DIRECT AND INDIRECT MEASURES OF GENE FLOW IN THREE TROPICAL DRY FOREST TREE SPECIES IN SOUTHWESTERN PUERTO RICO.

by

BRIAN KEEGAN DUNPHY

(Under the Direction of James L. Hamrick)

ABSTRACT

The ability to exchange genes across the fragmented dry forest life zone of Puerto Rico's south coast was examined in three species of tropical trees. Two of the species, *Bursera simaruba* and *Hymenaea courbaril*, are native to the island, while the third, *Albizia lebbek*, was introduced from tropical Asia in historic times. Indirect gene flow estimates, which are relatively easy to obtain, were compared against direct gene flow estimates, which although requiring more effort, entail fewer assumptions. One of the critical assumptions of the indirect approach, the presence of migration-drift equilibrium, was violated by all three species. Possible causes of the absence of equilibrium are discussed. Mating system analyses indicated that the three species were essentially completely outcrossed. Results from the direct gene flow study showed that for all three species, a substantial portion of a tree's seed production is the result of gene-flow pollen, with between 42% and 100% of seeds being sired by foreign pollen. Gene flow was highest in *A. lebbek*, and lowest in *H. courbaril*. This contradicts a priori expectations of relative gene flow rates based on pollinator flight capacities, where the bat-pollinated *H. courbaril* was expected to have the highest gene flow rates. Asynchronous flowering and a lower population density are likely explanations for the relatively low gene flow estimates in *H. courbaril*. No relationship was seen between gene flow rates and either distance to nearest neighbor or stand size. In *B. simaruba*, however, germination rates of seeds increased with increasing stand size, with an almost total absence of germination in stands of three trees or less. A study of pollen movement within a stand of 21 adult *H. courbaril* trees showed that pollen usually did not come from nearest neighbors, and that relatively few trees contributed the majority of the pollen leading to successful pollinations. Genetic diversity values were higher in *B. simaruba* and *A. lebbek* than in other species with similar life history characteristics, while they were lower in *H. courbaril*. Possible causes of these patterns are discussed.

INDEX WORDS: gene flow, habitat fragmentation, dry forest, *Albizia lebbek*, *Bursera simaruba*, *Hymenaea courbaril*, Puerto Rico, Caribbean.

DIRECT AND INDIRECT MEASURES OF GENE FLOW IN THREE TROPICAL DRY FOREST TREE SPECIES IN SOUTHWESTERN PUERTO RICO.

by

BRIAN KEEGAN DUNPHY

B.S., University of Illinois-Urbana, 1992

M.S., Michigan State University, 1996

A Dissertation Submitted to the Graduate Faculty of The University of Georgia in Partial Fulfillment of the Requirements for the Degree

DOCTOR OF PHILOSOPHY

ATHENS, GEORGIA

2003

DIRECT AND INDIRECT MEASURES OF GENE FLOW IN THREE TROPICAL DRY FOREST TREE SPECIES IN SOUTHWESTERN PUERTO RICO.

by

BRIAN KEEGAN DUNPHY

Major Professor:	James Hamrick
Committee:	Michael Arnold John Avise Russell Malmberg Chris Peterson

Electronic Version Approved:

Maureen Grasso
Dean of the Graduate School
The University of Georgia
August 2003

DEDICATION

To Stephanie and VB.

ACKNOWLEDGEMENTS

I would like to thank Dr. James Hamrick for giving me the freedom to pursue research of my own interest, while offering valuable advice when necessary. This research would not have been possible without the generous technical and logistical support provided by the staff of the Caribbean National Wildlife Refuge in Cabo Rojo, Puerto Rico, with special thanks to Joe Schwagerl. I would like to thank Mindy Burke for her generous assistance while I first made my way around an allozyme lab. Melanie Smith and Andy Tull did a great job keeping seedlings alive in the greenhouse long enough for me to harvest them for genetic analysis.

Fellow graduate students William Carromero, Jenny Cruse, Eva Gonzalez, Eleanor Pardini, Dorset Trapnell, and David Erickson provided valuable support and advice. Further thanks are owed my committee members, Dr. Michael Arnold, Dr. John Avise, Dr. Russell Malmberg, and Dr. Chris Peterson, for advice and critical comments.

Financial support for this project was provided by the Plant Biology Department at the University of Georgia, the Center for Latin American and Caribbean Studies (CLACs) at the University of Georgia, Sigma Xi, and National Science Foundation Grant DEB-9610275 to Dr. Hamrick.

I would also like to thank Ken Robertson for providing me my first opportunity for fieldwork, and Pete Murphy for providing me my first fieldwork opportunity in Puerto Rico. Finally, I would like to thank my family, especially my parents, Dr. James and Jean Dunphy, for support over the years, and special thanks to my wife, Theresa, for her support and patience in being kept far from family for so many years.

TABLE OF CONTENTS

LIST OF TABLES

Page

LIST OF FIGURES

Page

CHAPTER 1

INTRODUCTION AND LITERATURE REVIEW

Habitat loss is a primary cause of species extinction (Saunders et al. 1991). If a plant species survives the initial clearing of land, its ability to persist in the newly fragmented habitat depends upon its abilities to endure the altered conditions of the landscape. Survival depends both upon the ability to endure altered environmental conditions (e.g., increased "edge" effects), as well as ecological changes affecting reproduction. As habitat is cleared, remaining individuals are often isolated in small patches, well-separated from other groups of conspecifics. Pollinator visits among patches may be discouraged by the increased distances between plants (Groom 2001; Knapp et al. 2001; Fischer and Matthies 1997), or by changes in the intervening habitat (e.g., Vandermeer et al. 2001).

Within smaller populations, reduced size and lowered pollination visitation can lead to increased genetic drift, augmenting the negative effects of inbreeding (Fischer and Matthies 1997; Ledig et al. 1997; Kuang et al. 1999). These factors can reduce population growth and increase the risk of local extinction (Menges and Dolan 1998). Conversely, with gene flow between fragments, the genetic effects of both inbreeding and genetic drift will be reduced.

The ability to successfully deal with fragmentation will depend to a great degree on the response of pollinators to the altered landscapes. Some pollinators are lost quickly in fragmented landscapes (Kremen and Ricketts 2000; Steffan-Dewenter and Tscharntke

1999) while others, such as bats, operate freely in open areas (Law et al. 1999). In many cases, pollinators persist, but visits between patches are decreased (e.g., Aizen and Feinsinger 1994). Smaller, more isolated populations, which because of their small size are even more dependent upon foreign pollen to maintain genetic diversity, may be harder for pollinators to find, resulting in decreased pollen movement among those populations.

Accurate estimates of gene flow are essential to our understanding of how species will endure habitat fragmentation. Genetic markers that are rapid to screen have substantially increased our ability to acquire accurate estimates of gene flow and the breeding structure of populations (Hamrick and Nason 2000). Techniques employing genetic markers to estimate gene flow can be divided into two classes: direct measures and indirect measures. In paternity analysis, an example of a direct measure, the pollen donors for a large number of seeds from a population are determined. The percentage of seeds for which no pollen donor can be found in the population provides an initial, unambiguous, assessment of gene flow. The main drawback to this approach is the time and effort, both in the field and the laboratory, needed to carry out the study. It also only allows gene flow estimates for a single reproductive event. Analyses over several reproductive events would be needed to ensure that the results are not just due to the idiosyncrasies of a particular field season.

These direct approaches, despite their drawbacks, provide a powerful alternative to traditional, indirect estimates of gene flow. These latter methods use genetic markers to describe spatial patterns of genetic variation for a set of populations. Population genetic models, especially those of Wright (1951), then are used to describe the historical

level of gene flow necessary to produce the observed distribution of genetic variation. These techniques are easier and quicker to employ than direct methods, but rely on more assumptions, most importantly that the counterbalancing forces of migration and genetic drift be in equilibrium (Wright 1951; Slatkin 1993). This equilibrium will be evidenced by a pattern of isolation by distance, where the farther two populations are apart, the lower the amount of gene flow between them.

Several features of habitat fragmentation may create problems for the use of indirect measures. Migration-drift equilibrium may be disrupted shortly after fragmentation, as genetic drift within fragments disrupts relationships with neighboring populations (Nason et al. 1997). Possibly more problematic is the fact that a set of populations immediately after fragmentation may demonstrate migration-drift equilibrium that is reflective of gene flow levels prior to fragmentation. If gene flow rates have changed subsequent to fragmentation, it may take many generations for a new equilibrium to be achieved (e.g., Neel and Ellstrand 2001), and until that time, indirect gene flow estimates may yield inaccurate estimates. Direct measures may therefore be the only reliable means to estimate contemporary gene flow in a recently fragmented system.

The implementation of direct measures of gene flow have shown gene flow to be more extensive than previously thought, in both wind- and insect-pollinated trees. Dow and Ashley (1998) found long-distance pollen movement in the wind-pollinated temperate oak, *Quercus macrocarpa*, with 57% of the acorns resulting from pollen originating outside of their stand of 62 adult trees. Schuster and Mitton (2000) found that

6.5% of the pollen received by a stand of limber pine (*Pinus flexilis*), also wind pollinated, came from at least 2 km.

Recent studies utilizing paternity analysis procedures have demonstrated that insect-pollinated tropical trees may be capable of dispersing pollen distances great enough to reduce the threat of genetic isolation. Pollen flow frequently exceeds 25%, with pollen often traveling more than one kilometer (Hamrick and Nason 2000). Chase et al. (1996) used highly variable simple sequence repeats (SSRs) to conduct a paternity-analysis study of gene flow in the hawkmoth-pollinated tropical tree *Pithecellobium elegans*. They showed that pollen moved an average of 142 m, with the greatest movement being 350 m. Approximately 29% of the pollen came from outside the study area. In *Calophyllum longifolium*, 62% of pollen traveled over 200 m (Stacy et al. 1996). In the tropical tree *Tachigali versicolor*, 25% of pollen came from more than 500 m (Loveless et al. 1998). In the extreme, three species of *Ficus* in Panama received over 90% of their pollen from more than 1000 m (Nason and Hamrick 1997).

White et al. (2002) examined gene flow in several forest fragments of the rare tree species, *Swietenia humilis* in Honduras. Highly variable SSRs allowed them to determine that between 24% and 100% of pollen traveled more than 900m. In one case, an isolated tree received 71% of its pollen from over 4.5 km.

This study was conducted in the Subtropical Dry Forest life zone of southwestern Puerto Rico (sensu Holdridge 1967). This life zone has been cleared to a greater extent than the more-publicized rainforest life zone (Janzen 1986). Soils tend to be richer in nutrients. Forests are easier to clear both because of their less-developed stature and because fire can be used easier than in wetter forests. Additionally, human vectors of

disease are less prevalent, and the climate is considered more comfortable for human habitation.

Unlike many Caribbean islands of the which were cleared in the 17th and 18th century, the forests of Puerto Rico did not experience widespread clearing until the 19th century (Wadsworth 1950). Although some land was cleared for sugar cane, the main cause of land clearing for many islands, this did not occur in Puerto Rico until the early 20th century, and the cane fields did not persist for more than a few decades. Rather, the main reason for land clearing was the creation of pastureland. Between 1828 and 1899, the amount of pasture doubled from 634,000 acres to 1,200,000, or nearly 55% of the island (Wadsworth 1950). Vegetation along the south coast was especially impacted. Remnant vegetation is found in scattered patches of forest, strips of vegetation along fencerows and ravines, and isolated trees left in pastures.

In this study, I investigated gene flow estimates in three tree species, using both direct and indirect measures. This is the first study to examine levels of gene flow in Caribbean forests, and one of the few studies examining genetic variation of plants in the region (e.g., Negron-Ortiz and Hickey 1996). We determine whether populations are in migration-drift equilibrium, and therefore amenable to indirect estimates of gene flow. Using direct measures, we determine the proportion of pollinations that involve immigrant pollen, and whether population size and nearest-neighbor distance influence rates of pollen flow.

Two native and one introduced tree species were used in this study. The native trees were *Bursera simaruba* and *Hymenaea courbaril*, and the introduced species was *Albizia lebbek*. All are common to the Dry Forest life zone of Puerto Rico.

Bursera simaruba is a polygamodioecious tree with flowers that are pollinated by small insects. The perfect flowers of *Hymenaea courbaril* are pollinated by bats. The exotic *Albizia lebbek* also has perfect flowers, pollinated by moths, butterflies, and bees.

Based upon expectations of pollinator movements, the greatest gene flow was expected for *H. courbaril*, followed by *A. lebbek* and then *B. simaruba*. In regards to the indirect measures, we expected the native species, *B. simaruba* and *H. courbaril*, to be in migration-drift equilibrium, and consequently amenable to indirect estimates of gene flow. Equilibrium would likely be absent in *A. lebbek*, because of its relatively recent introduction to the island, and its colonizing habit.

References

Aizen MA, Feinsinger P (1994). Forest fragmentation, pollination, and plant reproduction in a Chaco dry forest, Argentine. Ecology 75:330-351.

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

Dow BD, Ashley MV (1998). High levels of gene flow in bur oak revealed by paternity analysis using microsatellites. Journal of Heredity 89: 62-70.

Fischer M, Matthies D (1997). Mating structure and inbreeding and outbreeding depression in the rare plant *Gentianella germanica* (Gentianaceae). American Journal of Botany 84: 1685-1692.

Groom M (2001). Consequences of subpopulation isolation for pollination, herbivory, and population growth in *Clarkia concinna* (Onagraceaea). Biological Conservation 100: 55-63.

Hamrick JL, Nason JD (2000). Gene flow in forest trees. In: Boyle TJB, Young A, Boshier D (eds) *Forest Conservation Genetics: Principles and Practice*, CIFOR and CSIRO: Australia. Pp. 81-90

Holdridge LR (1967). Life zone ecology. Tropical Science Center, San José, Costa Rica. 206 pp.

Janzen DH (1986). *Guanacaste National Park: Tropical ecological and cultural restoration*. Editorial Universidad Estatal a Distancia: San José, Costa Rica.

Knapp EE, Goedde MA, Rice KJ (2001). Pollen-limited reproduction in blue oak: Implications for wind pollination in fragmented populations. Oecologia 128: 48-55.

Kremen C, Ricketts T (2000). Global perspectives on pollination disruptions. Conservation Biology 14: 1226-1228.

Kuang H, Richardson TE, Carson SD, Bongarten BC (1999). Genetic analysis of inbreeding depression in plus tree 850.55 of *Pinus radiata* D. Don. II. Genetics of viability genes. Theoretical and Applied Genetics 99: 140-146.

Law BS, Anderson J, Chidel M (1999). Bat communities in a fragmented forest landscape on the south-west slopes of New South Wales, Australia. Biological Conservation 88: 333-345.

Ledig FT, Jacob-Cervantes V, Hodgskiss PD, Eguiluz-Piedra T (1997). Recent evolution and divergence among populations of a rare Mexican endemic, Chihuahua spruce, following Holocene climatic warming. Evolution 51: 1815-1827.

Loveless MD, Hamrick JL, Foster RB (1998). Population structure and mating system in *Tachigali versicolor*, a monocarpic neotropical tree. Heredity 81: 134-143.

Menges ES, Dolan RW (1998). Demographic viability of populations of *Silene regia* in Midwestern prairies: Relationships with fire management, genetic variation, geographic location, population size and isolation. Journal of Ecology 86: 63-78.

Nason JD, Hamrick JL (1997). Reproductive and genetic consequences of forest fragmentation: two case studies of Neotropical canopy trees. J Hered 88: 264-276.

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

Neel MC, Ellstrand NC (2001). Patterns of allozyme diversity in the threatened plant *Erigeron parishii* (Asteraceae). American Journal of Botany 88: 810-818.

Negron-Ortiz V, Hickey RJ (1996). The genus *Ernodea* (Rubiaceae) in the Caribbean Basin. I. Allozyme variation and mating systems. Systematic Botany 21: 433-443.

Saunders DA, Hobbs RJ, Margules CR (1991). Biological consequences of ecosystem fragmentation: a review. Conservation Biology 5:18-32.

Schuster WSF, Mitton JB (2000). Paternity and gene dispersal in limber pine (*Pinus flexilis* James). Heredity 84: 348-361.

Slatkin M (1993). Isolation by distance in equilibrium and nonequilibrium populations. Evolution 47: 264-279.

Stacy EA, Hamrick JL, Nason JD, Hubbell SP, Foster RB, Condit R (1996). Pollen dispersal in low-density populations of three neotropical tree species. American Naturalist 148: 275-298.

Steffan-Dewenter I, Tscharntke T (1999). Effects of habitat isolation on pollinator communities and seed set. Oecologia 121: 432-440.

Vandermeer J, Hoffman B, Krantz-Ryan S (2001). Effect of habitat fragmentation on gypsy moth (*Lymantria dispar* L.) dispersal: The quality of the matrix. American Midland Naturalist 145: 188-193.

Wadsworth FH (1950). Notes on the climax forests of Puerto Rico and their destruction and conservation prior to 1900. Caribbean Forester 11: 38-47.

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

Wright S (1951). The genetical structure of populations. Annals of Eugenics 15: 323-354.

CHAPTER 2

GENE FLOW INTO FRAGMENTED POPULATIONS OF *BURSERA SIMARUBA* (L.) SARG. IN SOUTHWESTERN PUERTO RICO[1]

[1] Dunphy, B.K. and J.L. Hamrick. To be submitted to *American Journal of Botany*.

Abstract

Gene flow was estimated in the tropical tree, *Bursera simaruba*, using indirect, F_{ST}-based measures, and a direct, paternity-based analysis. Estimates of allozyme genetic diversity for several spatially isolated populations in southwestern Puerto Rico were slightly higher than those reported for species with similar life histories. Indirect measures of gene flow yielded an estimate of 3.57 migrants per generation, but a key assumption that the populations are in migration-drift equilibrium was probably violated in this highly fragmented landscape. Direct estimates of gene flow, which can be used regardless of the demographic history of a species, demonstrated that between 45% and 100% of the effective pollen came from outside of a given population. The populations examined were therefore not genetically isolated, given the distances examined. In the most geographically isolated population, 66% of pollen moved at least one km. Despite the potential for long-distance pollen movement, seed abortion was high, especially in stands with less than four trees. Although the cause of abortion remains unknown, inbreeding depression and self-incompatibility are likely causes. Population size, rather than isolation distance, appears to be the factor limiting reproduction in the populations examined.

Introduction

Habitat fragmentation can substantially alter the demographic processes of a species (Gibbs 2001), potentially increasing the risk of local extinction. As habitat is cleared, the distance between conspecifics often increases, potentially reducing visits by pollinators in entomophilous plant species (Groom 2001; Knapp et al. 2001; Fischer and Matthies 1997). Change in intervening habitat may further discourage pollinator movement (e.g., Vandermeer et al. 2001). Smaller populations may present a smaller "target" for pollinators, thus being more easily passed over either intentionally or by chance (Gigord et al. 1999; Groom 2001). Within smaller populations, reduced population size and lowered pollination visitation can increase genetic drift, which can augment the negative effects of inbreeding (Fischer and Matthies 1997; Ledig et al. 1997; Kuang et al. 1999). These factors can reduce population growth and increase the risk of local extinction (Menges and Dolan 1998). Conversely, with gene flow between fragments, the genetic and demographic effects of both inbreeding and genetic drift will be reduced.

The recent increased availability of genetic markers that are rapid to screen has greatly enhanced our ability to estimate gene flow (Hamrick and Nason 2000). One common technique uses these markers to determine the spatial distribution of genotypes for a set of populations. Classic population genetics models, especially those of Wright (1951), describe the historical level of gene flow necessary to produce the observed distribution. These "indirect" approaches, however, rely on assumptions that may not hold in fragmented environments. The most important assumption is that genetic divergence among populations by genetic drift is balanced by the homogenizing effect of

gene flow (Wright 1951; Slatkin 1993). Even if a continuous set of populations exhibits migration-drift equilibrium, however, the equilibrium could conceivably be disrupted within a short time after a fragmentation event. This would occur as genetic drift within remnant populations disrupted relationships with neighboring populations (Nason et al. 1997).

Additionally, if fragmentation lowers pollen movement among remnant populations, a new equilibrium between migration and drift may be established, but it could be several generations before this level is reached (Figure 2.1)(e.g., Neel and Ellstrand 2001). An estimate made in such a situation, therefore, could greatly overestimate current gene flow rates, and lead to faulty predictions of future genetic changes.

Direct measures provide another means of measuring gene flow. In one approach, paternity analysis, a direct estimate of pollen flow is obtained by determining the pollen donors of a large number of seeds from a population. Although requiring more field and lab work, this approach requires fewer assumptions than indirect methods, and may provide the only reliable estimates of current gene flow under circumstances where the validity of assumptions cannot be ascertained.

Direct measures have shown gene flow to be more extensive than previously thought, in both wind- and insect-pollinated trees. Dow and Ashley (1998) found long-distance pollen movement in the wind-pollinated temperate oak, *Quercus macrocarpa*, with 57% of the acorns resulting from pollen originating outside of the stand of 62 adult trees. Schuster and Mitton (2000) found that 6.5% of the pollen received by a stand of limber pine (*Pinus flexilis*), also wind pollinated, came from at least 2 km.

Recent studies utilizing paternity analysis have demonstrated that insect-pollinated tropical trees may be capable of dispersing pollen distances sufficient to reduce the threat of genetic isolation. Pollen flow frequently exceeds 25%, with pollen often traveling more than a kilometer (Hamrick and Nason 2000). Chase et al. (1996) used highly variable simple sequence repeats (SSRs) to conduct a paternity-analysis study of gene flow in the hawkmoth-pollinated tropical tree, *Pithecellobium elegans*. They showed that pollen moved an average of 142 m, with the greatest movement being 350 m. Approximately 29% of the pollen came from outside the study area. In *Calophyllum longifolium*, 62% of the pollen traveled over 200 m (Stacy et al. 1996). In the tropical tree *Tachigali versicolor*, 25% of the pollen came from more than 500 m (Loveless et al. 1998). In the extreme, three species of *Ficus* in Panama received over 90% of their pollen from more than 1000 m (Nason and Hamrick 1997). White et al. (2002) examined gene flow in the rare tree species, *Swietenia humilis* among several forest fragments in Honduras. Highly variable SSRs allowed them to determine that between 24% and 100% of pollen traveled more than 900m. In one case, an isolated tree received 71% of its pollen from distances greater than 4.5 km.

In this study, we investigate estimates of gene flow in the neotropical tree, *Bursera simaruba*, using both direct and indirect measures. In so doing, we hope to uncover the role gene flow may play in allowing the species to persist in a highly fragmented landscape. The suite of pollinators for *B. simaruba* are similar to those of *S. humilis* (White et al. 2002), leading to the possibility that gene flow in the former may be as extensive as that of the latter. We determine whether populations are in migration-drift equilibrium, and therefore amenable to indirect estimates of gene flow. Using direct

measures, we determine the proportion of pollinations that involve immigrant pollen, and whether population size and nearest-neighbor distance influence rates of pollen flow. In particular, we note populations for which pollen flow is minimal, and consequently most at risk for loss of genetic diversity, and possible extinction.

Materials And Methods

All study populations were in southwestern Puerto Rico, situated near the Sierra Bermeja mountain range (Figure 2.2). The region is classified as subtropical dry forest (sensu Holdridge 1967), receiving less than 1000 mm of rain per year. Much of the land has been cleared for agriculture and cattle pasture. Remnant vegetation is found in scattered patches of forest, strips of vegetation along fencerows and ravines, and isolated trees in fields. Forest clearing in Puerto Rico, largely to create cattle pasture, occurred primarily during the 19th century (Wadsworth 1950). The location of the study site, however, was one of the first to be colonized by the Spanish in the 16th century, and so may have a longer disturbance history.

Bursera simaruba (L.) Sarg. (Burseraceae), known locally as Almácigo, is a polygamodioecious tree found frequently in dry through moist tropical and subtropical forests. It occurs from western Mexico, Florida and the Bahamas south to northern South America. It is found throughout Puerto Rico on soils derived from limestone, and is frequently planted as a fence row tree (Little and Wadsworth 1964). It has small (5-7 mm in diameter), green flowers that are pollinated by small flies, cerambycid beetles and other small insects (Stevens 1983).

The bird-dispersed fruits are preferred only by a small number of bird species, most notably vireos (Greenberg et al. 1995). Seeds ripen for 7-8 months on the tree (Becerra and Venable 1999), with the embryo rapidly filling the seed cavity in the final week (Stevens 1983). No obvious changes in external appearance accompany this change.

Genetic diversity and indirect estimates of gene flow

To obtain indirect gene flow estimates, six populations of 21 to 24 individuals were sampled (Figure 2.2). All populations had minimal forest cover. They were embedded in a matrix of mostly graminoid vegetation, with some shrubs and small trees. Trees averaged 5-8 m in height, and 10-20 cm in diameter, with a few as large as 30-40 cm in diameter. Aerial photographs from the Refuge population (site #5, Figure 2.2) showed the area to be free of trees as late as 1960, indicating that the trees, which are typical in size for trees at other sites, can be no more than 40 years old.

The minimum distance separating any two populations was 2.5 km, and the maximum was 25.2 km. At least 20 cm^2 of leaf tissue was collected from each tree. Leaves were shipped on ice to the University of Georgia within 48 hours of collection. In the lab, leaves were crushed in a potassium phosphate extraction buffer (Mitton et al. 1979) and the resulting extracts were stored on filter paper wicks at -70° C.

Horizontal starch gel electrophoresis was used to assay allozyme diversity. Three gel buffer, gel-electrode combinations and eleven enzyme stains resolved 15 loci on 11.0% starch gels. Gel and electrode buffer recipes followed Soltis et al. (1983). Enzymes stained (and loci resolved) on System 4 (a Tris/citrate gel and tray buffer) were

isocitrate dehydrogenase (Idh-1), shikimate dehydrogenase (Skdh-1), and UTP-glucose-1-phosphate (Ugpp-1). System 11 (a histidine gel and citric acid tray buffer) was stained for malate dehydrogenase (Mdh-1, Mdh-2, and Mdh-3) and 6-phosphogluconate dehydrogenase (6Pgd-1). System 6 (a Tris-citric acid gel buffer and sodium hydroxide and boric acid tray buffer) resolved diaphorase (Dia-1), glutamate dehydrogenase (Gdh-1), menadione reductase (Mnr-1), phosphoglucoisomerase (Pgi-1 and Pgi-2), triose-phosphate isomerase (Tpi-1 and Tpi-2) and peroxidase (Per-1). Stain recipes were modified from Soltis et al. (1983) except for diaphorase (Cheliak and Pitel 1984) and UTP-glucose-1-phosphate (Manchenko 1994). For enzymes with more than one locus, isozymes were numbered sequentially, with the lowest number assigned to the most anodal.

Standard measures of genetic variation were calculated for each population (subscript p) and at the species level (subscript s) by pooling all the individuals. Genetic diversity parameters estimate percentage of polymorphic loci (P), the mean number of alleles per locus (A) and per polymorphic locus (AP), the effective number of alleles per locus (A_e), observed heterozygosity (H_o) and expected heterozygosity (H_e).

Indirect gene flow estimates, in terms of the effective number of immigrants per generation (N_em), were made with the following equation (Wright 1951):

$$N_em = (1-F_{ST})/4F_{ST}$$

where N_e is the effective number of individuals in the population, m is the immigration rate and F_{ST} is the proportion of total genetic diversity due to differences among populations (Wright 1951). A key assumption for using this approach is that the populations are in migration-drift equilibrium. This assumption was tested by looking for

a statistically significant positive relationship between pair-wise estimates of $F_{ST}/(1-F_{ST})$ and geographic distance separating the pair (Rousset 1997).

Direct measures of gene flow

Eight small spatially isolated stands of trees were chosen from the same geographic location as the populations used for the indirect analysis (Figure 2.2). The stands varied in degree of isolation from 80 m to more than 1 km (Table 2.1). A total of 3,997 seeds were collected from the eight stands. Due to the long ripening time of the fruit on the trees with minimal external signs of maturation, seeds were collected from the ground beneath trees. Canopies of trees rarely overlapped. In cases where they did, seeds from the overlap zone were avoided. Seeds were brought to the Plant Biology Greenhouses at the University of Georgia in Athens, GA. Seed scarification techniques did not improve germination, so seeds were buried without treatment under 1.5-2 cm of soil. Most seeds that germinated did so within a week, with some seedlings taking up to three weeks to break the soil surface. Once true leaves appeared, roughly 20 cm^2 of leaf tissue was collected and processed for allozyme analysis as described above.

Multi-locus estimates of outcrossing were computed with MLTR (Ritland 2002). Paternity analysis was used to estimate gene flow levels. In short, each seedling's genotype was compared with those of the potential sires in the seedling's population. An initial estimate of gene flow was made by determining the percentage of seedlings that could not have been sired by any of the potential fathers. Allozyme trials yielded an exclusion probability of 0.67 for *B. simaruba*, indicating that for a given seed, 67% of potential sires will be correctly identified as non-sires (Chakroborty et al. 1988). The low

level of potential pollen donors within each site (3-9) produced little paternal ambiguity with this exclusion probability. The possibility remains, however, that a father from outside the population could produce a pollen gamete indistinguishable from that created by a sire from within the population. In such a case, the resulting seedling would be falsely identified as having originated within the population, leading to an underestimation of gene flow. The sum of these events comprise "cryptic" gene flow, and their estimation is essential to obtain unbiased estimates of total gene flow rates.

To estimate total gene flow, we used the program GFLOW (available from B.K.D.), which employs a slight modification of the technique developed by Devlin and Ellstrand (1990). This technique uses maximum-likelihood to determine the total (apparent + cryptic) gene flow rate that would be most likely to yield the observed apparent gene flow rate (Devlin and Ellstrand 1990). A simulation was run over all possible total gene flow levels, with the minimum value being set by the apparent gene flow rate, and the highest being 100%. For each gene flow level (in increments of 1%), five thousand populations were created. These populations had the same number of adult trees, adult genotypes, and numbers of seedlings for each maternal tree as the actual population. Seedling genotypes were generated using a "pollen gamete" from either a randomly chosen adult from the population (non-gene flow), or from the "pollen pool" (gene flow). The pollen pool represents the theoretical range of likely pollen genotypes that could be produced by trees in the vicinity of the population under study. For this study, genotype frequencies in the pollen pool were simply the average frequency of genotypes from the six populations sampled for the indirect gene flow study.

Once generated, the apparent gene flow rate was calculated for each of the 5,000 populations. The number of times an apparent gene flow rate of a simulated population matched that of the actual population was tabulated. The total gene flow rate was then incremented by 1%, and another 5,000 populations were generated.

The total gene flow rate with the greatest number of matches with the actual apparent gene flow rates was the most likely total gene flow rate. The overall gene flow rate for the population was calculated as the weighted average of gene flow for all the trees from the population.

Results

Genetic diversity statistics and population structure

Of the 15 loci surveyed, 11 (73.3%) were polymorphic in at least one of the populations sampled (Table 2.2). Four loci were monomorphic across all populations. At the species level, there were 2.73 alleles per polymorphic locus (AP_s), the effective number of alleles (A_{es}) was 1.50, and the expected heterozygosity (H_{es}) was 0.244.

On average, 60.0% of the loci were polymorphic for the six populations, with the highest percentage, 66.7%, shared by populations at either extreme, New_ac and Pargue. The lowest percentage, 55.3%, was shared by R303B and Refuge. There was an average of 1.87 alleles per locus (range 1.80-2.00), and 2.45 alleles at polymorphic loci (range 2.30-2.56). Genetic diversity (H_{ep}) ranged from 0.193 (Refuge) to 0.245 (Antenna), with an average of 0.222 (Table 2.2). No relationship existed between any of the diversity measures and distance to nearest neighboring population.

Patterns of polymorphism (or lack thereof) at four loci yielded hints of geographic patterning. The New_ac population, the westernmost population, was polymorphic for 6PGD1, which was monomorphic in the remaining populations. Conversely it was monomorphic for IDH, which was polymorphic in all the other populations except its closest neighbor, Refuge. Similarly, Pargue, the easternmost population, was polymorphic for DIA, which was monomorphic for all other populations except its closest neighbor, Tinaja. SKDH was polymorphic for all populations except R303B and Tinaja, which were the two populations closest to each other.

The percent of the variation distributed among populations (G_{ST}) was 6.6% (Table 2.3). The highest G_{ST} (0.114) was found for UGPP1, which also had a relatively low F_{IT} (0.026) due to an excess of heterozygotes within populations (F_{IS}=-0.100). The lowest G_{ST} was for DIA, which also had a slight excess of heterozygotes (F_{IS}=-0.036), and substantially lower total genetic diversity (H_T) than the average (0.020 vs. 0.332). The locus with the second lowest G_{ST}, 6PGD1, also had a low total genetic diversity (0.014). In fact, there was a significant positive correlation between total genetic diversity (H_T) and G_{ST} (R^2=0.388, P<0.01).

A deficit of heterozygotes, indicated by positive F_{IS} values (Table 2.3), was found for five loci, while six loci demonstrated an excess of heterozygotes. The mean value across loci (-0.024) is not significantly different from zero (X^2=14.35; P<0.10), however, indicating a match to what would be expected under Hardy-Weinberg expectations. This is also demonstrated by the close match of observed heterozygosity (0.220) to expected (0.222). With all six populations considered together, there were fewer heterozygotes than would be expected from Hardy-Weinberg expectations (F_{IT}=0.042), due primarily to

a Wahlund effect (G_{ST}=0.066). A X^2 analysis for allele frequency heterogeneity among populations revealed significant differences among populations for eight of 11 polymorphic loci (all except DIA, MNR1, and 6PGD1).

Three populations had a fixation index for a single locus that was significantly different from zero: MNR1 for Refuge ($p<0.05$), IDH for New_ac ($p<0.001$), and MDH1 for R303B ($p<0.01$). This number is close to the 3.3 significant differences expected by chance alone for the 66 fixation indices calculated. Therefore, there is probably no biological meaning to the three significant differences.

Indirect measures of gene flow

The five highest pairwise G_{ST} values were between the Refuge population and all other populations (range 0.122-0.211). There appeared to be a relationship between genetic and geographic distance, once the Refuge population was removed. The next highest pairwise G_{ST} values were between New_ac and Pargue (0.063), which also are the furthest apart. Next was R303B and Pargue (0.059), which were the sixth furthest apart, and then came Anten and Pargue (0.038), the third furthest apart. The lowest pairwise G_{ST} was between Antenna and R303B (0.009), which had the second lowest geographic separation.

Despite the apparent relationship between genetic and geographic distance, there was only a very slightly positive, yet not statistically significant, relationship between pairwise $F_{ST}/(1- F_{ST})$ and interpopulation distances (Figure 2.3), as would be expected under migration-drift equilibrium (r=0.006, P=0.349). Removing the Refuge population made the relationship slightly more positive, and increased the total variation in genetic

distance explained from 0.2% to 11.5%, yet the relationship was still not statistically significant (r=0.010, P=0.238).

Direct measures of gene flow

The multi-locus estimate of outcrossing (t_m) for the species was 0.985 (s.e.=0.042). The single-locus estimate was 0.904, leading to an estimate of biparental inbreeding (t_m-t_s) of 8.1%. Seed germination rates varied from 0% to 40.8% (Table 2.1). Due to low germination rates, only three populations had enough seedlings to allow paternity analyses of individual maternal trees. In two populations, which had 17 (DC) and 19 seeds (JFH), total gene flow rates were calculated at the population level only.

Estimated total gene flow rates were 66% (R307), 69% (JFL), and 82% (JFE) (Figure 2.4). In the two populations with less than 20 seeds, estimated gene flow levels were 100% (JFH) and 47% (DC). Apparent gene flow rates varied between 50-60% of the total gene flow rate, with the exception of Tree 5 from R307, where it was 70.4% of total gene flow rate (Table 2.4); and trees from JFL, which had values from 26% to 39%.

Within stands of trees, the greatest variation was seen in R307, with total gene flow estimates ranging from 45% to 98% for the five maternal trees (Table 2.4). Similar rates were found in JFL, with total gene flow rates of 53%, 69% and 100%. Rates were slightly higher for JFE, 64% and 89%.

When the DC stand was removed, there was a significant negative relationship between total gene flow rate and number of adult trees ($y= -0.111x+1.646$, $R^2=0.937$, $p<0.05$). Including the DC population rendered this relationship non-significant. The relationship with distance to nearest conspecific tree ($y= -0.0004x+1.044$, $R^2=0.861$,

p>0.05) and number of alleles in common (y= -0.617x+1.925, R^2=0.434, p>0.10) were also negative, but not statistically significant.

Discussion

Diversity statistics and population structure

Diversity statistics for *B. simaruba* were comparable to or exceeded those reported for other long-lived, perennial, outcrossing species by Hamrick and Godt (1996). The percent of polymorphic loci was slightly higher (73.3% versus 65.5%), while gene diversity values (H_e) was substantially higher (0.244 versus 0.180). The proportion of variation among populations (G_{ST}) was low for *B. simaruba* (0.066 versus 0.094), a value even lower than that reported for long-lived perennial wind-dispersed species (0.086).

The high pollen movement found in this study, combined with almost total outcrossing, is likely responsible for the low G_{ST} and relatively high P and H_e values found within populations. A possibility to consider for most populations, however, is that the populations under study may not have undergone enough generations of drift since fragmentation to appreciably decrease genetic diversity, considering that fragmentation of Puerto Rican forests occurred mainly in the 19th century (Wadsworth 1950). Low G_{ST} values may also be due to the chosen spatial scale of the study. Pollen, and possibly seeds, may move readily over the distances involved (ca. 25 km). A collection of populations from over a broader geographic range may be needed to make valid comparisons of G_{ST} with other species.

The role of seed dispersal in introducing variability is illustrated by the Refuge population, which aerial photos show to have been established after 1960. High levels of

H_e and a low G_{ST} indicate that these sites were established by seeds from several individuals and/or populations (immigrant pool model, Slatkin 1977) rather than from only a few source individuals (propagule pool model).

Indirect estimates of gene flow

The absence of migration-drift equilibrium limits the use of indirect, F_{ST}-based estimates of gene flow. Disregarding the migration-drift assumption could lead to inaccurate estimates of gene flow, namely Nm(W)=3.57. Since Wright (1931) has demonstrated that gene flow levels of less than one migrant per generation can allow genetic drift to increase genetic distance among populations, erroneous predictions of further genetic changes could be drawn if the true gene flow rate was lower. Regardless of the actual Nm values, the relatively low levels of differentiation among populations indicate that historical levels of gene flow must have been relatively high. Low levels of genetic divergence (average G_{ST}=0.055) were also found for 16 common woody species on the 6 km x 6 km Barro Colorado Island in Panama (Hamrick and Loveless 1989). Direct evidence that pollen moved over 500 m in one species, *Tachigali versicolor*, and over 200 m for several other species, are consistent with high gene flow causing low G_{ST} values.

There were some hints of a pattern in geographic variation, such as the proximity of populations monomorphic for loci that were polymorphic for most other populations (or vice-versa), and a slightly positive relationship between genetic and geographic distances. If the landscape is still responding to habitat fragmentation, these relationships may strengthen over time. Rousset (1997) noted, however, that populations separated by

small geographic distances generally will not follow expectations of the model used to determine isolation by distance. For *B. simaruba*, then, migration-drift equilibrium may only be detectable over larger distances. This further suggests that gene flow is high over the distances examined, a notion supported by the results of the paternity analyses.

Direct measures of gene flow

Stands that produced the most fruit yielded seeds with the highest germination rates. For those stands with lower fruit production, it took much longer to find seeds on the ground, and the vast majority of the seeds were hollow. There were aborted seeds in larger stands of trees (Figure 2.5), as well as the presence of seeds with full-grown embryos that did not germinate. Since embryos fill the seeds only in the last week of an 8-9 month development period, these seeds must have been alive at least that long.

The high pollen flow levels observed suggest that even stands with as few as 5-9 trees receive immigrant pollen, even when the nearest *B. simaruba* tree is further than 1 km (R307). Similar pollen flow levels were reported from Barro Colorado Island in Panama (Nason and Hamrick 1997). Three species of *Ficus* had pollen flow levels greater than 90% and small populations of *Spondias mombin* on islands in Lake Gatun had pollen flow rates of 60% to 100%.

The slightly negative relationship between pollen flow rates and distance to nearest neighboring population was not statistically significant. The lack of isolation by distance in the paternity analysis coincides with what was found in the indirect gene flow analysis. Kaufman et al. (1998) found a similar result in *Cecropia obtusifolia* from

Mexico, with isolation by distance absent over shorter distances, but present over greater distances.

The size of a stand of trees, although not strongly correlated with total pollen flow, did influence seed germination rates. Of 1,834 seeds collected from four sites with three trees each, only 17 (0.9%) germinated. The three stands with 7-9 trees had germination rates ranging from 20-41%. Similarly, Nason and Hamrick (1997) found significant reductions in germination in stands of the tropical tree *Spondias mombin* on small islands in Lake Gatun, Panama as compared to larger fragment and continuous-forest populations. In *B. simaruba* and *S. mombin*, both self-incompatible species, these findings suggest that larger stands may receive more foreign pollen, thus losing fewer seed from self-incompatibility, and experiencing less inbreeding. This may be analogous to what has been reported in other insect-pollinated (Platt et al. 1974; Willson and Rathcke 1974), bat-pollinated (Heithaus et al. 1982) and bird-pollinated (Carpenter 1976) plant species, where an increase in flower number led to increased pollination success.

The highest percentage of within-population pollinations occurred in R307, with 34% of its progeny being sired from within, as opposed to 10% and 18% of the progeny for the JFL and JFE populations. As a percentage of total seed production, R307 had the second highest number of externally-sired seeds (Figure 2.5), so the internally-sired progeny appear to be in addition to those sired from outside. This would suggest that the R307 stand, the largest of the eight, was more tolerant of internal breeding, rather than limited in the amount of outside pollen that it received.

Since so many seeds were aborted, the estimated total pollen flow rate would yield a very biased estimate of pollen flow if it were used for that purpose.

Consequently, the values given should be considered "effective pollen flow" rates. This concept was also illustrated in a direct analysis of pollen flow in fragmented stands of a rare species of Mahogany (*Swietenia humilis*) from Honduras. Stands of varying size showed an increase in foreign pollen received as population size decreased (White et al. 2002). Local pollen was presumably still being received in smaller stands, but a self-incompatiblity system filtered much of this pollen out before fertilization. In *B. simaruba*, seed abortion likely played a similar role in increasing the representation of foreign sires among seedlings.

White et al. (2002) suggested that self-incompatible species may be more tolerant of habitat fragmentation than mixed-mating species, since the latter could have an increase in selfed seeds as population sizes decrease and become more geographically isolated. The current study on *B. simaruba* suggests in some cases, however, that this benefit has to be weighed against the potentially high loss of reproductive output due to aborted seeds.

The almost total absence of selfing (t_m=0.985) and the increase in seed abortion in the smaller stands of trees suggests that *B. simaruba* may have a post-fertilization self-incompatibility mechanism. In a hand-pollination experiment of the putatively self-incompatible tree, *Dombeya acutangula* ssp. *acutangula*, an endemic tree from the island of La Reunion in the Indian Ocean, Gigord et al. (1998) found that seed set from forced outcrossing increased the farther two trees were apart. The authors attributed this to a self-incompatibility system rather than inbreeding depression because of the dramatic variation in seed set among trees. The substantial variation in germination seen in the present study may present similar evidence, especially since the genetic relatedness of

trees in these small stands is high (an average of 1.69 alleles in common per locus versus 1.53 alleles in common per locus for populations with at least 21 individuals).

Seed set increased with hand pollination in three of eight species of *Inga* in the lower montane wet forest of Monteverde, Costa Rica (Koptur 1984). Despite large inflorescences and an abundance of pollinator visits, seed set in *Inga* species tends to be low. Koptur (1984) found a gametophytic incompatibility system that may explain the low seed set, and might also explain why the successful pollinations, in terms of numbers of fruit set, tend to come from pollen that was collected from at least a kilometer away.

Hufford and Hamrick (2003) provide an alternative explanation for the high seed abortion rates. When pollen is not limiting, trees often produce more seeds than can be supported. Subsequent abortions thin the numbers to a level that maternal resources can support. Although large numbers of fruits are regularly produced (B. Dunphy, pers. obs.), any thinning that occurs probably is more strongly based on self-incompatiblity or inbreeding depression than resource limitation. This is best illustrated by the small stands of trees, where between 98.4% and 99.4% of seeds were aborted. It is unlikely that these trees could not support more seeds.

In several stands of trees, there were non-germinating seeds with mature embryos. Hufford and Hamrick (2003) found evidence of substantial inbreeding depression in offspring of the tropical tree, *Platypodium elegans.* Selection against selfed seedlings occurred mostly at the seed to seedling transition. A similar timing of selection against selfed offspring in *B. simaruba* could explain the fate of these non-germinating seeds. Alternatively, as is prevalent in dry tropical forest species, where mechanical or acid scarification is necessary to overcome seed coat dormancy (Khuruna and Singh 2001),

these seeds may have a dormancy mechanism preventing germination. It is also possible that they are mature seeds that simply died while sitting on the ground.

To compare indirect and direct measures of gene flow, an estimate of effective population size (N_e) was combined with a paternity-derived estimate of migration rate (m) to yield N_em, which was compared to the Nm derived through the indirect, F_{ST}-based approach. The migration rate (m) was estimated as half of the total gene flow rate estimated via paternity analysis, since the paternal contribution represents half of a seedling's genotype, and the maternal contribution (non-migrating in the current study, since seeds were taken from known maternal trees) the other half. Effective population size was calculated as $N_e=1/\Sigma((c_i+p_i)/2)^2$, where c_i and p_i are the relative female and male reproductive contribution of the *i*th individual, respectively (Crow and Denniston 1988). Relative fruit production yielded estimates of relative female reproductive success. Relative male reproductive success was calculated by assigning parentage to seedlings using CERVUS 2.0 (Marshall et al. 1998), and determining the relative number of pollen contributions by each male.

Interestingly, two of the N_em values, 2.97 (R307) and 3.00 (JFH) were only within 0.5 of 3.57, the F_{ST}-based value of Nm. The values of N_em for JFL (2.40) and JFE (2.05) were slightly lower. The N_em for DC was the lowest (0.71), although having just three trees limited its maximum possible N_em to 1.41. Wright (1931) demonstrated that when Nm values are between 1 and 4, both gene flow and genetic drift can shape the genetic composition of the populations. With the exception of DC, all of the cases above, along with the F_{ST}-based measure, place the N_em values within this range. Two important questions that remain are whether seed movement would significantly raise

N_em, and whether direct estimates of N_em for larger populations would be similar to those of the smaller stands.

Ecological implications

Tropical trees tend to display a clumped dispersion pattern, often with substantial distances separating neighboring groups of conspecific trees (Hubbell and Foster 1983). If trees in a clump are related, and if they possess a self-incompatibility mechanism, as do a large number of tropical species (Bawa 1974), then selection should be strong for floral strategies that support the energetic needs of long-distance pollinators. Koptur (1984) suggested that *Inga* species in wet montane forests of Costa Rica address this need to supply pollinators by producing a large number of relatively unspecialized flowers over an extended blooming season. These flowers are visited by a wide array of pollinators. Most of the pollen likely comes from neighboring trees or the maternal individual itself since only a small percentage of flowers actually set fruit. Enough long-distance pollinations probably occur, however, to justify the large expenditure of energy in flower production.

This may be the reproductive strategy employed by *B. simaruba*. The virtual absence of effective reproduction in very small stands of *B. simaruba*, however, suggests that these stands may present too small of a target for pollinators from other populations (Gigord et al. 1999; Groom 2001). Once pollinators arrive, they may spend too much time within the stand, distributing essentially non-viable pollen among related trees. Despite the negative effects on very small stands of trees, the promotion of long-distance pollen movement by generalist pollinators may help explain why *B. simaruba* tolerates

habitat fragmentation. White et al. (2002) suggest that obligate outcrossers will endure fragmentation better than species with a mixed-mating strategy because of this forced reliance on pollen from outside sources.

The low seed production in smaller populations (i.e., less than five trees) may pose an ecological problem to the survival of those populations. Dry forests present a particularly harsh environment for seedling survival (Ray and Brown 1995). If microsites for seedling survival are scarce, then it may be difficult for the relatively small number of seeds to find a suitable microsite and none may survive. This would be a threat not just to remnant populations in a fragmented landscape, but could also greatly slow the spread of this bird-dispersed species. If a few related seeds establish a population in a new location, there could be reproductive problems if this new population, with just a few trees, is exposed to pollen primarily from related trees. The abundance of *Bursera simaruba* across its range suggests that this problem is not restricting the distribution of the species. One possible explanation could involve the mechanism of facultative apomixis, which has been reported from other *Bursera* species (Becerra and Venable 1999). Under this scenario, trees capable of apomixis would have a selective advantage in small populations with restricted pollen flow, and would increase in number. As the population grows, pollinators become more plentiful, leading to a higher frequency of foreign pollen, which in turn yields more outbred, viable seeds. These recombinant individuals would introduce variability into the newly established population, and within-population pollinations would eventually be possible.

In summary, *Bursera simaruba* does not appear to be spatially isolated, at least for stands of more than four trees. High genetic diversity values in larger populations,

and relatively low genetic divergence among populations, supports the notion that gene flow occurs over substantial distances. Very small groups of trees (less than five individuals), however, may be at risk. A violation of the migration-drift equilibrium precluded the use of indirect measures of gene flow over the spatial scale used in this study. This may be due to extensive pollen movement, which the direct paternity-based estimates of gene flow support. A larger spatial scale may be needed to use indirect measures, or it may be that in a fragmented system in which patches of trees are of relatively recent origin, their use is not feasible. Paternity analysis provides a robust estimate of gene flow that relies on fewer assumptions and thus, is more suitable for gene flow estimation in a fragmented system.

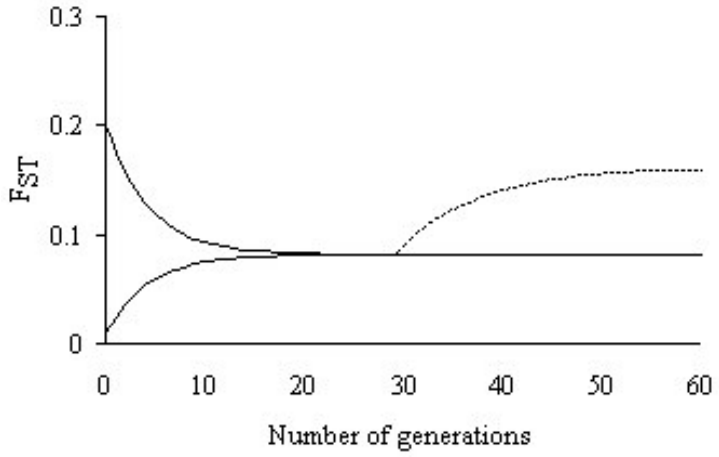


Figure 2.1. Theoretical change in F_{ST} over time in a set of populations. Initially, F_{ST} will approach the same equilibrium level, no matter the initial F_{ST}. This equilibrium will be set by gene flow between populations and genetic drift within populations. If at a certain time (generation 30 in this example), a fragmentation event occurs which lowers pollinator visits, the equilibrium F_{ST} may increase (dotted line). These lines were generated by iterating the following equation (Wright 1951):

$$F_t = \left(\frac{1}{2N}\right)(1-m)^2 + \left(1-\frac{1}{2N}\right)(1-m)^2 F_{t-1}$$

where N is population size, m is migration rate, F is shorthand for F_{ST}, and t is time in generations. In the present example, N=24, and initially m=0.1. After 30 generations, m was cut in half to 0.05 to generate the dotted line.

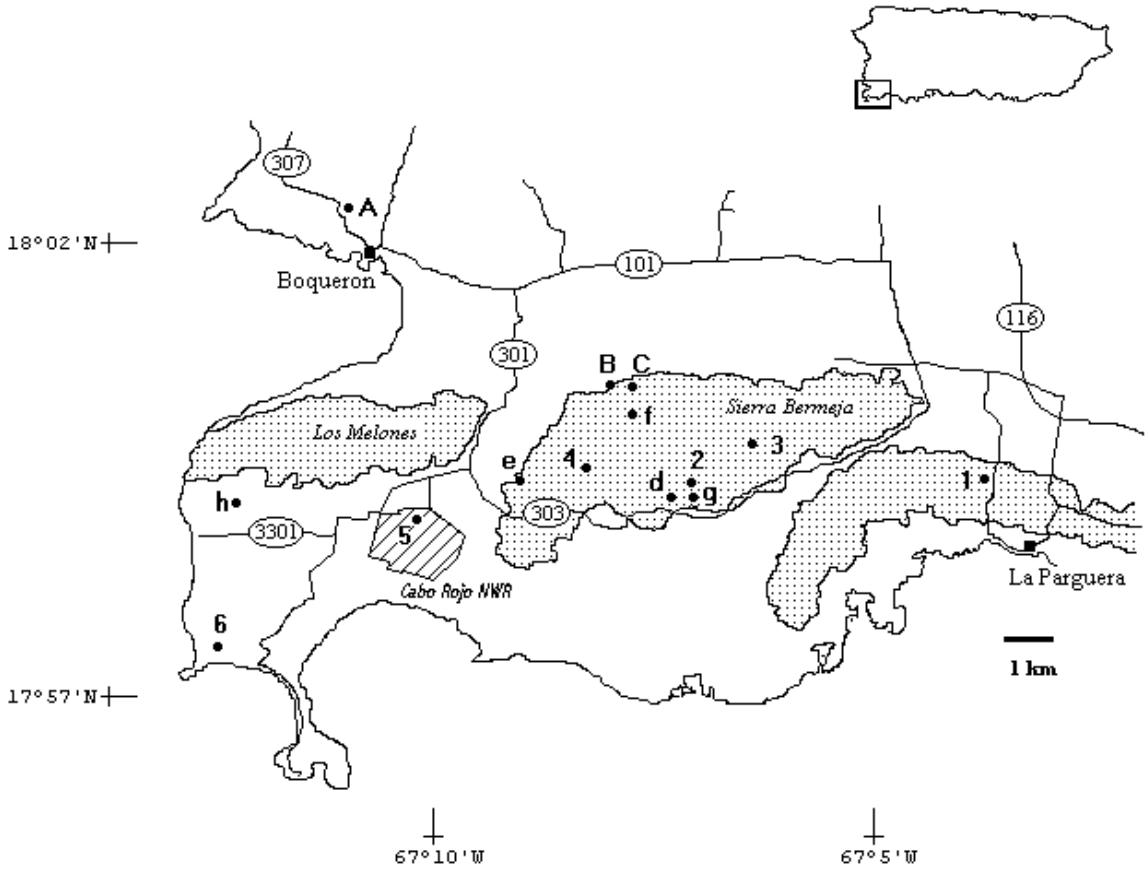


Figure 2.2. Location of study populations of *Bursera simaruba* in southwestern Puerto Rico. Populations of adult trees used in indirect estimates of gene flow are numbered 1-6. The eight stands of trees used for direct measures of gene flow are lettered A-h. Upper-case letters indicate stands where at least 139 seedlings were analyzed. Stippled regions are low mountains (maximum height of 200-225 meters). Diagonal lines show the location of the Cabo Rojo National Wildlife Refuge.

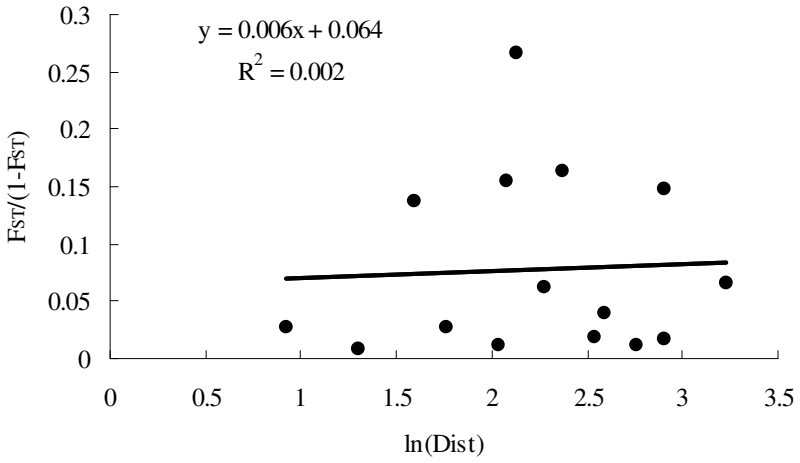
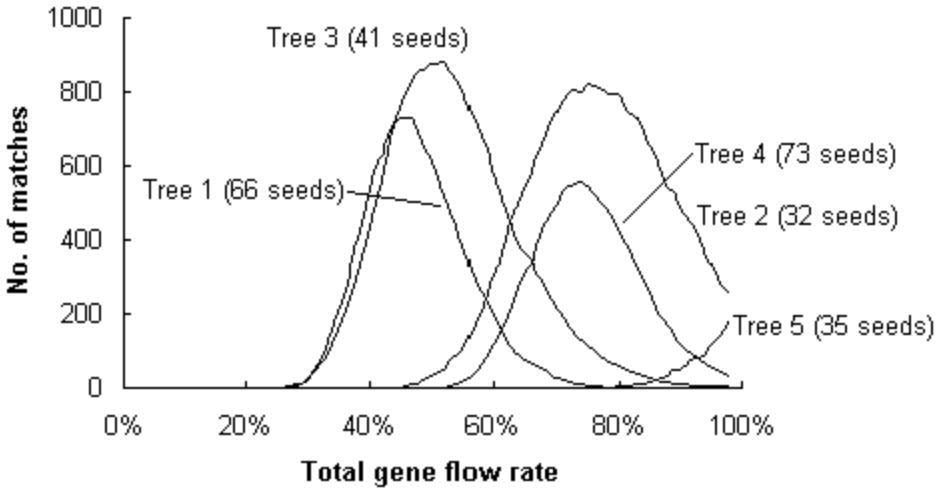
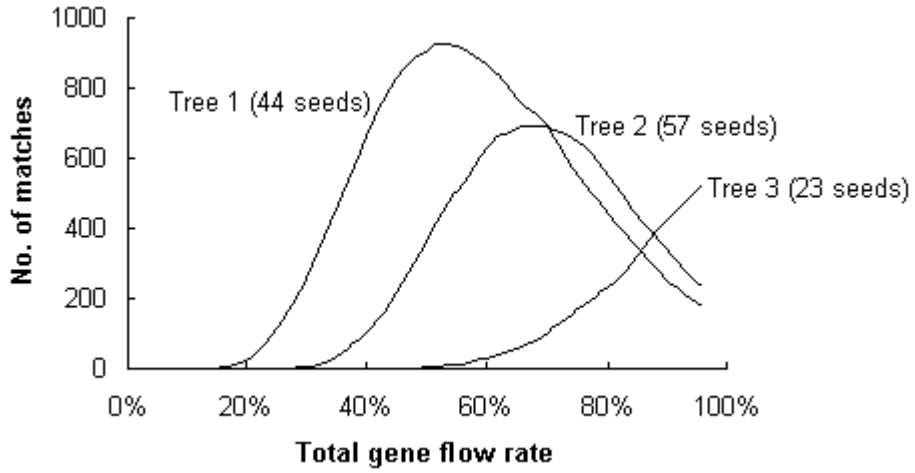


Figure 2.3. Examination of isolation by distance in all possible pairs of six *Bursera simaruba* populations from southwestern Puerto Rico. A significantly positive slope would have indicated migration-drift equilibrium.

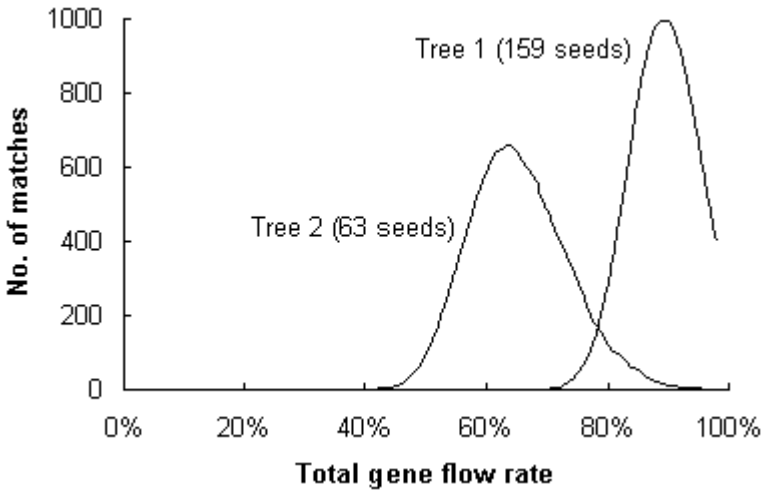


(a) R307

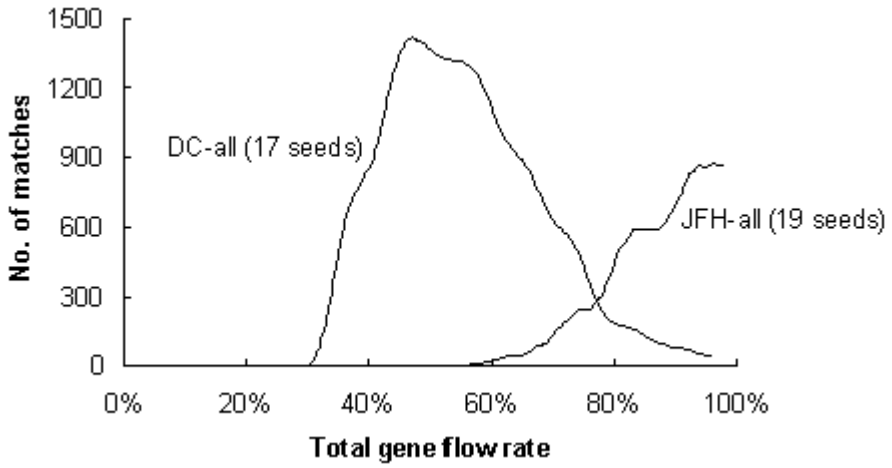


(b) JFL

Figure 2.4. Number of matches between a maternal tree's (or stand's) apparent gene flow rate and the apparent gene flow rates estimated in 5,000 simulated populations experiencing a given total gene flow rate (X axis). The maximum for each line represents the maximum-likelihood estimate of total gene flow for the corresponding tree or stand. Lines were smoothed with a moving-average function.



(c) JFE



(d) Smaller pops

Figure 2.4, cont.

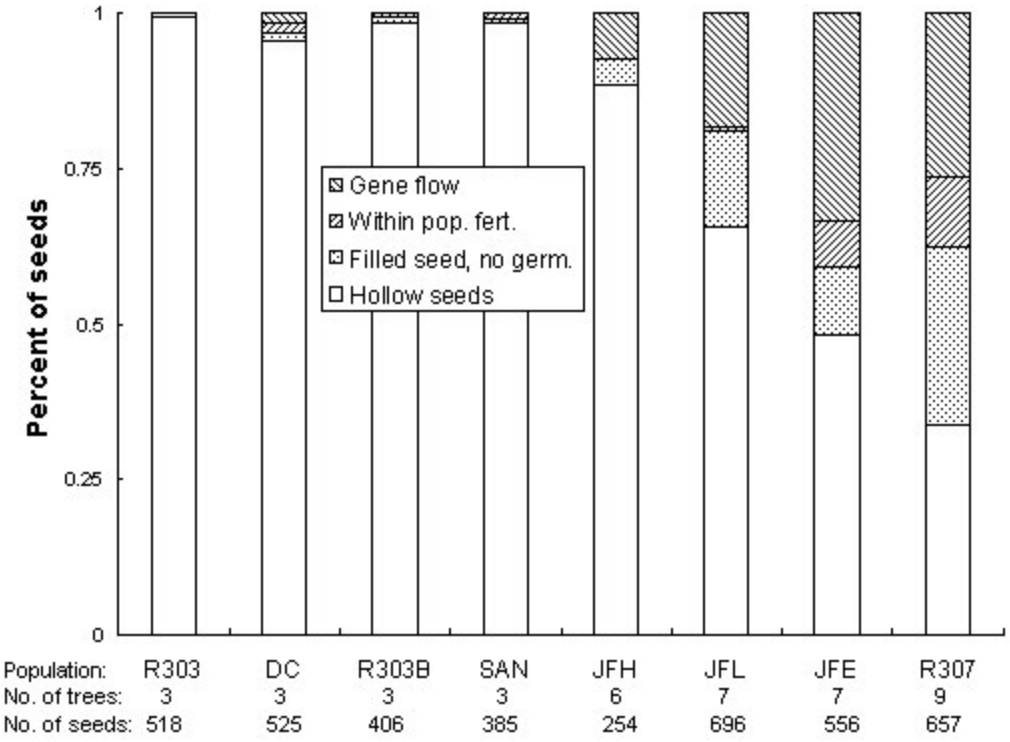


Figure 2.5. For each stand of trees examined: percentage of hollow seeds; seeds with enlarged embryo, but which did not germinate; seeds sired from within a population; and seeds that were the result of gene flow. Hollow seeds may have been aborted. Seeds that were filled, but did not germinate, had to have been alive at least 8-9 months, since embryos do not expand until the week before fruit ripening.

Table 2.1. The number of seeds collected for paternity analysis from stands of eight *Bursera simarunba* trees in southwestern Puerto Rico, the percent that germinated, and the percentage of those seeds that were aborted. The distance of each stand to the nearest conspecific tree is also presented.

Population	# adults	Isolation distance (m)	# seeds[a]	% aborted[b]
R303	3	200	518(0.0%)	99.4%
DC	3	300	525(3.2%)	95.6%
R303B	3	80	406(0.5%)	98.5%
SAN	3	>1000	385(0.8%)	98.4%
JFH	6	250	254(7.5%)	88.6%
JFL	7	380	696(19.9%)	65.5%
JFE	5	380	556(40.8%)	48.4%
R307	9	>1000	657(37.6%)	33.8%

[a] Percentage of seeds that germinated is in parentheses.

[b] Percentage of seeds which were hollow, with no apparent embryo. Viable seeds contained embryos that nearly filled the seed cavity.

Table 2.2. Genetic diversity statistics[a] for six populations of *Bursera simaruba*.

Population	N	*P*	*AP*	*A*	A_e	H_O (SD)	H_e (SD)
New_ac	24	66.67	2.50	2.00	1.50	0.202(0.069)	0.234(0.066)
Refuge	24	53.33	2.50	1.80	1.38	0.228(0.064)	0.193(0.061)
Anten	24	60.00	2.33	1.80	1.49	0.250(0.074)	0.245(0.063)
R303B	24	53.33	2.50	1.80	1.47	0.208(0.069)	0.227(0.065)
Tinaja	24	60.00	2.56	1.93	1.45	0.211(0.070)	0.213(0.064)
Pargue	24	66.67	2.30	1.87	1.40	0.223(0.074)	0.221(0.056)
MEAN		60.00	2.45	1.87	1.45	0.220	0.222
SD		5.13	0.10	0.08	0.05	0.029	0.026
Species estimate	144	73.30	2.73	2.27	1.50		0.244

[a] *P* is the percentage of polymorphic loci, *AP* is the mean number of alleles per polymorphic locus, *A* is the mean number of alleles per locus, A_e is the effective number of alleles, H_O is the observed heterozygosity and H_e is gene diversity, or expected heterozygosity.

Table 2.3. Estimates of genetic diversity parameters[a] for polymorphic loci surveyed for six populations of *Bursera simaruba*. Mean values are based on polymorphic loci.

Locus	H_T	H_S	F_{IS}	G_{ST}
DIA	0.021	0.020	-0.036	0.025
IDH	0.222	0.206	0.190	0.073
MDH1	0.455	0.423	0.195	0.070
MNR1	0.282	0.272	-0.249	0.035
MNR2	0.244	0.233	-0.220	0.044
PGI2	0.580	0.536	0.126	0.077
TPI1	0.558	0.504	0.001	0.096
UGPP1	0.520	0.461	-0.100	0.114
6PGD1	0.014	0.013	-0.044	0.035
SKDH	0.117	0.107	-0.168	0.086
PER	0.641	0.599	0.044	0.066
Avg:	0.332	0.307	-0.024	0.066

[a] H_T, total genetic diversity; H_S, genetic diversity found within populations; F_{IS}, deviations from Hardy-Weinberg expectations within individual populations; G_{ST}, the proportion of total genetic diversity found among populations.

Table 2.4. Estimated apparent gene flow rates, total gene flow rates (with standard deviation in parentheses), and migration rates (N_em) for five stands of *Bursera simaruba* from southwestern Puerto Rico.

Population	No. of fruits on tree[a]	# seeds analyzed	Isol. dist.(m)[b]	Apparent gene flow rate	Total gene flow rate
R307 (n=9)					
Tree 1	1625	66	>1 km	24%	45%(8%)
Tree 2	6375	32	>1 km	44%	75%(12%)
Tree 3	6290	41	>1 km	27%	52%(10%)
Tree 4	0	73	>1 km	42%	74%(9%)
Tree 5	125	35	>1 km	69%	98%(13%)
Mean (weighted)	–	247		39%	66%
JFL (n=7)					
Tree 1	750	44	880 m	14%	53%(17%)
Tree 2	0	57	780 m	19%	69%(16%)
Tree 3	5170	23	700 m	39%	100%(20%)
Mean (weighted)	–	124		21%	69%
JFE(n=5)					
Tree 1	1250	159	700 m	54%	89%(6%)
Tree 2	125	63	600 m	38%	64%(8%)
Mean (weighted)	–	222		49%	82%(5%)
JFH (n=6)					
Mean (weighted)	–	19	190 m	53%	100%(11%)
DC (n=3)					
Mean (weighted)	–	17	300 m	29%	47%(13%)

[a] Fruit production for individual tree, averaged over two years (R307, JFE) or three (JFL).

[b] Distance from a tree to the nearest conspecific tree from outside its population.

References

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

Becerra JX, Venable DL (1999). Nuclear ribosomal DNA phylogeny and its implications for evolutionary trends in Mexican *Bursera* (Burseraceae). American Journal of Botany 86: 1047-1057.

Carpenter FL (1976). Plant-pollinator interactions in Hawaii: pollination energetics of *Metrosideros collina* (Myrtaceae). Ecology 57: 1125-1144.

Chakroborty P, Meagher TR, Smouse PE (1988). Parentage analysis with genetic markers in natural populations. 1. Paternity exclusion and expected proportions of offspring with unambiguous paternity. Genetics 118: 527-536.

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

Cheliak WM, Pitel JA (1984). *Techniques for starch gel electrophoresis of enzymes for forest tree species*. Petawawa National Forestry Institute, Information Report P1-X-42. Canadian Forestry Service, Agriculture: Chalk River, Ontario.

Crow JF, Denniston C (1988). Inbreeding and variance effective population numbers. Evolution 42: 482-495.

Devlin B, Ellstrand NC (1990). The development and application of a refined method for estimating gene flow from angiosperm paternity analysis. Evolution 44: 248-259.

Dow BD, Ashley MV (1998). High levels of gene flow in bur oak revealed by paternity analysis using microsatellites. Journal of Heredity 89: 62-70.

Fischer M, Matthies D (1997). Mating structure and inbreeding and outbreeding depression in the rare plant *Gentianella germanica* (Gentianaceae). American Journal of Botany 84: 1685-1692.

Gibbs JP (2001). Demography versus habitat fragmentation as determinants of genetic variation in wild populations. Biological Conservation 100: 15-20.

Gigord L, Lavigne C, Shykoff JA (1998). Partial self-incompatibility and inbreeding depression in a native tree species of La Reunion (Indian Ocean). Oecologia 117: 342-352.

Gigord L, Picot F, Shykoff JA (1999). Effects of habitat fragmentation on *Dombeya acutangula* (Sterculiaceae), a native tree on La Reunion (Indian Ocean). Biological Conservation 88: 43-51.

Greenberg R, Foster MS, Marquez-Valdelamar L (1995). The role of the white-eyed vireo in the dispersal of *Bursera* fruit on the Yucatan Peninsula. Journal of Tropical Ecology 11: 619-639.

Groom M (2001). Consequences of subpopulation isolation for pollination, herbivory, and population growth in *Clarkia concinna* (Onagraceae). Biological Conservation 100: 55-63.

Hamrick JL, Godt MJW (1996). Effects of life history traits on genetic diversity in plant species. Philosophical Transactions of the Royal Society of London Series B 351: 1291-1298.

Hamrick JL, Loveless MD (1989). The genetic structure of tropical tree populations: associations with reproductive biology. In: Bock JH, Linhart YB (eds) *The Evolutionary Ecology of Plants*, Westview Press: Boulder, CO. Pp 129-146.

Hamrick JL, Nason JD (2000). Gene flow in forest trees. In: Boyle TJB, Young A, Boshier D (eds) *Forest Conservation Genetics: Principles and Practice*, CIFOR and CSIRO: Australia. Pp 81-90.

Heithaus ER, Stashko E, Anderson PK (1982). Cumulative effects of plant-animal interactions on seed production by *Bauhinia ungulata*, a neotropical legume. Ecology 63: 1294-1302.

Holdridge LR (1967). *Life zone ecology.* Tropical Science Center: San José, Costa Rica.

Hubbell SP, Foster RB (1983). Diversity of canopy trees in a neotropical forest and implications for the conservation of tropical trees. In: Sutton SL, Whitmore TC, Chadwick AC (eds) Tropical Rain Forest: Ecology and Management, Blackwell: Oxford. Pp 25-41.

Hufford KM, Hamrick JL (2003). Viability selection at three early life stages of the tropical tree, *Platypodium elegans* (Fabaceae, Papilionoideae). Evolution 57: 518-526.

Kaufman SR, Smouse PE Alvarez-Buylla ER (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.

Khurana E, Singh JS (2001). Ecology of seed and seedling growth for conservation and restoration of tropical dry forest: a review. Environmental Conservation 28: 39-52.

Knapp EE, Goedde MA, Rice KJ (2001). Pollen-limited reproduction in blue oak: Implications for wind pollination in fragmented populations. Oecologia 128: 48-55.

Koptur S (1984). Outcrossing and pollinator limitation of fruit set: breeding systems of neotropical *Inga* trees (Fabaceae: Mimosoideae). Evolution 38: 1130-1143.

Kuang H, Richardson TE, Carson SD, Bongarten BC (1999). Genetic analysis of inbreeding depression in plus tree 850.55 of *Pinus radiata* D. Don. II. Genetics of viability genes. Theoretical and Applied Genetics 99: 140-146.

Ledig FT, Jacob-Cervantes V, Hodgskiss PD, Eguiluz-Piedra T (1997). Recent evolution and divergence among populations of a rare Mexican endemic, Chihuahua spruce, following Holocene climatic warming. Evolution 51: 1815-1827.

Little EL, Wadsworth FH (1964). *Common trees of Puerto Rico and the Virgin Islands*. USDA Forest Service, Agr. Handbook 249.

Loveless MD, Hamrick JL, Foster RB (1998). Population structure and mating system in *Tachigali versicolor*, a monocarpic neotropical tree. Heredity 81: 134-143.

Manchenko GP (1994). *Handbook of detection of enzymes on electrophoretic gels*. CRC Press: Ann Arbor, MI.

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

Menges ES, Dolan RW (1998). Demographic viability of populations of *Silene regia* in Midwestern prairies: Relationships with fire management, genetic variation, geographic location, population size and isolation. Journal of Ecology 86: 63-78.

Mitton JB, Linhart YB, Sturgeon KB, Hamrick JL (1979). Allozyme polymorphisms detected in mature needle tissue of ponderosa pine. Journal of Heredity 70: 86-89.

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

Nason JD, Hamrick JL (1997). Reproductive and genetic consequences of forest fragmentation: two case studies of Neotropical canopy trees. Journal of Heredity 88: 264-276.

Neel MC, Ellstrand NC (2001). Patterns of allozyme diversity in the threatened plant *Erigeron parishii* (Asteraceae). American Journal of Botany 88: 810-818.

Platt WJ, Hill GR, Clark S (1974). Seed production in a prarie legume (*Astragalus canadensis* L.). Oecologia 17: 55-63.

Ray GJ, Brown B (1995). The structure of five successional stands in a subtropical dry forest, St. John, U.S. Virgin Islands. Caribbean Journal of Science 31: 212-222.

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

Rousset F (1997). Genetic differentiation and estimation of gene flow from F-statistics under isolation by distance. Genetics 145: 1219-1228.

Schuster WSF, Mitton JB (2000). Paternity and gene dispersal in limber pine (*Pinus flexilis* James). Heredity 84: 348-361.

Slatkin M (1977). Gene flow and genetic drift in a species subject to frequent local extinctions. Theoretical Population Biology 12: 253-262.

Slatkin M (1993). Isolation by distance in equilibrium and nonequilibrium populations. Evolution 47: 264-279.

Soltis DE, Haufler CH, Darrow DC, Gastony GJ (1983). Starch gel electrophoresis of ferns: a compilation of grinding buffers, gel and electrode buffers, and staining schedules. American Fern Journal 73: 9-27.

Stacy EA, Hamrick JL, Nason JD, Hubbell SP, Foster RB, Condit R (1996). Pollen dispersal in low-density populations of three neotropical tree species. American Naturalist 148: 275-298.

Stevens GC (1983). *Bursera simaruba* (Indio Desnudo, Jinocuave, Gumbo Limbo). In: Janzen DH (ed) *Costa Rican natural history*, University of Chicago Press: Chicago. Pp 201-202.

Vandermeer J, Hoffman B, Krantz-Ryan S (2001). Effect of habitat fragmentation on gypsy moth (*Lymantria dispar* L.) dispersal: The quality of the matrix. American Midland Naturalist 145: 188-193.

Wadsworth FH (1950). Notes on the climax forests of Puerto Rico and their destruction and conservation prior to 1900. Caribbean Forester 11: 38-47.

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

Willson MF, Rathcke BJ (1974). Adaptive design of the floral display in *Asclepias syriaca* L.. American Midland Naturalist 92: 47-57.

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

Wright S (1951). The genetical structure of populations. Annals of Eugenics 15: 323-354.

CHAPTER 3

DIRECT AND INDIRECT MEASURES OF GENE FLOW IN THE INTRODUCED TREE, *ALBIZIA LEBBEK*, IN SOUTHWESTERN PUERTO RICO[1]

[1] Dunphy, B.K. and J.L. Hamrick. To be submitted to *Heredity*.

Abstract

In this study we estimate gene flow and patterns of genetic diversity in *Albizia lebbek*, an invasive leguminous tree species in the Dry Forest of southwestern Puerto Rico. Genetic diversity estimates were calculated for ten populations of 24 trees each. These populations appear to have been formed from multiple introductions. The presence of unique genotypes in the northernmost populations suggests that trees with novel genotypes are still moving into the area. This combination of individuals from disparate source locations has led to high estimates of genetic diversity (H_e=0.266, P=0.67). Indirect estimates of gene flow indicate that only 0.69 migrants per generation move between populations, an amount that leads to the prediction that genetic diversity will decrease over time due to genetic drift. This estimate needs to be viewed with caution, however, since a key assumption, namely the presence of migration-drift equilibrium, was not met. Direct estimates of gene flow using paternity analysis were also obtained. This approach relies on fewer assumptions, and is probably more appropriate for fragmented populations. The percentage of pollen that was foreign in four stands of small trees (n<11) ranged from 44% to 100%. This high level of pollen flow is due not only to the flight capacities of its pollinators (several species of bees and butterflies), but also by a high level of outcrossing (t_m=0.979), forcing the trees to rely on outside pollen, rather than that of its neighboring, and presumably related, trees. Nevertheless, the regular production of pods indicates that sufficient outcross pollen is received. The role of gene flow in facilitating the spread of an invasive plant species is discussed.

Introduction

The rate that an introduced plant species spreads across a landscape depends largely upon the species' mode of reproduction and its dispersal abilities. Selfing species, for instance, are reproductively independent since a single individual can establish a population, allowing rapid colonization of new sites (Baker 1965). For outcrossing species, however, seed production is limited by the availability of compatible pollen. The greater the amount of compatible pollen, the more seeds will be available for increasing population size and the colonization of new sites. High levels of gene flow between established and newly colonized sites can provide the genetic material necessary to exploit new selective regimes as populations encounter novel environments (Sakai et al. 2001). Gene flow may also lead to hybrids between populations derived from different sources, increasing overall genetic diversity levels and perhaps leading to the formation of invasive genotypes (Ellstrand and Schierenbeck 2000). Characterizing levels of gene flow in exotic species, therefore, should increase our understanding of why certain species can successfully invade available habitats.

The development of genetic markers that are rapid to screen has greatly increased our ability to accurately estimate demographic and population-genetic measures such as gene flow. The easiest, and most widely used, approach to estimating gene flow using genetic markers begins by determining the spatial distribution of genetic diversity for a set of populations. Population-genetic models, especially that of Wright (1951), then describe the level of gene flow necessary to produce the observed distribution of genetic diversity. This approach relies on the assumption that the homogenizing force of migration balances genetic divergence among populations by genetic drift (Wright 1951;

Slatkin 1993), and consequently, the genetic structure seen among populations will be primarily affected by this equilibrium between migration and genetic drift.

Among introduced species, genetic structure is often heavily influenced by human activities (Godt and Hamrick 1991; Sakai et al. 2001). Since it can take many generations for these historical footprints to disappear, and for migration-drift equilibrium to be achieved (Slatkin 1993), indirect measures of gene flow may not be appropriate, at least for recently introduced species.

Direct measures provide an alternate means of measuring contemporary levels of gene flow. In one approach, paternity analysis, a direct estimate of gene flow via pollen is obtained by determining the pollen donor (i.e., the 'father") of a large number of seeds from a given population. Although requiring considerably more work, both in the field and the lab, this approach requires fewer assumptions than indirect methods, and may be the only reliable means of estimating contemporary gene flow for species not in migration-drift equilibrium. Gene flow, studied directly, has been found to be extensive in many tree species, with pollen often traveling more than a kilometer (Hamrick and Nason 2000).

In this study, we investigate estimates of gene flow into isolated populations of the introduced tree *Albizia lebbek*, using both indirect and direct measures. We determine whether its populations are in migration-drift equilibrium, and therefore amenable to indirect estimates of gene flow. Using direct measures, we determine the proportion of pollinations that involve immigrant pollen, and whether population size and distance to nearest neighbor influence rates of pollen flow. We also examine patterns of genetic diversity for clues to the success of *A. lebbek* as an invasive species.

Materials And Methods

Study populations were located near the tip of southwestern Puerto Rico in the vicinity of the Sierra Bermeja mountain range (Figure 3.2). The mostly flat topography is crossed by several low mountain ranges (200 to 225 m in height). The region is classified as subtropical dry forest (sensu Holdridge 1967), receiving less than 1000 mm of rain per year. Forest vegetation is scattered across a landscape that has been largely cleared for agriculture and cattle pastures. Like many formerly forested tropical areas, much of the site is covered by exotic grasses (Cabin et al. 2001; Daehler 1998). Especially common is African Guinea Grass (*Panicum maximum* Jacq.; J. Schwagerl, pers. comm.), which can burn in the dry season, although fires are not widespread.

Albizia lebbek (L.) Benth. (Fabaceae, Mimosoideae), a tree native to tropical Asia (Little and Wadsworth 1964), was introduced into islands of the British Caribbean in 1782 (Howard 1954). It has been in Puerto Rico at least since the beginning of the 20th century (Urban 1905). The species is used for shade and as an ornamental (Little and Wadsworth 1964), but has escaped cultivation and is now found in a variety of natural and disturbed habitats in Puerto Rico.

A. lebbek shares a number of traits with other successful invading tree species, including perfect flowers, generalist pollinators, a long residence time of the fruit on the tree (Reichard and Hamilton 1997), nitrogen-fixation (Daehler 1998), use of pioneer habitat, high seed production, high growth rates (Kolar and Lodge 2001), and multi-seeded fruits (Sakai et al. 2001). These traits presumably have helped *A. lebbek* invade this extensively-grazed island landscape, an environment particularly susceptible to invasion by non-native species (Lonsdale 1999).

The perfect flowers are visited mainly by a variety of generalist pollinators, predominantly bees and butterflies (B. Dunphy, pers. obs.). The fruits are long pods (10-20 cm) that typically contain between six and nine seeds. The light pods containing the seeds can be blown several hundred meters across open habitats with a strong wind, although a large number typically fall beneath the tree (B. Dunphy, pers. obs.).

Indirect estimates of gene flow

To obtain indirect gene flow estimates, ten populations of 24 individuals each were sampled (Figure 3.2). All populations were taken from low-elevation open fields with minimal forest cover. The minimum distance separating two populations was 0.5 km, and the maximum distance was 22.3 km. A minimum of 20 cm^2 of leaf tissue was collected from each tree. Leaves were shipped on ice to the laboratory at the University of Georgia within 48 hours of collection. Leaves were crushed in a potassium phosphate extraction buffer (Mitton et al. 1979) and resulting extracts were stored on filter paper wicks at -70° C.

Horizontal starch gel electrophoresis was used to assay allozyme diversity. Five gel buffer systems gel-electrode combinations and thirteen enzyme stains were used to resolve 21 loci on 11.0% starch gels. Gel and electrode buffer recipes followed Soltis et al. (1983). The enzymes stained (and loci resolved) on System 4 (a Tris/citrate gel and buffer) were isocitrate dehydrogenase (Idh), diaphorase (Dia), and phosphoglucoisomerase (Pgi-1 and Pgi-2). System 7 (a Tris-citric acid buffer gel and lithium hydroxide and boric acid tray buffer) was stained for amino acid transferase (Aat). A modified System 8 (a discontinuous LiOH system) was used to resolve leucine

aminopeptidase (Lap) and glutamate dehydrogenase (Gdh). System 11 (a histidine gel and citric acid buffer) was stained for malate dehydrogenase (Mdh-1, Mdh-2, and Mdh-3), 6-phosphogluconate dehydrogenase (6Pgd-1 and 6Pgd-2), UTP-glucose-1-phosphate (Ugpp-1 and Ugpp-2), shikimate dehydrogenase (Skdh), and Fructose-1,6-di-phosphatase (F16-1 and F16-2). System 6 (a Tris-citric acid buffer gel and sodium hydroxide and boric acid tray buffer) was used to resolve fluorescent esterase (Fe-1 and Fe-2) and triose-phosphate isomerase (Tpi-1 and Tpi-2). Stain recipes were modified from Soltis et al. (1983) except for diaphorase (Cheliak and Pitel 1984) and UTP-glucose-1-phosphate (Manchenko 1994). For enzymes with more than one locus, isozymes were numbered sequentially, with the lowest number assigned to the most anodal banding zone.

Standard measures of genetic variation were calculated for each population and the species level by pooling all the individuals analyzed. Genetic diversity parameters estimate percentage of polymorphic loci (*P*), the mean number of alleles per locus (*A*) and per polymorphic locus (*AP*), the effective number of alleles per locus (A_e), observed heterozygosity (H_o) and expected heterozygosity (H_e).

Indirect gene flow estimates, in terms of the effective number of immigrants per generation (N_em), were made with the following equation (Wright 1951):

$$N_em = (1-F_{ST})/4F_{ST}$$

where N_e is the effective number of individuals in the population, m is the immigration rate and F_{ST} is the proportion of total genetic diversity due to differences among populations (Wright 1951). A key assumption for using this approach is that the

populations are in migration-drift equilibrium. The presence of equilibrium is indicated by a statistically significant positive relationship between pair-wise estimates of $F_{ST}/(1-F_{ST})$ and the geographic distance separating the pair (Rousset 1997).

Direct measures of gene flow

Four stands of trees were chosen from the same geographic location as the populations used for the indirect estimates of gene flow (Figure 3.2). Distances from each stand to their nearest conspecific tree were 300 m (JH), 450 m (Manola), 60 m (R100) and 180 m (TX). From all stands, a total of 651 pods were analyzed.

Correlated mating, in which all seeds from a fruit are sired by a single pollen donor, is common in mimosoid legumes. To avoid problems of pseudoreplication that this may cause, a single seed was chosen from each pod, with the remaining seeds stored in labeled vials. Seeds chosen for analysis were soaked for 15 minutes in concentrated sulfuric acid, rinsed with water and planted at the Plant Biology Greenhouses at the University of Georgia in Athens, GA. Germination rates exceeded 90% and most seedlings emerged within 2-3 weeks. If a seed did not germinate, or if a seedling died, a replacement was taken from the same pod. Once true leaves appeared, roughly 20 cm^2 of leaf tissue was collected and processed for allozyme analyses as described above. A large number of seeds had exit holes created by the bruchid beetle, *Merobruchus paquetae* Kingsolver. These "pre-scarified" seeds were directly planted into the soil, and germinated at about the same rate as the acid-scarified seeds without exit holes.

Multi-locus estimates of outcrossing were computed with MLTR (Ritland 2002). Paternity analysis was used to directly estimate gene flow levels. Genotypes of all trees

from a seedling's stand that could potentially have sired the seed were compared against the seedling's genotype. An initial estimate of gene flow was made by determining the percentage of seedlings that could not have been sired by any of the potential fathers within the stand (i.e., apparent gene flow). Allozyme trials yielded an exclusion probability of 0.91 for *A. lebbek*, indicating that for a given seed, 91% of potential sires will be correctly identified as non-sires (Chakroborty et al. 1988). Since a father from outside the population could produce a pollen gamete indistinguishable from that created by a sire from within the population, however, the initial estimate will underestimate the true gene flow rate. The total number of seeds sired from outside a stand of trees, but seemingly consistent with having been sired from within, comprises "cryptic" gene flow, and its estimation is essential to obtain accurate and unbiased estimates of the total rate of gene flow.

Cryptic gene flow is difficult to estimate directly. We used the program GFLOW (available from B.K.D.), which employs a slight modification of the technique developed by Devlin and Ellstrand (1990). This technique uses maximum-likelihood to determine the total (apparent + cryptic) gene flow rate that would be most likely to yield the observed apparent gene flow rate (Devlin and Ellstrand 1990). A simulation was run over all possible total gene flow levels, with the minimum value being set by the apparent gene flow rate, and the highest being 100%. For each gene flow level (in increments of 1%), five thousand populations were created. These populations had the same number of adult trees, adult genotypes, and numbers of seedlings for each maternal tree as the actual population. Seedling genotypes were generated using a "pollen gamete" from either a randomly chosen adult from the population (non-gene flow), or from the "pollen pool"

(gene flow). The pollen pool represents the theoretical range of likely pollen genotypes that could be produced by trees in the vicinity of the population under study. For this study, the frequency of genotypes in the pollen pool was simply the average pooled frequency of genotypes from the 10 populations sampled for the indirect gene flow study.

Once generated, the apparent gene flow rate was calculated for each of the 5,000 populations. The number of times an apparent gene flow rate of a simulated population matched that of the actual population was tabulated. Total gene flow was then incremented by 1%, and another 5,000 populations were generated. The total gene flow rate with the greatest number of matches with the actual apparent gene flow rates was the most likely total gene flow rate. The overall gene flow rate for the whole population was calculated as the weighted average of gene flow for all the trees from the population.

Results

Diversity statistics and population structure

Fourteen of the 21 loci surveyed (66.7%) were polymorphic in at least one of the populations sampled (Table 3.1). At the species level, there were 3.29 alleles per polymorphic locus (AP_s), the effective number of alleles (A_{es}) was 1.58, and expected heterozygosity (H_{es}) was 0.266. At the population level, 48.9% of the loci were polymorphic for the ten populations, with R100S and R306 having the highest values (66.7% and 61.9%, respectively). Three of the ten populations (N_Refuge,Vivarium and Tinaja) shared the lowest value, 36.8%. There was an average of 1.36 effective alleles per locus (range 1.24-1.62), and 2.37 alleles at polymorphic loci (range 2.14-2.64). Values of genetic diversity (H_e) ranged from 0.151 (Parguera) to 0.293 (R100S), with an

average of 0.189 (Table 3.1). The highest diversity values were for three populations north of Rte. 101 (Figure 3.1).

Populations R100S and R306, and to some degree R100N differed in many respects from the other populations, in terms of common alleles and the presence of rare alleles. The most common alleles for populations R100S and R306 at Aat, Lap, Skdh and 6pgd-2 were found only in a few other populations, and at frequencies below 15%. Additionally, there were alleles at Dia, 6pgd-2, Lap and Fe-2 present only in population R100S, with the allele for Fe-2 being the most common for the population. Populations R100S and R306 were polymorphic at Tpi-1, Skdh and 6pgd-2; loci that tended to be monomorphic in most other populations.

Rare alleles (found at a species-level frequency of less than 0.05) were also found in Tex (Pgi-2, Aat, Skdh, Mdh-2 and 6pgd-2), R116W (Aat) and JFinca (Fe-1, Lap and 6pgd-2). These alleles were often the same as alleles that were common in populations R100S and R306.

Within R100S and R306, there was minimal mixing of alleles that were rare at a species-wide level, but common to R100S and R306, with alleles common to the other eight populations. Individuals tended to have either one or the other. A rare allele at Pgi-2 and one at Aat, in particular, were only found in homozygotes, despite the presence of common alleles within the same population. There was a similar finding with Fe-2, Dia, Lap, Skdh and 6pgd-2, although for each of these loci, there were a few individuals that were heterozygous for both rare and common alleles. Two alleles in 6pgd-2 formed only heterozygotes with the '3' allele, an allele mostly restricted to R100S and R306. Finally, individuals within R100S and R306 tended to either be free of rare alleles, or to have

them at multiple loci. Specifically, individuals with rare alleles averaged 6.6 (95% C.I., 6.10-7.05) loci with rare alleles.

Not surprisingly, R100S and R306 grouped together in the UPGMA phenogram, well separated from the other eight populations (Figure 3.3). Six populations were grouped in a close cluster. There did not appear to be any correlation between relationships in the phenogram and geographic proximity. For instance, N_Refuge and Vivarium are close together geographically, but did not group together in the phenogram.

Populations R100S and R306 exerted a strong influence over the population statistics. With the exception of *P*, which was unchanged with removal of these two populations from the analysis, all values decreased, with *AP* changing from 3.29 to 2.79, *A* dropping from 2.52 to 2.19, A_e decreasing from 1.58 to 1.39, and H_e lowering from 0.266 to 0.206.

A deficit of heterozygotes, indicated by positive F_{IS} values (Table 3.2), was found for ten loci, while four loci demonstrated an excess of heterozygotes. Across loci, there was a significant deficit of heterozygotes (X^2=145.54; P<0.001). Pooling populations together leads to substantially fewer heterozygotes than expected from Hardy-Weinberg expectations (F_{IT}=0.368), largely due to significant among population differentiation (G_{ST}=0.266). A X^2 analysis for allele frequency heterogeneity among populations yielded highly significant (P<0.001) differences among populations for all polymorphic loci except Idh, which was significantly different, but not as strongly as the rest (P<0.025).

Very high F_{IS} values were found for Fe-1, Lap, Skdh and 6pgd-2, loci that had novel genotypes in the two northern populations, R100S and R306. The very low F_{IS}

observed for Tpi-1, a locus monomorphic for most populations, was due to an excess of heterozygotes in R100S (F_{IS}=-0.345) and especially R306, where the excess was statistically significant (F_{IS}=-0.469, X^2=5.27, $p<0.05$).

Of 96 fixation indices calculated, 29 (30%) were significantly different from zero ($p<0.05$). Twenty-seven of these 29 indices were positive, indicating a deficit of heterozygotes. All populations except N_Refuge had at least one fixation index significantly different from zero. R306 had the most, with eight loci, followed by R100S with five loci.

Indirect measures of gene flow

The percent of the variation distributed among populations (G_{ST}) was 26.6.% (Table 3.1). Considered pairwise, only two G_{ST} values were between 0.1 and 0.2; the remainder were either higher or lower. At the low end, 28.9% of G_{ST} values were below 5%. Two population pairs had values below 1%, despite a minimum of 3.2 km separating each pair. At the other extreme, 33.3% of pairs had G_{ST} values above 30% (13.3% were above 40%). All pairwise G_{ST} values above 0.300 involved either R100S or R306, two of the three populations north of R101. With these two populations removed, overall G_{ST} drops to 0.120.

No apparent relationship existed between genetic and geographic distances. The genetic distance between the two most geographically distant populations (Tex and Parguera) was 0.0187, a value less than 77.8% of other pairwise genetic distance values. Plotting $F_{ST}/(1-F_{ST})$ against interpopulation distances (Figure 3.2) yielded a non-significant, slightly negative slope, rather than the positive slope expected under

migration-drift equilibrium (R^2=0.039, P>0.10). Removing populations R100S and R306 increased the total variation in genetic distance explained to 6.6%, yet the relationship was still not statistically significant.

Direct measures of gene flow

The multi-locus estimate of outcrossing (t_m) was 0.979 (s.e.=0.052). Gene flow rates were high, with values of 44% for JH, 78% for Manola, 100% for R100 and 81% for the Tex population Table 3.3). In Manola and Tex, the stands where seeds were available from multiple maternal trees, variation within stands was greater for the former, which had trees with gene flow estimates as low as 65% and as high as 100%. The two values for Tex were 80% and 83%.

An estimate of effective population size (N_e) was made for Manola and Tex, the two populations for which fruit production was recorded. N_e was calculated as $N_e=1/\Sigma((c_i+p_i)/2)^2$, where c_i and p_i are the relative female and male reproductive contribution of the *i*th individual, respectively (Crow and Denniston 1988). Values of N_e were 7.26 for Manola, and 4.33 for Tex. Combining N_e with the gene flow rates estimated above yields estimates of N_em of 2.83 (Manola) and 1.75 (Tex), although low effective population sizes limited N_em to 3.63 and 2.17, respectively.

Discussion

Gene Flow Rates

Gene flow rates were high for all four stands, demonstrating that pollen flows readily over the isolation distances examined, and therefore pollen limitation should not

prevent reproduction and spread of *A. lebbek*. Sakai et al. (2001) point out that the rate and extent of a species' invasion can depend upon patterns of gene flow, which brings genetic diversity to the edge of a species range. This genetic diversity may facilitate adaptation to local environmental conditions. In self-incompatible species, high gene flow also ensures large seed crops, further facilitating the colonization of new sites.

Estimates of gene flow from the indirect and direct estimators lead to slightly different predictions. According to population genetic theory (Wright 1931), the indirect estimate, Nm=0.69, predicts that populations should grow more dissimilar over time. With the potentially problematic genotypes from populations R100S and R306 removed, the value of Nm increased to 1.28, slightly within the key range of 1-4, where both gene flow and drift can influence future change in genetic diversity. The direct estimates, N_em=1.75 and 3.04, fall within this range as well. The apparent absence of migration-drift equilibrium for this invasive species suggests that the direct estimates are more reliable.

The high availability of suitable pollen is reflected in the consistent production of large numbers of seed pods. This was the case even in years when fruit production in other species was low (B. Dunphy, pers. obs.). A high gene flow rate acts therefore not only to exchange genes among populations, but also ensure a large fruit crop, providing seeds for both individual population expansion and the colonization of new sites. Long duration of fruit on the tree has been cited as a factor promoting success as an invasive species (Reichard and Hamilton 1997). In *A. lebbek*, where pods remain on the branch for weeks or months, this may increase exposure to occasional wind bursts that have the potential to move pods substantial distances. Wind speeds in the region average 11.2

m.p.h., but wind gusts between 28.4-34.2 m.p.h. occur almost every month of the year (NOAA 1995). Thus, pods are available when these wind bursts occur, allowing distant dispersal of seeds.

No relationship was seen between nearest neighbor distance or stand size and gene flow. This is consistent with pollen moving readily across the distances involved in this study. The high gene flow rates reported for *A. lebbek* are consistent with what has been found for other tree species. Lone trees of the congeneric species, *Albizia julibrissin*, isolated by at least 1 km, produced large crops of outcrossed seeds, indicating pollen movement over long distances (Godt and Hamrick 1997). Similar gene flow levels were reported from Barro Colorado Island in Panama (Nason and Hamrick 1997). Three species of *Ficus* had gene flow levels greater than 90% and small populations of *Spondias mombin* on islands in Lake Gatun had pollen flow rates of 60% to 100%. White et al. (2002) examined gene flow in several forest fragments of the rare tree species, *Swietenia humilis* in a highly disturbed landscape in Honduras. Highly variable SSRs allowed them to determine that between 24% and 100% of pollen traveled more than 900m. In one case, an isolated tree received 71% of its pollen from over 4.5 km. Thus, long-distance pollen movement for insect-pollinated species such as *A. lebbek* appears to be the rule rather than the exception.

Patterns of Genetic Diversity

Diversity statistics for *A. lebbek* were comparable or exceeded those reported for other long-lived perennial outcrossing tree species by Hamrick and Godt (1996). The percent of polymorphic loci was similar (66.7% versus 65.5%), while the gene diversity

value (H_e) was substantially higher (0.266 versus 0.180). Over a quarter of this diversity is due to variation among populations (G_{ST}=0.266 versus 0.094), a value normally seen with species that undergo some degree of selfing. Since *A. lebbek* is almost entirely outcrossing, the high G_{ST} values are likely due to historical population founding events, rather than a direct consequence of its mating system. Over time, the high rates of gene flow observed will likely homogenize populations, leading to a decrease of G_{ST}. Based upon direct estimates of N_em, G_{ST} should reach an equilibrium value of 0.097 (Manola) or 0.125 (Tex), values more consistent with what has been recorded in species with similar life history traits (Hamrick and Godt 1996).

There was a substantial lack of heterozygotes in both R100S (F_{IS}=0.416) and R306 (F_{IS}=0.438). Two possible causes of this pattern are the presence of a reproductively-isolated cryptic species or the recent introduction of new genotypes into these two sites. The latter explanation seems more likely, since there were heterozygotes, although uncommon, that combined rare alleles with ones common in the other eight populations, indicating that mating between genotypes occurs. At other polymorphic loci, R100S and R306 also shared common alleles with other populations. This would not be likely if a cryptic species was present. The most likely explanation is that individuals with unusual genotypes have only arrived recently in the area. This idea is further supported by the presence of loci with excess heterozygotes and loci with a deficit of heterozygotes (indicated by positive and negative F_{IS} values, respectively), a situation that would not normally be seen in an established population, where all loci are exposed to the same demographic forces. This would provide further evidence that these

populations are not yet in migration-drift equilibrium, and that current genetic structure is still influenced by founding events.

The very different genotypes found in two of the three northern populations suggest that multiple source populations contributed to the present genetic composition of the species in the area of the study site. This would explain the very high genetic diversity values (0.266), a value higher than the average reported for other species with similar life history characteristics (Hamrick and Godt 1996). Introduced populations, derived from several source populations, often will be more diverse than any single source population (Barrett and Husband 1990). A high level of genetic diversity (0.290) was also found for the introduced species, kudzu (*Pueraria lobata*), presumably due to historically documented cases of multiple introductions (Pappert et al. 2000).

The lack of migration-drift equilibrium may be due to the relatively recent introduction of *A. lebbek* into southwestern Puerto Rico, as the presence of unusual alleles in R100S and R306 suggests. Over time, as these alleles move through the landscape, it is unclear whether migration-drift equilibrium will be achieved. In the relationship between F_{ST} and geographic distance (Figure 3.2), there was a cloud of points, representing a range of geographic distances, with low F_{ST} values. These points all involve population-pairs with at least one of the pair from populations Tinaja, Parguera, R100N or R116W. These low F_{ST} values could be due to recent establishment of these populations from a common source, or from high gene flow between these populations. Since the four populations are along roadsides (Parguera, R100N, R116W) or in a fallow field (Tinaja), it is unlikely that they were deliberately planted. High gene flow, both via pollen as demonstrated in the current study, and via seed flow seems more

likely. If this is the case, then the low F_{ST} values indicate that these must be older, more established populations for which enough time has passed for gene flow to homogenize differences in allele frequencies due to the original founder effects.

With the cloud of points removed, the relationship between F_{ST} and geographic distance becomes positive. The relationship may be accidental, however, since JFinca, which is the most remote population in this group of four, also tends to have high F_{ST} values with all other nine populations, including its nearest neighbors. This is reflected in the UPGMA phenogram, where JFinca was separated from the other populations. In addition, a pattern of isolation by distance tends to first appear among nearest neighbors (Slatkin 1993). On average, however, an *A. lebbek* population is more closely related to 4.3 other populations than to its nearest neighbor, so there is no evidence that isolation by distance is developing on the spatial scale examined.

Albizia lebbek's success as an invasive exotic.

Albizia lebbek has a number of reproductive characteristics that make it a good invader, including even and large fruit production, large amounts of seed, and fruits that remain attached to the tree for long periods of time (Sakai et al. 2001). The same species, however, can have both invasive and non-invasive populations (e.g., Sakai et al. 2001), indicating a strong genetic component to invasiveness. Although some preadpatation must occur to allow an invader to colonize a site, subsequent adaptation may be just as important in allowing the invader to spread into new territory (Sakai et al. 2001). If there are few colonists, and especially if all are drawn from the same source, genetic bottlenecks can lead to inbreeding depression (Ellstrand and Elam 1993; Newman and

Pilson 1997), and can lower the genetic diversity that may be needed to allow populations to adapt to an invaded habitat (Sakai et al. 2001). Conversely, a large amount of genetic diversity may allow a species to endure a wide range of selective regimes while it expands its range (Antonovics 1976; Crawley 1986; Hengeveld 1990).

Morrison and Molofsky (1999) found a significant genotype x environment interaction in introduced populations of reed canary grass (*Phalaris arundinacea*), where the level of competition determined which genotype had the highest growth rate. In particular, one genotype grew faster than others under conditions of strong competition, while a different genotype had the fastest growth rate when competition was minimal. The maintenance of genetic diversity may therefore be an important factor allowing this grass to successfully invade new habitats, where the level of competition may vary within and across sites.

Invasiveness often develops only after multiple introductions, a potential source of substantial amounts of genetic diversity (e.g., Pappert et al. 2000). In particular, hybridization between individuals from source populations well-isolated in their native range may stimulate invasiveness. Extensive gene flow may bring together novel combinations of alleles, allowing invasive genotypes to form (Ellstrand and Schierenbeck 2000). In self-incompatible species, such as *A. lebbek*, multiple invasions may also introduce a large number of self-incompatibility alleles, thus overcoming a potential barrier to population expansion for these species, especially given small founder population sizes.

The frequent occurrence of bruchid beetles, which do not appear to harm the embryo, may also actually increase the ecological success of *A. lebbek*. Species with

germination requirements met over a range of environmental conditions tend to be more successful (Baker 1974), especially those that lack a pre-germination requirement (Reichard and Hamilton 1997). Two types of seeds are effectively produced in *A. lebbek*: seeds with exit holes, capable of immediate germination; and intact seeds that require scarification before they can germinate. These two types of seeds, then, cover a range of possible conditions. The former group of seeds can take advantage of any favorable conditions soon after production. The latter group, conversely, can wait in the seed bank for longer periods of time.

Summary

In conclusion, pollen appears to move freely over the distances examined. The *Albizia lebbek* populations in southwestern Puerto Rico are not in migration-drift equilibrium, so indirect measures of gene flow can not be used. Although considerable genetic structure exists, this is probably due to the recent arrival of foreign genotypes from north of the study site. Migration-drift equilibrium may only appear over larger geographic distances. Because trees are not pollen limited, the capacity for seed dispersal, and the ability to deal with local environmental regimes, should dictate the rate of spread of the species in the area of the study site. High gene flow rates maintain high genetic diversity levels, which may aid continued spread of the species by providing the environment-specific genotypes needed for a wide range of conditions.

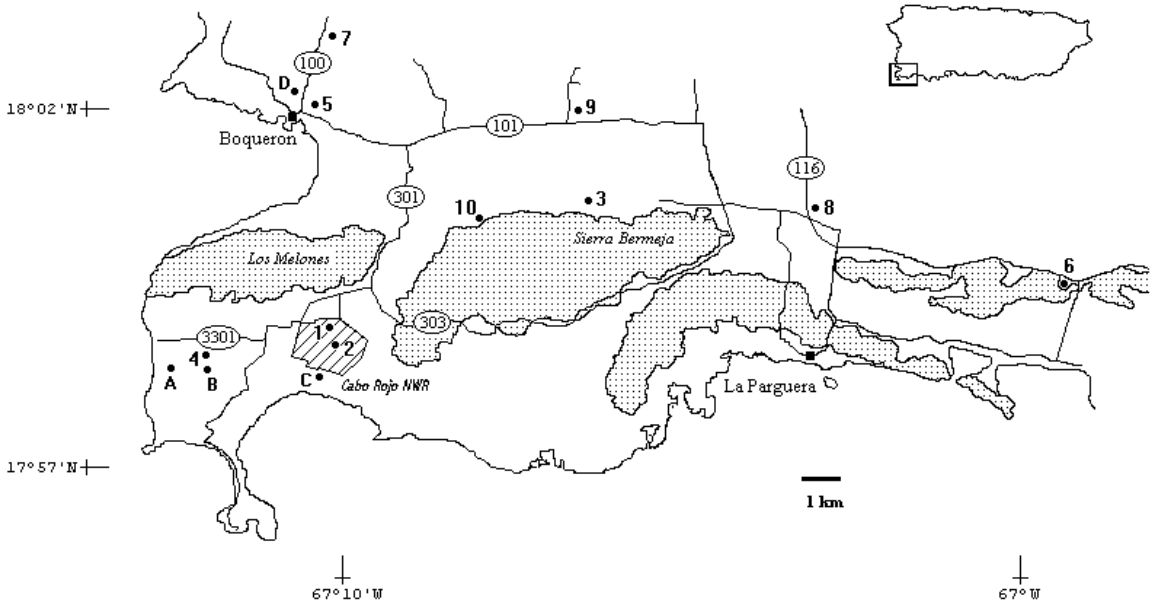


Figure 3.1. Study populations of *Albizia lebbek* in southwestern Puerto Rico. Populations of adult trees used in indirect estimates of gene flow are numbered 1-10. The four stands of trees used for direct measures are lettered A-D. Stippled regions are low mountains (maximum height of 200-225 meters). Diagonal lines show the location of the Cabo Rojo National Wildlife Refuge.

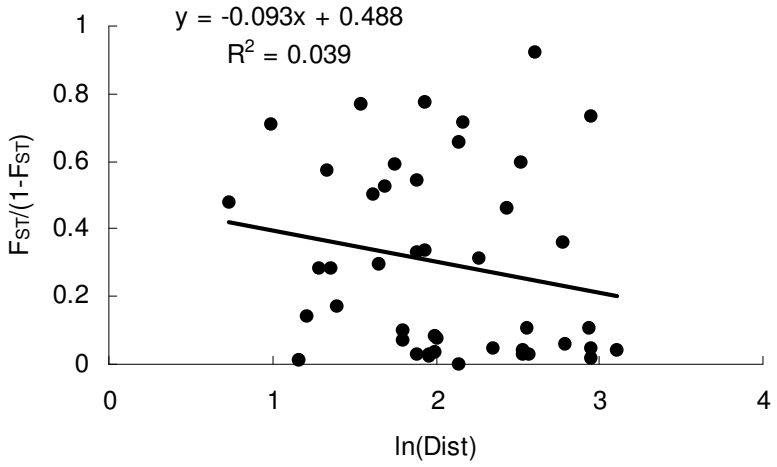


Figure 3.2. Examination of isolation by distance in all possible pairs of ten *Albizia lebbek* populations from southwestern Puerto Rico. A significantly positive slope would have indicated migration-drift equilibrium.

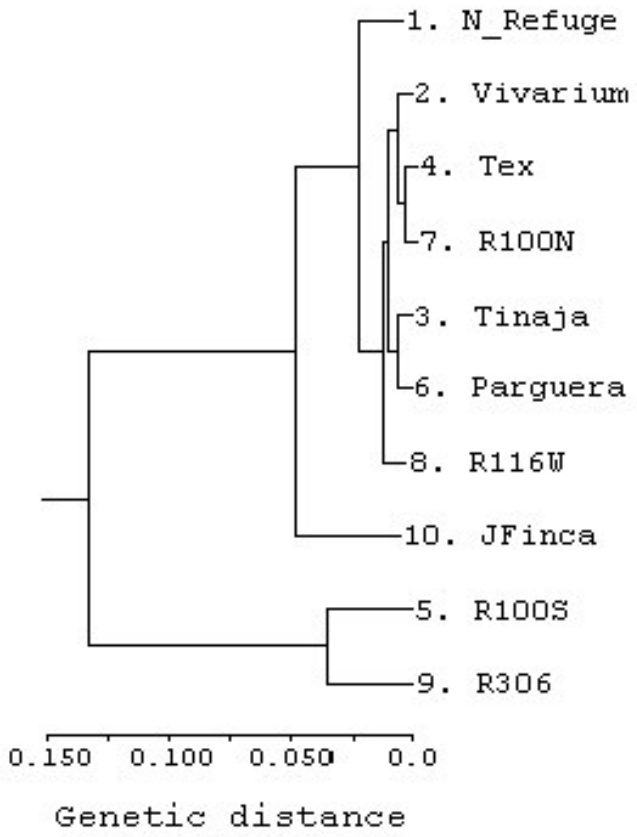


Figure 3.3. UPGMA phenogram of genetic distances among ten introduced *Albizia lebbek* populations in southwestern Puerto Rico.

Table 3.1. Genetic diversity statistics[a] for ten introduced populations of *Albizia lebbek* in southwestern Puerto Rico.

Population	N	*P*	*AP*	*A*	A_e	H_O (SD)	H_e (SD)
1. N_Refuge	21	36.8	2.43	1.53	1.30	0.170 (0.064)	0.162 (0.051)
2. Vivarium	24	36.8	2.29	1.47	1.34	0.126 (0.056)	0.157 (0.055)
3. Tinaja	23	36.8	2.14	1.42	1.28	0.152 (0.057)	0.153 (0.050)
4. Tex	11	47.4	2.56	1.74	1.40	0.163 (0.087)	0.186 (0.055)
5. R100S	20	66.7	2.64	2.10	1.62	0.171 (0.075)	0.293 (0.055)
6. Parguera	24	47.6	2.20	1.57	1.24	0.147 (0.061)	0.151 (0.040)
7. R100N	24	60.0	2.50	1.90	1.37	0.158 (0.063)	0.205 (0.048)
8. R116W	24	42.9	2.33	1.57	1.33	0.155 (0.063)	0.179 (0.049)
9. R306	24	61.9	2.38	1.86	1.38	0.122 (0.055)	0.217 (0.046)
10. JFinca	24	52.4	2.27	1.67	1.37	0.183 (0.064)	0.190 (0.051)
MEAN		48.9	2.37	1.68	1.36	0.155	0.189
SD		3.5	0.16	0.21	0.10	0.021	0.016
Species estimate	219	66.7	3.29	2.52	1.58		0.266

[a] *P* is the percentage of polymorphic loci, *AP* is the mean number of alleles per polymorphic locus, *A* is the mean number of alleles per locus, A_e is the effective number of alleles, H_O is the observed heterozygosity and H_e is gene diversity, or expected heterozygosity.

Table 3.2. Estimates of genetic diversity parameters[a] for polymorphic loci surveyed for ten populations of *Albizia lebbek*. Mean values are based on polymorphic loci.

Locus	H_T	H_S	F_{IS}	G_{ST}
Idh	0.226	0.214	−0.127	0.055
Pgi-2	0.587	0.414	0.221	0.294
Fe-1	0.146	0.079	0.484	0.458
Fe-2	0.532	0.387	0.190	0.272
Tpi-1	0.123	0.093	−0.411	0.241
Dia	0.502	0.448	0.106	0.109
Aat	0.610	0.477	0.046	0.219
Lap	0.705	0.534	0.543	0.242
Skdh	0.239	0.131	0.896	0.451
6pgd-2	0.281	0.144	0.557	0.489
Ugpp-1	0.541	0.463	0.044	0.144
Ugpp-2	0.498	0.226	−0.169	0.546
Mdh-2	0.498	0.443	0.190	0.109
Mdh-3	0.095	0.086	−0.158	0.091
Avg:	0.399	0.296	0.172	0.266

[a] H_T, total genetic diversity; H_S, genetic diversity found within populations; F_{IS}, deviations from Hardy-Weinberg expectations within individual populations; G_{ST}, the proportion of total genetic diversity found among populations.

Table 3.3. Apparent gene flow rates, total gene flow rates (with standard deviation in parentheses) and isolation distances for nine trees from four stands of *Albizia lebbek* in southwestern Puerto Rico.

Population	# seeds analyzed	Apparent gene flow rate	Total gene flow rate	Isol. dist.(m)
A. Tex(n=8)				
Tree 1	152	24%	80%(13%)	180
Tree 2	43	21%	83%(23%)	180
Mean (weighted)			80.7%	
B. JH (n=3)				
Tree 1	81	26%	44%(5%)	300
C. Manola(n=11)				
Tree 1	23	22%	100%(33%)	450
Tree 2	44	30%	65%(23%)	450
Tree 3	20	25%	93%(22%)	450
Tree 4	133	12%	76%(8%)	450
Mean (weighted)			77.9%	
D. R100(n=2)				
Tree 1	142	44%	100%(8%)	60

References

Antonovics J (1976). The nature of limits to natural selection. Annals of the Missouri Botanical Garden 63: 224-247.

Baker HG (1965). Characteristics and modes of origin of weeds. In: Baker HG, Stebbins GL (eds) *The Genetics of Colonizing Species*, Academic Press: New York. Pp 147-172.

Baker HG (1974). The evolution of weeds. Annual Review of Ecology and Systematics 5: 1-24.

Barrett SCH, Husband BC (1990). The genetics of plant migration and colonization. In: Brown AHD, Clegg MT, Kahler AL, Weir BS (eds) *Plant Population Genetics, Breeding and Genetic Resources*, Sinauer: Sunderland, MA. Pp 254-277.

Cabin RJ, Weller SG, Lorence DH, Flynn TW, Sakai AK, Sandquist D, Hadway LJ (2001). Effects of long-term ungulate exclusion and recent alien species control on the preservation and restoration of a Hawaiian tropical dry forest. Conservation Biology 14: 439-453.

Chakroborty P, Meagher TR, Smouse PE (1988). Parentage analysis with genetic markers in natural populations. 1. Paternity exclusion and expected proportions of offspring with unambiguous paternity. Genetics 118: 527-536.

Cheliak WM, Pitel JA (1984). *Techniques for starch gel electrophoresis of enzymes for forest tree species*. Petawawa National Forestry Institute, Information Report P1-X-42. Canadian Forestry Service, Agriculture: Chalk River, Ontario.

Crawley MJ (1986). The population biology of invaders. Philosophical Transactions of the Royal Society of London Series B 314: 711-729.

Crow JF, Denniston C (1988). Inbreeding and variance effective population numbers. Evolution 42: 482-495.

Daehler CC (1998). The taxonomic distribution of invasive angiosperm plants: ecological insights and comparison to agricultural weeds. Biological Conservation 84: 167-180.

Devlin B, Ellstrand NC (1990). The development and application of a refined method for estimating gene flow from angiosperm paternity analysis. Evolution 44: 248-259.

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

Ellstrand NC, Schierenbeck KA (2000). Hybridization as a stimulus for the evolution of invasiveness in plants. Proceedings of the National Academy of Sciences USA 97: 7043-7050.

Godt MJW, Hamrick JL (1991). Genetic variation in *Lathyrus latifolius* (Leguminosae). American Journal of Botany 78: 1163-1171.

Godt MJW, Hamrick JL (1997). Estimation of mating system parameters of *Albizia julibrissin* (Fabaceae). Forest Genetics 4: 217-221.

Hamrick JL, Godt MJW (1996). Effects of life history traits on genetic diversity in plant species. Philosophical Transactions of the Royal Society of London Series B 351: 1291-1298.

Hamrick JL, Nason JD (2000). Gene flow in forest trees. In: Boyle TJB, Young A, Boshier D (eds) *Forest Conservation Genetics: Principles and Practice*, CIFOR and CSIRO: Australia. Pp 81-90.

Hengeveld R (1990). *Dynamic Biogeography*. Cambridge Univ. Press: London.

Holdridge LR (1967). *Life zone ecology.* Tropical Science Center: San José, Costa Rica.

Howard RA (1954). A history of the Botanic Garden of St. Vincent, British West Indies. Geographical Review 44: 381-393.

Kolar C, Lodge DM (2001). Progress in invasion biology: predicting invaders. Trends in Ecology and Evolution 16: 199-204.

Little EL, Wadsworth FH (1964). *Common trees of Puerto Rico and the Virgin Islands*. USDA Forest Service, Agr. Handbook 249.

Lonsdale WM (1999). Global patterns of plant invasions and the concept of invasibility. Ecology 80: 1522-1536.

Manchenko GP (1994). *Handbook of detection of enzymes on electrophoretic gels*. CRC Press: Ann Arbor, MI.

Mitton JB, Linhart YB, Sturgeon KB, Hamrick JL (1979). Allozyme polymorphisms detected in mature needle tissue of ponderosa pine. Journal of Heredity 70: 86-89.

Morrison SL, Molofsky J (1999). Environmental and genetic effects on the early survival and growth of the invasive grass *Phalaris arundinacea*. Canadian Journal of Botany 77: 1447-1453.

Nason JD, Hamrick JL (1997). Reproductive and genetic consequences of forest fragmentation: two case studies of Neotropical canopy trees. Journal of Heredity 88: 264-276.

National Oceanic and Atmospheric Administration (NOAA) (1995). *Wind frequency statistics and monthly peak gusts for Ponce, Puerto Rico (1972-1990).* Unpublished data from the Southeast Regional Climate Center, Columbia, South Carolina.

Newman D, Pilson D (1997). Increased probability of extinction due to decreased genetic effective population size: experimental populations of *Clarkia pulchella*. Evolution 51: 354-362.

Pappert RA, Hamrick JL, Donovan LA (2000). Genetic variation in *Pueraria lobata* (Fabaceae), an introduced, clonal, invasive plant of the southeastern United States. American Journal of Botany 87: 1240-1245.

Reichard SH, Hamilton CW (1997). Predicting invasions of woody plants introduced into North America. Conservation Biology 11: 193-203.

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

Rousset F (1997). Genetic differentiation and estimation of gene flow from F-statistics under isolation by distance. Genetics 145: 1219-1228.

Sakai A, Allendorf FW, Holt JS, Lodge DM, Molofsky J, With KA, Baughman S, Cabin RJ, Cohen JE, Ellstrand NC, McCauley DE, O' Neill P, Parker LM, Thompson JN, Weller SG (2001). The population biology of invasive species. Annual Review of Ecology and Systematics 32: 305-332.

Slatkin M (1993). Isolation by distance in equilibrium and nonequilibrium populations. Evolution 47: 264-279.

Soltis DE, Haufler CH, Darrow DC, Gastony GJ (1983). Starch gel electrophoresis of ferns: a compilation of grinding buffers, gel and electrode buffers, and staining schedules. American Fern Journal 73: 9-27.

Urban I (1905). Symbolae Antillanae seu Fundamenta Florae Indiae Occidentalis. In: *Flora Portoricensis,* Fratres Borntraeger: Leipzig. Vol 4, pp 193-352.

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

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

Wright S (1951). The genetical structure of populations. Annals of Eugenics 15: 323-354.

CHAPTER 4

DIRECT AND INDIRECT MEASURES OF GENE FLOW IN THE BAT-POLLINATED TREE, *HYMENAEA COURBARIL*, IN SOUTHWESTERN PUERTO RICO[1]

[1] Dunphy, B.K., J.L. Hamrick and J. Schwagerl. To be submitted to *International Journal of Plant Sciences*.

Abstract

Gene flow was estimated in the tropical tree, *Hymenaea courbaril*, using indirect, F_{ST}-based measures, and direct, paternity analyses. Outcrossing estimates (t_m=1.07) are consistent with suggestions that *H. courbaril* is self-incompatible. Direct estimates of pollen gene flow into two small (n<6) stands of trees indicate that 42% and 62% of the pollen came from outside the stand. Despite this level of external gene flow, genetic drift is expected to cause divergence of populations over time (N_em =0.40 and 0.63), mainly due to the small number of adult trees within each population. Indirect estimates of gene flow based on nine larger populations of 21 to 48 trees gave estimates of 2.91 migrants per generation, but the absence of migration-drift equilibrium renders this estimate unreliable. Analysis of pollen movement among 21 adult trees indicated that the effective number of fathers was only 6.9, with the majority of pollinations coming from just three trees. Asynchronous flowering, and perhaps restricted use of flowering resources by bats, the primary pollinators of *H. courbaril*, is the likely cause of this result. Genetic diversity values, based on 31 allozyme loci, were lower than those reported for other outcrossing, gravity-dispersed tree species. Few effective fathers, and a relatively low population density, are the likely reasons.

Introduction

The capacity for long-distance pollen movement may be crucial for plants in a fragmented landscape, where inbreeding and genetic drift threaten the continued survival of small relictual populations. Plants pollinated by bats in particular may be expected to experience relatively high levels of gene flow, since bats are known to be strong fliers capable of traveling long distances. For example, two species of bats in southwestern Australia, fitted with radio transmitters, moved up to 6.9 km nightly from roosting to foraging sites (Lumsden et al. 2002). Even in the face of habitat fragmentation, many species of bats continue to visit plants in forest remnants, readily flying over open areas (Law et al. 1999). In fragmented Australian tropical rainforest, bats flew up to 5.8 km (mean 1 km) across cleared land (Law and Lean 1999). Home ranges encompassed 12-1796 ha, and often contained multiple forest fragments. Bats fed only 1.2 minutes per tree (yet carried more pollen than birds), and frequently moved over 200 m between trees, suggesting the possibility of long-distance pollen movement.

For the Costa Rican tree, *Bauhinia ungulata*, Heithaus et al. (1982) found that the distance that bats flew from a roosting site to a flowering tree (range 270-1420 m) did not affect pollination success, suggesting that trees could obtain adequate pollen over all distances examined. In 39.4% of flowers, over 50% of the stigmatic surface was covered by pollen, a situation usually followed by normal fruit production. In Mexico, the Jamaican fruit bat (*Artibeus jamaicensis*), has been found to fly nightly 8 ($\pm$2) km from roosting to foraging areas (Morrison 1978). Handley et al. (1991) captured a female of the same species in Panama in a mist net at its preferred tree, *Ficus insipida*, two hours after capturing it in a different *F. insipida* tree 1.2 km away. In Costa Rica, 22 bats of

five species carrying pollen of *Ceiba aesculifolia* were captured an average of 0.56 (±0.32) km from the lone tree of that species in the study site (Heithaus et al. 1975). Two larger bat species (*A. jamaicensis* and *Phyllostomus discolor*) were often caught over 1 km away.

Pollen from multiple species of plants is frequently found on bats. In Costa Rica, for instance, 42.3% (varying from 15.8% to 77.5%, depending upon species of bat examined) of pollen loads were mixed (Heithaus et al. 1975). Since bats groom thoroughly at their roosts, this suggests that multiple tree species, and perhaps multiple trees of the same species, are visited in a single evening (Heithaus et al. 1975). In Costa Rica, 43.0% of Jamaican fruit bats had mixed pollen loads on their fur. On average, 14.2% of Jamaican fruit bats were recaptured, at an average distance of 673.5 m from their first capture point.

Among six factors reducing reproductive output in the Costa Rican tree, *Bauhinia ungulata*, Heithaus et al. (1982) found that lack of pollination was the most important. Given the long distance bats are capable of traveling, we expect that plants pollinated by bats should be able to receive pollen over long distances.

Recent studies of gene flow using techniques such as paternity analysis have provided direct evidence in several tree species that gene flow can occur over substantial distances (Dow and Ashley 1988; Loveless et al. 1998; Nason et al. 1996; Stacy et al. 1996). In many cases, over a quarter of all successful pollinations involves pollen from outside a population, with pollen often traveling more than one kilometer (Hamrick and Nason 2000). In a study of three hermaphroditic, insect-pollinated tropical tree species, Stacy et al. (1996) found that in 62% of the progeny examined, effective pollen

movement occurred up to 210 m from the pollen donor to the maternal tree for one species (*Calophyllum longifolium*), and in another species (*Spondias mombin*), pollen in limited cases moved distances greater than 300 meters (2.5% and 5.2% of pollination events over two breeding seasons). Similarly, Chase et al. (1996) found in a Costa Rican population of the tree *Pithecellobium elegans* that the average distance pollen moved was 142 m, with a maximum of 350 m. In the tropical tree *Tachigali versicolor*, 25% of the pollen came from more than 500 m (Loveless et al. 1998).

In an illustration of the potential for paternity-based gene flow measures to detect pollination over very long distances, Nason et al. (1996) found that fig wasps in Panama regularly move up to 10 km between flowering fig trees; and three species of *Ficus* in Panama received over 90% of their pollen from more than 1000 m (Nason and Hamrick 1997). White et al. (2002) examined gene flow in several forest fragments of the rare tree species, *Swietenia humilis* in Honduras. Highly variable SSRs allowed them to determine that between 24% and 100% of the pollen traveled more than 900 m. In one case, an isolated tree received 71% of its pollen from over 4.5 km. These studies demonstrate the resolution and certainty attainable with paternity-based gene flow estimates.

Earlier, and still more frequently used, measures of gene flow use neutral or nearly-neutral genetic markers to estimate gene flow indirectly (Bossart and Prowell 1998). These estimates employ measures of genetic structure, most commonly Wright's (1951) F statistics, to describe the genetic structure of a group of populations. Population genetic models then describe the gene flow necessary to produce the observed genetic structure. These techniques are easier to use than direct, paternity-based methods. This

advantage is contrasted, however, by a greater number of assumptions, most notably the need for the counterbalancing forces of gene flow and genetic drift to be in equilibrium.

In this study, we report on patterns of gene flow in the bat-pollinated, neotropical tree species, *Hymenaea courbaril*, in the fragmented dry-forest life zone of southwestern Puerto Rico. Seed dispersers are absent over many parts of the contemporary range of *H. courbaril*, and the consequent limited seed dispersal may increase the possibility of biparental inbreeding (Brown 1989). Bat-mediated long-distance pollen movement may therefore be especially important to maintaining genetic diversity in the progeny of this species.

We estimate gene flow using both direct and indirect measures of gene flow. The presence of migration-drift equilibrium, essential to accurate indirect estimates of gene flow, is assessed. We estimate pollen movement directly into two small stands of trees, and describe the movement of pollen among isolated trees scattered across an old field.

Materials And Methods

Study populations were located in southwestern Puerto Rico, some in low-lying areas and others in the Sierra Bermeja mountain range (maximum height of 200 m to 225 m) (Figure 4.1). The region is classified as subtropical dry forest (sensu Holdridge 1967), receiving less than 1000 mm of rain per year. This forest type is dominant along the southern coast of Puerto Rico, accounting for 14% of the total area of the island (Ewel and Whitmore 1973). The dry forests of Puerto Rico were largely intact until extensive clearing for agriculture and cattle pasture began at the beginning of the 19th

century (Wadsworth 1950). Remnant vegetation is found in scattered patches of forest, strips of vegetation along fencerows and ravines, and isolated trees left in pastures.

Hymenaea courbaril L., known locally as algorrobo, is a tree found on most of the islands of the Lesser and Greater Antilles and throughout Central and northern South America, extending as far south as Southern Brazil (Lee and Langenheim 1975). In Puerto Rico, this large, usually evergreen tree is found throughout forests and pastures of dry and moist coastal regions (Little and Wadsworth 1964), and was a dominant tree in the present study site prior to forest clearing (J. Schwagerl, pers. comm.). At a study site in northwestern Costa Rica, *H. courbaril* was the most abundant flowering resource, but no more than one or two trees flowered simultaneously over the course of a flowering season (Heithaus et al. 1975).

Bawa (1974) reported *H. courbaril* to be self-incompatible based on success of self- versus cross-pollen in hand pollinations of flowers. The lack of fruit-production observed in selfed-flowers may not be due to classic gametophytic self-incompatibility. In a related species, *H. stigonocarpa*, self-pollen was accepted, and pollen tubes grew into the ovary (Gibbs et al. 1999). The resulting ovule, however, remained smaller than out-crossed ovules, and consistently died within 8 days of anthesis. Gibbs et al. (1999) suggest a similar mechanism is likely in *H. courbaril*.

The perfect, apparently protogynous (Crestana and Mariano 1985) flowers of *H. courbaril* are bat pollinated (Carvalho 1961; Crestana and Mariano 1985; Heithaus et al. 1975; Vogel 1958). Among the six bat species in Costa Rica that utilize its flowers, most visited flowers almost exclusively in the dry season (Heithaus et al. 1975), when fruit resources of the landscape were at a minimum. *Hymenaea courbaril* was the fifth most

visited flower for the Jamaican fruit bat, and it was the most-favored flower for *Phyllostomus discolor*. Among all individuals from all six species of bats, 7% carried *H. courbaril* pollen. Hawkmoths were found to regularly visit flowers of the related Brazilian species (*H. stigonocarpa*), but did not appear to make effective contact with either the anthers or the stigma (Gibbs et al. 1999).

Four of the 13 bats reported from Puerto Rico feed on nectar or pollen (Woods 1996). One of the most prevalent bats in Puerto Rico, the Jamaican fruit bat (Woods 1996), is mainly a frugivore, but will use pollen and nectar heavily in the dry season (Heithaus et al. 1975). Bats are capable of flying substantial distances, yet actual distances flown will often depend upon local conditions. Jamaican fruit bats in Mexico will fly 8 km (±2 km) from roosting to foraging areas, whereas conspecifics from Barro Colorado Island in Panama will fly only 0.6 km (±0.4 km)(Morrison 1978).

In Central and South America, seeds are dispersed by agoutis (Asquith et al. 1999), which have been reported to move seeds up to 225 m (Hallwachs 1986), and peccaries (Janzen 1983); animals absent from Caribbean islands. Pods may readily disperse between islands, however, on ocean currents, and viable seeds have been found even after long periods at sea (Gunn 1968; Lee and Langenheim 1975). Mammals of Puerto Rico that may have dispersed seeds in the past include Indian Hutia (*Isolobodon portoricensis*), large, herbivorous rodents (>1 kg) introduced by Amerindians long before the arrival of Europeans; the Puerto Rican Giant Hutia (*Elasmodontomys obliquus*); and two species of Spiny Rats (*Heteropsomys* sp.), which were anatomically, and perhaps ecologically, similar to Hutias (Woods 1996). All have long since been extirpated, although it is believed that there may be relict populations of the Indian Hutia in remote

parts of Puerto Rico (Woods 1996). There are living relatives of Hutias, mainly in Cuba (Alvarez and Gonzalez 1991). Little information is available about foraging habits of these extant members of the capromyid family, although many are known to feed on fruits (e.g., Campbell et al. 1991). It is conceivable, then, that the hutias, and hutia-like mammals, of Puerto Rico may have dispersed the seeds of *Hymenaea*. The only remaining mammals native to the island, however, are bats (Woods 1996), and consequently most seeds do not move far from the parent tree.

Indirect estimates of gene flow

Gene flow was estimated indirectly using eight populations of 21 to 24 trees, and one population (Sant) with 48 trees (Figure 4.1). The minimum distance separating two populations was 0.5 km, and the maximum was 22.8 km. From each tree, at least 20 cm^2 of leaf tissue was collected, and shipped on ice within 48 hours of collection to the laboratory at the University of Georgia. In the lab, leaves were crushed in a potassium phosphate extraction buffer (Mitton et al. 1979) and the resulting extracts were stored on filter paper wicks at -70° C.

Horizontal starch gel electrophoresis was used to assay allozyme diversity. Six buffer systems gel-electrode combinations and 15 enzyme stains were used to resolve 31 loci on 11.0% starch gels. Gel and electrode buffer recipes were from Table 1 in Soltis et al. (1983). The enzymes stained (and loci resolved) on System 4 (a Tris/citrate gel and buffer) were isocitrate dehydrogenase (Idh-1 and Idh-2 and aconitase (Aco-1 and Aco-2). System 11 (a histidine gel and citric acid buffer) was stained for adenylate kinase (Ak-1, Ak-2 and Ak-3) and UTP-glucose-1-phosphate (Ugpp-1 and Ugpp-2). System 6 (a Tris-citric acid buffer gel and sodium hydroxide and boric acid buffer) was used to resolve

phosphoglucoisomerase (Pgi-1 and Pgi-2), alcohol dehydrogenase (Adh), and phosphoglucomutase (Pgm-1 and Pgm-2). A modified System 8 (a discontinuous LiOH system) was used to resolve fluorescent esterase (Fe-2, Fe-3, Fe-4 and Fe-5), triose-phosphate isomerase (Tpi-1, Tpi-2 and Tpi-3), and diaphorase (Dia). System 7 (a Tris-citric acid buffer gel and lithium hydroxide and boric acid buffer) was stained for menadione reductase (Mnr-1 and Mnr-2), malic enzyme (Me) and amino acid transferase (Aat). System MC (a morpholine citrate gel and buffer) was stained for 6-phosphogluconate dehydrogenase (6Pgd-1 and 6Pgd-2) and malate dehydrogenase (Mdh-1, Mdh-2, and Mdh-3). Stain recipes were modified from Soltis et al. (1983) except for diaphorase (Cheliak and Pitel 1984) and UTP-glucose-1-phosphate (Manchenko 1994). For enzymes with more than one locus, isozymes were numbered sequentially, with the lowest number assigned to the most anodal.

Standard measures of genetic variation were calculated, which included percentage of polymorphic loci (*P*), the mean number of alleles per locus (*A*) and per polymorphic locus (*AP*), the effective number of alleles per locus (A_e), observed heterozygosity (H_o) and expected heterozygosity (H_e).

Gene flow estimates, in terms of the effective number of immigrants per generation (N_em), were made with the following equation (Wright 1951):

$$N_em = (1-F_{ST})/4F_{ST}$$

where N_e is the effective number of individuals in the population, m is the immigration rate and F_{ST} is the proportion of total genetic diversity due to differences among

populations (Wright 1951). A key assumption in using this approach is that the populations are in migration-drift equilibrium. This assumption was tested by looking for a statistically significant positive relationship between pair-wise estimates of $F_{ST}/(1-F_{ST})$ and geographic distance separating the pair (Rousset 1997).

Direct measures of gene flow

Two stands of trees were chosen for direct estimates of gene flow from the same geographic location as the populations used for the indirect analysis (Figure 4.1). Distances from each stand to the nearest conspecific tree were 800 m (Manola) and 600 m (Salinas).

Seeds were nick-scarified and planted at the Plant Biology Greenhouses at the University of Georgia in Athens, GA. Germination rates averaged 50-60% and most seedlings emerged within four weeks. Once true leaves appeared, roughly 20 cm^2 of leaf tissue was collected and processed for allozyme analysis as described above.

Multi-locus estimates of outcrossing were computed with MLTR (Ritland 2002). Paternity analysis was used to estimate gene flow levels. Genotypes of all trees from a seedling's stand that could potentially have sired the seed were compared against the genotype of the seedling. An initial estimate of gene flow was made by determining the percentage of seedlings that could not have been sired by potential fathers within the stand.

In genetic trials, an exclusion probability of 0.94 was found based on eleven polymorphic allozyme loci, indicating that for a given seed, 94% of potential sires will be correctly identified as non-sires (Chakroborty et al. 1988). Since a father from outside

the population could produce a pollen gamete indistinguishable from that created by a sire from within the population, however, this initial estimate will underestimate the true gene flow rate. The total number of seeds sired from outside a stand of trees, but seemingly consistent with having been sired from within, comprises "cryptic" gene flow, and its estimation is essential to obtain unbiased estimates of total gene flow.

Cryptic gene flow is difficult to estimate directly. We used the program GFLOW (available from B.K.D.), which employs a slight modification of the technique developed by Devlin and Ellstrand (1990). This technique uses maximum-likelihood to determine the total (apparent + cryptic) gene flow rate that would be most likely to yield the observed apparent gene flow rate (Devlin and Ellstrand 1990). A simulation was run over all possible total gene flow levels, with the minimum value being set by the apparent gene flow rate, and the highest being 100%. For each gene flow level (in increments of 1%), five thousand populations were created. These populations had the same number of adult trees, adult genotypes, and numbers of seedlings for each maternal tree as the actual population. Seedling genotypes were generated using a "pollen gamete" from either a randomly chosen adult from the population (non-gene flow), or from the "pollen pool" (gene flow). The pollen pool represents the theoretical range of likely pollen genotypes that could be produced by trees in the vicinity of the population under study. For this study, the frequency of genotypes in the pollen pool was simply the average frequency of genotypes from the 10 populations sampled for the indirect gene flow study.

Once generated, the apparent gene flow rate was calculated for each of the 5,000 populations. The number of times an apparent gene flow rate of a simulated population

matched that of the actual population was tabulated. Total gene flow rate was then incremented by 1%, and another 5,000 populations were generated.

The total gene flow rate that had the greatest number of matches with the actual apparent gene flow rates was the most likely total gene flow rate. The overall gene flow rate for the whole population was calculated as the weighted average of gene flow for all the trees from the population.

Pollen movement among a group of 21 trees

In a site consisting of twenty-one trees spread over roughly 13 ha (Figure 4.2), seeds were collected from nine trees. Since a paternity analysis involving twenty-one potential sires, given the observed exclusion probability, would leave paternity unassigned for too many trees, the movement of pollen among the 21 trees, rather than pollen movement from outside, was studied. Seedling leaf tissue from was analyzed at the allozyme loci mentioned above. Likelihood scores, generated with CERVUS 2.0 (Marshall et al. 1998), were computed by comparing genotypes of seedlings against those of all twenty-one adult trees. The tree with the highest likelihood score is assigned paternity for the given seedling.

Results

Genetic diversity statistics and population structure

Of the 31 loci surveyed, 11 (35.5%) were polymorphic in at least one of the populations sampled (Table 4.1). At the species level, there were 2.64 alleles per

polymorphic locus (AP_s), the effective number of alleles (A_{es}) was 1.20, and the expected heterozygosity (H_{es}) was 0.118.

On average, 32.5% of loci were polymorphic for the nine populations (Table 4.1), with a narrow range of individual values, from 29.0% (R303B) to 35.5% (TinMid and Hormigueros). There was an average of 1.39 alleles per locus (range 1.32-1.48), and 2.20 alleles at polymorphic loci (range 2.00-2.50). Genetic diversity (H_{ep}) ranged from 0.071 (Bermeja) to 0.132 (Sant), with an average of 0.101 (Table 4.1). No relationship existed between any of the diversity measures and distance to nearest neighbor.

The percent of the variation distributed among populations (G_{ST}) was 7.9% (Table 4.2). There was a narrow range of G_{ST} values, between 0.026 and 0.080, except for PGI1 (0.345). With PGI1 removed, G_{ST} averaged 0.053. There were a substantial deficit of heterozygotes across populations at the PGI1 locus (F_{IT}= 0.312), which was due mainly to a Wahlund effect, rather than a deficit of heterozygotes within populations (F_{IS}= -0.042).

A deficit of heterozygotes, indicated by positive F_{IS} values (Table 4.2), was found for four loci, while seven loci demonstrated heterozygote excess. The mean value across loci (F_{IS}=-0.051) is not significantly different from zero (X^2=2.32; P>0.10), however, indicating a match to Hardy-Weinberg expectations. This is also demonstrated by the close match of observed heterozygosity (0.104) to expected (0.101). With all nine populations considered together, there were fewer heterozygotes than would be expected from Hardy-Weinberg expectations (F_{IT}=0.032), due primarily to a Wahlund effect (G_{ST}=0.079). A X^2 analysis for allele frequency heterogeneity among populations

revealed significant differences among populations for 10 of 11 polymorphic loci (all except TPI3).

Of 90 fixation indices, five were significantly different from zero ($p<0.05$): two in Sant (FE2, MNR2); one each in Hormigueros (IDH2), TinLow (FE2) and TinHigh (PGM2). Since there were 4.5 significant differences expected by chance alone, there is no biological meaning to the five significant differences observed.

Direct measures of gene flow

The multi-locus estimate of outcrossing (t_m) for the species was 1.065 (s.d.=0.156). The single-locus estimate was 1.100 (s.d.=0.208). Taken together, there is no evidence for biparental inbreeding (t_m-t_s=–0.035, s.d.= 0.093).

Apparent gene flow rates were 26% (Salinas) and 43% (Manola), while total gene flow rates were 42% (Salinas) and 62% (Manola)(Table 4.3). The stand with the higher gene flow rate, Manola, was actually farther from neighboring populations than the Salinas population (800 m vs. 600 m). The nearest stand to Manola had more trees at a greater density, however, whereas the nearest neighbors for the Salinas population were fewer and more scattered. Estimates of N_em fell below 1.0 (0.40 for Manola and 0.63 for Salinas) due to the low N_e values for the two populations (1.29 and 1.30, respectively). Population-genetic theory therefore predicts that over time both stands of trees will diverge genetically from surrounding populations (Wright 1931).

Indirect measures of gene flow

The majority of the pairwise F_{ST} values were below 0.015 (Figure 4.3). The scatter of points above 0.015 almost all involved the Sant population. The relationship between pairwise $F_{ST}/(1- F_{ST})$ and interpopulation distances was positive, but not significantly different from zero (Mantel test; P=0.258), which would be expected under migration-drift equilibrium.

The indirect estimate of gene flow (Nm) was 2.91. This value suggests, in contrast to predictions from direct estimates of N_em, that gene flow will homogenize populations over time, albeit to a limited degree (Wright 1931).

Paternity assignment in Tex population

Genotypes for 337 seedlings from nine maternal trees were obtained. Unique paternity was assigned to 223 seedlings.. Of twelve trees which were found to be pollen donors, 59% of the pollen came from only three trees (8, 9, and 16; Figure 4.2), with 26% coming from tree #8 alone. This disproportionate production of pollen lowered the effective number of sires to 6.9. Pollen did not always come from the nearest neighboring tree. On average, there were 9.1 trees whose distance to a maternal tree was shorter than the distance between the maternal tree and the tree that actually sired a given seed.

Discussion

The analysis of pollen movement within the Tex population demonstrated that pollinations rarely involved nearest neighbors. This may explain why there were no signs of biparental inbreeding, a phenomenon often seen in species with gravity-dispersed seeds (Brown 1989). Chase et al. (1996) noted a similar occurrence in *Pithecellobium elegans* in Costa Rica, where the average distance to known fathers was 142 m, despite an average distance of 27 m separating trees. Just as *H. courbaril* trees received most of their pollen from an average of 9.1 trees away, seeds of *P. elegans* likewise received the majority of pollen contributions from more than 10 trees away.

In the Tex population, most pollen leading to successful pollinations did not come from a tree's nearest neighbor. Non-overlapping flowering phenology was suggested by Chase et al. (1996) to be the reason why pollen received by *P. elegans* did not usually come from nearest neighbors. This is probably also the case for *H. courbaril*, for which a strong asynchrony in flowering times was noted by Heithaus et al. (1975) at a study site in northwestern Costa Rica.

Three trees in the Tex population did the majority of the pollinations, leading to an effective number of sires of only 6.9. This disproportionate production of pollen by a few trees, lowering effective population size, may be a reason why genetic diversity is lower in *H. courbaril* than in other tree species with similar life history characteristics (Hamrick and Godt 1996).

The two stands of *H. courbaril* in which paternity analysis was conducted tended to have lower gene flow rates than those estimated in two other species of trees, *Bursera simaruba* (L.) Sarg. and *Albizia lebbek* L., studied in the same area of southeastern Puerto

Rico (Table 4.3). The value for Salinas (42%) was the lowest for any population of the three species examined, and only two populations, one each from *B. simaruba* and *A. lebbek*, had lower gene flow rates than Manola (62%). These results are surprising since bats, the pollinator for *H. courbaril*, would be expected to distribute pollen over a greater distance than the small insects which are pollinators for *B. simaruba*, or the bees and butterflies that are pollinators for *A. lebbek*.

It is possible that the long-distance flight capabilities of bats may not always lead to higher levels of gene flow than those seen with other pollinators. Most of the long-distance flights reported for bats are from roosting to feeding sites (e.g., Lumsden et al. 2002). Food in the vicinity of the day roost is often ignored (Handley and Morrison 1991), so these long flights may not involve the movement of pollen. If new trees are visited on subsequent feeding excursions from the day roost, pollen from previous nights may not be present, since bats groom extensively while roosting (Heithaus et al. 1975).

Bats tend to feed within the same area, often from just a few trees (e.g, Handley et al. 1991). Foraging groups (numbering 10-15) of the Brazilian bat, *Phyllostomus discolor*, spend between 10-20 minutes on trees with many (>60) flowers (Sazima and Sazima 1977). This extended foraging time per tree may reduce long-distance pollen movement, and potentially increase the risk of geitonogamous matings in self-compatible species. Jamaican fruit bats feeding on fig trees on Barro Colorado Island in Panama return to the same tree for feeding for up to eight (4.3 ± 1.8) consecutive nights (Handley and Morrison 1991), often bypassing other trees in fruit in favor of those near a familiar feeding roost. This roost is usually 25 to 200 m from the tree from which fruit is being gathered. When switching to new trees, bats often visit areas that had been explored

earlier. Handley and Morrison (1991) suggest that Jamaican fruit bats may be creatures of habit that prefer familiar sites along well-known flight paths. These bats are capable of moving around all of Barro Colorado Island, but radio-tracking showed that individual bats had different parts of the island that they chose to utilize (Handley et al. 1991). In Costa Rica, mist netting of bats suggested that three of six species (including *A. jamaicensis*) are consistent in their use of home ranges, which appear often to be confined to single rivers (Heithaus et al. 1975).

Even if bats were to distribute pollen evenly among trees, it may be that there are not enough simultaneously flowering *H. courbaril* trees to allow high levels of gene flow. In northwestern Costa Rica, no more than two *H. courbaril* trees were found to flower at the same time over the course of a flowering season (Heithaus et al. 1975). Therefore, the nearest flowering neighbor may be much further than the nearest conspecific tree. This would help explain the low effective number of fathers at the Tex site (6.9). The problems presented by a low number of flowering trees would be exacerbated by the low abundance of *H. courbaril* trees in the region, relative to *B. simaruba* and *A. lebbek*.

Although stands of *H. courbaril* received foreign pollen, gene flow into smaller stands was low enough that genetic drift may cause further loss of diversity over time. This, combined with limited seed movement, may cause the loss of smaller stands of trees and further genetic isolation of the remaining populations.

Indirect measures of gene flow

Combined with effective population sizes, there was an estimate of 0.40 migrants per generation (N_em) for Manola, and 0.63 for Salinas. These estimates, each being less

than one, suggest that these small stands of trees will be more genetically dissimilar to surrounding populations over time. The indirect estimate of gene flow (Nm=2.91) falls within the range of 1-4 migrants per generation, where theory can not accurately predict future changes (Wright 1931). The opposite pattern was seen in the introduced species, *A. lebbek*, growing in the same region, where direct estimates of N_em were between 1 and 4 (2.08 and 3.04), and the indirect estimate was below one (Nm=0.69). Only in *B. simaruba*, which demonstrated isolation by distance (albeit on a larger scale than the one used to indirectly estimate gene flow), did predictions based on direct and indirect estimates of Nm approximately match. Four of five estimates of N_em determined directly (1.49, 1.50, 1.20, 1.03) yielded predictions similar to that of the indirect estimate (3.57). The lone value outside of 1-4 (N_em=0.36) occurred in a small stand of three trees, where a pollen flow rate of 47% limited N_em to a maximum possible value of 0.71.

It must be pointed out that most of the stands of trees for which direct measures were made had fewer than eight trees, the minimum number needed to obtain an N_em estimate of 4 (assuming N_e=8 and a 100% pollen flow rate), the value above which gene flow will act to homogenize populations. It is possible that seed flow could increase overall gene flow (i.e., m), especially in *A. lebbek*, which has wind-dispersed pods, and *B. simaruba*, which has bird-dispersed fruits, but probably not enough to alter predictions of future population change. Given that substantial amounts of pollen are foreign, it might be expected that N_em would be greater than four in larger stands of trees. If m and the ratio of effective to actual population size (N_e/N) were independent of population size, then 20 trees in the Manola stand, and 19 trees in the Salinas stand would yield N_em values of 4. If gene flow rates decrease with increasing population size, as they did in a

rare mahogany species (*Swietenia humilis*) in Honduras (White et al. 2002) and in a tree from Panama (Nason and Hamrick 1997), then N_em may stay low, unless the decrease in gene flow rate with increasing population size is small.

Since *Hymenaea* has a long geological history in the Greater Antilles (Lee and Langenheim 1975), enough generations have certainly passed to allow migration-drift equilibrium to develop. A possibility to consider is that migration-drift equilibrium is present, but on a larger scale than used in the present study, as was the case for *B. simaruba* (B. Dunphy, unpubl. data).

Genetic diversity levels

Hymenaea courbaril had a lower H_e (0.118 vs. 0.152) and *P* (35.5% vs. 50.2%) than other outcrossing species with gravity-dispersed fruits (Hamrick and Godt 1996). There was a lower G_{ST} in *H. courbaril* (0.079 vs. 0.189) which would suggest higher gene flow levels, or that the distances separating the *H. courbaril* populations in this study are less than those separating populations of the other species. Diversity values were also lower than those of *B. simaruba* (H_e=0.244; *P*=73.3%) and *A. lebbek* (H_e=0.266; *P*=66.7%).

A relatively low population density may contribute to the low genetic diversity values of *H. courbaril*. With the exception of the Sant population, stands of *H. courbaril* tend to be small (<30 trees) and well separated, with few intervening trees. With a lower gene flow rate than that of *B. simaruba* or *A. lebbek*, there are smaller effective population sizes, leading to both a faster decrease of genetic diversity and a faster loss of polymorphic loci (lost at a rate 1/2 N_e; Kimura and Ohta 1969).

Although pollen flow accounts for the majority of gene movement (Ennos 1994), the minimal amount of seed movement in *H. courbaril* may be an additional factor in the species' low genetic diversity values. In Central America, where seed dispersers are more abundant, however, the genetic diversity values were actually lower (J. Hamrick, unpubl. data), suggesting that a factor other than seed dispersal (e.g., history) is responsible. Nevertheless, this limited ability to disperse its seeds could greatly limit the ability of *H. courbaril* to spread into new habitats. In the last few decades, forested areas in Puerto Rico have increased as old pastures and farmland are left fallow (Aide et al. 1995). Limited dispersal abilities may hamper the ability of *H. courbaril* to colonize this newly available territory. Species spread will occur as a slow-moving front, and the presence of even minimal barriers (e.g., a wide road) may severely hamper the species' expansion. Human intervention may be essential to re-introduce the species into parts of its former range.

Although a number of mammals feed on the pods, only agoutis are considered effective seed dispersers in parts of the range where mammalian seed predators are present (Asquith et al. 1999). They bury scatter-hoarded pods and seeds, which lead to higher germination rates than those for seeds left unburied. Where agoutis are absent, pods tend to remain below the tree, and seedlings are rare (Asquith et al. 1999). When present, seedlings often have defective root systems (Janzen 1983). In areas with seed predators, Hallwachs (1986) predicts that *H. courbaril* will go locally extinct without agoutis to bury the pods and seeds.

Two primate species introduced in the 1970s, Rhesus monkeys (*Macaca mulatta*) and African Patas monkeys (*Erythrocebus patas*), are reported to feed on *H. courbaril*

(González-Martínez 1995). In Costa Rica, capuchin monkeys break open and disperse seeds (Oliveira et al. 1995). These seeds are left on the soil surface, however, where they are exposed to mammalian seed predators. The degree to which Rhesus and Patas monkeys effectively disperse seeds, then, may depend upon the degree of predation from introduced mammals, most notably black rats (*Rattus rattus*) and Norwegian rats (*R. norvegicus*).

Summary

Although stands of trees received appreciable amounts of external pollen, gene flow levels were still at the low end of what has been seen in other tropical tree species. This, combined with low seed dispersal and uneven pollen production, may have led to the relatively low genetic diversity values measured in *Hymenaea courbaril*. Over time, diversity may be lost to genetic drift in smaller populations. This loss can be minimized if recruitment increases population sizes and if new stands of trees become established, either naturally or via human mediation, in the landscape separating current populations. Recent increases in fallow land in Puerto Rico may facilitate this process.

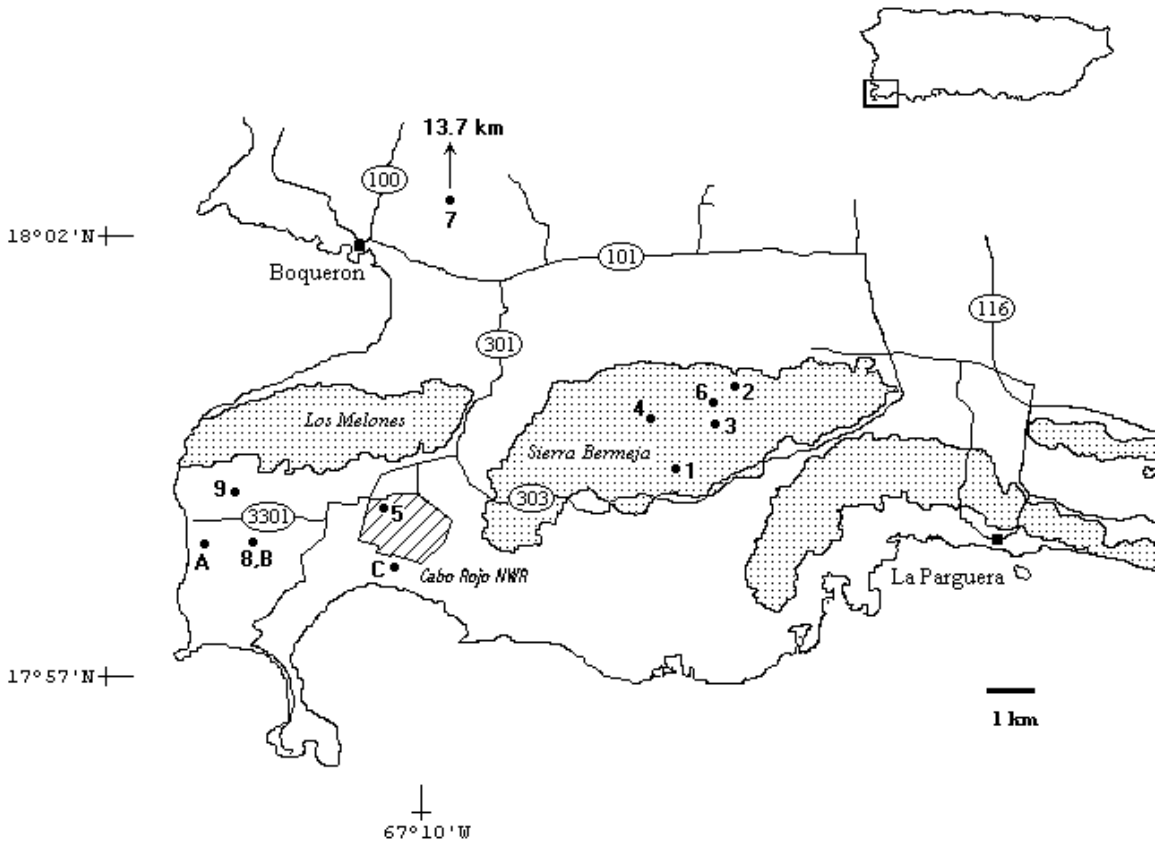


Figure 4.1. Location of study populations of *Hymenaea courbaril* in southwestern Puerto Rico. Populations of adult trees used in indirect estimates of gene flow are numbered 1-9. The three stands of trees used for direct measures of gene flow are lettered A-C. Stippled regions are low mountains (maximum height of 200-225 meters). Diagonal lines show the location of the Cabo Rojo National Wildlife Refuge.

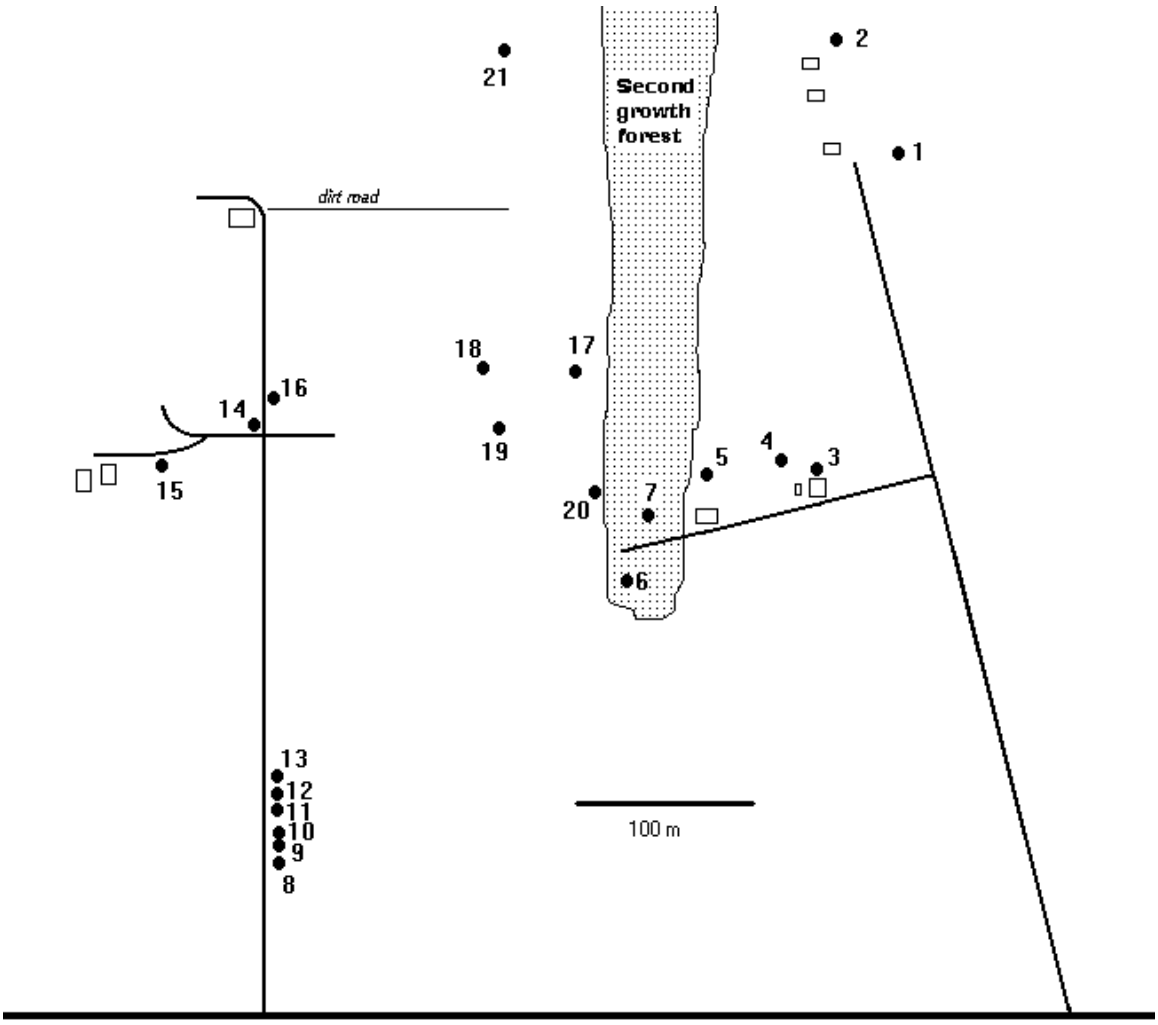


Figure 4.2. Location of *Hymenaea courbaril* trees in the Tex study site in southwestern Puerto Rico.

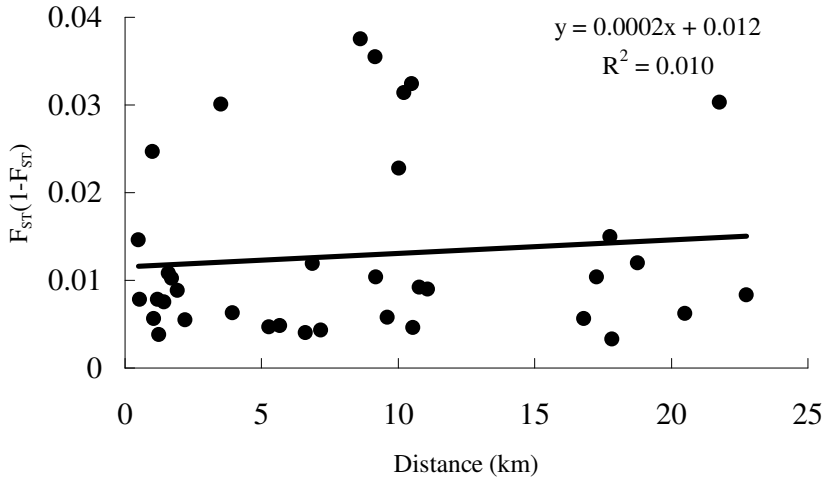


Figure 4.3. Examination of isolation by distance in all possible pairs of nine *Hymenaea courbaril* populations from southwestern Puerto Rico. A significantly positive slope would have indicated migration-drift equilibrium.

Table 4.1. Genetic diversity statistics[a] for nine populations of *Hymenaea courbaril* in southwestern Puerto Rico.

Population	N	P_P	AP_P	A_P	A_{ep}	H_O	(SD)	H_e	(SD)
1. R303	24	29.03	2.22	1.35	1.16	0.103	0.048	0.091	0.031
2. TinLow	24	32.26	2.10	1.35	1.19	0.073	0.047	0.101	0.034
3. TinHigh	24	32.26	2.20	1.39	1.18	0.108	0.050	0.105	0.033
4. Bermeja	24	30.00	2.22	1.37	1.10	0.068	0.046	0.071	0.024
5. Refuge	24	33.33	2.00	1.33	1.16	0.111	0.055	0.098	0.029
6. TinMid	24	35.48	2.18	1.42	1.14	0.087	0.048	0.086	0.029
7. Hormigueros	24	35.48	2.36	1.48	1.20	0.112	0.053	0.113	0.033
8. Tex	21	32.26	2.00	1.32	1.18	0.116	0.057	0.110	0.032
9. Sant	48	32.26	2.50	1.48	1.25	0.158	0.037	0.132	0.038
MEAN		32.49	2.20	1.39	1.17	0.104		0.101	
SD		2.81	0.16	0.06	0.04	0.016		0.011	
Species estimate	237	35.48	2.64	1.58	1.20			0.118	

[a] *P* is the percentage of polymorphic loci, *AP* is the mean number of alleles per polymorphic locus, *A* is the mean number of alleles per locus, A_e is the effective number of alleles, H_O is the observed heterozygosity and H_e is gene diversity, or expected heterozygosity. The "p" subscript indicates population means.

Table 4.2. Estimates of genetic diversity parameters[a] for polymorphic loci surveyed for nine populations of *Hymenaea courbaril*. Mean values are based on polymorphic loci.

Locus	H_T	H_S	F_{IS}	G_{ST}
6PGD2	0.507	0.491	0.014	0.032
ACO1	0.215	0.204	-0.111	0.048
ADH1	0.484	0.448	0.018	0.075
FE2	0.496	0.472	-0.204	0.049
FE4	0.430	0.396	-0.061	0.080
IDH2	0.081	0.077	0.016	0.045
MNR2	0.468	0.442	-0.007	0.056
PGI1	0.371	0.243	-0.042	0.345
PGI2	0.191	0.178	-0.109	0.069
PGM2	0.327	0.312	0.003	0.046
TPI3	0.086	0.084	-0.075	0.026
Avg:	0.332	0.304	-0.051	0.079

[a] H_T, total genetic diversity; H_S, genetic diversity found within populations; F_{IS}, deviations from Hardy-Weinberg expectations within individual populations; G_{ST}, the proportion of total genetic diversity found among populations.

Table 4.3. Estimated total gene flow rates (with standard deviation in parentheses) and migration rates (N_em) for *Hymenaea courbaril* and two other tree species (*Bursera simaruba* and *Albizia lebbek*) in southwestern Puerto Rico.

Population	# seedlings analyzed	Isol. dist. (m)[a]	Total gene flow rate	N_em
H. courbaril				
Manola (n=4)	66	800	62% (11%)	0.40
Salinas (n=6)	102	600	42% (11%)	0.63
B. simaruba[b]	629	300->1000	73% (10%)	1.24
A. lebbek[c]	655	60-800	79% (12%)	2.32

[a] Distance from a tree to the nearest conspecific tree from outside its population.

[b] Five populations, with between three and nine trees per population.

[c] Four populations, with between two and eleven trees per population.

References

Aide TM, Zimmerman JK, Herrera L, Rosario M, Serrano M (1995). Forest recovery in abandoned tropical pastures in Puerto Rico. Forest Ecology and Management 77: 77-86.

Alvarez VB, Gonzalez A (1991). The critical condition of hutias in Cuba. Oryx 25: 206-208.

Asquith NM, Terborgh J, Arnold AE, Riveros CM (1999). The fruits the agouti ate: *Hymenaea courbaril* seed fate when its disperser is absent. Journal of Tropical Ecology 15: 229-235.

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

Bossart JL, Prowell DP (1998). Genetic estimates of population structure and gene flow: limitations, lessons, and new directions. Trends in Ecology and Evolution 13: 202-206.

Brown AHD (1989). Genetic characterization of plant mating systems. In: Brown AHD, Clegg MT, Kahler AL, Weir BS (eds) *Plant population genetics, breeding and genetic resources*, Sinauer: Sunderland, Massachusetts. Pp 145-162.

Campbell DG, Lowell KS, Lightbourn ME (1991). The effect of introduced hutias (*Geocapromys ingrahami*) on the woody vegetation of Little Wax Cay, Bahamas. Conservation Biology 5: 536-541.

Carvalho CT de (1961). Sobre os hábitos alimentares de Phillostomídeos (Mammalia, Chiroptera). Revista de Biologia Tropical 9: 53-60.

Chakroborty P, Meagher TR, Smouse PE (1988). Parentage analysis with genetic markers in natural populations. 1. Paternity exclusion and expected proportions of offspring with unambiguous paternity. Genetics 118: 527-536.

Chase MR, Moller C, and Kesseli R, Bawa KS (1996). Distant gene flow in tropical trees. Nature 383: 398-399.

Cheliak WM, Pitel JA (1984). *Techniques for starch gel electrophoresis of enzymes for forest tree species*. Petawawa National Forestry Institute, Information Report P1-X-42. Canadian Forestry Service, Agriculture: Chalk River, Ontario.

Crestana CSM, Mariano IS (1985). Ecologia de polinizacão de *Hymenaea stilbocarpa* Hayne, o Jatobá. Silvicultura São Paulo 17/19: 31-37.

Devlin B, Ellstrand NC (1990). The development and application of a refined method for estimating gene flow from angiosperm paternity analysis. Evolution 44: 248-259.

Dow BD, Ashley MV (1998). High levels of gene flow in bur oak revealed by paternity analysis using microsatellites. Journal of Heredity 89: 62-70.

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

Ewel JJ, Whitmore JL (1973). *The ecological life zones of Puerto Rico and the U.S. Virgin Islands*. U.S.D.A. Forest Service, Institute of Tropical Forestry: Rio Piedras, Puerto Rico.

Gibbs PE, Oliveira PE, Bianchi MB (1999). Postzygotic control of selfing in *Hymenaea stigonocarpa* (Leguminosae-Caesalpinioideae), a bat-pollinated tree of the Brazilian Cerrados. International Journal of Plant Sciences 160: 72-78.

González-Martínez J (1995). Ecology of the introduced free-ranging patas and rhesus monkeys of southwestern Puerto Rico. PhD Dissertation, University of Colorado.

Gunn R (1968). Stranded seeds and fruits from the southeastern shore of Florida. Garden Journal (March/April)43: 54.

Hallwachs W (1986). Agoutis *(Dasyprocta punctata*): the inheritors of guapinol (*Hymenaea courbaril*: Leguminosae). In: Estrada A and Fleming TH (eds) *Frugivores and seed dispersal*, W. Junk Publishers: Dordrecht, Netherlands. Pp 285-304.

Hamrick JL, Godt MJW (1996). Effects of life history traits on genetic diversity in plant species. Philosophical Transactions of the Royal Society of London Series B 351: 1291-1298.

Hamrick JL, Nason JD (2000). Gene flow in forest trees. In: Boyle TJB, Young A, Boshier D (eds) *Forest Conservation Genetics: Principles and Practice*, CIFOR and CSIRO: Australia. Pp 81-90.

Handley CO Jr, Gardner AL, Wilson DE (1991). Core Areas. In: Handley CO Jr, Wilson DE, Gardner AL (eds) *Demography and Natural History of the Common Fruit Bat,* Artibeus jamaicensis*, on Barro Colorado Island, Panama*, Smithsonian Inst. Press: Washington, DC. Smithsonian Contributions to Zoology No 511, pp 89-130.

Handley CO Jr, Morrison WM (1991). Foraging Behavior. In: Handley CO Jr, Wilson DE, Gardner AL (eds) *Demography and Natural History of the Common Fruit Bat,* Artibeus jamaicensis*, on Barro Colorado Island, Panama*, Smithsonian Inst. Press: Washington, DC. Smithsonian Contributions to Zoology No 511, pp 137-140.

Heithaus ER, Fleming TH, Opler PA (1975). Foraging patterns and resource utilization in seven species of bats in a seasonal tropical forest. Ecology 56: 841-854.

Heithaus ER, Stashko E, Anderson PK (1982). Cumulative effects of plant-animal interactions on seed production by *Bauhinia ungulata*, a neotropical legume. Ecology 63: 1294-1302.

Holdridge LR (1967). *Life zone ecology.* Tropical Science Center: San José, Costa Rica.

Janzen DH (1983). *Hymenaea courbaril* (Guapinol, Stinking toe). In: Janzen DH (ed) *Costa Rican natural history*, University of Chicago Press: Chicago. Pp 253-256.

Kimura M, Ohta T (1969). The average number of generations until fixation of a mutant gene in a finite population. Genetics 61: 763-771.

Law BS, Anderson J, Chidel M (1999). Bat communities in a fragmented forest landscape on the south-west slopes of New South Wales, Australia. Biological Conservation 88: 333-345.

Law BS, Lean M (1999). Common blossom bats (*Syconycteris australis*) as pollinators in fragmented Australian tropical rainforest. Biological Conservation 91: 201-212.

Lee YT, Langenheim JH (1975). A systematic revision of the genus *Hymenaea* (Leguminosae; Caesalpinioideae; Detarieae). University of California Publications in Botany 69: 1-109.

Little EL, Wadsworth FH (1964). *Common trees of Puerto Rico and the Virgin Islands*. USDA Forest Service, Agr. Handbook 249.

Loveless MD, Hamrick JL, Foster RB (1998). Population structure and mating system in *Tachigali versicolor*, a monocarpic neotropical tree. Heredity 81: 134-143.

Lumsden LF, Bennett AF, Silins JE (2002). Location of roosts of the lesser long-eared bat *Nyctophilus geoffroyi* and Gould' s wattled ba*Chalinolobus gouldii* in a fragmented landscape in south-eastern Australia. Biological Conservation 106: 237-249.

Manchenko GP (1994). *Handbook of detection of enzymes on electrophoretic gels*. CRC Press: Ann Arbor, MI.

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

Mitton JB, Linhart YB, Sturgeon KB, Hamrick JL (1979). Allozyme polymorphisms detected in mature needle tissue of ponderosa pine. Journal of Heredity 70: 86-89.

Morrison DW (1978). Influence of habitat on the foraging distances of the fruit bat, *Artibeus jamaicensis*. Journal of Mammalogy 59: 622-624.

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

Nason JD, Herre EA, Hamrick JL (1996). Paternity analysis of the breeding structure of strangler fig populations: evidence for substantial long-distance wasp dispersal. Journal of Biogeography 23: 501-512.

Oliveira PS, Galetti M, Pedroni F, Morellato LPC (1995). Seed cleaning by *Mycocepurus goeldii* ants (Attini) facilitates germination in *Hymenaea courbaril* (Caesalpiniaceae). Biotropica 27: 518-522.

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

Rousset F (1997). Genetic differentiation and estimation of gene flow from F-statistics under isolation by distance. Genetics 145: 1219-1228.

Sazima I, Sazima M (1977). Solitary and group foraging: two flower-visiting patterns of the Lesser Spear-Nosed Bat (*Phyllostomus discolor*). Biotropica 9: 213-215.

Schnabel A, Hamrick JL (1995). Understanding the population genetic structure of *Gleditsia triacanthos* L.: The scale and pattern of pollen gene flow. Evolution 49: 921-931.

Soltis DE, Haufler CH, Darrow DC, Gastony GJ (1983). Starch gel electrophoresis of ferns: a compilation of grinding buffers, gel and electrode buffers, and staining schedules. American Fern Journal 73: 9-27.

Stacy EA, Hamrick JL, Nason JD, Hubbell SP, Foster RB, Condit R (1996). 1996. Pollen dispersal in low-density populations of three neotropical tree species. American Naturalist 148: 275-298.

Vogel S (1958). Fledermausblumen in Sudamerika. Osterreichische botanische zeitschrift 104: 491-530.

Wadsworth FH (1950). Notes on the climax forests of Puerto Rico and their destruction and conservation prior to 1900. Caribbean Forester 11: 38-47.

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

Woods CA (1996). The land mammals of Puerto Rico and the Virgin Islands. In: Figueroa-Colon JC (ed) *The Scientific Survey of Puerto Rico and the Virgin Island. An eighty-year reassessment of the Islands' natural history*, Annals of the New York Academy of Sciences: New York. Vol 776, pp 131-148.

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

Wright S (1951). The genetical structure of populations. Annals of Eugenics 15: 323-354.

CHAPTER 5

CONCLUSIONS

Direct gene flow estimates

All three species show an almost total absence of selfing. The mechanism for *A. lebbek* is unknown, but the almost total absence of selfing (t_m=0.985) and the increase in seed abortion in smaller stands of trees suggests that *B. simaruba* may have a post-fertilization self-incompatibility mechanism. Bawa (1974) reported that *H. courbaril* was self-incompatible, a conclusion supported by our data. As in *B. simaruba*, the mechanism in *H. courbaril* is not a traditional gametophtyic self-incompatibility system, but rather post-zygotic in nature (Gibbs et al. 1999). White et al. (2002) suggested that self-incompatible species may be more tolerant of habitat fragmentation than mixed-mating species, since the latter could have an increase in selfed seeds as population sizes decrease and become more geographically isolated. The benefit of forced outcrossing (assuming that enough immigrant pollen is available), needs to be weighed against the potentially high loss of reproductive output through aborted seeds. This point was illustrated by the high loss of seeds in small stands of *B. simaruba*. Similarly, small, isolated populations of the Panamanian tree, *Spondias mombin*, had lower fruit production and higher rates of seed abortion than larger populations in continuous forest (Nason and Hamrick 1997).

Direct estimates of gene flow for all three species were relatively high, ranging between 42% and 100%. The bat-pollinated native tree species, *H. courbaril*, tended to

have the lowest estimates. These results are surprising since bats, the pollinator for *H. courbaril*, would be expected to distribute pollen over a greater distance than the small insects which are pollinators for *B. simaruba*, or the bees and butterflies that are pollinators for *A. lebbek*. It appears that predictions of relative gene flow rates can be in error if too much reliance is placed on a priori knowledge of pollinator flight capacities. It is possible that the specific flight patterns of bats during feeding bouts may not move pollen as much as would be indicated by measures of their long distance flights from roosting to feeding areas.

In none of the three species was a relationship seen between gene flow and distance to the nearest neighbor. The only significant relationship that was detected involved stand size and germination rates in *B. simaruba*, where germination rates of seeds, normally between 20-41%, dropped below 1% in four stands of three trees each. This suggests that larger stands may receive more foreign pollen, thus minimizing inbreeding or avoiding problems of self-incompatibility.

None of the populations examined, from all three species, had an absence of foreign pollen, even when the nearest neighbor was over 1 km away (e.g., R307 in *B. simaruba*). Pollen flowed readily, with the lowest gene flow rate being 42% (*H. courbaril*). These must be considered effective gene flow rates, however, especially in the case of *Bursera simaruba*, where in stands of less than four trees the majority of seeds did not germinate, indicating that amounts of cross-pollination may be very low, relative to self-pollination. Similarly, gene flow into small stands of *H. courbaril*, although high, was not high enough to prevent future loss of genetic diversity, due to small effective

population sizes. Combined with limited seed movement, this loss of diversity may lead to the loss of smaller stands of trees, further isolating the remaining populations.

Indirect gene flow measures

Migration-drift equilibrium, a requirement for the use of indirect measures of gene flow, was not detected in any of the three species. There was a slightly positive relationship between genetic and geographic distances in *B. simaruba*, but the relationship was not significant. In *A. lebbek*, equilibrium may have been disrupted by the recent arrival of foreign genotypes into the area of the study site, suggesting that not enough generations have passed for equilibrium to be achieved. The main factor, however, for all three species may have been the spatial scale of the study. In *B. simaruba*, for instance, migration-drift equilibrium first shows up when population pairs are separated by 25-30 km, just beyond the greatest separation distances used in this study (B. Dunphy, unpubl. data).

Unlike *B. simaruba* and *H. courbaril*, for which enough time presumably has passed for migration-drift equilibrium to develop, it is possible that given enough time, equilibrium may develop in *A. lebbek*. This possibility is contradicted by observed patterns of variation, however. The first sign of equilibrium is a high degree of relatedness among nearest-neighboring populations (Slatkin 1993), a pattern not apparent in *A. lebbek*. This is consistent with what was noted by Rousset (1997), who points out that populations separated by small geographic distances may not follow expectations of the model used to determine isolation by distance.

Direct vs. indirect estimates of Nm

Despite trouble with migration-drift equilibrium, there was frequent agreement between direct and indirect measures in regards to predictions of future genetic change. In *B. simaruba*, four of five direct estimates of N_em were between 1 and 4, indicating ambiguity as to future genetic change. The indirect measure yielded the same finding. The same prediction was found for direct and indirect measures in *A. lebbek*, given that genotypes that appear to have recently arrived are removed from the analysis. Only in the native tree *H. courbaril* did predictions differ, where direct estimates indicate that genetic drift will decrease genetic diversity over time, while predictions from the indirect measures are less certain.

Although pollen movement is considered to impact gene movement much more than seed movement (Ennos 1994), it is possible that seed movement may increase N_em enough to alter predictions of future genetic change. More work would be needed to characterize seed movement in the species studied. More work is also needed to characterize how gene flow changes with increasing population size, specifically whether N_em rises above 4 (*B. simaruba* and *A. lebbek*) or 1 (*H. courbaril*) as population size increases.

Comparison of diversity statistics to other woody species

Diversity statistics, especially genetic diversity values (H_e), for two of the species (*Bursera simaruba* and *Albizia lebbek*) were higher than those reported for other species with similar life history traits (Hamrick and Godt 1996). In the case of *A. lebbek*, an old-world species introduced in historic times to Puerto Rico, multiple introductions are

likely responsible for the high genetic diversity values. Very high gene flow levels in the native tree *B. simaruba*, shown both directly and suggested by low G_{ST} values, act to maintain high H_e values. There is a possibility that the low G_{ST} values in *B. simaruba* may also be due to the chosen spatial scale of the study. Pollen, and possibly seeds, may flow readily over the distances involved (ca. 25 km). A collection of populations from over a broader geographic range may be needed to make valid comparisons of G_{ST} with other species.

Diversity statistics for *Hymenaea courbaril* were lower than those reported for other species with similar life history traits. Low G_{ST} values suggest that decreased pollen movement is not the cause of the low diversity values. Instead, the lower values may be due to demographic factors, including a relatively low population density and low numbers of simultaneously flowering trees. In a population of 21 trees where pollen movement was closely analyzed, almost all pollen came from just three trees. This imbalance in pollen production would lead to lower effective population sizes, leading to a more rapid loss of genetic diversity over time.

Future genetic changes

Fragmentation does not hinder pollen movement in any of the three species examined, although ecological factors, especially in *H. courbaril*, may threaten their continued survival. Estimates of N_em suggest that small stands of *H. courbaril* will lose diversity over time, but yield ambiguous predictions about future change in *B. simaruba* and *A. lebbek*.

High levels of genetic diversity and relatively low G_{ST} values in *B. simaruba* suggest that this species tolerates fragmentation well. Despite decreased seed production in very small stands of trees, the promotion of long-distance pollen movement by generalist pollinators may explain why *B. simaruba* can tolerate habitat fragmentation. White et al. (2002) suggest that obligate outcrossers will endure fragmentation better than species following a mixed-mating strategy because of this forced reliance on pollen from outside sources. A possibility to consider in regards to the high diversity levels, however, is that the populations under study may not have undergone enough generations of drift since fragmentation to appreciably decrease genetic diversity, considering that the fragmentation of Puerto Rican forests occurred mainly in the 19th century (Wadsworth 1950). A key factor maintaining genetic diversity in a fragmented habitat may be the ability of seed-dispersers to move substantial amounts of seeds around the landscape, as was suggested by the recent establishment of the Refuge population in an area that was devoid of trees 40 years ago. Low seed production in smaller stands of trees could pose a threat to species spread in *B. simaruba*, although facultative apomixis may allow continued spread.

Albizia lebbek has a number of reproductive characteristics that make it a good invader, including even and large fruit production, large amounts of seed, and fruits that remain attached to the tree for long periods of time. The high levels of genetic diversity in *A. lebbek*, possibly due to multiple introductions, may provide this already-successful invasive species the genetic diversity to adapt to new habitats. High gene flow also may produce larger fruit crops, supplying many seeds for further species expansion. Over time, the high rates of gene flow observed will likely homogenize populations, leading to

a decrease in the very high current estimates of G_{ST}. Based upon direct estimates of N_em, G_{ST} should reach an equilibrium value of 0.076 (Manola) or 0.125 (Tex). These values are close to 0.094, the value of G_{ST} determined by indirect measures.

In *H. courbaril*, small population sizes, limited recruitment, asynchronous flowering and localized seed dispersal may threaten the continued survival of the species in the area of the study site. The degree to which recently introduced mammals, notably monkeys, can disperse seeds, in contrast to the impact of introduced seed predators (e.g., rats), will influence the fate of established populations, as well as the spread of the species. Unlike *B. simaruba*, which seems to tolerate fragmentation, and *A. lebbek*, which prefers disturbed habitats, *H. courbaril* in Puerto Rico may benefit substantially from the increase in fallow farmland in Puerto Rico. Forested land is increasing (although so are urbanized areas), relative to what was present in the beginning of the 20th century (Aide et al. 1995). Human intervention may be necessary, however, to introduce this species with gravity-dispersed seeds into newly available habitats.

References

Aide TM, Zimmerman JK, Herrera L, Rosario M, Serrano M (1995). Forest recovery in abandoned tropical pastures in Puerto Rico. Forest Ecology and Management 77: 77-86.

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

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

Gibbs PE, Oliveira PE, Bianchi MB (1999). Postzygotic control of selfing in *Hymenaea stigonocarpa* (Leguminosae-Caesalpinioideae), a bat-pollinated tree of the Brazilian cerrados. International Journal of Plant Sciences 160: 72-78.

Hamrick JL, Godt MJW (1996). Effects of life history traits on genetic diversity in plant species. Philosophical Transactions of the Royal Society of London Series B 351: 1291-1298.

Nason JD, Hamrick JL (1997). Reproductive and genetic consequences of forest fragmentation: two case studies of Neotropical canopy trees. Journal of Heredity 88: 264-276.

Rousset F (1997). Genetic differentiation and estimation of gene flow from F-statistics under isolation by distance. Genetics 145: 1219-1228.

Slatkin M (1993). Isolation by distance in equilibrium and nonequilibrium populations. Evolution 47: 264-279.

Wadsworth FH (1950). Notes on the climax forests of Puerto Rico and their destruction and conservation prior to 1900. Caribbean Forester 11: 38-47.

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