Genetic dynamics across early life stages in the tropical tree *Prunus africana* (Rosaceae)

Dissertation zur Erlangung des Grades

Doktor der Naturwissenschaften

am Fachbereich Biologie

der Johannes Gutenberg-Universität in Mainz

Dana Berens

Mainz, im Oktober 2010

Dekan:

Erste Berichterstatterin:
Zweiter Berichterstatter:

Tag der mündlichen Prüfung: 17.12.2010

You ever notice that trees do everything to git attention we do, except walk?

(Alice Walker, The Color Purple)

1 SUMMARY

In many plant species, the genetic template of early life-stages is formed by animal-mediated pollination and seed dispersal and has profound impact on further recruitment and population dynamics. Understanding the impact of pollination and seed dispersal on genetic patterns is a central issue in plant population biology. In my thesis, I investigated (*i*) contemporary dispersal and gene flow distances as well as (*ii*) genetic diversity and spatial genetic structure (SGS) across subsequent recruitment stages in a population of the animal-pollinated and dispersed tree *Prunus africana* in Kakamega Forest, West Kenya. Using microsatellite markers and parentage analyses, I inferred distances of pollen dispersal (father-to-mother), seed dispersal/maternal gene flow (mother-to-offspring) as well as paternal gene flow (father-to-offspring) for four early life stages of the species (seeds and fruits, current year seedlings, seedlings ≤ 3yr, seedlings > 3yr). Distances of pollen and seed dispersal as well as paternal gene flow were significantly shorter than expected from the spatial arrangement of trees and sampling plots. They were not affected by the density of conspecific trees in the surrounding. At the propagule stage, mean pollen dispersal distances were considerably (23-fold) longer than seed dispersal distances, and paternal gene flow distances exceeded maternal gene flow by a factor of 25. Seed dispersal distances were remarkably restricted, potentially leading to a strong initial SGS. The initial genetic template created by pollination and seed dispersal was extensively altered during later recruitment stages. Potential Janzen-Connell effects led to markedly increasing distances between offspring and both parental trees in older life stages. This showed that distance and density-dependent mortality factors are not exclusively related to the mother tree, but also to the father. Across subsequent recruitment stages, the pollen to seed dispersal ratio and the paternal to maternal gene flow ratio dropped to 2.1 and 3.4, respectively, in seedlings > 3yr. The relative changes in effective pollen dispersal, seed dispersal, and paternal gene flow distances across recruitment stages elucidate the mechanisms affecting the contribution of the two processes pollen and seed dispersal to overall gene flow. Using the same six microsatellite loci, I analyzed genetic diversity and SGS across five life stages, from seed rain to adults. Levels of genetic diversity within the studied *P. africana* population were comparable to other *Prunus* species and did not vary across life stages. In congruence with the short seed dispersal distances, I found significant SGS in all life stages. SGS decreased from seed and early seedling

stages to older juvenile stages, and it was higher in adults than in late juveniles of the next generation. A comparison of the data with direct assessments of contemporary gene flow patterns indicate that distance- or density-dependent mortality, potentially due to Janzen-Connell effects, led to the initial decrease in SGS. Intergeneration variation in SGS could have been driven by variation in demographic processes, the effect of overlapping generations, and local selection processes. Overall, my study showed that complex sequential processes during recruitment contribute to the spatial genetic structure of tree populations. It highlights the importance of a multistage perspective for a comprehensive understanding of the impact of animal-mediated pollen and seed dispersal on spatial population dynamics and genetic patterns of trees.

2 GENERAL INTRODUCTION

Dispersal is a central concept in ecology. At least once during their life, almost every organism disperses. Individuals may disperse from one population to another, or their propagules may disperse to start a new generation (Cousens et al. 2008). Thus, dispersal plays a central role in population dynamics and has consequences at the individual level, as well as for populations and whole communities (Clobert et al. 2001, Cousens et al. 2008). Further, dispersal is of special interest for genetic processes, as it is ultimately linked to gene flow and provides genetic exchange within and between populations (Ross 2001).

2.1 Dispersal in plants

The lack of mobility of plants implicates decisive limitations at critical life stages during recruitment, including the significant processes of sexual reproduction and offspring dispersal, and, thus, of gene flow in general (Boucher et al. 1982, Pellmyr 2002, Savolainen et al. 2007). In a large number of species, this obstacle has been mitigated by mutualistic interactions with mobile animal dispersers. Animals provide dispersal of pollen grains and seeds, and, consequently, dispersal of genes, and at the same time are rewarded by access to nutritious nectar and fleshy fruits (Boucher et al. 1982, Pellmyr 2002, Ghazoul 2005).

Pollen dispersal by animal vectors provides advantages for both plant and animal dispersers. From the plant's perspective, animals increase the chance of a pollen grain from one flower to reach the receptive stigma of a conspecific flower (Pellmyr 2002, Ghazoul 2005). Thus, animal-pollination increases the probability of outcrossing, which in many species is essential for successful fruit set (Barret 2002). Animals guarantee pollen gene flow over long distances even in habitats where wind-dispersal is hampered, e.g. in close-canopy rainforests (Boucher et al. 1982, Pellmyr 2002).

Seed dispersal provides the advantage of moving offspring away from the maternal parent (Howe & Smallwood 1982). Thus, seeds are able to escape the zone of chemical influence and of high conspecific densities near the mother. Predation rates and pathogen attacks are usually higher near the parent tree, which means that removal away from the zone of influence enhances the survival probability for dispersed offspring (Janzen 1970, Connell 1971, Howe & Smallwood 1982). At the same time,

they have the chance of colonizing new microhabitats and of being dispersed to sites suitable for germination (Howe & Smallwood 1982).

From an ecological perspective, pollen and seed dispersal determine the successful reproduction and recruitment of plant species. They form the template upon which all further recruitment processes are based (Jordano & Godoy 2002). This template influences regeneration dynamics of the focal plant species as well as dynamics in whole plant communities, e.g. the maintenance of high species diversity in tropical forest ecosystems (Janzen 1970, Connell 1971). From an evolutionary perspective, pollination and seed dispersal, as the two vectors of gene flow in plants, shape the genetic composition and structure of plant populations (Hamrick et al. 1993). They have a strong impact on the maintenance of genetic diversity and the formation of genetic structure. Consequently, they determine the adaptation potential of plant populations to microenvironmental or broad-scale ecological conditions (e.g. climate change) (Hamrick et al. 1993, Davis & Shaw 2001, Dutech et al. 2002, González-Martínez et al. 2002).

Overall, a comprehensive understanding of animal-mediated pollination and seed dispersal processes is a central issue in plant population biology, especially in the tropics, where up to 90% of plant species rely on pollination and seed dispersal by animals (Howe & Smallwood 1982, Bawa 1990, Jordano 2000). Knowledge on the two processes and their consequences for recruitment and genetic makeup of populations is essential for the conservation of plant species as well as an understanding of ecosystem dynamics.

2.2 Animal-mediated gene flow

Gene flow is an important process modulating the action of genetic drift and selection, and thus determining the genetic structure of populations (Ross 2001). It can influence a host of relevant variables, e.g. the size of a population, genetic diversity, local adaptation, and ultimately speciation.

As a result of their immobility, gene flow is always of passive nature from a plant's perspective (Nathan 2006). The movement and deposition site of pollen grains and seeds, and, thus, the spatial patterns of instantaneous gene flow, are determined by animal movement (Schupp et al. 2002, Ghazoul 2005). For the majority of species, pollen-mediated gene flow seems to be more extensive in space, and hence more

important than seed-mediated gene flow (Levin & Kerster 1974, Petit et al. 2005). However, particularly in species with pollination by insects and seed dispersal by vertebrates, seed dispersal can play a significant role in gene flow as well (Petit et al. 2005). Dispersal of pollen and seeds is only the first step towards gene flow in plant species and, in a way, forms the template for further recruitment processes. Effective gene flow, i.e. the successful establishment of genes in the next generation, is the result of a multistage process. It requires successful fertilization, germination, and survival across subsequent recruitment stages (Savolainen et al. 2007). Thus, to get a comprehensive idea on of how dispersal affects gene flow, multistage analyses are indispensable.

Quantifying gene flow has always been a challenging issue in plant population biology (Ouborg et al. 1999). Instantaneous patterns of gene dispersal have been assessed by following the physical movement of pollen and seeds with the help of traps, dyes, and labeling, or by following the movement of dispersers (Levin & Kerster 1969, Greenwood 1986, Linhart et al. 1987). Besides being time-consuming and labor-intensive, these methods have the drawback that they can only measure dispersal, but provide no information on effective gene flow, i.e. the realization of contemporary gene flow in established offspring (*sensu* Meagher & Thompson 1987). During the last decades, molecular genetic techniques have opened up a new set of tools for assessing pollen and seed gene flow in plants. For several years, the field has been dominated by so-called “indirect” measures, which infer gene flow from the steady-state population genetic structure (Hardy et al. 2006). Thereby, gene flow is indirectly assessed by measures of genetic variation, e.g. the variation of genetic relatedness across geographical distances (Wright 1943, Hamrick et al. 1993). The use of highly variable markers such as microsatellites has facilitated the advance of “direct methods”. They reconstruct gene flow from exact genotypes of individual propagules without assuming equlilibrium conditions (Godoy & Jordano 2001, Hardy et al. 2006). The most direct approaches to measure gene flow are parentage assignments, whereby individual offspring are allocated to parents within the parental population (Jones & Ardren 2003). They offer the unique potential of directly measuring instantaneous distances of pollen and seed dispersal and gene flow. Further, effective gene flow in different life stages is traceable with these methods, thereby allowing for tracking changes of the initial gene flow template across recruitment (e.g. Isagi et al. 2000, González-Martínez et al. 2006).

2.3 Consequences of animal-mediated gene flow for genetic structuring

Spatial genetic structure (SGS) in populations is defined as the non-random distribution of genetic variation among individuals (Vekemans & Hardy 2004). It is determined by the interaction of genetic drift, natural selection, colonization history, the spatial distribution of individuals within populations, and, to a large part, by gene flow (Epperson 1993, Hamrick & Nason 1996). Especially at fine spatial scales, the most prevalent cause for SGS is the formation of local pedigree structures due to restricted gene flow (Vekemans & Hardy 2004). With gene flow being limited, genetic similarity among neighboring individuals is higher than similarity among distant individuals, leading to isolation by distance effects also within populations (Wright 1943, Hamrick et al. 1993). Due to the immobility of adult plants, dispersal of their propagules is more or less restricted. As a consequence, SGS is common in plant species (Vekemans & Hardy 2004).

Studies of SGS in plant populations can reveal the operation of key evolutionary processes. Genetic subdivision leads to decreasing effective population sizes, which influence genetic dynamics within the whole population (Crow & Kimura 1990, Chesser 1991). Further, clustering of related individuals affects viability of individuals by influencing predator- or pathogen-mediated survival probabilities (Hammond & Brown 1998, Mangan et al. 2010). Last, genetic subdivision can promote local adaptation to microenvironmental conditions varying within populations (Epperson 1992, Kalisz et al. 2001). Thus, an understanding of SGS in plant population can be crucial for management and conservation of genetic resources of these populations (Ng et al. 2004).

Vekemans and Hardy (2004) analyzed a data set of 47 plant species and tried to relate SGS in these species to life-history traits and ecological conditions. According to their results, plant breeding system and life form (e.g. tree, shrub) highly significantly affect SGS. Predominantly selfing species show higher levels of SGS than predominantly outcrossing species, and herb populations are in general more strongly genetically structured than trees. Animal-dispersed species are more strongly genetically structured than wind-dispersed species and less structured than gravity-dispersed species (Vekemans & Hardy 2004). The non-random spatial patterns of gene dispersal imposed by heterogeneous movement of animal pollinators and seed dispersers provoke highly structured genetic diversity within populations of animal-

dispersed species (Godoy & Jordano 2001, Jordano & Godoy 2002). Clumping of seeds in habitats favored by dispersal agents as resting or foraging sites can lead to clustering of related genotypes in these patches (Jordano & Godoy 2002, Zhou & Chen 2010). The persistence of the initial genetic make-up formed at the seed stage is influenced by various ecological and evolutionary processes across life-stages (Ng et al. 2004). If mortality with age is random, the initial SGS in a population is likely to persist to the adult stage (Pacheco & Simonetti 2000). If mortality is non-random due to selection processes, the patterns of SGS can change during recruitment (Hamrick et al. 1993, Kalisz et al. 2001, Jones & Hubbell 2006, Zhou & Chen 2010). Therefore, to understand the underlying ecological and evolutionary processes that lead to the formation of SGS within a population, a consideration of different life stages is indispensible (Aldrich et al. 1998, Schnabel et al. 1998, Kalisz et al. 2001).

2.4 Aims of the thesis

In the present study, I investigated animal-mediated gene flow and spatial genetic structure of the tropical tree species *Prunus africana*. I conducted my study in a mid-altitudinal rainforest in tropical western Kenya, because animal pollination and seed dispersal are especially important in tropical ecosystems (Howe & Smallwood 1982, Jordano 2000). The endangered African cherry tree *P. africana* was chosen as a model system for my study. It relies on pollination by insects and seed dispersal by birds and monkeys (Hall et al. 2000, Farwig et al. 2006). *Prunus africana* is abundant in Kakamega Forest and the local conditions allowed for the study of a semi-isolated population suitable for the methodology I used. A set of microsatellite loci had already been applied in an earlier study (Farwig et al. 2008) and was therefore available for genetic analyses.

One aim of my thesis was to assess gene flow via pollen and seed dispersal and to track changes of the initial patterns created at the seed stage across subsequent recruitment stages. A second aim was to relate my direct observations on gene flow to genetic diversity and SGS within the population across different life-stages.

My thesis consists of two major chapters. They are structured like a journal publication with an introduction, details on methodology, a results section, and a discussion. A synthesis of the results is given in the conclusions following the two chapters.

In the first chapter, I investigated pollen dispersal, seed dispersal, maternal and paternal gene flow of *P. africana* for four subsequent life stages, from seed rain to older juvenile stages. I made use of highly polymorphic microsatellite markers and parentage analyses to directly assess contemporary and effective gene flow distances and compare them across life stages.

In the second chapter, I focused on genetic diversity and the formation of SGS across five recruitment stages (seed rain to adults). In combination with the knowledge on contemporary gene flow from the first chapter, I wanted to elucidate the mechanisms leading to the emergence of SGS within the *P. africana* population.

3 DYNAMICS OF CONTEMPORARY GENE FLOW BY POLLEN AND SEEDS ACROSS SUCCESSIVE LIFE STAGES

3.1 Introduction

Pollination and seed dispersal are the two main processes of gene flow in plants. They form an initial spatial genetic template for further recruitment and thus impact genetic structuring, the maintenance of genetic diversity, as well as the adaptation capability of plant populations (Hamrick et al. 1993, Dutech et al. 2002, González-Martínez et al. 2002). Animal-mediated pollen and seed dispersal are common worldwide, especially in the tropics, where up to 90% of all plant species depend on animals for the dispersal of their genes within and between populations (Howe & Smallwood 1982, Jordano 2000). Hence, the behavior of pollinators and seed dispersers directly influences the spatial patterns of gene movement.

Pollinators and seed dispersers show spatially heterogeneous movement patterns (e. g. Ghazoul 2005, Russo et al. 2006, Garcia et al. 2007, Spiegel & Nathan 2007) influenced by variable factors. Movement patterns of pollinators can be affected by the availability of floral resources, e. g. floral display size (Burczyk et al. 2004, Ghazoul 2005). The distribution of pollinator flight distances is mostly leptokurtic with a majority of short flights (Levin & Kerster 1974, Bittencourt & Sebbenn 2007). However, due to high pollen carryover, i.e. the transport of pollen of the source tree to several subsequently visited individuals, pollinator flight distances do not necessarily reflect pollen dispersal patterns (Schaal 1980, Rasmussen & Brodsgaard 1992). Seed density also decreases leptokurtically with distance from the seed source, with a fat tail of long-distance dispersal events (Schupp et al. 2002, Nathan et al. 2008). However, seed dispersal curves created by animal dispersers can often be multimodal, as movement patterns of seed dispersers are affected by, for example, the spatial distribution of fruiting trees (Schupp et al. 2002, Clark et al. 2004, Kwit et al. 2004, Russo et al. 2006). Clumps of seeds are often deposited in microhabitats favored by dispersal agents, such as resting, nesting or foraging sites (Schupp et al. 2002, Kwit et al. 2004, Russo & Augspurger 2004). Thereby, the presence of conspecific fruiting trees in the surrounding of the foraging tree can affect post-foraging movement patterns (Izakhi et al. 1991, Jordano & Schupp 2000). Different traits and requirements of pollen

and seed dispersal agents can be reflected in different movement patterns, which in turn can have variable impact of pollen and seed dispersal on gene flow patterns.

Whereas effective seed dispersal (mother-to-offspring) and maternal gene flow (mother-to-offspring) reflect the same process, paternal gene flow (father-to-offspring) is a two step process of pollen dispersal (father-to-mother) and seed dispersal (figure 3.1F). For the majority of species, pollen-mediated gene flow seems to be more extensive in space, and hence more important than seed-mediated gene flow (Levin & Kerster 1974, Petit et al. 2005). However, particularly in species with pollination by insects and seed dispersal by vertebrates, seed dispersal can play a significant role in overall gene flow as well (Petit et al. 2005). Molecular tools like parentage analyses allow for directly disentangling the contribution of the two processes to genetic patterns (Jordano & Godoy 2002, Geng et al. 2008). Studies investigating pollen and seed dispersal by animals have revealed variable contributions of the two processes to gene flow patterns. For example, studies of insect-pollinated vertebrate-dispersed tree species report all possible combinations of (*i*) longer pollen dispersal in *Sorbus torminalis* (Oddou-Muratorio & Klein 2008), (*ii*) similar seed and pollen dispersal distances in *Simarouba amara* (Hardesty et al. 2006) and (*iii*) longer seed dispersal in *Prunus mahaleb* (Garcia et al. 2007).

Various factors alter the spatial patterns of gene movement initially created by pollination and seed dispersal across sequential recruitment stages. Biotic factors are predation pressure on dispersed seeds and herbivory attacks on seedlings (Asquith et al. 1997, Pulido & Diaz 2005). These factors are often distance and density-dependent, leading to an increased survival of seeds and seedlings farther away from the mother tree (Janzen 1970, Connell 1971). Consequently, if such "Janzen-Connell (J-C) effects" operate in subsequent recruitment stages, the mean effective seed dispersal distance (identical to maternal gene flow distance) is expected to increase as the recruitment process progresses, whereas both effective pollen dispersal and paternal gene flow distances (as defined in figure 3.1F) are expected to vary randomly with no particular trend. Alternatively, if J-C effects do not play an overriding role, the variation in abiotic conditions among different microhabitats would determine germination, establishment, and survival of juvenile plants (Schupp 1995, Alvarez-Buylla et al. 1996), and, therefore, subsequent recruitment patterns will be heterogeneous and uncoupled (Jordano & Herrera 1995, Schupp 1995). Overall, investigating initial gene flow patterns at the seed stage should be corroborated by examining how they translate into

gene flow patterns in subsequent recruitment stages (Imbert & Levèvre 2003, Burczyk et al. 2006).

Several studies have investigated effective pollen and seed dispersal patterns in established seedlings or saplings (e.g. Isagi et al. 2000, González-Martínez et al 2002, Oddou-Muratorio & Klein 2008). However, to date, few studies have focused on a multistage perspective (but see González-Martínez et al. 2006, Bittencourt & Sebbenn 2007), and, to my knowledge, the change in paternal gene flow distance has not yet been examined across several stages.

The aim of my study was to assess patterns of animal-mediated gene flow across four early life-stages in the tropical tree *Prunus africana*. My first objective was to compare observed distances of effective pollen dispersal, effective seed dispersal (= maternal gene flow) and paternal gene flow to distances created by animal pollen and seed dispersal agents to distances expected under unlimited uniform dispersal, and to elucidate how the density of conspecific trees affects these distances. I hypothesized that pollen and seed dispersal as well as paternal gene flow patterns are non-random and predicted that dispersal distances decline under high conspecific tree density. I also aimed to disentangle the contribution of pollen and seed dispersal to gene flow. My study tree is pollinated by insects and dispersed by vertebrates, mostly by birds. I, thus, expected to find longer seed than pollen dispersal distances due to the higher mobility of birds. Furthermore, I aimed to assess how dispersal and gene flow distances change across subsequent recruitment stages. Effective seed dispersal distances were expected to increase due to J-C effects, whereas pollen dispersal and paternal gene flow distances were expected to vary randomly across recruitment stages.

3.2 Methods

3.2.1 Study species

Prunus africana Hook f. (Rosaceae) is an evergreen monoecious tree species. Its range expands across East Africa, Madagascar and the Comores (Hall et al. 2000). It is listed on CITES appendix II due to bark exploitation for medicinal purposes.

Prunus africana has small white hermaphroditic protogynous flowers which are mainly pollinated by small insects (hymenoptera, diptera). The species is reported to be mainly outcrossing with the potential for self-fertilization (Munjuga et al., unpublished data). The one-seeded fruits turn from green to purple while ripening. *Prunus africana* seeds are dispersed by differently-sized bird species, for example the Black-and-white-casqued Hornbill (*Bycanistes subcylindricus*), Common Bulbul (*Pycnonotus barbatus*), or wintering Blackcaps (*Sylvia atricapilla*) (Farwig et al. 2006). Furthermore, three different monkey species have been identified as dispersers (Farwig et al. 2006). The seed embryo is covered by a woody endocarp of maternal origin. Seedlings germinate immediately after the first rainfall in the rainy season of the same year without seed dormancy. The woody maternal endocarp stays attached to the germinating seedlings for about one month (D. G. Berens, personal observations), which allows for identifying the mother tree of established seedlings.

3.2.2 Study site

My study area was Kakamega Forest in western Kenya (figure 3.1). The mid-altitudinal rainforest (1,500 – 1,700 m asl, KIFCON 1994) is the eastern most remnant of the former Guineo-Congolian rainforest belt (Kokwaro 1988). The forest comprises a main forest block and five fragments with a total area of 13,000 ha (Lung 2004). Mean daily temperatures range from 10.6 to 27.7°C (Tsingalia 1990). Average yearly precipitation is about 2000 mm/year (averaged from Forest Department records at Isecheno Forest Station from 1982 to 2005).

The study area within Kakamega Forest is a 120 ha forest peninsula in the northwest of the main forest block extending into the agricultural matrix (0°21´34.1´´ North, 34°51´31.0´´ East; figure 3.1). It comprises both primary and secondary forest.

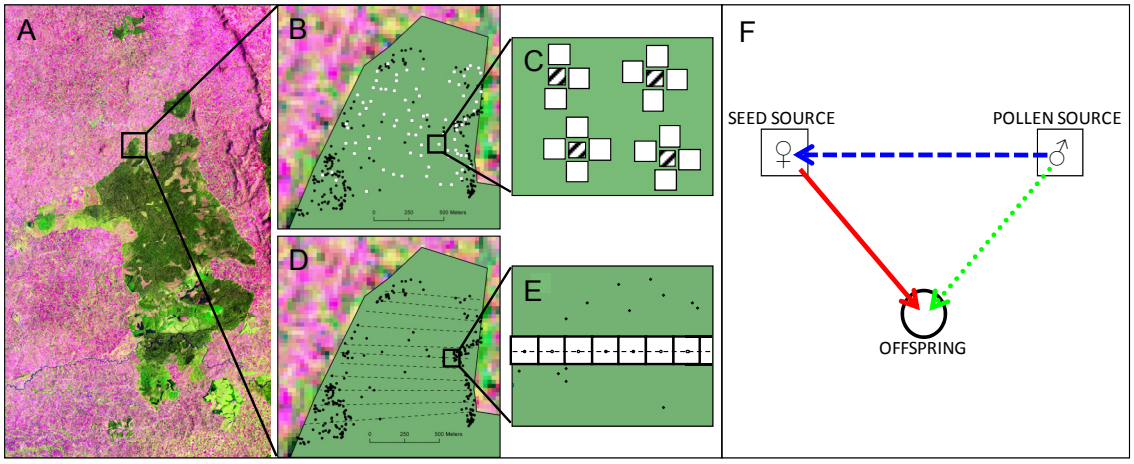

Figure 3.1. (A) Satellite image showing Kakamega Forest with forested area (green) and farmland (pink) and the location of the study area within the forest; (B) spatial distribution of adult *Prunus africana* trees (n=261, black dots) and of sampling plots (n=86, white squares) for sampling of fruits and seeds (SDFRT) and current-year seedlings (SLYNG); (C) each sampling plot consisted of four seed traps (0.5m², hatched squares) and 15 seedling sub-plots (1m², white squares); (D) spatial distribution of adult *P. africana* trees (black dots) and of eleven transects (dashed lines) for sampling of middle old seedlings (SLMID) and old seedlings (SLOLD); (E) arrangement of virtual plots (white squares) for sampling of SLMID and SLOLD. (F) Schematic representation of dispersal and gene flow distances in plants. Effective seed dispersal (red solid line) describes net displacement from mother tree to offspring; effective pollen dispersal (blue dashed line) describes net displacement from father to mother tree. Maternal gene flow is identical to effective seed dispersal (red solid line), whereas paternal gene flow (green dotted line) describes net displacement from father to offspring.

3.2.3 Sampling of plant material

In 2005, I established 22 parallel transects crossing the study area from the Eastern to the Western forest edge and covering a total area of ca. 87 ha (figure 3.1). Transects were between 694 and 1225 m long and were separated by 40 m. In total, 261 *P. africana* trees (dbh > 10 cm) were detected in the transect area and an adjacent 100 m northern and southern buffer zone. GPS coordinates of each tree were taken. Leaf material of all adult trees was collected between January and May 2006.

In the same year, I randomly established 86 sampling plots along the transects (figure 3.1B). I took GPS coordinates of each plot. Each plot consisted of an array of four seed traps (0.7 x 0.7 m) made of a wire frame covered with mesh (figure 3.1C). I collected propagules from seed traps on a weekly basis between January and April 2006 during a fruiting peak of *P. africana* in my study site. I assigned them to fruits (with pulp) or seeds (without pulp). Fruit pulp was manually removed for storage and all samples were dried with silica gel. Of the 8,544 propagules (5785 fruits, 2759 seeds) that were counted in traps, I used a sample of 311 (62 fruits, 249 seeds) for genetic

analyses. Frequency distributions of fruit and seed dispersal distances did not differ (Chi²-test: p-values > 0.05); therefore, data of fruits and seeds were pooled as propagule stage (SDFRT hereafter) for further analyses.

I collected current year seedlings (SLYNG hereafter) at the same 86 sampling plots between March and May 2006. Fifteen 1m² sub-plots were placed within each plot (figure 3.1C) and newly emerged seedlings were counted and collected on a biweekly basis. I dried leaf samples and woody endocarp attached to the roots with silica gel. A sample of 309 of the 48,991 seedlings was used for genetic analyses.

Older seedling stages were sampled between January and February 2008 along every second of the 22 transects crossing the study area (figure 3.1D). The area covered along transects was divided into virtual 10 x 10 m sampling plots (figure 3.1E; plots were termed virtual plots because they were not marked on the ground), thereby covering a total sampling area of ten ha, and the centroid of each plot was used as its geographic position. I recorded each *P. africana* seedling encountered in these virtual plots. I took a leaf sample of each individual, and seedling age was determined by an experienced field assistant, considering the degree of lignification of the stem axis. A sample of 298 of the 861 seedlings up to three years of age (SLMID hereafter) and 301 of the 368 seedlings older than three years (SLOLD hereafter) was taken for genetic analyses.

For SDFRT and SLYNG, I used both biparental and maternal tissue for genetic analyses. Biparental tissue was taken from the seed embryo for SDFRT and from leaf samples for SLYNG. Genotyping of maternal endocarp tissue of SDFRT and SLYNG potentially allowed for a discrimination of the pollen donor and the seed source for both stages. For SLMID and SLOLD, I genotyped biparental tissue taken from leaf samples.

3.2.4 Molecular analyses

All plant material (leaves of adult trees, SLYNG, SLMID, SLOLD, embryo tissue of SDFRT, endocarp material of SDFRT and SLYNG) was ground to fine powder with a Retsch mixer mill (Type MM 301, Retsch, Haan, Germany). DNA was extracted following the protocol described by Wang et al. (1993). I modified the standard protocol by using 40µl of NaOH and 95 µl of Tris buffer for each sample. Samples were genotyped at six microsatellite loci: locus UDP97-403 (Cipriani et al. 1999), locus P12A02 (Sosinski et al. 2000), locus BPPCT-002 (Dirlewanger et al. 2002), loci UDP 96-005 and UDP98-410 (Schueler et al. 2003), and locus EMPaS06 (Vaughan &

Russell 2004), developed for *Prunus persica and Prunus avium.* Primers were fluorescent 5'-end-labeled (6-FAM, NED, HEX; Applied Biosystems, Foster City, CA, USA). Polymerase chain reaction protocols followed Farwig et al. (2008) with annealing temperatures ranging from 54 to 62°C and number of cycles ranging from 44 to 46. For loci UDP97-403 and EMPaS06, I added a final extension step of 60°C for 30 min to the standard protocol. Samples were genotyped on ABI 3730 and 3130xl capillary sequencers. Allele scoring was performed using GeneMapper 4.0 (Applied Biosystems, Foster City, CA, USA). For all loci, base shifts between scorings of the different sequencers were calibrated by running a number of samples on both instruments. All samples, except that of one adult tree, were successfully amplified. Loci were highly polymorphic with a mean number of 20.0 ± 2.6 SD alleles per locus (range: 3.4 – 29.2).

3.2.5 Parentage analyses

Cumulative exclusion probabilities for all loci exceeded 99.0 %. Null allele frequencies were obtained using CERVUS 3.0.3 (Kalinowski et al. 2007). The null allele frequency at locus BPPCT-002 was slightly higher than 0.2 for SLMID and SLOLD and was thus considered to be problematic for parentage analyses (Dakin & Avise 2004). Therefore, for SLMID and SLOLD, homozygotes on this locus were coded as heterozygotes with one allele set to missing data to avoid wrong parentage exclusion. The maternal endocarp material of SDFRT and SLYNG reliably amplified on only four of the six loci (loci UDP97-403, P12A02, UDP 96-005, UDP98-410), and direct assignment of the seed source through identity analyses was not possible. Therefore, I used the biparental embryo or leaf material to assign two parents to each offspring with parentage analysis in a first step and, as a second step, compared the endocarp genotype of the offspring with the parental genotypes to distinguish between seed and pollen source. A discrimination of pollen and seed source was possible for 40 (30/76) and 21 (27/131) % of assignments for SDFRT and SLYNG, respectively, whereby the seed source was the nearer of the two parental trees in 97 (29/30) % and 93 (25/27) % of all cases. I furthermore tested the three scenarios that a) the mother, unambiguously assigned with the endocarp material, is considered to be the seed source; b) the nearer of the two assigned parental trees is the seed source, and c) the farther of the two parents is considered to be the seed source. Scenarios a) and c) had significantly different distance distributions (Kolmogorov-Smirnov (KS) tests: all p-values < 0.001), while

distributions for a) and b) were not significantly different (KS tests: all p-values > 0.05). Therefore, the nearest parental tree was considered to be the seed source for all stages in all analyses. For those offspring where no parent pair was assigned, a paternity analyses was conducted to see if at least one parent (either pollen or seed source) could be found inside the study area.

All parentage analyses were done in FAMOZ (Gerber et al. 2003) which is especially suitable for parentage analyses in non-isolated stands. FAMOZ conducts maximum-likelihood assignments based on log of the odds (LOD) scores, i.e. the likelihood of a parent pair of an offspring relative to the likelihood of an arbitrary parent couple (Marshall et al. 1998, Gerber et al. 2000). The LOD threshold for parentage assignment was obtained from two simulation procedures ("inside" and "outside" the stand). For simulations "inside" the stand, offspring is generated from random samples of specific gametes of parents inside the stand; simulations "outside" the stand generate offspring from randomly generated gametes according to allele frequency data of the population. The LOD threshold is set at the intersection of the two LOD distributions. All parent pairs with a LOD score higher than the threshold are assigned. The parameter set used for parentage analyses, comprising the simulation and calculation error rate as well as departure from Hardy-Weinberg equilibrium, is applied to simulated data to obtain statistical confidence. Simulating 10,000 offspring, I tested varying parameter sets, starting from an error rate of 0.001. Eventually, an error rate of 0.001 was chosen as parameter set as it yielded less than 20 % α-error and the highest rate of correct parent pair choices.

3.2.6 Assessment of pollen dispersal, seed dispersal, and gene flow distances

For all offspring with assigned parentage (SDFRT: n = 76, SLYNG: n = 131, SLMID: n = 36, SLOLD: n = 36), I derived pollen and seed dispersal as well as paternal gene flow distances from x and y coordinates of adult *P. africana* trees and sampling plots (i.e. the sampling plots in figure 3.1B, C or virtual sampling plots in figure 3.1D, E). Pollen dispersal distance is the distance between pollen and seed source trees. Seed dispersal (= maternal gene flow) is the distance between the seed source tree and the sampling plot, and paternal gene flow is the distance between the pollen donor and the sampling plot (figure 3.1).

3.2.7 Accounting for sampling biases

Sampling and genotyping of a fraction of all offspring counted at each plot, as well as the spatial arrangement of sampling plots and trees can introduce bias to the frequency distributions of dispersal distances. If, for example, a low proportion of offspring in close proximity to the mother tree and a high proportion at farther distances is sampled, seed dispersal distances will be overestimated. The spatial arrangement of sampling plots relative to the source trees can also bias the estimation of dispersal events (Robledo-Arnuncio & Garcia 2007). The frequency of dispersal events in distance classes with higher numbers of mother-plot pairs will be overestimated. The same holds for pollen dispersal, where the spatial arrangement of trees affects the estimation of pollen dispersal distances. If, for example, the number of potential pollen source trees is high in close vicinity of the mother tree, the probability of observing short dispersal distances by chance alone is high, and pollen dispersal distances will be underestimated.

To account for these biases I applied two types of weighting factors. First, I took into account the sampling bias caused by sampling and genotyping a variable proportion of offspring, and, second, I considered the spatial position of plots and trees to remove the spatial arrangement bias. In the following, seed traps on plots, sub-plots on plots (used to count and collect SLYNG), and virtual plots (used to count and collect SLMID and SLOLD) are all referred to as "plots". As a first step, each genotyped offspring was weighted by the proportion of the total number of offspring that was genotyped at each plot according to weighting factor *w1*:

$$w_1 = \frac{n_{total}}{n_{sample}} \quad (1),$$

where n_{total} is the total number of offspring counted at the plot and n_{sample} is the sample size used for genetic analyses. If, for example, 35 propagules were recorded at a plot and 15 thereof were genotyped, each genotyped propagule of this plot got the weight 35/15 = 2.33 for all analyses.

Second, to remove the spatial arrangement bias for pollen dispersal, I defined discrete distance intervals of 40 m around each tree that acted as a seed source. For each of these seed source trees, the number of potential pollen donors in each distance interval was determined. Each genotyped offspring of this tree was weighted with weighting factor *w2*:

$$w_2 = \frac{1}{n_{trees}} \quad (2),$$

with n_{trees} being the number of *P. africana* trees in the respective distance interval. If, for example, the pollen dispersal distance for one offspring was 70 m, it was weighted by the number of trees within a distance interval of 40 to 80 m from the seed source tree. With five trees occurring in this distance interval, for instance, the offspring got the weight 1/5 = 0.2. Consequently, with few trees occurring in this distance interval, the probability for the observed dispersal event to occur by chance alone was low. Thus, the offspring got a higher weight in the calculation of dispersal distance distributions.

To account for the spatial arrangement bias in seed dispersal/paternal gene flow distances, I also worked with discrete distance intervals of 40 m. For each tree that acted as a seed/pollen source I determined the total area of plots (separately for seed traps on plots, sub-plots in plots, and virtual plots) in each distance interval. Then, each genotyped offspring of this tree was weighted with weighting factor *w3* according to the total plot area in this distance interval:

$$w_3 = \frac{1}{area_{total}} \quad (3),$$

where $area_{total}$ is the total plot area in the respective distance interval. If, for example, the seed dispersal distance for one offspring was 70 m, it was weighted by the total area covered by plots within a distance interval of 40 to 80 m from the seed source tree. With a total area of 10m² of seed traps within this distance interval, a genotyped seed got the weight 1/10 = 0.1. Consequently, with a smaller plot area occurring in this distance interval, the probability for the observed dispersal event to occur by chance alone was lower. Thus, the offspring got a higher weight in the calculation of dispersal distance distributions.

For comparison of observed distributions with null models (see below), observed values were only weighted by sampling weighting factor w_1. For all other analyses, dispersal and gene flow distances of an offspring were weighted for sampling bias and spatial arrangement bias. To do this, each offspring was weighted by the product of the two weighting factors, i.e. by the product of w_1 and w_2 for pollen dispersal distances and w_1 and w_3 for seed dispersal distances and paternal gene flow distances.

3.2.8 Statistical analyses

Frequency distributions as well as means of pollen and seed dispersal and paternal gene flow distances were compared applying Kolmogorov-Smirnov (KS) tests and Welch two-sample tests, respectively. Weighting of samples for the analyses inflates type I

errors. Therefore, I followed a conservative approach for all statistical analyses and determined significance applying D-values (KS test) and degrees of freedom (Welch test) of the original sample size.

In order to analyze pollen and seed dispersal as well as paternal gene flow distances, observed frequency distributions and means were compared to null model distributions with KS tests and Welch-tests, respectively. Observed values were weighted by w_1 (sampling weighting factor) for the analyses. For pollen dispersal, null model distributions were inferred from all tree-to-tree distances in my study area. For seed dispersal and paternal gene flow, null model distributions were inferred from all tree-to-plot distances (distinguishing between seed traps on plots, sub-plots in plots, and virtual plots). This null model assumes uniform and unlimited dispersal/gene flow within my study area; that is, pollen/seed dispersers have equal probability to deposit pollen/seeds at any point within the study area (as defined in figure 3.1). Therefore, this null model represents the upper bound, expected if dispersal is uniform and unlimited. An extreme lower bound of very limited dispersal would yield zero dispersal distances for all pollen and seeds. Unfortunately, I do not do have sufficient knowledge on disperser and landscape attributes in this system to compare the observed patterns with mechanistically-derived ones.

In order to analyze the contribution of pollen and seed dispersal to gene flow patterns in my study system, frequency distributions and means of pollen and seed dispersal distances as well as paternal gene flow distances were compared with KS tests and Welch-tests for each life stage. Samples were weighted both for sampling and spatial arrangement bias, i.e. by the product of w_1 and w_2 for pollen dispersal distances and by w_1 and w_3 for seed dispersal and paternal gene flow distances.

I tested for a potential effect of conspecific tree density at the vicinity of the mother tree on pollen and seed dispersal distances. Thus, the number of *P. africana* trees in a 50 m radius around each seed source tree was calculated. The weighted (by the product of weighting factors w_1 and w_2 for pollen and w_1 and w_3 for seed dispersal) mean pollen and seed dispersal distance across all offspring originating from this tree was then calculated for each seed source tree. Weighted mean distances were tested for a rank correlation with conspecific tree density in the surrounding.

Furthermore, for a comparison of distances across successive life stages, pollen and seed dispersal as well as paternal gene flow distances were compared among successive offspring categories with KS tests and Welch-tests. Again, samples were

weighted for both sampling and spatial arrangement bias (by the product of weighting factors w_1 and w_2 for pollen and w_1 and w_3 for seed dispersal and paternal gene flow). All p-values were evaluated after Bonferroni correction.

3.3 Results

3.3.1 Parentage analyses

Parentage assignment rates were 24 % for SDFRT (n = 76/311), 42 % for SLYNG (n = 131/309), and 12 % for SLMID (n = 36/298) and SLOLD (n = 36/301), respectively. Adding paternity assignments, i.e. assignments of a single parent within the study area, assignment rates of at least one parent within the study site added up to 68 % for SDFRT (n = 212/311), 85 % for SLYNG (n = 364/309), 45 % for SLMID (n = 136/298), and 65 % for SLOLD (n = 195/301). Thus, gene immigration rates were 54 % for SDFRT, 36 % for SLYNG, 66 % for SLMID, and 61 % for SLOLD. Selfing accounted for a percentage of 11, 5, 13 and 8 % in genotyped samples of the different life stages.

3.3.2 Pollen and seed dispersal and gene flow distances

Comparing the observed distances to those expected under uniform and unlimited dispersal for the given spatial arrangement of trees and plots, I found that pollen and seed dispersal distance distributions as well as paternal gene flow distance distributions differed highly significantly, as anticipated, from the respective null model distributions in all life stages (KS-tests: all p-values < 0.001; figure 3.2, 3.3, 3.4). Distances were always much shorter than these upper bound expectations (Welch-tests: all p-values < 0.001).

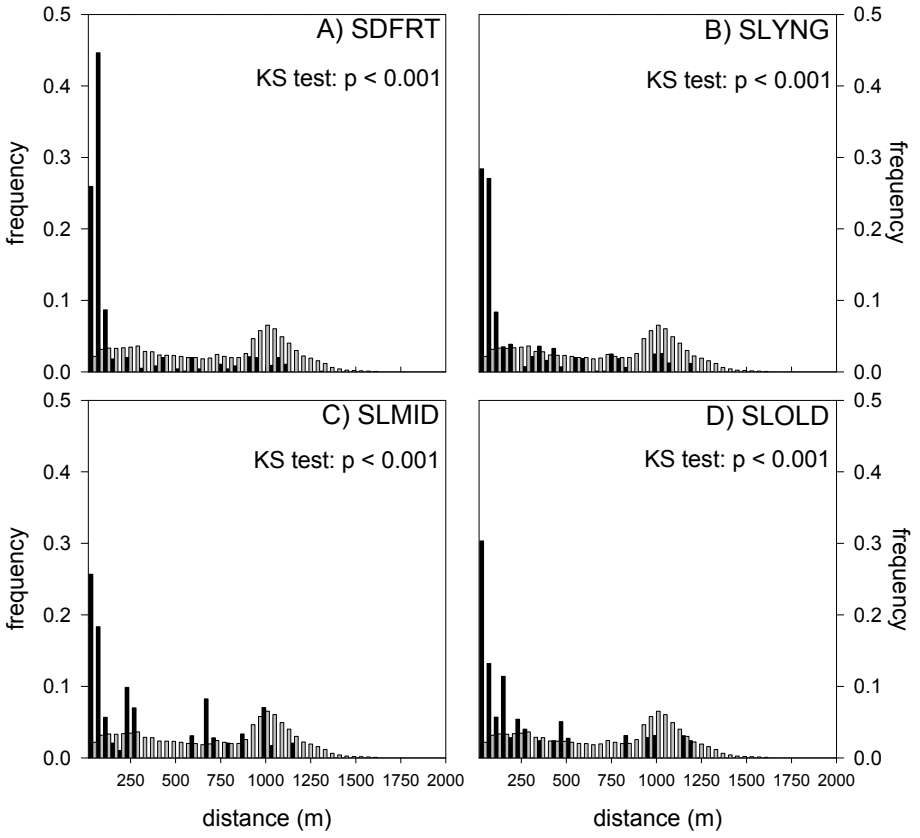

Figure 3.2. Histogram showing the observed frequency distribution of **pollen dispersal** distances (black bars) and distances expected under unlimited uniform dispersal according to all possible tree-tree distances (gray bars) for A) seeds and fruits (SDFRT, n = 76), B) current year seedlings (SLYNG, n = 131), C) middle-old seedlings (SLMID, n = 36, one to three years) and D) old seedlings (SLOLD, n = 36, older than three years). Frequency distributions of observed pollen dispersal are weighted regarding sampling bias (see methods for details). P-values of Kolmogorov Smirnov tests are given.

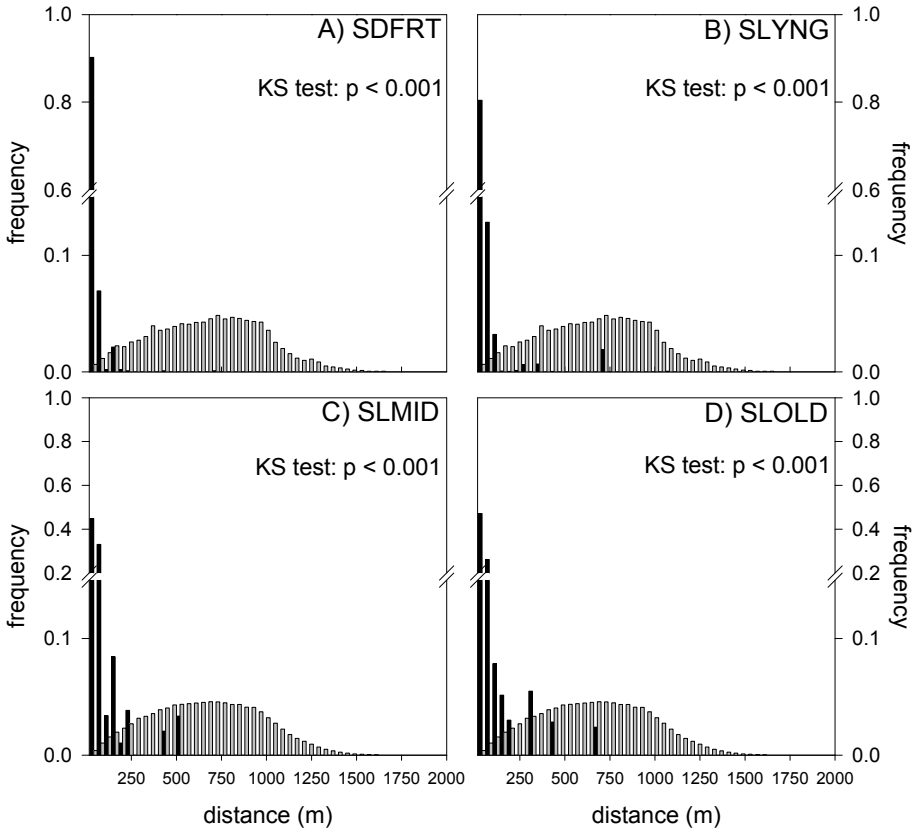

Figure 3.3. Histogram showing the observed frequency distribution of **seed dispersal/maternal gene flow** distances (black bars) and distances expected under unlimited uniform dispersal according to all possible tree-plot (A & B) and tree-virtual plot distances (C & D) (gray bars) for A) seeds and fruits (SDFRT, n = 76), B) current year seedlings (SLYNG, n = 131), C) middle-old seedlings (SLMID, n = 36, one to three years) and D) old seedlings (SLOLD, n = 36, older than three years). Frequency distributions are weighted regarding sampling bias (see methods for details). P-values of Kolmogorov Smirnov tests are given.

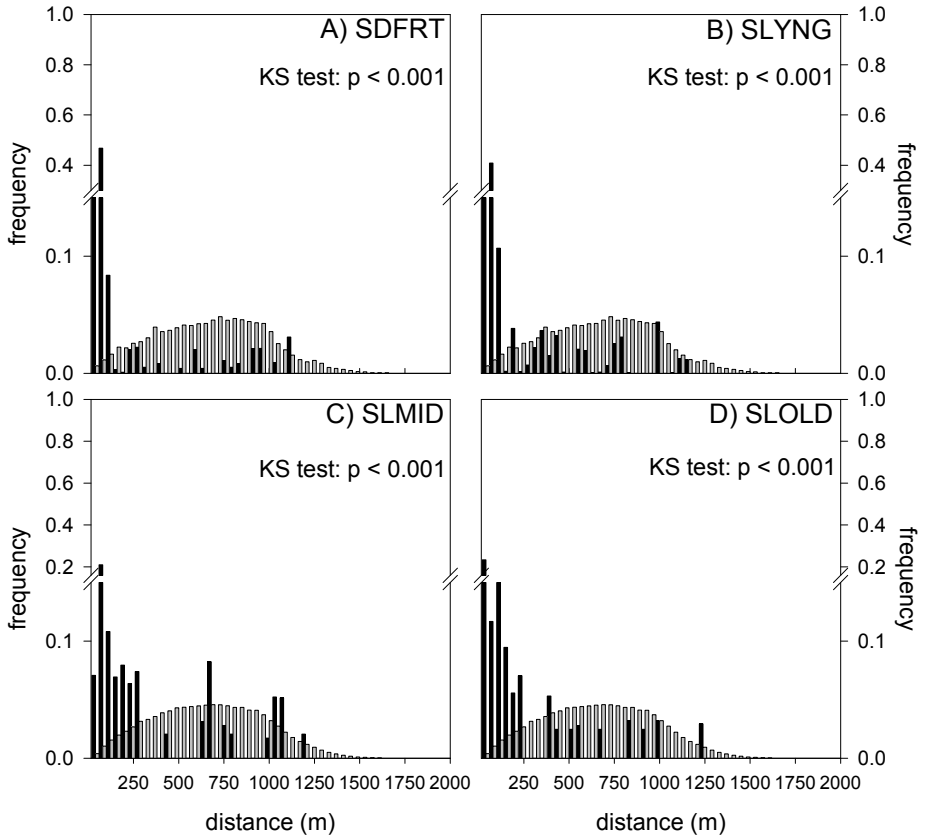

Figure 3.4. Histogram showing the observed frequency distribution of **paternal gene flow** distances (black bars) and distances expected under unlimited uniform dispersal according to all possible tree-plot (A & B) and tree-virtual plot distances (C & D) (gray bars) for A) seeds and fruits (SDFRT, n = 76), B) current year seedlings (SLYNG, n = 131), C) middle-old seedlings (SLMID, n = 36, one to three years) and D) old seedlings (SLOLD, n = 36, older than three years). Frequency distributions are weighted regarding sampling bias (see methods for details). P-values of Kolmogorov Smirnov tests are given.

3.3.3 Contribution of pollen and seed dispersal to gene flow

For all stages, distributions and means of pollen and seed dispersal distances differed significantly (table 3.1). The only exception was the distribution of pollen and seed dispersal distances of SLOLD, which did not differ significantly after Bonferroni correction. Pollen dispersal always occurred over significantly longer distances than seed dispersal (table 3.1). Mean ± SE dispersal distances were 114 ± 25 m (SDFRT), 123 ± 19 (SLYNG), 325 ± 55 (SLMID), and 175 ± 48 m (SLOLD) for pollen, and 5 ± 3 m (SDFRT), 15 ± 6 (SLYNG), 73 ± 8 (SLMID), and 84 ± 11 m (SLOLD) for seeds (table 3.1; figure 3.5, 3.6). The ratio of mean pollen to mean seed dispersal distance was 22.8 for SDFRT, 8.2 for SLYNG, 4.5 for SLMID, and 2.1 for SLOLD (figure 3.8). Between 13.8% (SDFRT) and 54.1% (SLMID) of pollen dispersal movements crossed distances of more than 120 m (figure 3.5). Seed dispersal events of more than 120 m

were considerably less frequent, with proportions from 0.7% (SDFRT) to 8.7% (SLMID) of all dispersal events (figure 3.6).

Table 3.1. Comparison of **pollen and seed dispersal** distance in m (upper table) and **paternal and maternal gene flow** distance in m (lower table) distributions for each stage class. Presented are weighted mean, standard error, range, as well as D-values of two-sample Kolmogorov-Smirnov (KS) tests and t and df values of Welch two-sample tests. Significance (after Bonferroni correction) is indicated with *, n.s. = non-significant. Significant values are bold. SDFRT = seeds or fruits, SLYNG = current year seedlings, SLMID = middle-age seedlings (one to three years), SLOLD = old seedlings (older than three years). Subscript numbers are sample sizes.

	stage	pollen/paternal		seed/maternal		KS-test	Welch-test	
		mean	range	mean	range	D	t	df
Dispersal	$SDFRT_{76}$	114 (25)	0-1308	5 (3)	0-698	**0.63***	**626.03***	89.14
	$SLYNG_{131}$	123 (19)	0-1193	15 (6)	0-1058	**0.60***	**1809.52***	185.8
	$SLMID_{36}$	325 (55)	0-1128	73 (8)	10-489	**0.53***	**209.90***	41.55
	$SLOLD_{36}$	175 (48)	0-1172	84 (11)	1-655	$0.36^{n.s.}$	**39.84***	45.47
Gene flow	$SDFRT_{76}$	124 (26)	0-1091	5 (3)	0-698	**0.85***	**923.57***	90.41
	$SLYNG_{131}$	168 (23)	0-1136	15 (6)	0-1058	**0.79***	**874.22***	186.98
	$SLMID_{36}$	290 (64)	11-1200	73 (8)	10-489	**0.57***	**17.78***	41.04
	$SLOLD_{36}$	289 (77)	2-1376	84 (11)	1-655	**0.39***	**9.30***	44.35

Testing for an influence of conspecific tree density on pollen and seed dispersal distances, I found that neither pollen nor seed dispersal distances depended on conspecific tree density within a radius of 50 m from the mother tree. This was the case for all recruitment stages (Spearman's rank correlation coefficients: $-0.2 \leq Rho \leq 0.1$).

Pollen and seed dispersal distances were mirrored in gene flow distances: distributions and means of maternal and paternal gene flow distances also differed significantly, whereby paternal gene flow always occurred over significantly longer distances than maternal gene flow (identical to seed dispersal; table 3.1). Mean ± SE paternal gene flow distances were 124 ± 26 m (SDFRT), 168 ± 23 m (SLYNG), 290 ± 64 m (SLMID), and 289 ± 77 m (SLOLD), compared to 5 ± 3 m (SDFRT), 15 ± 6 m (SLYNG), 73 ± 8 m (SLMID), and 84 ± 11 m (SLOLD) for maternal gene flow (identical to seed dispersal; table 3.1; figure 3.6, 3.7). The ratio of paternal to maternal gene flow distances was 24.8 for SDFRT, 11.2 for SLYNG, 4.0 for SLMID, and 3.4 for

SLOLD (figure 3.8). Between 19.6% (SDFRT) and 45.0% (SLMID) of paternal gene flow crossed distances of more than 120 m (figure 3.7), as compared to proportions from 0.7% (SDFRT) to 8.7% (SLMID) in maternal gene flow (figure 3.6)

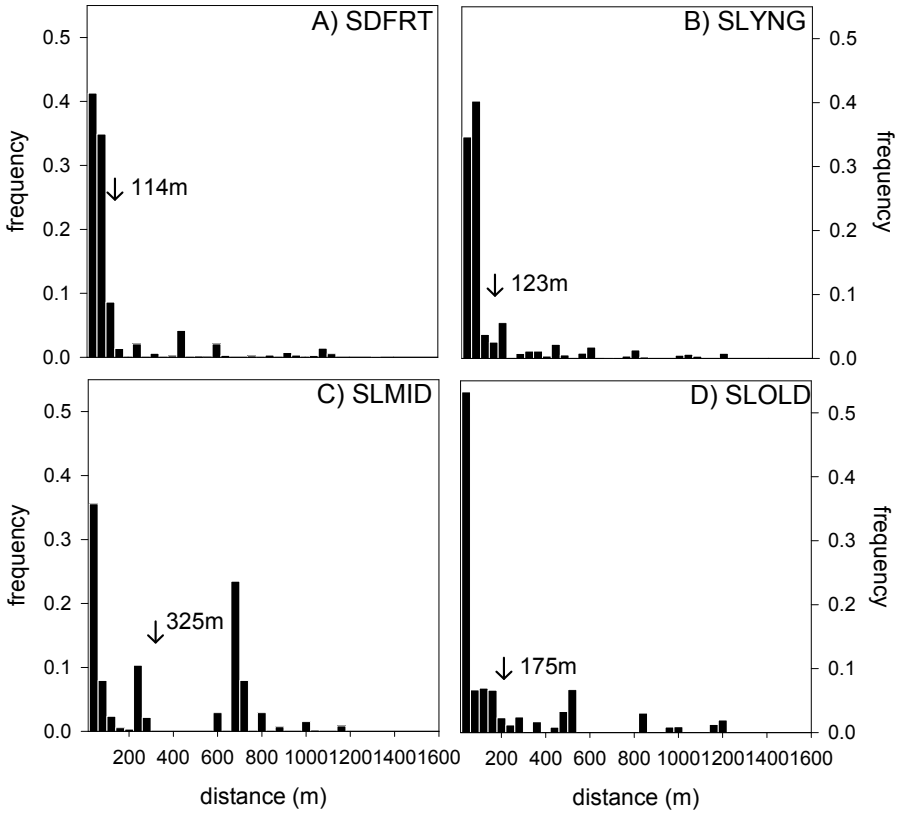

Figure 3.5. Histogram showing the frequency distribution of **pollen dispersal** distances in A) seeds and fruits (SDFRT, n = 76), B) current year seedlings (SLYNG, n = 131), C) middle-old seedlings (SLMID, n = 36, one to three years) and D) old seedlings (SLOLD, n = 36, older than three years). Frequency distributions are weighted regarding sampling bias and spatial arrangement of trees (see methods for details); weighted mean dispersal distances are given.

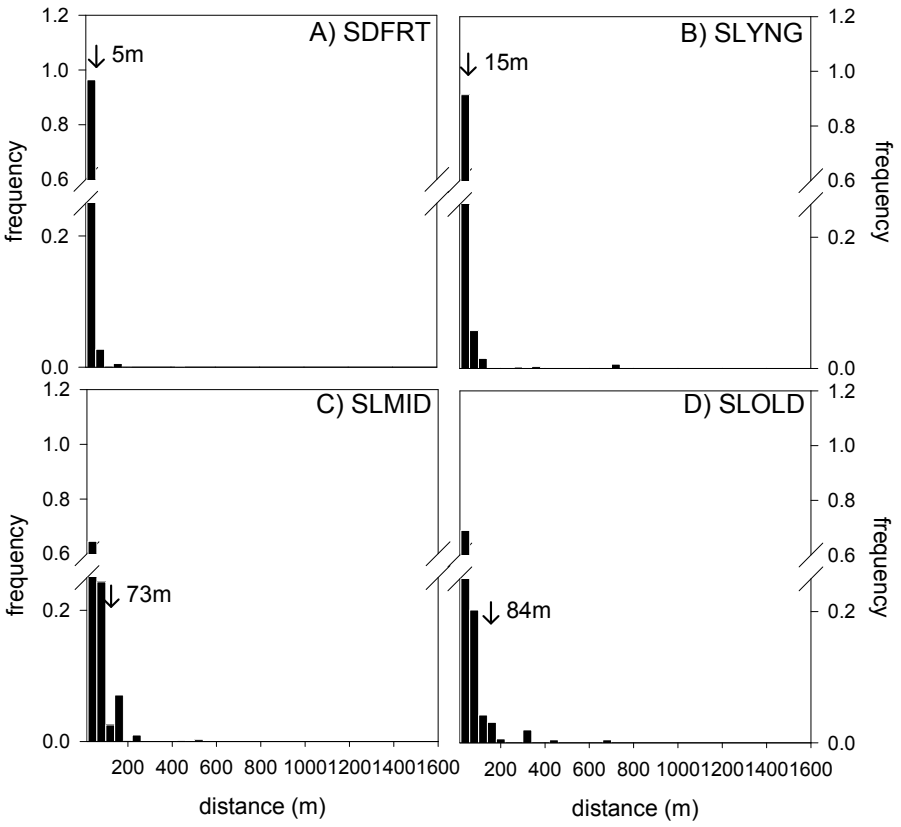

Figure 3.6. Histogram showing the frequency distribution of **seed dispersal/maternal gene flow** distances in A) seeds and fruits (SDFRT, n = 76), B) current year seedlings (SLYNG, n = 131), C) middle-old seedlings (SLMID, n = 36, one to three years) and D) old seedlings (SLOLD, n = 36, older than three years). Frequency distributions are weighted regarding sampling bias and spatial arrangement of trees and plots (see methods for details); weighted mean dispersal distances are given.

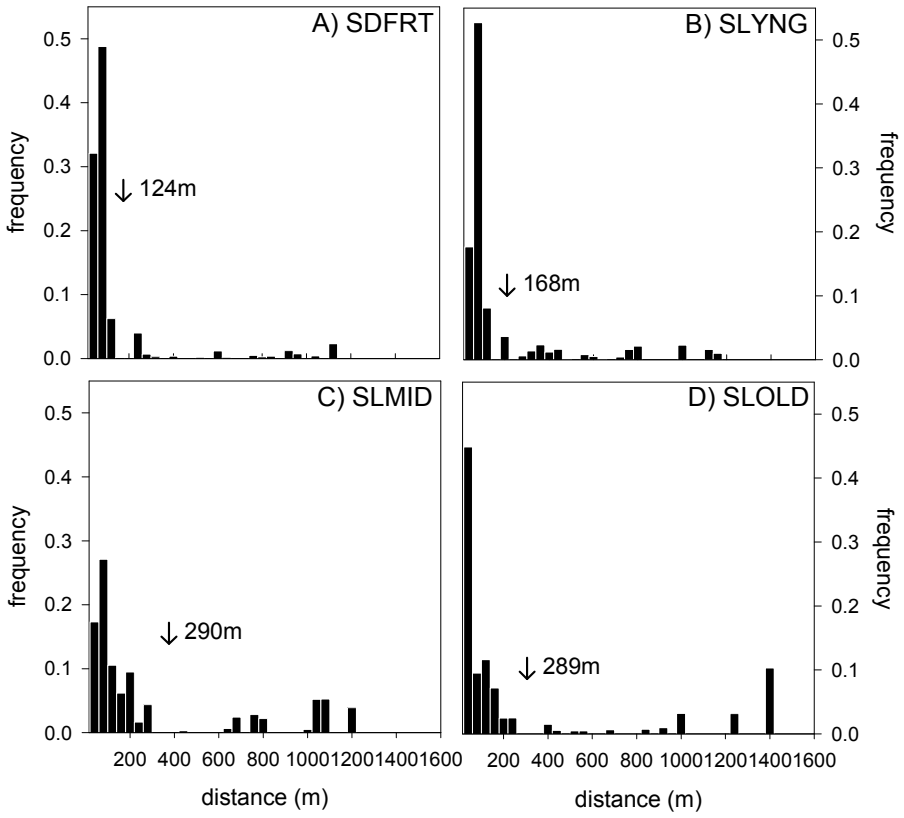

Figure 3.7. Histogram showing the frequency distribution of **paternal gene flow** distances in A) seeds and fruits (SDFRT, n = 76), B) current year seedlings (SLYNG, n = 131), C) middle-old seedlings (SLMID, n = 36, one to three years) and D) old seedlings (SLOLD, n = 36, older than three years). Frequency distributions are weighted regarding sampling bias and spatial arrangement of trees and plots (see methods for details); weighted mean gene flow distances are given.

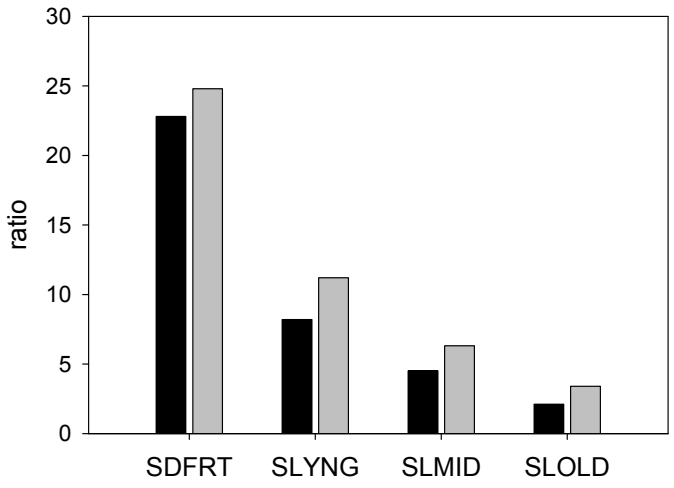

Figure 3.8. Ratios of mean pollen to seed dispersal distances (black bars) and paternal to maternal gene flow distances (grey bars) across life stages, for seeds and fruits (SDFRT, n = 76), current year seedlings (SLYNG, n = 131), middle-old seedlings (SLMID, n = 36, one to three years) and old seedlings (SLOLD, n = 36, older than three years).

3.3.4 Dispersal and gene flow distances across recruitment stages

Pollen dispersal distance distributions were significantly different between SLYNG and SLMID, while no difference in distance distributions was found between SDFRT and SLYNG, as well as between SLMID and SLOLD (table 3.2). However, mean pollen dispersal distances were highly significantly different between all stages, with increasing distances across the first three stages and a decrease in SLOLD compared to SLMID (table 3.1, 3.2).

Seed dispersal/maternal gene flow distance distributions were significantly different between SLYNG and SLMID, while, again, no difference in distance distributions was found between SDFRT and SLYNG as well as between SLMID and SLOLD (table 3.1, 3.2). Mean distances increased from one recruitment stage to the next, from 5 m over 15 m to 73 m and 83 m, whereby the increase in distance to the seed source between SLMID and SLOLD was not significant (table 3.2).

Paternal gene flow distance distributions were again significantly different only for the SLYNG *versus* SLMID comparison (table 3.2). Mean paternal gene flow distances increased significantly across the first three stages (124 to 464 m) and remained constant for SLMID and SLOLD (table 3.1, 3.2).

Table 3.2. Tests for differences in dispersal distance distributions and means of pollen dispersal, seed dispersal/maternal gene flow distances, and paternal gene flow distances between life stages. Presented are D-values of two-sample Kolmogorov-Smirnov tests for comparisons of distributions, as well as t-values of Welch two-sample tests for comparisons of means. Significance (after Bonferroni correction) is indicated with *, n.s. = non-significant. Significant values are bold. SDFRT = seeds or fruits, SLYNG = current year seedlings, SLMID = middle-age seedlings (one to three years), SLOLD = old seedlings (older than three years).

	SDFRT - SLYNG	SLYNG - SLMID	SLMID - SLOLD
Pollen dispersal	0.13 $^{n.s.}$ / **46.13***	**0.40** * **/ 359.9***	0.33 $^{n.s.}$ **/ 132.84***
Seed dispersal/maternal gene flow	0.18$^{n.s.}$ **/ 321.49***	**0.80** * **/ 20.30***	0.23 $^{n.s.}$ / 0.59$^{n.s.}$
Paternal gene flow	0.19 $^{n.s.}$ **/ 204.66***	**0.33** * **/8.73***	0.32 $^{n.s.}$ **/** 0.13$^{n.s.}$

3.4 Discussion

Animal-mediated movement of *P. africana* pollen and seeds is spatially heterogeneous, with pollen dispersal distances being much longer than seed dispersal distances, and paternal much longer than maternal gene flow distances. Furthermore, distances found at initial life stages increased considerably across subsequent stages.

3.4.1 Pollen and seed dispersal and gene flow distances

The comparison of frequency distributions of dispersal and gene flow distances with respect to null models revealed significant limited non-uniform patterns for pollen and seed dispersal/maternal gene flow as well as paternal gene flow. As anticipated, pollen and seed dispersal events, as well as paternal gene flow, occurred over significantly shorter distances than expected under unlimited uniform dispersal. Indeed, these results are in line with findings of other studies on wind-dispersed and other animal-dispersed species that also found highly skewed pollen and seed dispersal distance distributions (e. g. González-Martínez et al 2002, Robledo-Arnuncio & Gil 2005, Garcia et al. 2007, Geng et al. 2008, Carneiro et al. 2009). Obviously, both pollinators as well as seed dispersers in my study system moved over a restricted distance range, potentially mostly within patches of attractive food sources, and rarely travelled large distances. Pollinators are known to move preferentially to neighboring plants under high local resource availability (Rasmussen & Brodsgaard 1992, Ghazoul 2005). Regarding seed dispersal, several studies report short flight distances, especially of small-bodied avian seed dispersers, to neighboring conspecific trees (e.g. Jordano & Godoy 2002, Jordano et al. 2007).

In contrast to my hypothesis, I found no effect of conspecific tree density on dispersal distances, neither for pollination nor for seed dispersal. Thus, movement distances of pollen and seed dispersers are triggered by other factors besides conspecific tree density, for example by overall resource availability, by intra- and interspecific interactions or by behavioral preferences (Schupp et al. 2002, Ghazoul 2005, Jordano et al. 2007). For example, the overall availability of concurrently flowering and fruiting heterospecific trees in the surrounding of my study trees could determine the movement patterns of dispersers (Schupp et al. 2002, Clark et al. 2004, Ghazoul 2005). Pollinators typically are generalists when selecting flowers if no species is dominating the floral community (Kunin 1993). This probably holds for my study area, since *P. africana* did not dominate the floral community, resulting in pollinators flying also to heterospecific

flowering species. For seed dispersal, a number of studies have shown that heterospecific fruiting trees can attract frugivores and can in turn become foci for the dispersal of concurrently fruiting species (Schupp et al. 2002, Clark et al. 2004, Kwit et al. 2004).

3.4.2 Contribution of pollen and seed dispersal to gene flow

For all recruitment stages, I found significantly longer pollen than seed dispersal distances. These findings reflect differences in traits, behavior, and resources of the two groups of dispersal agents (Rasmussen & Brodsgaard 1992, Ghazoul 2005).

Compared to other studies with animal-mediated pollination, mean pollen dispersal distances found here (table 3.1) were similar or longer, depending on the recruitment stage considered (propagules: my study: 114 m, e.g. Garcia et al. 2007: 62.9 m, Geng et al. 2008: 15 m; seedlings: my study: 123 – 325 m, e.g. Konuma et al. 2000: 188-195 m, Hardesty et al. 2006: 345 m). However, seed dispersal distances in my study system were remarkably short, with a mean distance of only 5 m at the propagule stage, and 15 – 84 m at the early to late seedling stages (table 3.1). If I consider only seeds (SDFRT) dispersed beyond the crown of the mother tree (> 0 m dispersal distance), the weighted mean dispersal distances in my study are still rather short (30 m ±5.1 SE m). These distances are much shorter than those found for other animal-dispersed species at the propagule stage (e.g. Grivet et al. 2005: ~ 100m, Garcia et al. 2007: 145 m) and also at the seedling stage (Hardesty et al. 2006: 392 m, Oddou-Muratorio & Klein 2008: 135 m), yet are comparable to the estimates of Valbuena-Carabana et al. (2005, 14-42 m). The relatively short seed dispersal distances quantified in my study system might be attributed to differences in disperser species with different life history (e.g. body size) characteristics, as well as different behavioral preferences. Further, the spatial arrangement of *P. africana* trees, the availability of food resources in the study area etc. can have a strong impact on disperser movement and could potentially lead to the short distances I find here (Schupp et al. 2002, Jordano et al. 2007).

In *P. africana*, pollen dispersal distances were always considerably longer than seed dispersal distances. For seeds or fruits (SDFRT), pollen dispersal distances exceeded seed dispersal distances by a factor of 23. This is in line with the general assumption that gene flow by pollen is spatially more extensive than gene flow by seeds (Levin & Kerster 1974, Petit et al. 2005). However, it contradicts my hypothesis and the

findings of some other studies, where seeds dispersed by vertebrates were found to move over longer distances than pollen dispersed by insects (Garcia et al. 2007), or where the contribution of both was similar (Hardesty et al. 2006, Geng et al. 2008). Insect pollination distances are usually expected to be restricted (Burczyk et al. 2004, Petit et al. 2005). Therefore, I expected insect pollinators to transport pollen over short distances. However, insect pollinators were found to have the capability of moving long distances in other studies as well (Isagi & Kanazashi 2002, Dick et al. 2003, Carneiro et al. 2009). Furthermore, pollen dispersal distances do not necessarily correlate with pollinator flight distances between trees (Rasmussen & Brodsgaard 1992). Apparently, extensive pollen dispersal distances may be a result of frequent pollen carryover from one tree to several subsequent trees (Rasmussen & Brodsgaard 1992).

The short seed-dispersal distances I recorded contradicted my expectations. Birds and monkeys, which are the main seed dispersers in my study system, are highly mobile and can transport seeds over several kilometers (Holbrook et al. 2002, Jordano et al. 2007). Thus, I had expected seed dispersal distances to exceed distances crossed by insect pollinators in my study system. Inferring seed dispersal patterns only from movement patterns of dispersal agents might be misleading, since seed and fruit handling behavior (e.g. dropping of seed and fruits) can yield short seed passage times and might result in short distance dispersal also for highly mobile species (Nathan et al. 2008). It is also important to consider vector displacement velocity instead of its movement capacity (Nathan et al. 2008). *Prunus africana* is known to be very attractive for the local frugivore community, leading to high fruit removal rates (Farwig et al. 2006, Kirika et al. 2008a, Kirika et al. 2008b). Yet, this attractiveness, which potentially results from its nutrient content (Farwig et al. 2006), might imply that animals stay longer in each tree. Hence their displacement velocity is low, leading to short-distance dispersal either beneath the fruiting tree or within local patches of neighboring trees. In fact, monkeys, who are among the main dispersers of *P. africana*, were observed to stay long in a focal tree for foraging (D. G. Berens, pers. observations). Further, they tend to disgorge part of the fruits while foraging, which also causes many seeds to be dropped under the crown.

3.4.3 Dispersal and gene flow distances across recruitment stages

My results show that the initial dispersal distances created at the propagule stage change substantially through further recruitment stages for pollen and seed dispersal as well as paternal gene flow.

According to my hypothesis, effective seed dispersal/maternal gene flow distances changed among recruitment stages. Distances to the seed source continuously and substantially increased across successive stages from 5 m to 15 m, 73 m, and 84 m, suggesting strong J-C effects within my study population. This pattern implies increased mortality of offspring in close proximity to the mother tree, presumably due to distance – or density-dependent predator attack or pathogen infections (Janzen 1970, Connell 1971). Thus, enhanced recruitment probabilities farther away from the mother tree lead to increasing average mother-offspring distances in older stages (Hammond & Brown 1998). My results revealed a significant increase in average distance to the mother plant as early as at the seed-seedling transition. In addition, my results suggest that spacing mechanisms, such as distance- or density-dependent predation, herbivory and pathogens, do not act predominantly on seeds (Janzen 1970) or newly emerged seedlings (Connell 1971, Hyatt et al. 2003) but also across older recruitment stages.

In contrast to my expectations, patterns of pollen dispersal and paternal gene flow distances also differed among recruitment stages. An explanation for this could be that the seedlings I analyzed for older stages might results from different cohorts. Flowering phenology differs between years, resulting in large inter-annual variance in resource distributions (Berens et al., unpublished data) and potentially in variable movement patterns of dispersers (Ghazoul 2005). Because it is difficult to envision a mechanism that can generate such notable difference in pollen dispersal among recruitment stages, I propose that my findings reflect this inter-annual variability. Therefore, estimating pollen dispersal distances from material collected at different recruitment stages should provide more reliable estimates of the average (long-term) distribution of pollen dispersal distances in the population. The unexpected increase of paternal gene flow distances across the first three successive recruitment stages suggests that J-C effects potentially affect paternal contribution to gene flow as well. Mortality might not only be enhanced near the mother tree, but also in close proximity to the father. One potential mechanism that could lead to increased mortality near the father tree is the spread of specific pathogens, as shown in previous studies (e.g. Augspurger & Kelly 1984, Packer & Clay 2000, Mangan et al. 2010). Augspurger and Kelly (1984)

found that seedlings incurred a greater risk of pathogen attack when transplanted under their parents than under other conspecific trees. Pathogens potentially not only act species-specific, but also genotype-specific (Augspurger & Wilkinson 2007). If the susceptibility of the paternal genotype towards pathogens is high, pathogen infection may lead to high levels of mortality in closely related offspring individuals. Thus, J-C effects lead to a higher survival probability of the offspring further away from both parents, entailing a complex reshuffling of the genetic template created by initial dispersal across subsequent recruitment stages.

The relative changes in effective pollen dispersal, seed dispersal, and paternal gene flow distances across recruitment stages elucidate the mechanisms affecting the contribution of the two processes pollen and seed dispersal to gene flow. At the initial seed rain stage, pollen dispersal distances exceed seed dispersal distances by a factor of 22.8. However, across subsequent recruitment stages, the pollen to seed dispersal ratio dropped over 8.2 (SLYNG), 4.5 (SLMID) to 2.1 (SLOLD) (figure 3.8). The same was true for the paternal to maternal gene flow-ratio, which dropped from 24.8 (SDFRT) over 11.2 (SLYNG), 6.4 (SLMID) to 3.4 (SLOLD) (figure 3.8). This implies that during successive recruitment stages, offspring location is shifted away from mother trees at relatively greater rates than from father trees. The paternal gene flow distances are in the order of several hundred meters, with a mean distance of >100 m already at the initial stage of seed dispersal. In comparison, the maternal gene flow distances are initially in the order of 5 m extending to nearly 100 m at later stages. Thus, the area of approximately 100 m radius around parent *Prunus* trees may constitute a high mortality zone for offspring, within which J-C effects are most intense. The long-term consequences of such rearrangement of successive gene flow patterns for local genetic population structure are yet to be evaluated.

Overall, my results show that the interpretation of gene flow patterns strongly depends on the recruitment stage these estimates were based on. Patterns inferred from gene flow estimation from a single stage (most commonly from early life stages) might be misleading, regarding interpretations on the genetic make up of populations or implications for population structure and dynamics.

3.4.4 Implications for population genetic structure

Both pollination and seed dispersal showed spatially restricted dispersal distances within my study area. Highly restricted seed dispersal is expected to yield marked

small-scale spatial genetic structuring within the population, whereby closely related individuals are clustered (Hamrick et al. 1993, González-Martínez et al. 2002, Geng et al. 2008). Spatial aggregation of siblings can lead to fine-scale spatial genetic structuring despite extensive pollen dispersal (Hamrick & Nason 1996). The clustering of related individuals and establishment of spatial genetic structure can have significant impact on population dynamics of the plant species. It can lead to genetic subdivision of the population, influence offspring viability through predation and competition, and promote the adaptation to microenvironmental variation (Schnabel et al. 1998). The patterns found for different recruitment stages in my study population suggest that the initial genetic template created by pollination and seed dispersal undergoes extensive alterations during the multiphase process of recruitment. The initial pattern of highly restricted seed dispersal dissolves during later recruitment stages, which might once more alter the genetic population structure.

Despite the restricted gene flow patterns I documented within my study population, I recorded considerably high levels (36-66%) of gene immigration from the adjacent forest. The level of gene immigration I found at the seed and fruit stage in *P. africana* is far higher than that found in *P. mahaleb* (18.5 %, Garcia et al. (2007)), presumably due to the low isolation of my study stand. Both pollination and seed dispersal have the potential to facilitate gene immigration, and thus can counteract within-stand processes (Burczyk et al. 2004, Jordano et al. 2007, Carneiro et al. 2009). Regarding the limited within-stand seed dispersal distances, I assume that in my study system most of the gene immigration occurs due to long-distance pollination events.

3.5 Conclusions

I found that pollination and seed dispersal patterns mediated by animal dispersal agents are highly spatially heterogeneous, which is in concordance with many previous studies (e.g. Ghazoul 2005, Russo et al. 2006, Garcia et al. 2007). On the one hand, I found that mother-to-offspring (seed dispersal) distances increased at successive recruitment stages, as expected by J-C effects. On the other hand, my results question common simplifications such as inferring dispersal from the identity of the dispersal vectors (e.g., the common assumption that vertebrates disperse genes farther than insects). I also found an apparently unexpected increase in paternal gene flow distances among recruitment stages, which might hint on genotype-specific J-C effects, and, consequently, I call for an exploration of this process. Overall, both pollination and seed

dispersal processes need to be investigated across multiple recruitment stages, to understand and predict gene flow patterns and genetic structuring of animal-dispersed tree populations.

4 FINE-SCALE SPATIAL GENETIC DYNAMICS ACROSS SUCCESSIVE LIFE STAGES

4.1 Introduction

Gene flow via pollen and seed dispersal is a key factor in determining the genetic make-up of plant populations (Loveless & Hamrick 1984, Ouborg et al. 1999). Besides selection processes and historical factors, it is probably the most prevalent factor influencing fine-scale spatial genetic structure (SGS), i.e. the non-random spatial distribution of genotypes within populations (Vekemans & Hardy 2004). If gene flow is restricted, fine-scale SGS can develop within a few generations (Hamrick et al. 1993, Epperson 1995, Vekemans & Hardy 2004). Several field studies have shown this direct relationship between distance-limited pollen and seed dispersal and rapid SGS formation (e.g. Dutech et al. 2002, Ng et al. 2004, Isagi et al. 2007, Voigt et al. 2009). Dutech et al. (2002) found marked SGS at small spatial scales in *Vouacapoua americana*, a neotropical species with restricted dispersal (seed dispersal < 10 m, pollen dispersal < 100 m). In *Shorea leprosula*, Ng et al. (2004) suggested that pollination by insects with weak flying abilities resulted in SGS. In contrast, for most species, the major driver of SGS is rather restricted seed dispersal, because it results in spatial aggregations of siblings in close vicinity of the mother trees (Kalisz et al. 2001, Chung et al. 2003, Trapnell & Hamrick 2005).

The persistence of the initial genetic make-up formed after the seed rain is influenced by various ecological and evolutionary processes across life-stages (Ng et al. 2004). Starting from the genetic template created in the seed rain, mortality due to seed predation or seedling herbivory, pathogen attack or competition, as well as the impact of environmental heterogeneity and microhabitat conditions lead to the thinning of the juvenile population. This is especially true in the tropics, where the influence of mortality factors can be complex (Hammond & Brown 1998, Harms et al. 2000, Augspurger & Wilkinson 2007). If mortality randomly varies with age, the initial SGS in a population is likely to persist to the adult stage (Pacheco & Simonetti 2000, but see González-Martínez et al. 2002). If mortality is non-random and selection processes are acting, the patterns of SGS can change (Hamrick et al. 1993, Kalisz et al. 2001, Jones & Hubbell 2006). One important driver of non-random mortality is the so-called Janzen-Connell effect (Janzen 1970, Connell 1971). Usually, most seeds are dispersed to

distances close to the mother and are therefore deposited in the neighborhood of closely related individuals (Nathan & Muller-Landau 2000, Schupp et al. 2002). Due to density- and distance-dependent predation and pathogen pressure, juveniles may encounter a higher mortality in the vicinity of the mother tree (Janzen 1970, Connell 1971, Hammond & Brown 1998), leading to a thinning of the sib clusters (in cases up to single juveniles; Kalisz et al. 2001) that dilutes SGS (Degen et al. 2001). Other scenarios of postdispersal mortality that can affect SGS can be found in Nathan & Casagrandi (2004).

Studying genetic diversity and SGS of different life stages of a species helps to understand the role of population dynamics and selection in the formation of genetic structure within adult plant populations. So far, studies investigating genetic dynamics across several recruitment stages have found variable results. Chung et al. (2000) and Jacquemyn et al. (2009) neither found changes in genetic diversity nor in SGS across recruitment stages in *Neolitsea sericea* and *Orchis mascula*, respectively. Other studies found that SGS decreases with increasing age, potentially as a result of compensatory mortality due to environmental heterogeneity and competitive thinning (e.g. Hamrick et al. 1993, Ueno et al. 2002, Ng et al. 2004, Yamagishi et al. 2007, Zhou & Chen 2010). SGS increased with life history stage in other study systems, explained by local selection processes or historic events (e.g. Kalisz et al. 2001, Latouche-Hallé et al. 2003, Jacquemyn et al. 2006). However, only a few studies that assessed SGS from a multistage perspective incorporated the initial seed rain stage in their analyses (e.g. Jones & Hubbell 2006, Zhou & Chen 2010). Including this stage is of particular interest, as it reveals the direct influence of pollen and seed dispersal patterns on SGS, before processes like seed predation or selective germination change these patterns. Further, knowledge on the contemporary dispersal patterns within the study population, gained with the help of e.g. parentage analyses, can strongly improve the interpretation of the observed SGS as it provides insights on mechanisms for the development of the genetic make-up of a population (Kalisz et al. 2001, Jacquemyn et al. 2006, Zhou & Chen 2010).

Tree species allow the study of long-term temporal and spatial variation in population genetic structure, as SGS can be modified at variable points of time during the long life span of trees (Chung et al. 2003). However, due to overlapping generations in trees, SGS analyses on pooled life stages can be misleading (Chung et al. 2003). Thus, a multistage perspective can lead to a higher resolution of the underlying

processes and is essential to understand the formation of SGS in trees (Alvarez-Buylla et al. 1996, Chung et al. 2003). However, especially in tropical tree species, multistage analyses of SGS are rare (Ng et al. 2004, Zhou & Chen 2010).

Here, I investigated the development and persistence of SGS in the animal-pollinated and dispersed tropical tree *Prunus africana*. I used a multistage approach, starting from the seed stage that, combined with knowledge on contemporary dispersal patterns, allowed to assess the influence of gene movement on SGS formation in this species. Parentage analyses at different life-stages of *P. africana* have revealed marked distance-restricted seed dispersal within my study population. Further, Janzen-Connell effects apparently have led to increasing distances between offspring and mother trees in older recruitment stages (Berens et al., unpublished manuscript, chapter 3). I thus expected (*i*) to find a marked SGS at the seed stage and (*ii*) that these patterns would dissolve during later recruitment stages due to Janzen-Connell effects. However, during later stages, i.e. at the sapling to adult transition, factors like selection due to microhabitat heterogeneity may become important and, thus, final SGS patterns remain uncertain.

4.2 Methods

4.2.1 Study species

Prunus africana Hook f. (Rosaceae) is an evergreen monoecious tree species whose range expands from East Africa to Madagascar and the Comores (Hall et al. 2000). Due to bark exploitation for medicinal use, the species is listed on CITES appendix II. *Prunus africana* has small white hermaphroditic protogynous flowers. Main pollinators are hymenoptera and diptera. The species seems to be mainly outcrossing (Munjuga et al., unpublished data). Parentage analyses of the species revealed selfing rates of 5 to 13 %, depending on the life stage (Berens et al., unpublished manuscript, chapter 3). Main seed dispersers of the one-seeded fleshy fruits are birds and monkeys (Farwig et al. 2006). Farwig et al. (2008) studied genetic diversity and SGS of *P. africana* adults and seedlings in eight populations over the whole Kakamega Forest including adjacent forest fragments. They found high levels of genetic diversity (H_e) within populations (ca. 0.80) and weak genetic structure among populations, with F_{st} values of 0.009-0.020, depending on the life-stage (Farwig et al. 2008).

4.2.2 Study area

Kakamega Forest is located in western Kenya (figure 4.1). This mid-altitudinal rainforest (1,500 – 1,700 m asl, KIFCON 1994) is supposed to be the easternmost remnant of the Guineo-Congolian rainforest belt (Kokwaro 1988). The forest comprises a main forest block and five fragments with a total area of 13,000 ha (Lung 2004). Mean daily temperatures range from 10.6 to 27.7°C (Tsingalia 1990). Average yearly precipitation is about 2000 mm/year (averaged from Forest Department records at Isecheno Forest Station from 1982 to 2005).

The study area within Kakamega Forest is a 120 ha forest peninsula in the northwest of the unfragmented main forest block extending into the agricultural matrix (0°21´34.1´´ North, 34°51´31.0´´ East; figure 4.1). It comprises primary and secondary forest and is situated in a well-protected area with little human disturbance (Bleher et al. 2006).

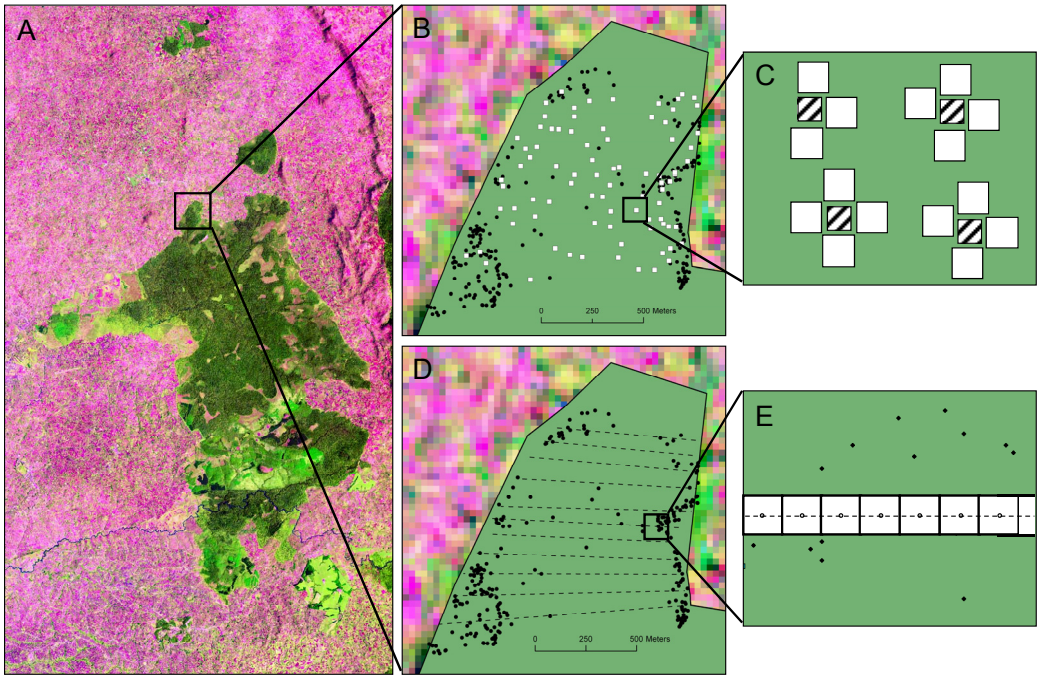

Figure 4.1. (A) Satellite image showing Kakamega Forest with forested area (dark grey) and surrounding farmland (light grey) and the location of the study area within the forest; (B) spatial distribution of adult Prunus africana trees (n = 261, black dots) and of sampling plots (n = 86, white squares) for sampling of fruits and seeds (SDFRT) and current-year seedlings (SLYNG); (C) each sampling plot consisted of four seed traps (0.5 m², hatched squares) and 15 seedling sub-plots (1 m², white squares); (D) spatial distribution of adult P. africana trees (black dots) and of eleven transects (dashed lines) for sampling of middle age seedlings (SLMID) and old seedlings (SLOLD); (E) arrangement of virtual plots (white squares) for the quantification of SLMID and SLOLD.

4.2.3 Plant material

In 2005, I established 22 parallel transects crossing the study area from the eastern to the western forest edge and covering ca. 87 ha (figure 4.1). Transects were between 694 and 1,225 m long and were separated by 40 m. In total, I detected 261 *P. africana* adult trees (diameter at breast height > 10 cm; ADLTS hereafter) in the transect area and an adjacent 100 m northern and southern buffer zone. I took GPS coordinates of each tree. Leaf material of all adult trees was collected between January and May 2006.

In the same year, 86 geopositioned sampling plots were randomly established along the transects (figure 4.1B). Each plot consisted of an array of four seed traps (0.7 x 0.7 m) made of a wire frame covered with mosquito mesh (figure 4.1C). I collected propagules (SDFRT) from seed traps on a weekly basis between January and April 2006 during a fruiting peak of *P. africana* in my study site. They were assigned to fruits (with pulp) or seeds (without pulp) and pooled for further analyses. I manually removed fruit pulp for storage and dried all samples with silica gel. Of the 8,544 propagules that were counted in traps, I randomly selected a sample of 311 for genetic analyses.

Current year seedlings (SLYNG hereafter) were collected at the same 86 sampling plots between March and May 2006. I placed fifteen 1m² sub-plots within each sampling plot (figure 4.1C) and counted and collected newly emerged seedlings on a biweekly basis. Leaf samples were dried with silica gel. A sample of 309 of the 48,991 seedlings was used for genetic analyses.

I sampled older seedling stages between January and February 2008 along every second of the 22 transects crossing the study area (figure 4.1D). The area covered along these eleven transects was divided into virtual 10 x 10 m sampling plots (figure 4.1E; the plots were termed virtual plots because they were not marked on the ground), thereby covering an area of ten ha, and the centroid of each plot was used as its geographic position. Each *P. africana* seedling encountered in these virtual plots was recorded. A leaf sample of each of these individuals was taken, and seedling age was determined by an experienced field assistant, considering the degree of lignification of the stem axis. A sample of 298 of the 861 seedlings up to three years of age (SLMID hereafter) and 301 of the 368 seedlings and saplings older than three years (SLOLD hereafter) was taken for genetic analyses.

4.2.4 Molecular analyses

All plant material (leaves of ADLTS, SLYNG, SLMID, SLOLD, and embryo tissue of SDFRT) was ground to fine powder with a Retsch mixer mill (Type MM 301, Retsch, Haan, Germany). I extracted DNA following the protocol described by Wang et al. (1993). The standard protocol was modified by using 40 µl of NaOH and 95 µl of Tris buffer for each sample. Samples were genotyped at six microsatellite loci that were originally developed for *Prunus avium* and *Prunus persica*: UDP97-403 (Cipriani et al. 1999), P12A02 (Sosinski et al. 2000), BPPCT-002 (Dirlewanger et al. 2002), UDP96-005 and UDP98-410 (Schueler et al. 2003), and EMPaS06 (Vaughan & Russell 2004). Primers were fluorescent 5'-end-labeled (6-FAM, NED, HEX; Applied Biosystems, Foster City, CA, USA). Polymerase chain reaction protocols followed Farwig et al. (2008) with annealing temperatures ranging from 54 to 62°C and number of cycles ranging from 44 to 46. For loci UDP97-403 and EMPaS06, a final extension step of 60°C for 30 min was added to the standard protocol. I genotyped samples on ABI 3730 and 3130xl capillary sequencers. Allele scoring was performed using the software GeneMapper 4.0 (Applied Biosystems, Foster City, CA, USA). For all loci, base shifts between scorings of the different sequencers were calibrated by running 48 samples per locus (3%) on both instruments. All samples, except that of one adult tree, were successfully amplified.

4.2.5 Genetic diversity and SGS

Allelic richness, which is a measure of number of alleles per locus independent of sample size, was assessed for each locus and life stage using the software FStat 2.9.3.2 (Goudet 1995). I computed observed (H_o) and expected heterozygosity (H_e) in CERVUS 3.0.3 (Kalinowski et al. 2007). Differences in H_o and H_e within life-stages were tested with paired t-tests. I conducted ANOVAs to test for differences in allelic richness, H_o, and H_e among life stages. All statistical analyses were done in R 2.9.0 (The R Foundation for Statistical Computing).

Fine-scale spatial genetic structure (SGS) for each life stage was assessed based on pairwise kinship coefficients between individuals (f_{ij}, Loiselle et al. 1995) using SpaGeDi (Hardy & Vekemans 2002). The kinship coefficient f_{ij} measures the extent of similarity between pairs of individuals i and j relative to the mean similarity of two randomly drawn individuals from the population (Hardy & Vekemans 2002). A regression between f_{ij} values and the logarithm of geographic distances between

individuals (as correspond to bidimensional space) was calculated to provide the regression slope *b* and to test for significant SGS patterns. For graphical representation, multilocus f_{ij} values were averaged over distance classes. I set class boundaries every 50 m for the first 200 m, every 100 m up to 500 m and every 200 m up to the maximum distance between individual pairs, i.e. ca. 1,700 m. Standard errors of the slope and of average kinship coefficients were computed using a jackknife procedure over loci. Values were tested for significance using 10,000 permutations. The extent of SGS was also quantified using *Sp*-statistics (Vekemans & Hardy 2004), calculated as $-b/1\text{-}F_{(1)}$, whereby *b* is the regression slope defined above and $F_{(1)}$ is the average kinship of neighboring individuals (approximated here by the average kinship coefficient of the closest distance interval).

To test whether the different sampling methods for SDFRT/SLYNG (figure 4.1B, C) and SLMID/SLOLD (figure 4.1D, E) affect SGS estimates, I used a resampling procedure. A random sample of 100 individual pairs in each distance class was drawn with replacement, and the kinship coefficients of these pairs were averaged. The procedure was repeated 1,000 times. This was also done for *Sp*-values, whereby mean resampled values for *b* and $F_{(1)}$ were used. Overall, means of resampled kinship coefficients per distance class were significantly different from original values only in distance class 700-900 m for SDFRT and SLYNG, where mean resampled kinship coefficients were slightly higher than original values. All other mean resampled values as well as *Sp*-values based on mean resampled *b* and $F_{(1)}$ values were not significantly different from the original values (all *p*-values > 0.05). Thus, I consider the effect of different sampling schemes on my results as negligible.

To better understand the relative contributions of pollen and seed dispersal to overall gene flow, I analyzed the shape of the kinship-log(distance) correlograms at short distances as outlined in Heuertz et al. (2003). Shapes of SGS curves were estimated by fitting polynomial functions of the third degree (i.e., $y = a + bx + cx^2 + dx^3$) on the logarithm of distance to the standardized f_{ij} residuals. In short, a concave shape of kinship-log(distance) regression at short distances points to more restricted seed than pollen dispersal. Conversely, a convex shape indicates longer seed gene flow as compared to pollen gene flow (Heuertz et al. 2003, De Lucas et al. 2009).

4.2.6 Sibling pairwise distances across life stages

Finally, to provide insights on mechanisms behind changes in SGS across life stages, I used data on contemporary gene flow patterns in my *P. africana* population. In a different study, I conducted parentage analyses for the four life stages SDFRT, SLYNG, SLMID, and SLOLD (Berens et al., unpublished manuscript, chapter 3). Assignment rates of at least one parent within the ADLTS population were 68 % for SDFRT (n = 212/311), 85 % for SLYNG (n = 364/309), 45 % for SLMID (n = 136/298), and 65 % for SLOLD (n = 195/301). Using this information, I was able to group the individuals into family groups, i.e. groups of offspring that had at least one parent in common (siblings hereafter). Assuming random mortality with age, the distribution of distances between siblings, as well as the mean distance should be consistent across life stages, given large enough sample sizes. With Janzen-Connell effects driving potential changes in SGS across life stages, distances between siblings should increase in older life stages. I identified 37, 49, 51, 50 family groups for SDFRT, SLYNG, SLMID and SLOLD, respectively, each consisting of two to 28 siblings (SDFRT: mean 6.0 ± 5.4 SD, SLYNG: mean 6.9 ± 7.3 SD, SLMID: mean 2.8 ± 1.3 SD, SLOLD: mean 3.3 ± 1.9 SD). Mean pairwise distances between all members of a family group were computed and mean distances within family groups were compared across life stages using ANOVAs and post-hoc pairwise t-tests as implemented in R 2.9.0 (The R Foundation for Statistical Computing).

4.3 Results

4.3.1 Genetic diversity

The microsatellite loci were highly polymorphic with a mean allelic richness of 18.4 ± 1.4 SD (from 16.7 in SDFRT to 29.2 in ADLTS; table 4.1). Observed heterozygosity (H_o) ranged between 0.66 (SDFRT) and 0.77 (ADLTS), while gene diversity was very similar (0.82-0.83) across all life stages (table 4.1). In addition, H_o was marginally or significantly lower than H_e in all cases (table 4.1). Neither allelic richness, nor H_o or H_e differed among life stages as shown by ANOVA (all p-values > 0.05).

Table 4.1. Mean allelic richness, observed heterozygosity (H_o), and expected heterozygosity (H_e) across six microsatellite loci, t-values of pairwise t-tests for difference in H_e and H_o with significance, mean kinship of neighboring (< 50 m distance) individuals, $F_{(1)}$, and standard error, SE, regression slope, *b*, of the kinship-log(distance)-regression and *Sp*-statistics in five different life stages of *Prunus africana*. SDFRT = seeds and fruits, SLYNG = current-year seedlings, SLMID = middle-old seedlings (one to three years old), SLOLD = old seedlings (more than three years old), ADLTS = adult trees (dbh > 10cm). *** = p<0.001, * = p<0.05, $^{+}$ = 0.05<p<0.1

	Allelic richness	H_o	H_e	t	$F_{(1)}$ ±SE	*b*	*Sp*
SDFRT	16.67	0.66	0.82	7.1***	0.10 ±0.001	-0.014	0.015
SLYNG	18.06	0.75	0.82	2.3^{+}	0.10±0.002	-0.015	0.017
SLMID	19.64	0.69	0.82	2.4^{+}	0.04±0.001	-0.007	0.008
SLOLD	19.06	0.67	0.82	3.5*	0.02±0.001	-0.005	0.005
ADLTS	20.30	0.77	0.83	2.7*	0.05±0.002	-0.013	0.013

4.3.2 Fine-scale spatial genetic structure (SGS)

I found pronounced SGS in all life stages. The slope *b* of the regression between f_{ij} and the logarithm of geographic distance was significantly negative in all cases (figure 4.2). Both, the regression slope *b* and $F_{(1)}$, i.e. the average kinship of neighboring individuals, were higher for early life stages (SDFRT and SLYNG) than for older seedlings (SLMID and SLOLD), whereas adults (ADLTS) showed intermediate values (figure 4.2, table 4.1). This was mirrored in the *Sp*-statistics, where values for SDFRT and SLYNG were two to three times higher than those for SLMID and SLOLD (figure 4.3, table 4.1), indicating stronger SGS in these life stages. Finally, fitting of polynomials to kinship-log(distance)-regressions revealed a concave shape at short distances for all life stages. This suggests that historical seed dispersal has been more restricted than pollen dispersal in my study system.

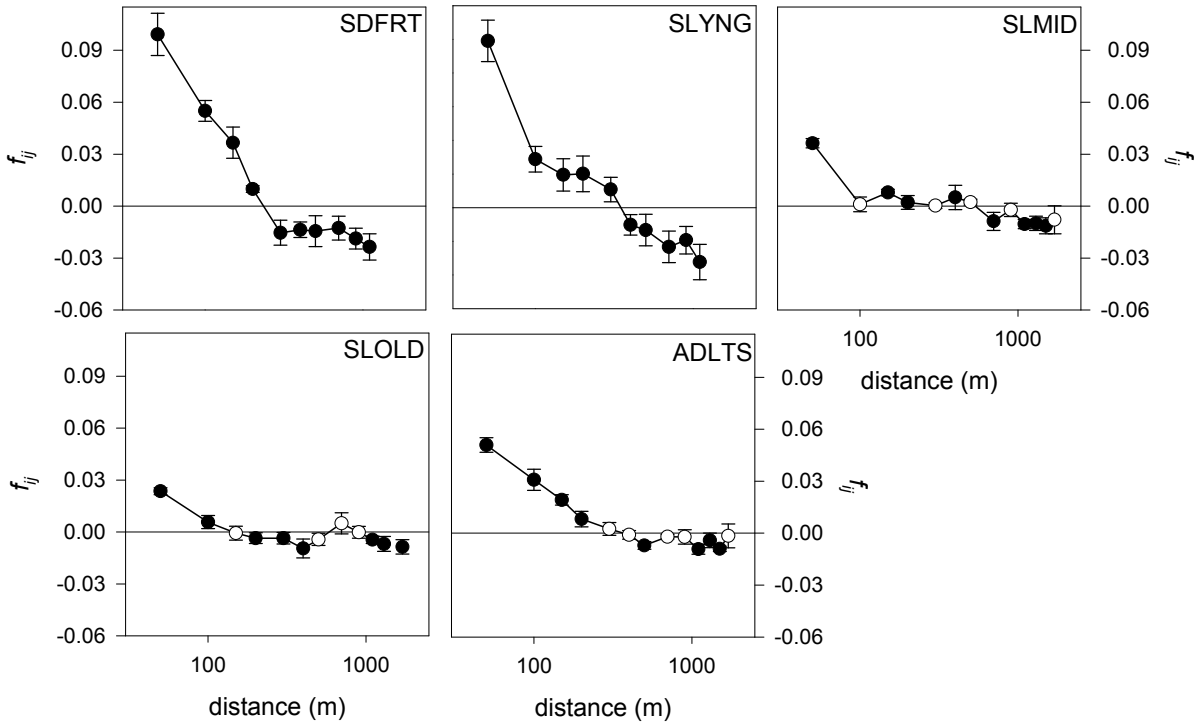

Figure 4.2. Correlograms of average pairwise kinship coefficients (f_{ij}) per distance class (see methods for details on distance classes) and log geographic distance for five life stages of *Prunus africana*. Displayed are multilocus jackknife-values and standard errors. Filled circles are significantly different from zero (after 10000 permutations), open circles are not. SDFRT = seeds and fruits, SLYNG = current-year seedlings, SLMID = middle-old seedlings (one to three years old), SLOLD = old seedlings (more than three years old), ADLTS = adult trees (dbh > 10cm).

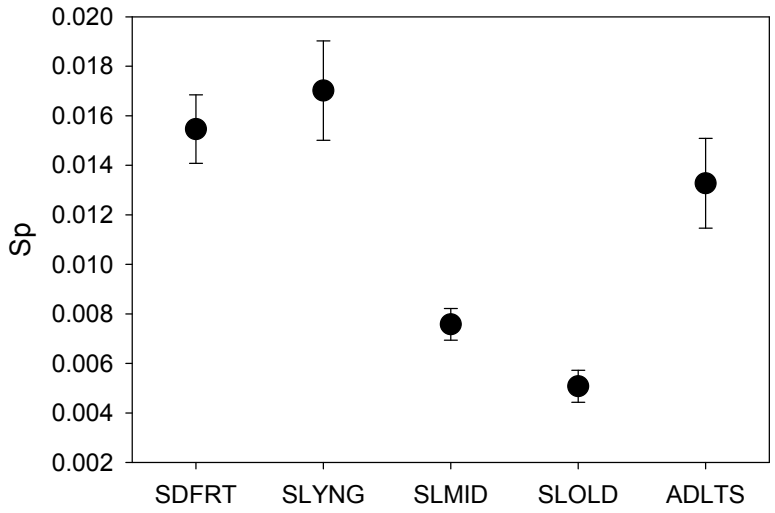

Figure 4.3. Sp-statistics ± standard error, representing the extent of fine-scale spatial genetic structure for five life stages of *Prunus africana*. SDFRT = seeds and fruits, SLYNG = current-year seedlings, SLMID = middle-old seedlings (one to three years old), SLOLD = old seedlings (more than three years old), ADLTS = adult trees (dbh > 10cm).

4.3.3 Sibling pairwise distances across life stages

Testing for the prevalence of Janzen-Connell effects revealed that distances between siblings within family groups were significantly affected by life-stage (ANOVA, $F_{3,182}$ = 2.8, p = 0.039). Thereby, pair-wise distances between siblings were marginally significantly longer for SLMID and SLOLD, respectively, as compared to SDFRT (post-hoc pairwise t-test, $p_{SLOLD\text{-}SFDFRT}$ = 0.08; $p_{SLMID\text{-}SDFRT}$ = 0.06; all other comparisons p > 0.1). Thus, in younger stages, more siblings are found in close proximity to each other, while remaining siblings at older stages are found further away from close relatives.

4.4 Discussion

All life stages of *P. africana* from Kakamega forest showed fine-scale SGS. While genetic diversity did not differ across life stages, the extent of SGS changed during the life cycle of *P. africana*. In addition, and also for all life stages, I found lower observed (H_o) than expected (H_e) heterozygosity. Null alleles, which are one potential cause for this pattern, were not frequent in my data (frequencies around or below 0.2) and are therefore considered unproblematic (Dakin & Avise 2004). Therefore, it seems more likely that the pattern I observed was driven by inbreeding. Although selfing does not seem to be common in the species (Munjuga et al., unpublished data, own unpublished

results), the significant $F_{(1)}$-value for adults indicates a close genetic relationship of reproductive nearby individuals. As a large proportion of pollination events normally occurs within patches (Berens et al., unpublished manuscript, chapter 3), biparental inbreeding (i.e. mating between relatives) may explain the observed heterozygosity deficit.

Genetic diversity, in terms of allelic richness and heterozygosity, was similar throughout life stages in *P. africana*. This is in line with other studies that have also shown that genetic diversity is generally similar across life stages within populations (e.g. Chung et al. 2003, Kelly et al. 2004, Jacquemyn et al. 2006). It shows that effective population sizes were large enough at each life stage to prevent the effects of genetic drift. The levels of genetic diversity in my study were similar to those of another study on *P. africana* in Kakamega Forest (Farwig et al. 2008) and to those of European *Prunus* species (Stoeckel et al. 2006, *P. avium*: H_e: 0.448-0.918), and higher than diversity in an Asian *Prunus* species (Pakkad et al. 2003, *P. cerasoides*: H_e: 0.292-0.689).

The extent of SGS, as estimated by *Sp*-statistics (0.005 – 0.017), found in *P. africana* is comparable to that of other predominantly outcrossing species (Vekemans & Hardy 2004). Further, *Sp*-values are within the range of other animal-pollinated and animal-dispersed species, such as *Psychotria nervosa* (*Sp*: 0.012), *Carapa procera* (*Sp*: 0.011), or *Neolitsea serica* (*Sp*: 0.005) (Vekemans & Hardy 2004). The preponderance of SGS I found in all stages can potentially be ascribed to limited seed dispersal in *P. africana*. The concave shape of the kinship-distance curves suggests that historical seed dispersal might have been more restricted than pollen dispersal. A direct assessment of contemporary pollen and seed dispersal distances using parentage analyses has revealed a 22.8-fold difference (Berens et al., unpublished manuscript, chapter 3). The mean seed dispersal distance was 5 m, whereas pollen dispersal occurred over mean distances of 114 m. Thus, the patterns of SGS I found in the present study agree with these direct observations. Further, the results of my study are in line with other studies that have also found a correlation between restricted seed dispersal and SGS (e.g. Dutech et al. 2002, Isagi et al. 2007).

In contrast to genetic diversity, the extent of SGS differed across life stages in *P. africana*. Early stages (SDFRT and SLYNG) had higher levels of SGS when compared to the reproductive ADLTS population. Then, SGS decreased from these initial values towards older juvenile stages (SLMID and SLOLD), but increased again

towards ADLTS. Other studies looking at SGS from a multistage perspective have either found a consistent increase (Kalisz et al. 2001, Latouche-Hallé et al. 2003, Jacquemyn et al. 2006), decrease (e.g. Parker et al. 2001, Ueno et al. 2002, Chung et al. 2003, Zhou & Chen 2010), or no change in SGS with age (Berg & Hamrick 1995, Chung et al. 2000, Jacquemyn et al. 2009). This absence of a common pattern highlights the importance of species- and site- specific postdispersal and early selection effects. To my knowledge, only one study found a variable pattern comparable to the one I found in the present study. In their study on *Jacaranda copaia*, Jones & Hubbell (2006) detected a decrease in SGS from young seedlings to saplings, and an increase from saplings to adults.

One potential cause for the observed decrease in SGS from early to late recruitment stages could be interannual variation in allele frequencies within the older juveniles (Parker et al. 2001). Indeed, SLMID and SLOLD stages comprise individuals from different cohorts. Variable mating patterns of *P. africana* trees in these years could have led to an overall veiling of the SGS formed in progeny of a single year. Nevertheless, the decrease in SGS can also reflect non-random thinning processes processes (Zhou & Chen 2010). Higher mortality of seeds and young seedlings near the mother trees (i.e. Janzen-Connell effects) is strongly supported by parentage analyses (Berens et al., unpublished manuscript, chapter 3). Both, distances among siblings within family groups (present study) and between mother trees and their progeny (as identified by parentage analyses) tend to increase with increasing age. Several studies, carried out mostly on tropical trees, have shown that mortality due to seed predation and seedling herbivory can be enhanced in close vicinity of the parent tree and in areas of high conspecific density (Janzen 1970, Connell 1971, reviewed in Hammond & Brown 1998). Furthermore, mortality in close vicinity to relatives could be caused by selective pathogens (Augspurger & Kelly 1984, Mangan et al. 2010). If susceptibility of a certain genotype towards pathogens is high, pathogen infection might lead to high levels of mortality in individuals related to this genotype, and thus, to thinning of genetically related clusters.

Surprisingly, SGS was higher in adult trees than in late juvenile stages. Such an increase in spatial genetic structure can have several explanations. First, a higher SGS in adults than in late juveniles could be caused by historical processes. For example, a founder effect with few, related individuals in the founding generation could lead to significant SGS in the adults and strong SGS in the descendants (Sezen et al. 2005,

Jones & Hubbell 2006). This strong genetic structure would dissolve with time as the gene pool is homogenizing (Jones et al. 2006). Such a process is unlikely in my case as the *P. africana* population that I studied is situated in the main forest block of Kakamega Forest, in an area of mostly near-natural forest that has not been clear-cut at least since 100 years (Mitchell et al. 2009). Furthermore, if the adult population had established from a founder generation, adults should show a lower allelic richness as compared to younger stages, as allelic richness would have increased with time due to gene flow (Jones & Hubbell 2006). However, this is not the case in my study. Further, changes in dispersal processes of pollen and seeds with time could lead to differences in SGS between juveniles and adults. Farwig et al. (2008) suggested that genetic differentiation among *P. africana* adults and their seedlings was a result of reduced gene flow in seedlings due to anthropogenic disturbances. However, this explanation seems unlikely for my findings, because genetic diversity in my case is comparable across life stages.

Second, the increase in spatial genetic structure from juveniles to adults could be due to intergeneration variation in demographic processes. Jones & Hubbell (2006) observed a five- to ten-fold rise in relatedness in the smallest distance class from small to large diameter classes. They allocate this pattern to non-equilibrium demographic processes within their study population of *Jacaranda copaia*. This species (and many others, e.g. *Tetragastris panamensis*, *Quararibea asterolepis*, De Steven & Wright 2002; Williamson & Ickes 2002) shows strong temporal variation in reproduction, so that they assumed their adult trees to be descendants of one or few past recruitment events, leading to high SGS among adults. Phenology data suggest that there is no reason to suppose that my adult generation was based on fewer recruitment events than the new one (Berens et al., unpublished data). Still, the range of dispersal and density-dependent selection processes may have been different at the time the adult generation established, due to e.g. anthropogenic influences or different microhabitat conditions (Farwig et al. 2008). Thus, the increase in SGS from SLOLD to ADLTS could reflect differences in dispersal and establishment processes across generations.

Finally, the increase in SGS from juveniles to adults could be driven by microhabitat heterogeneity and local selection (Linhart & Grant 1996, Kalisz et al. 2001, Jones & Hubbell 2006). Related individuals might have similar microhabitat requirements. Thus, spatial heterogeneity in abiotic conditions, e.g. soil nutrients or light environment, allow for the survival of different groups of related individuals

(Kalisz et al. 2001, Jones & Hubbell 2006). Furthermore, spatially structured diseases or genotype-dependent interactions with mycorrhiza could lead to the reclustering of related individuals (Kalisz et al. 2001). For long-living trees, overlapping generations can also drive SGS. If related individuals require similar microhabitat conditions, and generations of related individuals overlap, neighboring individuals can be a mixture of related parents and progeny, leading to pronounced SGS (Latouche-Hallé et al. 2003). For *P. africana*, this appears to be the most likely explanation.

4.5 Conclusions

In this study I showed that changes in fine-scale genetic structure across the life stages of a tree can be variable and complex. Changes in genetic structure can be seen as a multi-step process. Provoked mostly by distance-restricted seed dispersal, the seed rain and early seedling stages showed significant SGS, resulting in spatial clusters of closely related individuals. During the next transition to older seedling and sapling stages, density- and distance-dependent non-random mortality appears to have led to the thinning of family clusters and, eventually, to the reduction of SGS. In addition, across generations, spatial heterogeneity in microhabitat conditions, overlapping generations, and non-equilibrium processes may lead to differences in SGS. These results highlight the importance of a multistage perspective to understand genetic patterns within populations. Further, the comparison of data on SGS and on contemporary gene flow across different life stages helped to assess the mechanisms that shape the genetic structure of trees.

5 GENERAL CONCLUSIONS

Pollination and seed dispersal are key processes for the dispersal of plant species, thereby laying the foundation for further demographic and genetic processes within populations. Understanding how animals shape gene flow and genetic structure, and how these initial patterns influence further recruitment, is a central issue in plant population biology. In my thesis, I studied animal-mediated gene flow distances and resulting spatial genetic structuring (SGS) across several life stages in the animal-pollinated and dispersed tree species *Prunus africana* in western Kenya.

5.1 Dynamics of contemporary gene flow by pollen and seeds across successive life stages

In a first approach, I investigated dispersal and contemporary gene flow distances in *P. africana* across four subsequent early life stages, from seed rain to older juveniles, applying microsatellites and parentage analysis. I aimed to assess pollen and seed dispersal and resulting maternal and paternal gene flow distances, tried to disentangle the relevance of pollen and seed dispersal for gene flow, and tracked the changes of the initial patterns during the process of recruitment. First, I compared observed distances of pollen and seed dispersal as well as maternal and paternal gene flow to patterns expected under the assumption of unlimited uniform dispersal, and tried to elucidate how the density of conspecific trees affects these distances. Further, I compared the extent of pollen and seed dispersal and resulting gene flow. Last, I tested in how far initial gene flow patterns changed across three later life stages. Observed pollen dispersal, paternal gene flow, and seed dispersal/maternal gene flow distances were significantly distance-limited and non-uniform. Both pollinators and seed dispersers in my study system seemed to move over a restricted distance range. Movement distances were not affected by conspecific tree density in the vicinity. The comparison of pollen and seed dispersal distances revealed significantly longer pollen dispersal distances than seed dispersal distances. Pollen dispersal distances considerably exceeded seed dispersal distances by a factor of 22.8 (114 m: 5 m) at the initial seed rain stage. The same was true for gene flow, where pollen-mediated paternal gene flow was 24.8 times longer than seed-mediated maternal gene flow. Thus, pollen dispersal was the major vector of long-distance gene flow in my study population. This was unexpected, as seed

dispersers of *P. africana* are highly mobile species, e.g. bulbuls, hornbills and different monkey species that can transport seeds over several kilometers (Holbrook et al. 2002, Jordano et al. 2007). It shows that inferring seed dispersal from movement patterns of dispersing animals may be misleading (Nathan et al. 2008). The initial patterns changed substantially across further recruitment stages. Seed dispersal or maternal gene flow distances, i.e. distances between mother tree and offspring, increased significantly across successive stages, from 5 m in the seed rain stage to 84 m in old juveniles. Obviously, strong Janzen-Connell effects were acting within my *P. africana* population, whereby survival of offspring increases with increasing distance to the mother tree (Janzen 1970, Connell 1971). I could further show that effective pollen dispersal distances were also variable across recruitment, which stresses the importance of considering multiple recruitment stages for a sound evaluation of dispersal processes and gene flow within plant populations. Surprisingly, I also found an increase in paternal gene flow distances across recruitment stages. This indicates that mortality might not only be enhanced near the mother tree, but also in close proximity to the father. Selection due to genotype-specific pathogens might be a mechanism driving this pattern. The non-random mortality in close vicinity to mother and father tree led to a decrease of the pollen to seed dispersal ratio, as well as the paternal to maternal gene flow ratio from 22.8 and 24.8 to 2.1 and 3.4, respectively, in old juveniles. The relative changes in effective pollen dispersal, seed dispersal, and paternal gene flow distances elucidate the changes of the contribution of the two processes pollen and seed dispersal to overall gene flow during recruitment, which can have long-term consequences for local genetic population structure.

5.2 Fine-scale spatial genetic dynamics across successive life stages

In a second approach, I did a multistage analysis to study genetic diversity and the formation of SGS in the same *P. africana* population. I analyzed five life stages, from seed rain to adult trees, with the same six microsatellite loci applied in the first approach. Combining these data with contemporary gene flow data obtained in my first approach, I tried to understand the influence of pollen and seed gene flow patterns on the formation and persistence of SGS. Genetic diversity, in terms of allelic richness, as well as observed and expected heterozygosity, did not vary across life stages. This shows that the effective population sizes were large enough to counteract random genetic drift. I found SGS in all five life stages, most likely due to the distance-

restricted seed dispersal I had found in my previous approach. Levels of SGS were comparable to other animal-pollinated and dispersed species (Vekemans & Hardy 2004). Further, SGS differed between life stages, being higher in early life stages than in adults, and decreasing towards older juvenile stages. I allocated this decrease in SGS from early recruitment stages to older juveniles to non-random mortality due to Janzen-Connell effects within my study population. I was able to underpin this conclusion with data obtained from parentage analyses, where I could show that both distances between offspring and parent trees, as well as distances between siblings increased in older recruitment stages. Surprisingly, SGS was higher in adult trees than in older juveniles. This increase in SGS could be caused by several factors. I do not consider historical processes, such as founder effects, to be a likely explanation for the observed pattern in my study population. I rather assume that the increase in SGS towards adults could be caused by intergeneration variation in demographic processes. It could reflect differences in dispersal and establishment processes across generations (Farwig et al. 2006). Further, microhabitat selection could be a mechanism driving the pattern (Kalisz et al. 2001, Jones & Hubbell 2006). Related individuals may have similar microhabitat requirements for long-term survival, leading to the spatial grouping of relatives in the adult stage. In summary, the analysis of SGS across life stages, as well as the combination of these data with contemporary gene flow patterns allowed me to disentangle some of the complex processes and mechanism leading to genetic structure within a tree population.

5.3 Synthesis

Overall, in my thesis, I was able to gain significant findings of the consequences of animal pollination and seed dispersal for plant populations. On the one hand, my results question common simplifications, such as inferring dispersal from the identity and movement capacity of the dispersal vectors, e.g. the common assumption that vertebrates disperse genes farther than insects. My results showed that, despite high mobility of seed disperser species, factors like fruit and seed handling behavior, as well as actual movement patterns of animal dispersers need to be considered to get reliable estimates of dispersal and gene flow. Here, molecular genetic techniques, such as parentage analysis, proved to be an elegant tool to assess the actual dispersal patterns without time-consuming field observations, feeding and telemetry studies. Further, in my study, I was able to utilize the unique strength of genetic methods in distinguishing

the specific mother and father trees of offspring at different locations and different life cycle stages. This allowed resolving some of the processes acting across recruitment. In fact, my findings of strong Janzen-Connell effects within the study population related to both mother and father tree showed that factors like non-random mortality can severely alter the initial gene flow template created by dispersers. Gene flow distances changed significantly from the seed stage to older recruitment stages, which was reflected in changes in SGS across life stages. Further, Janzen-Connell effects relating not only to the maternal, but also to the paternal tree add another level of complexity to plant community dynamics. The fact that offspring shift away from maternal and paternal trees at different rates, can have long-term consequences for local genetic population structure that are yet to be evaluated. Furthermore, combining my results of direct gene flow estimates with an analysis of SGS across life stages facilitated the interpretation of mechanisms driving the formation of genetic structure within the population. The gene flow patterns I directly assessed with the help of parentage analysis, e.g. the strong Janzen-Connell effects, were mirrored in patterns of spatial genetic structure among life stages. Here, again, using a multistage analysis proved to be a valuable approach to track the formation of genetic structuring across recruitment.

In my thesis, I could show that for assessing genetic patterns within plant populations, analyses across the life-cycle of the plant species are indispensible. Further, the combination of contemporary gene flow data and SGS analysis is a powerful approach that allows for a high resolution of genetic processes within populations, and can help to interpret consequences for population dynamics and evolutionary processes.

5.4 Outlook

A future challenge will be to extend the approach I used for the model species *P. africana* to broader research directions. Investigations should be extended to different plant species, in order to compare the influence of life-history traits, such as life form, mating system, or different attraction strategies for pollinators or seed dispersers, on genetic dynamics within and between generations. Garcia-Cruz et al. (2009) for example, showed that pollinator attraction strategies of different *Govenia* species have an effect on reproductive success, gene flow, and genetic structure of the species. Multi-stage analyses could help to elucidate the mechanisms leading to these differences, and especially, help to understand how they transfer into long-term differences in genetic

patterns of the species. Further, investigating the role of different dispersal vectors for gene flow and long-term dynamics of plant populations is a challenging issue for future research. Considering multiple dispersal vectors acting at different scales is important for the understanding of large-scale dynamics of plant populations (Jordano et al. 2007, Nathan et al. 2008). Thereby, consequences of dispersal by non-standard dispersal vectors, i.e. dispersal vectors different from the ones that can be inferred from the phenotypic characteristics of the plant, for gene flow processes is mostly unclear (Nathan et al. 2008). Further, a sound combination of genetic and ecological methods is still rarely applied to investigate the role of plant-animal interactions in shaping genetic patterns in plants (e.g. Zhou & Chen 2010), even though an integration of methods could vastly advance our knowledge. Tracking disperser movement by telemetry studies or assessing disperser behavior, and relating these data to seed deposition patterns, resulting regeneration dynamics and effects on genetic patterns can lead to a complete picture of consequences of plant-animal interactions for plant communities (e.g. Zhou & Chen 2010). Last, in my opinion, a major future challenge will be to apply the basic knowledge we gain on plant population dynamics to "real-life problems" in applied ecology (Cain et al. 2000). Climate change, as an example, poses increasing pressure on many plant species that can only be mitigated by either migration or local adaptation (Jump & Penuelas 2005). Understanding how dispersal and gene flow processes, genetic structure, and the maintenance of genetic diversity influence the migration or adaptation capability of plant species, is essential to predict and understand species' responses to climate change. Ongoing habitat destruction and fragmentation threaten plant communities by changing population densities, loss of dispersal agents, or the disruption of plant-animal interactions (Dick 2001, Cordeiro & Howe 2003). Consequences of these environmental changes are hard to predict. Dick (2001), for example, showed that pollination by African honeybees maintains high genetic diversity in fragmented populations of the tropical tree *Dinizia excelsa*, despite the loss of native pollinators. In contrast to that, Dubreuil et al. (2010) found high levels of genetic structure and inbreeding in the temperate tree *Taxus baccata* in fragemented areas, despite the high dispersal potential of the species. Understanding the mechanisms behind these patterns is essential for the integration of our knowledge into conservation plans and management strategies. Thereby, assessing consequences across multiple life stages can be valuable, especially in long-lived tree species, as due to their longevity, a prompt evaluation of long-term consequences of changing ecological conditions is

difficult. Overall, analyzing processes across life stages is a promising approach to meet future challenges in understanding plant-animal interactions and their consequences for plant communities.

6 REFERENCES

Aldrich, P.R., J.L. Hamrick, P. Chavarriage, and G. Kochert. 1998. Microsatellite analysis of demographic genetic structure in fragmented populations of the tropical tree *Symphonia globulifera*. Molecular Ecology 7:933-944.

Alvarez-Buylla, E.R., A. Chaos, D. Pinero, and A.A. Garay. 1996. Demographic genetics of a pioneer tropical tree species: patch dynamics, seed dispersal, and seed banks. Evolution 50:1155-1166.

Asquith, N.M., S.J. Wright, and M.J. Clauss. 1997. Does mammal community composition control recruitment in neotropical forests? Evidence from Panama. Ecology 78:941-946.

Augspurger, C.K, and C.K. Kelly. 1984. Pathogen mortality of tropical tree seedlings – experimental studies of the effects of dispersal distance, seedlings density, and light conditions. Oecologia 61:211-217.

Augspurger, C.K., and H.T. Wilkinson. 2007. Host specificity of pathogenic *Pythium* species: implications for tree species diversity. Biotropica 39:702-708.

Barret, S.C.H. 2002. Sexual interference of the floral kind. Heredity 88:154-159.

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

Berg, E.E., and J.L. Hamrick. 1995. Fine-scale genetic structure of a Turkey oak forest. Evolution 49:110-120.

Bittencourt, J.V.M., and A.M. Sebbenn. 2007. Patterns of pollen and seed dispersal in a small, fragmented population of the wind-pollinated tree *Araucaria angustifolia* in southern Brazil. Heredity 99:580-591.

Bleher, B., D. Uster, and T. Bergsdorf. 2006. Assessment of threat status and management effectiveness in Kakamega Forest, Kenya. Biodiversity and Conservation 15:1159-1177.

Boucher, D.H., S. James, and K.H. Keeler. 1982. The ecology of mutualism. Annual Review of Ecology and Systematics 13:315-347.

Burczyk, J., S.P. DiFazio, and W.T. Adams. 2004. Gene flow in forest trees: how far do genes really travel? Forest Genetics 11:179-192.

Burczyk, J., W.T. Adams, D.S. Birkes, and I.J. Chybicki. 2006. Using genetic markers to directly estimate gene flow and reproductive success parameters in plants on the basis of naturally regenerated seedlings. Genetics 173:363-372.

Cain, M.L., B.G. Milligan, and A.E.Strand. 2000. Long-distance seed dispersal in plant populations. American Journal of Botany 87:1217-1227.

Carneiro, F.D., B. Degen, M. Kanashiro, A.E.B. de Lacerda, and A.M. Sebbenn. 2009. High levels of pollen dispersal detected through paternity analysis from a continuous *Symphonia globulifera* population in the Brazilian Amazon. Forest Ecology and Management 258:1260-1266.

Chesser, R.K. 1991. Influence of gene flow and breeding tactics on gene diversity within populations. Genetics 129:573-583.

Chung, M.G., M.Y. Chung, G.S. Oh, and B.K. Epperson. 2000. Spatial genetic structure in a *Neolitsea sericea* population (Lauraceae). Heredity 85:490-497.

Chung, M.Y., B.K. Epperson, and M.G. Chung. 2003. Genetic structure of age classes in *Camellia japonica* (Theaceae). Evolution 57:62-73.

Cipriani, G., G. Lot, W.G. Huang, M.T. Marrazzo, E. Peterlunger, and R. Testolin. 1999. AC/GT and AG/CT microsatellite repeats in peach [*Prunus persica* (L) Batsch]: isolation, characterisation and cross-species amplification in *Prunus*. Theoretical and Applied Genetics 99:65-72.

Clark, C.J., J.R. Poulsen, E.F.Connor, and V.T. Parker. 2004. Fruiting trees as dispersal foci in a semi-deciduous tropical forest. Oecologia 139:66-75.

Clobert, J., E. Danchin, A.A. Dhondt, and J.D. Nichols, editors. 2001. Dispersal. Oxford University Press, Oxford.

Connell, J.H. 1971. On the role of natural enemies in preventing competitive exclusion in some marine animals and in rain forest trees. Pages 298-310 in P.J. denBoer, and G.R. Gradwell, editors. Dynamics of Populations. Centre for Agricultural Publishing and Documentations, Wageningen.

Cordeiro, N.J., and H.F. Howe. 2003. Forest fragmentation severs mutualism between seed dispersers and an endemic African tree. Proceedings of the National Academy of Sciences 100:14052-14056.

Cousens, R., C. Dytham, and R. Law. 2008. Dispersal in Plants. A population perspective. Oxford University Press, Oxford.

Crow, J.F., and M. Kimura. 1990. Male and female fertility variation in wild radish, a hermaphrodite. The American Naturalist 136:87-107.

Dakin, E.E., and J.C. Avise. 2004. Microsatellite null alleles in parentage analysis. Heredity 93:504-509.

Davis, M.B., and R.G. Shaw. 2001. Range shifts and adaptive responses to quaternary climate change. Science 292:673-679.

Degen, B., H. Caron, E. Bandou, C. Dutech, P. Breyne, M. Van Montagu, L. Maggia, and A. Kremer. 2001. Small scale spatial genetic structure of six tropical tree species in French Guiana. Pages 75-92 in G. Müller-Starck, and R. Schubert, editors. Genetic Response of Forest Systems to Changing Environmental Conditions. Kluwer Academic Publishers, Dordrecht.

De Lucas A.I., S.C. González-Martínez, G.G. Vendramin, E. Hidalgo, and M. Heuertz. 2009. Spatial genetic structure in continuous and fragmented populations of *Pinus pinaster* Aiton. Molecular Ecology 18:4564-4576.

De Steven, D., and S.J. Wright. 2002. Consequences of variable reproduction for seedling recruitment in three neotropical tree species. Ecology 83:2315-2327.

Dick, C.W. 2001. Genetic rescue of remnant tropical trees by an alien pollinator. Proceedings of the Royal Society B 268:2391-2396.

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

Dirlewanger, E., P. Cosson, M. Tavaud, M.J. Aranzana, C. Poizat, A. Zanetto, P. Arus, and F. Laigret. 2002. Development of microsatellite markers in peach [*Prunus persica* (L.) Batsch] and their use in genetic diversity analysis in peach and sweet cherry (*Prunus avium* L.). Theoretical and Applied Genetics 105:127-138.

Dubreuil, M., M. Riba, S.C. González-Martínez, G.G. Vendramin, F. Sebastiani, and M. Mayol. 2010. Genetic effects of chronic habitat fragmentation revisited: strong genetic structure in a temperate tree, *Taxus baccata* (Taxaceae), with great dispersal capability. American Journal of Botany 97:303-310.

Dutech, C., J. Seiter, P. Petronelli, H.I. Joly, and P. Jarne. 2002. Evidence of low gene flow in a neotropical clustered tree species in two rainforest stands of French Guiana. Molecular Ecology 11:725-738.

Epperson, B.K. 1992. Spatial structure of genetic variation within populations of forest trees. New Forests 6:257-278.

Epperson, B.K. 1993. Recent advances in correlation analysis of spatial patterns of genetic variation. Evolutionary Biology 27:95-155.

Epperson, B.K. 1995. Spatial distributions of genotypes under isolation by distance. Genetics 140:1431-1440.

Farwig, N., K. Böhning-Gaese, and B. Bleher. 2006. Enhanced seed dispersal of *Prunus africana* in fragmented and disturbed forests? Oecologia 147:238-252.

Farwig, N., C. Braun, and K. Böhning-Gaese. 2008. Human disturbance reduces genetic diversity of an endangered tropical tree, *Prunus africana* (Rosaceae). Conservation Genetics 9:317-326.

Garcia, C., P. Jordano, and J.A. Godoy. 2007. Contemporary pollen and seed dispersal in a *Prunus mahaleb* population: patterns in distance and direction. Molecular Ecology 16:1947-1955.

Garcia-Cruz, J., J. Gonzalez-Astorga, V. Sosa, and O. Hernadez-Gonzalez. 2009. Genetic diversity in six *Govenia* (Orchidaceae) species with different pollinator attraction strategies. International Journal of Plant Sciences 170:894-905.

Geng, Q.F., C.L. Lian, S. Goto, J.M. Tao, M. Kimura, M.D.S. Islam, and T. Hogetsu. 2008. Mating system, pollen and propagule dispersal, and spatial genetic structure in a high-density population of the mangrove tree *Kandelia candel*. Molecular Ecology 17:4724-4739.

Gerber, S., S. Mariette, R. Streiff, C. Bodenes, and A. Kremer. 2000. Comparison of microsatellites and amplified length polymorphism markers for parentage analysis. Molecular Ecology 9:1037-1048.

Gerber, S., P. Chabrier, and A. Kremer. 2003. FAMOZ: a software for parentage analysis using dominant, codominant and uniparentally inherited markers. Molecular Ecology Notes 3:479-481.

Ghazoul, J. 2005. Pollen and seed dispersal among dispersed plants. Biological Reviews 80:413-443.

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

González-Martínez, S.C., S. Gerber, M.T. Cervera, J.M. Martinez-Zapater, L. Gil, and R. Alia. 2002. Seed gene flow and fine-scale structure in a Mediterranean pine (*Pinus pinaster* Ait.) using nuclear microsatellite markers. Theoretical and Applied Genetics 104:1290-1297.

González-Martínez, S.C., J. Burczyk, R. Nathan, N. Nanos, L. Gil, and R. Alia. 2006. Effective gene dispersal and female reproductive success in Mediterranean maritime pine (*Pinus pinaster* Aiton). Molecular Ecology 15:4577-4588.

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

Greenwood, M.S. 1986. Gene exchange in loblolly pine: the relation between pollination mechanisms, female receptivity and pollen availability. American Journal of Botany 73:1443-1451.

Grivet, D., P.E. Smouse, and V.L. Sork. 2005. A novel approach to an old problem: tracking dispersed seeds. Molecular Ecology 14:3585-3595.

Hall, J.B., E.M. O´Brien, and F.L. Sinclair. 2000. *Prunus africana* - a monograph. School of Agricultural and Forest Sciences, Publication Number 18, University of Wales, Bangor.

Hammond, D.S., and V.K. Brown. 1998. Disturbance, phenology and life-history characteristics: factors influencing distance/density-dependent attack on tropical seeds and seedlings. Pages 51-78 in D.M. Newbery, H.H.T. Prins, and N.D. Brown, editors. Dynamics of tropical communities. The 37th symposium of the British Ecological Society. Cambridge University. 1996. Blackwell Science Ltd, Oxford.

Hamrick, J.L., D.A. Murawski, and J.D. Nason. 1993. The influence of seed dispersal mechanisms on the genetic structure of tropical tree populations. Plant Ecology 107-108:281-297.

Hamrick, J.L., and J.D. Nason. 1996. Consequences of dispersal in plants. Pages 203-236 in O.E.J. Rhodes, R.K. Chesser, and M.H. Smith, editors. Population dynamics in ecological space and time. The University of Chicago Press, Chicago.

Hardesty, B.D., S.P. Hubbell, and E. Bermingham. 2006. Genetic evidence of frequent long-distance recruitment in a vertebrate-dispersed tree. Ecology Letters 9:516-525.

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

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

Harms, K.E., S.J. Wright, O. Calderon, A. Hernandez, and E.A. Herre. 2000. Pervasive density-dependent recruitment enhances seedling diversity in a tropical forest. Nature 404:493-495.

Heuertz, M., X. Vekemans, J.F. Hausman, M. Palada, and O.J. Hardy. 2003. Estimating seed vs. pollen dispersal from spatial genetic structure in the common ash. Molecular Ecology 12:2483-2495.

Holbrook, K.M., T.B. Smith, and B.D. Hardesty. 2002. Implications of long-distance movements of frugivorous rain forest hornbills. Ecography 25:745-749.

Howe, H.F., and J. Smallwood. 1982. Ecology of seed dispersal. Annual Review of Ecology, Evolution and Systematics 13:201-228.

Hyatt, L.A., M.S. Rosenberg, T.G. Howard, G. Bole, W. Fang, J. Anastasia, K. Brown, R. Grella, K. Hinman, J.P. Kurdziel, and J. Gurevitch. 2003. The distance dependence prediction of the Janzen-Connell hypothesis: a meta-analysis. Oikos 103:590-602.

Imbert, E., and F. Lefèvre. 2003. Dispersal and gene flow of *Populus nigra* (Salicaceae) along a dynamic river system. Journal of Ecology 91:447-456.

Isagi, Y., T. Kanazashi, W. Suzuki, H. Tanaka, and T. Abe. 2000. Microsatellite analysis of the regeneration process of *Magnolia obovata* Thunb. Heredity 84:143-151.

Isagi, Y., and T. Kanazashi. 2002. Gene flow analysis in *Magnolia obovata* Thunb. using highly variable microsatellite markers. Pages 257-269 in N. Matsumoto, editor. Diversity and interaction in a temperate forest community: Ogawa Forest Reserve of Japan. Springer-Verlag, Tokyo.

Isagi, Y., D. Saito, H. Kawaguchi, R. Tateno, and S. Watanabe. 2007. Effective pollen dispersal is enhanced by the genetic structure of an *Aesculus turbinata* population. Journal of Ecology 95:983-990.

Izhaki, I., P.B. Walton, and U.N. Safriel. 1991. Seed shadows generated by frugivorous birds in Eastern Mediterranean scrub. Journal of Ecology 79:575-590.

Jacquemyn, H., R. Brys, K. Vandepitte, O. Honnay, and I. Roldan-Ruiz. 2006. Fine-scale genetic structure of life history stages in the food-deceptive orchid *Orchis purpurea*. Molecular Ecology 15:2801-2808.

Jacquemyn, H., T. Wiegand, K. Vandepitte, R. Brys, I. Roldan-Ruiz, and O. Honnay. 2009. Multigenerational analysis of spatial structure in the terrestrial, food-deceptive orchid *Orchis mascula*. Journal of Ecology 97:206-216.

Janzen, D.H. 1970. Herbivores and number of tree species in tropical forests. American Naturalist 104:501-528.

Jones, A.G, and W.R. Ardren. 2003. Methods of parentage analysis in natural populations. Molecular Ecology 12:2511-2523.

Jones, F.A., and S.P. Hubbell. 2006. Demographic spatial genetic structure of the neotropical tree, *Jacaranda copaia*. Molecular Ecology 15:3205-3217.

Jones, F.A., J.L. Hamrick, C.J. Peterson, and E.R. Squiers. 2006. Inferring colonization history from analyses of spatial genetic structure within populations of *Pinus strobus* and *Quercus rubra*. Molecular Ecology 15:851-861.

Jordano, P., and C.M. Herrera. 1995. Shuffling the offspring - uncoupling and spatial discordance of multiple stages in vertebrate seed dispersal. Ecoscience 2:230-237.

Jordano, P. 2000. Fruits and frugivory. Pages 125-162 in M. Fenner, editor. Seeds: the ecology of regeneration in plant communities. CABInternational, Wallingford.

Jordano, P., and E.W. Schupp. 2000. Seed disperser effectiveness: the quantity component and patterns of seed rain for *Prunus mahaleb*. Ecological Monographs 70:591-615.

Jordano, P., and J.A. Godoy. 2002. Frugivore-generated seed shadows: a landscape view of demographic and genetic effects. Pages 305-321 in D.J. Levey, W.R. Silva, and M. Galetti, editors. Seed dispersal and frugivory: ecology, evolution and conservation. CABInternational, Wallingford.

Jordano, P., C. Garcia, J.A. Godoy, and J.L. Garcia-Castano. 2007. Differential contribution of frugivores to complex seed dispersal patterns. Proceedings of the National Academy of Sciences of the United States of America 104:3278-3282.

Jump, A.S., and J. Penuelas. 2005. Running to stand still: adaptation and the response of plants to rapid climate change. Ecology Letters 8:1010-1020.

Kalinowski, S.T., M.L. Taper, and T.C. Marshall. 2007. Revising how the computer program CERVUS accommodates genotyping error increases success in paternity assignment. Molecular Ecology 16:1099-1106.

Kalisz, S., J.D. Nason, F.M. Hanzawa, and S.J. Tonsor. 2001. Spatial population genetic structure in *Trillium grandiflorum*: The roles of dispersal, mating, history, and selection. Evolution 55:1560-1568.

Kelly, B.A., O.J. Hardy, and J.-M. Bouvet. 2004. Temporal and spatial genetic structure in *Vitellaria paradoxa* (shea tree) in an agroforestry system in southern Mali. Molecular Ecology 13: 1231-1240.

KIFCON. 1994. Kakamega guide. The official guide. Kenya Indigenous Forest Conservation Programme, Nairobi.

Kirika, J.M., N. Farwig, and K. Böhning-Gaese. 2008a. Effects of local disturbance of tropical forests on frugivores and seed removal of a small-seeded Afrotropical tree. Conservation Biology 22:318-328.

Kirika, J.M., B. Bleher, K. Böhning-Gaese, R. Chira, and N. Farwig. 2008b. Fragmentation and local disturbance of forests reduce frugivore diversity and fruit removal in *Ficus thonningii* trees. Basic and Applied Ecology 9:663-672.

Kokwaro, J.O. 1988. Conservation status of the Kakamega forest in Kenya: the easternmost relic of the equatorial rainforests of Africa. Monograph Systematics of Missouri Botanical Garden 25:471-489.

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

Kunin, W.E. 1993. Sex and the single mustard: population density and pollinator behavior effects on seed-set. Ecology 74:2145-2160.

Kwit, C., D.J. Levey, and C.H. Greenberg. 2004. Contagious seed dispersal beneath heterospecific fruiting trees and its consequences. Oikos 107:303-308.

Latouche-Hallé, C., A. Ramboer, E. Bandou, H. Caron, and A. Kremer. 2003. Nuclear and chloroplast genetic structure indicate fine-scale spatial dynamics in a neotropical tree population. Heredity 91:181-190.

Levin, D.A., and W.H. Kerster. 1969. Density-dependent gene dispersal in *Liatris*. The American Naturalist 103:61-74.

Levin, D.A., and W.H. Kerster. 1974. Gene flow in seed plants. Evolutionary Biology 7:139-220.

Linhart, Y.B., W.H. Busby, J.H. Beach, and P. Feinsinger. 1987. Forager behavior, pollen dispersal and inbreeding in two species of hummingbird-pollinated plants. Evolution 41:679-682.

Linhart, Y.B., and M.C. Grant. 1996. Evolutionary significance of local genetic differentiation in plants. Annual Review of Ecology and Systematics 27:237-277.

Loiselle, B.A., V.L. Sork, J. Nason, and C. Graham. 1995. Spatial genetic structure of a tropical understory shrub, *Psychotria officinalis* (Rubiaceae). American Journal of Botany 82:1420-1425.

Loveless, M.D., and J.L. Hamrick. 1984. Ecological determinants of genetic structure in plant-populations. Annual Review of Ecology and Systematics 15:65-95.

Lung, T. 2004. Landbedeckungsänderungen im Gebiet "Kakamega Forest und assoziierte Waldgebiete" (Westkenia) - multispektrale Klassifikation von Landsat-Satellitenbilddaten und Auswertung mittels Methoden im Raster-Gis. Karlsruher

Geowissenschaftliche Schriften - Reihe A 15. Karlsruhe University of Applied Sciences, Karlsruhe.

Mangan, S.A., S.A. Schnitzer, E.A. Herre, K.M.L. Mack, M.C. Valencia, E.I. Sanchez, and J.D. Bever. 2010. Negative plant-soil feedback predicts tree-species relative abundance in a tropical forest. Nature 466:752-756.

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

Meagher, T.R., and E. Thompson. 1987. Analysis of parentage for naturally established seedlings of *Chamaelirium luteum* (Liliaceae). Ecology 68:803-812.

Mitchell, N., G. Schaab, and J.W. Wägele. 2009. Kakamega Forest ecosystem: an introduction to the natural history and the human context. Karlsruher Geowissenschaftliche Schriften, Reihe A, Band 17. Karlsruhe University of Applied Sciences. Karlsruhe, Germany.

Nathan, R., and H.C. Muller-Landau. 2000. Spatial patterns of seed dispersal, their determinants and consequences for recruitment. Trends in Ecology & Evolution 15:278-285.

Nathan R., and R. Casagrandi. 2004. A simple mechanistic model of seed dispersal, predation and plant establishment: Janzen-Connell and beyond. Journal of Ecology 92:733–746.

Nathan, R. 2006. Long-distance dispersal of plants. Science 313: 786-788.

Nathan, R., F.M. Schurr, O. Spiegel, O. Steinitz, A. Trakhtenbrot, and A. Tsoar. 2008. Mechanisms of long-distance seed dispersal. Trends in Ecology & Evolution 23:638-647.

Ng, K.K.S., S.L. Lee, and C.L. Koh. 2004. Spatial structure and genetic diversity of two tropical tree species with contrasting breeding systems and different ploidy levels. Molecular Ecology 13:657-669.

Oddou-Muratorio, S., and E.K. Klein. 2008. Comparing direct vs. indirect estimates of gene flow within a population of a scattered tree species. Molecular Ecology 17:2743-2754.

Ouborg, N.J., Y. Piquot, and J.M. Van Groenendael. 1999. Population genetics, molecular markers and the study of dispersal in plants. Journal of Ecology 87:551-568.

Pacheco, L.F., and J.A. Simonetti. 2000. Genetic structure of a mimosoid tree deprived of its seed disperser, the spider monkey. Conservation Biology 14:1766-1775.

Packer, A., and K. Clay. 2000. Soil pathogens and spatial patterns of seedling mortality in a temperate tree. Nature 404:278-281.

Pakkad, G., C. James, F. Torre, S. Elliot, and D. Blakesley. 2003. Genetic variation of *Prunus cerasoides* D. Don, a framework tree species in northern Thailand. New Forest 27:189-200.

Parker, K.C., J.L. Hamrick, A.J. Parker, and J.D. Nason. 2001. Fine-scale genetic structure in *Pinus clausa* (Pinaceae) populations: Effects of disturbance history. Heredity 87:99-113.

Pellmyr, O. 2002. Pollination by animals. Pages 157-184 in C.M. Herrera, and O. Pellmyr, editors. Plant-animal interactions. An evolutionary approach. Blackwell Publishing, Malden.

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

Pulido, F.J., and M. Diaz. 2005. Regeneration of a Mediterranean oak: A whole-cycle approach. Ecoscience 12:92-102.

Rasmussen, I.R., and B. Brodsgaard. 1992. Gene flow inferred from seed dispersal and pollinator behavior compared to DNA analysis of restriction site variation in a patchy population of *Lotus corniculatus* L. Oecologia 89:277-283.

Robledo-Arnuncio, J.J., and L. Gil. 2005. Patterns of pollen dispersal in a small population of *Pinus sylvestris* L. revealed by total-exclusion paternity analysis. Heredity 94:13-22.

Robledo-Arnuncio, J.J., and C. Garcia. 2007. Estimation of the seed dispersal kernel from exact identification of source plants. Molecular Ecology 16:5098-5109.

Ross, K.G. 2001. How to measure dispersal: the genetic approach. The example of fire ants. Pages 29-42 in J. Clobert, E. Danchin, A.A. Dhondt, and J.D. Nichols, editors. Dispersal. Oxford University Press, Oxford.

Russo, S.E., and C.K. Augspurger. 2004. Aggregated seed dispersal by spider monkeys limits recruitment to clumped patterns in *Virola calophylla*. Ecology Letters 7:1058-1067.

Russo, S.E., S. Portnoy, and C.K. Augspurger. 2006. Incorporating animal behavior into seed dispersal models: Implications for seed shadows. Ecology 87:3160-3174.

Savolainen , O., T. Pyhäjärvi, and T. Knürr. 2007. Gene flow and local adaptation in trees. Annual Review of Ecology, Evolution and Systematics 38:595-619.

Schaal, B.A. 1980. Measurement of gene flow in *Lupinus texensis*. Nature 284:450-451.

Schnabel, A., J.D. Nason, and J.L. Hamrick. 1998. Understanding the population genetic structure of *Gleditsia triacanthos* L.: seed dispersal and variation in female reproductive success. Molecular Ecology 7:819-832.

Schueler, S., A. Tusch, M. Schuster, and B. Ziegenhagen. 2003. Characterization of microsatellites in wild and sweet cherry (*Prunus avium* L.) - markers for individual identification and reproductive processes. Genome 46:95-102.

Schupp, E.W. 1995. Seed seedling conflicts, habitat choice, and patterns of plant recruitment. American Journal of Botany 82:399-409.

Schupp, E.W., T. Milleron, and S.E. Russo, 2002. Dissemination limitation and the origin and maintenance of species-rich tropical forests. Pages 19-34 in D.J. Levey, W.R. Silva, and M. Galetti, editors. Seed dispersal and frugivory: ecology, evolution and conservation. CABInternational, Wallingford.

Sezen, U.U., R.L. Chazdon, and K.E. Holsinger. 2005. Genetic consequences of tropical second-growth forest regeneration. Science 307:891-891.

Sosinski, B., M. Gannavarapu, L.D. Hager, L.E. Beck, G.J. King, C.D. Ryder, S. Rajapakse, W.V. Baird, R.E. Ballard, and A.G. Abbott. 2000. Characterization of microsatellite markers in peach [*Prunus persica* (L.) Batsch]. Theoretical and Applied Genetics 101:421-428.

Spiegel, O., and R. Nathan. 2007. Incorporating dispersal distance into the disperser effectiveness framework: frugivorous birds provide complementary dispersal to plants in a patchy environment. Ecology Letters 10:718-728.

Stoeckel, S., J. Grange, J.F. Férnández-Majarres, I. Bilger, N. Frascaria-Lacoste, and S. Mariette. 2006. Heterzygote excess in a self-incompatible and partially clonal forest tree species – *Prunus avium* L. Molecular Ecology 15:2019-2118.

Trapnell, D.W., and J.L. Hamrick. 2005. Mating patterns and gene flow in the neotropical epiphytic orchid, *Laelia rubescens*. Molecular Ecology 14:75-84.

Tsingalia, M.H. 1990. Habitat disturbance, severity and patterns of abundance in Kakamega Forest, Western Kenya. African Journal of Ecology 28:213-226.

Ueno, S., N. Tomaru, H. Yoshimaru, T. Manabe, and S. Yamamoto. 2002. Size-class differences in genetic structure and individual distribution of *Camellia japonica* L. in a Japanese old-growth evergreen forest. Heredity 89:120-126.

Valbuena-Carabana, M., S.C. González-Martínez, V.L. Sork, C. Collada, A. Soto, P.G. Goicoechea, and L. Gil. 2005. Gene flow and hybridisation in a mixed oak forest (*Quercus pyrenaica* Willd. and *Quercus petraea* (Matts.) Liebl.) in central Spain. Heredity 95:457-465.

Vaughan, S.P., and K. Russell. 2004. Characterization of novel microsatellites and development of multiplex PCR for large-scale population studies in wild cherry, *Prunus avium*. Molecular Ecology Notes 4:429-431.

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

Voigt, F.A., R. Arafeh, N. Farwig, E.M. Griebeler, and K. Böhning-Gaese 2009. Linking seed dispersal and genetic structure of trees: a biogeographical approach. Journal of Biogeography 36:242-254.

Wang, H., M.Q. Qi, and A.J. Cutler. 1993. A simple method of preparing plant-samples for PCR. Nucleic Acids Research 21:4153-4154.

Wang, B.C., and T.B. Smith. 2002. Closing the seed dispersal loop. Trends in Ecology and Evolution 17:379-385.

Williamson, G.B., and K. Ickes. 2002. Mast fruiting and ENSO cycles – does the cue betray a cause? Oikos 97:459-461.

Wright, S. 1943. Isolation by distance. Genetics 28:114-138.

Yamagishi, H., H. Tomimatsu, and M. Ohara. 2007. Fine-scale spatial genetic structure within continuous and fragmented populations of *Trillium camschatcense*. Journal of Heredity 98:367-372.

Zhou, H.-P., and J. Chen. 2010. Spatial genetic structure in an understorey dioecious fig species: the roles of seed rain, seed and pollen-mediated gene flow, and local selection. Journal of Ecology 98:1168-1177.