection loads resulting from unbalanced degrees of selfing can accelerate the loss of adaptational capacity by the loss of genetic variation.

Among the three primary indicators, the third is probably most difficult to study comprehensively in long-lived organisms like trees. It can, however, be very informative to consider defined phases of the reproductive cycle separately. As an example, this is the case for early stages, where the drastic reductions of the population size following seed production allow for strong selective adaptation. In this case, a considerable fraction of the cohort selection load is attributable to early developmental stages. Postzygotic incompatibility constitutes an important special form of this load. It is therefore useful to restrict studies of the effects of mating and viability selection on the selection load to special phases.

A more comprehensive but strongly model-dependent idea of the overall cohort selection load can be obtained from a comparison of adult trees with their seed production. Assuming that the predominant characteristics of the mating relations do not change essentially over the generations, the adult genotypic structures can be considered to have resulted from viability selection that acted on zygotic frequencies which were similar to those observable among the offspring of these adults. Thus, reverting the actual succession of mating and selection in this way, the assumption permits computation of selection loads.

The requisites for the determination of selection loads are genotypic frequencies at two successive stages, the first of which is close to the zygotic stage. While this is experimentally feasible in most situations, the requisites for the direct determination of indicator (2), the reproductively effective population size, are more difficult to realize. The reason is that for direct determination, methods are required which allow the identification of the parents of each offspring in a sample (as detailed in the next chapter on "parentage distributions"). With the perfection of methods of DNA-analysis, the chances for obtaining such direct information are likely to improve considerably. Yet, since reproduction means identical multiplication of individual genetic information, model-dependent methods of estimating reproductively effective sizes in plants will always be required.

The commonly applied methods of estimating effective population sizes rely on quite restrictive model assumptions that are hard to verify or test (see e.g. Schoen and Brown 1991, for an application of such a method). This problem is aggravated by the fact that in many studies, the characteristics for which effective sizes are defined (such as inbreeding, variance, drift, reproduction, etc.) are not clearly stated. In fact, reproductively effective sizes are occasionally subsumed under some of these characteristics, and it appears that the above definition is not yet explicitly applied in theoretical or experimental work.

Some mating system characteristics may affect several indicators of intactness simultaneously. This is true for the above-mentioned modes of pollination, which affect indicators (1) and (3) via determination of the reproductive success and the selection load. In more complex situations like this, the question of how mating systems affect the indicators of intactness can frequently be answered only with the help of simulation scenarios, based on models.

Characterizing mating systems by parentage distributions

When a decision is to be made on the mating system characteristics to be studied, the above explanations suggest that their potential effects on the intactness of mechanisms of the mating system should be taken into consideration. For the applied models, this requires that their design and parameterisation should enable inferences on the status of intactness of the addressed mechanisms. Since the basis for these inferences is provided by the three primary indicator variables of intactness, the model should supply information on the differential reproductive success as part of the mating process. This can be achieved in two ways. The direct approach consists of designing a model to yield estimates of the indicators, after calibration for the observations. Otherwise,

the model design should at least provide for results which can be used in other models that produce the indicator variables under realistic scenarios. Estimates of a proportion of self-fertilization could, for example, be obtained with the help of a model that contained no assumptions on the numbers of zygotes produced by the various genotypes (this is true for the classical mixed mating model, for example). Another model, in which such numbers are explicitly taken into account, could be compatible with the former model, such that the selfing estimates could be adopted and an analysis of intactness of the mating system mechanisms could be carried out. In any case, a clear concept of the kind of observations that would be useful may account for all of these aspects.

Inclusion of intactness aspects into the analysis can be realized in an ideal manner, if experimental and model designs focus on each individual offspring as representing a unique and successful mating event. The characterization of a mating episode would thus be based on offspring (ideally zygotes) as units of observation, where for each offspring two "traits" are scored, one indicating its maternal and the other of its paternal parent. These observations specify a *parentage distribution* for each collection of offspring. The distribution refers to any frequency distribution on the collection of offspring (see Table 1). The parentage distribution thus summarizes all of the information relevant for the estimation of frequencies of mating types and reproductive successes.

Table 1. Parentage distribution with Z_{ij} : = number of zygotes with the i-th individual as maternal and the j-th individual as paternal parent. The Z_{ii}'s of the diagonal represent numbers of zygotes resulting from self-fertilization. If the i-th individual is a female, then $Z_{ji} = 0$ for all j, i.e. the i-th column in the table consists of zeros. Dioecious species are therefore characterized by the fact that if the i-th row contains positive elements then the i-th column consists only of zeros and vice versa.

	♂–parent				
♀–parent	1	2	3	4	...
1	Z_{11}	Z_{12}	Z_{13}	Z_{14}	...
2	Z_{21}	Z_{22}	Z_{23}	Z_{24}	...
3	Z_{31}	Z_{32}	Z_{33}	Z_{34}	...
4	Z_{41}	Z_{42}	Z_{43}	Z_{44}	...
...	...	...	...	...	...

Paternity analysis (Hamrick and Schnabel, 1985) is a well known example for the utilization of parentage distributions. This method accounts for the fact that

in plants seeds are usually collected before dispersal from each of a sample of individuals. In this case the maternal parent of each offspring is known and the paternal parent is to be inferred with the help of gene markers. The relevant methods are mainly based on paternity exclusion complemented by likelihood estimation procedures of paternity (see e.g. Weir 1996, p.209ff). It is well known that the precision of the resulting inference depends heavily on the available marker, the sample of potential paternal parents and the samples of seeds. Models have a substantial part in the likelihood estimation procedures, even though they are frequently not explicitly mentioned (mostly concerning free recombination, Mendelian inheritance, stochastic independence among loci, absence of postzygotic selection, or random mating). Particularly the assumption of random mating in the model-dependent estimation of parentage is problematic, if the estimates are used in an analysis of mating relations, because of the danger of circular reasoning.

The set of zygotes on which an analysis of mating system characteristics is to be based depends on the problem to be studied. In the following sections, this will be demonstrated by addressing a few problems of elementary significance. Suggestions for an integrative approach to their treatment with the help of parentage distributions will be made.

Demarcation of populations

Even if the totality of a population's seed production could be representatively sampled prior to dispersal, only the maternal contribution to each seed can be definitely stated to have originated from this population. The paternal contribution could result via pollen immigration from other populations. Moreover, a pollen grain produced by a population member could fertilize an ovule produced by the member of another, neighbouring population. Seed dispersal could bring the resulting seed back into the vicinity of its paternal parent. Strictly speaking, this raises the question as to the population to which such a seed should be assigned and thus brings to attention the demarcation of a population as a reproduction community. Since a reproduction community is, in turn, defined by the mating relations of its members, mating systems can be conceived of as fundamental determinants of populations.

To approach this problem, the above notion of parentage distributions will be generalized to include the situation of a group of potential parents, together with all zygotes with at least one parent from this group. All of these zygotes, but no others, result from mating relations realized by the members of the group. They must therefore be part of the group's mating system, irrespective of the place where they become established (grow). To simplify wording, zygotes with both parents from the group will be called "homodemic" with respect to this group,

and zygotes with only one parent from the group will be called "heterodemic" (for an illustration see top of Figure 1). There is probably wide consent to call the group a population if all zygotes are homodemic with respect to the group. Yet, since populations are also generally accepted to be open systems, sufficiently small fractions of heterodemic zygotes are counted as a result of gene flow by mating into the group, without questioning its status as a population.

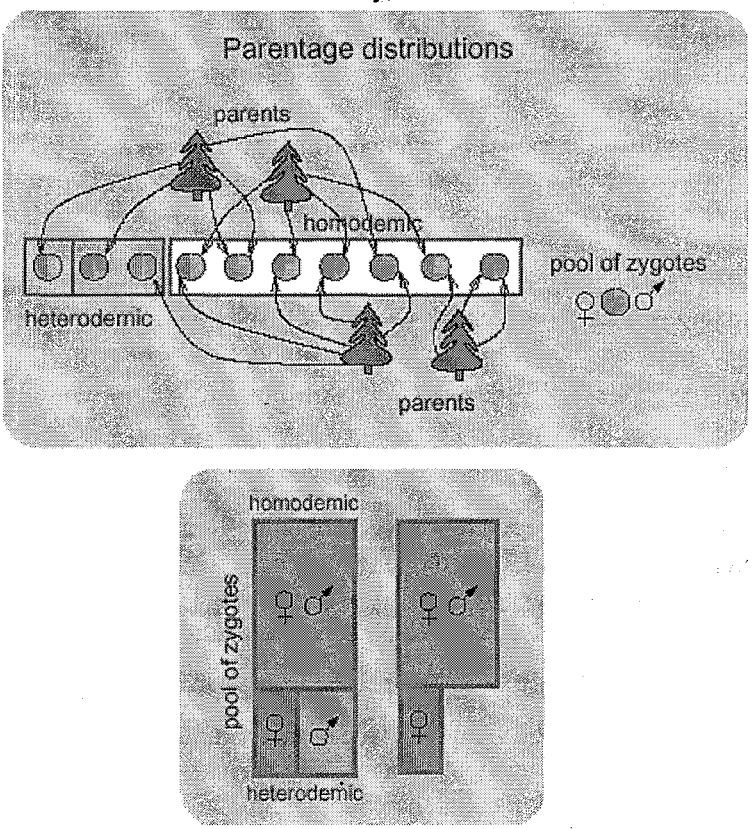


Figure 1. Top: Illustration of homodemic and heterodemic matings in parentage distributions. Bottom: Illustration of amounts of gene flow by mating considering both maternal and paternal parents (left) and considering only maternal parents (right) among the heterodemic matings.

With increasing fractions of heterodemic zygotes the population concept does gradually become blurred and no clear delineation can be made. It could therefore be meaningful and is probably closer to reality, if one specifies the *degree to which a group behaves as a closed population* by the fraction of

homodemic zygotes among all zygotes (homo- and heterodemic) of the group. This fraction becomes 1 for a completely closed population (or reproductively isolated group) and it reaches a value of 0 if the members of the group mate only with individuals outside the group. A trivial example of a group with a zero degree of its population status is provided by any set of males in a dioecious species. To realize a positive degree, some females are to be added to this group.

Gene flow by mating

For a given population, its gene flow by mating is described by the set of heterodemic zygotes. To enable a formal representation, denote by Z the number of zygotes with at least one parent from the population, by Z_{hom} and Z_{het} the numbers of homodemic and heterodemic zygotes, and by Z_{het}^{f} and Z_{het}^{m} the numbers of heterodemic zygotes with maternal and paternal parent, respectively, from the population. Then $Z_{het} = Z_{het}^{f} + Z_{het}^{m}$ and $Z = Z_{hom} + Z_{het}$. The fraction of zygotes resulting from gene flow thus equals Z_{het}/Z (also note that this equals the complement of the above degree Z_{hom}/Z to which a group behaves as a closed population, see Figure 1). This obvious measure of the overall amount of gene flow by mating does not, however, underly the common measures of gene flow. The reason is to be found in the fact that in the common concept an offspring is assigned to the population to which either its paternal or (mostly) its maternal parent belongs. Only from this point of view can gene flow by mating be conceived of as being directed into or out of a population. If neither gametic sex can be assigned a sessile role, gene flow cannot be analysed for its direction.

In plants, seeds are almost exclusively assigned to the population of their maternal parents. Consequently, gene flow into a population by mating is measured by the fraction of heterodemic seeds, but excluding all those heterodemic seed with paternal parent from within the population and maternal parent from outside the population (i.e. Z_{het}^{m}, see bottom of Figure 1). With the above notation, this fraction amounts to $Z_{het}^{f}/(Z - Z_{het}^{m})$. Clearly, this measure of gene flow can be used to describe external mating relations of population members functioning as female parents. The external mating relations of population members functioning as male parents, which would represent the gene flow out of the population, are here completely ignored. This marks the strong bias observable in the great majority of experimental and theoretical studies (see e.g. the review of Adams and Birkes, 1991, which still provides a good account of the prevailing basic approaches to modelling and estimation of mating system characteristics of forest tree populations; Willson, 1994, reviews sexual selection in plants as being female governed; Gregorius et al. 1987, demonstrate the differential effects of both sexes on the measurement of self-fertilization).

The tacit assumption that female parents are the predominant determinants of plant mating systems still awaits an experimental verification. Concerning external mating relations, this would require at least a comparison between maternal and paternal parents with respect to their amounts of heterodemic zygotes. The fraction of heterodemic zygotes characterizes the extent to which the three fundamental functions of mating systems are realized through reproductive contacts with individuals from outside the group. Therefore, and because of the differences between genetic information transmitted by male and female gametes (chiefly concerning extra-nuclear information such as that residing in mitochondria and plastids), it is important to have reliable information about a possible asymmetry between the sexes in their external mating relations ($Z_{het}{}^{f} \neq Z_{het}{}^{m}$). The development of methods for the estimation of amounts $Z_{het}{}^{m}$ of external matings by paternal parents is a big challenge.

In fact, the fraction Z_{het}/Z of overall gene flow by mating always exceeds the female oriented fraction $Z_{het}{}^{f}/(Z - Z_{het}{}^{m})$ as can be taken from

$$\frac{Z_{het}}{Z} - \frac{Z_{het}^{f}}{Z - Z_{het}^{m}} = \frac{Z_{het}^{m} \cdot Z_{\hom}}{Z \cdot [Z - Z_{het}^{m}]}$$

Hence, even strong (but with the presently available experimental means, hardly verifiable) assumptions such as sexual symmetry among heterodemic mating relations ($Z_{het}{}^{f} = Z_{het}{}^{m}$) cannot compensate for this difference in the measurement of gene flow by mating.

Gene flow, in the sense of heterodemic matings, involves intactness considerations through its effects on the reproductive success of the population members, on the reproductively effective population size, and on the genotypic structure among the offspring. The assessment of reproductive success and effective size is complicated by the fact that heterodemic matings involve only one parent from the population. Gene flow by mating may thus affect the intactness performance of all three fundamental functions of mating systems.

Reproductively effective number of parents

There are two approaches to the reproductively effective numbers of parents. One approach focuses on a group of potential parents, the reproductive output of which is represented by the totality of their successful gametes. In terms of zygotes, the totality of successful gametes is contained in the pool of all zygotes with at least one parent from the group of potential parents. An effective number of parents is thus defined for a specified group of potential parents. The other approach focuses on a pool of zygotes as the reproductive output of their parents. An effective number of parents is here defined for a specified pool of zygotes.

The following derivations will be formulated such that they apply equally to both approaches.

Again denoting by Z the number of zygotes under consideration, it follows that the effective number of parents of these zygotes cannot exceed a number of $2 \cdot Z$. More specifically, let Z_{ij} as specified in Table 1, so that $\sum_{i,j} Z_{ij} = Z$. The number of zygotes produced by self-fertilization of the i-th parent equals Z_{ii}. If all parents would reproduce solely by self-fertilization, the effective number of parents could not be more than Z, which is half of the maximum number realizable without self-fertilization.

It follows that the i-th parent occurs $\sum_j (Z_{ij} + Z_{ji})$ times as a contributor of a gamete to the pool of zygotes under consideration. In terms of relative frequencies, the i-th parent therefore has a share of $g_i := \sum_j (Z_{ij} + Z_{ji})/(2 \cdot Z)$ among all gametes contributed to the zygotes. Note that $\sum_i g_i = 1$ since all parents of the pool of zygotes are taken into consideration. These relative frequencies allow us to relate the concept of effective number to that of diversity. One of the most frequently applied measures of diversity is $v_2 = (\sum_i g_i^2)^{-1}$, which, considering the definition of the g_i's, specifies the reproductively effective number of parents in terms of the diversity of parents in their contribution of gametes to the pool of zygotes. The measure v_2 is one from a continuum of measures v_a (given by $v_a = (\sum_i g_i^a)^{\frac{1}{1-a}}$), where the index a runs from 0 to ∞, and where v_0 equals the number of types (parents) found in a collection and $v_\infty = (\max_i g_i)^{-1}$ For a given frequency distribution, v_a decreases with increasing a, so that v_0 and $v\infty$ constitute the largest and smallest diversity measure in this continuum (Gregorius 1978).

The smallest measure $v\infty$ is distinguished by a property which has particular intuitive appeal as a measure of the reproductively effective number. Considering the parent that contributes the most to the pool of successful gametes as a reference for effectiveness, it is of immediate interest to know how many of such parents would have sufficed to produce all of the gametes of the pool. This number would ideally reflect the notion of a number of reproductively effective parents. In fact, if G denotes the overall number of successful gametes, then $G \cdot \max_i g_i$ equals the maximum number of successful gametes contributed by a single parent. If each parent would contribute this number to the pool of successful gametes, it would require $G/(G \cdot \max_i g_i) = (\max_i g_i)^{-1}$ such parents to account for all the gametes.

The effective number of parents alone makes an incomplete statement as to the intact performance of the second fundamental function of mating systems. It remains to relate the pool of zygotes and their parents to a group of individuals considered as a population of potential parents, as is pointed out in the first of the two above-mentioned approaches. These potential parents must, of course, comprise at least one of the actual parents for each zygote from the considered pool of zygotes. By this, it is guaranteed that all zygotes are offspring of members of the group, with the possibility that not all zygotes have both parents from the group (existence of heterodemic zygotes) and that not all of the successful gametes of the group are represented in the pool of zygotes. The reproductively effective size of the group is defined relative to the considered pool of zygotes, and it equals the above effective number of parents only if all zygotes are homodemic for the group.

Otherwise, each heterodemic zygote is represented by only one parent from the group, so that the maximum number of parents from the group contributing to the zygotes reduces from $2 \cdot Z$ to $2 \cdot Z - Z_{het}$. Then the i-th group member has a share of $g_i = \sum_j (Z_{ij} + Z_{ji}) / (2 \cdot Z - Z_{het})$ among all gametes contributed to the zygotes by the group. Note that the subscript j runs over the whole set of parents of the zygotes (thus including parents from outside the group), while the subscript i refers only to group members. Hence, summation of the g_i's only over group members yields 1. The above indices ν_a, when applied to these g_i's, are again possible measures of the reproductively effective number of group members.

The same principles apply to the determination of sex-specific reproductively effective numbers. Thus, the reproductively effective number of maternal parents rests on the fractions $g_i^f = \sum_j Z_{ij} / (Z - Z_{het}^m)$ of gametes contributed by the i-th group member to all successful female gametes of the group. The pertinent fractions for the reproductively effective number of paternal parents are $g_j^m = \sum_i Z_{ij} / (Z - Z_{het}^f)$ where j refers to the j-th group member.

The above explanations apply to any given group of individuals and pool of zygotes, provided each zygote from the pool has at least one parent from the group. This allows us to treat a large variety of situations in terms of reproductively effective numbers of parents. For example, if the seed of a single tree is to be analysed with respect to its effective number of paternal parents, this can be done on the basis of the above frequencies g_j^m. In this case $g_j^m = Z_{ij}/Z$ and $Z_{het}^f = 0$, since only one maternal parent (individual i) is considered and since all paternal parents of the zygotes are included in the group of potential paternal parents. This can be extended to any group of potential parents, in combination with the pool of all zygotes with maternal parent from the group. It

is also possible to invert the point of view by consideration of the pool of all zygotes with paternal parent from the group. This case will, however, be hardly possible to study in plants, because the totality of ovules fertilized by a pollen parent can normally not be sampled.

Population structure due to mating preferences in continuous populations

The structure of metapopulations is closely related to subpopulation differentiation due to mating preferences. This type of differentiation can again most consistently be analyzed if parentage distributions can be assessed. Parentage distributions are to be specified for the differentiation criterion to be studied. For example, if the differentiation criterion is membership of locally defined groups, each offspring is characterized by the group membership of its maternal and paternal parent. Since location is a spatial characteristic, the above category (*i*) of mating system determinants applies. Parentage distributions are then specified by the frequencies of offspring with parents belonging to the same group (homodemic offspring) or belonging to different groups (heterodemic offspring). If external matings are considered not to contribute to population subdivision, only offspring with both parents from the total population (homodemic for the total population) enter the analysis.

Subpopulation differentiation due to preferential mating among members of the same group can then be inferred by comparison of the frequencies of homodemic and heterodemic offspring, since these correspond to matings within and between groups. If no obvious subdivision of the population into disjoint groups is observable (i.e. if the population is continuous), an analysis of mating systems is usually aimed at the detection of relationships between spatial distance and mating. A mere analysis of frequency distributions of mating distances (frequencies of zygotes with given spatial distance between their parents), however, may not be satisfactory. Spatially heterogeneous distribution of potential mating partners may feign preferential mating among neighbours simply because of clumped occurrence. The situation is similar to the bias of spatial autocorrelation analyses by clumped spatial distributions of population members. Such pitfalls can be avoided when the analysis is based on a clear concept of mating preferences, as will be demonstrated in the following section.

Spatial distance and individual mating preferences

As a rule, preferential mating among spatial neighbours cannot produce discrete subpopulation structures if neighbourhood is distributed more or less evenly in space. In this case, an analysis of subpopulation differentiation may be inadequate,

since no *a priori* discrete subpopulation structure can be identified. A more adequate approach to studies of population structure is then suggested by an analysis of the mating preferences of each parent as a function of the distance between the parent and its mates.

More concretely, consider all zygotes which have the i-th individual as female parent, the frequency of which is $\sum_l Z_{il}$. The j-th potential male parent appears at a proportion $Z_{ij} / \sum_l Z_{il}$ as contributor to these zygotes, and for each such male parent its distance from the reference female parent i is recorded. Note that individuals are regarded as potential parents only to the degree to which they contribute to the pool of zygotes. In order to detect special preferences with respect to distance, all paternal contributors to the total pool of zygotes have to be drawn upon, together with their distances from the reference female parent, for comparison. On this set of potential mates, in which the j-th paternal parent is represented with a proportion $\sum_k Z_{kj} / Z$, the mating preferences are based. The preference of the female parent for a particular mate is then obtained by computing (a) the frequency of this mate among all mates of the female and (b) the frequency with which the mate occurs among the paternal contributions to the total pool of zygotes. Division of (a) by (b) yields the desired measure of preference. The mating preference $U_{j<i}$ of the i-th maternal parent for the j-th paternal parent is therefore given by

$$U_{j<i} = \frac{Z_{ij}}{\sum_l Z_{il}} \Big/ \frac{\sum_k Z_{kj}}{Z} = \frac{Z_{ij} \cdot Z}{[\sum_l Z_{il}] \cdot [\sum_k Z_{kj}]}$$

(for the concept of mating preferences see Gregorius 1989). For each mate of the reference female parent, its mating preference and its spatial distance are thus known, and this permits an analysis of neighbourhood mating by plotting distance against mating preference. The resulting graphs are directly interpretable (consult Figure 2 for an illustration of isotropic - i.e. independent of direction - mating preferences).

This procedure can be repeated for each female parent and would provide an impression of the degree to which spatial distance determines female mating preferences. The same can be done with the mating preferences of male parents, in which case $U_{i<j}$ denotes the preference of the j-th paternal parent for the i-th maternal parent. The symmetry of the preferences, i.e. $U_{i<j} = U_{i<j}$, follows directly from the above definition of mating preferences. This symmetry is a consequence of considering individuals as potential mating partners, to the degree to which they contribute to zygotes.

In such an analysis, spatial distance between parents can in fact be replaced by any other measure of difference between parental characteristics, without having to change the principle of the analysis. This is easily realized for the categories (*iii*) and (*v*) of mating system determinants, when considering differences between parental pairs with respect to their phenotypes, genotypes, or with respect to their common ancestry, measured in terms of coefficients of kinship.

A problem frequently arising in parentage analyses of co-sexual plants consists in the lack of means to distinguish ovule from pollen contributions to zygotes or embryos. Parents of zygotes may still be identifiable, but they cannot be distinguished with respect to their maternal and paternal functions. In this situation observations are restricted to the symmetric frequencies $Z^{\circ}_{ij} := Z_{ij} + Z_{ji}$ for $i \neq j$ and $Z^{\circ}_{ii} := Z_{ii}$Obviously $Z^{\circ}_{ij} = Z^{\circ}_{ji}$ and $\sum_{i \leq j} Z^{\circ}_{ij} = Z$. Under these restrictions, the frequency with which the i-th individual contributes as (maternal or paternal) parent to the zygotes equals $\sum_j (Z_{ij} + Z_{ji}) = \sum_j (1+\delta_{ij}) \cdot Z^{\circ}_{ij}$, where $\delta_{ii} = 0$ and $\delta_{ij} = 1 = 1$ for $i \neq j$. Therefore, among all zygotes with the i-th individual as parent, a proportion $Z^{\circ}_{ij} / \sum_k (1+\delta_{ik}) \cdot Z^{\circ}_{ik}$ has the j-th individual as second parent, and this parent appears at a proportion of $\sum_k (1+\delta_{jk}) \cdot Z^{\circ}_{jk} / (2 \cdot Z)$ among the $2 \cdot Z$ contributions of all parents to the set of zygotes. By the above definition, one now arrives at a mating preference $U_{j<i}$ of the i-th for the j-th parent, given by

$$U_{j<i} = \frac{Z^{\circ}_{ij}}{\sum_k (1+\delta_{ik}) \cdot Z^{\circ}_{ik}} \Big/ \frac{\sum_k (1+\delta_{jk}) \cdot Z^{\circ}_{jk}}{2 \cdot Z}$$

$$= \frac{2 \cdot Z^{\circ}_{ij} \cdot Z}{\left[\sum_k (1+\delta_{ik}) \cdot Z^{\circ}_{ik} \right] \cdot \left[\sum_k (1+\delta_{jk}) \cdot Z^{\circ}_{jk} \right]}$$

All of the above principles for an analysis of spatial distance as a determinant of sex-specific mating preferences apply equally to the symmetrical preferences.

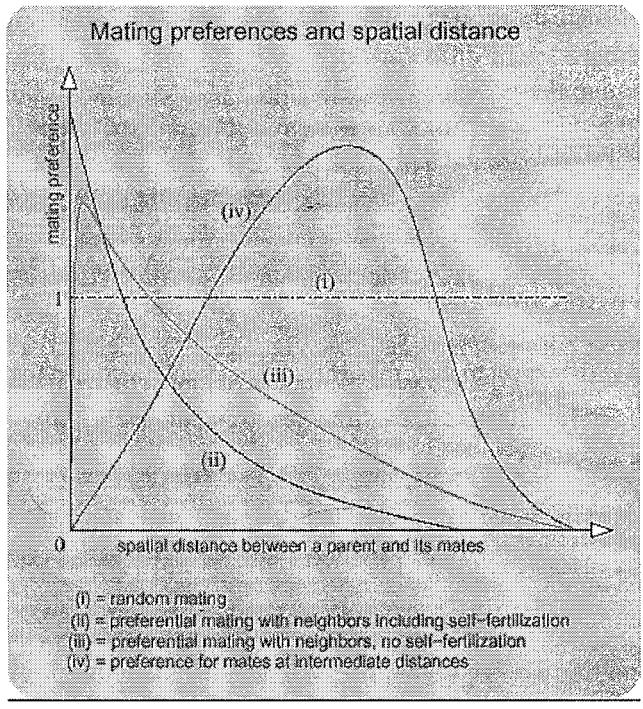


Figure 2. Illustration of isotropic mating preferences as a function of distance between mating partners for different mating systems.

Discrete population structure, due to mating preferences

Treatment of a broader spectrum of problems requires extension of the measurement of mating preferences to arbitrary traits of the parents. The individual parent, which was focused on in the last chapter, is then replaced by the set of all parents with the same trait state. In principle, any of the categories (*i*) to (*v*) of mating system determinants can be treated on this basis. In most cases, the traits of interest are of a discrete type or can be classified into such types so that distinguishable groups or demes of potential parents can be specified. Matings (as determined by the parentage distribution) can then again be characterized by the affiliation of mating partners (parents) to groups, and the frequencies of matings within and between the groups (homodemic and heterodemic matings) can, for example, form the basis for an analysis of differentiation among groups with respect to their mating relations. To provide a conceptual basis for such an analysis, each zygote is now considered to express

two traits, specified by properties of its maternal and paternal parent. Variables X and Y will be used to identify the maternal and paternal trait, respectively.

The basic frequencies which are required for the computation of mating preferences are then given by the frequency $P(Y=y|X=x)$ of zygotes with paternal parents (Y) of type y among all zygotes with maternal parents (X) of type x and the frequency $R(Y=y|X=x)$ of paternal parents of type y among all potential mates of maternal parents of type x. With this notation, the mating preference of maternal parents of type $X=x$ for paternal parents of type $Y=y$ reads

$$U_{Y=y<X=x} = \frac{P(Y=y|X=x)}{R(Y=y|X=x)}.$$

The inverse quantities $P(X=x|Y=y)$, $R(X=x|Y=y)$ and $U_{X=x<Y=y}$ are defined analogously.

If, as was done in the last chapter for individual mating preferences, the potential mates are equated to their overall maternal and paternal contributions to the pool of zygotes, then $R(Y=y|X=x) = P(Y=y)$ and $R(X=x|Y=y) = P(X=x)$, where $P(Y=y)$ and $P(X=x)$ equal the frequencies or probabilities of zygotes with paternal parent (Y) of type y and maternal parent (X) of type x, respectively.

Given the frequencies of the potential mating partners, it follows that the mating preference $U_{Y=y<X=x}$ is bounded from above by $R(Y=y|X=x)^{-1}$ since, $P(Y=y|X=x) \leq 1$. This upper bound is reached if maternal parents of type $X=x$ mate exclusively with males of type $Y=y$, which is indeed the highest preference an individual can realize among its potential mates. Hence, to make more apparent the concept of preference and rejection of potential mates and to allow its quantification over the whole range from complete preference via indifference to complete rejection, it is desirable to normalize the measures U such that they vary symmetrically between +1 and -1. Symmetry around 0 is required to enable comparison of the extents of rejection and preference. This is achieved by the following normalization

$$\hat{U}_{Y=y \lhd X=x} := \begin{cases} U_{Y=y \lhd X=x} - 1 & \text{if } U_{Y=y \lhd X=x} \leq 1 \\ \dfrac{U_{Y=y \lhd X=x} - 1}{R(Y=y \mid X=x)^{-1} - 1} & \text{if } U_{Y=y \lhd X=x} > 1 \end{cases}$$

As desired, $\hat{U}_{Y=y<X=x} = -1$ for complete rejection of $Y=y$ mates by $X=x$ females ($U_{Y=y<X=x} = 0$), $\hat{U}_{Y=y<X=x} = 1$ for exclusive mating of $X=x$ females with $Y=y$ males, and $\hat{U}_{Y=y<X=x} = 0$ for indifference of $X=x$ females towards $Y=y$ males (random mating, $U_{Y=y<X=x} = 1$).

$\hat{U}$ can be written in a more compact form if one considers that for $U_{Y=y<X=x} > 1$, one obtains

$$\hat{U}_{Y=y \lhd X=x} = \frac{R(Y=y \mid X=x) \cdot (U_{Y=y \lhd X=x} - 1)}{1 - R(Y=y \mid X=x)}$$
$$= \frac{P(Y=y \mid X=x) - R(Y=y \mid X=x)}{R(Y \neq y \mid X=x)}$$
$$= \frac{R(Y \neq y \mid X=x) - P(Y \neq y \mid X=x)}{R(Y \neq y \mid X=x)} = 1 - U_{Y \neq y \lhd X=x}.$$

With the help of this equation, $\hat{U}$ can be rewritten in the form

$$\hat{U}_{Y=y<X=x} = \max\{1 - U_{Y \neq y<X=x}, 0\} - \max\{1 - U_{Y=y<X=x}, 0\}.$$

The inverse normalized preferences are defined analogously.

The preferences U are symmetrical for potential mates, given by the actual mating frequencies $P(Y=y)$ and $P(X=x)$. In contrast, $\hat{U}$ is symmetrical for these potential mates only when it is negative. Otherwise, both $\hat{U}_{Y=y<X=x}$ and $\hat{U}_{X=x<Y=y}$ are positive, and symmetry is realized only for $P(Y=y) = P(X=x)$.

In the following examples of application, in which population structure is considered solely for the reproducing members of populations, frequencies of potential mates will be assumed to equal their actual mating frequencies.

Metapopulations: reproductive isolation and coherence

In particular, if only membership of a group is relevant for maternal parents X and paternal parents Y, and if groups are denoted by z, then $\hat{U}_{Y=z<X=z}$ specifies the degree to which maternal parents from group z prefer to mate with paternal parents from the same group (homodemic mating preferences). Positive values of $\hat{U}_{Y=z<X=z}$ state that, on the average, maternal parents from group z prefer matings with paternal parents from their own group over matings with paternal parents from outside the group. This is reversed for negative values of $\hat{U}_{Y=z<X=z}$. Analogous statements hold for the preferences $\hat{U}_{Y=z<X=z}$ of paternal for maternal parents. In other words, group z is reproductively isolated from other groups and thus forms a subpopulation, to the degree to which both homodemic preferences $\hat{U}_{Y=z<X=z}$ and $\hat{U}_{X=z<Y=z}$ are positive.

Hence, if the $\hat{U}_{Y=z<X=z}$ and $\hat{U}_{X=z<Y=z}$ are strictly positive for all groups z, a clear tendency towards formation of a metapopulation, with respect to the chosen group structure, can be stated. Looking at metapopulation structures from the opposite point of view, i.e. in terms of degrees to which matings are performed among groups, the reproductive coherence or gene flow among groups is to be quantified. The relevant measures are the heterodemic preferences $\hat{U}_{Y\neq z<X=z}$ and $\hat{U}_{X\neq z<Y=z}$. The exact complementarity of both views is reflected by the mathematical identity $\hat{U}_{Y=z<X=x} = -\hat{U}_{Y\neq y<X=x}$. Positive values for $\hat{U}_{Y\neq z<X=z}$ and $\hat{U}_{X\neq z<Y=z}$ therefore indicate the absence of metapopulation structures for the chosen grouping criterion.

To arrive at a single *measure of metapopulation structure*, the differences between the sexes must again be taken into account. This suggests that we should distinguish between the average homodemic mating preferences of maternal for paternal parents, i.e. $\sum_z \hat{U}_{Y=z<X=z} \cdot P(X=z)$, and of paternal for maternal parents, i.e. $\sum_z \hat{U}_{X=z<Y=z} \cdot P(Y=z)$.

The special case of self-fertilization

When group structure is broken down to the level of the individual, so that each individual is considered a group of its own, the above homodemic mating preferences correspond to self-fertilization. It is thus meaningful in this case to talk about a measure of genealogical "self-preference" (the computation of U follows in this case the rules stated in the previous chapter, in connection with effects of spatial distance on mating preferences). The asymmetry in the measure $\hat{U}$ takes care of the possibility that self-fertilization has a female and a male component which may differ. Thus $\hat{U}_{Y=z<X=z} > \hat{U}_{X=z<Y=z} > 0$ states that parent z self-fertilizes more than at random and it does so to a larger degree as a maternal than as a paternal parent. At an extreme, when parent z is completely self-incompatible or produces for other reasons no offspring by self-fertilization, one obtains $\hat{U}_{Y=z<X=z} = \hat{U}_{X=z<Y=z} = -1$, which correctly reflects the implied complete rejection of homodemic matings.

Note that negative values of self-preference do not exclude self-fertilization but rather state that the reference individual mates more frequently with other individuals than with itself. In other words, despite a partial self-incompatibility, estimates of the proportion of offspring resulting from self-fertilization may be positive. This is an important mating system characteristic that does not show up in the common estimates of proportions of self-fertilization. Estimates of

large proportions of self-fertilization may also be the result of random fertilization combined with low reproductively effective population sizes. In this case, high proportions of self-fertilization would be due to non-intact performance of the second fundamental function of mating systems. The third function is not directly affected by low reproductively effective population sizes. The inbreeding depression that might show up in the next generation is not due to the mating preferences but rather to the loss of adaptational capacity due to the drift effects associated with the small reproductive population size. Misinterpretations of this kind are, in fact, ruled out by consideration of the values of self-preference, since these would be zero because of the random fusion of the gametes.

Indirect (model-dependent) estimation

In the introductory remarks, it was indicated that the legitimacy of indirect estimates rests on tests of the validity of the underlying model. Models for which no experimental methods of testing are available are of limited practical relevance. This applies particularly to models which are used for the estimation of characteristics of real systems. Thus, if the mixed mating model is used in the estimation of the proportion of self-fertilization, the estimate is arbitrary if no test as to the validity of the model was performed. Estimates obtained from models which did not pass a statistical test are without substance. The estimation of amounts of gene flow among populations with the help of F_{ST} was mentioned in this context.

It is, of course, common practice to discuss unexpected indirect estimates in terms of the validity of the underlying model. Occasionally, sensitivity analyses are performed to rule out the possibility of substantial misinterpretation of data on the basis of the model. Most frequently, however, detailed arguments refer to more tractable problems, such as those arising from sample variance of estimators or from estimation algorithms connected with the respective model. Problems of testing the validity of the applied model are then of lesser concern (a more recent account of this situation is given e.g., by Ivey & Wyatt, 1999). This is at odds with the system analytic requirement for the joint consideration of model-dependent parameter estimation and model testing (see e.g. the paper of Gregorius, 1999, which demonstrates this requirement for the non-equilibrium and equilibrium version of the mixed mating model; a more comprehensive account of the system analytic approach is found in Gregorius 1998).

Intactness of mating mechanisms

The utilization of model-dependent methods of estimating mating system characteristics enforces integration of elements of parameter estimation and

intactness analysis in the same model. This should be considered in the model design. For example, if the classical mixed mating model (selfing and random cross-fertilization) is used for the estimation of individual selfing rates, and if these estimates are used in a model of the effects of neighbourhood mating including selfing, the assumptions on the form of cross-fertilization are contradictory between the two models. Inferences as to the intactness of the neighbourhood mating mechanisms can thus not be consistently made, even if the neighbourhood mating model would include the information on mating success required in an analysis if intactness.

Similar caution has to be taken with experiments in which the parentage (usually paternity) analysis is restricted to offspring of special origin such as seed from a few seed trees. Since intactness inferences of mating system mechanisms are to be based on samples that are representative of the population's offspring (zygote) production, the seed tree sample should represent the population, and the seed samples should represent each tree's seed production. Particularly the representativity of seed samples of the individual trees' seed production is frequently difficult to realize. This problem is more easily settled if seed is sampled after dispersal. Yet, this is achieved at the expense of reliable information on the maternal parent. However, as long as affiliation of seeds to population is unambiguous and representativity is guaranteed, the difference between sampling strategies affects the analysis only through statistical precision.

Acknowledgment

The author is indebted to Bernd Degen for providing the means to present and discuss this paper at the 2000 Symposium: "Modelling and Experimental Research on Genetic Processes in Tropical and Temperate Forests". The thoughtful suggestions and linguistic advice of Peter Smouse helped to improve the argumentation of the paper.

References

Adams, W.T, Birkes, D.S. (1991): Estimating mating patterns in forest tree populations. pp. 157-172 in Fineschi et al. 1991

Fineschi, S., Malvolti, M.E., Cannata, F., Hattemer, H. H. (Eds.) (1991): Biochemical markers in the population genetics of forest trees. SPB Academic Publishing bv

Gregorius, H.-R. (1978): The concept of genetic diversity and its formal relationship to heterozygosity and genetic distance. Math. Biosci. 41: 253-271

Gregorius, H.-R. (Ed) (1985): Population Genetics in Forestry. Lecture Notes in Biomathematics 60, Springer-Verlag

Gregorius, H.-R. (1989): Characterization and analysis of mating systems. Ekopan Verlag, Witzenhausen, Germany. http://webdoc.sub.gwdg.de/ebook/y/2001/gregorius/matesys.pdf

Gregorius, H.-R. (1998): The system analytical approach to the study of hypotheses. http://webdoc.gwdg.de/ebook/y/1998/gregorius/sysana.pdf

Gregorius, H.-R. (1999): Qualified testing of models and character estimation: the system analytic approach. Grüne Reihe der Sektion Biometrie und Informatik des Deutschen Verbandes Forstlicher Forschungsanstalten, 11. Jahrestagung: 253-260

Gregorius, H.-R., Degen, B. (1994): Estimating the extent of natural selection in seedlings from different *Fagus sylvatica* (L.) populations: Application of new measures. J. of Heredity 85: 183-190

Gregorius, H.-R,. Ziehe, M., Ross, M.D. (1987): Selection caused by self-fertilization. I. Four measures of self fertilization and their effects on fitness. Theor. Pop. Biol. 31: 91-115

Haldane, J.B.S. (1954): The measurement of natural selection. Proc. 9th Int. Congr. Genetics: 480-487

Hamrick, J.L., Schnabel, A. (1985): Understanding the genetic structure of plant populations: Some old problems and a new approach. Pp. 50-70 in Gregorius 1985

Ivey, C.T., Wyatt, R. (1999): Family outcrossing rates and neighborhood floral density in natural populations of swamp milkweed (*Asclepias incarnata*): potential statistical artefacts. Theor. Appl. Genet. 98: 1063-1071

Lloyd, D.G. (1979): Some reproductive factors affecting the selection of self-fertilization in plants. Am. Nat. 113: 67-79

Schoen, D. J, Brown A.H.D. (1991): Intraspecific variation in population gene diversity and effective population size correlates with the mating system in plants. Proc. Natl. Acad. Sci. USA 88: 4494-4497

Steiner, W., Gregorius, H.-R. (1997): Reinforcement of genetic coherence: a single-locus model. BioSystems 43: 137-144

Wallace, A. R. (1889): Darwinism - An exposition of the theory of natural selection with some of its applications. Macmillan, London

Weir, B.S. (1996): Genetic Data Analysis II. Sinauer Associates, Sunderland

Whitlock, M.C., McCauley, D.E. (1999): Indirect measures of gene flow and migration: $F_{ST} \neq 1/(4Nm + 1)$. Heredity 82: 117-125

Willson, M.F. (1994): Sexual selection in plants: Perspective and overview. Am. Nat. 144: 13-39

Ziehe, M., Gregorius, H.-R. (1988): Selection caused by self-fertilization. II. Ecological constraints on selfing advantage. J. evol. Biol. 1: 233-253

Appendix: *Selection load*

A concept of viability selection load generalizing the classical approach of Haldane (1954) was developed by Gregorius and Degen (1994). An analogous generalization of this concept to include reproduction can be obtained by considering all adult individuals to have the same maximum ability (capacity) to reproduce. Reproductive output (number of offspring) can be measured in terms of zygotes, gametes, or successful gametes. Whenever the actual number of offspring of an individual falls below its reproductive capacity, impairment of

the reproductive output due to environmental challenges or mating system effects can be stated. Making use of the notation

n_i := number of i-type adults
f_i := number of offspring of i-type adults
n := $\Sigma_i\, n_i$
f := $\Sigma_i\, f_i$
p_i := n_i / n
p_i' := f_i / f
c := individual capacity of reproduction in terms of numbers of offspring,

the absolute reproductive impairment of the i-th type amounts to $n_i \cdot c - f_i$ offspring. By definition, c is always sufficiently large to assure non-negativity of this difference for all types. Hence, $c \geq f_i/n_i$ or all i, so that

$$c \geq \max_i \frac{f_i}{n_i} = \frac{f}{n} \cdot \max_i \frac{p_i'}{p_i}$$

It follows that the minimum reproductive capacity $c = c^*$, which must have been realized per individual to produce the observed numbers of offspring of the various types, equals $c^* = \max_i(f_i / n_i)$. The absolute total reproductive impairment across all types sums to $\sum_i (n_i \cdot c - f_i) = n \cdot c - f$, so that the relative impairment yields a fraction $\dfrac{n \cdot c - f}{n \cdot c}$ by which the total reproductive output is reduced relative to the total reproductive capacity. Since c^* is the minimum individual reproductive capacity that must have been realized to explain the differences in reproductive output between the types, c^* must be substituted for c to obtain the minimum relative reproductive impairment. One therefore arrives at

$$L_R := \frac{n \cdot c^* - f}{n \cdot c^*} = 1 - \frac{f}{n \cdot c^*} = 1 - \frac{1}{\max_i \dfrac{p_i'}{p_i}} = 1 - \min \frac{p_i}{p_i'}$$

as the minimum reduction in total reproductive output required to arrive at the realized differences in reproductive output between types. The fraction L_R is thus meaningfully addressed as the *reproduction selection load* to distinguish it from the viability selection load.

Comparison with the viability selection load L as stated in Gregorius and Degen (1994) shows that L is formally identical to the present reproduction load L_R, if the number of offspring of a type is measured by the number of individuals of this type remaining after selection. Hence, the above concept of reproduction selection load can indeed be generalized to include all vegetative and generative stages. One only has to refer the n_i's and f_i's to any two successive developmental stages, the first of which being a vegetative (including zygotic) and the second a vegetative or generative stage. This justifies denotation of L_R simply as the selection load.

Using oaks to compare microsatellites and AFLP markers for parentage analysis

S. GERBER, S. MARIETTE, R. STREIFF[1], C. BODÉNÈS and A. KREMER

INRA, Laboratoire de génétique et amélioration des arbres forestiers BP 45, 33611 Gazinet Cedex, France [1]Present address: INRA - URLB, Laboratoire de Modélisation et de Biologie Evolutive 488 rue de la Croix-Lavit, 34000 Montpellier, France

Key words: gene flow, markers, dominant markers, codominant markers

Abstract: Our aim was to compare the properties of dominant markers, like AFLPs, to those of codominant multiallelic markers, like microsatellites, in reconstructing parentage. These two kind of markers were used to look for both parents of an individual without prior knowledge of their relationships, by calculating likelihood ratios based on genotypic data, including mistyping or not. Experimental data on 89 oak trees genotyped for 6 microsatellite markers and 159 polymorphic AFLP loci were used as a starting point for simulations and tests. Both sets of markers produce high exclusion probabilities, and among dominant markers those with dominant allele frequencies in the range 0.1 to 0.4 are more informative. Such codominant and dominant markers can be used to construct powerful statistical tests to decide whether a genotyped individual (or two individuals) can be considered as the true parent (or parent pair). Gene flow from outside the study stand (GFO) inferred from parentage analysis with microsatellites overestimates the true GFO whereas with AFLPs, it is underestimated. As expected, dominant markers are less efficient than codominant markers to achieve this task, but can still be used with good confidence, especially when loci are purposely selected according to their allele frequencies.

The results presented in this paper have been published in Molecular Ecology (Gerber, S., Streiff, R., Bodénès, C., Mariette, S., Kremer, A. (2000): Comparison of microsatellites and AFLP markers for parentage analysis. 9: 1037-1048).

Introduction

Gene flow is an important feature of population genetics, shaping diversity of species. Contemporary gene flow can be studied with the help of genetic markers, by reconstructing relationships between parental and offspring generations (i.e. paternity or parentage analysis). Paternity or parentage assignment can be achieved by any type of genetic marker provided it is polymorphic enough. For that reason, microsatellites are usually be preferred to isozymes. The development of microsatellite markers for a given species is an expensive task. Markers based on random amplification of DNA fragments, like random amplified

polymorphism DNA (RAPD) or amplified fragment length polymorphism (AFLP) are easier to perform, but show dominant-recessive inheritance.

To reconstruct parentage, genotypes can be simply compared. However, when several loci and many potential parents are available, statistical analysis of the results are necessary. When paternity and maternity are analysed at the same time, likelihoods can be calculated, as described for instance in Meagher and Thompson (1986) for codominant markers. In the present study, our aim was to extend statistical analysis to dominant markers and thus to compare codominant microsatellite to dominant AFLP markers for parentage analysis in a population.

The same set of oak trees (*Quercus petraea and Q. robur*) genotyped for both types of markers was used as a starting point for simulations to compare the ability of codominant and dominant markers to reconstruct parentage. Looking for both parents of a given offspring, LOD score (log-likelihood ratio) calculations and statistical approaches were used. Since only a small subset of all potential parents were available and genotyped, a rationale had to be set to decide whether a given individual could be considered as a true parent or not and whether a given pair of individuals could be considered as a true parent pair or not. For this purpose, simulations based on a theoretical large, random mating population were performed. This allowed us to build empirical statistical tests minimizing both type I and II errors. It also permitted to measure the ability of those tests to make the correct decisions concerning parentage, and to compare their impact on the evaluation of gene flow with both types of markers.

Plant material and molecular markers

The white oak trees (*Quercus petraea*, *Quercus robur*) of the present study are located on a 5.76 ha stand in the northwest of France (part of the "Petite Charnie" forest (Streiff et al. 1999). The 296 mature trees (about 100-years old) originated from natural regeneration, and 89 of them were included in the present study. The 89 trees of the study have been genotyped for 6 microsatellite loci by Streiff et al. (1998). The same 89 trees were genotyped for 214 AFLP loci, among which 159 were polymorphic.

Exclusion probabilities

An exclusion probability can be defined as the average capability of any marker system to exclude any given relationship. Three types of exclusion probabilities can be calculated. The most commonly used concerns paternity, in which a mother-offspring pair is compared to a potential father. The second concerns a

single parent compared to an offspring (without any information on the other parent), and the third concerns a pair of potential parents compared to an offspring. Exclusion probabilities formulae are given in Table 1 for codominant markers (Jamieson and Taylor 1997) and dominant markers.

Table 1. Exclusion probabilities for one locus. For K independent loci, the overall exclusion probabilities are calculated as: $P = 1 - \prod_{j=1}^{K}(1 - P_j)$

Marker type	Exclusion probabilities P_j		
	paternity	single parent	parent pair
Codominant	For a locus with n alleles (allele i in frequency p_i) let a_k be : $a_k = \sum_{i=1}^{n} p_i^k$		
	$1 - 2a_2 + a_3 + 2a_4 - 3a_5 - 2a_2^2 + 3a_2a_3$	$1 - 4a_2 + 2a_2^2 + 4a_3 - 3a_4$	$1 + 4a_4 - 4a_5 - 3a_6 - 8a_2^2 + 8a_2a_3 + 2a_3^2$
Dominant	Two phenotypes [+] and [-] at the locus, p: frequency of the dominant allele +		
	$p(1-p)^4$	always zero	$p(2-p)(1-p)^4$

The exclusion probabilities computed for 30 theoretical AFLP loci, where p frequencies (in the range 0 to 1) of the «presence» allele at each locus were equal are given in Figure 1. The exclusion probabilities have their highest values for p varying between 0.1 and 0.4. Among the 159 polymorphic AFLP loci of our sample, 45 loci (28.3%) exhibited p values in the range [0.1, 0.4]. Comparing exclusion probabilities, this subset of loci is almost as informative as the total set of loci, showing that the different loci do not contribute equally to the total exclusion capacity.

The exclusion probabilities are higher for microsatellites, and the single parent exclusion probability is only available for this type of markers (0.998233), because neither combination of parent-offspring AFLP phenotypes allows exclusion. However, probabilities are high and very similar for both types of markers (paternity exclusion is 0.996820 for AFLPs and 0.999910 for microsatellites; pair exclusion is 0.999959 for AFLPs and 1.000000 for microsatellites).

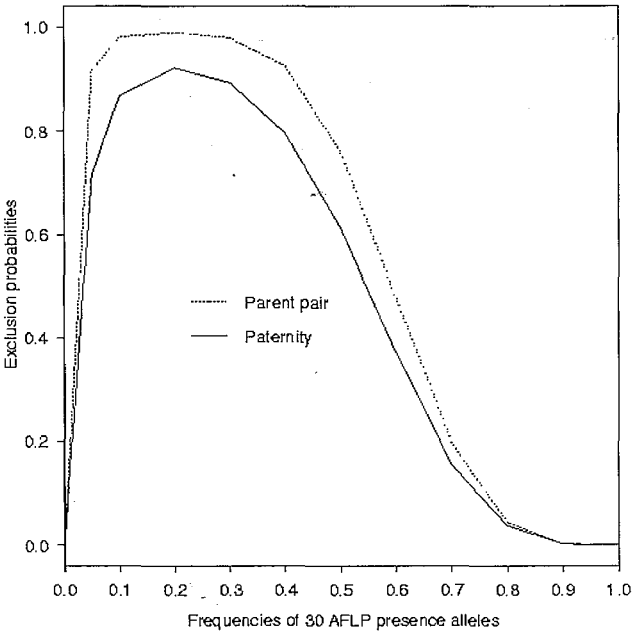


Figure 1 .Exclusion probabilities cumulated over 30 AFLP loci, assuming the same frequency for the "presence" allele at each locus

Likelihood ratio and simulation

When genotypes are available for a set of offspring and a set of potential parents, the most probable single parents and parent pairs can be determined by the calculation of log-likelihood ratios (LOD scores). These ratios compare the likelihood of an individual (or a couple) being the parent (or the parent pair) of a given offspring divided by the likelihood of these two individuals (or three individuals) being unrelated. These LOD scores are additive over independently inherited loci.

The impact of scoring error was taken into account by introducing the probability of error in genotyping (mistyping), into the LOD score calculation (Marshall et al. 1998).

Simulations were used first to build empirical statistical tests to decide whether a given tree (or pair of trees) would be chosen as the true parent (or true parent pair), following the scheme drawn in Figure 2. For both types of markers, 10,000 offspring were generated with both parents either inside or outside the stand.

The two potential parents and the parent pair from the stand giving the highest LOD scores were recorded and the distributions of these LOD scores were plotted. In order to minimize type I and type II errors, thresholds of LOD scores were chosen at the intersection of the two distributions.

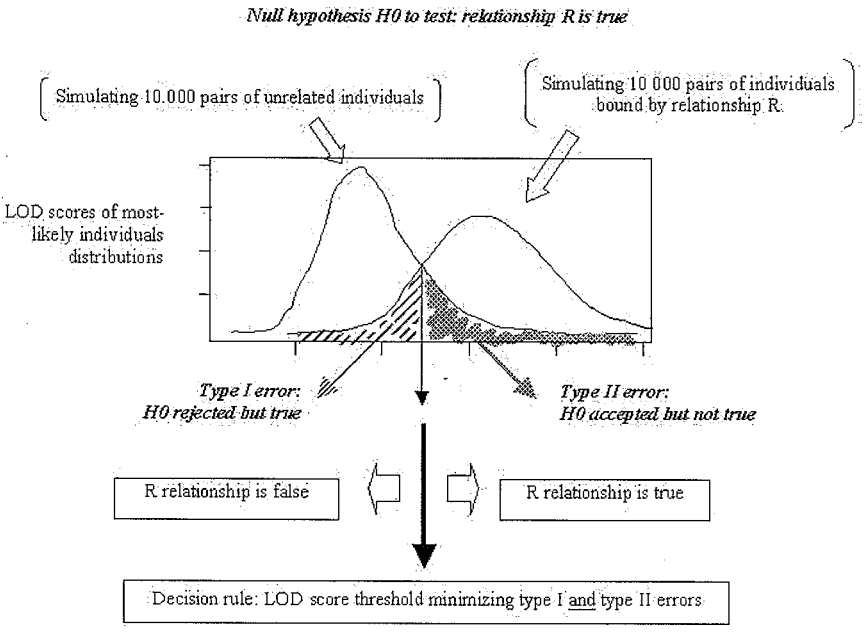


Figure 2. Using simulation for building an empirical statistical test to make a decision concerning relationship R

Simulations were used secondly to measure the ability of those tests to make the correct decisions concerning parentage and thirdly to compare the impact of the tests on the evaluation of gene flow. For these two objectives, we simulated a "true" population, where each offspring could have zero, one or two parents among the genotyped individuals. Ten thousand offspring were generated by picking their parent randomly among a population of N trees (N=500 or N=1,000). Among these N trees, the first 89 were our genotyped trees. If the random [1:N] number exceeded 89, the alleles of the offspring were randomly chosen among the alleles of the locus considered according to their frequencies. Our empirical test was applied to these data, and the number of correct decisions provided by the test was recorded.

The percentage of correct decisions increases slightly with the size N of the total reproducing population, but decreases with a non-zero mistyping (Table 2). Microsatellites are more affected by mistyping, but the choise of parent pair based on AFLP markers is also affected. Nonetheless, parent pairs are correctly chosen in nearly 100% of the cases, whatever the marker type. The smallest percentages of correct decisions are close to 90%.

Table 2 Results of the test simulations, according to the mistyping rate and to the size *N* of the simulated population for both kinds of markers. [a]The decision made by the test corresponds to the true situation.

			Correct decisions (%) [a]	
	Mistyping	*N*	Microsatellites	AFLPs
Single parent	0%	500	95.04	89.05
		1000	96.7	91.27
Single parent	0.1%	500	91.20	89.31
		1000	91.26	90.69
Parent pair	0%	500	100	99.65
		1000	100	99.77
Parent pair	0.1%	500	99.8	94.39
		1000	99.93	96.69

With the same kind of simulations, the test was applied and a decision was made assigning, for each offspring, zero, one or two parents among the genotyped individuals. For each simulated set we deduced from the results the different gene flow events observed:

- true gene flow from outside the study stand (GFO) events, i.e. the actual number of times a parent from outside the stand produced one of the offspring,
- apparent GFO events, i.e. according to the results of the statistical tests, the number of times neither parent from inside the stand was detected for the simulated offspring.
- cryptic gene flow events, i.e. according to the results of the statistical tests, the number of parents that had been detected inside the stand whereas the true parents were outside the stand.

With microsatellites, apparent GFO overestimated true GFO by 4.8% on average (Table 3). This overestimation decreases with the increase in population size N, and decreases with the increase in mistyping. The situation is different with AFLPs; the true GFO is underestimated by the apparent GFO by 4.3% on

average. This underestimation increases with the increase in population size N, and increases with the increase in mistyping. The cryptic gene flow is smaller for microsatellites and decreases for a non-zero mistyping (1% to 0.4%). Cryptic gene flow averaged 0.76 % of the apparent GFO for microsatellites and 6.99 % for AFLPs. There is a systematic overestimation of true GFO by apparent GFO with microsatellite (+ 4.8%), and a systematic underestimation of true GFO by apparent GFO with AFLPs (- 4.3%). This underestimation increases with an increase in population size, and with an increase in mistyping. Attributing a parent inside the stand to an offspring is more likely with AFLP marker tests, creating an apparent gene flow from inside the stand event, and decreasing apparent GFO. Microsatellite marker tests give rise to opposite results.

Table 3. Estimation of gene flow from outside the stand (GFO) for 10,000 simulated offspring.

			Number of GFO events					
Markers	Mis-typing	Population size *N*	expected	true (*t*) [b]	apparent (*a*) [c]	% *a/t*	cryptic (*c*) [d]	% *c/a*
Microsatellites	0%	500	16,440	16,464	17,052	3,57	172	1.01
AFLPs	0%	500	16,440	16,408	16,128	-1,71	843	5.23
Microsatellites	0%	1,000	18,220	18,263	18,439	0,96	220	1.19
AFLPs	0%	1,000	18,220	18,208	17,414	-4,36	1,054	6.05
Microsatellites	0.1%	500	16,440	16,491	18,177	10,22	70	0.39
AFLPs	0.1%	500	16,440	16,395	15,543	-5,20	1,390	8.94
Microsatellites	0.1%	1,000	18,220	18,163	18,973	4,46	81	0.43
AFLPs	0.1%	1,000	18,220	18,205	17,134	-5,88	1,327	7.74

[a] $\frac{N-89}{N} \times 20{,}000$

[b] actual number of times a parent from outside the stand produced one of the offspring

[c] number of times neither parent from inside the stand was detected for the simulated offspring, according to the statistical tests.

[d] number of times a parent had been detected inside the stand whereas the true parent was outside the stand, according to the statistical tests.

Conclusion

Most paternity or parentage studies are conducted using codominant markers, like microsatellites. The present contribution shows that parentage analysis can be performed with dominant markers, like AFLPs, with favourable results.

The exclusion probabilities provided by our dominant markers are very high, a prerequisite to any parentage analysis, and are close to those provided by codominant markers. Thus, in the present study, fewer than 10 highly polymorphic

codominant loci or between one and two hundred dominant loci (and even less if they are first selected according to their allele frequencies : see Figure 1) are adequate for parentage studies

As mistyping is very likely to occur, LOD score calculation should include it, but at a low rate.

The tests performed on simulated data allowed a large proportion of correct decisions to be made, with zero, one or two parents of a given offspring being correctly assigned in the great majority of cases. The proportion of correctly identified parent pairs was always high, whatever the type of marker, codominant or dominant.

Microsatellites, characterised by high levels of codominant polymorphism, give rise to highly accurate assignments of parentage. Comparatively, the loss of information encountered with AFLPs due to dominance is counterbalanced by a high number of polymorphic loci. The methods used in this paper could be applied in other situations where codominant or dominant markers will probably have different levels of polymorphism related to the breeding size, to the mating system, and to the history of the population. These levels determine exclusion probabilities, and hence the ability of resolving parentage. In each population, simulations have to be performed to determine the thresholds to be used, the associated error levels and power of the tests, the impact of scoring errors, and the cryptic gene flow. A software package that is currently being developed will allow users to compute these different elements.

References

Gerber, S., Streiff, R., Bodénès, C., Mariette, S., Kremer, A. (2000): Comparison of microsatellites and AFLP markers for parentage analysis. Molecular Ecology 9: 1037-1048

Jamieson, A., Taylor, SS. (1997): Comparisons of three probability formulae for parentage exclusion. Animal Genetics, 28, 397-400.

Marshall, T.C., Slate, J., Kruuk L. E. B., Pemberton J.M. (1998): Statistical confidence for likelihood-based paternity inference in natural populations. Molecular Ecology, 7, 639-655

Meagher, T.R., Thompson, E. (1986): The relationship between single parent and parent pair genetic likelihoods in genealogy reconstruction. Theoretical Population Biology, 29, 87-106

Streiff, R., Ducousso, A., Lexer, C., Steinkellner, H., Gloessl, J., Kremer, A. (1999): Pollen dispersal inferred from paternity analysis in a mixed stand of *Quercus robur* L. and *Quercus petraea* (Matt.) Liebl. Molecular Ecology, 8, 831-841

Streiff, R., Labbe, T., Bacilieri, R., Steinkellner, H., Glossl, J., Kremer, A. (1998): Within-population genetic structure in *Quercus robur* L. and *Quercus petraea* (Matt.) Liebl. assessed with isozymes and microsatellites. Molecular Ecology, 7, 317-328

Genetic conservation within managed forests in Amazonia: Dendrogene project

M. KANASHIRO[1], I. S. THOMPSON[2] and B. DEGEN[3]

Address: [1]Embrapa Amazonia Oriental, C.P. 48, 66095-100 Belém-PA, Brazil; [2]United Kingdom Department for International Development (DFID), C.P. 48, 66095-100 Belém-PA, Brazil; [3]UMR CIRAD INRA ENGREF, Campus agronomique, BP 709, 9738 - Kourou cedex, French Guiana

Key words: Genetic diversity, harvesting planning, logging impact, reproductive biology, species identification, tropical trees

Abstract: The genetic structure of tropical tree species forms the basis for adaptation to present environmental conditions. Moreover, genetic variation improves the adaptability to future conditions. Hence, conservation of forest genetic resources is an important issue for sustainable forest management in Amazônia. This paper describes an applied research project which will integrate indicators for genetically sustainable management into forest logging plans. The novelty of the project, entitled Dendrogene, is the interdisciplinary approach which results from the integration of different software packages: a simulation model (ECO-GENE) to study temporal and spatial dynamics of genetic structures of tree populations; a database (DENDROBASE) summarizing the ecological and genetic information known for tropical tree species; and tree mapping and forest management utilities (TREMA) using the concept of minimum stock by species to influence harvest tree selection. The DENDROBASE application will identify groups of species with similar genetic systems. For these groups simulation studies with ECO-GENE will help to estimate the dynamics of genetic structures for different logging scenarios. Information on species groups and their responses to different logging intensities will be added to TREMA for further integration into logging plans. Perspectives on implementation of genetic indicators within management plans are discussed.

Parts of this paper represent an extract from the publication of Kanashiro et al. (2002) in Unasylva (Vol. 52), No. 2, (Issue No. 209): http://www.fao.org/forestry/FODA/UNASYLVA/unasyl-e.stm

Introduction

Although tropical lowland rain forests cover only 6 to 7 per cent of the Earth land's surface, they accommodate more than 50 percent (possibly as much as 90 per cent) of all living species (Linsenmair 1997). Amazonia is acknowledged to be one of the richest and most biologically diverse forest ecosystems in the world. In Brazil alone the Amazon forest encompasses 3.5 million km^2. Despite the enormous international pressure due to rates of deforestation and intensive

selective logging operations, more than 2500 mills were operating in 1996/97, consuming an equivalent of 27.8 millions m^3 of tropical timber (Nepstad et al. 1999).

In the last two decades the forest legislation for Brazilian Amazon forests has evolved considerably, yet many advances still need to be implemented either at a technical and/or an operational level (Zachow 1996). The current Sustainable Forest Management Plans (SFMP) still do not take into account the characteristics of individual commercially-harvested species, although an understanding of their reproductive traits is increasingly well-studied, at least for some timber tree species.

In a recent evaluation process, IBAMA, (The Brazilian Environmental Agency) determined that, out of 4.204 Sustainable Forest Management projects, only 1759 (41.8%) were properly prepared, and thus approved to continue (www.ibama.gov.br). About the same number were temporarily interrupted, for not presenting appropriate documents, and others were completely cancelled. However, there are few initiatives trying to show that tropical hardwood can be economically harvested and managed on a sustainable basis such as promoting low impact logging and selection criteria (Blate 1997; Silva 1997; Kageyama 1998; Bihun 1999). As noted by Thompson and Yared (1997), in addition to public pressure for sound environmental practices, and in particular sustainable management of tropical forests, there is an increasing demand for practical measures which can be adopted by companies and monitored by regulatory authorities, expecting to move away from destructive logging practices towards sustainable management

More recently timber certification as a means of furthering sustainable forestry has been intensively debated within many countries as well as at the international level. Although the creation of incentives for moving towards a wise and sustainable use of forest resources are positive steps, we still lack much basic information to inform such attempts at sustainable use (Lisenmair 1997). Several initiatives towards discussing, developing, and testing criteria and indicators are under way (van Bueren and Blom 1997; Flinn and Franc 1998). However ecological criteria directed towards processes which maintain genetic variation were recently waived in field testing due to the difficulties of fair assessment in monitoring and auditing activities (Sabogal 1999, personal communication). Forest certification may increase the product value provided the forest operations are approved for certification, but it still may not guarantee genetic sustainability, herein taken as the maintenance of intra-and inter-population species diversity as a basis for future adaptation.

Szaro (1998) reports that there is a considerable body of knowledge available on tropical forests, but it needs to be synthesised and put into a form usable by managers and policy makers. The first step is to ensure that what we know is

used, and the second is to prioritise where additional resources for research can make a critical contribution to sustainable forest management and to a global dialogue on forest policy. Namkoong (1998) stressed that forest scientists needed to make special effort to come together and communicate effectively. He noted that forest policy was not interested in complexity and was often based on little scientific information. In that respect, modelling information provided by the adapted version of ECO-GENE and parameter estimates and opportunities for model verification for predictive purposes, will be important to link forest practices to their effects on the parameters of population size, inbreeding, selection and migration that can be experimentally estimated. Forest management decisions can then be made on the basis of the effects of practices on evolution through their effects on genetic processes, and this may represent one of the major contributions of Dendrogene to sustainable forest management decisions.

In this paper we give a general description of the Dendrogene project which tries to integrate indicators for genetically sustainable management into forest logging plans. We introduce the different software packages which are important tools of the integrative approach of the project. Thereafter we comment on some perspectives on implementation of genetic indicators within management plans.

The *Dendrogene* project

Hosted at the Embrapa Eastern Amazon research station in Belém, Pará, Brazil, the Dendrogene project depends on a multidisciplinary approach and multi-institutional participation. The Department for International Development (DFID) of the United Kingdom supports the project (2000 to 2004) through the Brazil-United Kingdom Technical Assistance Programme. Many of the initiatives in Dendrogene are based on earlier activities in the Rainforest Silviculture Research Project (1993 to 1998), also supported by DFID.

The project focus is to develop mechanisms to apply scientific knowledge (species composition, reproductive health and genetic diversity of populations) to promote sustainable rainforest management in the Brazilian Amazon. The idea is to link forest management in the field to ongoing scientific research, as an attempt to contribute to a goal of achieving sustained use and conservation of genetic resources in the region's natural forest.

One of the novelties of this approach is the integration of several software packages such as ECO-GENE, DENDROBASE and TREMA, in order to generate meaningful information to be translated into sustainable forest management at an operational level.

ECO-GENE simulation model

The temporal and spatial dynamics of genetic structures of tree populations are the results of different processes like gene flow, mating system, selection and random genetic drift. Logging and forest fragmentation influence several processes simultaneously. Thus their overall impact on genetic variation of tropical tree species requires the understanding of the dynamic within a complex system. Modelling and simulation studies are very helpful to analyse complex systems. The simulation model ECO-GENE has been developed to study temporal and spatial dynamics of genetic structures of tree populations (Degen et al. 1996). Overlapping or separated generations can be created and different processes like gene flow, mating systems, selection, random drift can be implemented. The model can be run with empirical and fictitious input data. It have been successfully applied to study the impact of different silvicultural practices and the effect of air pollution on the genetic structure of temperate tree populations (Degen and Scholz 1996; Degen et al. 1997; Degen and Scholz 1998; Geburek and Mengel 1998). Since 1998 it have been adapted to the specific conditions of tropical tree species. Modules on pollen and seed dispersal, and a new module on flowering phenology have been added to ECO-GENE. The integration of the forest simulation model Symfor (www.symfor.org) offers improved modelling of growth processes and management impacts.

The ECO-GENE model will be validated through comparison with real data collected on an intensive study plot (500 ha) in the National Forest of Tapajós. Dendrogene is carrying out this work in cooperation with the project "Sustainable Forest Management for Timber" of the Brazilian Institute for Environment and Natural Renewable Resources (Instituto Brasileiro do Meio Ambiente e dos Recursos Naturais Renováveis, IBAMA) and the International Tropical Timber Organization (ITTO). Studies on the genetics, reproductive biology and ecology of seven timber species are under way and will continue after the area is logged in 2003. Using the ECO-GENE model, the project seeks to project the impacts of management alternatives on the genetic make-up of tree populations 50 to 100 years into the future. The ECO-GENE model will be used to predict the sustainability of alternative forest management scenarios to support decision-making in the development of public policies that concern forest resources. Some possible uses include testing of:

- the use of alternative criteria for the selection of trees for harvesting;
- the impact of incorrect identification of species;
- the impact of different intensities of logging;
- the impact of different spatial distributions of logging;

- the impact of different layouts of felling coupes within the management unit;
- the impact of riverside buffer zones and permanent reserve areas;
- the impact of different scales of management, from small community forest holdings to large holdings and landscapes;
- the impact of fragmentation of forest areas;
- the usefulness of different management indicators, such as percentage of commercial tree stock retained;
- species vulnerability to harvesting practice and possible species-specific controls.

An example for the application of ECO-GENE to estimate the impact of logging and forest fragmentation on the genetic diversity of Jacaranda copaia (one of the model species) is given in the contribution of Degen et al. (this issue).

DENDROBASE – Database on the genetic system of tropical tree species

The DENDROBASE is a databank which integrates genetic and ecological data as well as results of experimental field plots of tropical tree species (Degen 1999). The basic idea is to bring together existing knowledge in order to determine, for different key species or groups of tree species, the critical thresholds of minimal stock levels after logging. Therefore the data base contains tables with information about the breeding systems, level of inbreeding, reproductive system, agents of pollen and seed dispersal, parameters of genetic variation and genetic differentiation between populations as well as information about species abundance and spatial distribution. Where no experimental data for the simulation model ECO-GENE are available, the databank serves as a basis for the generation of model input data and defines the meaningful range of model parameters.

The efficiency of the database and the significance of data analysis and evaluation of species increase with a greater level of knowledge stored in the different tables. Therefore open access of the database on a network would be the best means to build up the system. Gathering together the knowledge generated by different institutional groups would be an important step forward. Controlled operation of the databank on the internet, maintaining correct usage, is the grand challenge for the future. Even if modelling shows limitations in the future, this database by itself will be of great importance for the generation of indicators and thresholds of genetic sustainability for forest management, as

well as for collecting together existing knowledge on genetic systems of tropical tree species. It will also contribute to improve the planning, coordination and prioritisation of research while improving access to existing information.

TREMA – Tree Mapping and Utilities

TREMA is a software tool being developed to manage and map tree data (http://www.trema.co.uk). It can be conceived as managing four basic types of data, species, geographic, tree and log data. It is developed in open architecture so that additional specific modules can be written and interfaced with the TREMA shell and core elements. It has many applications but core functions include the management of species names, planning and monitoring of forest harvesting, and mapping forest stand information. It is an appropriate mechanism to enable genetic and reproductive ecology knowledge to inform operational forest management (Hawthorne et al. 1999).

Field identification by common names, the normal practice in forest management, needs to correspond to scientific species identification, as species is the basis for biological reproduction and for access to technical information such as timber properties, ecology and economic value. To this end TREMA interfaces with the botanical research and the herbarium management software, BRAHMS. Species names need in turn to be related to timber trade names for marketing purposes. Cross-referencing and standardisation of names are complex but necessary steps to improve information use and management ability. The IBAMA Forests Products Laboratory have made their list of tree species of Brazil available for use in TREMA (Camargos J.A.A et al. 2001).

Tropical forest management systems are commonly based on selective logging practices. The ability to plan and monitor this selection in intensity and distribution at the species population level, within economically viable limits, is fundamental for sustainable forest management. The harvest planning module facilitates the application of multiple criteria to the forest stand information to determine selection of trees to be felled (harvested) or reserved.

The criteria used to select individual trees depend on factors such as local logging regulations; the current market for the forest produce, and short and long term economic and ecological considerations. Information on the actual harvest result in terms of which trees where actually felled can be incorporated for the monitoring of operational efficiency. Log production can be related to tree harvesting, creating the potential to evaluate economic results and trace chain of custody to the processing industry. Stand tables and harvest reports for management planning can be produced

Integral to the harvest planning module is the capability to display information spatially using maps. These have many applications e.g., to analyse spatial

distribution during tree selection, to plan extraction routes, or as basis for the operational control of tree felling and extraction operations.

Dendrogene impact pathways

Providing tools which offer the potential to take account of genetic and ecological evidence in policy and operational decision-making is the core activity for the Dendrogene project. However the project is also concerned with the pathways by which this potential will be realized. This is conceived in three steps:

- Participatory development of alternative policy and management scenarios
- Analysis of ecological sustainability impacts of alternative scenarios
- Participatory assessment of ecological impacts and examination of trade-offs with social and economic impacts

The success of this effort depends on an effective communication programme based on establishing effective two-way communication channels with key stakeholders. Many of the incentives to establishing effective communication are beyond the project's control e.g. legislative context, but such risks are minimized using a strategy of cooperation allied to a clear vision of the project's role in the broader sectoral and regional development scenario.

Acknowledgements

The authors acknowledge the Organising Committee of the Symposium for the opportunity in presenting this project overview, and for Dr. Namkoong and all participants´ contributions during Dendrogene Workshop in Belém-Pará, Brazil, in May 1998. The first author sincerely expresses his appreciation to Dr. Namkoong for all the growing opportunities he has been exposed to during the doctorate degree (and afterwards) under his guidance. The critical thinking, respect and support for contributions of others, made a tremendous impact on him at the personal and professional level.

The authors also express their thanks to DFID and Embrapa for the financial support of the project as well as to all members of the project team and partners for their participation in the project with special mention to Peter Coventry for his work in establishing the Tapajós field site. We are also grateful to Lyn Loveless for her constructive comments on a former version of this manuscript.

References

Bihun, Y. (1999): Unlocking Amazonas. Timber & Wood Products International (jun.99): 33-45.

Blate, G. (1997): Sustainable forest management in Brazil. Tropical Forest UPDATE, ITTO Newsletter, 7(3): 14-15

Camargos J.A.A.; Coradin V.T.R; Czarneski, C.M; De Oliveira D.; Meguerditchian.I. 2001 Catálogo de árvores do Brasil (2ª edição revisada e atualizada) Brasília:Ed.IBAMA, 2001 896p

Degen, B., Gregorius, H.-R., Scholz, F. (1996): ECO-GENE, a model for simulation studies on the spatial and temporal dynamics of genetic structures of tree populations. Silvae Genetica 45: 323-329

Degen, B., Scholz, F. (1996): Der Einsatz des Simulationsmodells ÖKO-GEN zur Erarbeitung von Entscheidungshilfen für eine nachhaltige Forstwirtschaftl. pp. 284-299 in: Müller-Starck. G. (ed.) Biodiversität und nachhaltige Forstwirtschaft. Ecomed - Verlagsgesellschaft, Landsberg

Degen, B., Scholz, F. (1998): Ecological genetics in forest ecosystems under stress as analysed by the simulation model ECO-GENE. Chemosphere 36: 819-824

Degen, B., Streiff, R., Scholz, F., Kremer, A. (1997): Analyzing the effects of regeneration regime on genetic diversity and inbreeding in oak populations by use of the simulation model ECO-GENE. pp. 9-21 in: Steiner K. C. (Ed.) Diversity and adaptation in oak species. Pennstate, College of Agricultural Sciences, Pennsylvania, University Park, Pennsylvania.

Degen, B. (1999): Dendrobase: Genetic system of tropical tree species. Users's manual for the data base. http://kourou.cirad.fr/genetique/Software/Dendrobase/manual.html. 29 pp.

Flinn, D.W., Franc, A. (1998): Conference overview. pp. 1-61. in: Flinn, D., Dolman, G., and Russel Haines (Eds.) Proceedings of international conference on indicators for sustainable forest management. Melbourne, 24-28 August, 1998.

Geburek, T., Mengel, M. (1998): Considerations in view of identity control of forest reproductive material and minimum number of seed trees required by law as exemplarity described in European larch. Centralblatt für das gesamte Forstwesen 115: 211-228

Hawthorne,W.D. , Filer, L D., Thompson, I. S (1999): O desenvolvimento e implementação do aplicativo TREMA de mapeamento e do planaejamento de colheita para seleção de arvores na Embrapa Amazônia Oriental. Simpósio Silvivultura na Amazônia Oriental:Contribuições do projeto EMBRAPA/DFID, Resumos Expandidos. 133-137. 303pp.

Kageyama, P. (1998): Manejo de florestas tropicais: O que há de novo?. Cad. biodivers. 1: 7-11.

Kanashiro, M., Thompson, I. S., Yared, J. A. G., Loveless, M.D., Coventry, P., Martins-da-Silva R.C.V., Degen, B., Amaral, W. (2002): Improving conservation values of managed forests: the Dendrogene Project in the Brazilian Amazon. Unasylva 53: http:// www.fao.org/forestry/FODA/UNASYLVA/unasyl-e.stm

Lisenmair, K. E. (1997): Biodiversity and sustainable management of tropical forests. In: Focus: Tropical Forests edited by Natural Resources and Development (45/46): 13-27

Namkoong, G., Boyle, T., Gregorius, H.R., Joly, H., Savoleinen, O. (1996): Testing criteria and indicators for assessing the sustainability of forest management: genetic criteria and indicators. CIFOR Working Paper No 10. Bogor, Indonesia.

Namkoong, G . (1998): Dendrogene; Genetic Conservation within Managed Forests in Amazônia, Workshop Report, Belém-PA (11-13 May, 1998). 34 pp.

Nepstad, D.C., Veríssimo, A., Alencar, A., Nobre, C. Lima, E. Lefebvre, P., Schelesinger, P., Potter , C., Moutinho, P., Mendonza, E. Cochrane, M., Brooks, V. (1999): Large scale impoverishment of Amazonian forests by logging and fire. Nature 398: 505-508

Silva, N. M. (1997): Manejo de florestas de terra-firme da Amazônia Brasileira. Curso de Manejo Sustentável. Embrapa Florestas, Curitiba-Pr, Brazil. 40 pp.

Szaro, R. C. (1998): Biological Diversity. pp. 60-61. in: Flinn, D., Dolman G., Russel Haines (Eds.) Proceedings of international conference on indicators for sustainable forest management. Melbourne, 24-28 August, 1998.

Thompson, I. S., Yared, J. A. G. (1997): Three selection: a key decision for sustainable forest management for timber production in terra-firme forests of the Eastern Amazônia. Proceedings of the XI World Forestry Congress. Antalya, Turkey (13-22 Oct. 1997). CD-ROM.

van Bueren, E.M.L., Blom, E. M. (1997) Principles criteria indicators, a hierarchical framework for formulation of sustainable forest management standards. The Tropenbos Foundation, The Netherlands.

Zachow, R. (1996): Os planos de manejo das florestas naturais no Brasil. Curso de Manejo Sustentável. Embrapa Florestas, Curitiba-Pr, Brazil.

Part 2

Modelling approaches

The effects of autocorrelated patterns among adults on pollen pool differentiation

R. J. DYER and V. L. SORK

Department of Biology, University of Missouri - St. Louis, St. Louis, Missouri 63121-4999, USA

Keywords: pollen pool, genetic differetiaiton, gene flow.

Abstract: In this paper, we examine how the genetic structure of adult pollen donors influences the differentiation of spatially separated pollen pools. We utilize Monte Carlo methods to simulate adult populations of 10,000 individuals on a 100x100 unit landscape under various levels of spatial genetic autocorrelation. Spatial autocorrelation was simulated by imposing allele frequency gradients on the adult genotypes. The strength of the allele frequency gradient, measured by the change in allele frequency across the landscape, ∂p, was simulated as 0.00, 0.10, 0.20, and 0.30 corresponding to adult genetic differentiation of ϕ_{ST} = 0.000, 0.023, 0.112, and 0.235, respectively. For each gradient strength, ∂p we simulated mating events to produce 20 offspring for 36 focal mother trees using mean pollen dispersal distances, γ, of 5, 10, 15, and 20 units. The genetic structure of spatially separated pollen pools was analyzing using the Two-Generation analysis. Results of our simulations suggest that spatially autocorrelated genetic structure significantly biases the among pollen pool measures of differentiation (ϕ_{FT}) upward. However, the magnitude of the bias was relatively small, ranging from near zero for $\partial p = 0.10$ up to $\partial\phi_{FT}$ = 0.021 for $\partial p = 0.30$. These results suggest that while non-uniformity in the adult genetic structure upwardly biases pollen pool differentiation, the magnitude of adult spatial autocorrelation patterns must be relatively large to be of concern. While the magnitude of the bias is relatively small, we present a simple method based upon the Mantel test for identifying the presence of spatial variation in adult genetic structure. We close this paper by applying these methods to a *Quercus alba* data set where we know *a priori* that spatial genetic structure exists in the pollen donor population.

Introduction

For many plant species, pollen dispersal is the primary vector of gamete exchange (e.g., Ennos 1994). With uniform genetic structure in the adult population, specific predictions can be made as to the differentiation of spatially separated pollen pools (e.g., Levin & Kerster 1974, Malécot 1973, Smouse et al. 2001). However, in the real world, natural plant populations typically exhibit spatial genetic structure, and examining the effects of this structure on pollen pool has serious implications for the analysis of gene flow patterns.

The magnitude of pollen pool differentiation is determined primarily by pollen dispersal distance and secondarily by adult genetic structure. Increasing the distance that pollen is dispersed will reduce pollen pool differentiation because maternal individuals have a greater probability of sampling the same pollen donor (e.g., Austerlitz and Smouse 2001). Conversely, pollen pool differentiation will increase with increasing genetic structure in the pollen donor population as spatially separated maternal individuals will sample pollen from genetically diverse subsets of fathers.

In this paper we present a portion of a larger work examining the effects of autocorrelated spatial genetic structure on pollen pool differentiation. Here, we show how autocorrelated structure in the pollen donor population significantly biases the observed pollen pool structure. We utilized Monte Carlo simulations to create populations both with (hereafter gradient) and without (static) adult spatial autocorrelated structure. Using the static populations as the null case, we show how allele frequency gradients upwardly bias the among pollen pool measure of genetic variation across a range of pollen dispersal distances. Using the maternal location as a surrogate to map the allele frequency gradient in the adults, we present a method based upon the Mantel test to identify spatial structure in the pollen donor population. We close by analyzing a *Quercus alba* data set where we know *a priori* that there is a mutlilocus gradient in adult genetic structure.

Methods

To determine the effects of spatial autocorrelated genetic structure on pollen pool differentiation, we simulated populations of 10,000 multilocus individuals following the simulation methods in Smouse et al. (2001) and Dyer et al. (in prep). All individuals were arrayed on a static 100 x 100 lattice, with 8 polymorphic loci and 3 alleles per locus in Hardy-Weinberg equilibrium. Pollen donor selection followed Austerlitz and Smouse (2001) using a bivariate exponential function. A total of 36 focal mothers were selected, from each of which 20 offspring were simulated during each simulation run. A total of four pollen dispersal distances were simulated, γ = [5, 10, 15, 20] during each simulation.

Two types of adult populations were created: populations with static allele frequencies (i.e., $p = q = r$; Fig. 1a) and populations with a gradient in allele frequencies (i.e., $p \neq q \neq r$; Fig. 1b). For the gradient populations, only the frequencies of the first and second allele were altered with $\delta p = -\delta q$, whereas the frequency of the third allele r was static. The strength of the gradient (δp) was simulated at three different levels, δp = [0.10, 0.20, 0.30], corresponding to

a mean genetic structure in the adults, as measured from one end of the gradient to the other, of ϕ_{st} = [0.023, 0.112, 0.235], respectively (AMOVA; $P < 0.01$ in all cases; Excoffier et al. 1992).

Analysis of pollen pool differentiation followed the Two-Generation analysis method presented elsewhere in this volume as well as in Smouse et al. (2001), where the measure of pollen pool differentiation, ϕ_{FT}, is analogous to the measure of adult differentiation, ϕ_{ST}, in Excoffier et al. (1992). The change in a pollen pool structure due to the gradient, $\delta\phi_{FT}$, was calculated as the difference between that observed in the gradient and the static simulations for the same pollen dispersal distance (i.e., $\delta\phi_{FT} = \phi_{FT\text{-gradient}} - \phi_{FT\text{-static}}$). Each simulation was repeated 1000 times for each combination of γp and γ, resulting in 16,000 $\delta\phi_{FT}$ estimates.

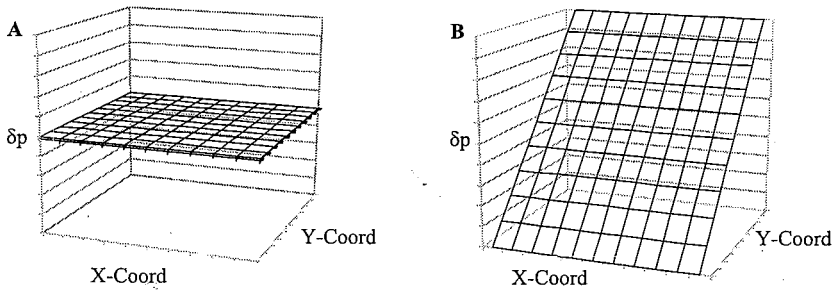


Figure 1. Schematic of change in allele frequencies (δp) across the landscape. A. The null case with δp = 0.0. B. The test cases where an allele frequency was imposed on the pollen donor population (i.e., δp > 0.0).

Results and discussion

Spatial autocorrelated genetic structure in the adult population does significantly upwardly bias the estimates of among mother genetic variation (Fig. 2). As predicted, the bias is proportional to the strength of the allele frequency gradient, with higher bias observed in the populations with stronger allele frequency gradients. The bias was also inversely proportional to the pollen dispersal distance. The magnitude of $\partial\phi_{FT}$ due to the allele frequency gradient ranged from near zero for $\partial p = 0.10$ up to $\partial\phi_{FT} = 0.021$ for $\partial p = 0.30$.

The magnitude of the bias attributable to allele frequencies in the adults, while significant, is relatively small. For example, genetic differentiation of pollen pools (ϕ_{FT}) for the static populations was $\phi_{FT} = 0.116$ for a mean dispersal distance, $\gamma = 5$ (data not shown), whereas the bias introduced by $\partial p = 0.30$ under these conditions was only 0.021 (Figure 2). While the permutation

confidence intervals clearly do not overlap those for $\partial p = 0.00$, the difference between $\phi_{FT} = 0.116$ and $\phi_{FT} = 0.137$ is relatively small.

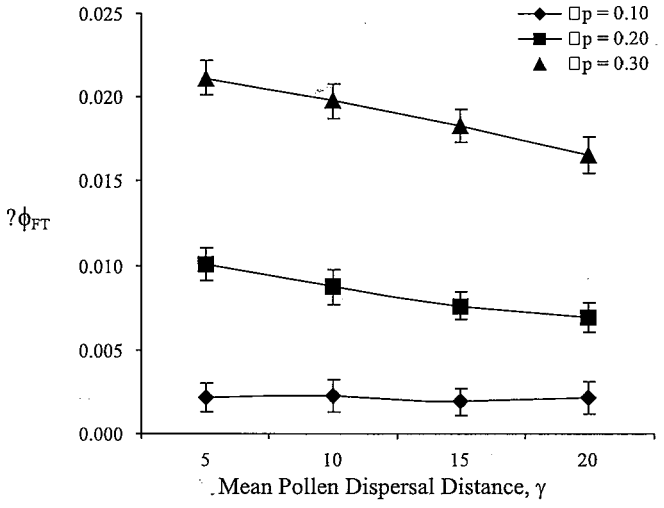


Figure 2. Change in ϕ_{ST} due to a gradient in adult genetic structure, δp. Error bars represent permutation 95% confidene intervals.

Clearly, non-uniformity in the adult population has an effect on the estimate of pollen pool differentiation, regardless of its magnitude. However, the investigator has no *a priori* knowledge of the presence of this bias by simply examining the pollen pool itself. But if there is spatial autocorrelation in the adult population, then the maternal locations will map that structure directly. Therefore, it is possible to use the maternal locations in conjunction with the observed array of paternal haplotypes to detect non-uniform adult genetic structure.

One of the unique features of the Two-Generation Analysis is its use of a pair-wise genetic distance matrix (see Excoffier et al. 1992; Smouse et al. 2001). For a set of focal mothers, one can construct a corresponding pair-wise physical distance matrix. The relationship between physical and genetic distances can then be evaluated using a non-parametric Mantel test (Mantel 1967). Significant correlation between maternal placement and paternal pollen donor identity would then suggest structure within the pollen donor population.

To evaluate the ability of the Mantel test to detect an allele frequency gradient, we analyzed a subset of a data set for *Quercus alba*. A full description of this data set is presented in Smouse et al. (2001). Previous analyses of the adult individuals on this landscape revealed a significant gradient in genetic

structure whose major axis runs from the southwest to the northeast (Koop and Sork, unpub.). A total of 945 offspring were sampled from 35 maternal trees (see Smouse et al., 2001). Significant differentiation in the pollen pools was observed (Two-Generation Analysis; $\phi_{FT} = 0.06$, $P < 0.01$). The correlation between maternal position along the north/south axis and the genetic differentiation of sampled pollen pools was $\rho = 0.32$ (Mantel; $P < 0.01$). In this case, the approximately linear gradient in the adult genetic structure was clearly observed in the corresponding pollen pool sampled by the focal mothers.

Here we have presented a brief overview of the effects that a non-uniform genetic structure has on pollen pool differentiation. The most important finding of these simulations is that pollen pool differentiation is biased upward by non-uniform adult genetic structure. Although the magnitude of the bias is relatively small, and would lead to underestimates of gene flow among maternal individuals, it can be used to identify underlying structure in the pollen donor population. Furthermore, we have briefly touched upon one approach for using the pollen pool to make inferences regarding the adult genetic structure. An alternate approach to identifying the presence of a bias as well as correcting for it requires a complete re-parameterization of the AMOVA model and is presented in Dyer et al. (in prep.).

Acknowledgements

We would like to thank Peter Smouse, Robert Westfall, and two anonymous reviewers as well as the Missouri Department of Conservation.

References

Austerlitz, F., Smouse P.E. (2001): Two-generation analysis of pollen flow across a landscape. II. Alternative dispersal functions and sampling strategies. Genetics 157: 851-857

Ennos, R.A., (1994): Estimating the relative rates of pollen and seed migration among plant populations. Heredity 72: 250-259

Excoffier, L. Smouse P.E., Quattro, J.M., (1992): Analysis of molecular variance inferred from metric distances among DNA haplotypes: Application to human mitochondrial DNA restriction data. Genetics 131: 479-491

Levin, D.A. Kerster, H.W. (1974): Gene flow in seed plants. Evol. Biol. 7: 139-220

Malécot, G. (1973): Isolation by distance. pp. 72-75 in: Morton, N.E. (Ed.). Genetic structure of populations. University of Hawaii Press, Honolulu.

Mantel, N.A. (1967): The detection of disease clustering and a generalized regression approach. Cancer Res. 27: 209-220

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

Effects of colonization processes on genetic diversity: differences between annual plants and tree species

F. AUSTERLITZ[1,2], S. MARIETTE[1], N. MACHON[3], P.-H. GOUYON[1] and B. GODELLE[1]

Address: [1] Laboratoire Evolution et Systématique, Université Paris-Sud, F-91405 Orsay, France; [2] Laboratoire de Génétique et d'Amélioration des Arbres Forestiers, INRA, F-33611 Gazinet, France; [3] Conservatoire Botanique du Bassin Parisien, Muséum National d'Histoire Naturelle, F-75005 Paris, France.

Key words: F_{ST}, trees, life-cycle, colonization, diversity

Abstract: Tree species are striking for their high within-population diversity and low among-population differentiation for nuclear genes. In contrast, annual plants show much more differentiation for nuclear genes but much less diversity than trees. The usual explanation for this difference is that pollen flow, and therefore gene flow, is much higher for trees. This explanation is problematic since it relies on equilibrium hypotheses. Since trees have very recently recolonised temperate areas, they have experienced many founder events, which usually reduce within-population diversity and increase differentiation. Only extremely high levels of gene flow could counterbalance these successive founder effects. We develop a model to study the impact of the life-cycle of forest trees, in particular the length of their juvenile phase, on genetic diversity and differentiation during the glacial period and the following colonization period. We show that both a reasonably high level of pollen flow and the life-cycle characteristics of trees are needed to explain the observed structure of genetic diversity. We also show that gene flow and life-cycle both have an impact on maternally-inherited cytoplasmic genes, which are characterized both in trees and annual species by much less diversity and much more differentiation than nuclear genes.

Introduction

The present distribution of genetic variability in temperate forest tree species poses many questions. Isozyme data (Hamrick et al. 1992; Hamrick and Godt 1996) on a large number of species indicate that trees maintain a significantly higher level of genetic diversity within species (H_e = 0.177 on average) and within populations (H_e = 0.148 on average) than annual plants (respectively H_e = 0.154 and H_e = 0.101 on average) for nuclear genes. Forest trees also show a lower level of genetic differentiation among populations. Hamrick and Godt (1990) calculated average values of G_{ST} of 0.084 for woody long-lived perennial species and 0.355 for annual plants. On the other hand, maternally-inherited cytoplasmic markers give completely different results: a clear geographic

structure is observed (Ennos 1994). For instance, Gst is 0.905 for *Quercus robur* and 0.925 for *Q. petraea* (Petit et al. 1993).

The high within-population diversity and low differentiation of trees' nuclear genes is unexpected. During the last glacial period of about 100,000 years (Andersen and Borns 1994), tree species from the Northern Hemisphere in Europe were confined to a few southern refuges completely isolated from one another. About 15,000 years ago, they began to recolonize all temperate areas, and some species have only recently reached their modern range limits (Huntley and Birks 1983; Huntley 1990). Our previous results (Austerlitz et al. 1997) showed that the successive founder events that occur during colonization yield strong genetic differentiation and low within-population diversity. This is the classically denoted founder effect.

The usual explanation for the difference between nuclear and maternally-inherited cytoplasmic genes is that there is much more migration through pollen than through seeds. Under equilibrium hypotheses and an island model, Ennos (1994) showed that this difference could be explained by a 200:1 ratio between pollen and seed migration rate for oaks. Equilibrium hypotheses seem unrealistic since recolonization is recent for forest tree species, and an island model on the scale of most temperate areas is doubtful. Le Corre (1997) simulated the process of European recolonisation with a 2D stepping-stone model. She had to use an even higher ratio of pollen flow versus seed flow to achieve results consistent with the data.

None of these studies took into account the main characteristics of the tree life-cycle: overlapping generations and a long juvenile phase. The first aim of this work is to show the impact of this life-cycle on genetic diversity and population structure during colonization. The second aim is to study the impact of the persistence of several refuges, isolated one from another during the glacial period, on contemporary population genetic structure.

We modeled the evolution of genetic structure for nuclear and maternally-inherited cytoplasmic genes of plants with a tree life-cycle (juvenile phase and overlapping generations), during and after a colonization period, following long isolation. They are compared with plants having no juvenile phase or non-overlapping generations, with either a one-year generation (annual plants) or the same generation time as trees. Here we present the model and the results very synthetically. For more details see Austerlitz et al. (2000).

Material and methods

Demographic model within each population: We simulated a tree life-cycle of several size classes. We structured populations according to the size of the

individuals rather than the age, as in Lefkovitch (1965). If A is the annual transition matrix and N(t) and N(t+1) the vectors of the numbers of trees in each size class, respectively, at years t and t+1, we write N(t+1) = A·N(t).

$$\mathbf{A}=\begin{pmatrix} P_{11} & 0 & 0 & \mathrm{L} & f \\ P_{12} & P_{22} & 0 & \mathrm{L} & 0 \\ 0 & P_{23} & P_{33} & & \mathrm{M} \\ \mathrm{M} & & \mathrm{O} & \mathrm{O} & \mathrm{M} \\ 0 & \mathrm{L} & 0 & P_{k-1,k} & P_{kk} \end{pmatrix}$$

where P_{ii} is the proportion of trees that stay alive and remain in class i, P_{ii+1} the that stay alive and move from i to i+1 and f the fecundity of the individuals of the last class, which are the only ones that can reproduce. We also included a density-dependent growth regulation function, adapted from Buongiorno et al. (1995).

The number of classes (k) is 25, so trees can never reproduce before the age of 25, and fecundity is high in the absence of density dependence (f = 250). The other parameters were chosen so that at equilibrium an individual reached the adult class on average at 50 years and the average age of the individuals in the adult class was 100 years. The results obtained for trees were compared with those obtained for plants with no juvenile phase and non-overlapping generations, with similar growth rate and effective population size (N_e).

Colonization models: We first modelled colonization for a one-dimensional stepping-stone metapopulation. At the beginning of a simulation, all sites were empty, except for the refuge, within which there was a population at demographic and genetic equilibrium. The total number of sites (d) was 15. Every 100 years a new site was opened to colonization, making the total colonization period last 1400 years. Subsequently, we let the populations evolve after the colonization period until time t = 5000 years, with the same seed and pollen flow.

The second model of colonization was a two-dimensional stepping-stone model. Since the process is highly computer intensive, the total number of populations was set at only 102 (6 along the X axis times 17 along the Y axis). We started with three refuge areas, each containing four populations. Before starting the colonization, we let the 12 populations evolve for 80,000 years, with no gene flow either by seeds or pollen among them. After the isolation period, the colonization was started, with one new area of populations being open for colonization every 500 years. This yielded a colonization period of 7,000 years. Then the populations evolved, remaining at their equilibrium size, for a post-colonization period of 8,000 years. The entire process lasted 95,000 years. Effective population sizes were 100 or 1,000, seed migration rate ranged from 0.0002 to 0.005 and pollen migration rate ranged from 0 to 0.1.

Genetic parameters: The mutation rate was set at 10^{-6} for nuclear and cytoplasmic genes. We used a Markovian approach to calculate the probabilities of identity by descent iteratively from one generation to the next, assuming an infinite allele model. This allowed us to calculate the average within-population diversity H_S, the expected global diversity ($H_T = 1 - f_T$) and the expected F_{ST}.

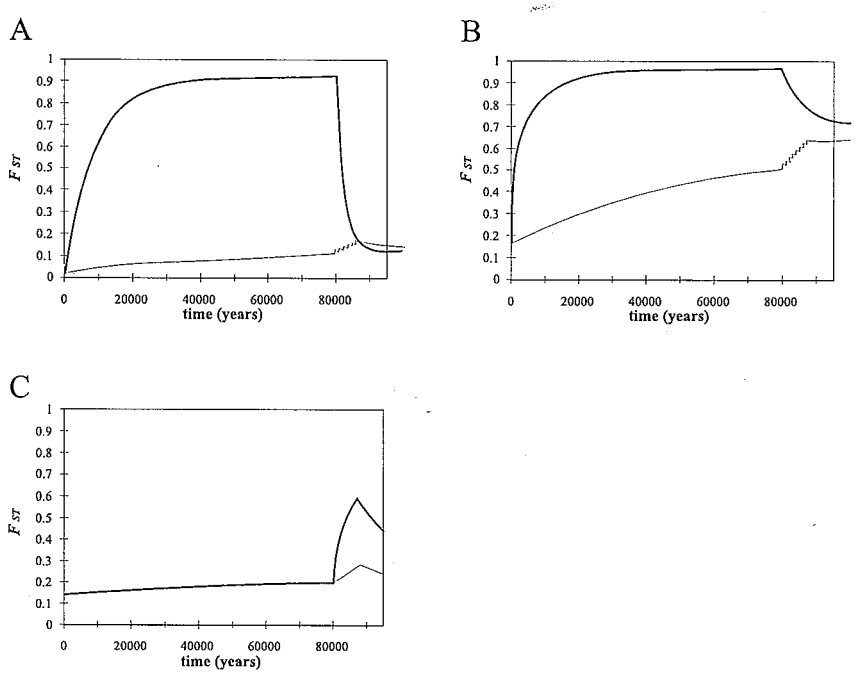


Figure 1. F_{ST} plotted against time for the 2D colonization process, with 6x17 populations for trees (solid lines) and annual plants (bold lines A-B) or plants with non-overlapping generations and the same generation length as trees (bold lines C). A: nuclear genes with seed migration rate $m_s = 0.0005$, pollen migration rate $m_p = 0.01$, effective size: $N_e = 1000$ B: Cytoplasmic genes: $m_s = 0.0005$, $N_e = 1000$, C: nuclear genes: $m_s = 0.0005$, $m_p = 0.01$, $N_e = 1000$. At 80,000 years, colonization processes were initiated.

Results

One-dimensional colonization process: For annual plants there was a very strong founder effect: diversity in each population decreased substantially and F_{ST} increased strongly during the first generations after each colonization event. On the other hand, for trees, genetic diversity decreased only slightly in the

newly-founded populations, and F_{ST} also increased very slowly. There was almost no founder effect.

Even with no pollen flow in the tree model, the F_{ST} value at the end of the colonization period was only 0.036, much lower than the final equilibrium value. For cytoplasmic genes, the increase in differentiation was slightly larger in the colonization period (Figure 1).

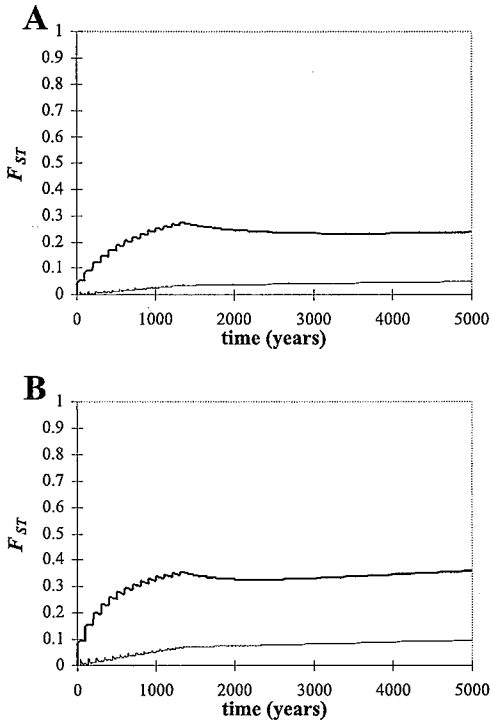


Figure 2. F_{ST} plotted against time since the beginning of the colonization period in the 1D colonization process, for trees (solid lines) and annual plants (bold lines), for nuclear genes (A) or maternally-inherited cytoplasmic genes (B). In all cases, d = 15, ms = 0.0002, Ne = 1000. For nuclear genes, the pollen flow mp was 0.

Two-dimensional colonization process: Several features appeared when two-dimensional processes were taken into account. For nuclear genes, F_{ST} first increased during the isolation period (Figure 2A). The increase was much stronger for annual plants in all cases. Second, F_{ST} for annual plants decreased

during the colonization period that followed. For trees, it decreased with high pollen flow or low effective population size, otherwise it increased slightly. Third, during the post-colonization period, there was a decrease of F_{ST} that was very sharp in the case of annual plants but smooth for trees. As a consequence, the F_{ST} for annual plants fell quickly below that of trees.

Concerning cytoplasmic genes (Figure 2B), even if we started with a low F_{ST} value, differentiation of trees for cytoplasmic genes increased much more than for nuclear genes during the isolation period. It increased even more sharply during the colonization period and then slowly approached the equilibrium.

For plants with non-overlapping generations and a generation time similar to that of trees (Figure 2C), the isolation period had the same impact as on forest trees. The colonization period yielded a stronger increase in F_{ST} of nuclear genes for these plants than that for the tree species. This difference was very low when pollen flow was very high. It was much higher and remained so long after the colonization period was over, when pollen flow had a lower value.

Discussion

In modelling the one-dimensional process, we showed that the founder effect seems to be much more limited for trees than for annual plants. This can be explained as follows: when a tree population is founded, growth for the first several years is because of new juvenile migrants and not because of reproduction. Therefore, when the first trees reach reproductive age in the newly-founded population, a non-negligible proportion of the site occupied by this population is already occupied by juveniles from seeds that arrived years before.

Our model also helps to explain the general pattern of genetic diversity observed on a continental scale for temperate forest tree species, both the differences with annual plants, as well as the opposite patterns of diversity of nuclear and cytoplasmic genes. On this last point, the two-dimensional colonization process shows that the differences between these two kinds of differentiation reflect not only differential gene flow during and after the colonization period, but also the fact that we have much differentiation between refuges for cytoplasmic genes and very little differentiation for nuclear genes during the isolation period. This is because of the lower effective population size for cytoplasmic genes and also the fact that pollen flow maintains connections between populations inside each refuge, allowing them to keep more of their diversity. The fact that this pattern is the same in Europe and Northern America (see e.g. Latta and Mitton 1997) indicates that even if the refuges were larger in Pleistocene North America, it did not allow tree species to maintain more of their cytoplasmic genetic diversity.

Provided that tree populations were already only slightly differentiated for nuclear genes when they entered the refuges and that the effective population sizes in the refuges were large enough, the tree populations might have maintained this low level of differentiation throughout the glacial period. Since recolonization does not yield a strong increase in differentiation for trees, F_{ST} would remain relatively low, without invoking a high level of pollen flow (between 1% and 10% of pollen coming from the neighbor populations, that is, between 20 and 200 times more than seed flow).

Our models reveal the complimentarity of all these processes. If tree life-cycle was not taken into account, a huge level of pollen flow would be necessary to explain observed differentiation. Nevertheless, with the same level of seed and pollen flow, nuclear genes' F_{ST} for annual plants falls below that of trees at the end of the process because of the rapidity of homogenization of annual plant populations. Therefore, higher levels of pollen flow or larger effective population sizes for trees than for annual plants are also necessary to explain the observed differences. Additionally, the two-dimensional process appears to be a key explanation for the lack of genetic differentiation as well.

References

Andersen, B.G., Borns, H.W., Jr. (1994): The Ice Age world: an introduction to Quaternary history and research with emphasis on North America and northern Europe during the last 2.5 million years. Oslo: Scandinavian University Press.

Austerlitz, F., Jung-Muller, B., Godelle, B., Gouyon, P.-H. (1997): Evolution of coalescence times, genetic diversity and structure during colonization. Theor. Pop. Biol. 51: 148-164.

Austerlitz, F., Mariette, S., Machon, N., Gouyon, P.-H., Godelle, B. (2000): Effects of colonization processes on genetic diversity: differences between annual plants and tree species. Genetics 154: 1309-1321.

Buongiorno, J., Peyron, J.-L., Houllier, F., Bruciamacchie, M. (1995): Growth and management of mixed-species, uneven-aged forests in the French Jura: Implications for economic returns and tree diversity. Forest Sci. 41: 397-429.

Ennos, R.A. (1994): Estimating the relative rates of pollen and seed migration among plant populations. Heredity 72: 250-259.

Hamrick, J.L., Godt, M.J.W. (1996): Effects of life history traits on genetic diversity in plants. Phil. Trans. R. Soc. Lond. B 351: 1291-1298.

Hamrick, J.L., Godt, M.J.W., Sherman-Broyles, S.L. (1992): Factors influencing levels of genetic diversity in woody plant species. New For. 6: 95-124.

Huntley, B. (1990): European vegetation history: paleovegetation maps from pollen data - 13000 yr BP to present. J. Quat. Sci. 5: 103-122.

Huntley, B., Birks, H.J.B. (1983): An atlas of past and present pollen maps of Europe: 0-13000 years ago. Cambridge: Cambridge University Press.

Latta, R. G., Mitton J.B. (1997): A comparison of population differentiation across four classes of gene markers in limber pine (*Pinus flexiblis* James). Genetics 146: 1153-1163.

Le Corre, V. (1997): Organisation de la diversité génétique et histoire postglaciaire des chênes blancs européens: approche expérimentale et par simulation. Ph.D. Thesis. Institut National Agronomique Paris-Grignon.

Lefkovitch, L.P. (1965): The study of population growth in organisms grouped by stages. Biometrics 21: 1-18.

Petit, R.J., Kremer, A., Wagner, D.B. (1993): Geographic structure of chloroplast DNA polymorphisms in European white oaks. Theor. Appl. Genet. 87: 122-128.

The response of forest tree populations to natural selection: variability and differentiation of adaptive traits and their underlying loci

V. LE CORRE[1], and A. KREMER[2]

[1] *INRA, Unité de Malherbologie et Agronomie, B.P. 86510, 21065 Dijon Cedex, France;* [2] *INRA, Laboratoire de Génétique et d'Amélioration des Arbres Forestiers, BP 45, 33611 Gazinet Cedex, France.*

Key words: selection, genetic variance, gametic disequilibrium, *Qst*, *Fst*.

Abstract: Recent advances in the detection of markers that are linked to the genes underlying phenotypic traits, or *QTLs*, will make it possible in the near future to measure adaptive genetic variation in forest tree populations. In this paper, we explore the relationship between allelic differentiation at *QTLs* (measured by *Gst*) and that for adaptive traits (measured by *Qst*), using theoretical predictions and simulations. Different types of natural selection are modelled and their consequences for the genetic structure of a metapopulation are described. The relationship between *Qst* and *Gst* is shown to depend on the extent of gametic disequilibrium between the QTLs.

Introduction

The management of genetic resources of forest trees relies upon two kinds of information: first, provenance tests are used to describe patterns of population differentiation for adaptive traits; second, biochemical and molecular markers provide insight into the demography and history of populations. These latter markers usually do not show patterns of differentiation that are concordant with provenance tests, because their evolution is driven mainly by neutral forces, while adaptive traits would experience diversifying selection (Latta 1998).

While recent progress has been made in the molecular identification of genes involved in the expression of adaptive traits, we still lack detailed understanding of how selective forces acting on a trait will affect allelic frequencies at the loci controlling the trait (hereafter referred to as "*LCATs*"). Previous studies by Latta (1998) and Kremer et al. (2000) have shown that for a multilocus trait, diversifying natural selection is more likely to act via the maintenance of among-population gametic disequilibria, rather than via allele frequency shifts at those same loci. In the present paper, these results are extended to various kinds of

selective pressure (uniform to highly diversifying, and weak to strong). We first present theoretical expectations, then present simulation.

Analytic theory

Consider a continuous trait that is controlled by n independent loci, each with two alleles and having symmetrical additive genetic effects $-a/2$ and $+a/2$ (Falconer and Mackay 1996, p.109). We consider a set of identical demes of constant size. Migration is conservative (Whitlock 1999). Mating is at random within each deme. The allelic diversity at *LCAT*s is subdivided according to Nei (1987) into the within-population expected diversity *Hs* and the among-population expected diversity *Dst*. The differentiation is measured by $Gst = Dst/(Hs + Dst)$. The genetic variance for the trait is decomposed into its within-population component *Vw*, and its among-population component *Vb*. A measure of differentiation for the trait is $Qst = Vb / (Vb + 2Vw)$ (Spitze, 1993).

For a neutral additive trait, *Qst* is equal to *Gst* (Spitze 1993; Whitlock 1999). This result assumes the absence of linkage disequilibrium among loci, based on the hypotheses of random mating within demes and conservative migration. Under selection, however, this assumption no longer holds, and the genetic variance consists of two components: (1) the contribution of allelic variation at each locus, or genic variance, and (2) the covariance of allelic effects among loci. Considering first-order gametic disequilibrium only,

$$V = \Sigma_i \sigma^2_i + \Sigma_i \Sigma_{j \neq i} C_{ij}$$

where σ^2_i is the genic variance at the i^{th} locus and C_{ij} is the covariance between the i^{th} and j^{th} loci. We introduce the parameter $\theta = \Sigma_i \Sigma_{j \neq i} C_{ij} / \Sigma_i \sigma^2_i$, so as to simplify the previous expression into $V = (1 + \theta) \Sigma_i \sigma_i$.

Under the diallelic model the genic variance at each locus is proportional to its expected genetic diversity, so that (Falconer and Mackay 1996, formula 8.5):

$Vw = n\, a^2\, (1 + \theta_W)\, Hs$ and $Vb = 2\, n\, a^2\, (1 + \theta b)\, Dst$,

which leads to the following relationship between *Qst* and *Gst*:

$$Qst = (1 + \theta b)\, Gst / [\, (\theta b - \theta w)\, Gst + (1 + \theta w)\,].$$

Under the diallelic model, *Qst* will therefore take the same value as *Gst* if:

1. there is linkage equilibrium among *LCAT*s (Latta 1998); and
2. linkage disequilibrium among *LCAT*s contributes equally to the within and among-population variances for the trait ($\theta b = \theta w$).

Qst will be greater than *Gst* if among-population disequilibrium is larger than the within-population disequilibrium, which is expected when diversifying selection drives populations to different optima. *Qst* will be smaller than *Gst* if within-population disequilibrium is larger, which is expected when selection drives all populations to the same optimum (Kremer et al. 2000).

Simulation results

The simulation model

The simulation model METAPOP used for this study has been described elsewhere (Kremer et al. 2000). Here, we consider a trait controlled by 10 independent *LCAT*s, with allelic effects drawn from a Gaussian distribution with mean zero and variance 5. The phenotypic value of a given genotype is the sum of independent genetic and environmental contributions, $Y = G + E$, where E is distributed normally with zero mean and variance set to 1. The genetic value G is the sum of allelic effects over all LCATs. The metapopulation consisted of 25 demes, of size 1000 each, and a stepping-stone migration model, with rates of 0.01 for pollen and 0.0001 for seeds. At the within-population level, stabilizing Gaussian selection towards a local optimum *Zopt*(*i*) determines the relationship between phenotype Z and fitness:

$$Wi(Z) = \exp[-(Z - Zopt(i))^2 / 2\omega^2]$$

The strength of selection ω^2 was set to vary between 1 (strong selection) and 1000 (very weak selection). The diversifying action of selection is scaled by *V(Zopt)*, the variance of *Zopt*(*i*) over populations. In the present study, *Zopt*(*i*) was set to vary according to a one-dimensional linear gradient on a grid of 5 x 5 populations. *V(Zopt)* varied between 0 (uniform selection) and 100 (highly diversifying selection). Initial genotypes were taken at random from a population of size 25,000 at mutation-drift equilibrium. Each simulation then lasted 10,000 generations.

Genetic variability at the within-population level

Both the genetic diversity at *LCAT*s and the genetic variance for the trait were reduced by selection (Table 1). The value of *Hs* was determined by the selection intensity ω^2. The within-population genetic variance resulted from an interaction between ω^2 and *V(Zopt)*. This is because strong selection depletes genetic variation in each population, while gene flow in association with diversifying selection restores some variation. As expected (Falconer and Mackay, 1996), stabilizing selection induced negative gametic disequilibrium that accounted for up to half the total genic variance under very strong selection but became insignificant for mild to weak selection.

Table 1. Genetic variability at *LCAT*s and at the trait under contrasting selection regimes

	Neutrality ($\omega^2 = 10^9$)	Uniform selection (*V(Zopt)* = 0)		Highly diversifying selection (*V(Zopt)* = 100)	
		Weak ($\omega^2 = 1000$)	Strong ($\omega^2 = 1$)	Weak ($\omega^2 = 1000$)	Strong ($\omega^2 = 1$)
H_S	0.469	0.192	0.045	0.310	0.032
V_W	47.78	0.843	0.098	8.328	0.208
G_{ST}	0.091	0.100	0.834	0.282	0.927
Q_{ST}	0.083	0.057	0.016	0.847	0.996

Values are means over 5 replicate runs for each selection regime.

Genetic differentiation among populations

The lowest values of allelic differentiation at *LCAT*s were observed for uniform and weak selection, while the highest values were observed for strong selection, either uniform or diversifying. Thus *Gst* depends essentially on ω^2. The quantitative differentiation *Qst* showed quite a different pattern, as it varied predominantly with *V(Zopt)*. This was because the within-population variance did not vary much, always being low, whereas the among-population variance was equal to *V(Zopt)*.

Gametic disequilibrium at the among-population level varied widely with selection parameters. Under uniform selection, negative genetic covariance eliminates among-population variance (that is, $\theta b = -1$). Populations might therefore show no differentiation for the trait, whereas allelic differentiation was maintained at *LCAT*s. Under diversifying selection, the among-population

disequilibrium was negative when *V(Zopt)* was small, or positive when *V(Zopt)* was large. Thus, some combinations of ω^2 and *V(Zopt)* should exist, for which the response to selection is achieved via allele frequency change ($\theta b = 0$), and *Qst* may be either greater or smaller than *Gst*. But the values of the selection parameters for which *Qst* = *Gst* are not necessarily the same as those for which $\theta b = 0$ (contrarily to the model of Latta, 1998), because some within-population disequilibrium may also affect *Qst*.

Conclusion

In accordance with previous results (Latta 1998; Kremer et al. 2000), the difference between genetic differentiation at an adaptive trait versus that at loci controlling it was found to be determined mainly by the amount of among-population gametic disequilibrium. The amount of within-population gametic disequilibrium also influences this relationship, at least theoretically, but we showed that its effect is meaningful only under strong selection. According to our simulation results, the amount of among-population disequilibrium, and thus the discrepancy between *Qst* and *Gst*, is determined by the variance of local optima for selection. Under diversifying selection, *Qst* is expected to be greater than *Gst* and the difference between the two increases with the variance of local optima. Since different traits may be subject to different kinds of selection pressure, according to their ecological function and impact on fitness, the amount of differentiation at a trait, compared with that at the underlying loci, is not predictable in a general way (Table 2).

Table 2. Putative patterns of genetic differentiation between forest tree populations for traits, *LCAT*s and molecular markers.

Selection:	Uniform or weakly diversifying	Diversifying
Weak	*Qst* < *Gst* for *LCAT*s ≈ *Gst* for markers Height, architecture ?	*Qst* >> *Gst* for *LCAT*s > *Gst* for markers Others traits?
Strong	*Qst* < *Gst* for markers << *Gst* for *LCAT*s Traits correlated with fitness?	*Qst* ≈ *Gst* for *LCAT*s >> *Gst* for markers Phenology ?

References

Falconer, D.S., Mackay, T.F.C. (1996): Introduction to quantitative genetics. Longman Group Ltd

Kremer, A., Le Corre V., Mariette S. (2000): Population differentiation for adaptive traits and their underlying loci in forest trees: theoretical predictions and experimental results. pp. 59-74 in: Matyas, C. (Ed.) Forest genetics and Sustainability, Volume 63. Kluwer Academic Publishers

Latta, R.G. (1998): Differentiation of allelic frequencies at quantitative trait loci affecting locally adaptive traits. Am. Nat. 151: 283-292

Nei, M. (1987): Molecular evolutionary genetics. Columbia University Press, New-York.

Spitze, K. (1993): Population structure in *Daphnia obtusa*: quantitative genetics and allozymic variation. Genetics 135: 367-374

Whitlock, M.C. (1999): Neutral genetic variance in a metapopulation. Genet. Res. Camb. 74: 215-221

Impact of selective logging and forest fragmentation on the seed cohorts of an insect- pollinated tropical tree : a simulation study

B. DEGEN[1], D. W. ROUBIK[2] and M. D. LOVELESS[3]

Address: [1]UMR CIRAD INRA ENGREF, Campus agronomique, BP 709, 9738 - Kourou cedex, French Guiana; [2]Smithsonian Tropical Research Institute, Apartado 2072, Panama City, Panama; [3]Department of Biology, College of Wooster, Wooster, OH 44691 USA

Key words: genetic diversity, model, bees, animal pollination, forest exploitation, *Jacaranda copaia*

Abstract: We used the population genetics simulation model ECO-GENE to evaluate the potential effects of two different disturbance regimes on the genetics of seeds produced by a tropical forest tree. The model simulated pollination by a large bee, moving pollen between individuals of *Jacaranda copaia*. The spatial distribution of the modelled population was derived from field data collected in the Tapajós National Forest (Pará, Brazil). Two alternative scenarios, one for selective logging (removal of all trees larger than a minimum diameter) and one for fragmentation (interruption of the forested population by areas without forest) were tested against the control scenario. The genetic makeup of seed arrays produced under these conditions was described with a variety of standard genetic measurements. Both disturbance scenarios resulted in an increase in the distance of pollen movement in the populations, and a decrease in the effective population sizes represented in the seed arrays. All genetic diversity measures decreased in the disturbance scenarios compared to the control, but some measures of genetic change were more sensitive than others to logging or fragmentation. In the disturbance scenarios, the genetic distance was nearly doubled among allele frequencies of adults and seed arrays. Measures of heterozygosity, widely used to measure population genetic variability, were quite insensitive to these major population disturbances. On the other hand, measures of effective population size and of multilocus genotype diversity showed clear responses to post-disturbance adult distributions. Thus, documentation of genetic changes after disturbance needs to be measured with the appropriate genetic measures.

Introduction

In the last 20 years, tropical forests have been disappearing on a global scale, as a result of forest conversion to agriculture and other land uses. An estimated 100,000-200,000 km² are deforested every year (Katzman and Cale 1990). Forest loss during the decade from 1981 to 1990 was 15.4%, or an annual loss of 0.81% of the total forest area present in 1980 (Whitmore 1997). In addition to forest conversion, various less intensive land uses are altering the makeup of

tropical forests. In most tropical forests, selective logging is common. In the Brazilian Amazon, which covers 285 million ha (IBGE 1996), an annual harvest of 28 million m^3 of timber is extracted. Such a harvest is unlikely to be sustainable over the long term.

The result of these logging and land use practices is either a forest with a reduced density of reproducing trees or a fragmented forest, with local patches of degraded woodlands separated by intervening, non-forested habitat. Recent studies have focused on trying to understand the impact of such changes in the forest on genetic variation in tree populations. Reduction of population sizes after logging may lead to genetic erosion, especially the loss of rare alleles (Barrett and Kohn 1991; Ellstrand and Elam 1993; White and Boshier 2001). In addition, changes in adult densities in the logged forest may alter pollen movement and other breeding parameters among residual individuals (Loveless and Hamrick 1984; Nason et al. 1997; Cascante et al. 2002). This could reduce cross-pollination and gene flow by impeding pollinator movement between trees. On the other hand, evidence from some studies suggests that pollinating animals may be able to alter their behaviours to adapt to this reduced density, and to continue to provide effective gene flow (Chase et al. 1996; Nason and Hamrick 1997; Aldrich and Hamrick 1998; Dick 2001, and this issue). However, logging also reduces mean tree sizes in the residual population, and as remaining trees become increasingly reproductively isolated, more frequent inbreeding and lower seed production is likely to occur (Stacy et al. 1996; Nason and Hamrick 1997; Ghazoul et al. 1998; Cascante et al. 2002). Thus, the impacts of disturbance and fragmentation on the genetic structure of tree populations may vary, depending on the ecological details of each tree species.

In this study, we used the simulation model ECO-GENE (Degen et al. 1996) to elucidate the varied processes and genetic consequences associated with logging and forest fragmentation. The model integrates population genetic and demographic processes to simulate the impact of different management scenarios on genetic makeup, particularly on the genetics of seeds produced after logging. Recently a new module to simulate animal pollination has been added to the model to further improve its ability to simulate real mating processes in tropical trees (Degen and Roubik, submitted).

The objectives of our study were: (1) to analyse the impact of selective logging and forest fragmentation on the genetic diversity of post-logging seed cohorts, (2) to determine pollen dispersal distances under changed post-disturbance conditions, and (3) to evaluate experimental methods for monitoring the potential negative impact of logging and forest fragmentation on population genetic structure.

Materials and methods

The simulation model ECO-GENE

The simulation model ECO-GENE was originally developed to evaluate human influences on the genetics of tree populations. It simulates the temporal and spatial dynamics of allele and genotype frequencies within a spatially and genetically defined population. As a general model, ECO-GENE includes modules that simulate tree growth, population genetics, and population dynamic processes. Genetic dynamics include single locus and multi-locus genotypes, and bi-parentally and uni-parentally inherited genes. Spatial and temporal genetic dynamics also are included, and overlapping or separate generations can be specified. A variety of mating systems can be implemented. Further description of the model and its application can be found in Degen et al. (1996), Degen et al. (1997), and Degen and Scholz (1998). Here we present simulations that include a module that attempts to include the most important components of animal pollination.

Modelling animal pollination

Pollination occurs during foraging flights of the pollinator. Our simulation generates a sample of foraging flights that lead to pollination and seed production. Variables taken into consideration are the temporal differentiation of flowering among trees, their spatial distribution, differences in inflorescence densities among trees of different sizes, levels of genetic self-compatibility of individual trees, and pollen carryover produced by a pollinator as it moves from flower to flower. Selected variables also are modelled for pollinators: spatial distribution of nests or home bases, maximum flight distance, and differing attraction to different tree crowns, based on the size of the floral display or proximity of the tree to the nest or individual forager. Further details are given elsewhere (Degen and Roubik submitted).

Data set

ECO-GENE requires input data on tree spatial positions, stem diameters, and genotypes. We generated an artificial data set, based on forest inventory data (diameter and density) for *Jacaranda copaia* from a 100 ha experimental plot in the Tapajós National Forest south of Santarém (Pará, Brazil). The field data were collected by the Dendrogene Project, a research effort jointly supported by the Brazilian government (EMBRAPA) and the British government

(Department for International Development). The inventory was made by a forest inventory team employed by Maflops, the timber company which holds the logging concession for this site, under the supervision of IBAMA. All trees of *Jacaranda copaia* with a minimum diameter of 10 cm were measured and mapped in 100 ha. An artificial data set (scenario I, control), was extrapolated from these data. The control data set included 473 trees on a 2000 m x 2000 m stand (400 ha, Fig. 1). Trees were randomly distributed, in accordance with the spatial patterns measured at Tapajós.

Jacaranda copaia is a light-demanding pioneer tree species, germinating and growing in gaps within the forest, and eventually recruiting into the canopy. It is a common component of both primary and secondary forests throughout Central and South America. Each flowering tree produces a large, showy display of lavender flowers above the crown. The hermaphroditic, zygomorphic flowers are pollinated by medium to large-sized bees. Seeds are wind dispersed. Genetic studies have shown that the species is self-incompatible (James et al. 1998).

To test the effects of various logging processes on this population, we modified the original (control) tree distribution to produce two different disturbance scenarios. In the logged stand (scenario II), we eliminated all trees greater than 31 cm dbh, throughout the entire 400 ha, simulating a very strong selective logging event. For the fragmentation model (scenario III), we reduced the area of closed forest in this plot from 400 ha to 140 ha, distributed randomly within several fragments. Outside the fragments, all trees disappeared. Within the fragments, the original tree distributions, sizes, and densities were retained. (Fig. 1).

Because we had no empirical information on the genetics of this species, we used a data generation routine to simulate a reasonable genetic structure for our study population. Multi-locus genotypes of these trees were generated using information on allele frequencies for three microsatellite loci, analysed for *Symphonia globulifera* in French Guiana (Degen et al. unpublished). Genotypes were assigned to the simulated population in such a way that it had the same genetic characteristics (allele frequencies and F_{is}) as the measured *Symphonia* population. As with the empirical data, the generated genotypes had an excess of homozygotes (F_{is} (fixation index) = 0.18). This excess indicates a certain level of inbreeding which can be explained by bi-parental inbreeding for a self-incompatible species.

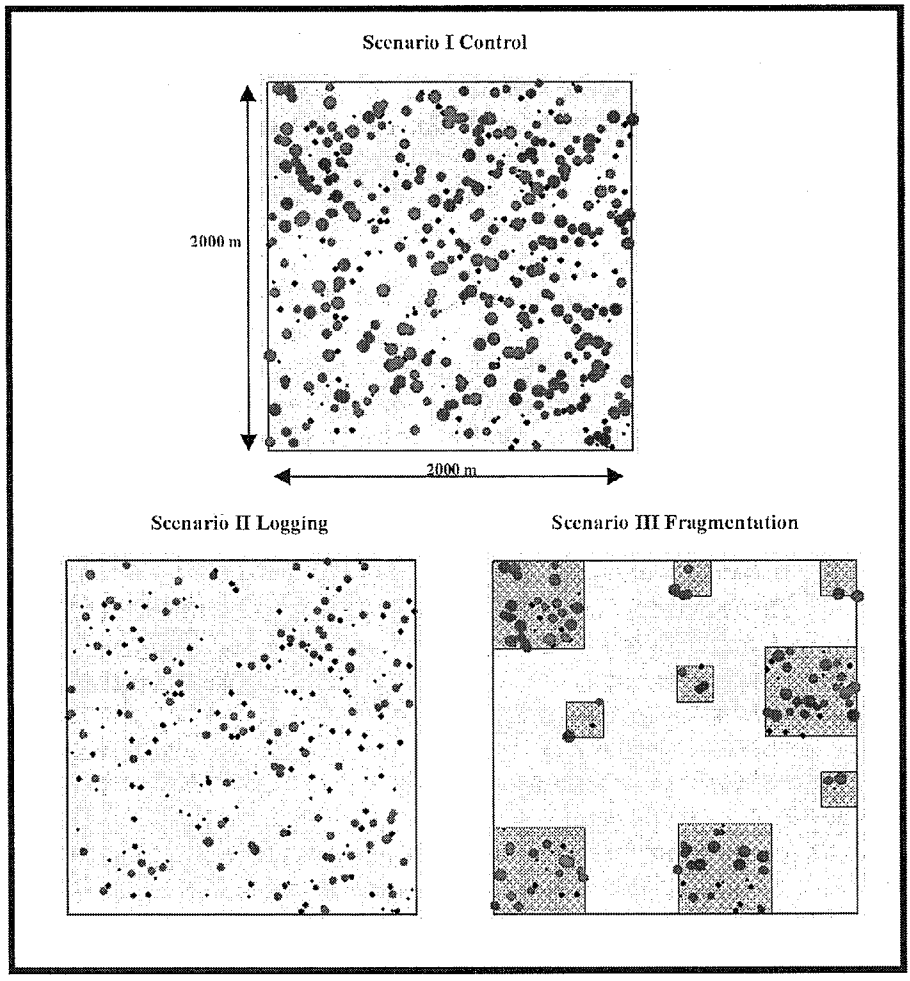


Figure 1. Spatial distribution of the trees in the three scenarios: scenario I (control), scenario II (logging) and scenario III (fragmentation). Each population covers an area of 400 ha. In the control, there are 308 trees above the size (20 cm) of hypothesized flowering. In both the logging and fragmentation populations, 90 trees above flowering diameter remain. In the logging scenario, these are all small trees (20-30 cm dbh). In the fragmented population, larger trees remain, but the trees are separated by areas of non-forest habitat.

Scenarios

In all three scenarios, the parameters for individual flowering phenology remained constant (starting date: $\mu = 0$, $\sigma = 10$ days; duration: $\mu = 5$, $\sigma = 2$ days). All trees above a minimum diameter of 20 cm flowered and produced fruit. We

used a standard parameter set to simulate the process of animal pollination and set the self-incompatibility to 100% (Degen and Roubik, submitted). Among the variables for pollinators were level of social organization and capacities of communication, search rules, flower preferences, foraging constancy, departure rules from flower patches, and aggressive territorial behaviour. Parameters were chosen to simulate the probable movement patterns of mid- to large-sized tropical bees (Roubik, pers. comm.). Resulting pollen distribution curves are shown in figure 2.

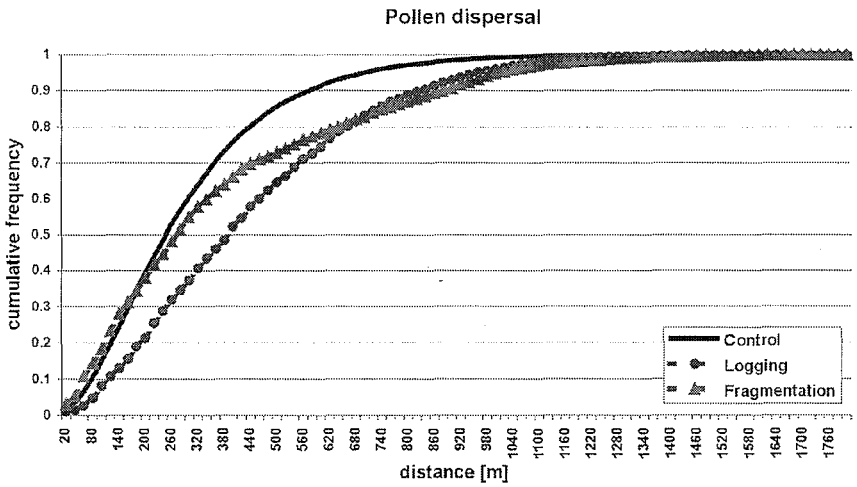


Figure 2. Cumulative distribution of effective pollen dispersal in the three scenarios. The curves are based on mean values from 100 repetitions.

All parameters were kept constant for the three scenarios. Only the data on *Jacaranda* density and distribution used for model initiation differed between scenarios. The control, Scenario I, with 308 reproducing trees, served as a standard setting against which to measure genetic changes. In both scenario II and scenario III, 90 reproducing trees were left in the post-disturbance populations.

Data analysis

We ran 100 replicate simulations of each scenario and computed means and standard deviations for the following measures, based on the 300 seeds produced in each simulation:

Effective population size was calculated as:

$$1 \leq N_e = \frac{1}{\sum w_i^2} \leq N$$

where:
w_i = relative contribution of successful haplotypes of tree i (pollen, ovules) to the sampled seeds
N = absolute number of adult trees in the population

We used the number of different multi-locus genotypes (*N* Geno) found in the seed sample as measures of multi-locus genetic diversity. In addition, the mean values for the expected heterozygosity (H_e) under Hardy-Weinberg equilibrium at the three microsatellite loci were computed. Other measures included mean distance of effective pollen movement (PolDis), the cumulative number of different alleles at all three loci (*A*) and the mean of the effective number of alleles per locus ($A_e = 1/(1-H_e)$). Moreover, we computed genetic distance (D_G) among allele frequencies of adults and seed arrays (Gregorius 1978).

Results

All genetic variation measures were smaller in the logging and fragmentation scenario compared to the control (Table 1).

Table 1. Mean values and standard deviations for measures of effective population size (N_e), mean pollen movement distance (PolDis), number of different multilocus genotypes (*N*), expected heterozygosity (H_e), absolute and effective number of alleles (A, A_e), and genetic distance of Gregorius (D_G) among allele frequencies of adults and seeds calculated using seed arrays of 300 seeds from each of 100 trials per scenario)

	Control	Logging	Fragmentation
N_e	134.65 (7.21)	43.21 (4.06)	42.43 (3.30)
PolDis [m]	311.9 (19.24)	456.8 (31.9)	389.3 (35.4)
***N* (geno.)**	283.2 (4.7)	222.6 (10.4)	203.7 (10.2)
H_e	0.77 (0.01)	0.74 (0.01)	0.75 (0.01)
A	30.9 (0.92)	30.59 (0.98)	29.12 (0.68)
A_e	4.41 (0.19)	3.93 (0.26)	4.08 (0.30)
D_G	0.067 (0.01)	0.128 (0.02)	0.118 (0.02)

The strongest differences were observed for the parameters measuring reproductive effective population sizes N_e. These values were reduced by about 68% under logging and 79% by fragmentation. Large differences in the opposite direction were found for the mean distances of pollen flow, which increased 46% in the logging scenario and 25% in fragmented landscape (fig 2). The multilocus diversity in the seed cohort (*N* geno) was reduced 22% by logging and 29% by fragmentation. The calculated values for heterozygosity, and the absolute and effective number of alleles actually changed little among scenarios (1 to 11% differences compared to the control).

Discussion

Realism of the model

The realism and utility of our model can only be validated by comparison of modelled results with real data sets of empirically measured outcomes or impacts. Real systems are likely to contain significantly higher levels of complexity, in both ecological and genetic processes, than can be captured in a simulation. One might expect to see frequency-dependence or density-dependence in individual mortality, as well as non-linear responses and relationships among parameters. For example, a mixture of different pollinators visiting the same tree species, each with differing competitive abilities, flight ranges, foraging patterns, and responses to changes in plant density, may introduce considerable variation in genetic outcomes. Our discussion here illustrates potential applications of the model to scenario analysis, and evaluates our simulated outcome based on published studies.

Gene flow

In our simulations, the mean distance of pollen flow increased after both logging and fragmentation. Hence, given the initial parameters we used, the pollinators could buffer, to a certain extent, the reduced density of flowering trees by flying longer distances. This observation is in agreement with recent studies on gene flow in fragmented forests. In *Pithecellobium elegans*, Chase et al (1996) demonstrated that trees in a pasture "fragment" received pollen from distant, isolated trees, as well as from adults outside the mapped population. White and Boshier (2001) compared the gene flow in fragmented populations of *Swietenia humilis* in Honduras with the gene flow of the same species in intact, closed forest. They found that the proportion of long-distance pollen flow increased

with a decrease in the population size of the forest fragments. In this region *Swietenia humilis* is pollinated by small bees, butterflies and other insects. They observed pollen flow over distances larger than 4.5 km. Similar observations were made by Dick (2001) for *Dinizia excelsa* in the Brazilian Amazon. He found higher fecundity and greater distances of pollen flow to trees in pasture and forest fragments than for trees in nearby, pristine forest. The long-distance pollen movement in *Dinizia* was explained not by changes in pollinator flight distances, but by a shift in the species makeup of pollinator assemblages between forest fragments and pristine forest. The pollen vectors in disturbed habitats were Africanized honeybees (*Apis mellifera*), which recruited in large numbers to visit flowers on isolated trees within pastures. In this study, Dick demonstrated pollen movement between trees separated by up to 3.2 km. Nason and Hamrick (1997) found that up to 100% of the pollen that fertilized ovules in small fragmented populations of *Spondius mombin* came from outside the fragment. In addition, however, they found reductions in fruit production and in germination for fruits from individuals in the fragments, in comparison to trees from nearby, continuous forest. Finally, Ghazoul et al. (1998) examined pollination and fruit production in the Dipterocarp *Shorea siamensis* in western Thailand. As was true for *Spondius*, these authors also detected a reduction in fruit production with decreasing population density, under strong logging pressure. In addition, however, and in contrast with the previous studies, they found that inter-tree movements of small, pollinating *Trigona* bees declined with increasing distance between flowering trees. These small bees were not behaviourally compensating for changes in adult densities. Clearly, the behaviour of pollinators in disturbed habitats will be highly dependent on their willingness to cross gaps of non-forested habitat, as well as their ability to detect conspecific trees at further distances, and their degree of fidelity to any particular species. In order to model these behaviours accurately, we need more field studies on pollinator behaviours under different field conditions.

The (in)sensitivity of diversity measures

Our study shows that measures of effective population size and multi-locus diversity were most sensitive to the impact of logging, whereas the absolute number of alleles and measures of heterozygosity were relatively insensitive. The effective population size was calculated directly during simulations and not estimated from genotype arrays, as is the usual practice. It might be possible to calculate the direct measure via paternity analysis, where such data are available. The effective population size, although evidently a good measure for population genetic impact, according to our simulation results, might be problematic to use for experimental studies. This is because the sampling design and sample size both impose restrictions that might be difficult to overcome in the field.

Highly variable microsatellites are increasingly being used to study genetic structure and gene flow. Their advantage for paternity analysis is evident, but their usefulness in monitoring forces that reduce genetic variation is questionable. Single locus measures, heterozygosity for example, are usually high for hypervariable microsatellites, and therefore less sensitive to bottleneck effects. This was observed in our simulation study and is also reported by White et al. (1999) in their experimental study of *Swietenia humilis*. Using microsatellites, they observed an immediate effect of fragmentation through loss of low-frequency alleles, but the level of heterozygosity was not significantly affected. Accordingly, we would expect a stronger effect of logging on the absolute number of alleles. But in our simulation, only fragmentation led to a stronger reduction in the numbers of alleles (-6%). A population of 90 reproducing adults, combined with an effective gene flow system, seemed sufficient to conserve allelic variation under the logging scenario, at least in this simulation study.

Short term and long term genetic impacts

Most tropical rain forest tree species have more individuals that are below the minimum size for timber exploitation than above. Where this is so, the advanced regeneration might already contain a large fraction of the genetic diversity of the adults, thus buffering the species against genetic erosion under moderate logging scenarios (Jennings et al. 2001). Thus, even if selective logging has a significant impact on the genetic diversity of seed cohorts produced in the immediately post-logged forest, the long term impact on population-level genetic variation might be small. Most measures of tropical tree diversity base their sampling on genotypes of adult trees. There are few data to indicate the degree to which advanced regeneration reflects the genetics of the adult, reproducing population but one could reasonable assume that they are pretty highly correlated. Such information would be valuable in order to predict the degree to which demography might buffer genetic erosion. However, over the longer term, seeds produced in a forest after disturbance events will affect the genetic makeup of a future forest. These seeds will provide the ongoing new recruitment that will dominate the forest 40 or more years in the future. Thus, changes in the genetics of seed cohorts are relevant to long-term genetic structure and the maintenance of genetic variability. Changes in pollination and seed dispersal following fragmentation have been shown by Aldrich and Hamrick (1998) to produce a bottleneck effect in newly-established seedlings of *Symphonia globulifera*, with potential long-term consequences for regional genetic variation. Seeds of *Symphonia* produced after fragmentation were derived from many fewer parents than those produced (and established) before the fragmentation event. In order to improve our ability to simulate this long term impact, we are adding

modules on demographic dynamics, mortality and inter-annual flowering patterns to ECO-GENE. We see the challenge as being able to apply this modelling tool to predict long term effects of selective logging or other disturbance and management scenarios on the genetic diversity of tropical trees. In this context, genetic modelling can serve as a means to identify critical limits for sustainable tropical forest management (Kanashiro et al. 2002).

References

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

Barrett, S. C. H., Kohn, J. R. (1991): Genetic and evolutionary consequences of small population size. In: Falk, D. A. and Holsinger, K.E. (Eds.) Genetics and Conservation of Rare Plants. Oxford Univ. Press, NY. Pp. 3-30

Cascante, A., Quesada, M, Lobo, J. J., Fuchs, E. A. (2000): Effects of dry tropical forest fragmentation on the reproductive success and genetic structure of the tree *Samanea saman*. Conservation Biology 16:137-147

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

Degen, B, Roubik, D. W. (submitted) Genetic variation in seed cohorts produced by animals in tropical trees - a simulation approach

Degen, B, Gregorius, H-R, Scholz, F. (1996): ECO-GENE, a model for simulation studies on the spatial and temporal dynamics of genetic structures of tree populations. Silvae Genetica 45, 323-329

Degen, B, Scholz, F. (1998): Ecological genetics in forest ecosystems under stress as analysed by the simulation model ECO-GENE. Chemosphere 36, 819-824

Degen, B., Streiff, R., Scholz, F., Kremer, A. (1997): Analyzing the effects of regeneration regime on genetic diversity and inbreeding in oak populations by use of the simulation model ECO-GENE. pp. 9-21 in: Steiner, K. C. (Eds.) Diversity and Adaption in Oak Species. University Park, Pennsylvania. Pennstate, College of Agricultural Sciences, Pennsylvania

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

Ellstrand, N. C., Elam, D. R. (1993): Population genetic consequences of small population size: Implications for plant conservation. Ann. Rev. Ecol. Syst. 24:217-242

Ghazoul, J., Liston, K. A., Boyle, T. J. B. (1998): Disturbance-induced density-dependent seed set in *Shorea siamensis* (Dipterocarpaceae), a tropical forest tree. Journal of Ecology, 86, 462-473

Gregorius, H.-R. (1978): The concept of genetic diversity and its formal relationship to heterozygosity and genetic distance. Math. Bioscience, 41, 253-271

James, T., Vege, S., Aldrich, P., Hamrick, J.L. (1998): Mating systems of three tropical dry forest tree species. Biotropica 30, 587-594

Jennings, S. B., Brown, N. D., Boshier, D. H., Whitmore, T. C., Lopes, J. d. C. A. (2001): Ecology provides a pragmatic solution to the maintenance of genetic diversity in sustainably managed tropical rain forests. Forest Ecology and Management 154, 1-10

Kanashiro, M., Thompson, I. S., Yared, J. A. G., Loveless, M.D., Coventry, P., Martins-da-Silva R.C.V., Degen, B., Amaral, W. (2002): Improving conservation values of managed forests: the Dendrogene Project in the Brazilian Amazon. Unasylva 53: http://www.fao.org/DOCREP/004/Y3582E/Y3582E3500.HTM

Katzman, M. T., Cale, W. G. (1990): Tropical forest preservation using economic incentives. Bioscience, 40, 827-833

Loveless, M. D., Hamrick, J. L. (1984): Ecological determinants of genetic structure in plant populations. Ann. Rev. Ecol. Syst. 15: 65-95

Nason, J. D., Hamrick, J. L. (1997): Reproductive and genetic consequences of forest fragmentation: two case studies of Neotropical canopy trees. Jour. Heredity 88: 264-276

Nason, J. D., Aldrich, P. R., Hamrick, J. L. (1997): Dispersal and the dynamics of genetic structure in fragmented tropical tree populations. pp. 304-320 in: Laurance, W. F., Bierregaard, R.O, Jr. (Eds.) Tropical Forest Remnants: Ecology, Management, and Conservation of Fragmented Communities. Univ. Of Chicago Press, Chicago

Sachs, L. (1992): Angewandte Statistik. Springer-Verlag, Berlin, Heidelberg, New York

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

White, G. M., Boshier, D. H. (2001): Fragmentation in Central American dry forests: genetic impacts on *Swietenia humilis* (Meliaceae). pp. 293-311 in: Young, A G. and Clark, G. M. (Eds.) Genetics, demography and viability of fragmented populations. Cambridge University Press

White, G. M., Boshier, D. H., Powell, W. (1999): Genetic variation within a fragmented population of *Swietenia humilis* Zucc. Molecular Ecology, 8, 1899-1909

Whitmore, T. C. (1997): Tropical forest disturbance, disappearance, and species loss. pp. 3-12 in: Laurance, W. F. and Bierregaard, R. O. Jr. (Eds.) Tropical forest remnants - Ecology, Managment, and Conservation of Fragmented Communities. University of Chicago Press, Chicago.

Conceptual and experimental elements to model natural inter-specific hybridisation between two mountain southern beeches (*Nothofagus* spp)

L. GALLO

Forest Genetics Unit, INTA Bariloche, C.C. 277, 8400 Bariloche, Argentina

Key words: prezygotic incompatibilities, postzygotic incompatibilities, hybrid fitness, pollination barriers

Abstract: Natural hybridisation has been analysed and modelled mainly from a deterministic point of view within the traditional conception of evolution through which it should always lead to a more stable evolutionary state like a species. Examples of natural hybridisation in trees are scarce and incomplete probably because of the difficulty of identifying species-specific gene markers. Hybridisation between two sympatric mountain southern beeches of the South Andes, *Nothofagus nervosa* and *N. obliqua,* where two isozyme species-specific gene marker alleles, at the protein loci, *Adh* and *Pgi,* have been detected, might prove to be an exception. In this paper, the main conceptual and experimental elements for developing a model of the natural hybridisation between these two species are considered in the framework of their evolutionary importance. To test theoretical aspects of hybridisation species-specific gene markers are needed. The main conceptual elements to model hybridisation are found in the low frequency of creation of F1 individuals, their fitness, and the contribution of their novel genetic combinations to the next hybrid generation. Three types of pre-zygotic incompatibilities are identified acting at three different spatial scales: watershed, slope and site. The inferred unidirectional mating *N. nervosa* (female) x *N. obliqua* (male) is probably a consequence of the better pollen competition ability of the latter species. Post-zygotic incompatibilities are expressed in the scarcity of *F1* hybrids, the drastic negative exogenous selection (low temperatures) they are subjected to (elimination of more than 40 % during the first three years), and the big variation on their fitness. Some of the few *F1* individuals that reach reproductive age are fertile and backcross with both parental species; consequently unidirectional introgression seems unlikely. Different evolutionary stages are inferred in the hybridisation process in different populations across the whole distributional range of both species in Argentina. The combination of all these elements could lead to a long-term evolutionary equilibrium maintained through hybridisation. Natural inter-specific hybridisation should be considered as another important source of genetic variation and the combined genetic system it generates should be treated as a whole when dealing with conservation and management measures.

The problem

Natural hybridisation has been considered from and within the species level as a secondary and deterministic event that ends in better defined or understood

evolutionary status. Different models of natural hybridisation have been developed for annual or biannual plants (Arnold 1997) and their distinctive features have been enumerated (Carney et al. 2000). Hybrid speciation has been shown to occur primarily at secondary contact between populations after allopatric differentiation (Arnold 1997). Prezygotic incompatibilities between populations is assumed to be essential to assure a stable pathway to speciation whereas postzygotic incompatibilities can accelerate the reproductive isolation in sympatric evolution (Gregorius 1992). No model has been described for long-lived trees within the framework of evolutionary change. This contribution tries to point out the main conceptual and experimental elements in modelling the hybridisation process between two deciduous mountain Southern Beeches, *Nothofagus nervosa* (Phil.) Dim. et Mil. and *N. obliqua* (Mirb) Oerst., which occur naturally in the cool-temperate forest of Southwest Argentina.

The hybridisation example

Nothofagus nervosa and *N. obliqua* occur sympatrically in Chilean and Argentinean cool-temperate forests in the southern most part of South America. They are long-lived trees (up to 300 yrs) with very valuable wood and have been abusively logged in the past. These species possess a monecious, protandrous, and anemophilous mating system. Mature or over mature trees (150-250 yr. old) form most of this unmanaged mixed *Nothofagus* forest in Argentina (Chauchard et al. 1997) which is distributed through the glacier lake watersheds in a mountain region that originated during the Tertiary. This region is characterised by strong winds predominantly from the west that cause a drastic precipitation gradient over a very short distance (from 3000 mm a.a.p to 300 in just 150 km). This directional wind pattern also moves pollen mainly from west to east in these populations of *Nothofagus*. Whereas *N. nervosa* in Argentina covers a small but continuous area of just one and a half latitudinal degrees, *N. obliqua* has a fragmented range that extends over more than three degrees. In some of these watersheds they occur sympatrically (See Fig.3, Marchelli and Gallo, this issue). In the areas of overlap, there is an upper, well-defined zone with pure *N. nervosa* (from 800 up to 1,000-m a.s.l) and a lower zone (from 630 (lake level) up to 800 m a.s.l.) with mainly *N. obliqua* forest. In an intermediate zone (700-800 m a.s.l.) a mixed 1:1 forest is found. Finally, descending "tongues" of *N. nervosa* trees down to the lake level and into the *N. obliqua* forest can be observed in some areas.

Stewart (1979) and Tuley (1980) first reported natural crosses between *Nothofagus nervosa* and *N. obliqua* in plantations and arboreta of the United Kingdom. Lennon et al. (1987) proposed the name *N. x alpina* for the hybrid.

Donoso et al. (1990) postulated hybridisation between the same species in seedlings found in a natural population, based on anatomical traits of leaves and wood and flavonoid analysis. Some years later Gallo and collaborators (1997), verified the occurrence of natural hybridisation between both species using open pollinated families. These authors presented growth, morphological and allozymic evidence for natural hybridisation from nursery cultivated progeny. The main direction of the cross was found to be *N. nervosa* x *N. obliqua*, which indicated that it is mainly the latter species that acts as pollen donor. They estimated the self fertilisation rate in one isolated *N. nervosa* tree to be on average 14 % , varying from 6 to 27 % in the four years analysed. (Table 1, mother tree 7). It was also found that those families with the bigger proportions of morphological hybrids presented the higher average second year height and the higher average degree of heterozygosity for three isozyme loci. A detailed study of weekly growth, timing of bud burst in the spring, and termination of growth in the fall was carried out in the same progeny test (Crego 1999). Recently, Marchelli and Gallo (2000a) reported finding large-scale hybridisation by analysis the proportions of hybrids found in populations covering the entire distribution of *N. nervosa* in Argentina.

Requirements for modelling natural inter-specific hybridisation

Three types of information are needed for constructing a basic model of natural inter-specific hybridisation: 1) a clear detection of *species-specific gene markers* (quantitative and/or qualitative), 2) knowledge about probable *pre-mating barriers* (pre-pollination and post-pollination), and 3) knowledge about probable *post-mating barriers* (endogenous or intrinsic and exogenous or extrinsic selection effects). If this information is available a conceptual model for the first generation of hybridisation can be proposed. Further information on *F1* fertility and *F2* or backcrosses fitness would suggest the evolutionary significance of the hybridisation and how this process contributes to the generation and maintenance of intra-generic stable genetic variation.

Species-specific gene markers in *Nothofagus nervosa* and *N. obliqua*

New evidences from adult trees in the natural forest was recently presented by Gallo et al. (2000) to demonstrate species specificity of the previously detected

isozyme markers. Based on these data, a more complete and reliable determination of isozyme gene markers could be carried out in both species (Marchelli and Gallo 2000b, Azpilicueta and Gallo 2001). The species specificity of the gene markers has been therefore tested in different studies using 4500 embryos from 40 populations and 70 families, 540 adult trees from eight populations, and 334 two year old progenies from four families. *Nothofagus nervosa* possessed genotypes *Adh 2-2* and *Pgi 1-1* and *N. obliqua Adh 1-1* and *Pgi 2-2,* whereas *F1* hybrids were *Adh 1-2* and *Pgi 1-2.*

Pre-pollination barriers in Nothofagus nervosa and N. obliqua.

Transport of pollen by wind is normally assumed to be a more or less massive and stochastic mechanism that induces uniform and random pollination. Quantitative genetic estimations used to predict genetic gains from open pollinated families are based mainly on this assumption. Wind pollination has been thought possible over long distances (more than 300 km) in some New Zealand *Nothofagus* species and long-distance hybridisation has been proposed as a principal mechanism responsible for the present disjunct distributions of some species (Wardle et al. 1988). Wind frequency, direction, and intensity during pollination time may vary greatly with different physiographic conditions. This is of special importance in mountain regions like the eastern slopes of the South Andes, where *Nothofagus nervosa* and *N. obliqua* occur. In this region, about 85 % of the winds move from the West to the East during the period of pollination. These winds are also stronger than those coming from other directions. The air flow is further canalised by the East-West orientation of the glacier valleys where these species occur. Under these conditions it is likely that such unidirectional wind corridors result in unidirectional gene flow. Tree population probably receives little pollen from the eastern side.

Flowering phenology constitutes the main pre-mating barrier for wind pollinated plants. As has been described for other *Nothofagus* spp. and is accepted in general for temperate woody species, flowering time depends on temperature. In *Nothofagus* species flowering begins soon after leaf budding (Wardle 1984). Within the mountain species *Nothofagus pumilio* flowering was synchronous with leaf formation and began ten days earlier at a low altitude (1,225 m a.s.l.) than 165 m higher (Rusch 1993). Therefore leaf budding can be used as an indirect way to register differences in flowering time between individuals. *Nothofagus nervosa* occurs naturally in pure stands at higher altitudes (about 200 m) than typical *N. obliqua* habitats, and bud flushing is also normally later in nervosa forests. However phenological differences are reduced in the mid- altitude mixed 1:1 forest. A three year old progeny test where *N. nervosa* open pollinated families were grown together with a commercial sample

of *N. obliqua* at the same elevation showed no significant differences in the beginning of spring growth among pure individuals of both species or with their inter-specific hybrids (Crego 1999). Further, in seed samples from isolated *N. nervosa* trees surrounded by *N. obliqua*, the proportion of hybrids detected in *Adh* was very high and present in all four years for which samples were taken (Table 1). Thus flowering phenology should not be a barrier for mating between these species at sites where they occur together.

Table 1. Variation in the proportion of hybrids seeds from up to four sampled years produced by differentially isolated *N. nervosa* mother trees (+++ nearest *N. nervosa* trees at 1000 m;++ nearest *N. nervosa* trees at 100 m; + nearest *N. nervosa* trees at 10 m)

Mother tree	Isolation degree	Year	Adh genotypes		
			1/2	2/2	N
7	+++	1994	0.94	0.06	117
7	+++	1996	0.95	0.06	108
7	+++	1998	0.84	0.16	56
7	+++	2000	0.73	0.27	60
1	++	1994	0.52	0.48	159
1	++	1996	0.25	0.75	72
1	++	2000	0.78	0.22	74
12	++	1994	0.59	0.41	68
8	+	1994	0.06	0.94	142
16	+	1994	0.00	1.00	70

Post-pollination barriers

Natural hybridisation in plants may be limited by pollen germination capacity, penetration of the stigma surface, and pollen tube growth into the ovary (Arnold 1997). In the whole process leading up to fertilisation, competition between con-specific and hetero-specific gametes is likely to be very important in determining the frequency of hybrid progeny. In *N. nervosa* trees that are sufficiently isolated and surrounded by *N. obliqua* pollen donors, no apparent impediment to fertilisation is found, even when considering different years (Table 1). We can see that the proportion of hybrids (*Adh-1/2*) increases with the degree of isolation of *N. nervosa* trees in relation to con-specific individuals. Such a situation means that the proportion of hetero-specific gametes in the

pollen cloud also increases. On the other hand, recent analysis of 40 *N. obliqua* individuals from the 1:1 mixed forest showed no *F1* hybrid seeds in their progeny. This results highlights two important elements: 1) pollen competition must be the main limiting barrier for the production of *F1* hybrids and 2) *N. nervosa* pollen apparently has a lower competitive capacity than that of *N.obliqua* which could explain the directional mating found (*N. obliqua* acting as pollen donor).

Post mating barriers (Hybrids fitness)

Leaf margins of one and two-year old plants, growing at a nursery in a progeny test, were morphologically analysed. Those individuals belonging to the same *N. nervosa* open pollinated family (half sibs) that had leaf margins like *N. obliqua* were assumed to be hybrids. In a later analysis a tied correspondence between the proportion of "*obliqua*-morphotypes" (putative hybrids) and the proportion of the expected allozyme hybrid genotypes was found. This phenotypic-genotypic correspondence could not be shown in adult hybrid individuals in the forest. Analysis of about 100 adult trees in phenotypically pure stands of each species and about 200 in 1:1 mixed populations in the natural forest resulted in more *Pgi* allozymic hybrids (7.5 %) than those that could be detected through their intermediate bark traits (only 3 %). About 90 % of the *Pgi*-hybrids appeared in the mixed 1:1 forest and half of them were masked in *N. obliqua* phenotypes (Gallo et al. 2000). Assuming phenotypically hybrid expression as corresponding to double heterozygous individuals, then 3 % would be the maximum estimated proportion of putative *F1* adult individuals in the analysed mixed forest. This observation agrees with the distribution of the height data of young hybrids measured in the nursery. Although on average hybrids grew better in the nursery, the observed variation was also large. Sixty percent of the hybrids showed very low growth, 25 % showed middle values and 15 % were among the tallest three-year old trees. This same large variation in the growth of hybrids was recently observed in other forest trees species (e.g. in *Eucalyptus* spp. Lopez et al. 2000). In spite of the on-average better growth, many hybrids (up to 40 %) died during the first three years. Hybrids terminate their growth later in the fall and damage from early frosts was observed on terminal and axial shoots. Hybrids are eight times more common in seed samples of a progeny test than in adult trees of the same population. A strong selection against *F1* hybrids is taking place. However some hybrids are able to reach sexual maturity.

F1 fertility

Experimental data from a genetic differentiation study among seed populations of *N. nervosa* suggested that *F1* individuals are very scarce but fertile (Marchelli and Gallo in press). Seed genotypes *Adh1-2/Pgi1-2* (*F1* hybrids) were found just in three populations and with a frequency lower that 6.7 % (Table 2).

Table 2: Genotypic frequencies (%) in *Adh* and *Pgi* marker loci indicating probable *F1* and *Backcrosses* in *Nothofagus nervosa* seed populations Pure *N. nervosa* seeds have *Adh 2-2/ Pgi 1-1* genotypes * 2.7 % of the seed sample with genotypes *Adh 1-1, Pgi 1-1* (From Marchelli 2001, modiffied)

Population	N	F1	Backcross	
		Adh Pgi 1-2 1-2	Adh Pgi 2-2 1-2	Adh Pgi 1-2 1-1
Espejo chico	106	0	1	0
Pudu-Pudu	133	0	0	0
Pla. Rauli	116	0	0	0
Escondido	117	0	1.7	0
Bandurrias	115	3.5	1.7	0.9
Chidiak	104	6.7	1	0
Quilanla-hue	120	0	0.8	0
Hua-Hum	120	0.8	0	0
Queni	115	0	0	0
Lolog 17	103	0	0	17.5
Lolog 18*	112	0	4.5	24.1
Lolog 19	115	0	0	0
Curruhue 6	98	0	0	0
Curruhue 21	126	0	0.8	0
Huechulafquen	130	0	0	0
Bahia Azul	126	0	0	0
Paimun	99	0	0	0
Tromen 27	129	0	2.3	0
Tromen 28	108	0	0	0
Quillen	105	0	1	0

On the other hand, probable backcrosses (hybrids *Adh2-2/Pgi1-2* and/or *Adh1-2/Pgi1-1*) were found in 10 of the 20 analysed *N. nervosa* populations with a maximum frequency in one of them of about 30 %. Independent

segregation of the marker loci has to be assumed. The scarcity of F1 adult genotypes and the very low self-fertilisation rate found in these and other related species (about 10 %), suggest backcrossing with *N. nervosa* to be the most reliable explanation for these results. In one population seed genotypes *Adh 1-1/Pgi 1-1* were also found with a frequency of 2.7 % suggesting the fertility of the backcrosses-hybrids. Preliminary data obtained from pure *N. obliqua* adult tree sample populations showed allozyme genotypes *Adh1-1/Pgi1-2* and *Adh1-2/Pgi2-2* in almost 35 % of the analysed adult individuals. Further, very recently, viable seed was collected from one *N. obliqua* adult hybrid individual (Azpilicueta and Gallo 2001). This indicates that backcrosses of F1 to *N. obliqua* have also taken place and that F1 individuals act as a "Hybrid Bridge" between both parental species. Finally, the large variation detected in the proportion of *F1* and backcross hybridisation events in the analysed population suggests that evolutionary changes are taking place at different stages.

Concluding remarks for the conceptual model

The experimental data and observations presented here provide information which allows me to propose a conceptual "Model of hybridisation in *Mountain southern beeches*". Some of the experimental elements are considered in part of the different natural hybridisation models described elsewhere (e.g. Arnold 1997) but they are presented here for the first time in the hybridisation system of long-lived organisms. Moreover, particular elements of the mountain landscape, like variation of temperature with altitude or probable gene flow corridors, are taken into account to understand the hybridisation system in these wind-pollinated species. The proposed conceptual model may also be relevant to hybridisation in other wind-pollinated temperate tree species. Part of the main elements of the model is graphically presented in Figure 1. The fitness of the parental species and of the hybrid (*N. x alpina*) are considered in relation to the altitude in a typical slope of the sympatric area.

Pre-zygotic incompatibilities:

Three types of prezygotic incompatibilities are acting in different spatial scales:

- "Watershed scale" At this scale the unidirectional wind during pollination time would prevent pollination from eastward side located populations. (Pre-pollination barrier)
- "Slope scale": Temperature adapted flowering time would build a

phenological barrier between *N. nervosa-N. obliqua* and *N. nervosa-N. nervosa* populations located at different altitudes (pre-pollination barrier) (Fig. 1).

- "Site scale": In 1:1 mixed populations, pollen competition (post-pollination barrier) between hetero-specific and con-specific pollen grains determines the proportion of F1 hybrids generated. Because *N. obliqua* pollen seems to compete better, directional mating of *N. nervosa* (as female) x *N. obliqua* (as male) takes place.

Figure 1. Schematic presentation of the relation between fitness (*w*) and altitude in a typical slope where *N. nervosa* and *N. obliqua* occur sympatrically.

Post-zygotic incompatibilities

- The generation of *F1* individuals is rare (less than 7 %), even in the 1:1 mixed populations.
- There is large variation in the fitness of *F1* individuals (Fig. 1). No generalisation about their expected fitness can be made since some hybrids could be better adapted than the parental species whereas most of them do no even reach the reproductive age.

- Exogenous selection (probably a result of low temperatures) can be assumed. Such selection eliminates, on average at least 40 % of *F1* individuals during the first three years. Endogenous selection, such as disruption of co-adapted genes, from the fertilization up through the adult stages could be responsible for part of the *F1* mortality but could not be analysed here.

F1 fertility

- At least some *F1* individuals that reach the reproductive age are fertile and backcross with both parental species.
- Backcross hybrid seed can reach considerable proportions in some populations and backcross hybrids seem to be more fit than *F1* hybrids.

Evolutionary considerations

- Different evolutionary stages in the hybridisation process can be assumed in different areas of the natural distribution of both species.
- Unidirectional introgression seems unlikely since *F1* individuals seem to act as a "Hybrid bridge" between both parental species.
- An equilibrium between generation and elimination of *F1* individuals could maintain stable hybrid zones. These equilibrium frequencies could be modelled in order to simulate the long-term consequences and speculate as to the evolutionary paths expected after climate change.
- Selective preferential timber extraction of *N. nervosa* in the past has probably increased the proportion of *N. obliqua* pollen in some mixed populations and consequently the current generation of *F1* hybrid seeds. Their diminished fitness might reduce the whole fitness of some populations through limited recruitment of naturally regenerated seedlings.
- As a result of the hybridisation taking place in some areas, the local heterospecific gene pool may be more important for the evolution of both species than their own geographically or altitudinally separated gene pools.

Acknowledgements

I would like to thank Sheila Ward and Lyn Loveless for their comments and suggestions to improve the quality of this contribution. The present work was partially financial supported by different projects. I would like to thank IFS (International Foundation for Science), DAAD-GTZ (Germany), SAGPyA-BIRF and BID-FONTAGRO. My special thanks to the Silviculture Department of National Parks Administration in Patagonia, with whom we began this and

other related works. I am greately indebted to Paula Marchelli, Paula Crego, Lucas Oudkerk and María Marta Azpilicueta, from the Forest Genetics Unit, for their professional and technical assistance.

References

Arnold, M. L. (1997): Natural Hybridization and Evolution. Oxford University Press. 215 pp.

Azpilicueta, M. L., Gallo, L. A. (2001): Análisis de la variación genética en *Nothofagus obliqua* (Mirb.) Oerst. a través de marcadores génicos isoenzimáticos. XXX Congreso Argentino de Genética. 16-19 septiembre 2001, Mar del Plata, Argentina.

Carney, S.E., Wolf, D.E., Rieseberg, L.H. (2000): Hybridisation and forest conservation. pp. 167-181 in: A., Young, D. , Boshier, T., Boyle (Eds.) Forest Conservation Genetics. Principle and Practice. CSIRO Publishing-CABI Publishing.

Chauchard, L., Sbrancia, R., González Peñalba, M., Rabino, A., Maresca, L. y., Blacher, C. (1997): Dinámica y manejo del bosque de *Nothofagus*. Informe final de Proyecto. Universidad Nacional del Comahue-Administración de Parques Nacionales-Instituto Nacional de Tecnología Agropecuaria. 93 pp.

Crego, M.P. (1999): Variación genética en el comportamiento fenológico y el crecimiento juvenil de progenies puras e híbridas de raùlí, *Nothofagus nervosa* (Phil.) Dim. et Mil. Tesis de Grado. Universidad Nacional del Comahue. 105 pp.

Donoso, C., Morales, J. y., Romero, M. (1990): Hibridación natural entre roble (*Nothofagus obliqua*) (Mirb.) Oerst. y raulí (*N. alpina*) (Poepp. & Endl.) Oerst, en bosques del sur de Chile. Revista Chilena de Historia Natural 63: 49-60

Gallo, L.A., Marchelli, P., Breitembücher, A. (1997): Morphological and allozymic evidence of natural hybridization between two southern beeches (*Nothofagus* spp) and its relation to heterozygosity and height growth. Forest Genetics 4: 15-23.

Gallo, L., Marchelli, P., Crego, P. (2000): Hybridization dynamic between two southern beeches (*Nothofagus* spp.). XXI IUFRO World Congress- 7-12 August 2000, Kuala Lumpur, Malaysia.

Gregorius, H-R. (1992): Population genetic keys to speciation. Göttingen Research Notes in Forest Genetics 13, 19 p.

Lopez, G. A., Potts, B. M., Tilyard, P. A. (2000): F1 hybrid inviability in *Eucalyptus*: the case of *E.ovata x E. globulus*. Heredity 86: 242-250

Marchelli, P. (2001): Variabilidad genética en Raulí (*Nothofagus nervosa* (Phil.)Dim. et Mil., su relación con procesos evolutivos y la importancia en la conservación y utilización de sus recursos genéticos. Tesis Doctoral. Universidad Nacional del Comahue. 214.pp.

Marchelli, P., Gallo, L.A. (1999): Annual and geographic variation in seed traits of Argentinean populations of southern beech *Nothofagus nervosa* (Phil.) Dim. et Mil.. Forest Ecology and Management 121: 239-250

Marchelli, P., Gallo, L.A. (2000a): Large scale hybridisation process between two *Nothofagus* species from Southamerican Temperate Forests. III Southern Connection Congress. 17-22 January 2000, New Zealand. P. 48.

Marchelli, P., Gallo, L. (2000b): Genetic analysis of isozyme variants in open pollinated families of Southern beech *Nothofagus nervosa* (Phil.) Dim. et Mil.. Silvae Genetica 49: 90-98.

Marchelli, P., Gallo, L. (in press): Genetic diversity and differentiation in a southern beech subjected to introgressive hybridisation, Heredity

Rieseberg, L.H. (1995): The role of hybridization in evolution: old wine in new skins. American Journal of Botany 82: 944-953

Rusch, V.E. (1993): Altudinal variation in the phenology of *Nothofagus pumilio* in Argentina. Revista Chilena de Historia Natural 66: 131-141

Stewart, P.J. (1979): Le genre "*Nothofagus*" et son utilisation dans la sylviculture Britannique Rev. For. Fr. 31: 473-482.

Tuley, G. (1980): *Nothofagus* in Britain. Forestry Commission. Forest Record 122, 26 pp.

Wardle, J.A. (1984): The New Zealand Beeches. Ecology, Utilisation and Management. New Zealand Forest Service, 447 pp.

Wardle, P., Harris, W., Buxton, R. P. (1988): Effects of glacial climates on floristic distribution in New Zealand. 2. The role of long-distance hybridisation in disjunct distributions. New Zealand Journal of Botany 26: 557-564.

Part 3

Gene flow and population differentiation

Spatial pattern of gene flow in a scattered woody species, the wildservice tree

S. ODDOU-MURATORIO[1], F. AUSTERLITZ[2], S. GERBER[2], and B. DEMESURE[1].

[1] *Conservatoire Génétique des Arbres Forestiers, ONF, campus INRA F-45160 Ardon, France.*
[2] *INRA, Station de Recherches Forestières - BP 45 F-33611 Gazinet cedex, France.*

Key-words: **microsatellite, wildservice tree, genetic structure, pollen flow, paternity analysis.**

Abstract: We used microsatellite markers to analyze patterns of local genetic structure and pollen dispersal in the wildservice tree, a scattered, deciduous species with good colonization abilities. Within a local population covering 450 ha, the low adult density and their scattered distribution resulted in a strong pattern of isolation by distance, despite high gene dispersal abilities. Pollen mediated gene flow was directly evaluated through paternity analysis of 100 offspring collected on a single mother-tree, around which potential males were exhaustively sampled within a radius of 500m. Finally, these direct and indirect estimates of gene flow were compared.

Introduction

Patterns of gene flow within and between populations have been extensively investigated in many tree species, either through the study of genetic structure ("cumulative gene flow"), or more recently, by paternity and parentage analysis ("current gene flow"). Some general trends in pollen- and seed-mediated gene flow dynamics of woody species have been identified, and in particular the extensive level of pollen flow and/or the importance of population size. In some cases, these trends seem incompatible with long-standing assumptions of population biology: for example, tropical tree species have been predicted to be predominantly self-fertilizing or inbreeding, because of their scattered distribution and animal-mediated pollination. However, many studies of mating patterns in tropical trees have revealed that most of these species are outcrossed, and that apparent pollen flow greater than hundreds of meters can usually be observed (e.g., Stacy et al. 1996, Kaufman et al. 1998, Dick 2001).

Some tree species of temperate forests are ecologically very similar to tropical species: in particular, the so-called scattered species, which combine extensive

range and low local density with high colonization abilities, but low success in competing with other tree species. This is true of many broad-leaved and notably fruit trees, such as *Prunus, Pyrus, Malus* or *Sorbus* species. These life history traits result in typical extinction-recolonization dynamics, with nomadic sub-populations fluctuating over space and time. Population studies over the range of species have revealed high seed flow abilities of colonizing species (Raspe et al. 1999). In the case of *Sorbus torminalis*, an European scattered species, we could show that at the scale of one hundred kilometers, seed flow is at least as important as pollen flow (Oddou-Muratorio et al. 2001). We can thus expect that seed flow will also be a major component of gene flow at the local scale of a forest, or of a stand.

In 1998 a 450 ha intensively studied plot (ISP) was established to evaluate local population structure, as well as pollen and seed dispersal in the wild service tree (*S. torminalis* L. Crantz). In this short communication, we will present some initial results of this survey. We address the following questions: (1) Is there a limitation to pollen dispersal in species with such a patchy distribution and entomophilous pollination? (2) Does the balance observed at broader scale between pollen- and seed-mediated gene flow persist at local scale? For this purpose, we studied current pollen flow through paternity analysis. Then, we used the spatial structure of genetic diversity observed among adults to obtain an indirect estimate of gene dispersal. Finally, these direct and indirect estimates of gene flow were compared.

Materials and methods

Study organism

Sorbus torminalis is an entomophilous species, pollinated by a wide range of bees and Diptera. Its fleshy fruits are disseminated by birds and mammals, and it is usually described as a post-pioneer species.

The intensively studied plot (ISP) was chosen in the northwest of Rambouillet forest (Yvelines, France). This site is approximately rectangular, and it groups 22 parcels of mixed oaks and broad-leaved stands, at different stages of high forest treatment. Throughout these stands, we collected and mapped 185 reproducing wild service trees. The reproductive status of these adults was assessed on the basis of flowering observations. In autumn 1999, 100 seeds were sampled on one of these adult trees (Figure 1).

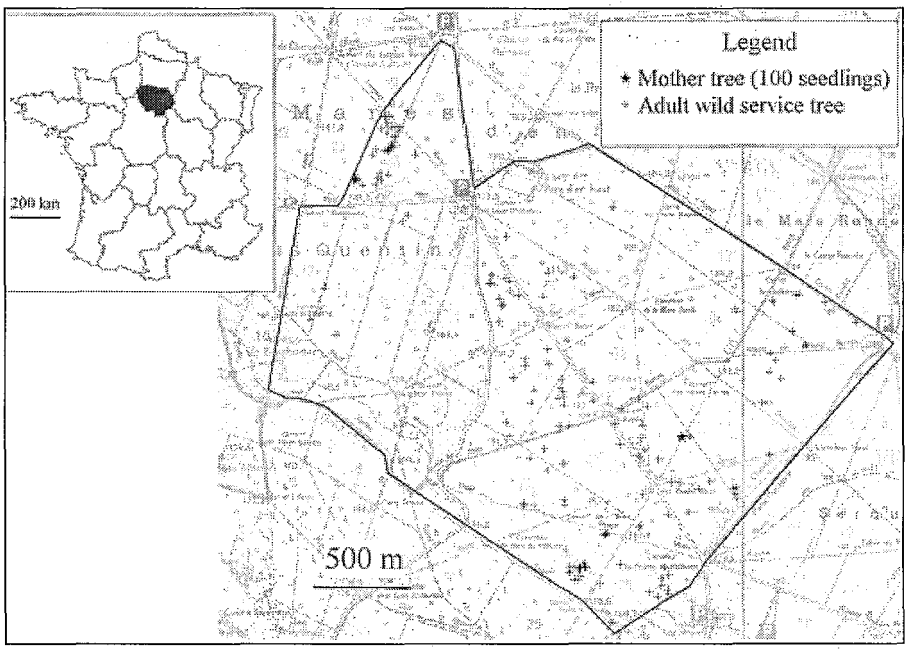


Figure 1. Map the Rambouillet intensive studied plot (450 ha) representing reproducing individuals and the sampled mother tree.

Genetic analysis

DNA was extracted from frozen leaves for the adult trees, and directly from embryo with the cotyledons for the seeds following the DNeasy procedure (Qiagen Inc.). All the individuals were genotyped at 6 microsatellite loci as described in Oddou-Muratorio et al. (2001).

Data analysis

Spatial structure: The spatial component of fine-scale genetic structure was assessed by mean of auto-correlation analysis, using as statistics the genetic distance measure proposed by Rousset (2000). This measure, called $F_0/1\text{-}F_0$, is an analogue of the $F_{ST}/1\text{-}F_{ST}$ ratio using pairs of individuals instead of pairs of populations (Rousset, 1997). The coefficients $F_0/1\text{-}F_0$ were computed between all possible pairs of reproducing individuals, grouped in ten distance classes (Figure 2). The graphic representation of pairwise $F_0/1\text{-}F_0$ coefficients against distance is called a correlogram. The matrix of pairwise $F_0/1\text{-}F_0$ coefficients was regressed on the matrix of the logarithm of the pairwise spatial distances, and a Mantel test was carried out to detect a correlation between the matrices (by 10000 random permutations of spatial locations among individuals).

Computation of the statistics and Mantel tests were performed using the software AUTOCORG, ver 3.0, developed by Hardy and Vekemans (available on request to O. Hardy, ohardy@ulb.ac.be).

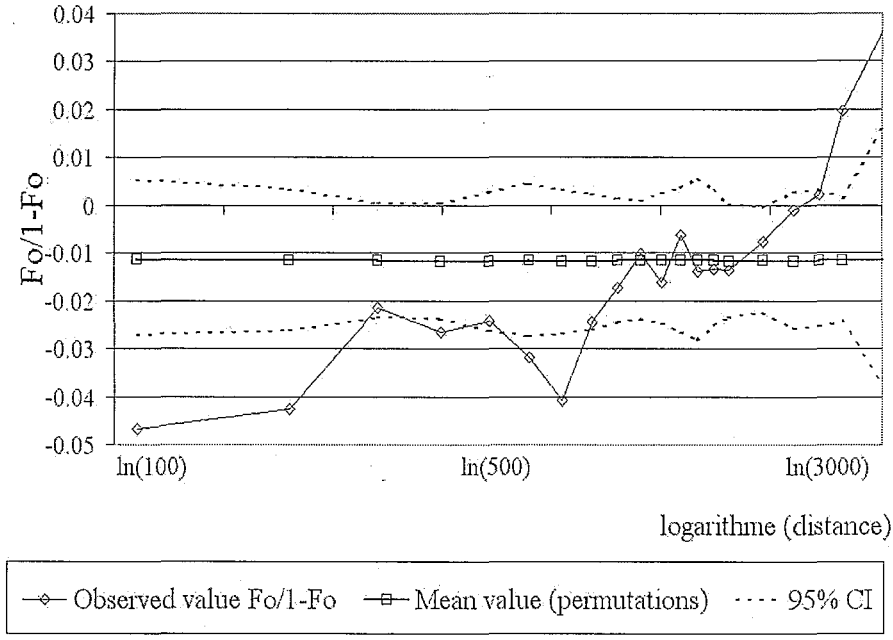


Figure 2. Average pairwise $F_0/1\text{-}F_0$ coefficients between reproducing individuals as a function of the spatial distance (in meter). Dashed lines represent the 95% confidence interval around the mean value of $F_0/1\text{-}F_0$ expected under the hypotheses that differentiation between individual does not depend upon distance. These null model was constructed by the mean of 5000 permutations of individual positions.

Indirect inference of gene flow through pattern of spatial genetic structure: Rousset (2000) showed that the slope of the regression of pairwise $F_0/1\text{-}F_0$ coefficients against spatial distances can be used to provide an estimate of gene dispersal distances. For a two-dimensional space, defining the neighborhood size N_b as $4\Pi . D\sigma^2$, where D is the "effective" population density and σ^2 the variance of gene dispersal distances (distances measured on an axial scale), we have the following relation: Nb=1/b, where b is the slope of the regression of pairwise $F_0/1\text{-}F_0$ coefficients against the logarithm of distance.

Such indirect estimates are to be taken with caution because they require assumptions (e.g. steady-state equilibrium) that are not necessarily met in nature. But when a linear relationship is observed between $F_0/1\text{-}F_0$ and the logarithm of distance, the slope expresses the degree of genetic structuring and contains most of the information regarding intra-locus structure for two-gene relationships.

Paternity analysis: for each offspring, knowing the maternal genotype, the genotype of each sampled reproductively mature individual in the ISP (including the mother) was checked for its compatibility as the paternal genotype. This simple exclusion procedure allowed us to make binary conclusions concerning the origin of the successful pollen grain: outside the ISP ("external") or inside the ISP ("local"). Moreover, in order to infer precise paternity relationships when multiple males were not excluded, we used likelihood paternity methods as originally proposed by Meagher (1986). Marshall et al. (1998) proposed an extension of this method, implemented in a software called CERVUS ver 1.0, that allows the user to assign paternity to the male parent with the highest likelihood probability after a specific likelihood test. This likelihood test is calibrated by simulations of a panmictic population and requires the following parameters: the allele frequencies at each locus (we used those of all reproductive adults collected), the error rate of genotype determination (assumed to be null), and the total number of reproducing individuals (n = 149 individuals with flowering corymbs in 1999). The last parameter to be given is the percentage of the reproducing individuals with a known genotype: it allows an average fraction of candidate males for which no genetic data are available, i.e. immigrant pollen flow, to be introduced. We obtained elsewhere an estimate of the total number of reproducing individuals that contributed to sire the analyzed progenies in 1999 (Oddou-Muratorio et al., in prep. 1), which shows that the sampling fraction of the breeding male population amounted to around 50%.

Comparison between cumulated and immediate estimate of pollen flow: analyses of spatial structure and paternity assignment provides us with two estimates for the pollen flow at a local scale:

- A direct estimate of the effective pollen pool size (N_{ep}), through the paternity analysis. Indeed, if pollen dispersal distances are normally distributed with equal variances along the two directions, the effective pollen pool size is a circle of radius $2\sigma_p$ (where σ_p^2 is the axial variance of pollen dispersal), of area A and within which 86.5% of the male parents will be found: $N_{ep} = 4\pi D\sigma_p^2$ (Austerlitz and Smouse 2001).
- An indirect estimate of N_{ep} based on the regression analysis of F_O /1-F_O estimates among adults. Indeed, for a nuclear gene in plants, the variance of gene dispersal distances is linked to the axial variance of pollen and seed dispersal (respectively σ_p^2 and σ_s^2) as follows:

$$\sigma^2 = \frac{1}{2}\sigma_p^2 + \sigma_s^2 \qquad \text{(Crawford 1984)}$$

and thus, under the hypothesis of equal pollen and seed dispersal ($\sigma_p^2 = \sigma_s^2$), we have:

$$N_{ep} = \frac{2}{3} N_b$$

To test this hypothesis of equal abilities of pollen and seed dispersal, we simply estimated the pollen neighborhood area from the spatial structure among adults. Then, we computed the observed percentage of male parents lying within this area around the maternal tree.

Results

Genetic structure among adults, and cumulated gene flow

The distribution of genetic diversity among adult individuals has been thoroughly described in Oddou-Muratorio et al. (in prep 2). Here, we will focus on the parameters that allow an indirect estimate of N_b (the neighborhood size) to be obtained. The $F_0/(1- F_0)$ values were smaller than expected by chance for distance up to 750 m (Figure 2). The overall slope was significant, and provided an N_b estimate of 38 individuals. Taking into account the density of adults observed in the Rambouillet ISP (D = 0.41 individuals/ha), the neighborhood has an area of 95 ha, and the standard deviation (σ) of gene dispersal distance is 274 m.

Paternity analysis, and current gene flow

Among the 100 analyzed offspring, 31 found no compatible father by simple exclusion from among the reproducing individuals collected. We will thus assume that the rate of seeds sired by males outside the experimental plot was 31%. Among the remaining 69 offspring, 53 found only 1 compatible father, and the 16 others matched with 2, 3, 4 or 5 adult genotypes. Among the 53 offspring assigned to one father alone, 36 were sired by the same male, the nearest neighbor of the mother-tree (at 40m). The most distant father stood at 2090 m from the mother-tree.

The maximum-likelihood procedure allowed us to solve 9 cases of multiple paternity out of the 16 we recorded. Bulking together the 62 offspring assigned to one father only by simple exclusion or maximum-likelihood method, the mean pollination distance was 290m (s.d. = 444 m) and the distribution of these distances showed an excess of short-distance pollination events (median = 40 m).

The distance between the mother-tree and the nearest limit of the experimental site was 500 m. Bulking together the offspring originating from external males (31), and the offspring sired by local mates located at more than 500 m from the mother tree (13), we could conclude that in 44 cases out of 93 (cases when the origin of the pollen can be precisely inferred) the successful pollen grain traveled more than 500m.

Comparison of cumulated and current gene flow estimate.

The spatial pattern of genetic structure provided us with an indirect, mean estimate of N_h = 38 individuals Under the hypothesis of balanced seed and pollen gene flow in the wildservice tree, the pollination neighbourhood around the mother-tree would thus be N_{ep} =2/3 N_h = 25 individuals. Using the density of flowering individual to estimate D ($D \approx 3.3\ 10^{-5}$), this would correspond to a circle of radius 246 m, 78 ha in area, centred at the mother in question. This circle theoretically encompasses 86.5% of the fathers of central progeny.

On the other hand, the direct study of pollen dispersal through paternity analysis allowed us to estimate the proportion of male parents found within a circle of radius 500m around a mother tree, that is precisely within this 78 ha area that would correspond to the pollination neighborhood. Within this area, only 49 offspring (that is 52% of the offspring with known pollen origin) found a compatible male parent.

Discussion

The objective of this first report on the local genetic dynamics in the wildservice tree was to evaluate the abilities of gene flow both indirectly, through spatial pattern of genetic structure, and directly, through paternity analyses.

A clear pattern of isolation by distance was detected within the adult population: individuals were more related than expected by chance within the first 750 m. This spatial pattern of genetic structure indicates high abilities of gene dispersal, combined with a very low effective population size (38 individuals, corresponding to 95 ha). In comparison, using geostatistical analysis of spatial distribution of allelic frequencies, Le Corre et al. (1998) found for *Quercus*

petraea (a widespread oak species) a neighbourhood size of 3000-4000 individuals, equivalent to a neighbourhood area between 15 and 20 ha. In the common ash (*Fraxinus excelsior*, a low-density species, that can be locally abundant though not dominant in forest stands) the neighbourhood size ranged between 38 and 126 individuals (for a set of 8 analysed populations), that is between 0.63 and 0.81 ha (Heuertz et al. 2001). These contrasting spatial patterns of genetic structure result from differences in dispersal modalities by pollen and by seed, but also from different patterns of local density and aggregation. Though obviously limited by distance, pollen and seed dispersal in the wildservice tree (entomophilous, with animal-dispersed seeds) appear at least as efficient as in oaks (anemophilous, with a combination of gravity and bird-mediated seed dispersal) and common ash (anemophilous, with heavy wind-dispersed seed). The combination of these high abilities of gene dispersal with the strong pattern of isolation by distance observed in the wildservice tree can be highlighted by considering the influence of low population density: as suggested by Hamrick et al. (1993), scattered distribution of individuals within stands enhances the genetic structure, due to weak overlapping of individual seed shadows.

This combination of extended dispersal and of preferential mating with the few close neighbors was also revealed by paternity analysis. Whereas 36% of the seeds were sired by the mother's nearest-neighbor, in more than 47 % of the observed pollination events, the identified male parents stood at more than 500 m from the mother tree. Pollen dispersal events at more than 1 km were not rare. The important extent of pollen flow observed in the wildservice tree is consistent with direct estimates of pollen dispersal obtained in other entomophilous tree species (mainly tropical species, e.g., Stacy et al. 1996, Kaufman et al. 1998, Dick, 2001).

Though preliminary, these results show the power of paternity analyses for estimating levels of pollen flow at an ecological time scale, and for evaluating how landscape characteristics influence pollen dispersal in plant populations. In this paternity study, the simple exclusion procedure gave rather good results, because the molecular markers we used yielded a high level of variation (Streiff et al. 1999). Besides the simple exclusion procedure, statistical methods of paternity assignment potentially allow us to account for various ecological factors affecting mating patterns (that can be included in the odds of paternity of each potential male; e.g. Devlin et al. 1988). Moreover, maximum-likelihood methods should thus provide more information than the simple exclusion method, as they allow us to assign paternity to more progeny. However, the use of these statistical methods also requires reliable estimates of the total size of the reproducing population, and of the proportion of that population that was sampled. In the CERVUS procedure, for instance, these parameters directly influence statistical confidence in the inferred father-offspring relationships. This implies that only

studies that combine paternity surveys with field observations, or use refined statistical methods of paternity assignment, are amenable to minimum-biased estimates of pollen flow (see Oddou-Muratorio et al., in prep 1 for a comparison of different procedures). This requirement may actually be quite salutary, as it is only by accounting for ecological and demographic population parameters that a global understanding of mating patterns in natural populations can be achieved.

Finally this study allowed us to compare a direct, contemporaneous estimate of pollen flow (estimated from paternity analysis between only two successive generations) with an indirect, mean value of gene flow over several generations derived from population structure. The levels of pollen flow estimated by paternity analysis appeared slightly higher than those inferred from population structure under the hypotheses of balanced gene flow by pollen and by seed. Less than 52 % of the fathers were found within the area corresponding to the indirect estimate of pollination effective size, although that is where 86.5% of the male parents were expected. In the few studies that have compared direct and indirect gene flow estimates at a local scale, good consistency was generally obtained (Rousset 2000, Hardy et al. in prep), though confidence intervals around estimates are usually large. Additional progenies sampled on different mother trees are needed to confirm the discrepancy between indirect and direct estimates of gene flow in this local population of the wild service tree. We present elsewhere results of paternity analysis acquired during two years of experimentation on different mother-trees (Oddou-Muratorio et al. in prep 1), which tend to show that the extent of pollen dispersal observed here may be the rule in the wildservice tree. It is also possible that our indirect estimate of gene flow through population structure may be biased, due to violation of some of the hypotheses underlying its computation. This estimates requires an equilibrium between migration and drift within the population, which cannot be assumed readily in a species with frequent extinction-colonisation events. However, the strong pattern of isolation by distance and the local scale of the investigation provide some reassurance that bias in the indirect estimate of gene dispersal is minimal (see also Oddou-Muratorio et al. 2001 for discussion). At first glance, it thus seems that the hypotheses of balanced gene flow by pollen and seed may not be valid at the local scale. With a pattern of seed dispersal depending upon distance, there could be preferential short distance seed dispersal, combined with extensive long-distance dispersal at the metapopulation scale. Under this scheme, pollen flow would be more important than seed flow at a local scale.

In conclusion, paternity analysis in *S. torminalis* revealed high levels of pollen flow. These results are consistent with the biology of the species, and notably its pollination by bee species, which have been shown to be very efficient in tropical tree species. Seed dispersal patterns might vary with distance, and

appear to combine rare, long-distance movements with preferential, short-distance dispersal. These questions will need to be specifically investigated either through parentage analysis, or through behavioural observations of birds and mammals activities during fruit dissemination.

Acknowledgement

We thank the organising committee of the symposium "Modelling and experimental research on genetic processes in tropical and temperate forests". Participation of S.O.M. at the symposium was supported by a grant from Silvolab Guyane.

References

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

Crawford, T. J. (1984): The estimation of neighbourhood parameters for plant populations. Heredity 52:273-283

Devlin, B., Roeder, K. Ellstrand N. C. (1988): Fractional paternity assignment: theoretical development and comparison to other methods. Theoretical and applied Genetics 76:369-380

Dick, C. (2001): Genetic rescue of remnant tropical trees by an alien pollinator. Proc. R. Soc. Lond. 268: 2391-2396.

Hamrick, J. L., Murawski, D. A., Nason, J. D. (1993): The influence of seed dispersal mechanisms on the genetic structure of tropical tree populations., pp. 281-297 in Estrada, A. and Fleming, T.H. (Eds.) Frugivory and seed dispersal: ecological and evolutionary aspects., Kluwer Academic Publisher, Dordrecht, The Nederlands

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

Heuertz, M., Hausman, J.-F., Tsvetkov, I., Frascaria-Lacoste, N., Vekemans, X. (2001): Assessment of genetic structure within and among Bulgarian populations of the common ash (*Fraxinus excelsior* L.). Molecular Ecology 10: 1615-1625

Le Corre, V., Roussel, G., Zanetto, A., Kremer, A. (1998): Geographic structure of gene diversity in Quercus Petrae (Matt.) Liebl. III. Patterns of variation identified by geostatistical analysis. Heredity 80: 464-473

Kaufman, S. R., Smouse, P. E., Alvarez-Buylla, E. R. (1998): Pollen-mediated gene flow and differential male reproductive success in a tropical pioneer tree, *Cecropia obtusifolia* Bertol. (Moraceae): a paternity analysis. Heredity 81: 164-173

Marshall, T. C., Slate, J., Kruuk, L. E. B., Pemberton, J. M. (1998): Statistical confidence for likelihood-based paternity inference in natural populations. Molecular Ecology 7: 639-655

Oddou-Muratorio, S., Petit, R. J., LeGuerroué, B., Guesnet, D., Demesure, B. (2001a): Pollen- versus seed-mediated gene flow in a scattered woody species. Evolution 55: 1123-1135

Oddou-Muratorio, S., Aligon, C., Decroocq, S., Plomion, C., Lamant, T., Demesure, B. (2001b): Microsatellite primers for *Sorbus torminalis* and related species. Molecular Ecology Notes 1: 297-299.

Oddou-Muratorio, S., Gerber, S., Demesure-Musch, B., Austerlitz, F. (in prep. 1). Real-time patterns of pollen flow in the wildservice tree, *Sorbus torminalis* (L.) Crantz. I. Evaluating the levels of apparent and cryptic pollen flow.

Oddou-Muratorio, S., Demesure-Musch, B., Pélissier, R., Gouyon, P.-H. (in prep. 2). The roles of colonization and history in the local genetic structure of a scattered temperate tree species

Raspe, O. (1999): Chloroplast DNA haplotype variation within and among populations of S. aucuparia L. (Rosaceae: Maloideae)

Rousset, F. (1997): Genetic differentiation and estimation of gene flow from F-Statistics under isolation by distance. Genetics 145: 1219-1228

Rousset, F. (2000): Genetic differentiation between individuals. J. Evol. Biol. 13:58-62

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

Streiff, R., Ducousso, A., Lexer, C., Steinkellner, H., Gloessl, J., Kremer, A. (1999): Pollen dispersal inferred from paternity analysis in a mixed oak stand of *Quercus robur* L. and *Q. petraea* (Matt.) Liebl. Molecular Ecology. 8:831-841.

Effect of pollinator composition on the breeding structure of tropical timber trees

C. W. DICK

Smithsonian Tropical Research Institute, Unit 0948, APO AA 34002, USA and Biological Dynamics of Forest Fragments Project, Instituto Nacional de Pesquisas da Amazônia, C.P. 478, Manaus, AM 69011-970, Brazil

Key words: *Dinizia excelsa*; *Apis mellifera*; microsatellites; fecundity; paternity inference; logging, habitat fragmentation, Amazonia

Abstract: Logging can alter the breeding structure of remnant tropical trees by increasing distances between potential mates and by disrupting pollination. This may result in inbreeding depression and low recruitment of desired species in altered habitats. Here I present the results of ecological and genetic studies of remnant populations of *Dinizia excelsa* (Fabaceae), a prominent Amazonian hardwood tree, which lost most of its native pollinators in highly disturbed habitats; nevertheless, fecundity was higher in pasture and forest fragments than in pristine forest. Genetic analyses of microsatellite genotypes showed that high rates of outcrossing and paternal allele diversity were maintained across habitat types. A paternity analysis further demonstrated pollen movement between trees separated by up to 3.2 km of pasture. The pollen vectors in disturbed habitat were feral African honeybees (*Apis mellifera scutellata*), which visited remnant trees in vast numbers. The effect of pollinator changes on reproduction in other tropical species is discussed, along with suggestions for further research.

Introduction

Foresters who work in lowland tropical rain forests may encounter a bewildering diversity of species and ecological interactions. Woody plants occur in assemblages of 300 species in a single hectare (Phillips et al. 1994; Romoleroux et al. 1997) and upwards of 1000 species in 50 hectares of pristine forest (Condit et al. 2000). Concomitant with high alpha diversity, tropical trees occur in low population densities, typically less than one individual per hectare (Hubbell and Foster 1983). Despite the potential difficulty of finding mates, self-fertilization is rare in undisturbed forests. In the tropical trees that have been studied, outcrossing is enforced through self-incompatibility (Bawa et al. 1985a), dioecy, and high genetic loads (Alvarez-Buylla et al. 1996). Animals, rather than wind, are presumed pollen vectors for over 99% of lowland rain forest species (Bawa et al. 1985b). Animals also disperse the fruits and seeds of most tropical trees.

Because market forces have created a demand for a mere handful of the thousands of potentially useful Neotropical timber species, most harvesting occurs in the form of selective logging. Selective logging has two immediate effects on the breeding structure of the remaining trees. First, it increases the distances between potential mates. Second, it can alter the species composition, abundances and foraging patterns of pollinators. Pollinator disruption can result in low fecundity due to pollen limitation (Levin 1995, Ghazoul et al. 1998), reduced outcrossing (Aldrich and Hamrick 1998; Murawski et al. 1994) and genetic erosion through drift (Young et al. 1996) if trees left in forest patches become reproductively isolated.

Despite the impact of selective logging on genetic processes, thus far only two studies have addressed this issue. Murawski et al. (1994) and Lee (2000) compared mating parameters of trees in logged and pristine Dipterocarp forests of Southeast Asia, and found increased levels of self-fertilization in the logged habitats. The majority of gene flow studies in tropical trees have focused on habitat fragmentation rather than logging and have demonstrated, contrary to early expectations (eg. Janzen 1986; Rathcke and Jules 1993), that gene flow by pollen can be high among isolated trees in tropical pastures (Chase et al. 1996; Aldrich and Hamrick 1998), and spatially isolated forest fragments (Nason and Hamrick 1997; White et al. 1999) indicating that spatially isolated trees do not readily experience reproductive isolation. Many of the trees that have been studied are timber species and are pollinated by small bees in undisturbed forest, although it is worth noting that none of these studies examined floral visitors in the disturbed habitats.

In this paper I shall discuss the effects of pollinator composition on the breeding structure of remnant timber trees. As a case study I present ecological and genetic results from a study of remnant *Dinizia excelsa* (Fabaceae), a giant Amazonian timber tree often left standing in pasture and forest fragments shortly after the felling of virgin forest for agriculture. *Dinizia excelsa* is a model species for studies of pollinator disruptions, since its generalist bee pollination system is shared by approximately 30% of lowland rainforest species (Bawa et al. 1985b). Although this study does not deal directly with the case of selective logging, the genetic effects of spatial isolation and pollination disruption can be extrapolated to logging scenarios. The effects of logging on trees with specialized bee pollination, or to trees not pollinated by bees (e.g. by bats, beetles, butterflies, or flies) await future studies.

Case study of *Dinizia excelsa* (Fabaceae)

Dinizia excelsa is one of the largest Amazonian trees, reaching 60 meters in height and two meters in girth (Ducke 1922), and it is one of the region's most

important timber species (Barbosa 1990). It is endemic to the Brazilian Amazon, where it occurs at natural densities of one adult tree per six hectares (Dick 2001a). *Dinizia excelsa* produces pale yellow flowers on racemes. The flowers are small (calyx 1-1.5 mm), hermaphroditic, and attract small insects with nectar and fragrance. The small seeds (~1 cm.) are borne in indehiscent pods and dispersed primarily by wind and gravity. Each pod contains 3 seeds on average (Dick 2001a). The study populations were mapped in the reserves of the Biological Dynamics of Forest Fragments Project (BDFFP; Lovejoy and Bierregaard 1990) located approximately 90 kilometers north of Manaus, Brazil (Figure 1).

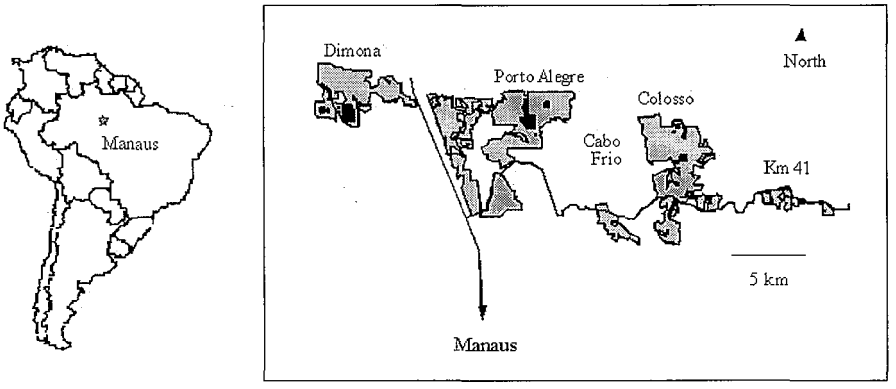


Figure 1. Reserves of the Biological Dynamics of Forest Fragments Project (2°30´S, 30°W). *Dinizia excelsa* (=40 cm DBH) were mapped in Dimona, Porto Alegre and Colosso ranches, and in undisturbed forest at Cabo Frio and Km 41. Populations were also mapped in Reserva Ducke (not shown) located~70 km to south. The ranches (stippled) are embedded in continuous forest. Black squares represent forest fragments.

The study trees (≥40 cm dbh) occurred in three habitat types: (1) 15-20 year-old pasture of Colosso ranch (N=10), (2) forest fragments in Colosso (N= 26), and (3) undisturbed forest located west and east of Colosso ranch (Cabo Frio [N= 30] and Km 41 [N=45]). Field studies were performed in 1995, 1996, and 1999. The trees were visited periodically to monitor phenology. A subset of the study population was then selected for canopy observations (N=7), for fecundity measures (N= 91 total), and for genetic analysis of progeny arrays (N= 11).

Canopy studies

Rope-climbing techniques (Laman 1994) were used to gain access to the canopy, at an average height of 25 meters. Flower-visiting insects were observed and collected on seven large trees to assess differences in pollination across habitat types. In addition, all flowering trees were observed from the ground with 10x binoculars during the course of phenological surveys. The climbed trees were located in the following habitats: (i) an isolated 10 ha reserve in Colosso ranch (two trees), (ii) isolated 100 ha reserves in Porto Alegre and Dimona (two trees), (iii) pasture in Colosso ranch (one tree), and (iv) undisturbed forest at site Km 41 (two trees). The intensively studied pasture tree of Colosso, "Col. 06", was separated from its nearest neighbour by over 600 meters of open field (Fig. 2). Continuous forest trees ("41.01" and "41.13") were 700 meters apart. In this study, all habitat fragments (10- and 100-ha) were considered as disturbed habitats.

In total, 52.5 hours were spent (on 17 visit days) in canopy observation during the morning (6 to 12 a.m.) when insect visitation was at its peak, and 10 additional hours were spent in observation (4 days) during the afternoon and evenings. Observations were evenly divided between undisturbed forest (27 hrs) and disturbed habitats (26.5 hrs.). The putative pollinators were captured with nets, mounted and identified by specialists (Dick 2001a).

Reproductive performance

Fecundity was used as the primary indicator of reproductive performance, and was measured as the total number of pods and seeds produced per tree. Pods were counted on large adult trees (=90 cm dbh), using binoculars, in Cabo Frio (N=20) and Km 41 (N=45) and Colosso ranch (N=26). The pasture and fragment populations of Colosso were analyzed separately and then combined for comparison with undisturbed populations. Seeds were counted from 2,177 fallen pods from trees in Colosso (N=25 trees), Km41 (N=45) and Cabo Frio (N=23) to obtain estimates of seed set per pod.

Microsatellite analyses

Leaf tissue was collected from all of the mapped adult trees in Colosso (N=36). DNA was extracted from adult leaves and from 11 seed families (25-55 seeds per family) of maternal trees located in the different habitat types. Genotypes were obtained for five microsatellite loci using primers developed for *D. excelsa* (Dick and Hamilton 1999). A total of 80 alleles in the seed and adult populations provided an average exclusion probability of greater than 99% (Table 1). The

outcrossing rate (t) for each maternal tree was calculated as the proportion of its seeds that contained non-maternal alleles. The minimum number of sires for each seed family was estimated by dividing the number of paternal alleles for the most variable locus by two, since the species is diploid. Seed paternity was assigned on the basis of multilocus segregation probabilities using the program Cervus 1.0 (Marshall et al. 1998).

Table 1. Microsatellite loci used in this study. The repeat motif is of the cloned sequence. The allele numbers shown are from 240 seed and tree genotypes.

Locus	Repeat sequence	No. of alleles
DE27	(AAG)8	4
DE37	(AC)20	12
DE44	(GT)13	7
DE48	(GA)27	34
DE54	(CT)39	23

Results

Pollinator Shifts

Only native insects were observed on the densely flowering trees of continuous forest. These included stingless bees in the tribe Meliponini (14 species in 12 genera), carnivorous wasps, and small beetles 2-5 mm in length (Dick 2001a). Stingless bees were judged to be the primary outcross pollinators of these trees, and they were the only insects observed foraging among branches and flying away from the trees. More than ten species of small beetles (eight families) were observed in the floral cups of *D. excelsa*. These beetles were not observed flying from flowering branches, but may effect self-pollination.

Similar species compositions and abundances of native insects were observed in continuous forest, 10 ha and 100 ha forest fragments (Dick 2001a). However, African honeybees were the predominant insects on flowering trees in pasture and forest fragments, where they greatly outnumbered all native insects. Honeybees were virtually the exclusive pollinating insects on the pasture tree Col. 06, which was devoid of beetles, and visited only rarely by the native bees *Frieseomelitta trichocerata* and *Trigona spp.* No behavioral interactions were observed between African honeybees and native insects in any of the trees. The co-existence of native and exotic bees in tree canopies in the forest isolates

suggests that African bees did not displace native insects from pasture trees. Rather, long distances of open pastures probably inhibit flights by native bees, which nest in standing forest (Roubik 1989).

Reproductive success

The trees in pasture and forest fragments (N=26) produced more than three times as many pods per tree as did trees from two natural populations (N=20 and N=45; Welsch test; $P<0.01$). No significant difference in fecundity was observed between trees in pasture (N=10) and forest fragments (N=16; $P>0.05$). The seed counts indicated that pods were an accurate proxy for seed set, with approximately 3 seeds per pod across habitat types. The mean seed set over all habitat types in Colosso was ~10,000 seeds/tree, with the most isolated pasture tree "Col. 06" also with ~10,000 seeds. Senescent pasture trees increased the variance in seed set at Colosso.

Genetic Results

Gene flow within the fragmented population was frequent and extensive. Even the most isolated trees incorporated a large fraction of the adult allele pool in their seed arrays (Dick 2001b). For example, twenty-five seeds from the gallery forest tree Col. 13 contained 11 out of 18 alleles found in the adult population for locus DE48. There was no significant difference in paternal allele diversity among progeny arrays, despite the spatial isolation of pasture trees. For example, the seed array of pasture tree Col.06 (N=35 seeds) had = 6 sires, and was separated from its two nearest neighbors by 600 and 1300 meters. In comparison, forest tree Km.41.13, with = 5 sires per 25 seeds, was surrounded by 18 potential mates in a 600 meter radius.

Outcrossing rates did not differ between forest fragments and continuous forest ($P>0.05$), but were reduced in pasture trees (t-test; $P<0.01$). The mean outcrossing rate in pasture was 0.85 (N=6) and 0.95 in forested habitat (N=7). Paternity inference of seed genotypes from all seven maternal trees in Colosso yielded paternity assignments with high exclusion probabilities for 77 out of 240 seeds, indicating that about two-thirds of the seeds were sired by unmapped trees in the surrounding forest (Figure 2). Twenty-two of the 77 inferred pollinations resulted from self-fertilization. Of the remaining 55 outcross pollinations, 26 occurred between trees separated by =1 km of pasture, with a range of 128 to 3200 meters (Figure 2). The mean pollination distance for the Colosso pasture and gallery forest trees was 1,288 meters (N= 45 seeds), which greatly exceeded the mean distance to nearest neighbour (DNN) of 235 meters ($P< 0.001$; N= 18 trees). The mean pollination distance in the 10-ha fragment

of 417 meters (N=10 seeds) also exceeded the mean DNN of 50.5 meters (N=12 trees). Prior phenological observations confirmed that nearest neighbours flowered in synchrony and thus were potential mates.

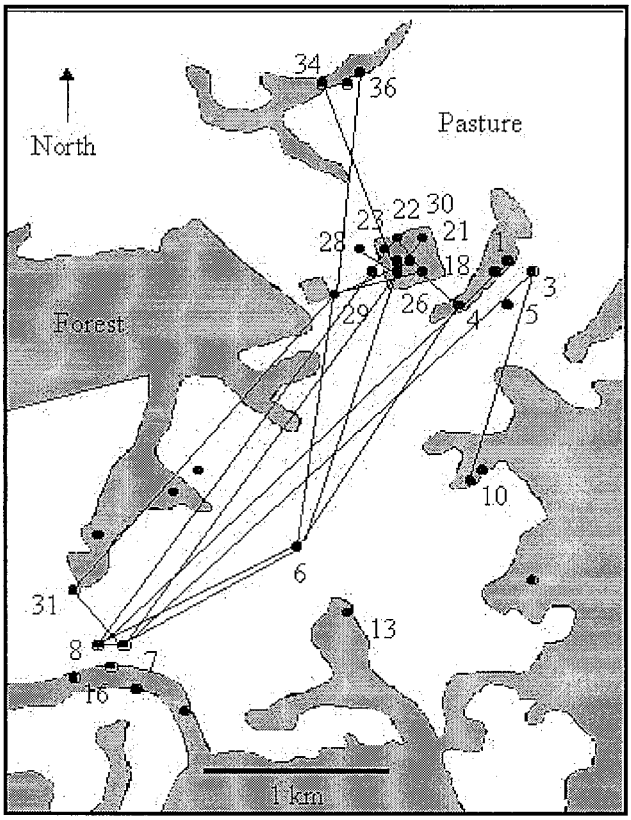


Figure 2. Map of D. excelsa at Colosso ranch highlighting some of the long distance pollinations. Pasture and secondary vegetation in white; shaded areas are primary forest. The numbers are the tree names referred to in the text.

Higher rates of self-fertilization have been reported for other canopy species remnant in pasture (Aldrich and Hamrick 1998) and selectively logged forest (Murawski et al. 1994). The negative demographic effects of inbreeding depression in isolated *D. excelsa* should be obviated, however, by the higher absolute number of outcrossed seeds produced. Moreover, regeneration of *D. excelsa* was found to be extensive in the abandoned pastures (Dick 2001b).

Discussion

Feral honeybees in the Neotropics

Although the honeybee *Apis mellifera* has a long and global history in association with humans, its appearance in the neotropics is relatively recent. Humid tropical climates are not ideal for European honeybees. In order to breed a tropical strain with improved honey production and climatic resilience, researchers imported bees from the African race *Apis mellifera scutellata* to southern Brazil in the early 1950s. It differs from the European honeybees *A. m. mellifera* in its aggressive behaviour, larger colony size, greater number of colonies per land area, and tendency to abscond (Michener 1975). Unlike European honeybees, which are often found dead in their hives during food shortages, colonies of African honeybees can travel over 200 km to resource-rich areas (Kigatiira 1988, cited in Roubik 1989).

In 1956, 26 queens and about 200 males escaped from apiaries in southern Brazil and spread rapidly. They initially covered ~320 km/yr, mating with all of the European honeybee queens in their path. The hybrids retained morphological and behavioural traits of the African race (Diniz and Malaspina 1996), and therefore are still called African honeybees. There were no *Apis mellifera* in the Amazon basin prior to the African honeybee invasion (Roubik 1989). African honeybees were reported in Belém in 1971 and in Manaus in 1974 (Prance 1976); they were observed in French Guiana by 1976 (Roubik 1978).

There are an estimated 50 to 100 million AHB colonies in Latin America today (Winston 1992). Estimates of colony density range from 10/km^2 (Taylor 1985) to 108/km^2 (Kerr 1971, in Michener 1975). African honeybees seem to thrive in mosaic habitats because of their eclectic use of nesting sites, and their use of primary forest, crop, and weedy plant species for nectar and pollen (Roubik 1989). They do, however, also occur in primary forest. A census of trees in secondary growth and dry, seasonally wet forest in Yucatan, Mexico, yielded 15-36 colonies per square kilometer (Quezada-Euan and May-Itza 1996).

Introduced honeybees visit 1/3 of local native plants but their effect on the genetic structure of native plants is unknown (Butz-Huryn 1997). They are thought to follow an optimal foraging strategy, which indicates that they should remain on a huge flowering canopy; and in rare instances of flights between densely flowering trees they should visit only nearest neighbors (Butz-Huryn 1997).

Effects of African bees on D. excelsa.

African honeybees appear to have promoted the high fecundity of *D. excelsa* in pasture and forest fragments. Changes in light and nutrient availability can enhance plant fecundity, but this is not a sufficient explanation for *D. excelsa*. The leaves of this canopy emergent are fully exposed to light regardless of habitat, and no differences in flower density were observed across habitat types. The soil environment differs most dramatically between pasture and forest fragments, but there was no significant difference in the fecundity of trees in these two habitats. African honeybees have been shown to increase seed set over natural levels in other tropical plants (Aizen and Feinsinger 1994; Roubik in press). Enhanced fecundity may contribute to the demographic growth of the remnant *D. excelsa* in the long term, because of its capacity for regeneration in abandoned pasture and forest edges.

These genetic data challenge conventional views about the foraging behaviour of *Apis mellifera* (Butz-Huryn 1997). First, the data strongly suggest that inter-tree foraging is not rare. Pasture tree Col.06 was visited almost exclusively by honeybees and it set roughly 8,000 outcrossed seeds, with a nearest neighbour located 600 meters away. Second, the data challenge the view that inter-tree foraging is limited to neighbouring trees. Col.13 and Col.10, for example, did not sire any seeds sampled from Col.06 (N=55), despite their proximity and overlap in flowering (Figure 2). Interestingly, Col.06 received pollen from the 10 ha fragment (Col.26), which contained a dense stand of flowering *D. excelsa*. This violation of the optimal foraging was inferred from genetic analyses of pasture trees in Costa Rica (Chase et al. 1996), although the pollinators in this study were not identified.

Ironically, African honeybees may increase the neighbourhood area of fragmented *D. excelsa* above natural levels. *Apis mellifera* have a vast foraging area (212 km^2) compared to native stingless bees (12.5 km^2) (Roubik 1989).

The large number of seeds sired by trees outside of the mapped population indicates that the African honeybees were not confined to the disturbed habitats. The fact that cross-fertilized seeds of pasture trees provide a broad sampling of the forest genotypes has practical conservation implications, since a guideline for tropical reforestation is to avoid collecting seeds from isolated trees because of their presumed low genetic quality.

Effects on other plant species

Although honeybees (including African honeybees) visit up to one-third of the plant species in local floras, they intensively visit and pollinate a much smaller proportion of the flora. Menezes and Camargo (1991) reported African

honeybees on 47 out of 187 plant species in Brazilian wooded savanna (*cerrado*), although only nine species received 65 percent of the visits. To the plants they visit, honeybees can be major pollinators, or secondary pollinators that pollinate less effectively than native insects, or floral parasites that pollinate only incidentally (Butz-Huryn 1997). Honeybees tend to be major pollinators of plants with simple flowers and rich nectar or pollen resources. They have been shown to take over pollination in the threatened Madagascar palm *Neodypsis decaryi* (Ratsirarson and Silander 1996), the cashew *Anacardium occidentale* (Freitas and Paxton 1996) and the avocado (*Persea americana*) (Isham and Eisikowitch 1993).

In a study of leguminous shrubs in dry forest fragments of Argentina (*Propopsis nigra* and *Cercidium australe*) Aizen and Feinsinger (1994) found that despite high visitation rates by African honeybees, the fragmented populations suffered slightly reduced seed set compared to continuous forest populations. In some cases, however, honeybees make up for lower efficiency by visiting flowers more frequently. In flowers of the dry forest annual *Kallstroemia grandiflora*, for example, African honeybees transferred 2.5 times fewer pollen grains per visit than did the native bee *Trigona nigra* but compensated by making 2.65 times as many visits (Osorioberistain et al. 1997).

African honeybees may reduce the viability of plants with specialized flowers, especially if significant quantities of floral resources are extracted without pollination. One such class of plants has poricidal anthers that require sonic vibrations to release their pollen (buzz pollination). Bees curl around the anther and grasp the stamens tightly with their wings. By rapidly flexing their indirect flight muscles, they transmit sonic vibrations (buzzing) that expel pollen from the anthers (Buchmann 1983). Although honeybees consume nectar from this class of plants, they are not capable of pollinating them. Poricidal *Mimosa pudica* (Mimosaceae) in French Guiana suffered a 26 percent decline in seed set when about 74 percent of the visitors were honeybees, compared to forest populations visited by high frequencies of native bees (Roubik 1996). Buzz-pollinated species are found in such prominent tropical families as the Fabaceae, Lauraceae, Elaeocarpaceae, Solanaceae, and Melastomataceae (Buchmann 1983).

Directions for future research

There are now numerous studies on the effects of habitat fragmentation on gene flow in tropical canopy trees, and these have generally concluded that gene flow among remnant trees is high. If this line of research continues, I would suggest that researchers focus on species that are most susceptible to

habitat disturbance. For example, the canopy observations on *D. excelsa* showed that isolated pasture trees lost their small beetle pollinators. This was of little consequence to *D. excelsa*, since its major pollinators are bees. However, these small beetles are primary pollinators for many economically important trees, such as *Virola surinamensis* (Myristicaceae) and other trees in the nutmeg family (Armstrong and Irving 1989). Virtually nothing is known about moth and butterfly pollination (Bawa 1990), yet these are probably important pollinators for timber trees such as *Cordia alliadora* (Boraginiaceae), *Swietenia macrophylla* and *Cedrela odorata* (Meliaceae). Even more obscure is the pollination of trees such as cacao, *Theobroma cacao* (Sterculiaceae), visited primarily by flies.

Geneticists are rightly concerned that their research be incorporated into plans for sustainable tropical forestry. But as the case study of *D. excelsa* clearly shows, ecological research will play a leading role in elucidating the causes of demographic and genetic change.

Acknowledgements

Lyn Loveless and Dave Roubik made valuable comments on the manuscript. This research was supported by a NSF Pre-Doctoral Fellowship, a Deland Award from the Arnold Arboretum, Sigma Xi, the Biological Dynamics of Forest Fragments Project, the Harvard Committee on Latin American Studies and the Department of Organismic and Evolutionary Biology at Harvard. This is publication number 381 in the BDFP technical series.

References

Aizen, M. A., Feinsinger, P. (1994): Forest fragmentation, pollination, and plant reproduction in a Chaco dry forest, Argentina. Ecology 75 (2): 330-351.

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

Alvarez-Buylla, E. R., García-Barrios, R., Lara-Moreno, C., Martínez-Ramos, M. (1996): Demographic and genetic models in conservation biology: applications and perspectives for tropical rain forest tree species. Ann. Rev. Ecol. Syst. 27: 387-421.

Armstrong, J. E., Irvine, A. K. (1989): Floral biology of *Myristica insipida* (Myristicaceae), a distinctive beetle pollination system. Am. J. Bot. 76 (1): 86-94.

Barbosa, R. I. (1990): Análise do setor madeireiro do estado de Roraima. Acta Amazonica 20: 193-209.

Bawa, K. S. (1990): Plant-pollinator interactions in tropical rain forests. Ann. Rev. Ecol. Syst. 21: 399-422.

Bawa, K. S., Bullock, S. H., Perry, D. R., Colville R. E., Grayum, M. H. (1985b): Reproductive biology of tropical lowland rainforest trees. II. Pollination mechanisms. Amer. J. Bot. 72 (3): 346-56.

Bawa, K. S., Perry, D. R., Beach, J. H. (1985a): Reproductive biology of tropical lowland rainforest trees. I. Sexual systems and incompatibility mechanisms. Am. J. Bot. 72: 331-345.

Buchmann, S. L. (1983): Buzz pollination in angiosperms. Pp. 73-113 in: Jones, C. E. and Little, R. J. (Eds.) Handbook of experimental pollination biology. Van Nostrand Rinhold, New York.

Butz-Huryn, V. M. (1997): Ecological impacts of introduced honeybees. Quart. Rev. Biol. 72: 275-297

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

Condit, R., Ashton, P. S., Baker, P., Bunyavejchewin, S., Gunatilleke, S., Gunatilleke, N., Hubbell, S. P., Foster, R. B., Itoh, A., LaFrankie, J. V., Lee, H. S., Losos, E., Manokaran, N., Sukumar, R., Yamakura, T. (2000): Spatial patterns in the distribution of tropical tree species. Science 288:1414-1418.

Dick, C. (2001a): Habitat change, African honeybees and fecundity in the Amazonian tree *Dinizia excelsa* (Fabaceae). Pp. 146-157 in: Bierregaard, R. O., Gascon, C., Lovejoy , T. E. and Mesquita, R. (Eds.) Lessons from Amazonia: The Ecology and Conservation of a Fragmented Tropical Forest. Yale University Press, New Haven.

Dick, C. (2001b): Genetic rescue of remnant tropical trees by an alien pollinator. Proceedings of the Royal Society of London B 268: 2391-2397.

Dick, C., Hamilton, M.B. (1999): Microsatellites from the Amazonian tree *Dinizia excelsa* (Fabaceae). Mol. Ecol. 8: 1765-66.

Diniz, J.A.F., Malaspina, O. (1996): Geographic variation of Africanized honey bees (*Apis mellifera*) in Brazil-multivariate morphometrics and racial admixture. Brazilian Journal of Genetics 19 (2): 217-224.

Ducke, A. (1922): Plantes nouvelles ou peu connues de la région amazonienne. Archivos do Jardim Botanico do Rio de Janeiro 3: 2-269.

Freitas, B.M., Paxton, R. J. (1996): The role of wind and insects in cashew (*Anacardium occidentale*) pollination in NE Brazil. Journal of Agricultural Science 126 (3): 319-326.

Ghazoul, J., Liston, K. A., Boyle, T. J. B. (1998): Disturbance-induced density-dependent seed set in *Shorea siamensis* (Dipterocarpaceae), a tropical forest tree. Journal of Ecology 86: 462-473.

Hubbell, S. P. and Foster, R. B. (1983): Diversity of canopy trees in a neotropical forest and implications for conservation. pp. 25-41 in: Sutton, S. L., Whitmore, T. C. and Chadwick, A. C. (Eds.) Tropical Rain Forest: Ecology and Management. Blackwell Scientific Publications , Oxford.

Isham, G., Eisikowitch, D. (1993): The behaviour of honey bees (*Apis mellifera*) visiting avocado (*Persea americana*) flowers and their contribution to its pollination. Journal of Apicultural Research 32 (3-4): 175-186.

Janzen, D. H. (1986): The future of tropical ecology. Ann. Rev. Ecol. Syst. 17: 305-324.

Laman, T. (1994): Safety recommendations for climbing rain forest trees with "single rope technique". Biotropica 27 (3): 406-409.

Lee, S.L. (2000): Mating system parameters of *Dryobalanops aromatica* Gaertn. f. (Dipterocarpaceae) in three different forest types and a seed orchard. Heredity 85 (4): 338-345

Levin, D. (1995): Plant outliers: an ecogenetic perspective. Am. Nat. 145 (1): 109-118

Lovejoy, T., Bierregaard, R. (1990): Central Amazonian forests and the minimum critical size of ecosystems project. pp. 60-75 in: Gentry, A. (Ed.) Four Neotropical Forests. Yale University Press, New Haven.

Marshall, T. C., Slate, J., Kruuk, L. E. B., Pemberton, J. M. (1998): Statistical confidence for likelihood-based paternity inference in natural populations. Mol. Ecol. 7: 639-655.

Menezes, P. S., Camargo, J.M.F. (1991): Interactions on floral resources between the Africanized honeybee *Apis mellifera* L. and the native bee community (Hymenoptera: Apoidea) in a natural "cerrado" ecosystem in southeast Brazil. Apidologie 22: 397-415.

Michener, C.D. (1975): The Brazilian bee problem. Ann. Rev. Entomol. 20: 399-416.

Murawski, D.A., Gunatilleke, I.A.U, Bawa, K. (1994): The effects of selective logging on inbreeding in *Shorea megistophylla* (Dipterocarpaceae) from Sri Lanka. Cons. Bio. 8 (4): 997-1002.

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

Osorioberistain, M., Dominguez, C.A., Eguiarte, L.E., Benrey, B. (1997): Pollination efficiency of native and invading Africanized bees in the tropical dry forest annual plant, *Kallstroemia grandiflora* Torr ex Gray. Apidologie 28 (1): 11-16.

Phillips, O. L., Hall, P., Gentry, A. H., Sawyer, S. A., Vasquez, R. (1994): Dynamics and species richness of tropical rain forests. Proc. Natl. Acad. Sci. USA 91: 2805-2809.

Prance, G. (1976): The pollination and androphore structure of some Amazonian Lecythidaceae. Biotropica 8 (4): 235-241.

Quezada-Euan, J. J., May-Itza, Y. W. De. J. (1996): Características morfométricas, poblacionales y parasitosis de colonias silvestres de *Apis mellifera* (Hymenoptera: Apidae) en Yucatán, Mexico. Folia Entomológica Mexicana 97: 1-19.

Rathcke, B. J., Jules, E. S. (1993): Habitat fragmentation and plant-pollinator interactions. Current Science 65:273-277.

Ratsirarson, J., Silander, J. A. (1996): Reproductive biology of the threatened Madagascar Triangle-Palm *Neodypsis decaryi* Jumelle. Biotropica 28 (4): 737-745.

Romoleroux, K., Foster, R., Valencia, R. Condit, R. Balslev, H., Losos, E. (1997): Especies leñosas (dap=1 cm) encontradas en dos hectáreas de un bosque de la Amazonía ecuatoriana. pp. 189-215 in: Valencia, R. and Balslev, H. (Eds.) Estudios sobre Diversidad y Ecología de Plantas. Pontificia Universidad Católica del Ecuador, Quito.

Roubik, D. W, (1989): Ecology and natural history of tropical bees. Cambridge University Press, Cambridge.

Roubik, D. W. (1996): African bees as exotic pollinators in French Guiana. pp 163-172 in: Matheson, A. Buchmann, S. L., O'Toole, C., Westrich, P., Williams, I. (Eds.) The conservation of bees. Academic Press, London.

Roubik, D. W. (in press): Feral African bees augment neotropical coffee yield. In Kevan, P. G. and Imperatriz-Fonseca, V. (Eds) International Workshop on the Conservation and Sustainable Use of Pollinators. International Bee Research Association, Cardiff.

Roubik, D.W. (1978): Competitive interactions between neotropical pollinators and Africanized honey bees. Science 201: 1030-1032.

Taylor, O.R. (1985): African bees: potential impact in the United States. Bull. Entmol. Soc. Am. 31:15-24.

White, G.M., Boshier, D.H., Powell, W. (1999): Genetic variation within a fragmented population of *Swietenia humilis* Zucc. Mol. Ecol. 8: 1899-1909.

Winston, M. L. (1992): The biology and management of Africanized honey bees. Ann. Rev. Entomol. 37: 173-93.

Young, A., Boyle, T., Brown, T. (1996): The population genetic consequences of habitat fragmentation for plants. Trends. Ecol. Evol. 11 (10): 413-418.

Detecting reliable parent-offspring matches in parentage analysis: a case study

S.C. GONZÁLEZ-MARTÍNEZ[1,2], S. GERBER[3], M.T. CERVERA[1,4], J.M. MARTÍNEZ-ZAPATER[1,4], L. GIL[2] and R. ALÍA[1]

[1]Departamento de Mejora Genética y Biotecnología, INIA, P.O. 8111, 28.080 Madrid, Spain; [2]Unidad de Anatomía, Fisiología y Genética, ETSIM, Ciudad Universitaria s/n, 28040 Madrid, Spain; [3]Laboratoire de génétique et amélioration des arbres forestiers, INRA, BP 45, 33611 Gazinet Cedex, France; [4]Departamento de Genética Molecular de Plantas, Centro Nacional de Biotecnología, CSIC, Campus Universidad Autónoma de Madrid en Cantoblanco, 28049 Madrid, Spain.

Key words: LOD-scores, probability levels, simulations, *Pinus pinaster*, SSRs.

Abstract: A simulation approach was recently developed to identify single parents and parent pairs of a given offspring at a certain probability level using molecular markers. In a previous paper, this method was applied to the study of seed gene flow in maritime pine (*Pinus pinaster* Ait.) using three highly polymorphic nuclear microsatellites. In the present communication, the accuracy of the method was tested using simulated populations. 'Assumed' probability levels were compared with probability levels computed using simulated populations ('true' probability levels). The 'assumed' probability levels were close to 'true' probability levels when single parents-offspring assignments were considered. In contrast, parentage analysis based on parent pairs showed 'true' probability levels lower than 0.01, whatever the 'assumed' probability levels (0.05, 0.10, 0.15).

Introduction

Parentage analysis, that is the identification of parents for offspring within a population, provides fundamental data for evolutionary and ecological research. Methods for parentage inference in plant species include single exclusion (Yazdani et al. 1989; Dow and Ashley 1996), categorical allocation of offspring to the most-likely parent or parent pair (Meagher and Thompson 1987; Gerber et al. 2000) and fractional allocation of offspring among all nonexcluded parents (Devlin et al. 1988). The identification of the most-likely parent or parent pair using likelihood methods has become common in parentage studies. This approach has received three major criticisms (Schnabel 1998; and references within): (1) parentage assignment is biased towards homozygotes, since parents with the highest number of homozygous loci have the highest transition probabilities, (2) the parent or parent pair with the highest likelihood score is typically chosen to

be the true parent without regard to the likelihood of other possible parents, and (3) because ties cannot be resolved, the analysis favoured the identification of parents with rare alleles.

To avoid statistical pitfalls related to most-likely methods, simulation approaches to evaluate confidence levels in parentage assignments and the restriction of analysis to only reliable parent-offspring matches have been proposed (Marshall et al. 1998; Gerber et al. 2000). One of these approaches was recently developed and used to estimate seed gene flow in maritime pine (González-Martínez et al. 2002). The aim of this paper is to evaluate: (1) the accuracy of the simulation-based probability levels computed in González-Martínez et al. (2002) and (2) the applicability of the method to other sets of data.

Material and Methods

Maritime pine: a case study

Maritime pine (*Pinus pinaster* Ait.) is a wind-pollinated outcrossing conifer with a wide distribution in the Mediterranean basin, where it spreads all over the western range. The development of markers in maritime pine has attracted considerable attention due to both the ecological and economic importance of the species. However, the development of nuclear microsatellites (nSSR) was quite unsuccessful and only three primer pairs are available to date (*Itph4516, Frpp91* and *Frpp94*; Mariette et al. 2001). Laboratory protocols using a LI-COR 4000 automatic sequencer (LI-COR Inc., Nebraska, USA) are described in Mariette et al. (2001) and a protocol using acrylamide denaturing gels for two of them (*Itph4516* and *Frpp91*) can be found in González-Martínez et al. (2002).

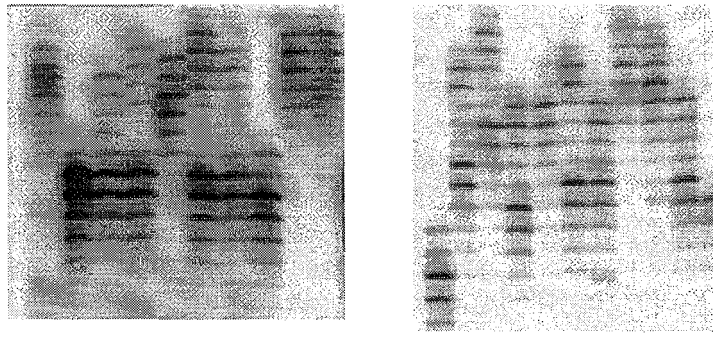


Figure 1. Nuclear microsatellites in Maritime pine (*Frpp91*, left; *Itph4516*, right) as revealed by acrylamide/bisacrylamide denaturing gels and $^{33}\gamma$P-ATP labelling.

The Intensive Sampling Plot (ISP) of Coca (Central Spain) is situated in a typical Mediterranean area, with annual average rainfall and temperature of 432 mm and 12.3°C, respectively, and a pronounced summer drought. A data set was analysed that included 76 mature trees within a circular plot 50 m in radius and all seedlings and saplings taller than 20 cm within a central subplot of 25 m in radius.

Data analysis

The most-likely parents and parent pairs were detected using log-likelihood ratios or LOD- scores (Meagher and Thompson 1986; Gerber et al. 2000) and a scheme to calculate probability levels for each match was done as follows: two sets of 10,000 offspring were generated: one with both parents inside the stand and the other with both parents from outside the stand (generated according allele frequencies in the whole population). Then, the distribution of the likelihoods of the cases in which the trees that had the highest likelihood were not the true parents, independently of being from inside or outside the plot, was used to compute the probability of the match. That is, the probability, p, of wrongly inferring a tree as parent of a given offspring when the potential parent with the highest likelihood is 'assumed' to be the true parent. Only mature trees with a good performance as single parents ($p<0.35$) were included in the calculation of probability levels for parent pairs. A more detailed description of the simulation scheme can be found in González-Martínez et al. (2002).

To measure the accuracy of the simulation-based probability levels computed in the preceding step, 10,000 offspring were simulated by picking their parents randomly among a population of 2,500 trees where the 76 first were the genotyped trees in our stand and the rest were generated by picking both alleles at each locus at random according to their frequencies in the whole population (Gerber et al. 2000). For each offspring, confidence levels were calculated following the preceding step ('assumed' probability levels). Offspring with a probability lower than a given value (0.05, 0.10, 0.15) of wrongly inferring a tree as its parent - when it is not - were retained and the number of wrong parent-offspring assignments in each set was counted. Then, 'assumed' probability levels were compared to the percentage of wrong parent-offspring assignments. A similar test was done to check the quality of parent pair-offspring assignments but only including mature trees with a good performance as single parents ($p<0.35$).

Results

The comparison between 'assumed' probability levels computed following González-Martínez et al. (2002) and probability levels obtained from simulated

populations is shown in Table 1. In single parent-offspring assignments, the 'assumed' probability levels were close to 'true' probability levels whereas percentages of correct parent-offspring assignments were higher than 99%, whatever the 'assumed' probability levels, in parent pair-offspring assignments.

Table 1. Percentage of wrong single parent/parent pair-offspring assignments in 10,000 offspring generated from a simulated population of 2,500 mature trees where offspring could have zero, one or two parents inside the plot. Only mature trees with a good performance as single parents were included in the parent pair analysis; *p* stands for the 'assumed' probability level computed following González-Martínez et al. (2002).

	'Assumed' probability level		
	$p<0.15$	$p<0.10$	$p<0.05$
Single parent	12.57%	10.09%	6.72%
Parent pair	0.32%	0.25%	0.35%

Discussion

Parentage analysis can be unsuccessful due to either limited resolution of the molecular markers (see, for instance, Taylor et al. 1997) or a high number of potential parents in the population (Chakraborty et al. 1988). In particular, parentage analysis in *Pinus* is difficult for two reasons. First, the development of codominant highly-polymorphic markers (e.g. nuclear microsatellites) is expensive and complicated due to the size and complexity of conifers' genome (Kinlaw and Neale 1997; Wakamiya et al. 1993). Second, pines are usually well adapted to long-distance gene flow so that isolated populations with only few potential parents are rare. In the case of maritime pine, with only three nuclear microsatellites available to date, the calculation of probability levels using simulation procedures seems adequate to detect reliable parent-offspring assignments in parentage analysis. However, it is striking to note that realised probabilities in parent pair-offspring analysis were higher than 'assumed' probability levels. In addition, a trend to increase the correct identification of parent pairs for offspring when the number of simulated potential parents increases has been found (although the data are not shown here). One possible explanation is obtained from the way the distribution to compute probability levels was constructed by González-Martínez et al. (2002). Because they pooled wrong assignments independently of the origin of the parents (parents from within the plot or simulated parents from outside the plot) and parent-pair analysis showed a higher power to detect wrong assignments that include parents from

outside the plot (i.e. putative parents with randomly generated genotypes), the reported trend could be expected.

In conclusion, simulation-based approaches to evaluate confidence levels in parentage assignments seems adequate and could be used extensively to study family structure in natural populations of conifers.

Acknowledgements

This paper represents a portion of the doctoral research of the first author who is supported by a FPU scholarship from MEC (Ministerio de Educación, Cultura y Deporte). The study was funded by the Cooperation project DGCN–INIA CC00-0035 and the INIA project SC97-118. Thanks to D. Agúndez, S. Mariette, G. Le Provost, J.A. Cabezas, A. Álvarez, A. Piñera and F. del Caño for field and technical assistance and to P.C. Grant who revised the language.

References

Chakraborty, R., Meagher, T.R., Smouse, P.E. (1988): Parentage analysis with genetic markers in natural populations. I. The expected proportion of offspring with unambiguous paternity. Genetics 118: 527-536

Devlin, B., Roeder, K., Ellstrand, N.C. (1988): Fractional paternity assignment: theoretical development and comparison to other methods. Theor. Appl. Genet. 76: 369-380.

Dow, B., Ashley, M. (1996): Microsatellite analysis of seed dispersal and parentage of saplings in bur oak, *Quercus macrocarpa*. Mol. Ecol. 5: 615-627

Gerber, S., Mariette, S., Streiff, R., Bodénès, C., Kremer, A. (2000): Comparison of microsatellites and AFLP markers for parentage analysis. Mol. Ecol. 9: 1037-1048

González-Martínez, S.C., Gerber, S., Cervera, M.T., Martínez-Zapater, J.M., Gil, L., Alía, R. (2002): Seed gene flow and fine-scale structure in a Mediterranean pine (*Pinus pinaster* Ait.) using nuclear microsatellite markers. Theor. Appl. Genet. 104: 1290-1297

Kinlaw, C.S., Neale, D.B. (1997): Complex gene families in pine genomes. Trends in Plant Science 2: 356-359

Mariette, S., Chagne, D., Decroocq, S., Vendramin, G.G., Lalanne, C., Madur, D., Plomion, C. (2001): Microsatellite markers for *Pinus pinaster* Ait. Annals of Forest Science 58: 203-206

Marshall, T.C., Slate, J., Kruuk, L.E.B., Pemberton, J.M. (1998): Statistical confidence for likelihood-based paternity inference in natural populations. Mol. Ecol. 7: 639-655

Meagher, T.R., Thompson, E.A. (1986): The relationship between single parent and parent pair genetic likelihoods in genealogy reconstruction. Theoretical Population Biology 29: 87-106

Meagher, T.R., Thompson, E.A. (1987): Analysis of parentage for naturally established seedlings of *Chamaelirium luteum* (Liliaceae). Ecology 68: 803-812

Schnabel, A. (1998): Parentage analysis in plants: mating systems, gene flow, and relative fertilities. Pp. 173-189 in: Carvalho, G.R. (Ed.) Advances in Molecular Ecology. IOS Press, The Netherlands

Taylor, A.C., Horsup, A., Johnson, C.N., Sunnucks, P., Sherwin, B. (1997): Relatedness structure detected by microsatellite analysis and attempted pedigree reconstruction in an endangered marsupial, the northern hairy-nosed wombat *Lasiorhinus krefftii*. Mol. Ecol. 6: 9-19

Wakamiya, I., Newton, R.J., Johnston, J.S., Price, H.J. (1993): Genome size and environmental factors in the genus *Pinus*. Am. J. Bot. 80: 1235-1241

Yazdani, R., Lindgren, D., Stewart, S. (1989): Gene dispersion within a population of *Pinus sylvestris*. Scandinavian Journal of Forest Research 4: 295-306

Genetic differentiation among Argentinean populations of southern beech *Nothofagus nervosa*

P. MARCHELLI and L. A. GALLO

Unidad de Genética Forestal, INTA EEA Bariloche, CC 277, 8400 Bariloche, Río Negro, Argentina

Key words: population genetics, allozymes, cpDNA, glaciation refugia, Nothofagaceae.

Abstract: *Nothofagus nervosa* is one of the main components of southern South American temperate forests. Genetic variation was analysed in 20 Argentinean populations of the species using allozymes and cpDNA markers. High levels of variation were observed considering the very restricted range of the species. The high levels of diversity and occurrence of two geographically differentiated chloroplast haplotypes suggest the existence of at least two refugia during the last glaciation. The distribution pattern of genetic variation is discussed in relation to hybridisation between *N. nervosa* and the closely related *N. obliqua* and as a consequence of recolonisation processes since the last glacial maximum.

Introduction

The southern temperate forests of South America lie on both sides of the Andes Mountains between 35° and 55° S. *Nothofagus nervosa* is one of the most important constituent species of these forests due to its high wood quality. In Argentina its natural area of distribution is very restricted, covering only 55,000 ha. (39° 20´- 41° 35´), compared to its extensive distribution in Chile. Because its range has been drastically reduced in the past due to overexploitation, overgrazing and recurrent forest fires, the species has been recently included in national conservation and breeding programs.

There are several factors that might be affecting the distribution of genetic variation in this species. In particular, its geographic distribution, which follows the west-east lake watersheds of glacial origin, together with strong winds of the same orientation, especially during pollination time, suggest the possibility of unidirectional gene flow occurring within watersheds. Although seeds are also wind-dispersed, seeds rarely travel more than 50 m from the seed tree (Donoso, 1993), so gene flow mainly refers to pollen movement. In addition, hybridisation between *N. nervosa* and *N. obliqua* has been demonstrated, and a process of introgression of *N. obliqua* genes into the gene pool background of *N. nervosa* has been suggested (Gallo et al. 1997a; 1997b; Gallo et al. 2000; Gallo this

issue). The occurrence of both species in some but not all watersheds could therefore be affecting the distribution of genetic variation. Finally, several studies strongly suggest that during the last glaciation in South America, many ice-free areas remained, allowing the occurrence of multiple refugia where species could survive (e.g. Markgraf et al. 1995; Bianchi 1997). If this is so, distribution of the current genetic variation should reflect recolonisation routes from multiple refugia.

In the present study, genetic variation among and within populations of *N. nervosa* sampled from its natural area of distribution in Argentina was analysed using allozymes and chloroplast DNA markers. Factors of geographic distribution, hybridisation in zones of sympatry, and potential Pleistocene refugia were considered to an attempt to identify the main influences on the distribution pattern of genetic variation within this species.

Material and Methods

Populations: Bulk seed collections were made from 20 Argentinean populations of *Nothofagus nervosa* by placing nets below the canopy at ca. 1.5 m off the ground, capturing seeds from at least 40 trees in each population. A minimum of 100 seeds per population were used for isozyme analyses, while leaves from two seedlings of each population were employed for chloroplast DNA markers. Seeds were kept at 4° until needed.

Isozymes: Eight previously characterised isozyme gene markers (Marchelli and Gallo 2000) were analysed for at least 100 seeds per population. From this previous work it is known that two of the surveyed loci carry alleles which are species-specific to *N. obliqua* and *N. nervosa* (*Adh* and *Pgi*). Enzyme systems, electrophoretic conditions and procedures are those described in Marchelli and Gallo (2000). Mean number of alleles per locus (A_L), genetic diversity (v, Gregorius 1978), actual and expected heterozygosity (H_a and H_e, Nei 1973), Gregorius genetic distance, amount of genetic differentiation (Dj) and mean level of genetic differentiation (δ; Gregorius and Roberds 1986) were estimated. GSED (Gillet 1994) and POPGENE (Yeh and Boyle 1997) were used for calculations.

cpDNA: DNA extraction and PCR-RFLP analysis were done according to Marchelli et al. (1998). In that previous study only eleven populations were surveyed. An additional nine populations (1, 4, 8, 11, 12, 14, 18, 23, 26) were analysed in this study, for two of the polymorphic fragments: primer DT (Demesure et al. 1995) digested with restriction enzymes *HaeIII* and *TaqI*.

Results

The observed values of gene diversity denote minor polymorphisms in most enzyme systems and the heterozygosity values were highly variable both among loci and among populations (Table 1).

Locus	Sample size	AL	υ	Ha	He
Mdh-B	4526	6	1.49 - 2.24	0.45	0.51
Mdh-C	4690	3	1.00 - 1.15	0.02	0.02
Idh	4388	2	1.05 - 1.69	0.17	0.18
Adh	4520	2	1.00 - 1.34	0.02	0.03
Got-A	4432	4	1.54 - 2.00	0.41	0.47
Got-B	4368	3	1.01 - 1.81	0.14	0.15
Got-C	4394	4	1.01 - 1.86	0.14	0.16
Pgi	4826	2	1.00 - 1.08	0.01	0.01
Mean	4518	3.25	1.13 - 1.38	0.17	0.19
St. Dev		1.39		0.17	0.20

Table 1. Number of alleles (A_L), genetic diversity (V), actual and expected heterozygosity (Ha and He) for the eight enzyme loci analysed in samples from 20 populations.

The greatest genetic distance was found between populations 17 and 30 (*do* = 0.103). Cluster analysis did not revealed any relationship among populations from the same watershed (Fig. 1). The genetic differentiation of each population compared to the complement formed by the remaining populations (Dj), was variable among populations and also within watersheds. The mean level of genetic differentiation for the gene pool (δ) was 5.2 % (Fig. 2). Four populations were found to be the most variable and also most differentiated (populations 5, 14, 17 and 18). In the case of populations 17 and 18, a large proportion of hybrid seeds were also observed (16.7 % and 22.7 % respectively; for further details see Gallo, this issue).

The two haplotypes previously found in *Nothofagus nervosa* (Marchelli et al. 1998) were also observed among the nine additional populations analysed here. Moreover, the geographic pattern reported in that study, with the existence of two distinctly separated groups in a north-south distribution, was maintained with the addition of more populations (Fig. 3).

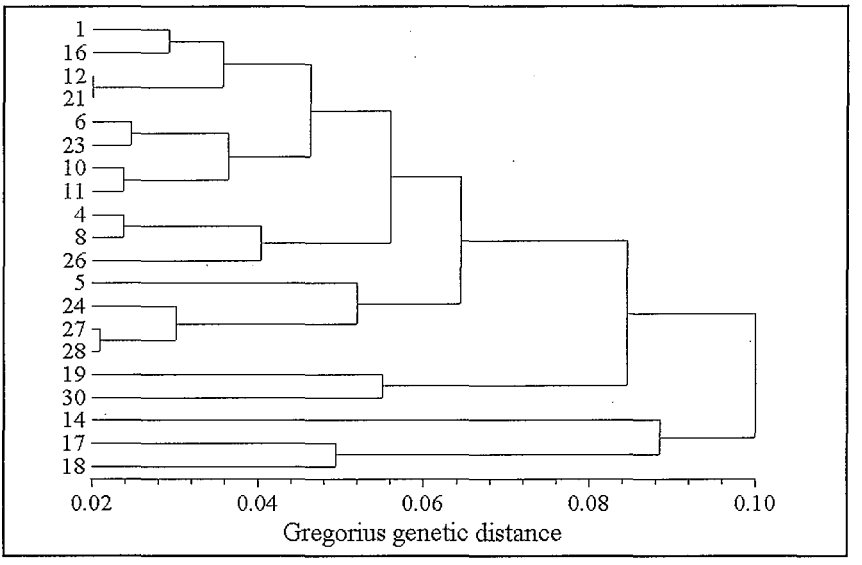


Figure 1: Cluster analysis using Gregorius genetic distance according to the UPGMA method.

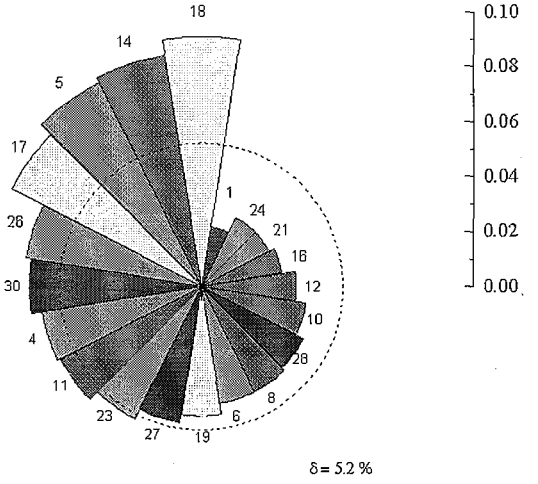


Figure 2. "Differentiation snail" (according to Gregorius and Roberds 1986) for the gene pool. The radius of the dotted circle represents the mean level of allelic differentiation (δ), each sector of the circle correspond to a population (same colours indicates same watersheds), their radius represents the level of genetic differentiation (Dj) and their angles denotes sample size.

Discussion

Considering the very limited and fragmented distribution range of *N. nervosa* in Argentina, the observed values of genetic differentiation among populations are relatively high. They show differentiation values similar to those *Nothofagus* species with restricted but continuous and larger distribution areas (e.g. *N. nitida*: G_{ST} = 4.7 %; Premoli 1997). Two of the most variable populations were those with the highest proportion of hybrid seeds, suggesting the importance of this process in generating genetic variation. As in a previous study including only eleven populations (Marchelli and Gallo 2001), high levels of differentiation among populations within watersheds were observed, indicating that the boundaries of a watershed were not strictly associated with genetic differentiation.

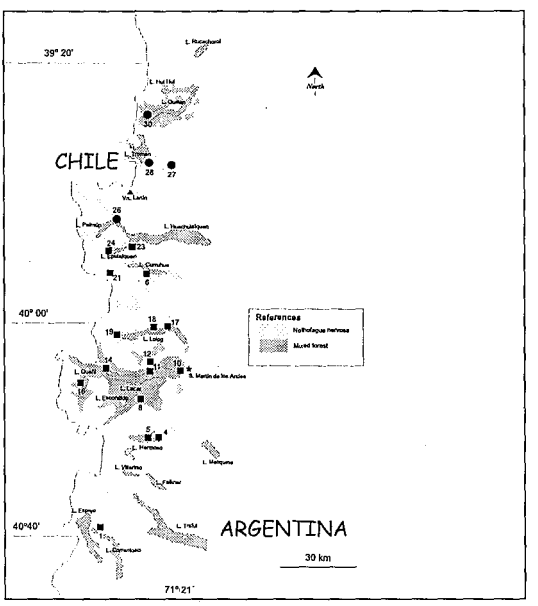


Figure 3. Geographic distribution of *Nothofagus nervosa* and the mixed forest with *N. obliqua* and distribution of the two cpDNA haplotypes

The occurrence of two different haplotypes, confirming previous results (Marchelli et al. 1998), suggests the existence of at least two glacial refugia. The high levels of allozyme diversity detected in some populations could be indicating the probable location of these refugia.

The results suggest a pattern for the distribution of the genetic variation which is mainly related both with hybridisation occurring in some populations,

and the recolonisation process from multiple refugia and, to a lesser extent, unidirectional gene flow within watersheds.

Acknowledgements

We would like to thank to the Silviculture Department of National Parks Administration, San Martín de los Andes, for their invaluable cooperation on seed collection, and to Lucas Oudkerk and Fernanda Izquierdo for technical assistence. This work was partially supported by a common project between National Parks Administration (APN) and National Institute of Agricultural Technology (INTA), by a grant from the IFS (International Foundation for Sciences) D/2458-1 and by projects from SAGPyA-BIRF and FONTAGRO.

References

Bianchi, M. M. (1997): Evidence of forest change during Late-Glacial-Holocene transition from lake sediment and peatbog pollen record at 41° South, East of the Andes. II Southern Connection Congress. Valdivia, Chile. pp 91-92.

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

Donoso, C. (1993): Bosques templados de Chile y Argentina. Variación, Estructura y Dinámica. Ecología Forestal. Editorial Universitaria, Santiago, Chile, 484 pp.

Gallo, L., Marchelli, P., Breitembücher, A. (1997 a): Allozymic heterozygosity in families with natural hibrids between two southern beeches (*Nothofagus spp.*). II Southern Connection Congress. 6 -11 January 1997, Valdivia, Chile. pp 152.

Gallo, L., Marchelli, P., Breitembücher, A. (1997b): Morphological and allozymic evidence of natural hybridization between two southern beeches (*Nothofagus spp.*) and its relation to heterozygosity and height growth. Forest Genetics, 4, 15-23

Gallo, L., Marchelli, P., Crego, P. (2000): Hybridisation dynamic between two southern beeches (*Nothofagus* spp.). XXI IUFRO World Congress, 7-12 August 2000, Kuala Lumpur, Malaysia. vol. 2 pp 43-44.

Gillet, E. (1994): Genetic Structures from Electrophoresis Data, Version 1.0. Abteilung für Forstgenetik und Forstpflanzenzüchtung. Universität Göttingen

Gregorius, H.-R. (1978): The concept of genetic diversity and its formal relationship to heterozygosity and genetic distance. Mathematical Biosciences, 41, 253-271

Gregorius, H.-R., Roberds, J. H. (1986): Measurement of genetical differentiation among subpopulations. Theor. Appl. Genet., 71, 826-834

Markgraf, V., McGlone, M., Hope, G. (1995): Neogene paleoenvironmental and paleoclimatic change in southern temperate ecosystems. A southern perspective. TREE, 10 (4): 143-147

Marchelli, P., Gallo, L., Scholz, F., Ziegenhagen, B. (1998): Chloroplast DNA markers revealed a geographical divide across Argentinean southern beech *Nothofagus nervosa* (Phil.) Dim. et Mil. distribution area. Theor. Appl. Genet., 97, 642-646

Marchelli, P., Gallo, L. (2000): Genetic analysis of isozyme variants in open pollinated families of Southern beech *Nothofagus nervosa* (Phil.) Dim. et Mil. Silvae Genetica, 49, 90-98

Marchelli, P., Gallo, L. (2001): Genetic diversity and differentiation in a southern beech subjected to introgressive hybridization. Heredity, 87: 284-293

Nei, M. (1973): Analysis of gene diversity in subdivided populations. Proc. of the National Academy of Sciences, 70, 3321-3323

Premoli, A. C. (1997): Genetic variation in a geographically restricted and two widespread species of South American *Nothofagus*. Journal of Biogeography, *24:* 883-892

Yeh, F. C., Boyle, T. (1997): POPGENE Version 1.2 Microsoft Window-based Software for Population Genetics Analysis. Alberta, Canada.

Past climatic changes and genetic diversity in the tropics: the example of *Vouacapoua americana*, a neotropical tree species – Preliminary results

C. DUTECH [1,2], L. MAGGIA [1], H. JOLY [3], and P. JARNE [2]

Address: [1]Laboratoire de Génétique et d'Ecologie Moléculaire – Silvolab, CIRAD-Forêt, BP 701, 97387 Kourou Cedex – French Guiana; [2]CNRS/CEFE, 1919 route de Mende, 34000 Montpellier – France; [3]CIRAD-Forêt, Campus international de Baillarguet, 34000 Montpellier Cedex 1 - France

Key words: chloroplast diversity, microsatellite, refuge hypothesis, tropical rainforest, *Vouacapoua americana*

Abstract: Several studies have suggested that climatic changes during the Pleistocene may have resulted in a reduction in tropical rainforest cover. During cool drier periods, populations of some tropical rainforest taxa may have been restricted to small and isolated areas where precipitation was sufficient to maintain this ecosystem (the refuge hypothesis). This hypothesis is supported by data sets on the spatial distribution of numerous taxa (insects, birds and trees) in South America and Africa. However, little data exist from the tropics to measure the impact of past climatic changes on genetic diversity. The refuge hypothesis makes two predictions for genetic diversity. First, a reduction in population sizes results in a decrease of genetic diversity within population, *i.e.* a bottleneck effect. Second, the geographic isolation results in an increase of genetic differentiation among populations. We studied the genetic diversity of *Vouacapoua americana*, a tree species of mature rainforests from the Guiana shield. Within French Guiana, we examined 22 populations for chloroplast diversity using a PCR-RFLP method, and 17 populations at nine microsatellite loci. A high genetic differentiation among populations exists for chloroplast diversity. Moreover, geographic clustering of chlorotypes strongly suggests the role of historical events such as population contractions, on the spatial organization of genetic diversity. In contrast, spatial genetic structure of the nuclear genome exhibited a pattern more consistent with the "isolation by distance model". In the light of these results, we discuss the possibility of post-Pleistocene recolonizations from several refuges in French Guiana and we suggest that reduction of population sizes in *V. americana* were probably ancient and/or slight.

Introduction

Among the factors that influence the genetic structure in natural populations, gene flow and history play a major role at large spatial scales. However, these two factors are difficult to dissociate in genetic analysis. Little data exist about this subject in neotropical tree species. Gene flow in neotropical tree species is

under the control of numerous and complex interactions with fauna. It is estimated that more than 80 % of tree species in the neotropics are pollinated by animals (Bawa 1990). Most tree species are also dispersed by mammals or birds (see for example Howe 1990). The number of animal species involved in pollination and seed dispersal are large and their behaviors are extremely diverse. Presently, the influence of these different parameters on genetic structure of tropical tree is not well measured and additional studies are needed to derive more general conclusions (Loveless 1992).

In terms of historical data, the effects of past climatic changes are widely debated (Willis 2000). Several authors have suggested that the cover of tropical rainforest shrank dramatically during the Pleistocene period (Prance 1982). Under this hypothesis, some residual fragments were isolated in areas where precipitation was sufficient to maintain this ecosystem. This hypothesis is called the refuge hypothesis. It has received substantial criticism (Colinvaux et al. 2000). Modifications of the global climate in the Pleistocene period certainly affected the tropics. But the impact of these climatic changes on rainforest populations was probably low, especially on the Guiana shield. For example, no palynologic data support the notion that Central Amazonian rainforest contracted substantially (Colinvaux et al. 2000). In this perspective, can any clear signals be detected in the genetic structure of a tree species which might be associated with possible ancient contraction of populations in South-America?

The method used here to study the effects of gene flow and historical factors on the genetic diversity in a species includes analysis of the diversity of the nuclear and the chloroplast genomes. Due to biparental transmission, nuclear genes are dispersed by a combination of seed and pollen dispersal. In contrast, the chloroplast genome is generally transmitted maternally, via seeds in angiosperm species. Comparison of the genetic structure of the two genomes provides insights on differential dispersal through seeds and pollen (Ennos 1994). Moreover, because of ploidy level, low mutation rates and low recombination rates, the chloroplast genome is more sensitive to bottlenecks than the nuclear genome, including a higher loss of diversity and a longer time to return to migration-drift equilibrium. Comparison of the diversity and its spatial organization between the two genomes offers the opportunity to consider the existence and the impact of ancient population contractions on the genetic structure of populations (see for example Le Corre et al. 1997).

Vouacapoua americana is a tree species with a geographic distribution limited to the Guiana shield. It ranges from Pará (Brazil) to the eastern part of Suriname, encompassing all of French Guiana. *V. americana* is a shade tolerant species of the mature rainforest in French Guiana, with low growth rates (Sabatier and Prévost 1990). If forest contractions were important in the past, this species should certainly be a good witness of these ancient events. Density

is of the order of 5 to 10 individuals per ha (data from Paracou Field Station French Guiana). Its spatial distribution is clustered in large patches of several hectares (Forget et al. 1999). This variation in local density is explained by edaphic conditions, availability of light (Forget 1997) and seed dispersal. The floral biology is characterized by small, fragrant yellow flowers, suggesting that pollination is carried out by small or medium-sized bees. We therefore assume that most pollen dispersal is probably limited to a few hundred meters. Seed are dispersed by rodents, Agouti (*Dasyprocta leporina*) and Accouchi (*Myorpocta exilis*). Most seeds are dispersed 10 to 30 m from the maternal trees (Forget 1990), but long distance dispersal events are possible. These biological traits predict a possible marked genetic structure at local and large spatial scale.

Materials and methods

The chloroplast diversity in *V. americana* was estimated in 22 populations, with three to six individuals sampled per population. Due to the low variability observed in populations in a preliminary study (less than four chlorotypes per population), it was not necessary to sample more individuals per population (Dutech et al. 2000a). The study on nuclear diversity was based on a sample of 17 populations that were already sampled for the study of the chloroplast genome. Twenty one to 24 individuals per population were collected.

The analysis of the chloroplast genome was performed using a polymerase chain reaction method with universal primers, followed by a random fragment length polymorphism analysis (Dutech et al. 2000a). We used eight pairs of primers and three endonucleases. The nuclear genome was studied using nine microsatellite loci which were cloned in *V. americana* (Dutech et al. 2000b).

Classical index of gene diversity (expected heterozygosity, number of alleles) and genetic differentiation among populations were estimated for the chloroplast genome with the Haplodiv software (available from R. J. Petit) following the method of Pons and Petit (1995), and for the nuclear genome using the Genepop package (Raymond and Rousset 1995).

Results and discussion

Five polymorphic fragments were detected by the PCR-RFLP method in the chloroplast genome. The combination between different mutations across loci produced seven chlorotypes observed in a sample of 22 populations dispersed

over all of French Guiana (Fig. 1). Genetic diversity within populations was low ($H_s < 0.15$). In contrast, the differentiation among populations was high, with a G_{st} value higher than 0.8. It is among the highest values encountered in the literature for chloroplast diversity in tree species (Raspé et al. 2000). This high differentiation between populations is consistent with the apparent low capacity of seed dispersal in the species (Forget 1990; Dutech et al. 2002). We detected large patches of identical chlorotypes in the north of French Guiana. These patches are roughly limited by the rivers. There are, however, some exceptions, such as for the "cross" haplotype and central chlorotype (see Fig 1).

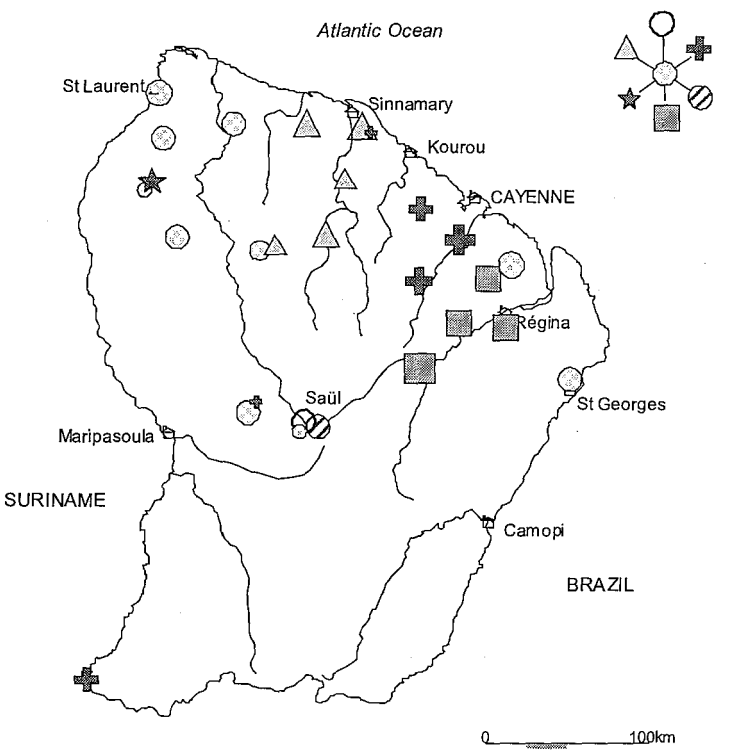


Figure 1. Geographic distribution of the seven cpDNA haplotypes in *V. americana* in French Guiana. The relationship between all the chlorotypes is given in the upper right corner. The size of the symbol is proportional to the number of individuals sampled.

Such homogeneity at a rather large geographic scale (30 to 60 km) strongly suggests an effect of historical factors. We can hypothesize that population regressions led to the fixation of one chlorotype in different small isolated populations. Recolonization from these populations produced large homogenous patches for single chlorotypes. Some patches are well associated with possible refuge areas, such as the "square" haplotype in Montagnes Tortues, the central

chlorotype in the lucifer - dekou-dekou and the "triangle" haplotype with Montagne Trinité (Fig. 1). These three areas are proposed as putative Pleistocene refuges (Granville 1982). However, spatial structure of haplotypes seems also to be organized relative to river systems. This genetic structure suggests some rainforest refuges on the borders of main rivers. This "riparian refuge" hypothesis has been discussed in recent papers (see for example, Aide and Rivera 1998). The present study was limited, however, to populations in northern French Guiana. Data from southern populations are needed, especially to discuss the influence of the central area in the recolonization process.

The mean number of alleles per locus and per population is low for microsatellite loci. Except for locus Wac1, with 10 alleles per locus, the eight other loci have less than five alleles per population (Fig. 2). The populations exhibit roughly similar allelic richness (Dutech 2001). This low variability could result from ancient dramatic bottleneck events, small effective population sizes, or alternatively, from poorly variable loci. Using other nuclear markers with different mutation rates would be informative in this respect.

The mean F_{is} for populations and loci is near 0 (see Dutech 2001). Tests of departures to Hardy-Weinberg expectation are non-significant for each locus in each population, except nine out of 150 tests. No selection, there were no null alleles, and no subpopulation structure was detected within populations. The overall genetic differentiation among populations is low ($F_{st} = 0.08$). It is very similar to F_{st} values estimated in other tropical tree species (Loveless, 1992). All the F_{st} values between pairs of population are significantly different from 0 except one pair (Dutech, 2001). This low differentiation among populations is not consistent with the assumed limited pollen dispersal by insects and long time divergence between refuge populations, but using overall genetic differentiation to infer gene flow must be done with caution (Whitlock and McCauley 1999).

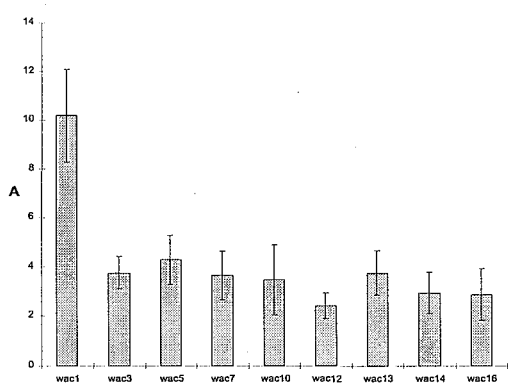


Figure 2. Average number of alleles (A) of nine microsatellite loci in 17 populations of *Vouacapoua americana* in French Guiana. Lines give standard errors.

We tested whether a model of isolation by distance could be fitted to data. The method used is based on the regression of $F_{st} / (1 - F_{st})$ over the logarithm of the geographic distance (see Rousset 1997). The probability of rejecting the null hypothesis (*i.e.* absence of isolation by distance) is significant ($p = 0.02$), which can explain part of the observed differentiation among populations. On the other hand, we were not able to relate the distribution of microsatellite diversity with any feature of the refuge hypothesis (Dutech 2001).

Conclusions

Although population retraction into refugia seems to have influenced the chloroplast diversity, we do not find a clear signal for this process for the nuclear genome. Different explanations can be suggested, such as differences in gene flow between seeds and pollen. Recent results using microsatellite loci, obtained at a smaller geographic scale, indicate a strong genetic structure, which is not in agreement with high gene flow through pollen (Dutech et al. 2002). Alternatively, the absence of a clear pattern of population grouping at larger scales might result from the lower sensitivity of the nuclear genome to bottleneck events. To further test the hypothesis of Pleistocene refuges, it would be interesting to conduct similar studies in other species. Studies in *Dicorynia guianensis* in French Guiana (Caron et al. 2000), which give similar results, are encouraging. Work at a larger geographical scale is also required, especially in Brazil, because historical factors at work in French Guiana are undoubtedly integrated into a larger pattern of recolonization by the different tree species in the neotropics.

References

Aide, T., Rivera, E. (1998): Geographic patterns of genetic diversity in *Poulsenia armata* (Moraceae) implications for the theory of Pleistocene refugia and the riparian forest. J. Biogeog. 25: 695-705

Bawa, K. S. (1990): Plant - pollinator interactions in tropical rain forests. Ann. Rev. Ecol. Syst. 21: 399-422

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

Colinvaux, P. A., de Oliveira, P. E. (2000): Palaeocology and climate of the Amazon basin during the last glacial cycle. J. Quaternary Sci. 15: 347-356

Dutech, C. (2001): *Diversité génétique et dynamique de la forêt tropicale humide: le cas d'une espèce ligneuse guyanaise, Vouacapoua americana (Caesalpiniaceae).* Thèse de Doctorat, Université Montpellier II

Dutech, C., Maggia, L., Joly, H. I. (2000a): Chloroplast diversity in *Vouacapoua americana* (Caesalpiniaceae) a neotropical forest tree. Mol. Ecol. 9: 1427-1432

Dutech, C., Amsellem, L., Billotte, N., Jarne, P. (2000b): Characterization of (GA)n microsatellite loci using an enrichment protocol in the neotropical tree species *Vouacapoua americana*. Mol. Ecol. 9: 1433-1449

Dutech, C., Seiter, J., Petronelli, P., Joly, H. I. , Jarne, P. (2002): Evidence of low gene flow in a neotropical tree species in two forest stands of French Guiana. Mol. Ecol. 11: 725-738

Ennos, R. A. (1994): Estimating the relatives rates of pollen and seed migration among plant populations. Heredity 72: 250-259.

Forget , P.-M. (1990): Seed dispersal of *Vouacapoua americana* (Caesalpiniaceae) by caviomorph rodents in French Guiana. J. Trop. Ecol 6: 459-468

Forget, P. M. (1997): Effect of microhabitat on seed fate and seedling performance in two rodent-dispersed tree species in rain forest in French Guiana. J. Ecol. 85: 693-703

Forget, P. M., Mercier, F., Collinet, F. (1999): Spatial patterns of two rodent-dispersed rainforest trees, *Carapa procera*(Meliaceae) and *Vouacapoua americana* (Caesalpiniaceae) at Paracou, French Guiana. J. Trop. Ecol. 15: 301-313.

Granville, J. (1982): Rain forest and xeric flora refuges in French Guiana. pp. 137-158 in: Biological diversification in the tropics - Columbia University Press, New York.

Howe, H. F. (1990): Seed dispersal by birds and mammals: implications for seedling demography. pp. 191-218 in: Bawa K. S. and Hadley M. (Eds.) Reproductive ecology of tropical forest plants,. Man and the Biosphere Series, Paris.

Le Corre, V., Dumolin-Lapègue, S., Kremer, A. (1997): Genetic variation at allozyme and RAPD loci in sessile oak *Quercus petraea* (Matt.) Liebl.: the role of history and geography. Mol. Ecol. 6: 519-529

Pons, O., Petit , R., J. (1995): Estimation, variance and optimal sampling of gene diversity. I Haploïd locus. T.A.G. 90: 462-470

Prance, G. T. (1982): Forest refuges: evidence from woody angiosperms. pp. 137-158 in : Prance, G. T. (Eds.) Biological diversification in the tropics. Columbia University Press, New York.

Raspé, O., Saumitou-Laprade, P., Cuguen, J., Jacquemart, A.-L. (2000): Chloroplast DNA haplotype variation and population differentiation in *Sorbus aucuparia* L. (Rosaceae : Maloideae). Mol. Ecol. 9: 1113-1122.

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

Rousset, F. (1997): Genetic differentiation and estimation of gene flow from F-statistics under isolation by distance. Genetics 145: 1219-1228.

Sabatier, D., Prévost, M.-F. (1990): Quelques données sur la composition floristique et la diversité des peuplements forestiers de Guyane française. B. F. T. 219: 31-51

Willis, K. J., Whittaker, R. J. (2000): The refugial debate. Science 287: 1406-1407

Whitlock, M. C., McCauley, D. E. (1999): Indirect measures of gene flow and migration: Fst ≠ 1/(1+4Nm). Heredity 82: 117-125.

Regional scale genetic structure within two Central American tree species the influence of geography, biology and geological history

S. CAVERS and A. LOWE

[1] *Centre for Ecology and Hydrology, Edinburgh Station, Bush Estate, Penicuik, EH26 0QB, UK.*

Key words: genetic diversity, phylogeography, regional variation, cpDNA, AFLP, *Cedrela odorata*, *Vochysia ferrugiea*.

Abstract: We examine spatial genetic structure occurring at a regional scale for two widely distributed tree species from Central America, *Vochysia ferruginea* and *Cedrela odorata*, and speculate on some of the geographic, edaphic and geological factors that may be responsible in a continent that has a complex geological history and is rich in natural resources. For *V. ferruginea*, the central mountain ranges of Costa Rica coincide with a major genetic differentiation and are probably acting to maintain this genetic split by reducing gene flow. A major genetic differentiation within Costa Rica is also apparent for *C. odorata*, for both nuclear and chloroplast markers. However, this split coincides with habitat type (dry vs. moist), not with geographic features and appears to be a case of ecotypic differentiation or possibly cryptic speciation. A cpDNA phylogeographic analysis of *C. odorata* across Central America reveals that the strong genetic/ ecotypic differentiation continues across a large portion of the region. The pattern of cpDNA variation is discussed within the context of species colonisation and extinction. However, a scenario that involves separate and repeated colonisation of Central American regions from diverse sources most closely fits the present pattern. Future work to examine cpDNA variation across the Caribbean and South America should shed more light on this issue. Finally the importance of such genetic data for prioritising the management of forest genetic resources is discussed.

Introduction

Mesoamerica is recognised as a global biodiversity hotspot due to the high number of species found in the region (e.g. approximately 24 000 plant species) and the current high rates of habitat destruction. The pressure on primary vegetation is extreme, and only 20% of the original extent of primary vegetation now remains (Myers *et al.* 2000). Commercially valuable timber species have been particularly heavily impacted due to targeted extraction combined with habitat loss. Conservation efforts are now being made, but protected land is not always coincident with the distribution of threatened species. Furthermore, the importance of managing genetic resources of such species is also generally

poorly appreciated. In particular, there is little information on the evolutionary history of species or the extent and structure of genetic diversity for species that cover a wide geographic area.

To develop realistic management strategies for the genetic resources of key species, it is essential to have an understanding of the structure of genetic diversity within existing populations. It is also essential to know how the dynamics of gene flow will affect variation. The structure of genetic variation can be influenced by many factors including: biology (breeding system, seed and pollen dispersal mechanisms and regeneration ecology), geography (topographic features which may reduce gene flow) and history (geology and biogeography). Studies that adopt a comparative approach, for species with similar ranges, will allow the identification of those factors that influence genetic structure.

Presented here are preliminary data on regional-scale genetic structure within two tree species from Costa Rica chosen for their economic importance and wide ranging distributions: *Cedrela odorata* and *Vochysia ferruginea*. In addition, *C. odorata* was subject to a cpDNA phylogeographic analysis across much of its Central American range to examine patterning of genetic variation at a continental scale.

Materials and methods

Cedrela odorata L. is a fast growing, light demanding tree, distributed from Mexico (26°N) to Argentina (28°S). It is monoecious, pollinated by small bees and wasps and has wind dispersed seed. It is a commercially valuable timber and is threatened by deforestation and over exploitation. Its modern day distribution is patchy, and individuals are often left as remnant trees on farmland established on cleared forest areas. It is capable of tolerating fairly dry conditions (1200-2000 mm rainfall p.a.) and grows up to 1200 m above sea level. *C. odorata* is recognised in the Costa Rican National Conservation Strategy and is listed by the IUCN as vulnerable.

Vochysia ferruginea Mart. is a typical canopy tree of lowland forest, and distributed from Nicaragua to Brazil. It is a pioneer species and recolonises cleared land, forming dense secondary stands. It prefers moist conditions (2500-4000 mm rainfall p.a.) and grows up to 800 m above sea level. The species tolerates high soil concentrations of aluminium and iron and is fast growing, which together with its good timber qualities, makes it ideal for reforestation or reclaiming degraded land. Individuals bloom synchronously in populations but its breeding system has been variably reported as outcrossing to strict autogamy,

including cleistantheric pollination. It has hermaphroditic flowers that are pollinated by birds and insects (mainly larger bees) and has wind dispersed seed.

Collections

For both species, collections comprised up to 20 individuals from each of ten Costa Rican populations (Table 1, 2). In each case, collections were from the full geographic range of the species within Costa Rica (Fig 1). For *C. odorata*, additional collections were made throughout Central America, with a total of 29 populations sampled from Mexico, Guatemala, Honduras, Nicaragua, Costa Rica and Panama. Each of these populations consisted of between 5 and 20 individuals (Table 1, 2).

PCR-RFLP screening of chloroplast DNA

Full screening is still underway but preliminary results are presented here. DNA extraction, PCR, restriction digestion and gel electrophoresis methods followed those of Demesure et al. (1995). Preliminary analysis to identify polymorphism within each species involved screening a single individual from each population for several primer/enzyme combinations. In total, 79 primer/enzyme combinations were analysed, and the combined length of PCR products covered approximately 38,000 bp of the chloroplast genome. For *C. odorata*, where five cpDNA haplotypes were identified, genetic relationships between haplotypes were resolved using minimum spanning methods.

AFLP analysis

Ten Costa Rican populations of up to 20 individuals of *C. odorata* were further screened for AFLP variation (Table 1) using the protocol of Vos *et al.* (1995). Genomic DNA was digested using the enzymes *Mse*I and *Eco*RI and preamplified using primers *Mse*I+AC and *Eco*RI+0. The selective amplification step used primers *Mse*I+ACAG and *Eco*RI+CC. Genetic distances between populations were analysed using Nei's distance algorithm based on population fragment frequencies (POPGENE, Yeh 1997). Pairwise genetic distances were clustered using the Neighbour-Joining algorithm (PHYLIP, Felsenstein 1993).

Table 1. List of collection sites, locations, sample sizes for cpDNA and AFLP studies, and cpDNA PCR-RFLP haplotypes for ***Cedrela odorata***.

Country	Population	Lat.	Long.	RFLP1	RFLP2	RFLP3	RFLP4	RFLP5	AFLP
				Central	Southern	Northern	Guatemala	Panama	
Costa Rica	Puriscal	9.93	-84.29	20	0	0	0	0	13
Costa Rica	Horizontes	10.74	-85.59	20	0	0	0	0	13
Costa Rica	Canas	10.2	-84.95	20	0	0	0	0	13
Costa Rica	Hojancha	10.09	-85.37	20	0	0	0	0	17
Costa Rica	Palo Verde	10.35	-85.35	20	0	0	0	0	1
Nicaragua	Masatepe	11.9	-86.14	20	0	0	0	0	
Nicaragua	Ometepe	11.49	-85.49	20	0	0	0	0	
Nicaragua	La Trinidad	12.99	-86.23	20	0	0	0	0	
Nicaragua	Wabule	12.88	-85.68	20	0	0	0	0	
Honduras	Comayagua	14.15	-87.62	20	0	0	0	0	
Honduras	Meambar	14.83	-88.1	20	0	0	0	0	
Honduras	La Paz	14.42	-87.05	20	0	0	0	0	
Honduras	Taulabe	14.83	-88.1	20	0	0	0	0	
Costa Rica	Jimenez	10.21	-83.61	0	20	0	0	0	14
Costa Rica	Talamanca	9.63	-82.85	0	20	0	0	0	13
Costa Rica	Upala	10.79	-85.03	1	19	0	0	0	14
Costa Rica	Pacifico Sur	8.54	-82.85	0	20	0	0	0	16
Costa Rica	Perez Zeledon	9.33	-83.65	0	20	0	0	0	10
Panama	Gualaca	8.58	-82.24	0	0	0	0	20	
Panama	Las Lajas	8.2	-81.86	0	16	0	0	4	
Panama	San Francisco	8.24	-80.98	0	0	0	0	20	
Guatemala	Los Esclavos	14.25	-90.28	0	0	4	1	0	
Guatemala	Tikal	17.23	-89.62	0	0	5	0	0	
Guatemala	San Jose	17.18	-89.86	0	0	5	0	0	
Guatemala	El Idolo	14.43	-91.38	0	0	5	0	0	
Mexico	Bacalar	18.29	-89.15	0	0	20	0	0	
Mexico	Calakmul	18.45	-88.32	0	0	20	0	0	
Mexico	Zona Maya	19.36	-88.02	0	0	20	0	0	
Mexico	Escarcega	18.4	-90.9	0	0	20	0	0	

Table 2. List of collection sites, locations, sample sizes and cpDNA PCR-RFLP haplotypes for ***Vochysia ferruginea.***

Country	Population	Lat.	Log.	RFLP1	RFLP2
Costa Rica	La Marta	9.76	-83.68	17	
Costa Rica	Penjamo	10.33	-84.48	9	
Costa Rica	Tirimbina	10.39	-84.14	13	
Costa Rica	Perez Zeledon	9.36	-83.69		20
Costa Rica	Coto Brus	8.9	-83.09		20
Costa Rica	Pindeco	9.2	-83.47		19
Costa Rica	Brasillea	11.03	-85.36	10	
Costa Rica	Osa	8.72	-83.49		11
Costa Rica	Puriscal	7.17	-89.36		13
Costa Rica	Cano Negro	10.95	-84.71	-	-

Results

For both species, two chloroplast haplotypes were found within Costa Rica, and distributions are plotted in Figure 1. For *V. ferruginea*, no within population variation was found, whereas for *C. odorata*, one population (Upala) was found to be polymorphic. *Vochysia ferruginea* exhibited significant intraspecific genetic differentiation that coincided with the central mountain range of Costa Rica, i.e. individuals from Pacific populations possessed a different haplotype to those from the Atlantic. By contrast, the major genetic differentiation for *C. odorata* did not correlate with the mountain barrier, but closely corresponds to habitat type (wet and dry; Figure 3).

Preliminary AFLP results for *C. odorata* are presented in Figure 2. Of 52 loci scored using a single primer combination, 40 were polymorphic. The pattern of variation is similar to that found for cpDNA analysis, as populations from drier regions (NW Costa Rica) clustered together and were strongly differentiated from those of moist regions. Further differentiation occurred between populations within the moist regions, and populations on the Atlantic side of the mountain chain (Jimenez, Talamanca and Upala) were distinct from those on the Pacific side (Pacifico Sur and Perez Zeledon).

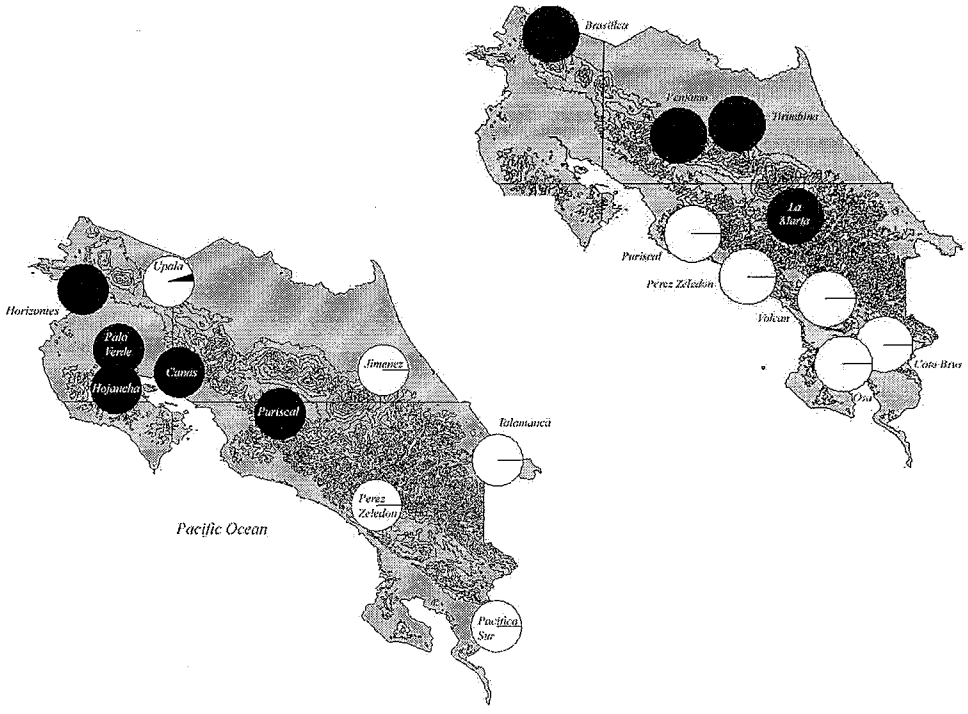


Figure 1. Distribution of cpDNA haplotypes for *C. odorata* (left) and *V. ferruginea* (right).

Finally the analysis of cpDNA variation across the range of *C. odorata* in Central America identified a total of five cpDNA types. Their location and minimum spanning relationship can be seen in Figure 3. Three commonly occurring haplotypes were found, restricted to three specific areas: 1. Dry regions of Honduras, Nicaragua and Costa Rica; 2. Moist regions of Costa Rica and Panama; and 3. Mixed habitat types in Mexico and Guatemala. Two other, rare haplotypes, 4 and 5, were found in Guatemala and Panama respectively (Figure 3). The minimum spanning relationship between the haplotypes is shown in the inset of Figure 3. Of the common haplotypes, haplotypes 1 and 2 were most similar (3 mutations differentiated them) but haplotype 3 was most similar to haplotype 2 (5 mutations) rather than the geographically proximate haplotype 1 (8 mutations). Of the less common haplotypes, 5 was most similar to 2, and 4 to 3, with one mutation differentiating each pair.

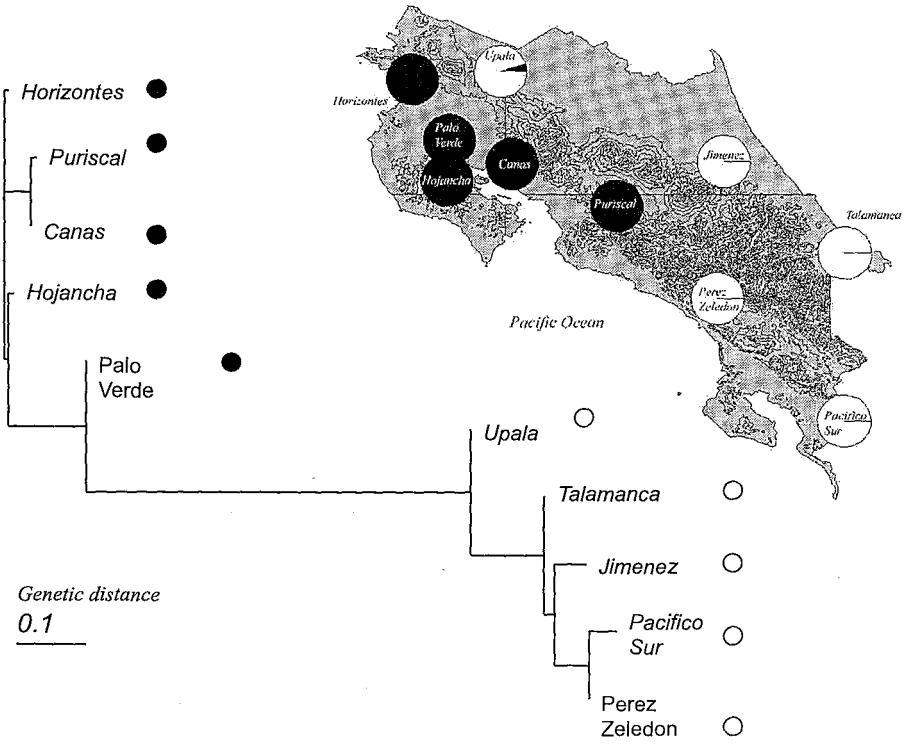


Figure 2. Neighbouring joining dendrogram based on shared presence of 52 AFLP fragments for 10 Costa Rican populations of C. odorata. Map indicates location of sampled populations.

Discussion

Geographic partitioning of haplotypes

Significant genetic differentiation within Costa Rica was uncovered for both species. For *V. ferruginea*, genetically differentiated populations were split by the central mountain ranges of the country; the Cordillera de Guanacaste, Cordillera de Tilaran, Cordillera Central and Cordillera de Talamanca (Figure 1). For *C. odorata*, AFLP analysis of moist region populations also demonstrated genetic differentiation of populations that were split by the central mountain ranges. Several reasons can be postulated to explain such genetic differentiation

including isolation and drift caused by the uplift of the mountain chain (at least 1 million years ago), or colonisation of the Pacific and Atlantic areas from different source populations. However, at this stage it is difficult to identify the reasons for the genetic differentiation of Atlantic and Pacific populations. What we can say is that this central mountain range represents a significant barrier to gene exchange and is probably responsible for the maintenance of some of the genetic differentiation within both species.

Ecological partitioning of haplotypes

In contrast to the situation found for *V. ferruginea*, the major intraspecific genetic differentiation within *C. odorata* coincides closely with rainfall and habitat, not the mountain barrier (Figure 1 and 3). This split is apparent from both cpDNA and AFLP analysis and indicates significant differentiation. This significant genetic divergence was also noted by an earlier RAPD study (Gillies *et al.* 1997) and individuals growing within the two habitats have significant morphological and physiological differences (Newton *et al.* 1997). Atlantic region populations have a higher growth rate, redder newly emergent leaves and smoother bark than those from the northern Pacific region (Guanacaste).

Significant differentiation associated with habitat is also seen for other species within this region. For example, the genus *Swietenia* exhibits two forms in Costa Rica, differentiated according to habitat, and in this case classified as separate species. *Swietenia macrophylla*, a widely distributed, moist forest species, inhabits Atlantic and southern Pacific areas similar to the 'wet' form of *C. odorata*, whereas, *S. humilis* is only found in the dry areas of Guanacaste and further north along the Pacific Coast of Central America. The work of Chase *et al.* (1995) on *Cordia alliodora* also demonstrates a significant genetic split along an Atlantic/Pacific axis.

The reasons for this strong habitat-correlated genetic divergence are more difficult to elucidate. Certainly there is some evidence for strong selection against the dry forms when they are reared in the wet regions and vice versa (Adrian Newton pers. comm.). The AFLP data show no evidence of hybridisation between the ecotypes, even in populations where they are sympatric, such as Upala. This suggests the presence of a post-zygotic breeding barrier maintaining the genetic difference between forms. However, further work will need to be done, particularly in the area of sympatry, to examine fertility and selection pressures on inter-ecotype crosses. At the moment these genetically and morphologically distinct populations are recognised as two ecotypes of the same species, but there may be good arguments to revise the taxonomy and recognise the two forms as separate species. Indeed the two forms may have been isolated for a considerable period of time, as evidenced by the considerable differentiation

for neutral characters, and are only now in secondary contact after evolution in allopatry. In such a case the two forms should be more appropriately described as separate species.

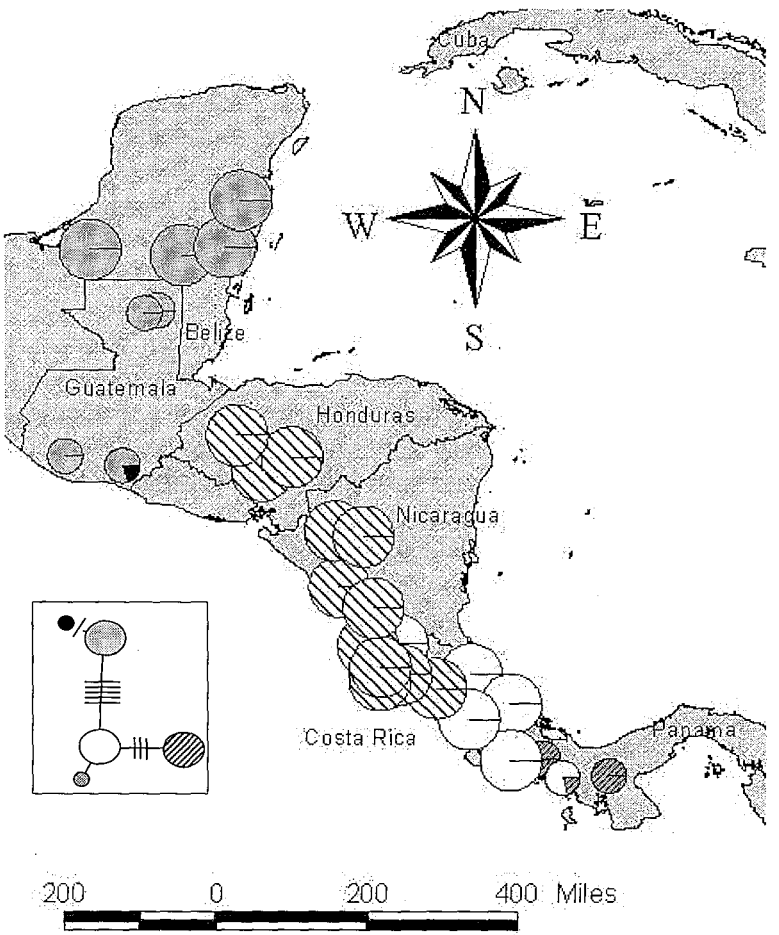


Figure 3. Distribution of five cpDNA haplotypes found within *C. odorata* samples collected from across Central America. Mean annual precipitation of locations is indicated. Primer enzyme combinations used to distinguish haplotypes and populations sample sizes are presented in Table 1. In this figure, Haplotype 1 = central, Haplotype 2 = southern, Haplotype 3 = northern, Haplotype 4 = Guatemalan and Haplotype 5= Panamanian.

Range wide phylogeography, influence of species history

The range of sampling of *C. odorata* was extended across most of its Central American distribution. The correlation between habitat type and genetic differentiation observed within Costa Rica is maintained throughout the continental sample (Figure 3) where haplotype 1 occurs in dry regions within Costa Rica, Nicaragua and Honduras and haplotype 2 occurs only in moist regions of Costa Rica and Panama. A third major haplotype (3) was found within Mexican and Guatemalan populations. It was found predominantly in wet regions but also occurred in dry areas of the Yucatan Peninsula. In addition, two rare haplotypes were found. Haplotype 4 was restricted to a single individual from Guatemala and haplotype 5 was found in central Panamanian populations. The distribution of the most common haplotypes and their genetic relationship is somewhat puzzling. Populations in Mexico possess a haplotype that is most closely related to the southern type (haplotype 1) and not the haplotype (2) found in proximate Honduran populations. One explanation for such a pattern could be that the Mexican material represents a different colonisation history to other Central American populations. Pollen of *Cedrela* species has been recorded from the Yucatan Peninsula in the early Miocene (20 MY; Graham 1999), well before the formation of the Isthmus of Panama. From this time, either the Caribbean Islands, which were located much closer to the continental masses of Central and South America, or an island arc along the line of present day Central America, could have provided routes by which *Cedrela* species could have established in the Yucatan. Following the rise of the Isthmus and the joining of Central and South America via Panama approx. 3 million years ago, further colonisations from different sources could have established *C. odorata* in southern Central America, creating the boundaries between different genetic types that we observe in the present day distribution. This idea will need further testing and will include sampling of the Caribbean Islands and South American populations. In particular we will need to establish the origin of this genus and potential source and colonised regions. This work is on-going and the hypothesis will be tested alongside other scenarios that include extinctions and recolonisations during past climatic changes.

Conclusions and application to genetic resource management

The structure of genetic diversity within a species range is established by complex interactions between the biology of the species and its geographic and environmental history. However, it is not yet possible to provide predictive estimates of how each of these factors contribute to establishing genetic structure.

Indeed in a survey of the isozyme literature, Hamrick et al. (1992) were only able to explain 34% of the variation in the data set by combining all biological and ecological characters. Thus individual case studies are still required when developing appropriate genetic resource management strategies for target species.

For the species examined here, the geographic regions of the Pacific and Atlantic need to be recognised as genetically distinct for *V. ferruginea*. For *C. odorata*, the major differentiation between the ecotypes is of most importance to be considered but also the significant subdivision across the mountain barrier within the range of the wet ecotype should be recognised. Further work is required to establish the taxonomic relevance of these differentiations. Whilst these are specific case studies, similar patterns of variation for other species are also seen within the region. It is certain that the particular pattern of topography and environment of the area has had an influence on many species. When considering other Costa Rican target species for which no genetic data is available it would be appropriate to include Atlantic and Pacific, and wet and dry regions when designing any preliminary sampling strategy.

The influence of species history is particularly difficult to predict. Whilst interesting observations are possible from the study of *C. odorata* it is not possible to use them to predict the evolutionary history of the other species in this study. The study of *C. odorata*, is however, informative with regard to the management of genetic resources, and gives insight into the time that the different areas of the species range may have been isolated and their potential connectivity.

It is clear from these case studies that many factors cause significant genetic differentiation within species, and need to be taken into consideration during the planning of genetic resource management. Such divergence can be due to a complex interaction of geographic, adaptive and historical processes (Newton et al. 1999), particularly in a geologically and environmentally dynamic region like Central America.

Acknowledgements

We would like to thank Carlos Navarro, Marvin Hernandez and Lionel Coto for making field collections and Kevyn Wightman (ICRAF) for providing provenance samples; Patrick Hopkins for undertaking lab work on *V. ferruginea*, Richard Ennos and Adrian Newton for providing useful feedback during the progress of the work; and two anonymous referees for constructive discussion. The financial support of NERC and the EU is gratefully acknowledged and this work was undertaken as part of an INCO-DC funded project 'Assessment of levels and dynamics of intra-specific genetic diversity of tropical trees for conservation and sustainable management', contract number ERBIC18CT970149.

References

Chase, M., Boshier, D.H., Bawa, K. S. (1995): Population genetics of *Cordia alliodora* (Boraginaceae), a neotropical tree. 1. genetic variation in natural populations. American Journal of Botany. 82: 468-475

Demesure, B., Sodzi, N., et al. (1995): A set of universal primers for amplification of polymorphic non-coding regions of mitochondrial and chloroplast DNA in plants. Molecular Ecology. 4: 129-131

Felsenstein, J. (1993): PHYLIP (Phylogeny Inference Package), version 3.5C. Distributed by the author. Department of Genetics, University of Washington, Seattle, WI.

Gillies, A.C.M., Cornelius, J.P., Newton, A.C., Navarro, C., Hernandez, M., Wilson, J. (1997): Genetic variation in Costa Rican populations of the tropical timber species *Cedrela odorata* L. assessed using RAPDs. *Molecular Ecology.* 6: 1133-1146

Graham, A. (1999): Studies in Neotropical paleobotany. XIII. An Oligo-Miocene palynoflora from Simojovel (Chiapas, Mexico). American Journal of Botany. 86: 17-31

Hamrick, J.L., Godt, M.J.W., Sherman-Broyles, S.L. (1992): Factors influencing levels of genetic diversity in woody plant species. New Forests 6: 95-124

Newton, A.C., Allnutt, T., Gillies, A.C.M., Lowe, A.J., Ennos, R.A. (1999): Molecular phylogeography, intraspecific variation and the conservation of tree species. Trends in Ecology and Evolution. 14: 140-145

Toledo, V. (1981): Pleistocene changes of vegetation in tropical Mexico. in: Pance, G.T. (Ed.). Biological Diversification in the Tropics. Columbia University Press, New York.

Vos, P., Hogers, R., et al. (1995): AFLP: a new technique for DNA fingerprinting. Nucleic Acids Research. 23: 4407-4414

Yeh, F.C. (1997): POPGENE, version 1.21. Distributed by the author. University of Alberta. http://www.ualberta.ca/~fyeh/fyeh/.

Preliminary genetic parameters from *Swietenia macrophylla* trials in the Yucatan, Mexico

S. E. WARD[1], K. WIGHTMAN[2], J. HAGGAR[2], and B. RODRIGUEZ SANTIAGO[3]

[1]International Institute of Tropical Forestry, PO Box 25,000, San Juan, Puerto Rico, [2]International Center for Research in Agroforestry, Chetumal Mexico. Current address: CATIE-MIP, Nicabox 112, PO Box 52-7444, Miami, Florida 33152-7444; [3]Instituto Nacional de Investigaciones Forestales y Agropecuarias A. Postal 182 Chetumal, Mexico 77000

Key words: *Swietenia macrophylla, Hypsipyla grandella*, Yucatan, multilocal genetic trial, heritability

Abstract: In 1996 several research organizations collaborated in a collection of germplasm of *Swietenia macrophylla* in the Yucatan peninsula for genetic characterization, conservation, and tree improvement for local reforestation programs. Five areas were sampled as provenances, and a total of 100 families were collected. From this collection, three genetic trials were established in the Yucatan at the San Felipe Bacalar Research Station, in the community of Zoh Laguna, and at a farm near Felipe Carillo Puerto in the Zona Maya. The purpose of these trials was to conserve genetic diversity as well as to provide improved seed for local nurseries. The trials were designed with 47 to 56 families per site, 4 to 5 trees per family plot, and 5 replicates per site. The measured variables included tree height, shoot borer attack at apex, branching at apex, and overall form. By 2 years, the average height of trees ranged from 129 to 151 cm per site. The shoot borer (*Hypsipyla grandella)* attacked from 7% to 68% of trees per site in the second year. For all three sites together, the variation components due to provenance and family effects explained less than 1% of the total variation. For each site considered separately, the block and block*family interaction effects generally explained the most variation in a trait, and family level variance components were usually much larger than those at the provenance level. At each site, the highest individual heritabilities at year 2 were for relative height increment (h^2= 0.14 to 0.55). According to these early results, minimal population differentiation appears to have occurred across the Yucatan. In addition, there appears to be limited potential for selection of material that will perform well across the conditions of the Yucatan. These results may be attributable to high environmental heterogeneity. The selection of families and individuals for conversion of these trials to seed orchards may be best made on a per site basis, for use on nearby properties. These trials are very young and may yet be strongly influenced by small scale environmental effects. They will be followed for several more years to adequately predict adult tree behaviour.

Introduction

The conservation status of Neotropical mahogany (*Swietenia* spp.) has been the focus of much concern (Lugo 1999). Big-leaf mahogany *(S. macrophylla)*

remains a harvested timber, but the natural range has diminished as populations have been heavily exploited in Mexico and Central America, with prospects for commercial extinction in other parts of the range (C. Navarro, T. Killeen, pers. comm.). The conservation of genetic resources is critical, both as a component of biodiversity, and for future human use. However, collections and trials of mahogany germplasm have been limited (but see Geary et al. 1973, Ward and Lugo 1998, in press).

The range of big-leaf mahogany extends into the Yucatan peninsula of Mexico. The political organization of Yucatan territory into community properties (ejidos) renders cooperation necessary with local people for conserving this species. Genetic improvement of mahogany would be valuable for the economic well being of these communities: in plantations, agroforestry, and forest enrichment. The central barrier to plantation success with mahogany may be damage by the mahogany shoot borer, *Hypsipyla grandella* (Newton et al. 1993, Mayhew and Newton 1998), which attacks the apical meristem of young trees. Identification of resistant material would improve the likelihood of success of mahogany plantings in the Yucatan region. To meet the objectives of conservation and providing improved seed to local users, a collection of mahogany germplasm was made in the Yucatan in 1997 and established in several genetic trials in 1998.

We present here the preliminary results from these recently established trials of big-leaf mahogany. We assessed the relative importance of different environmental and genetic effects in trait variation within and across sites, heritabilities, and the absolute amount of genetic variation for height growth rate, shoot borer attack, tree form, and branching.

Methods

Cooperating Organizations

The germplasm collection in the Yucatan peninsula was part of a larger big-leaf mahogany collection throughout Mexico and Central America by the Center for Tropical Agricultural Research and Education of Costa Rica (CATIE) and the Institute of Terrestrial Ecology of Scotland (ITE). Cooperators for the Yucatan segment of the project included the International Center for Research in Agroforestry (ICRAF), The Instituto Nacional de Investigaciones Forestales, Agricolas, y Pecuarias (INIFAP) of Mexico and the International Institute of Tropical Forestry (IITF) of the USDA Forest Service, as well as community organizations which helped identify seed areas, Xpujil Regional Agrosilvopastoral

Council (CRASX) and the Organization of Ejido Forest Producers for the Zona Maya (OEPFZM), land owners where the collections were made, and the Zoh Laguna ejido and a private land owner who agreed to trial establishment on their properties.

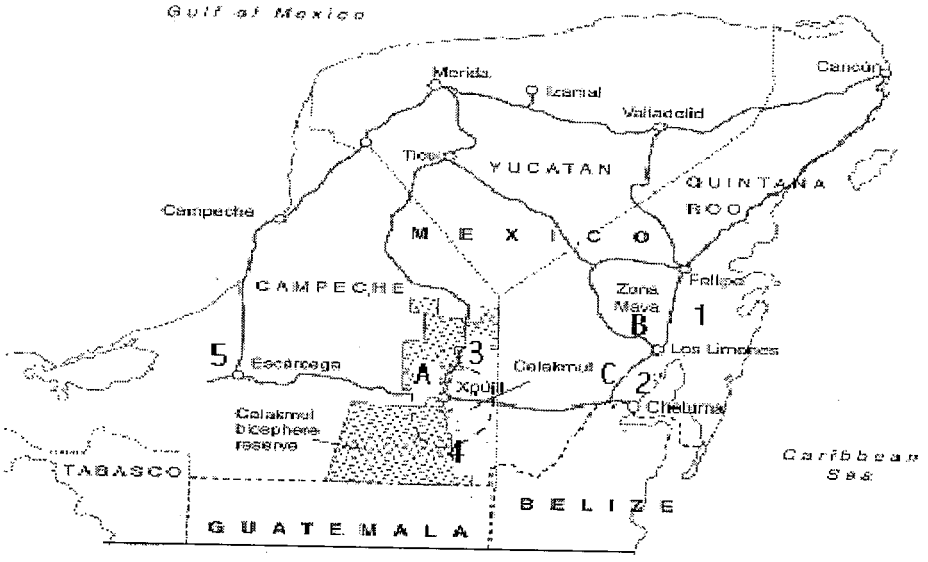


Figure 1. Map of collection and trial area. See Tables 1 and 2 for key.

Seed Collection and Trial Establishment

Open pollinated seeds were collected in February 1997 from about 100 trees. The seed collection was made throughout the natural distribution of mahogany in the Yucatan organized in five provenance areas (Table 1, Fig. 1) which could be distinguished by somewhat different climatic conditions (Table 2). Within provenances, seed trees were at least 100 m apart and mostly in natural forest stands.

Table 1. Provenance descriptions.

ProvenancePrecipitation	Temperature (mm)	Dry months (°C)	No. of families collected	Map key
Zona Maya: Naranjal and Laguna Kana, Quintana Roo	1200	24	3	20
Bacalar, Quintana Roo	1300	25	3	18 2
Nueva Becal, Campeche	1000	24	4	51 3
Carlos Madrazo, Campeche	1600	26	4	4 4
Escarcega, Campeche	1200	24	4	6 5

Seedlings were raised in the INIFAP San Felipe Bacalar Research Station (SFB Station). Germination rates varied from 20 to 80% due to prior handling and storage of seeds. Seedlings were transferred from a germination bed to black plastic bags at about 6 weeks old. Seedlings were 6 and 7 months old when planted in the community of Zoh Laguna and at a farm near Felipe Carillo Puerto in the Zona Maya, and 18 months when planted out at the SFB Station. Sites varied in previous land use. At each site, blocking attempted to take into account the local variation in soil characteristics. Some variation also occurred among sites in management of the trial after planting.

The community trials were planted as randomised complete block design with 5 repetitions and 4 tree row plots. Planting density was 3 x 3 m. For the SFB Station trial, a compact family design in which provenances were grouped within sub-blocks with 5 tree row plots was chosen. Approximately 50 families were planted at each site (Table 2), with 36 families planted at all three locations. Data were collected at time of planting, at 6 months, and at one and two years on tree height (cm; Ht), apical shoot borer attack (Hyps), tree form (on a scale of 1 to 3 for the SFB Station, 1 to 4 for the other sites, 1 being best form), and number of apical branches (Shoot). Growth rate (äHt) was calculated on an annual basis from difference between initial size at planting and final size. Relative growth rate (räHt) was calculated as growth rate divided by size at planting.

Table 2. Number of families planted at each site.

	Planting sites		
Provenance	Zoh Laguna	Zona Maya	SFB station
Map Key	A	B	C
Zona Maya	20	19	19
Bacalar	7	10	5
Nuevo Becal	20	9	26
Carlos Madrazo	4	4	4
Escarcega	2	5	2
TOTAL	53	47	56
Total number of trees	1060	940	1400

Data Analysis

Transformations were selected that provided the best compromise between homogeneity of variance and common use for all sites. Thus the square root transformation was used for äHt, räHt, and Form, and the inverse square root transformation was used for Shoot. Although attack was a binary variable, components of variation have been used on untransformed binary data with satisfactory results (Song and Goddard 1979), so the same approach was used here.

Analysis of variance and variance components analyses were performed across all three sites for the families in common, and separately on all families at each site. The model employed for across sites comparisons (with site effects dropped for within-site analysis) was:

$$Y_{ijklm} = u + S_i + B(S)_{ij} + P_k + F(P)_{kl} + SP_{ik} + SF(P)_{ikl} + B(S)P_{ijk} + B(S)F(P)_{ijkl} + e_{ijklm}.$$

Where Y_{ijklm} is the observed value for an individual, u is the overall mean, S_i is the effect of the ith site, $B(S)_{ij}$ is the effect of the jth block in the ith site, P_k is the effect of the kth provenance, $F(P)_{kl}$ is the effect of the lth family in the kth provenance, $B(S)P_{ij}$ is the interaction of the ith block and the jth provenance, SP_{ik} is the interaction of the ith site and the pth provenance, $SF(P)_{ikl}$ is the interaction of the ith site and the kth family in the lth provenance, $B(S)P_{ijk}$ is the interaction of the jth block of the ith site and the kth provenance, $B(S)F(P)_{ijkl}$ is the interaction of the ith block in the jth site and the kth family in the lth

provenance, e_{ijklm} is the residual of the deviation of the mth tree of the lth family of the kth provenance in the ith block of the jth site.

The significance of different effects was determined via the General Linear Model (SAS 6.12), with all effects considered to be random. Variance components were determined for all effects in the previous model using restricted maximum likelihood (SAS 6.12 PROC VARCOMP). Since provenance level effects were generally insignificant, variance components used for determining heritabilities retained provenance level variation and interactions within the corresponding component at the family level.

Individual (H_I^2) and family (H_F^2) heritabilities across sites were derived using approach and Kiss and Yeh (1988), and of Xie and Ying (1996), with the k values being coefficients of the expected mean squares (Searle 1971):

$$H_I^2 = \sigma_A^2 / (\sigma_w^2 + \sigma_{SF}^2 + \sigma_{BF}^2 + \sigma_F^2)$$
$$H_F^2 = 1/4\sigma_A^2 / ((\sigma_w^2 + k_7\sigma_{SF}^2 + k_8\sigma_{BF}^2) / k_9 + \sigma_F^2)$$

where σ_A^2 is the additive genetic variance and equal to $4\sigma_F^2$, σ_F^2 is due to among family variation, σ_{BF}^2 is due to the block-family interaction, σ_w^2 is the variation within plots, and the k_x coefficients were obtained from the type 3 expected mean squares from Proc GLM (SAS 6.12). Site level effects were excluded from the within-site heritabilities.

The additive genetic coefficient of variation (AGCV) for each trait were assessed using the formula $100(\sigma_A/x)$ where σ_A is the additive genetic standard deviation and x is the population mean to obtain an index of the absolute amount of genetic variation available for selection for a given trait.

Results and discussion

By two years after planting (one year for the SFB Station), the average height of trees ranged from 129 to 151 cm per site. Survival at two years after planting was 84% at Zoh Laguna, 72% at Zona Maya, and 95% after one year at the SFB Station (with replanting). Height growth rates (äHt) at 2 years probably contained residual nursery and mother tree effects.

From the combined analysis of the three sites, the site effects, which included the differences in planting dates, were usually the largest (Table 3a). Most of the variation among sites in height growth rate may have been due to differences in time of planting. The site effect was very small for the relative height growth rate (räHt), for which the effect of tree size at the time of planting had been

essentially removed. Variation in branching, attack, and form among sites were more likely to be due to large-scale spatial environmental differences than planting date. Provenance and family variance components were usually near zero and not significant. The large-scale genetic-by-environment interactions (site-by-provenance and site-by-family effects) were greater than 1% only for relative height growth rate. These interactions would limit the ability to select for family and provenance performance across sites for this trait.

Table 3. Variance components and significance of effects (from analysis of variance), individual and family heritabilities, genetic additive coefficients of variation, and means.

a. Across all sites	Components of variance as % of the total variance							Heritabilities			
Trait	S	B(S)	P	F(P)	S*P	S*F (P)	B*P	B*F(P)	E	H^2_I	H^2_F	AGCV
äHt (24-0)	45.9 ****	4.7 ****	0.0 ns	0.3 ns	0.8 *	0.6 ns	0.8 ns	7.8 ****	39.1	2.4	15.8	0
räHt (24-0)	1.9 ns	5.8 ****	0.0 ns	0.3 ns	1.7 *	4.8 ***	1.2 ns	13.1 ****	71.2	0.0	0.0	7.71
Hyps (24)	47.2 ****	5.5 ****	0.0 ns	0.0 ns	0.4 ns	0.8 ns	0.0 ns	1.1 *	45.0	0.0	0.0	14.89
Shoot (24)	15.1 ***	1.2 *	0.4 ns	0.0 ns	0.0 ns	0.6 ns	0.3 *	5.5 *	76.8	1.2	10.0	18.54

b. Zoh Laguna	Components of variance as % of the total variance					Heritabilities				
Trait	B	P	F	BI*P	B*F (P)	E	H^2_I	H^2_F	AGCV	Mean
äHt (24-0)	1.3 ns	0.0 ns	3.9 **	3.2 ***	3.8 ns	87.8	13.2	35.9	19.04	44.72
räHt (24-0)	0.9 ns	0.0 ns	6.7 ****	3.4 ***	2.9 ns	86.0	25.8	53.7	40.38	1.27
Hyps (24)	13.9 ****	0.7 ns	0.7 ns	0.0 ns	4.0 ns	80.6	5.9	17.6	15.92	0.68
Shoot (24)	2.1 *	0.3 ns	1.7 *	0.4 *	3.4 ns	92.0	8.1	22.9	29.95	1.06
Form	1.8 *	0.8 ns	4.1 *	0.6 ns	8.6 **	84.0	19.0	39.4	16.76	2.58

c. Zona Maya	Components of variance as % of the total variance					Heritabilities				
Trait	B	P	F	BI*P	B*F (P)	E	H^2_I	H^2_F	AGCV	Mean
äHt (24-0)	14.9 ****	1.2 ns	0.5 ns	0.4 ns	16.2 ****	66.8	7.7	19.7	28.22	51.60
räHt (24-0)	13.4 ****	1.3 ns	1.9 ns	0.3 ns	18.5 ****	64.7	14.2	30.6	30.91	1.35
Hyps (24)	1.7 **	0.0	0.0 ns	0.0 ns	7.0 **	91.3	0.0	0.0	0.00	0.07
Shoot (24)	1.4 ns	0.6 ns	0.7 ns	0.0 ns	10.6 *	86.6	5.0	11.9	68.05	0.18
Form	3.1 ns	0.6 ns	0.0 ns	0.1 ns	19.0 ****	77.3	0.0	0.0	0.00	1.69

d. Bacalar	Components of variance as % of the total variance					Heritabilities				
Trait	B	P	F	BI*P	B*F (P)	E	H^2_I	H^2_F	AGCV	Mean
äHt (24-0)	4.4 *	0.8 ns	5.7 ****	3.6 *	10.4 ****	75.2	27.4	53.3	27.42	69.75
räHt (24-0)	3.9 *	5.6 **	9.2 ****	0.5 ns	9.5 ****	71.3	55.3	72.2	47.98	1.09
Hyps (24)	2.4 **	0.0 ns	1.5 ns	0.3 ns	0.2 ns	95.6	6.1	26.8	76.85	0.10
Shoot (24)	1.0 *	1.8 ns	0.2 ns	1.2 ns	7.7 ***	88.1	6.1	21.1	33.21	0.71
Form	5.0 ***	0.0 ns	1.2 ns	0.8 ns	4.9 **	88.2	4.2	16.7	3.43	2.07

KEY:
S Site
B Block
P Provenance
F Family

hI^2=individual level
hF^2=family level
ns p>.05
* .05>=p>.01
**.01>=p>.001
***.001>=p>.0001
****.0001 >=p

For trait acronyms, see Methods.

Numbers in parenthesis refer to months after planting

For all trials together, the large site effect for shoot borer attack at the apex was apparently due to the much higher rate of attack at Zoh Laguna where 68% of the trees were hit (Table 3a). At the other two sites (Tables 3b and 3c), 10% or fewer of the trees were hit. Genetic variation for this trait at each site was generally very small. As in this study, shoot borer attack was found to vary widely among sites in a mahogany provenance trial in Puerto Rico, where genetic variation for attack was also limited (Ward and Lugo 1998, 2001). However, these Yucatan and Puerto Rican trials had only annual assessment of shoot borer attack. A clonal trial is now underway for *S. macrophylla* (Navarro et al. 2001) that includes more intensive assessment of attack and may indicate

more genetic variation in shoot borer susceptibility, as has been found for *Cedrela odorata* (Newton et al 1993).

The mean apical branching was relatively high at the SFB Station (Table 3c) and at Zoh Laguna (Table 3a), but not at Zona Maya (Table 3b). Attack, form, and apical branching also showed higher (although small) family level heritabilities at the SFB Station and at Zoh Laguna than at Zona Maya. Other researchers have considered that *Hypsipyla* attack affects branching therefore tree form (Newton et al. 1993). Although other analyses of this study did not reveal consistent positive genetic and phenotypic correlations between attack and branching (not shown), these correlations may have been positive with more intensive assessment of attack.

Within sites, the provenance effects were generally smaller than family effects, indicating that little population differentiation has occurred in *S. macrophylla* across the Yucatan peninsula (Tables 3a, 3b, and 3c). A similar pattern was found with RAPDs molecular markers in the Yucatan (Gillies et al 1997). The SFB Station site showed greater provenance level variation for relative height growth rate, which may have been due to the subblocking on provenances at this site.

For each site considered separately, the block effect and block by family interaction usually explained the most variation (Tables 3a, 3b, and 3c). These within-site sources of variation may have been due to spatial heterogeneity in soil characteristics. The Yucatan is considered to have highly variable soil (P. Negreros, pers. comm.) which probably depends on microsite drainage patterns.

Blocking effects were usually significant at all three sites. In Zoh Laguna height and diameter growth were worst in block 5 (poor drainage) where the number of apical attacks by *Hypsipyla grandella* was also greatest. This corresponds to the large block variance component for this trait (14%) at Zoh Laguna. In contrast, at Zona Maya and the SFB Station where growth was best (blocks 3 and 2, respectively), shoot borer attack was also greatest.

Heritabilities were higher for height growth traits than for branching, attack, and form, and tended to be somewhat higher at Zoh Laguna and the SFB Station than at Zona Maya. At two years of age, heritabilities for height growth rates at each site were sufficiently large for use in a genetic improvement program. Per site, individual heritabilities for relative height growth rates (.14-.55) were in the range of values seen for height heritabilities in other tree species (mean .28; Cornelius 1994). However, the extremely low across-site heritabilities derived from these early results indicate limited potential for selection of material that will perform well across the conditions of the Yucatan, and that selection would have to be on a per site basis. Even for relative height growth rate, where the effect of site (and planting date) was minimized, the cross-site heritabilities were zero. Within-site heritabilities at this early stage may contain

residual nursery effects, but when the trials are older, index selection at each site based on height growth and form, branching, or attack may be feasible.

The additive genetic coefficients of variation (AGCVs; indicative of absolute genetic variation) observed in this study are within the limits of those seen in other mahogany trials (Narvarro and Hernandez 1998). On the other hand, RAPDs markers have indicated that mahogany in the Yucatan peninsula was deficient in genetic variation (Gillies et al. 1997).

Conclusion

According to these early results, minimal population differentiation appears to have occurred across the Yucatan. In addition, there appears to be limited potential for selection of material that will perform well across the conditions of the Yucatan. This may be due to high environmental heterogeneity. The selection of families and individuals for conversion of these trials to seed orchards may be best made on a per site basis, for use on nearby properties. At two years, the height growth rates showed sufficient heritability to respond to selection, whereas shoot borer attack and the amount of branching did not. These trials were very young and still may have been strongly influenced by small-scale environmental and residual nursery effects. They will be followed for several more years to adequately predict adult tree behaviour.

Acknowledgements

Lionel Cotto and Carlos Navarro of CATIE helped collect the seed. We thank Jonathan Cornelius for comments on the manuscript. This work was largely funded by the Research and Scientific Exchange Division of the Foreign Agriculture Service of the US Department of Agriculture.

References

Cornelius, J. (1994): Heritabilities and additive genetic coefficients of variation in forest trees. Canadian Journal of Forest Research 24: 372-379

Geary, T. F., H. Barres, Ybarra-Coronado, R. (1973): Seed source variation in Puerto Rico and Virgin Islands grown mahoganies. Research Paper ITF-17. U. S. Department of Agriculture, Forest Service, Institute of Tropical Forestry, Rio Piedras, Puerto Rico.

Gillies, A. C. M., Navarro, C., Lowe, A. J., Newton, A. C., Hernández, M., Wilson, J., Cornelius, J. P. (1997): Genetic diversity in Mesoamerican populations of mahogany (*Swietenia macrophylla*) assessed using RAPDs. Heredity 83: 722-732

Kiss, G., Yeh, F. C. (1988): Heritability estimates for height for young interior spruce in British Colombia. Canadian Journal of Forest Research 18: 158-162

Lugo, A. E. (1999): Point-counterpoints on the conservation of big-leaf mahogany. USDA Forest Service. General Technical Report WO-64.

Mayhew, J. E., Newton, A. C. (1998): The silviculture of mahogany. CABI Publishing, New York, New York.

Navarro, C., Hernandez, M. (1998): Evaluación de la diversidad genética de especies tropicales de importancia económica y ecológica en Centro América y el Caribe implicaciones para la conservación, la utilización y utilización sostenible y el manejo. Primer Congreso Latinoamericano IUFRO Valdivia Chile.

Navarro, C., Cornelius, C., Ward, S. E. (2001): Domestication of mahogany and Spanish cedar in Mesoamerica: germplasm collection, genetic selection and development of improved plantation silviculture techniques: second annual project report to the USDA Foreign Agricultural Service: CATIE, Turrialba. Unpublished

Newton, A. C., Cornelius, J. P., Baker, P., Ramnarine, S., Mesén, J. F., Leakey, R. R. B. (1993): The mahogany shoot borer: prospects for control. Forest Ecology and Management 57: 301-328

Patiño, F. (1997): Genetic Resources of *Swietenia* and *Cedrela* in the Neotropics: Proposals for coordinated action. Based on contractual work for FAO by Kageyama, P.Y., Linares, C. , Navarro, C., Patiño. F.: Forest Resources Division, Forestry Department, FAO, Rome.

Searle, S. R. (1971): Linear models. John Wiley & Sons, Inc. New York

Song, S.-I., Goddard, R. E. (1979): Influence of infection percent on improvement of fusiform rust resistance in slash pine. Silvae Genetica 28: 173-180

Ward, S., Lugo, A. (in press): A study of 20 provenances of *Swietenia spp.* in differing environmental conditions in Puerto Rico and the US Virgin Islands. Big Leaf Mahogany: Ecology, Genetic Resources and Management. Springer-Verlag

Ward, S. E., Lugo, A. E. (1998): Efectos del ambiente y la genética en la variación del crecimiento, ataque del barrenador de la yema, sobrevivencia, y tamaño en 20 procedencias de caoba en Puerto Rico. Procedimentos del Primer Congreso Latinoamericano IUFRO, Valdivia, Chile

Xie, C.-Y., Ying, C. C. (1996): Heritabilities, age-age correlations, and early selection in logdepole pine (*Pinus contorta ssp. latifolia*). Silvae Genetica 45: 101-107

ISBN 85-87690-14-0