Conserv Genet (2012) 13:1447–1457
DOI 10.1007/s10592-012-0386-2

RESEARCH ARTICLE

A rare case of natural regeneration in butternut, a threatened forest tree, is parent and space limited

Sean M. Hoban · Scott E. Schlarbaum · Sunshine L. Brosi · Jeanne Romero-Severson

Received: 9 February 2012 / Accepted: 6 July 2012 / Published online: 28 July 2012

Abstract The genetic consequences of natural in situ recovery for rare or threatened species are not as well understood as the impact of population bottlenecks, fragmentation and admixture, particularly the mechanisms by which genetic diversity is lost or preserved as populations recover. Here we examine how mating patterns, dispersal and ecologically constrained regeneration influences genetic diversity and kinship in a naturally regenerating population of a threatened temperate forest tree, *Juglans cinerea* L. (butternut). Butternut regeneration is now rare throughout the native range due to the butternut canker, a lethal fungal disease from Asia, and land use changes. In this study of one of the only known regenerating patches large enough for kinship and parentage analysis, we used 12 microsatellite markers, direct and inferred parentage analyses and Bayesian clustering of 152 trees to show that natural regeneration at this site resulted in loss of allele richness due to a small number of parents, most of which are spatially proximal to the regenerants. Of the 116 potential parents tested, one contributed 20.8 % and the top four contributed 71.1 % of the gametes in 36 regenerants. Parent-parent and parent-offspring distances revealed limited pollen and seed dispersal (<100 m). Regenerants were highly related and spatially clustered in sibling groups. Proximity to the regenerating patch was the most significant factor in parental success. Our results suggest that in situ regeneration of forest trees with limited propagule dispersal and specific site requirements may be insufficient to preserve native genetic diversity in protected areas with few suitable sites.

Keywords Restoration genetics · Population structure · Parentage analysis · Kinship · *Juglans cinerea*

Electronic supplementary material The online version of this article (doi:10.1007/s10592-012-0386-2) contains supplementary material, which is available to authorized users.

S. M. Hoban
Fourier Laboratoired'Ecologie Alpine Biologie D, Université Joseph, Bureau 305 2233 Rue de la Piscine, 38041 Grenoble, France

S. E. Schlarbaum
Department of Forestry, Wildlife and Fisheries, The University of Tennessee, 282 Ellington Plant Sciences Bldg, Knoxville, TN 37996, USA

S. L. Brosi
201 Compton Science Center, Department of Biology, Frostburg State University, 101 Braddock Road, Frostburg, MD 21532-2303, USA

J. Romero-Severson (✉)
Department of Biological Sciences, University of Notre Dame, 327 Galvin Life Sciences Building, Notre Dame, IN 46556, USA
e-mail: jromeros@nd.edu

Introduction

Many investigators have studied genetic and ecological consequences of census decline and fragmentation in forest trees (Sork et al. 2002; Oddou-Muratorio et al. 2004; Bacles et al. 2005; Sork and Smouse 2006; Pardini and Hamrick 2008) but fewer have focused on dynamics of in situ demographic recovery, the consequences of which remain difficult to predict (Born et al. 2008; Pardini and Hamrick 2008). An understanding of the mechanisms that influence genetic shifts during rapid in situ regeneration, could improve restoration and management outcomes (Frankham 1995; Reusch et al. 2005; Pertoldi et al. 2007; Liu et al. 2008) for species currently in decline from biotic and abiotic stresses, suboptimal management practices, and climate shifts.

In natural populations of wind-pollinated, temperate forest trees, dispersal kernels generally show leptokurtosis (Vekemans and Hardy 2004; Dick et al. 2008), with substantial fractions of pollen dispersal often exceeding 1 km (Sork and Smouse 2006; Craft and Ashley 2007; Pluess et al. 2009). As expected given these results, many studies on forest disturbance and fragmentation have detected low differentiation among fragmented populations (Hamrick 2004; Bacles et al. 2005; Victory et al. 2006; Kramer et al. 2008), high diversity in founding populations (Born et al. 2008), and/or high gene flow between fragments (Craft and Ashley 2007; Fernández-Manjarrés and Sork 2007). In some cases, the genetic consequences of contemporary demographic fluctuations are negligible (Hamrick 2004; Bacles et al. 2005; Kramer et al. 2008).

However, once a species becomes increasingly rare, the spatial, temporal, and biological circumstances of local regeneration may profoundly influence genetic diversity. These include the number and spatial distribution of parents, variance in reproductive success across individuals and seasons, pollen and seed dispersal functions, size and spatial distribution of suitable recruitment sites, length of juvenile period, and depth of generation overlap (Vekemans and Hardy 2004; Oddou-Muratorio et al. 2005; Sork and Smouse 2006). These factors may in turn be influenced by competition, predation, disease, and recurring abiotic stress (e.g. drought), producing a range of genetic consequences (Loveless and Hamrick 1984; Sork et al. 2002; Grivet et al. 2009; Hampe et al. 2010). However, studies of recovering populations are divided in their findings; some suggest that diversity is maintained (Thomas et al. 1999; Rajora and Pluhar 2003), while others reach opposite conclusions (Oddou-Muratorio et al. 2004; Fernández-Manjarrés and Sork 2007).

The descriptive statistics for differentiation, heterozygosity and allele richness commonly used in studies of population dynamics do not reveal the spatial distribution of alleles or mating patterns within populations, critical considerations for genetic diversity studies in long-lived forest trees (Jones et al. 2007; Sork et al. 2002). Parentage analysis, realized dispersal patterns, pair-wise kinship, Bayesian clustering and other individual-based approaches can add mechanistic insights. For example, Born et al. (2008) observed high diversity in a spatially expanding juvenile cohort of *Aucoumea klaineana*, and used parentage assignment to show that >75 % of potential parents contributed successful offspring. In contrast, Grivet et al. (2009) found high genetic differentiation between *Quercus lobata* seedling patches and adults, attributable to very few (1–3) seed parents. Jones et al. (2007) used SGS analyses across juvenile and mature cohorts in *Eucalyptus globulus* and spatial interpolation of ordination axes to reveal spatial shifts in genetic variation due to prevailing winds in seed dispersal. These studies showed that the degree of genetic shift in regenerants can be a function of the proportion of parents contributing offspring, the evenness of their contribution, and the importance of seed dispersal mechanisms.

Here we examine the genetic consequences of a rare in situ regeneration in a population of *Juglans cinerea* (butternut), a wind-pollinated canopy tree. The lethal disease butternut canker has decimated butternut throughout its native range (Eastern United States, Ontario, Quebec and New Brunswick) during the last eighty years, causing regional census declines up to 90 % and widespread local extinctions (Fleguel 1996). Seedlings are especially susceptible and typically die within the first season if infected (Fleguel 1996; Schlarbaum et al. 1997). Our study site in Tennessee is one of the few locations within the native range with substantial regeneration.

A null model of genetic and spatial dynamics for in situ regeneration requires that (1) contributing parents represent a random sample of the potential parent population, (2) the number of offspring per parent tree is randomly distributed and (3) seeds disperse at random (Born et al. 2008; Pardini and Hamrick 2008). Our main objective was to test this null model by using two complementary methods of parentage analysis (the Bayesian inference in Parentage and the maximum likelihood method in Cervus) to quantify the number of parents, the number of offspring per parent, and the distribution of parent-offspring and sibling-cluster distances. If regenerants arise from few parents (who may themselves be related), the regenerating population may show strong genetic differentiation, even over short distances (Loveless and Hamrick 1984; Jones et al. 2007; Pardini and Hamrick 2008). Our synthesis of descriptive statistics, kinship estimates, and genetic structure analyses enables us to test the validity of the null model for the genetic impact of in situ regeneration on genetic diversity in an early successional forest tree. Our objective is to contribute to understanding of genetic change during forest tree establishment, knowledge which may help predict responses to management practices and to future environmental change.

Methods

Species and site

Butternut, native to Eastern North American riparian forests, experienced a 20th century decline due to land use and forest management changes, loss of habitat, and a lethal fungal disease, butternut canker (Ostry et al. 1994). Contemporary populations are scattered and small (tens to hundreds of trees per population). Butternut is a fast growing (age of maturity $\sim$ 10 years, depending on conditions), light-demanding, comparatively short-lived (60–70 years), early successional

species typically occurring along waterways, or in upland forest gaps. The large fruits are dispersed by gravity, water and rodents (Fleguel 1996). Butternuts are monoecious and heterodichogamous, so selfing is not expected. No published studies address seed or pollen dispersal for butternut, but the shade-intolerant seedlings are unlikely to survive beneath a closed canopy (Rink 1990; Schultz 2003). The population (Fig. 1) is located in central Tennessee, USA, in a riparian forest valley previously used for small-scale agriculture. There is no record of tree planting, harvest or other management practices since abandonment 50–75 years ago. This site contains ~240 butternuts, the largest known census of butternut for an equivalent area in the entire native range (Fig. 1). Our collection strategy was to obtain a representative sample of the size classes at the site. We classified individuals having DBH (diameter 1.4 m above ground) between 1 and 7.6 cm as juveniles and the rest as adults. We also recorded tree height. In black walnut (*J. nigra*), a riparian zone species which grows to similar sizes, this size class is considered immature (Bruckerhoff 2005). Based on this, we are relatively certain that the individuals classified as juveniles are immature. However, as reproductive maturation (the age of first flowering and successful fruiting thereafter) depends highly on local conditions, some individuals classified as adults may also still be immature. As open sites in a riparian forest are quickly occupied and butternut establishment depends on availability of these sites, this cohort likely stems from recruitment over a small number of seasons. Our study focused on successful recruitment, i.e. established juveniles likely to survive long enough to reproduce. We also recorded number of cankers from the butternut canker disease on the main trunk.

The stand occupies one side of a V-shaped valley, growing along ~4 km of a stream at the bottom of the valley. Most juveniles occupy a patch ~20 m × 100 m (Fig. 1). Natural disturbance (tree fall, stream incision) tends to occur in patches, so patches of regenerants are expected for early successional species. All but two of the 116 adults and two of the 36 juveniles genotyped for this study have UTM reference points. We know of no other butternut stands in this valley, but wind-pollination from isolated trees or distant small populations is a possibility.

Genotyping

We genotyped all sampled individuals successfully for 12 microsatellite markers as previously described (Hoban et al. 2010), with an average success rate of 11.7 markers per individual. Using previously described methods (Hoban et al. 2009; McCleary et al. 2009), individuals were tested for hybridization with the introduced *J. ailantifolia* (Japanese walnut). We detected five hybrids and excluded them from further analysis. We have previously tested these markers for gametic equilibrium and found no consistent patterns of disequilibrium among loci (Hoban et al. 2008).

Genetic diversity and population structure

We use FSTAT (Goudet 1995) to calculate number of alleles (A), allelic richness (A_r) and F_{IS} for juveniles and adults separately. We expect a difference in allele richness immediately following a bottleneck, but heterozygosity (assuming random mating) may decline slowly if at all, unless the effective population size remains small over many generations. As each statistic (except for allelic range) showed some departure from normality (Shapiro-Wilks test, data not shown), we used Wilcoxon sign-rank tests to detect significant differences between juveniles and adults for these statistics.

We calculated F_{ST} between juveniles and adults to reveal whether the juveniles represented a random sample of the adult gene pool. We expect the F_{ST} values to be similar if this is true. We calculated F_{ST} between each of two clumps of adult trees (b and e, Fig. 1) and the rest of the adult population to provide a useful comparison for level of differentiation among groups of trees within the population.

We used SPAGEDI (Hardy and Vekemans 2002) to calculate Nason's kinship coefficient, F_{ij}, which performs well for microsatellite data (Vekemans and Hardy 2004) for pair-wise relationships among juveniles only, adults only, and adult-juvenile pairs only (as in Hampe et al. 2010), and over all individuals. The kinship coefficient complements our other results by enabling us to calculate mean relatedness among juveniles, among adults, and between adults and juveniles (Hampe et al. 2010).

We complemented F_{ST} with TESS, a Bayesian clustering technique for visualizing population genetic structure (Chen et al. 2007). TESS simultaneously infers allele frequencies and assigns individuals to clusters. TESS and other clustering methods cluster genetically differentiated groups on scales ranging from families to subpopulations to species, as well as populations across time (Anderson and Dunham 2008; Lepais et al. 2009; Pearse et al. 2009; Fonseca et al. 2010). This method also utilizes spatial information as prior, useful in situations in which shared ancestry varies across space, a reasonable expectation for regenerating patches. TESS employs a formal statistical method for determining the "best K" (number of clusters), the Deviance Information Criteria (DIC), in which gain in model fit as parameters are added is balanced by a penalty, similar to model selection in multiple regression. Settings for TESS: admixture model, 125,000 MCMC sweeps (25,000 sweeps for burn-in), and 20 reps for each K from 1

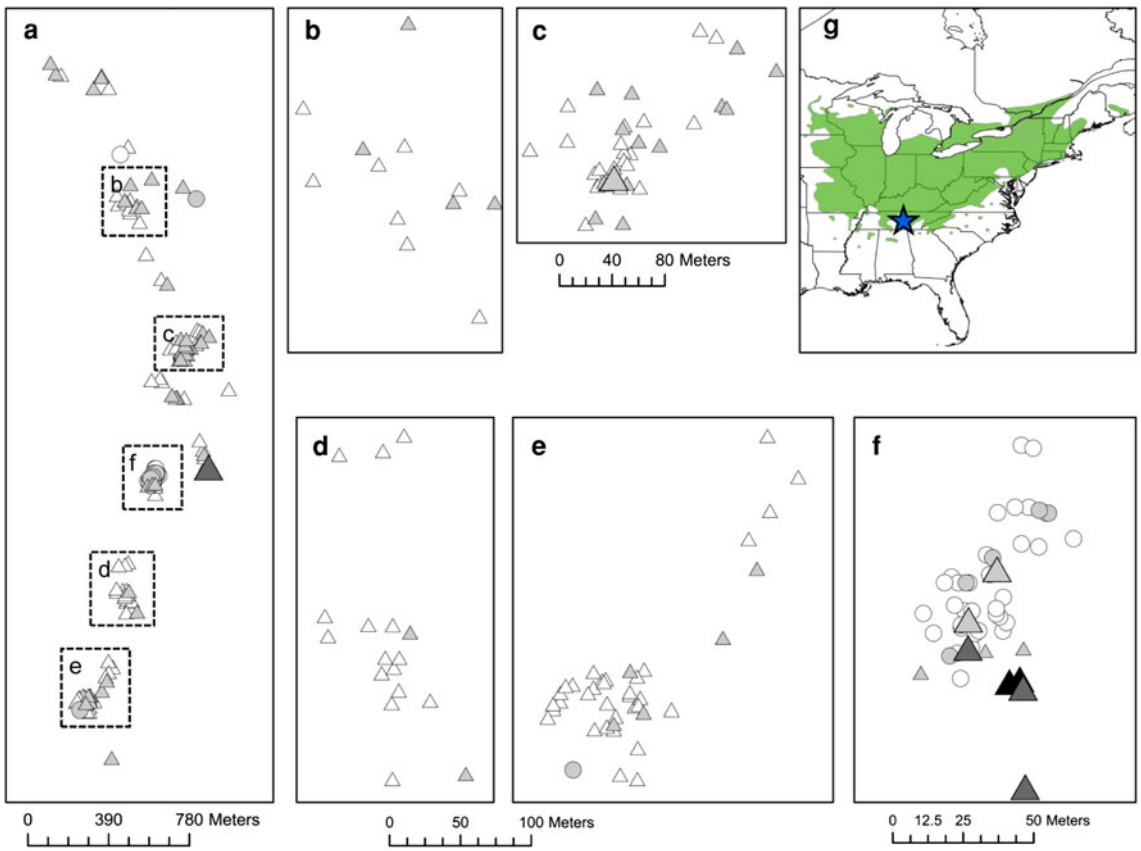

Fig. 1 Spatial distribution of adults (*open triangles*) and juveniles (*open circles*), unsampled individuals in *gray*. **a** all individuals, boxes **b–f** indicate areas shown in expanded view. **b–f** expanded views. **g** range map showing study site (*star*). *Large triangles* (all but one shown only in expanded view) indicate the number of single parent-offspring relationships assigned with 95 % confidence to the indicated individuals. *Light gray*- one assignment, *gray*- three to six, *black*- ten to 15

to 10. We also used STRUCTURE (Hubisz et al. 2009), using the LnP(D) and Delta K methods to evaluate K. Settings for STRUCTURE: admixture model, infer alpha (mixing parameter), allele frequencies correlated, assume different F_{ST} for each population pair, 500,000 MCMC sweeps (100,000 sweeps for burn-in), Q updated every 10 steps.

Parentage analysis

We used two approaches to analyze parentage of juveniles: (1) categorical parentage assignment using CERVUS (Kalinowski et al. 2007), in which genotypes of potential parents are evaluated against genotypes of progeny and the likelihood of parentage calculated and (2) Bayesian inference in PARENTAGE (Emery et al. 2001), which, given a set of offspring, determines the most likely number of mothers and fathers, and groups offspring into sibling clusters. Neither requires known parent-offspring relationships a priori. CERVUS assigns parentage to individual juveniles, permitting us to determine the number of offspring per parent and, using UTM coordinates, the geographic distance between parent-offspring pairs, between both parents, and among siblings. However, CERVUS provides no information on the number of parents that might have contributed to unassigned offspring, a disadvantage in the absence of exhaustive sampling. PARENTAGE provides additional information by using offspring genotypes to estimate the total number of contributing parents and number of offspring produced by each, without actually assigning particular relationships. PARENTAGE also reports a measure of shared parentage for each offspring pair, from which probable sibships can be reconstructed. Together the approaches give a more complete view than either approach yields independently.

The settings used for CERVUS were 10,000 simulated offspring and a genotype success rate of 97 %. As error rates in CERVUS may be influenced by the parameters of the simulations that guide its likelihood analysis (Herlin et al. 2007; Oddou-Muratorio et al. 2004), we tested a range of values for genotyping error (1, 5 and 10 %), proportion of parents sampled (55, 70, and 85 %) and level of inbreeding (0.0, 0.05 and 0.10), based on observations in range-wide populations (Hoban et al. 2010). Parentage analysis was performed using the allele frequencies of all adults,

because a priori we have no knowledge of where parents are located in the population. We report strict (0.95 probability) parentage assignments (Oddou-Muratorio and Klein 2008). For each set of input parameters, we observed <5 parentage assignments at the relaxed but not strict level, as parentage was usually assigned with very high or near zero likelihood. If, for a given juvenile, assignment could be made with strict probability to more than two parents, we assigned the parents with highest probability. This occurred <4 times for each parameter set. Using these assignments, we counted the number of parents contributing, and the number of offspring per parent.

We performed four sets of runs with PARENTAGE. First set: weak priors: uniform prior on number of fathers (1–50) and mothers (1–25), and a gamma distribution for mutations with parameters of (2, 0.0001), the span of microsatellite mutations observed in literature (Emery et al. 2001). Second set: prior weight on few parents: a normal distribution (mean = 2, standard deviation = 10). Third set: prior weight on many parents: a normal distribution (mean = 20, standard deviation = 10). Fourth set: population allele frequencies provided as prior. All analyses were done with 2×10^6 steps, 2×10^5 discarded as burn-in and samples taken every 400 steps, resulting in 4.5×10^3 samples from the posterior, similar to previous investigations (Emery et al. 2001).

Individual size, inter-individual distance, genetic incompatibilities, resource availability, and landscape features can cause differences in parental success. To determine whether distance and other factors were important in determining parental success, we performed Kolmogorov-Smirnov tests to compare observations to null distributions. (1) To examine whether success of adults is a function of distance to juveniles, we test whether the frequency distribution of distances between assigned parents and offspring is different from the frequency distribution of distances between all potential parents and offspring (the null model). We actually perform three tests here: one with all observed single parent assignments, one with inferred paternal assignments, and one with inferred maternal assignments. (2) To examine whether the success of parent pairs is explained well by distance between the parents in the pair, we test whether the frequency distribution of distances between assigned parents (realized parent pairs) is different from the null model of inter-parent distance, the frequency distribution of distances between all potential parent pairs. (3) To examine whether the success of adults is predicted well by three different measures (height, the number cankers on the main trunk and DBH), we test whether the frequency distribution of a given measure for each parent in an assigned parent-offspring relationship is different from the frequency distribution of that same measure for all potential parents.

Results

Genetic diversity and intra-population substructure

Size classes show a distinct bimodal distribution (Fig. 2) suggesting a recent (~10 years) burst of regeneration. Juveniles had significantly lower allele richness than adults (Table 1) but F_{is} and heterozygosity were not significantly different. Juvenile vs. adult F_{ST} (0.0451) was high, as was F_{ST} between clump b vs. the rest (0.0379) and clump e vs. the rest (0.0696). Results from STRUCTURE and TESS were congruent. DIC indicated an optimal K of 2, while Ln P(D) and Delta K indicate 6 and 5 respectively (Supplemental Fig. 1). As suggested by Pritchard, the weak support for K = 6 indicates hierarchical structure. In all cases, population structure clearly delineates adults and juveniles, but further structure emerges among adults (Supplemental Fig. 2). Juveniles show admixture and assignment probabilities consistent with the inference of parentage from those individuals identified as parents by CERVUS. Mean kinship was much higher among juveniles (0.097) than among-adults (0.034), between adults and juveniles (0.014) and among all individuals (0.031).

Parentage analysis

Nearly identical results were obtained using different sampling, mistyping, and inbreeding assumptions, even assuming only 55 % of potential parents were sampled (Table 2), indicating robust parentage assignment. Under the worst-case scenario (55 % of parents sampled and 10 % mistyping rates), we were able to assign with high confidence (mean LOD = 8.90) one or both parents to 32 (88.9 %) of the 36 offspring. Under relaxed confidence, 34 offspring (94.4 %) were assigned one or both parents. In the following description, we report relaxed confidence results in parentheses next to the strict confidence result. For 20 offspring (22), both parents were assigned. Among the 52 (56) successful single parent assignments, there were 10 (12) unique parents from the 116 potential parents genotyped. The remaining 106 contributed no gametes. We identified eight half-sib groups with size ranging from 15 to 3, and four full-sib groups, sized from 2 to 6. Most assigned individuals were members of at least a half sib cluster, and ~38 % (relaxed-47 %) were members of a full sib group. High variance was observed in parental success. One parent contributed 15 gametes, the top two parents contributed 25 gametes, and the top four parents contributed 37 gametes (20.8, 48.1, and 71.1 % of successful assignments).

Nearly all parent-offspring distances were <100 m, but five (seven) exceeded 500 m (Table 3, Fig. 3). For offspring with both parents assigned, if we assume that the shorter distance is the maternal parent–offspring distance,

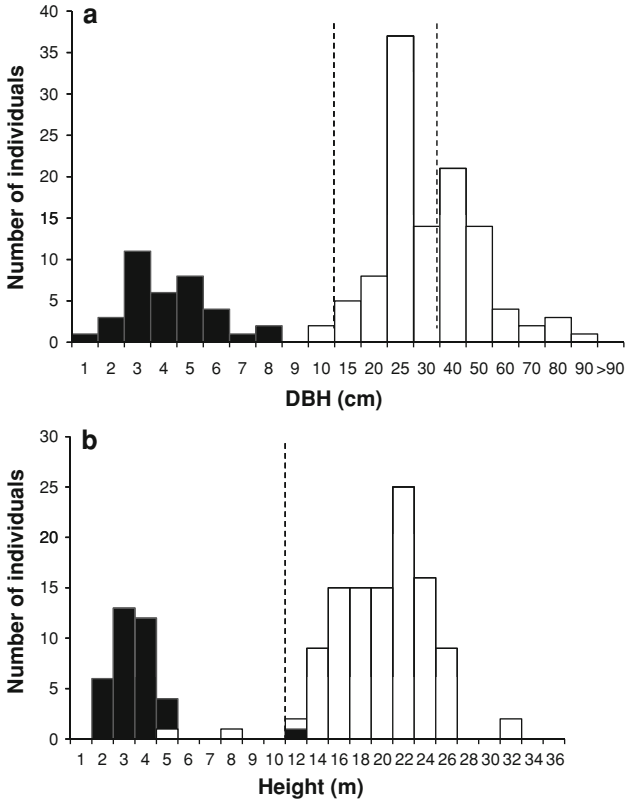

Fig. 2 DBH (**a**) and heights (**b**) for juveniles (*black*) and adults (*white*). *Dashed line* indicates change of scale on *x*-axis

Table 1 Descriptive statistics for juveniles and adults

Statistic	Juveniles (N = 36)		Adults (N = 116)		*p* value
	Mean[a]	SD	Mean	SD	
Observed heterozygosity (H_o)	0.737	0.159	0.782	0.124	0.1514
Wright's inbreeding (F_{is})	0.053	0.133	0.039	0.067	0.7334
Number of alleles (A)	9.17	3.24	14.7	5.19	0.0025
Allelic richness (Ar)	8.92	3.18	10.9	3.59	0.0015
Number of private alleles[b]	9		75		
Differentiation (F_{ST})	0.045				<0.05
Differentiation (R_{ST})	0.049				<0.05

[a] Means across loci

[b] Sum across loci

all seed dispersal distances were <100 m, with most under 40 m (mean = 25.8 m). Most pollen dispersal distances were <100 m (mean = 88.2 m), but four are >100 m, with one >500 m. While the observations are relatively few, leptokurtosis is apparent for both distributions.

Even within the small area of regeneration, offspring were markedly clustered. Mean distance between individuals assigned to the same half or full sib group was substantially smaller than the size of the patch, typically less than 20 m, with full sib clusters having the smaller distances (Table 3). We did not arbitrarily delimit an area in which to genotype regenerants but genotyped the trees we were able to sample, regardless of size class. Thus the marked difference in the area occupied by juveniles as opposed to adults is a characteristic of the site as we found it and not an artifact introduced by our sampling design.

Analysis with Parentage revealed that the most likely total number of parents contributing to the juveniles was 16–19 (Fig. 4a), regardless of priors (Supplemental Fig. 3), suggesting high information content in the data. Variance in parental success was high, with most parents contributing to one or two offspring, and a small proportion of parents contributing many (Fig. 4b). Four groups of 2, 3, 5, and 9 full sibs were supported with probability >0.999. Cervus and Parentage clearly indicate that that a small number of parents contribute most of the offspring.

The K–S test was significant for comparing the observed distribution of parent offspring distances to the null distribution of all potential parent-offspring distances, as well as for inferred paternal and maternal distributions (all *p* values <0.0001). The distribution of distances between two successful parents was also significantly different from the distribution of distances between all potential pair pairs ($p = 0.0001$). Height was not significant ($p = 0.6826$), DBH was significant ($p = 0.0026$), and number of cankers was not significant ($p = 0.9470$, Supplemental Fig. 1).

Discussion

The dynamics of demographic recovery will affect the long-term efficacy of protection, translocation, assisted regeneration, and habitat restoration efforts (Frankham 1995; Pertoldi et al. 2007). We have shown that extensive natural regeneration of a threatened forest tree at a single favorable site resulted in shifts in allele frequencies and a loss of diversity due to a small number of contributing parents, most of which are close to the patch of highly related and spatially clustered juvenile sibling groups. We note that greater age range in adult trees might potentially inflate diversity within their cohort, which should be kept in mind when interpreting the statistical comparison between adult and juvenile diversity. Our findings were robust to methodologies and demographic assumptions, including rates of inbreeding and relatedness in the population, assumed number of potential parents sampled and priors on parental contributions.

Table 2 CERVUS assignment for number of offspring per parent, given assumptions on proportion of parents genotyped (PPG), genotyping error rate (GER) and level of inbreeding

Category	Assumed values						
PPG	0.55	0.55	0.70	0.85	0.85	0.85	0.85
GER	1	10	1	1	1	5	1
Inbreeding	0.05	0.05	0.05	0.05	0.00	0.05	0.10
Parent/s	Number of offspring per parent						
1213[a]	0	1	0	0	0	1	0
1227	3	3	3	3	3	3	3
1229	4	4	4	4	4	4	4
1230	15	15	15	16	16	15	15
1231	10	10	10	10	10	10	10
1232	6	6	6	6	6	6	6
1303	1	1	1	1	1	1	1
1344	6	6	6	6	6	6	6
1367	1	0	1	0	0	1	1
1227/1230[b]	1	1	1	1	1	1	1
1227/1344	1	1	1	1	1	1	1
1229/1230	1	1	1	1	1	1	1
1229/1303	1	1	1	1	1	1	1
1230/1232	3	1	4	4	4	4	1
1230/1231	6	3	6	6	6	6	3
1231/1367	1	6	1	1	1	1	6
1231/1367	0	1	0	0	0	0	1

[a] One parent identified

[b] Both parents identified

Table 3 Number of siblings and distance between pairs of siblings in sibships with >3 sibs, as constructed by CERVUS and PARENTAGE

	Sibships based on actual parentage assignments from CERVUS							Sibships inferred from PARENTAGE			
	Half sib families[a]					Full sib families[a]		Full sib families[a]			
Descriptors	CH1	CH2	CH3	CH4	CH5	CF1	CF2	PF1	PF2	PF3	PF4
N	16	4	10	6	6	4	6	4	9	3	2
D Mean[b]	16.6	20.6	12.9	17.0	17.8	15.4	12.5	15.4	12.0	9.1	9.2
D Max[b]	44.6	34.5	25.7	26.4	35.1	20.6	24.2	20.6	24.2	13.2	9.2
D Min[b]	2.0	1.0	1.4	6.7	5.4	7.6	2.0	7.6	1.4	1.0	9.2
D sd[b]	7.0	13.5	6.7	6.4	8.7	4.5	6.7	4.5	6.5	7.0	9.2

[a] Relationships among sib groups are complex as one parent can be male or female across multiple groups, thus the number of genetically distinct groups is not a simple sum

[b] Mean, maximum, minimum and standard deviation of distance between pairs of individuals in that particular sibship

The loss of diversity and the clear population structure between juvenile and adult cohorts are consistent with some studies of restored or regenerating populations (Oddou-Muratorio et al. 2004), but not others (Thomas et al. 1999; Travis et al. 2002; Rajora and Pluhar 2003). However, these studies did not examine parent-offspring distances, parental success, and sibling clusters. Consistent with a study that did examine these factors (Hampe et al. 2010), we show that the parents of most offspring were close to the patch, likely due to limited seed dispersal, and close to each other. While distance is a major factor, it is plausible that among the closest trees, those that are largest are likely to produce more seed and shed more pollen, but distant trees seem to have a poor chance no matter their size. Further, small distances among siblings suggest co-dispersal, perhaps due to squirrels caching multiple seeds

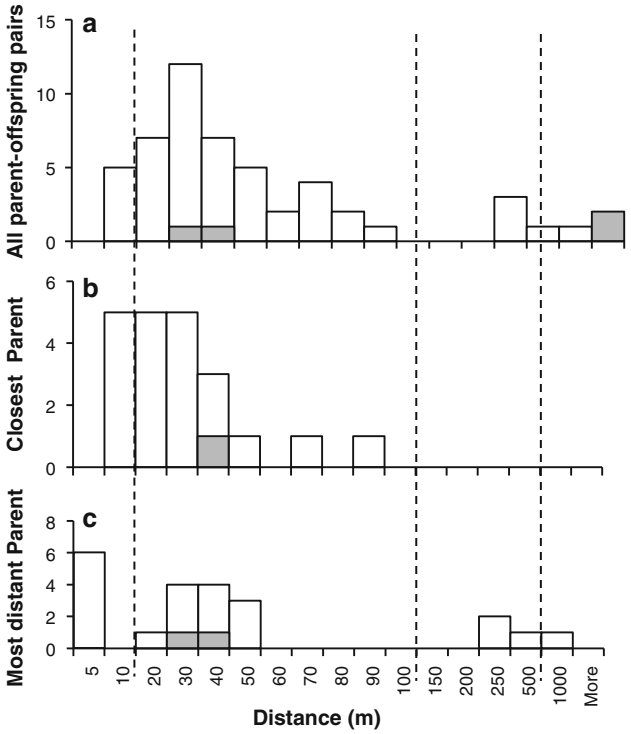

Fig. 3 Distribution of parent-offspring distances. *Dashed lines* indicate change of scale on *x*-axis, *x*-axis common to **a**, **b**, and **c**. **a** All parent-offspring pair assignments from CERVUS. For offspring with two parent assignments, closest (**b**) and most distant (**c**) parent-offspring pairs. *Shaded bars* are assignments made with 'relaxed' confidence

from the same tree. Grivet et al. (2009) also examined patch regeneration in oak, and found very few maternal parents (<2), more paternal parents (15), and strong sibling groupings in patches of similar numbers of offspring. From this we can infer that in situ regeneration, when ecologically or spatially constrained, is likely to result in loss of allele richness and heterozygosity. Key factors in these and related studies (Born et al. 2008) include space limitation, passive seed transport, and rodent caching.

If our study population is representative of butternut range-wide, co-dispersed siblings from few parents into favorable microsites may be a natural feature of butternut, contributing to differentiation over small scales and across generations. Jones et al. (2007) in an investigation of *Eucalyptus globulus*, another species subject to recurrent regeneration, refers to these as "spatially coherent cohorts." It is further possible, although certainly not proven, that a shifting mosaic of spatially and genetically coherent cohorts may be a natural phenomenon in butternut, *Eucalyptus*, and other early successional sub-dominant species. In these cases, which we call "windfall" situations, the opportunity for seed establishment may be rare but when the opportunity does arise, many related seeds may succeed simultaneously. From an evolutionary point of view, windfall dynamics may cause wide fluctuations in population allele frequencies, which could drive neutral, mildly deleterious, or potentially favorable alleles to high frequencies in the absence of strong selective pressures, an intriguing possibility.

The question of how contemporary in situ regeneration in this patch may differ from past intergenerational change at this and other sites is important. Regeneration opportunities are likely to be more restricted than they were in the past, due in part to combined threats of disease and environmental change. This site represents a rare combination of site favorability, the presence of seed-producing parents, land-use abandonment and subsequent protection, and the relatively low level of butternut canker. Under these circumstances, a relatively large cohort of progeny from a few parents may establish over a short time in the available space, a burst of regeneration that increases the census but decreases the effective population size if continued environmental pressures or restricted space provide adult trees with few or no other opportunities to produce successful descendants. While our site is only one case, the loss of sites favorable for regeneration and the shrinking size of those that remain is an increasingly common occurrence with rare and threatened species.

The results that we have shown in a small regenerating patch are similar to work demonstrating a loss of diversity in colonizing populations, where successful offspring in a small site descend from few parents (Pardini and Hamrick 2008). However, some studies have shown that recolonizing or expanding populations of wind-pollinated trees maintain the genetic diversity of source populations (Born et al. 2008). An explanation for some conflicting results in this field may be the spatial scale and the dispersal agents involved: colonization of distant sites, in situ bottlenecks, regeneration, and expansion into adjacent locations each may involve different demographic processes. Clearly, the location and number of adults, the temporal and spatial availability of suitable sites and species-specific dispersal kernels must be considered for predictive models of regeneration, in many tree species. Other critical factors during regeneration may include generation time, extent of overlap among generations, and period of occupancy of favorable sites. The relevance of each factor could be explored in future computer simulations (Hoban et al. 2012).

Considerations for management

Regenerating populations, while a positive sign for conservation efforts, may not represent past genetic composition. Our results suggest that in species with limited dispersal, a spatially or ecologically restricted naturally regenerating patch may possess significantly altered genetic composition. While this may be intuitive, the

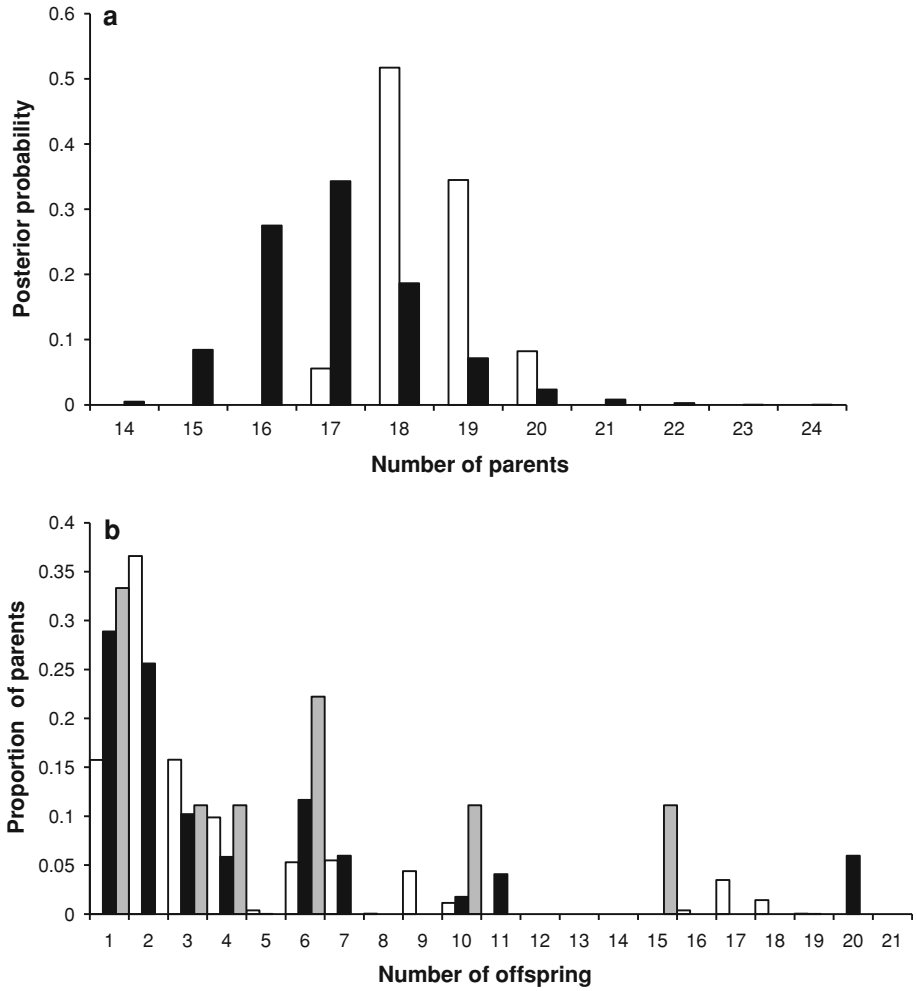

Fig. 4 Inferences from parentage. **a** Distribution of posterior probabilities for the number of parents contributing to juveniles, using uniform (*black*) and normal priors (*white*). For reference, CERVUS counted 10 (12, relaxed confidence) parents based on parentage assignments. **b** The proportion of reconstructed (single, not parent pairs) parents that produced a given number of offspring using uniform (*black*) and normal priors (*white*). *Gray*, observed proportion of parents producing a given number of offspring based on CERVUS assignments

number of parents and the variance in reproductive success have proven rather difficult parameters to estimate empirically in forest trees (Smouse and Sork 2004; Grivet et al. 2005, 2009; Sork and Smouse 2006).

Small regenerating patches in which kinship is high could ultimately suffer from inbreeding, as well as loss of diversity. It is possible that a shifting mosaic of regeneration sites over time and space could ameliorate these effects, providing different cohorts of parents the opportunity to contribute to regeneration. However, suitable regeneration sites may be increasingly rare for some early successional species. For example, in even aged stands, a common management practice, gaps may rarely occur from tree fall during the entire reproductive life of most trees. Other management practices or human impacts (e.g., dams, agricultural diversion, urban development) could also alter natural disturbance regimes and reduce the number of suitable sites for regeneration.

Natural regeneration is now almost unknown in this species. Butternut seedlings require water and sunlight, a constraint that limits this species to riparian zones subject to disturbance and incidental gaps in upland forests. The abandonment of formerly cleared land along a narrow riparian corridor at our study site provided butternut with a windfall, an opportunity for a sudden burst of regeneration over a relatively large area. In newly protected sites, such opportunities for a regenerative burst may be more frequent than in more "natural" settings but the long-term viability of this in situ regeneration is doubtful unless humans insure that the genetic base of the species is conserved. This may require deliberate introductions of seed from other areas in similar ecoregions and gap creation, approaches proven effective in other tree species (Ledig 1988; Frankham 1995). Until more regenerating sites are found, we cannot conclude that small size of the regenerating patch at our study site is typical for butternut. However, deliberately created large gaps or a series of small, spatially distributed gaps within propagule dispersal ranges could facilitate reproductive success by a greater number of adults. Empirical studies of regenerating sites in similar species and mathematical modeling should enable quantitative estimation of size of gaps or number of parents needed to

contribute representative diversity under particular situations. In the long term, forest management practices that facilitate naturally occurring gaps (rather than labor intensive felling) seem desirable, for this and other threatened early successional species.

No parent-offspring distances were less than five meters, in spite of large, heavy fruits, a result in agreement with shade intolerance in this species, a useful consideration for managed plantings of butternut. However, many juveniles were separated by less than 5 m, suggesting that intraspecific competition is weak enough to allow trees from the same cohort to grow together at this spacing for a time. Performing parentage analyses on naturally colonized and restored populations, (Travis et al. 2002; Liu et al. 2008; Fant et al. 2008; Cloutier et al. 2007) promises to reveal species-specific and well as more general responses to forest management and restoration projects. While our study is based on only one population out of necessity, we anticipate that comparative studies on multiple species across years and locations could reveal more about regeneration dynamics in early to mid-successional trees.

Conclusions

The ecological and economic services of forest biomes (carbon sequestration, habitat, sustainable timber supply) depend on successful long-term restoration and sustainable management practices. While continued study of declining and stable populations is needed to better understand human impact and set a baseline for future restoration, an understanding of consequences of restoration and conservation activities is a research priority, for guiding adaptive management and predicting consequences of future human activities. We demonstrated the utility of combining several recently developed techniques that offer fine resolution and a temporal-spatial perspective, for such investigations. The information we garnered illustrates the usefulness of diverse analytical approaches to quantify parentage, mating patterns and diversity in threatened species. Our findings indicate that multiple aspects of population biology, including location of potential parents, location and size of recruitment sites, and seed dispersal, are crucial considerations for ensuring a broad genetic base for adaptation. Accumulated data from additional studies will facilitate prediction and assessment of the consequences of alternative conservation strategies or environmental change over multiple generations, a critical advance given the threatened or endangered status of many long-lived taxa.

Acknowledgments We acknowledge Tim McCleary for technical assistance and Eagle Brosi, Ashton Berdine, Ami Sharp and Jennifer Murrow for field collection. We thank the US Forest Service, Natural Resources Canada, Barbara Boysen, Sandra Anagnostakis, Robert Anderson, Dan Borkowski, and Paul Berrang for contributions to the butternut project. Funding organizations include University of Notre Dame, University of Tennessee, the Arthur J Schmitt Foundation, and the Northern Nut Growers Association. This manuscript was completed at the Université Joseph Fourier (Grenoble, France) while SH was a postdoctoral associate funded by the Centre National de la Recherche Scientifique.

References

Anderson EC, Dunham KK (2008) The influence of family groups on inferences made with the program structure. Mol Ecol Resour 8:1219–1229

Bacles CFE, Burczyk J, Lowe AJ, Ennos RA (2005) Historical and contemporary mating patterns in remnant populations of the forest tree *Fraxinus excelsior* L. Evolution 59:979–990

Born C, Kjellberg F, Chevallier M-H, Vignes H, Dikangadissi J-T, Sanguié J, Wickings EJ, Hossaert-McKey M (2008) Colonization processes and the maintenance of genetic diversity: insights from a pioneer rainforest tree, *Aucoumea klaineana*. Proc R Soc B Biol Sci 275:2171–2179

Bruckerhoff DN (2005) Improving black walnut stands. Kansas State University Agricultural Experiment Station Publication L-718

Chen C, Durand E, Forbes F, François O (2007) Bayesian clustering algorithms ascertaining spatial population structure: a new computer program and a comparison study. Mol Ecol Notes 7: 747–756

Cloutier D, Kanashiro M, Ciampi AY, Schoen DJ (2007) Impact of selective logging on inbreeding and gene dispersal in an Amazonian tree population of *Carapa guianensis* Aubl. Mol Ecol 16:797–809

Craft KJ, Ashley MV (2007) Landscape genetic structure of bur oak (*Quercus macrocarpa*) savannas in Illinois. For Ecol Manage 239:13–20

Dick C, Hardy O, Jones F, Petit R (2008) Spatial scales of pollen and seed-mediated gene flow in tropical rain forest trees. Trop Plant Biol 1:20–33

Emery AM, Wilson IJ, Craig S, Boyle PR, Noble LR (2001) Assignment of paternity groups without access to parental genotypes: multiple mating and developmental plasticity in squid. Mol Ecol 10:1265–1278

Fant JB, Holmstrom RM, Sirkin E, Etterson JR, Masi S (2008) Genetic structure of threatened native populations and propagules used for restoration in a clonal species, American beachgrass (*Ammophila breviligulata* Fern.). Restor Ecol 16: 594–603

Fernández-Manjarrés JF, Sork VL (2007) Genetic variation in fragmented rorest stands of the Andean Oak *Quercus humboldtii* Bonpl. (Fagaceae). Biotropica 39:72–78

Fleguel RV (1996) A literature review of butternut and the butternut canker In: Eastern Ontario Model Forest, Information Report No 20. Eastern Ontario Model Forest, Ottowa, Ontario

Fonseca DM, Widdel AK, Hutchinson M, Spichiger SE, Kramer LD (2010) Fine-scale spatial and temporal population genetics of *Aedes japonicus*, a new US mosquito, reveal multiple introductions. Mol Ecol 19:1559–1572

Frankham R (1995) Conservation genetics. Annu Rev Genet 29:305

Goudet J (1995) FSTAT (Version 1.2): a computer program to calculate F-statistics. J Hered 86:485–486

Grivet D, Smouse PE, Sork VL (2005) A novel approach to an old problem: tracking dispersed seeds. Mol Ecol 14:3585–3595

Grivet D, Robledo-Arnuncio JJ, Smouse PE, Sork VL (2009) Relative contribution of contemporary pollen and seed dispersal to the effective parental size of seedling population of California valley oak (*Quercus lobata*, Née). Mol Ecol 18:3967–3979

Hampe A, El Masri L, Petit RJ (2010) Origin of spatial genetic structure in an expanding oak population. Mol Ecol 19:459–471

Hamrick JL (2004) Response of forest trees to global environmental changes. For Ecol Manage 197:323–335

Hardy OJ, Vekemans X (2002) SPAGeDi: a versatile computer program to analyse spatial genetic structure at the individual or population levels. Mol Ecol Notes 2:618–620

Herlin M, Taggart JB, McAndrew BJ, Penman DJ (2007) Parentage allocation in a complex situation: a large commercial Atlantic cod (*Gadus morhua*) mass spawning tank. Aquaculture 272:S195–S203

Hoban SM, Anderson R, McCleary TS, Schlarbaum SE, Romero-Severson J (2008) Thirteen nuclear microsatellite loci for butternut (*Juglans cinerea* L.). Mol Ecol Resour 8:643–646

Hoban SM, McCleary TS, Schlarbaum SE, Romero-Severson J (2009) Geographically extensive hybridization between the forest trees American butternut and Japanese walnut. Biol Lett 5:324–327

Hoban SM, Borkowski DS, Brosi SL, McCleary TS, Thompson LM, McLachlan JS, Pereira MA, Schlarbaum SE, Romero-Severson J (2010) Range-wide distribution of genetic diversity in the North American tree *Juglans cinerea*: a product of range shifts, not ecological marginality or recent population decline. Mol Ecol 19:4876–4891

Hoban S, Bertorelle G, Gaggiotti OE (2012) Computer simulations: tools for population and evolutionary genetics. Nat Rev Genet 73:2–14

Hubisz MJ, Falush D, Stephens M, Pritchard JK (2009) Inferring weak population structure with the assistance of sample group information. Mol Ecol Resour 9:1322–1332

Jones TH, Vaillancourt RE, Potts BM (2007) Detection and visualization of spatial genetic structure in continuous *Eucalyptus globulus* forest. Mol Ecol 16:697–707

Kalinowski ST, Taper ML, Marshall TC (2007) Revising how the computer program CERVUS accommodates genotyping error increases success in paternity assignment. Mol Ecol 16:1099–1106

Kramer AT, Ison JL, Ashley MV, Howe HF (2008) The paradox of forest fragmentation genetics. Conserv Biol 22:878–885

Ledig FT (1988) The conservation of diversity in forest trees: why and how should genes be conserved? Bioscience 38:471–479

Lepais O, Petit RJ, Guichoux E, Lavabre JE, Alberto F, Kremer A, Gerber S (2009) Species relative abundance and direction of introgression in oaks. Mol Ecol 18:2228–2242

Liu M-H, Chen X-Y, Zhang X, Shen D-W (2008) A population genetic evaluation of ecological restoration with the case study on *Cyclobalanopsis myrsinaefolia* (Fagaceae). Plant Ecol 197:31–41

Loveless MD, Hamrick JL (1984) Ecological determinants of genetic structure in plant populations. Annu Rev Ecol Syst 15:65–95

McCleary TS, Robichaud RL, Nuanes S, Anagnostakis SL, Schlarbaum SE, Romero-Severson J (2009) Four cleaved amplified polymorphic sequence (CAPS) markers for the detection of the *Juglans ailantifolia* chloroplast in putatively native *J. cinerea* populations. Mol Ecol Resour 9:525–527

Oddou-Muratorio S, Klein EK (2008) Comparing direct vs. indirect estimates of gene flow within a population of a scattered tree species. Mol Ecol 17:2743–2754

Oddou-Muratorio S, Demesure-Musch B, Pelissier R, Gouyon P-H (2004) Impacts of gene flow and logging history on the local genetic structure of a scattered tree species, *Sorbus torminalis* L. Crantz. Mol Ecol 13:3689–3702

Oddou-Muratorio S, Klein EK, Austerlitz F (2005) Pollen flow in the wild service tree, *Sorbus torminalis* (L.) Crantz. II. Pollen dispersal and heterogeneity in mating success inferred from parent–offspring analysis. Mol Ecol 14:4441–4452

Ostry ME, Mielke ME, Skilling DD (1994) Butternut- Strategies for managing a threatened tree. General technical report NC-165. U.S. Department of Agriculture, Forest Service, North Central Forest Experiment Station, St. Paul, MN

Pardini EA, Hamrick JL (2008) Inferring recruitment history from spatial genetic structure within populations of the colonizing tree *Albizia julibrissin* (Fabaceae). Mol Ecol 17:2865–2879

Pearse DE, Hayes SA, Bond MH, Hanson CV, Anderson EC, Macfarlane RB, Garza JC (2009) Over the falls? Rapid evolution of ecotypic differentiation in steelhead/rainbow trout (*Oncorhynchus mykiss*). J Hered 100:515–525

Pertoldi C, Bijlsma R, Loeschcke V (2007) Conservation genetics in a globally changing environment: present problems, paradoxes and future challenges. Biodivers Conserv 16:4147–4163

Pluess AR, Sork VL, Dolan B, Davis FW, Grivet D, Merg K, Papp J, Smouse PE (2009) Short distance pollen movement in a wind-pollinated tree, *Quercus lobata* (Fagaceae). For Ecol Manage 258:735–744

Rajora OP, Pluhar SA (2003) Genetic diversity impacts of forest fires, forest harvesting, and alternative reforestation practices in black spruce (*Picea mariana*). Theor Appl Genet 106:1203–1212

Reusch TBH, Ehlers A, Hämmerli A, Worm B (2005) Ecosystem recovery after climatic extremes enhanced by genotypic diversity. Proc Nat Acad Sci USA 102:2826–2831

Rink G (1990) *Juglans cinerea* L. Butternut. In: Burns RM, Honkala BH (ed) Agriculture handbook 654. U.S. Department of Agriculture, Forest Service, Washington, DC, pp 386–390

Schlarbaum SE, Hebard F, Spaine PC, Kamalay JC (1997) Three American tragedies: chestnut blight, butternut canker, and Dutch elm disease. In: Britton KO (ed) Proceedings, exotic pests of eastern forests. Tennessee Exotic Pest Plant Council, Nashville, TN, pp 45–54

Schultz J (2003) Conservation assessment for butternut or White walnut, Region FSE edn. United States Department of Agriculture, Forest Service, Milwaukee, WI

Smouse PE, Sork VL (2004) Measuring pollen flow in forest trees: an exposition of alternative approaches. For Ecol Manage 197:21–38

Sork V, Smouse P (2006) Genetic analysis of landscape connectivity in tree populations. Landscape Ecol 21:821–836

Sork VL, Davis FW, Smouse PE, Apsit VJ, Dyer RJ, Fernandez-M JF, Kuhn B (2002) Pollen movement in declining populations of California Valley oak, *Quercus lobata*: where have all the fathers gone? Mol Ecol 11:1657–1668

Thomas BR, Macdonald SE, Hicks M, Adams DL, Hodgetts RB (1999) Effects of reforestation methods on genetic diversity of lodgepole pine: an assessment using microsatellite and randomly amplified polymorphic DNA markers. Theor Appl Genet 98: 793–801

Travis SE, Proffitt CE, Lowenfeld RC, Mitchell TW (2002) A comparative assessment of genetic diversity among differently-aged populations of *Spartina alterniflora* on restored versus natural wetlands. Restor Ecol 10:37–42

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

Victory ER, Glaubitz JC, Rhodes OE Jr, Woeste KE (2006) Genetic homogeneity in *Juglans nigra* (Juglandaceae) at nuclear microsatellites. Am J Bot 93:118–126