USING GENETIC MARKERS TO DIRECTLY ESTIMATE MALE SELECTION GRADIENTS

Martin T. Morgan[1] and Jeffrey K. Conner[2]

[1]*School of Biological Sciences, Washington State University, Pullman, Washington 99164-4236 E-mail: mmorgan@wsu.edu*
[2]*Kellogg Biological Station, Department of Botany and Plant Pathology, Michigan State University, 3700 East Gull Lake Drive, Hickory Corners, Michigan 49060 E-mail: Conner@kbs.msu.edu*

Abstract.—We present an analysis of *Raphanus raphanistrum* and simulations illustrating the utility of directly estimating male phenotypic selection gradients using genetic markers. The method offers a much more refined characterization of selection than attempting to assign paternity to individual progeny. Our analysis of *R. raphanistrum* reveals selection on remarkably fine features of floral morphology, including anther exsertion, that were opaque to previous approaches. The new results also undermine a previous conclusion that selection on wild radish floral morphology acts primarily through female fitness. Simulation results show that selection gradients on the order of $\beta = 0.1$–0.2 can be readily detected with allozyme markers in moderate-sized (< 200 paternal individuals) populations. Highly polymorphic (e.g., microsatellite) markers will likely detect fine scale selection ($\beta < 0.1$) in larger populations (≥ 400 individuals). Increased progeny sample size, by sampling either additional maternal families or more progeny per maternal parent, partly compensates for low exclusion probability. Increasing the number of possible fathers without changing progeny sample size decreases the ability to detect selection, especially at lower exclusion probabilities. Sampling only some male genotypes reduces the power to detect selection and biases (underestimates) the magnitude of the selection gradient estimate.

Key words.—Male fertility, maximum likelihood, natural selection gradient, paternity analysis, *Raphanus raphanistrum*.

Male fertility represents exactly half the genetic contribution hermaphrodites make to future generations (Charnov 1982). Measuring male fertility is difficult, especially in plants where ecological measures are ''disappointingly crude'' (Snow and Lewis 1993, p. 332). Genetic marker data allow estimation of male fertility, but until recently methods have been limited in their ability to successfully partition paternity (Roeder et al. 1989; Morgan 1998). The difficulties of measuring male reproductive success thus pose serious obstacles to assessing fitness in hermaphroditic plants.

Most authors attempt to estimate individual male fertilities (e.g., Meagher 1986, 1987; Schoen and Stewart 1986; Broyles and Wyatt 1990; Devlin and Ellstrand 1990; Devlin et al. 1992; Conner et al. 1996b; Nason et al. 1996; Stacy et al. 1996; Campbell 1998; for reviews, see Snow and Lewis 1993; Luikart and England 1999; Sork et al. 1999). For some questions, however, properties of populations are more appropriate. For instance, many studies of selection in natural populations characterize the selection gradient (Lande 1976, 1979; Lande and Arnold 1983; Campbell et al. 1991; Johnston 1991; Morgan 1992; Conner et al. 1996b; Morgan and Schoen 1997). The selection gradient is the slope of the regression of relative fertility on trait value. Data usually consist of a set of points describing individual fertility and trait value. The parameter of interest, though, is the population-level selection gradient. Indirect selection gradient estimation procedures first estimate male fertilities and then fit the selection gradient to the estimated fertilities (e.g., Conner et al. 1996b). Estimating male fertilities of individuals requires as many parameters as there are males in the population.

Recent advances in estimation methodology greatly increase the utility of allozyme and other genetic markers for estimating selection gradients through male function. Adams and coworkers (Adams and Birkes 1991; Adams 1992; Ad-

ams et al. 1992; Burczyk et al. 1996; Burczyk and Prat 1997; see also Nason et al. 1996) directly estimate population characteristics. These authors use prior knowledge to restrict the pool of potential fathers (see also Campbell 1998). They then estimate the fraction of offspring sired from parents outside the neighborhood and fit a model relating plant characters (e.g., plant height, pollen fecundity, distance between putative parents) to reproductive success of individuals inside the neighborhood. Smouse et al. (1999) present a simpler version of Adams's model, ignoring fertilization events from outside the sampled population. Smouse et al. (1999) estimate a model relating plant characters to male fertility and equate estimated parameters to the selection gradients of quantitative genetics. Only a single parameter (one degree of freedom) is estimated for each phenotypic trait in the selection gradient context, many fewer than the number of parameters associated with indirect selection gradient estimation. Directly estimating the selection gradient may result in better estimates of relevant parameters while making more efficient use of genetic and phenotypic data. Nonetheless, Smouse et al. (1999) fail to detect significant selection in an analysis of previously published data from *Chamaelirium luteum* (Meagher 1986, 1987, 1991; Smouse and Meagher 1994).

Here we analyze male selection gradients on floral characters of *Raphanus raphanistrum* and investigate statistical properties of direct selection gradient estimates. We begin by describing our methods for maximum-likelihood estimation and for statistical assessment through parametric and simulation tests. We use these methods to directly estimate selection gradients in *R. raphanistrum,* comparing the efficacy of direct estimation with previously employed methods assigning male fertility to individual plants (Conner et al. 1996b). We then present simulations to investigate the statistical properties of selection gradient estimation. The sim-

ulation results provide important guidelines for realistic experimental design. Our results suggest that directly estimating selection gradients is a statistically powerful method for studying selection through male function in populations of 400 or more individuals.

<center>METHODS</center>

Existing Theoretical Context

Smouse et al. (1999) model the probability that offspring i is sired by paternal parent j. The model combines phenotypic and genetic probabilities of paternity. The phenotypic probability of paternity, λ_j, is the expected male fertility of the jth male divided by the male fertility of all individuals in the population, $\lambda_j = w_j/\Sigma_j w_j$. Expected male fertilities are modeled using the selection gradients of quantitative genetics (Lande 1976, 1979; Lande and Arnold 1983). Selection gradients are linear representations of the relationship between expected fertility and trait value, for example $w_j = \Sigma_t \beta_t z_{jt}$, where z_{jt} is the value of the tth trait measured on the jth male, β_t is a parameter common to all males describing the relationship between trait value and fertility (i.e., the selection gradient), and the summation is over all measured traits. The selection gradient parameters β_t can apply to continuous and discrete variables and reflect quadratic (stabilizing and disruptive) selection when terms have the form $\beta_t z_{jt}^2$. As in Lande (1976, 1979), weak directional and quadratic selection on relative (log) fitness approximates an exponential fitness function. The genetic probability of paternity, X_{ij}, is the probability that offspring genotype i results from known maternal genotype and potential paternal genotype j. Genetic information and the rules of Mendelian inheritance are used to calculate X_{ij}. The combined phenotypic and genetic probability that offspring i is sired by paternal parent j is $X_{ij} \lambda_j$.

Smouse et al. (1999) offered an important extension to the quantitative genetic context that incorporates distance between mates and other pairwise fitness determinants. For instance, pollen transfer is likely related to distance between both parental plants. Incorporating pairwise fitness determinants requires a measure of the distance z_{ijd} between the maternal parent of progeny i and the paternal parent j. Male fertility is specific to each maternal parent (i.e., the maternal parent of progeny i), $\lambda_{ij} = w_{ij}/\Sigma_{ij} w_{ij}$, with fertilities

$$w_{ij} = \exp\left[\sum_d \beta_d z_{ijd} + \sum_t \beta_t z_{jt}\right]. \quad (1)$$

The summation Σ_d is over all pairwise fitness components; β_d describes the regression of relative fertility on distance or other pairwise fitness components and is referred to as a diadic selection gradient (Smouse et al. 1999).

The model of Smouse et al. (1999) can be used to obtain maximum-likelihood estimates of male selection gradients β_t, β_d. The likelihood of observing the ith offspring is the sum of the probability of paternity over all males, $L_i = \Sigma_j X_{ij} \lambda_{ij}$. The likelihood of the entire dataset is the product, over all offspring, of the likelihood of each offspring,

$$L = \prod_i L_i = \prod_i \left[\sum_j X_{ij}\lambda_{ij}\right]. \quad (2)$$

The genetic probabilities X_{ij} and phenotypic trait values z_{jt}, z_{ijd} are obtained directly from the data. The goal is to identify selection gradients β_t, β_d that maximize L in equation (2). We used Newton-Raphson iteration to identify these selection gradients (computer program available from M. T. Morgan). This method converges more rapidly than the method used by Smouse et al. (1999) for the datasets used here, allowing simulation tests of statistical significance. With less genetic information or very strong selection, our Newton-Raphson method requires careful choice of initial parameter values for convergence.

Selection in Raphanus raphanistrum

We used data reported in Conner et al. (1996a, b). In 1991 and 1992, 57 and 60 plants, respectively, were positioned randomly in an experimental array. In 1993, seven groups of 16 plants each were exposed to pollinators over several days and at varying array positions. Each year, a suite of six floral morphological characters was measured on each plant. These were summarized as the first component (flower size) of a principal components analysis. Additional measures include anther exsertion, pollen production per flower, and total flower number. Each year, between 1866 and 2750 offspring were assayed for electrophoretic variation at eight allozyme loci.

Statistical significance of selection gradients was evaluated using parametric likelihood-ratio tests. The test involves calculation of the likelihood-ratio, LR. In the test, the numerator of LR is the estimated likelihood with a single parameter (selection gradient) constrained to zero. The denominator is the estimated likelihood of the full model. The likelihood ratio statistic, $-2 \ln LR$, is compared to a χ^2 distribution with one degree of freedom (Manly 1992). This tests the hypothesis that the constrained parameter equals zero. The procedure is repeated for each parameter in the model.

Nonparametric computer simulations were also used to evaluate parameter significance. Simulations are useful because the parametric test is only asymptotically distributed as χ^2. The simulations make fewer assumptions, particularly about gene frequency and linkage in the parental data, than the nonparametric test. The simulations do not address all assumptions, for instance, about the contribution of unsampled males and cryptic pollen flow. For these reasons the simulation tests are probably, although not certainly, more powerful than corresponding parametric tests. The simulations use the observed parental genotypes and phenotypes. Male fertilities of each genotype were assigned using equation (1), with parameters (selection gradients) estimated from the original data. The simulations created progeny for each female, with the number of progeny equal to the number of progeny of the corresponding female in the original dataset. Each progeny was created by choosing a male with probability in proportion to his assigned fertility. The progeny genotype was then constructed from the maternal and paternal genotypes based on the usual rules of Mendelian segregation. The simulated data was then subjected to the same estimation procedure as the original data. To evaluate the hypothesis that a particular parameter was equal to zero, male fertilities and consequently progeny genotypes were generated by replacing the parameter value estimated from the data with

zero. The simulated data (1000 replicates) give an empirical approximation to the distribution of the likelihood-ratio statistic. The proportion of values in the simulated distribution greater than the likelihood ratio of the original data approximates the P-value.

Investigation of Statistical Properties through Simulation

We performed simulations to explore the statistical properties associated with direct estimates of selection gradients. The simulations start by generating a hermaphroditic parental generation consisting of a specified number of individuals. Each parent consists of a multilocus genotype as well as phenotypic trait values. The multilocus genotype is generated by sampling alleles from a specified frequency distribution. Alleles are sampled independently within and between loci. Expected genotype frequencies in the parental population are therefore in Hardy-Weinberg proportions with linkage equilibrium between loci. Phenotypic trait values were assigned to each parental individual by sampling a value z from a standard normal (i.e., $\mu = 0$, $\sigma^2 = 1$) distribution. The values of z were then used, in conjunction with a specified value of β and equation (1), to determine male fertility of the individual.

The parental generation was used to simulate offspring genotypes. A specified number of individuals in the population were chosen at random and used as maternal parents. Each maternal parent produced an identical, specified number of progeny. Offspring for each maternal parent were generated independently of one another. A paternal parent was chosen for each offspring. The paternal parent was chosen using a random number generator, with the probability of choosing the jth male given by λ_j in equation (1). Self-fertilization was not explicitly incorporated, but occurred when by chance the selected paternal parent was the same individual as the current maternal parent. The genotypes of the maternal parent and of the paternal parent were then combined to generate the offspring genotype. The offspring genotype was generated by choosing one maternal and one paternal allele randomly at each locus, assuming that loci are unlinked. The choice of paternal parent and generation of offspring genotype was repeated for each offspring genotype of the maternal parent, and for all maternal parents from which offspring were sampled. The complete set of parental and offspring genotypes were then subjected to the estimation procedure outlined above.

The potentially large number of parameters (traits and their covariance structure, parameters of population, sampling, and genetic variation) precluded exhaustive investigation of statistical properties. Instead we focused on key properties that the investigator can anticipate. Properties investigated include variability of markers (number of equally frequent alleles per locus), number of paternal individuals, number of progeny and maternal genotypes sampled, and the fraction of the paternal population sampled. Populations consisted of 50, 100, 200, or 400 male parents. Progeny arrays of five, 10, 20, or 40 individuals were sampled from each of 50, 100, or 200 maternal parents. Exclusion probabilities investigated were $\epsilon \approx 0.810$, 0.917, 0.998. These probabilities were obtained by simulating eight, 12, or 30 loci with two equally

frequent alleles, respectively (Weir 1996, p. 210). A single selection gradient was estimated, with magnitude varying between zero and one in increments of 0.05. One thousand replicates were performed for each parameter combination.

The power to detect selection was measured as the selection gradient required for 95% of estimated gradients to exceed zero. The 1000 replicates for each parameter combination were ranked, and the 50th rank (i.e., the fifth percentile, separating the lowest 5% of estimated gradients) identified. This procedure was repeated for simulated selection gradients ranging from zero to one in steps of 0.05. The result is a series of pairs, each pair consisting of a simulated gradient (x) and the corresponding fifth percentile estimated gradient (y). The fifth percentile estimated gradients (i.e., y) were regressed on simulated gradients (i.e., x), and the x-intercept of the regression line determined. The x-intercept was used as an estimate of the selection gradient required for 95% of estimated gradients to exceed zero. The bias of the selection gradient estimator was measured as the average deviation of actual and estimated selection gradients.

RESULTS

Selection in Raphanus raphanistrum

Basic properties of the analysis are similar to those of Conner et al. (1996b). For instance, in 1992 genotypes were available for 60 maternal and paternal parents. Forty-seven paternal parents had measurements on all traits; all 60 paternal parents were measured for traits other than pollen number. A total of 2092 progeny were assayed. Genetic data excludes all but one paternal parent for 387 of the progeny (18.3% of all progeny, 3.49% of the 11,088 genetically possible parent-offspring triplets). A total of 243 progeny (11.6%) have genotypes consistent with none of the paternal parents and therefore do not contribute to selection gradient estimates. About 90% of fathers are excluded for each progeny.

Table 1 summarizes results of multivariate selection gradient analysis of plant characters. Conner et al. (1996b) were able to document significant selection only on flower number. Our analysis reveals selection on all traits measured, although the patterns of selection differ among years. The sign and magnitude of detected selection gradients are qualitatively similar between analyses of Conner et al. (1996b) and those presented in Table 1. This suggests that differences between analyses are mainly due to differences in statistical power.

Several results in Table 1 suggest caution in the interpretation of parametric significance tests, although in most cases parametric and simulation tests of significance agree. In the 1991 analysis the selection gradient associated with flower size has a likelihood ratio of 3.03. With one degree of freedom, this value is not statistically significant. Simulations show, however, that the gradient is significant at $P < 0.05$. Of greater concern is the likelihood ratio of 6.10 for squared anther exsertion in the 1992 analysis with all traits included. The parametric test indicates statistical significance ($P < 0.01$), but simulations show the parameter is not statistically significant ($P = 0.108$). The discrepancy likely results from violation of underlying assumptions (e.g., about the distribution of the fertility residuals) of the parametric test. Be-

TABLE 1. Maximum-likelihood estimates of male fertility selection gradients in experimental populations of *Raphanus raphanistrum,* without distance between mates in the model. In 1992, the all-plants column includes all 60 plants, but pollen number per flower is not in the analysis because this trait was not measured on 13 plants. The all-traits column includes the 47 plants for which all traits were measured. Superscripted significance values refer to χ^2 and simulation tests, respectively.

Character	1991	1992 All plants	1992 All traits	1993
Flower size	0.106[ns,*]	−0.018[ns,ns]	−0.163[**,***]	0.090[*,***]
(Flower size)^2	0.106[ns,ns]	−0.234[***,***]	−0.445[***,***]	0.016[ns,ns]
Anther exsertion	−0.050[ns,ns]	0.147[*,***]	0.026[ns,ns]	0.069[*,**]
(Anther exsertion)^2	−0.175[***,***]	−0.053[ns,ns]	−0.096[**,ns]	−0.008[ns,ns]
Pollen no./flower	—	—	0.278[***,***]	0.078[*,ns]
Flower production	1.480[***,***]	0.571[***,***]	0.451[***,***]	0.075[*,*]

[ns] $P > 0.05$; * $P < 0.05$; ** $P < 0.01$; *** $P < 0.005$.

cause the simulations address some, but not all, assumptions (as outlined in the Methods section), there is no guarantee that the simulations have greater power than the parametric tests.

Table 2 includes distance between plants as a pairwise factor in the selection gradient analysis. Data from 1993 are not included because these plants were placed in different positions on successive days (see Conner et al. 1996b). In all analyses, the estimated diadic selection gradient through distance is negative and highly significant, indicating that proximity increases probability of paternity. Distance generally does not have substantial influence on the interpretation of selection gradients. The exceptions are two of the gradients for anther exsertion in 1992 that were significant using simulation tests, but only when distance was included in the model.

There was significant directional selection for increased flower size in 1991 and 1993, with no significant curvature in the fitness function (Fig. 1; Tables 1, 2). In contrast, there was strong stabilizing selection (negative quadratic gradient) on flower size in 1992, with no significant linear term when all plants are included, and a significant negative linear term with pollen number per flower in the model (Fig. 2; Tables 1, 2).

The linear selection gradients for anther exsertion were negative and not significant in 1991, but significantly positive in 1992 and 1993 (Tables 1, 2; in 1992 the linear gradient is not significant with 47 paternal plants and distance excluded, but all other tests were significant). The quadratic selection gradients for anther exsertion were all negative, but not always significant: Parametric and simulation tests of quadratic selection on anther exsertion were significant in

1991 but neither was significant in 1993 (Tables 1, 2). In 1992, four of the eight tests were significant, including the more robust simulation test with all plants and distance included. The weight of evidence suggests stabilizing selection on anther exsertion in 1991, directional selection for increased exsertion in 1993 (Fig. 1), and a combination of the two in 1992, that is, a positive decelerating fitness function (Fig. 2; note that these plots are from models with distance excluded).

There was highly significant directional selection for increased pollen number per flower in 1992, and weaker and marginally significant selection for an increase in 1993 (Tables 1, 2). Similarly, there was very strong directional selection for increased flower number in 1991 and 1992, and weaker but still significant selection for an increase in 1993 (Tables 1, 2). The patterns of selection on flower number are in agreement with Conner et al. (1996b).

Investigation of Statistical Properties through Simulation

Table 3 summarizes simulations investigating key statistical properties of populations in which all males are genotyped. Each entry in the table is an estimate of the selection gradient for which 95% of the simulations exceed zero. This provides a guide to the standardized selection gradient likely to result in statistical significance. The table shows the expected result that greater exclusion probability, fewer fathers, more maternal parents, or more progeny per maternal parent each decrease the detectable selection gradient (i.e., increase statistical power). An encouraging result is that increased sample size can partly compensate for low exclusion probability. For instance, a four-fold increase in total progeny

TABLE 2. Maximum-likelihood estimates of male fertility selection gradients in experimental populations of *Raphanus raphanistrum,* with distance between mates included in the models. Superscripted significance values refer to χ^2 and simulation tests, respectively.

Character	1991	1992 All plants	1992 All traits
Flower size	0.095[ns,*]	−0.023[ns,ns]	−0.035[ns,***]
(Flower size)^2	0.066[ns,ns]	−0.259[***,***]	−0.446[***,***]
Anther exsertion	−0.085[ns,ns]	0.167[***,***]	0.132[*,**]
(Anther exsertion)^2	−0.139[*,*]	−0.085[**,***]	−0.129[***,ns]
Pollen no./flower	—	—	0.256[***,***]
Flower production	1.32[***,***]	0.505[***,***]	0.385[***,***]
Distance	−0.233[***,***]	−0.321[***,***]	−0.305[***,***]

[ns] $P > 0.05$; * $P < 0.05$; ** $P < 0.01$; *** $P < 0.005$.

FIG. 1. Multivariate selection on flower size (left panels) and anther exsertion (right panels) in 1991 and 1993. Linear (β) and quadratic (γ) regression coefficients and their significance levels are those in Table 1. Lines represent the relative fitness function estimated from the fitted model; points represent *estimated* fertilities under the fitted model. Note that the relative fitness scale differs between years.

sampled (e.g., from 10 to 40 from each of 50 maternal parents) almost halves the detectable selection gradient.

The total number of progeny sampled, rather than number of maternal parents or progeny size per maternal parent, determines the detectable selection gradient. For instance, in a population of 400 potential males, sampling 200 maternal families each for 10 progeny results in detectable selection gradients of 0.21. Nearly identical results are obtained when sampling 100 maternal families of 20 progeny (detectable selection gradient 0.23) or 50 maternal families of 40 progeny (detectable selection gradient 0.22). These results show that the ability to detect selection can be increased through addition of families or addition of progeny per family. This contrasts with selfing rate estimation, where addition of families is more important when progeny consist of six to 10 individuals (Brown and Allard 1970).

Increasing the number of possible fathers without changing progeny sample size decreases the ability to detect selection, especially at lower exclusion probabilities. For instance, with an exclusion probability $\epsilon = 0.81$ in a population of 50 potential male parents, sampling 50 maternal parents for 10 progeny each allows detection of a selection gradient equal to 0.18. In a population of 400 potential male parents, the detectable selection gradient is more than doubled, to 0.43.

At intermediate exclusion ($\epsilon = 0.917$), increased numbers of fathers decreases statistical power, although with reasonable progeny sizes selection gradients < 0.2 can still be detected. For polymorphism with $\epsilon = 0.998$ (Table 3), increasing the number of possible fathers eight-fold (from 50 to 400) causes very small changes in detectable selection gradients.

Partial sampling of male genotypes reduces the ability to detect significant selection (Table 4). Polymorphism with exclusion probability $\epsilon = 0.81$ is especially susceptible to reduced power from unsampled males. Sampling 90% of the male population (i.e., 50 of 55 males) increases the detectable selection gradient from 0.13 to 0.17. Sampling only one half of the potential male genotypes makes detection of all but the strongest selection impossible with exclusion probability $\epsilon = 0.81$. The effect of partial sampling is considerably reduced with exclusion probability $\epsilon = 0.998$, presumably because this level of exclusion reduces false paternity attribution while eliminating genetically impossible offspring from the analysis.

When all males are genotyped, the selection gradient estimator is virtually unbiased. For instance, the magnitude of the 1000-replicate average deviation between observed and expected values is less than 5% for almost all parameter combinations. This is illustrated by the near diagonal line in
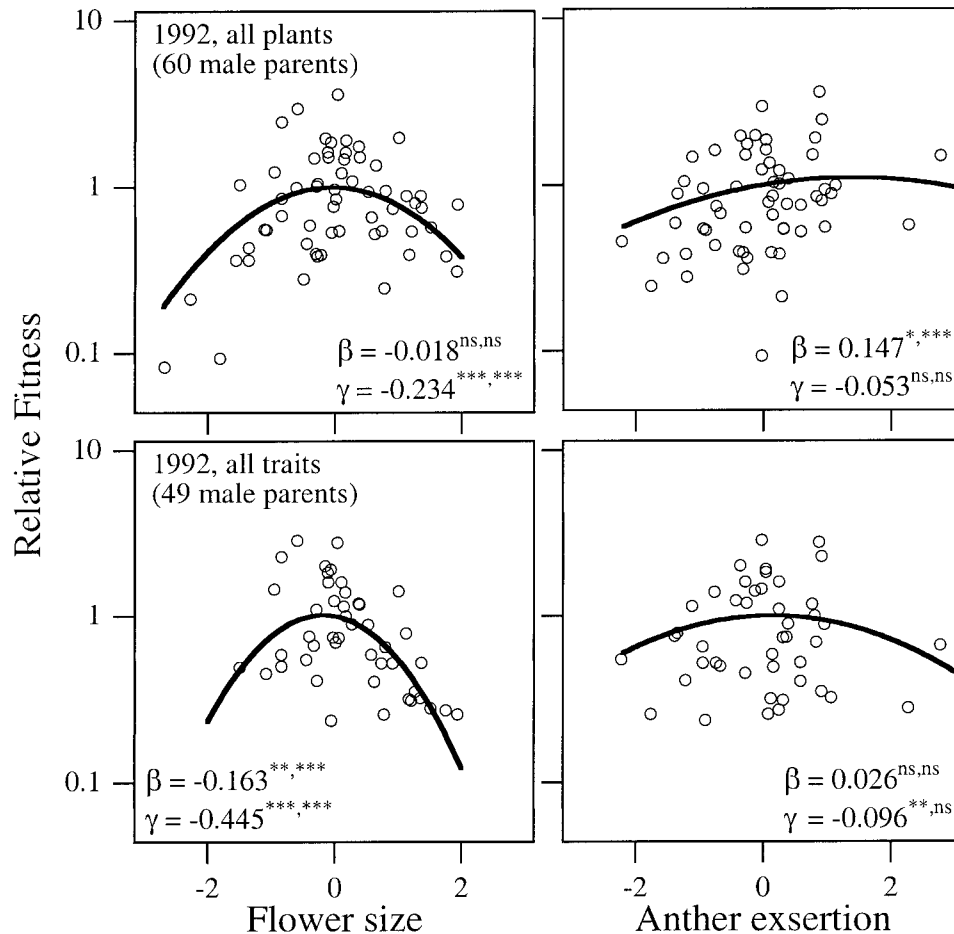
FIG. 2. Multivariate selection on flower size (left panels) and anther exsertion (right panels) in 1992 including all plants ($N = 60$ male parents, analysis excludes pollen number) or all traits ($N = 49$ male parents). Linear (β) and quadratic (γ) regression coefficients and their significance levels are those in Table 1.

Fig. 3 when all males (i.e., 50 male genotypes in a population of 50 potential paternal parents) are sampled. Partial sampling of male genotypes may bias estimates of selection gradients. This is shown in Fig. 3, where 50 male genotypes were determined in populations of 100 potential paternal parents. With only half the males genotyped, the estimator underestimates the selection gradient. The magnitude of the bias is proportional to the actual selection gradient. Moderate exclusion probabilities (e.g., $\epsilon = 0.81$) can lead to severely biased estimates. The reason for bias is that paternity is partly or entirely attributed to genotyped males, when the actual father is among those males who were not sampled. This partial attribution of paternity obscures the relationship between trait value and actual fertility, resulting in an underestimate of the true selection gradient.

DISCUSSION

Selection in Raphanus raphanistrum

Comparing the results of the original empirical study on radish (Conner et al. 1996b) with those presented here (Tables 1, 2; Figs. 1, 2) shows the increased statistical power of the direct selection gradient estimation method. In general, the

magnitudes of the gradients estimated were similar with the two methods, but many more of these gradients were statistically significant with direct analysis. We were able to find significant selection on floral morphology and pollen number per flower where the original analysis failed to do so. We also see differences in selection among years in flower size and perhaps anther exsertion.

Selection on floral morphology through differences in male fitness is usually assumed to be mediated through pollinators. In 1991, the majority of visitors were honey bees (55% of all visits) and small bees (32%), whereas in 1992 and 1993 honey bees were rare and small bees and syrphid flies dominated (40% of visits each in 1992, 55% and 38% in 1993 respectively; Conner et al. 1996a). All these taxa are effective pollinators of wild radish (Conner et al. 1995).

The directional selection for increased flower size in 1991 and 1993 makes sense based on the preferences of honey bees, small bees, and syrphid flies for large flowers in wild radish (Stanton and Preston 1988; Conner and Rush 1996). The strong stabilizing selection on flower size in 1992 is difficult to explain based on pollinator preferences. It suggests some cost to large flowers in this year, perhaps through trade-offs with pollen number, size, or quality. The pheno-

TABLE 3. Consequences of population, genetic, and sampling parameters for statistical power. Entries under progeny per maternal plant correspond to selection gradients where 95% of simulated datasets exceed zero. Additional details provided in the text.

Number of		Progeny per maternal plant			
Maternal plants	Paternal plants	5	10	20	40
$\epsilon = 0.81$					
50	50	0.24	0.18	0.13	0.09
	100	0.33	0.25	0.16	0.11
	200	0.43	0.29	0.23	0.15
	400	0.65	0.43	0.29	0.22
100	100	0.23	0.16	0.12	
	200	0.31	0.23	0.16	
	400	0.43	0.32	0.23	
200	200	0.25	0.15		
	400	0.30	0.21		
$\epsilon = 0.917$					
50	50	0.16	0.13	0.080	0.053
	100	0.19	0.14	0.098	0.069
	200	0.25	0.18	0.127	0.090
	400	0.32	0.23	0.174	0.122
100	100	0.14	0.10	0.068	
	200	0.18	0.13	0.093	
	400	0.23	0.17	0.119	
200	200	0.12	0.09		
	400	0.17	0.12		
$\epsilon = 0.998$					
50	50	0.108	0.077	0.057	0.039
	100	0.112	0.079	0.056	0.040
	200	0.107	0.080	0.054	0.043
	400	0.112	0.083	0.066	0.042
100	100	0.075	0.057	0.039	
	200	0.077	0.059	0.040	
	400	0.079	0.060	0.042	
200	200	0.051	0.040		
	400	0.054	0.047		

TABLE 4. Reduced statistical power introduced by partial sampling of male genotypes. The first column gives the number of paternal parents in the simulated population. In all cases, however, genetic marker information is available for only 50 of these. Entries in the table correspond to selection gradients where 95% of simulated datasets exceed zero. For all results, 50 maternal progeny arrays of 20 individuals each were sampled.

Paternal plants in population	Detectable selection gradient		
	$\epsilon = 0.81$	$\epsilon = 0.917$	$\epsilon = 0.998$
50	0.13	0.08	0.055
55	0.17	0.11	0.059
60	0.21	0.14	0.067
75	0.33	0.22	0.086
100	0.61	0.38	0.120

more exserted anthers. An important goal for future studies is to better understand the mechanisms and agents of selection on floral traits through male fitness differences.

Our results from the direct selection gradient analysis give new support for the hypothesis that the high correlation between filament and corolla tube seen in wild radish and some other Brassicaceae (Conner and Sterling 1995) is caused by stabilizing selection on anther exsertion (Conner and Via 1993; Conner 1997). There was strong evidence for stabilizing selection on anther exsertion in 1991, and some evidence for a negative quadratic term in 1992 (Tables 1, 2; Figs. 1, 2). It is possible that increased variation in exsertion would reveal more consistent evidence for stabilizing selection in this trait; this is a goal of current artificial selection experiments.

Conner et al. (1996b) concluded that selection through female function in *R. raphanistrum* is stronger than selection through male function. This was based on the observation of significant selection acting on morphology through female function but not through male function. The analysis pre-

typic correlation between flower size and pollen production per flower in 1992 was weakly positive ($r = 0.22$, $P = 0.12$), and the linear selection gradient on flower size was significantly negative when this correlation was corrected for by including pollen number in the model (Tables 1, 2; Fig. 2). This suggests that the trade-off is not with pollen number, because the negative effects of large flowers on male fitness become even stronger when pollen number is corrected for. A trade-off with pollen size or quality is not ruled out.

The differences in selection gradients across years were less clear for anther exsertion than for flower size—and not in directions that one would expect based on pollinator differences. Of the three major visitor taxa in this study, only honey bees removed more pollen from more exserted anthers in a previous study (Conner et al. 1995). However, honey bees were only common in the year (1991) in which there was no evidence for directional selection for increased exsertion. There was some evidence for higher pollen removal by small bees from flowers with intermediate anther exsertion (Conner 1997), which could explain the stabilizing selection on anther exsertion in 1991 and the negative quadratic term in 1993. The directional selection for increased anther exsertion in 1992 and 1993 is difficult to explain based on patterns of pollen removal in previous studies. In both these years small bees and syrphid flies made most of the visits, but neither of these pollinators removed more pollen from
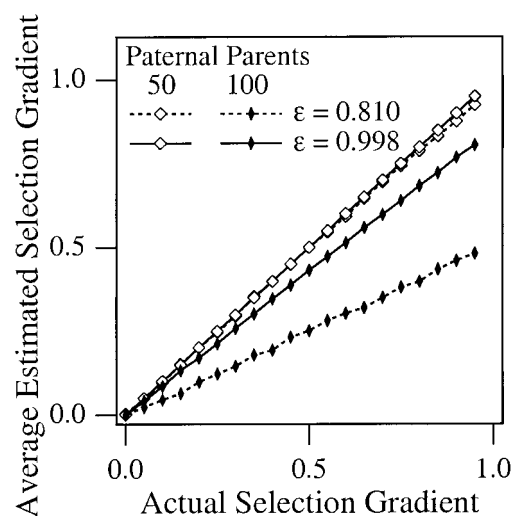

FIG. 3. Bias introduced by partial sampling of male genotypes. For all results, 50 maternal progeny arrays of 20 individuals each were sampled, with either $\epsilon = 0.81$ or $\epsilon = 0.998$ levels of polymorphism. The population consisted of either 50 or 100 males, but genotypes for only 50 were obtained.

sented here suggests a different conclusion, because a substantial amount of selection occurs through male function.

Comparison of male versus female selection is complicated by issues of statistical power and bias. The power associated with directly estimating female selection gradients (through direct observation of fertility) is likely to be greater than that for detecting male selection gradients (through the indirect inference outlined here). In addition, most studies of paternity include a fraction of offspring not sired by censused males. For instance, in the 1992 dataset of Conner et al. (1996b), 243 of the 2091 progeny ($\approx$11.6%) had genotypes inconsistent with any of the censused males. Some of these individuals might represent scoring or other processing errors. The remainder, however, indicate that a fraction of the male population was not censused. This and Fig. 3 imply that the gradients estimated here are underestimates. Table 4 indicates reduced statistical power when some males remain unsampled, so the assessment of statistical significance reported here is likely to be conservative. These issues of bias and power tend to decrease the magnitude and statistical significance of male gradients relative to female gradients. It is therefore possible that selection on floral morphology was stronger through male function than female function in this study of *Raphanus*. This is in agreement with earlier results for flower color in this species (Stanton et al. 1986, 1989).

Investigation of Statistical Properties through Simulation

Simulation results provide encouragement and guidance for marker-based inference of male fertility selection. Low exclusion probability (e.g., resulting from allele frequency distribution) can be partly compensated by sampling more offspring. Results indicate that more offspring can be sampled either through inclusion of additional maternal parents or of additional progeny from each maternal parent. This is in contrast to estimation of selfing rate, where progeny sizes of six to 10 per maternal family are suggested as optimal (Brown and Allard 1970). The suggested family size for selfing rate estimation is the result of a trade-off between inference of maternal genotype and estimation of parental genotype proportions. The models presented here require explicit knowledge of parental genotype, so the trade-off is not relevant. In addition, our assumption of random mating results in a likelihood that is the product over all progeny, regardless of maternal family (eq. 2). Progeny and families are exchangeable to the extent that maternal genotypes result in equal paternal exclusion probabilities.

Simulation results (Table 3) indicate that selection gradients can be estimated in populations consisting of several hundred individuals, provided sufficient genotypic information (e.g., $\epsilon = 0.998$) is obtained from all individuals in the population. Even with an exclusion probability of only 0.81 and 400 possible fathers, a selection gradient of 0.21 is detectable with 2000 offspring (the sample size in Conner et al. 1996b). The method used by Conner et al. estimates male fertilities on individual females, which requires that individual maternal progeny arrays are larger than the number of male fertilities estimated (see Roeder et al. 1989). This requirement was not always met in 1991 and 1992, so that male fertility estimates in these years may have inflated variances.

Each parameter estimated in this study uses information from all progeny, and fewer parameters are being estimated. Therefore, error about the estimates is reduced and statistical power increased.

The different exclusion probabilities for all simulations presented above were obtained by varying the number of biallelic loci. Additional simulations show that the results are not sensitive to underlying genetic details responsible for exclusion. For instance, use of four loci each with 10 equally frequent codominant alleles (e.g., multilocus microsatellite) or of 50 dominant loci with frequency of the dominant allele equal to 0.25 (e.g., Amplified Fragment Polymorphisms, or AFLPs) yields an exclusion probability ($\epsilon = 0.998$) and results nearly identical to those obtained with 30 biallelic loci. The first exclusion probability ($\epsilon = 0.81$) corresponds to a reasonable set of allozyme loci. The second ($\epsilon = 0.917$) is close to that anticipated with highly variable allozymes or a single microsatellite locus with 10 equally frequent alleles ($\epsilon \approx 0.87$). The third ($\epsilon = 0.998$) might result from multiple microsatellite loci (e.g., four loci, each with 10 alleles, $\epsilon \approx 0.998$), or an AFLP analysis (50 bands, dominant phenotype at frequency 0.25, $\epsilon \approx 0.996$). Deviation of allele or band frequencies from the optima in these scenarios will substantially decrease the exclusion probability and thus reduce ability of the investigator to detect selection (Weir 1996, p. 210).

Highly polymorphic markers enhance the statistical opportunities for assessing paternity, without obviating the need for methods to estimate male selection gradients directly from marker data. For instance, a recent study (Streiff et al. 1999) used six microsatellite loci to study paternity in *Quercus* species. The study consisted of 984 progeny from 13 females. Genotype information was available for 296 males. The authors used paternity exclusion to assign 310 progeny to unique parents. Seventeen additional progeny had ambiguous paternity, whereas the remaining 657 seedlings had no compatible male genotype. This large fraction of unassigned progeny implies a male population size many times larger than the approximately 300 males censused. Unbiased estimates of selection gradients would require sampling many more males. This, in turn, is likely to reduce the number of uniquely assigned seedlings and thus require the kind of statistical analysis developed here. A further refinement might involve sampling fewer progeny from more maternal families. Sampling in this manner randomizes estimated male fertilities over more females and, in contrast to previous paternity assignment methods, does not compromise the ability to detect selection (Table 3).

Smouse et al. (1999) illustrate the direct estimation of selection gradients with an analysis of *C. luteum* allozyme data. The data consist of 2294 seeds collected from 70 females (approximately 33 progeny per female), with genotypes (polymorphic allozyme loci with exclusion probability $\epsilon \approx 0.73$; Smouse and Meagher 1994) and trait values available for 273 males. Estimated selection gradients range from -0.02 to 0.12, but none are statistically significant. Smouse et al. (1999) argue that the absence of statistical significance is not due to small sample size, but do not support this argument with simulation or power analysis. Simulations of data comparable to that of Smouse et al. (Table 3; M. T. Morgan, unpubl. data) suggest that selection gradients with

magnitude < 0.2 would not be detectable with exclusion probability < 0.8 even if all males were sampled. Thus, statistical power might limit the ability to detect relatively strong selection through male function in *C. luteum*. In addition, Smouse et al. (1999) evaluate statistical significance by generating a distribution of likelihood ratios associated with shuffled male trait values. This method evaluates how the fit of the actual data compares with other permutations, rather than assessing fit compared with a null hypothesis involving constrained parameters. Both parametric and simulation methods were used to assess statistical significance.

Both methods compare the likelihood under the fitted model to the likelihood under a null hypothesis, $\beta_t = 0$. The most probable explanation for differences in the significance indicated by the methods relates to the distribution of likelihood ratios under the null model. The parametric test makes explicit assumptions about the form of this distribution, whereas the simulations do not. For this reason the simulations appear preferable, although conflicts in assessment of statistical significance should probably signal caution in conclusions drawn from the analysis.

Distance between mates is likely to influence fertility in plant populations. Including distance may therefore account for sources of variation in fertility and therefore increase statistical power to correctly infer selection on morphological characters, as seen in Table 2. One potential problem arises when uncensused males sire progeny. Uncensused males may represent long-distance gene flow, so the dispersal parameter estimate (e.g., Table 2) may underestimate actual dispersal. The underestimate of dispersal distance influences the paternity probability of males, so that estimates of selection gradients may be biased. Estimates of effective population size based on dispersal distance (e.g., Burczyk et al. 1996) are also likely to be underestimates. Including distance in selection gradient analysis can improve statistical power, but may introduce additional artifacts when some males remain uncensused.

CONCLUSIONS

The analysis and simulations presented here show that direct estimates of selection gradients offer substantial statistical power. Features of floral morphology in *R. raphanistrum* experience strong directional and stabilizing selection through male function. This contrasts with previous analyses of the same dataset using paternity assignment techniques as an intermediate step in documenting character selection (Conner et al. 1996b). The previous analyses found no significant selection on floral morphology through male function. Simulations provide guidance in experimental design, suggesting feasible studies in populations of several hundred individuals. Genetic variation from highly polymorphic markers may allow methods to be applied in populations of more than 400 individuals. An important caveat on this optimism is the requirement to assess the genotype of most fathers in the population.

ACKNOWLEDGMENTS

We thank P. E. Smouse and colleagues for sharing their manuscript prior to publication. K. Holsinger and anonymous reviewers provided very insightful comments. This research was supported by the National Science Foundation under grants DEB 9796183 and DEB 9796185 to JC. This is Kellogg Biological Station contribution no. 939.

LITERATURE CITED

Adams, W. T. 1992. Gene dispersal within forest tree populations. New For. 6:217–240.

Adams, W. T., and D. S. Birkes. 1991. Estimating mating patterns in forest tree populations. Pp. 157–172 *in* S. Fineschi, M. E. Malvolti, F. Cannata, and H. H. Hattemer, eds. Biochemical markers in the population genetics of forest trees. SPB Academic Publishing, The Hague.

Adams, W. T., A. R. Griffin, and G. F. Moran. 1992. Using paternity analysis to measure effective pollen dispersal in plant populations. Am. Nat. 140:762–780.

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

Broyles, S. B., and R. Wyatt. 1990. Paternity analysis in a natural population of *Asclepias exaltata*: multiple paternity, functional gender, and the ''pollen-donation'' hypothesis. Evolution 44:1454–1468.

Burczyk, J., and D. Prat. 1997. Male reproductive success in *Pseudotsuga menziesii* (Mirb.) Franco: the effects of spatial structure and flowering characteristics. Heredity 79:638–647.

Burczyk, J., W. T. Adams, and J. Y. Shimizu. 1996. Mating patterns and pollen dispersal in a natural knobcone pine (*Pinus attenuata* Lemmon.) stand. Heredity 77:251–260.

Campbell, D. R. 1998. Variation in lifetime male fitness in *Ipomopsis aggregata*: tests of sex allocation theory. Am. Nat. 152:338–353.

Campbell, D. R., N. M. Waser, M. V. Price, E. A. Lynch, and R. J. Mitchell. 1991. Components of phenotypic selection: pollen export and flower corolla width in *Ipomopsis aggregata*. Evolution 45:1458–1467.

Charnov, E. L. 1982. The theory of sex allocation. Princeton Univ. Press, Princeton, NJ.

Conner, J. K. 1997. Floral evolution in wild radish: the roles of pollinators, natural selection, and genetic correlations among traits. International J. Plant Sci. 158:S108–S120.

Conner, J. K., and S. Rush. 1996. Effects of flower size and number on pollinator visitation to wild radish, *Raphanus raphanistrum*. Oecologia 105:509–516.

Conner, J. K., and A. Sterling. 1995. Testing hypotheses of functional relationships: a comparative survey of correlation patterns among floral traits in five insect-pollinated plants. Am. J. Bot. 82:1399–1406.

Conner, J. K., and S. Via. 1993. Patterns of phenotypic and genetic correlations among morphological and life-history traits in wild radish, *Raphanus raphanistrum*. Evolution 47:704–711.

Conner, J. K., R. Davis, and S. Rush. 1995. The effect of wild radish floral morphology on pollination efficiency by four taxa of pollinators. Oecologia 104:234–245.

Conner, J. K., S. Rush, and P. Jennetten. 1996a. Measurements of natural selection on floral traits in wild radish (*Raphanus raphanistrum*). 1. selection through lifetime female fitness. Evolution 50:1127–1136.

Conner, J. K., S. Rush, S. Kercher, and P. Jennetten. 1996b. Measurements of natural selection on floral traits in wild radish (*Raphanus raphanistrum*). 2. Selection through lifetime male and total fitness. Evolution 50:1137–1146.

Devlin, B., and N. C. Ellstrand. 1990. Male and female fertility variation in wild radish, a hermaphrodite. Am. Nat. 136:87–107.

Devlin, B., J. Clegg, and N. C. Ellstrand. 1992. The effect of flower production on male reproductive success in wild radish populations. Evolution 46:1030–1042.

Johnston, M. O. 1991. Natural selection on floral traits in two species of *Lobelia* with different pollinators. Evolution 45:1468–1479.

Lande, R. 1976. Natural selection and random genetic drift in phenotypic evolution. Evolution 30:314–334.

———. 1979. Quantitative genetic analysis of multivariate evolution, applied to brain:body size allometry. Evolution 33: 402–416.

Lande, R., and S. J. Arnold. 1983. The measurement of selection on correlated characters. Evolution 36:1210–1226.

Luikart, G., and P. R. England. 1999. Statistical analysis of microsatellite DNA data. Trends Ecol. Evol. 14:253–256.

Manly, B. F. J. 1992. The design and analysis of research studies. Cambridge Univ. Press, Cambridge, U.K.

Meagher, T. R. 1986. Analysis of paternity within a single natural population of *Chamaelirium luteum*. 1. Identification of most-likely male parents. Am. Nat. 128:199–215.

———. 1987. Analysis of parentage for naturally established seedlings within a population of *Chamaelirium luteum* (Liliaceae). Ecology 68:803–812.

———. 1991. Analysis of paternity within a natural population of *Chamaelirium luteum*. II. Patterns of male reproductive success. Am. Nat. 137:738–752.

Morgan, M. T. 1992. The evolution of traits influencing male and female fertility in outcrossing plants. Am. Nat. 139:1022–1051.

———. 1998. Properties of maximum likelihood male fertility estimation in plant populations. Genetics 149:1099–1103.

Morgan, M. T., and D. J. Schoen. 1997. Selection on reproductive characters: floral morphology in *Asclepias syriaca*. Heredity 79: 433–441.

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

Roeder, K., B. Devlin, and B. G. Lindsay. 1989. Application of maximum likelihood methods to population genetic data for the estimation of individual fertilities. Biometrics 45:363–379.

Schoen, D. J., and S. C. Stewart. 1986. Variation in male reproductive investment and male reproductive success in white spruce. Evolution 40:1109–1120.

Smouse, P. E., and T. R. Meagher. 1994. Genetic analysis of male reproductive contributions in *Chamaelirium luteum* (L.) Gray (Liliaceae). Genetics 136:313–322.

Smouse, P. E., T. R. Meagher, and C. J. Kobak. 1999. Parentage analysis in *Chamaelirium luteum* (L.) Gray (Liliaceae): why do some males have disproportionate reproductive contributions? J. Evol. Biol. 12:1069–1077.

Snow, A. A., and P. O. Lewis. 1993. Reproductive traits and male fertility in plants—empirical approaches. Annu. Rev. Ecol. Syst. 24:331–351.

Sork, V. L., J. Nason, D. R. Campbell, and J. F. Fernandez. 1999. Landscape approaches to historical and contemporary gene flow in plants. Trends Ecol. Evol. 14:219–224.

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

Stanton, M. L., and R. E. Preston. 1988. Ecological consequences and phenotypic correlates of petal size variation in wild radish, *Raphanus sativus* (Brassicaceae). Am. J. Bot. 75:528–539.

Stanton, M. L., A. A. Snow, and S. N. Handel. 1986. Floral evolution: attractiveness to pollinators increases male fitness. Science 232:1625–1627.

Stanton, M. L., A. A. Snow, S. N. Handel, and J. Bereczky. 1989. The impact of a flower-color polymorphism on mating patterns in experimental populations of wild radish. Evolution 43: 335–346.

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

Weir, B. S. 1996. Genetic data analysis. vol. II. Sinauer, Sunderland, MA.

Corresponding Editor: K. Holsinger