Theor Appl Genet (1999) 98: 1063–1071

C. T. Ivey · R. Wyatt

Family outcrossing rates and neighborhood floral density in natural populations of swamp milkweed (*Asclepias incarnata*): potential statistical artifacts

Received: 28 September 1998 / Accepted: 27 October 1998

Abstract To evaluate how environmental and genetic factors influence mating-system evolution, accurate estimates of outcrossing rates of individual plants (families) are required. Using isozyme markers, we observed wide variation in family outcrossing rates in three natural populations of *Asclepias incarnata* using three statistical methods: (1) a multilocus maximum-likelihood procedure (t_m); (2) a multilocus method-of-moments procedure (t_a); and (3) a direct comparison of progeny phenotypes against maternal phenotypes (t_d). Neighborhood floral-display size was positively correlated with t_a in one population, but showed no relationship with any of the other estimates of outcrossing for any population. Monte-Carlo simulations revealed that statistical variation associated with these estimation procedures can be large enough to explain all of the observed variation in outcrossing. We also found that significant, spurious correlations with neighborhood floral display could arise, on average, 7% of the time by chance alone. Our observations suggest that it is difficult to obtain accurate estimates of outcrossing in naturally pollinated plants using the estimation procedures currently available. Moreover, we caution that attempts to interpret observed variation in family outcrossing estimates by observing variation in ecological parameters could be misleading.

Key words *Asclepias incarnata* · Flower-density family outcrossing · Monte-Carlo simulation · Self-compatibility

Communicated by P. M. A. Tigerstedt

C. T. Ivey (✉)
Department of Botany, University of Georgia, Athens, Georgia, 30602, USA, and Organization for Tropical Studies, Apartado 676-2050, San Pedro, Costa Rica
E-mail: chrisivey@cro.ots.ac.cr

R. Wyatt
Institute of Ecology, University of Georgia, Athens, Georgia, 30602, USA

Introduction

Plant density can affect the movements of pollinators between plants, which in turn can influence patterns of plant mating (e.g., Kunin 1993; Karron et al. 1995). Pollinators may visit a higher proportion of flowers on isolated plants than on plants with several closely spaced neighbors (Kunin 1993). As a consequence, relatively isolated plants are likely to experience a higher incidence of geitonogamous pollinations than are densely spaced plants. For plants with mixed mating systems, this may lead to an increase in the proportion of self-fertilized progeny produced by that plant.

Although pollinator behavior can be directly observed, breeding patterns in plants must be discerned through inference. Fortunately, in recent decades, the increased availability of genetic markers has facilitated the estimation of plant outcrossing rates at the population level (e.g., Clegg 1980; Barrett and Harder 1996; Cruzan 1998). Robust statistical methods have been developed to estimate mating-system parameters from allelic variation observed in progeny arrays (e.g., Ritland and Jain 1981; Shaw et al. 1981). In the past, workers relied on controlled experimental crosses to characterize plant mating systems, a method which provided incomplete information about the actual mating behavior of plants in nature. These more recent developments have allowed considerable progress toward understanding the behavior of mating systems in natural plant populations (Brown et al. 1985; Jarne and Charlesworth 1993; Barrett and Harder 1996).

Most models of population-level mating systems assume a constant outcrossing rate among maternal plants (Ritland and Jain 1981; Shaw et al. 1981; Schoen and Clegg 1984; Cruzan et al. 1994). In nature, however, plant density (Ellstrand et al. 1978; Watkins and Levin 1990; Karron et al. 1995), as well as other factors (Smyth and Hamrick 1984; Ritland and Ganders 1985), may lead to violations of this assumption. Thus, there has been considerable parallel interest in extending

these statistical techniques to estimate outcrossing rates of individual, mixed-mating plants in nature, in order to understand more fully the influence of environmental and genetic factors on the evolution of mating systems and associated floral characteristics (Smyth and Hamrick 1984; Ritland and Ganders 1985; Morgan and Barrett 1990; Motten and Antonovics 1992; Cruzan et al. 1994).

The most widely used estimation procedure is a multilocus maximum-likelihood technique (Ritland and Jain 1981) based on the mixed-mating model described by Fyfe and Bailey (1951). Analytical and simulation studies have shown that this model generates robust estimates of population-level mating-system parameters when at least 4–5 loci are included in the analysis (Ritland and Jain 1981). Some studies, however, have reported that family level outcrossing estimates generated by this program can be highly variable statistically (Ritland and Ganders 1985; Morgan and Barrett 1990). In addition, earlier versions of this program used a Newton-Raphson procedure to estimate outcrossing rate. When this procedure is employed, difficulties with estimate convergence are often observed, possibly due to families violating model assumptions of homogeneity among pollen allele frequencies or due to the small numbers of progeny sampled (Morgan and Barrett 1990; Perry and Knowles 1990; Cruzan et al. 1994; Burgess et al. 1996). More recent versions of the maximum-likelihood procedure offer an alternative estimation-maximization procedure, which has fewer convergence problems.

To address these concerns, Cruzan et al. (1994) developed a program to estimate family outcrossing rate that is based on an extension of the method-of-moments procedure described by Shaw et al. (1981) and which also evaluates multiple loci to estimate outcrossing rate. This procedure appears to make somewhat less efficient use of data in comparison to the maximum-likelihood model, but it relaxes assumptions about the frequency of genotypes in the outcross pollen pool (Ritland and Jain 1981; Cruzan and Arnold 1994; Cruzan et al. 1994). Perhaps as a consequence of these relaxed assumptions, Cruzan et al. (1994) reported that all families converged on outcrossing estimates between 0 and 1. Thus, Cruzan et al. (1994) suggested that this model is more stable under a wider variety of mating patterns, which would make it a useful procedure for estimating family outcrossing rates in natural populations. Unfortunately, the statistical variation associated with these estimates has not been estimated, either through simulation or analytical procedures.

Here, we compare family outcrossing estimates generated by both Ritland and Jain's (1981) model and Cruzan et al.'s (1994) procedure, using identical data from three natural populations of swamp milkweed (*Asclepias incarnata* L.). Specifically, we compare these estimates with respect to an ecological variable: the surrounding density of flowers. We also use Monte-Carlo simulation to characterize the accuracy of the estimates and to estimate the likelihood of spurious correlations of neighborhood flowering density with outcrossing estimates.

Materials and methods

Swamp milkweed is a perennial herb that occurs in wetlands throughout much of the United States (Woodson 1954). The three populations of swamp milkweed that we studied occurred in wet, abandoned pastures in the Shenandoah Valley of northern Virginia, USA. Two of these populations, located in Clarke and Frederick Counties, represent ssp. *incarnata*; the third, located in Fauquier County, represents ssp. *pulchra*. Voucher specimens from all three populations are housed in the University of Georgia Herbarium (Ga.). Most milkweeds are self-incompatible, but *A. incarnata* is unusual in that it is at least partially self-compatible (Kephart 1981; Wyatt and Broyles 1994; Ivey et al. 1998). Natural populations of swamp milkweed appear to be largely outcrossing, but there is variation among plants in self-fertility (Ivey et al. 1998). The basis of this variation is unclear (Ivey et al. 1998).

We have found that fruit-set is predicted by pollination success and that increased insect visitation is linked to increased pollination success in *A. incarnata* (Ivey et al. submitted). Furthermore, we have observed that the frequency of insect visitation increases with floral display size and that floral neighborhood size is strongly positively correlated with pollination success (unpublished data). These observations, combined with earlier observations of apparent variation in self-fertility among plants (Ivey et al. 1998), led us to predict that outcrossing rate should be influenced by the density and floral display size of neighboring plants. Pollinators are more likely to move between plants that are closely spaced or that have neighbors with large numbers of open flowers (e.g., Kunin 1993; Karron et al. 1995), leading to increased opportunities for outcrossing. Plants that are relatively isolated or surrounded by plants with few flowers are predicted to be more attractive to pollinators than their neighbors and, because pollinators may spend proportionately more time foraging on these plants, they are expected to experience increased levels of self-pollination and self-fertilization (Klinkhamer and de-Jong 1993).

Our choice of plants to sample was constrained by plant size, which may have introduced bias in our study. To increase the precision of our family outcrossing estimates, we selected plants that seemed likely to produce at least 25 fruits. Precision could not be increased by sampling multiple seeds within fruits because all seeds within milkweed fruits typically result from a single pollination event (Broyles and Wyatt 1990). Although swamp milkweed plants can produce hundreds of flowers within a season, only about 5–7% of those flowers will produce mature fruits (Wilbur 1976). Thus, the plants we sampled tended to be the larger plants in the populations. If plant size is correlated with outcrossing rate, then the plants we sampled may not represent the entire range of variation in outcrossing rate present in the population.

We collected an average of 23–30 fruits from each of 9–27 plants in three populations of swamp milkweed. One population (Clarke County) was sampled in both 1995 and 1996, whereas the other two populations (Frederick and Fauquier Counties) were sampled only in 1996. Two seeds from each fruit were germinated and grown to the seedling stage. Cotyledons and leaves from one of the two seedlings were then crushed, using a ceramic mortar and glass pestle, with a chilled extraction buffer modified from Broyles and Wyatt (1990) to include only 0.005% 2-mercaptoethanol. The extract was filtered through Miracloth and absorbed onto 8×3-mm paper wicks cut from Whatman 3MM chromatography paper. Sample wicks were stored at -70°C until electrophoresis was performed. Using the same procedure, we prepared extracts of maternal leaf tissue for

each plant used in the analysis and determined the maternal isozyme phenotype.

Three electrophoretic buffer systems were employed to resolve up to nine polymorphic allozyme loci. A continuous histidine-citrate system (pH 5.7: Stuber et al. 1977) was used to resolve aconitate hydratase (ACO: E.C. 4.2.1.3), isocitrate dehydrogenase (IDH: E.C. 1.1.1.42), and phosphoglucose isomerase (PGI: E.C. 5.3.1.9). A continuous tris-citrate system (electrode buffer: 0.3435 M Tris, 0.0715 M citric acid monohydrate, pH 8.0; gel buffer 6.68% dilution) was used to resolve fumarate hydratase (FUM: E.C. 4.2.1.2), leucine aminopeptidase (LAP: E.C. 3.4.-.-), phosphoglucomutase (PGM: E.C. 5.4.2.2), and fluorescent esterase (FLE: E.C. 3.1.1.-). A discontinuous buffer system with a lithium-borate electrode buffer and a tris-citrate gel buffer (pH 8.5: O'Malley et al., 1980) was used to resolve glutamate oxaloacetate transaminase (GOT: E.C. 2.6.1.1) and triosephosphate isomerase (TPI: E.C. 5.3.1.1). Protocols for resolving enzymes followed those described by Werth (1985).

We employed Ritland and Jain's (1981) maximum-likelihood, multilocus procedure to estimate the outcrossing rate of each family in all three populations, using joint estimations of outcrossing (t) and pollen allele frequencies (p). Several families from each population did not converge on an estimate between 0 and 1 when we used the Newton-Raphson estimation procedure of this model. We therefore re-analyzed the data using the estimation-maximization procedure; with this technique we found that all families converged on outcrossing estimates between 0 and 1. We refer to these estimates as "the maximum-likelihood outcrossing rate, t_m." We generated standard errors of these estimates via 100 bootstraps of the data. In addition, we analyzed our data using a program described by Cruzan et al. (1994) to estimate family level outcrossing rate in all three populations. The values we report for outcrossing rate using this model are the mean and standard error of ten jackknifed replicates of these estimates. To be consistent with the terminology employed by Cruzan et al. (1994), we refer to these estimates as "the apparent outcrossing rate, t_a." We also determined the family level outcrossing rate by calculating the proportion of progeny within each array that possessed an allele not represented in the maternal genotype. This third estimate, which, like Cruzan et al. (1994), we call "the detectable outcrossing rate, t_d" (note that others have referred to this parameter as "the apparent outcrossing rate": e.g., Morgan and Barrett 1990), will be an underestimate of outcrossing since some matings will involve pairs of plants with similar genotypes.

We also calculated an index of floral neighborhood size for each plant. During the week of peak flowering for each population, we recorded the distance to the three nearest neighbors for each plant and counted the number of umbels with open flowers on each of them. We calculated the floral neighborhood of each plant as

$$N = \sum_{i=1}^{3} \frac{U_i}{D_i^2}, \quad (1)$$

where U_i is the number of umbels on the nearest-neighbor i, and D_i is the distance of the nearest-neighbor i from the focal plant. Plants with large values of N occur in areas where floral display sizes are relatively large and/or where stands are dense, whereas plants with small neighborhood values are relatively isolated and/or have neighbors with small floral displays.

We calculated Spearman rank-sum correlation coefficients (r_s) between outcrossing rate and floral neighborhood size within each population. A positive relationship between floral neighborhood and outcrossing rate is expected if pollinator movements among plants are determined by the density of flowers on neighboring plants and if fruit-set is directly related to pollination.

Variation in family estimates of outcrossing rate could potentially arise from statistical variation associated with the estimation procedure employed. We estimated the amount of variance expected due to chance from both procedures using a Monte-Carlo simulation. We simulated the original data from each population by creating 40 data sets that were identical to the original data with respect to number of loci, number of families, and average size of families. Each data set also was assigned the same pollen allele frequencies, ovule allele frequencies, Wright's (1922) fixation indices, correlations for outcrossing and outcrossed paternal success among progeny pairs (r_t and r_p, respectively: Ritland 1989), and a population outcrossing rate that was estimated from the original data using Ritland and Jain's (1981) population-level estimation procedure. Thus, the simulated data were identical to the original data except that the outcrossing rate was known to be identical among all families. This is similar to the procedure employed by Morgan and Barrett (1990) to estimate statistical variation in t_m.

Family outcrossing rates were estimated in each of the simulated data sets using Ritland and Jain's (1981) program to generate estimates of t_m and the procedure of Cruzan et al. (1994) to estimate t_a. We also calculated t_d for each data set. The variance of the family outcrossing estimates in each simulated data set was then calculated. To determine the amount of variation expected by chance from using these estimation procedures, we calculated the mean of the variances from the simulated data sets. We also computed the 95th percentile of the variances from the simulated data sets to determine the upper boundaries of the amount of variation expected due to chance.

In addition, we used the simulated data sets to estimate the frequency with which spurious correlations between outcrossing rate and neighborhood floral density would be expected by chance. For each data set, we calculated r_s between the observed values for floral neighborhood size and family outcrossing rate estimates from the simulated data sets. Since the variation in these outcrossing estimates is expected to be zero, any significant correlation will be solely due to chance.

Results

Estimated allele frequencies revealed a large amount of polymorphism in all populations sampled (Table 1). We were unable to score *Aco-1* reliably in the Clarke County 1995 or Fauquier County populations, and *Fum-1* resolved well enough to be scored confidently only in the Clarke County 1995 population. Similarly, we did not score *Tpi-2* in the Fauquier County population because of poor resolution. The remaining loci, however, were well-resolved.

We found a wide range in values for family level outcrossing rate in all three populations (Table 2). For example, in the Clarke County 1995 population, outcrossing estimates ranged from 0.53 to 1.00; in the same population, estimates ranged from 0.25 to 1.00 in the following year. In the Frederick County population, family outcrossing estimates ranged from 0.32 to 1.00. The Fauquier County population had the broadest range in outcrossing estimates, from a low of 0.00 to a high of 1.00.

Estimates of t_m were higher than t_a for 58 of the 70 families studied, whereas estimates of t_a were higher than t_m for only eight families (Table 2). Cruzan et al. (1994) also reported that estimates of t_m tended to exceed those of t_a. Although the estimates of t_m and t_a were comparable for many of the families, 14 of the 70 families differed by > 0.2 between estimates of t_m and t_a. Estimates of t_a and t_d were significantly positively correlated in all populations, but estimates of t_a and t_m were significantly correlated only in the Frederick County population (Table 3). We estimated

Table 1 Estimated pollen and ovule allele frequencies of natural populations of *A. incarnata* located in northern Virginia. Standard errors (indicated in parentheses) were estimated by 100 bootstraps of the data. These estimated allele frequencies were used to generate the simulated data sets (see Materials and methods)

Locus	Allele	Clarke 1995		Clarke 1996		Frederick 1996		Fauquier 1996	
		Pollen	Ovule	Pollen	Ovule	Pollen	Ovule	Pollen	Ovule
Aco-2	*a*	–	–	–	–	0.055 (0.036)	0.034 (0.002)	–	–
	b	–	–	0.905 (0.095)	0.907 (0.032)	0.850 (0.048)	0.793 (0.067)	–	–
	c	–	–	0.095 (0.013)	0.093 (0.032)	0.095 (0.018)	0.172 (0.066)	–	–
Fle-1	*a*	0.365 (0.027)	0.450 (0.080)	0.244 (0.029)	0.481 (0.067)	0.181 (0.026)	0.207 (0.060)	0.757 (0.082)	0.722 (0.123)
	b	0.635 (0.027)	0.550 (0.080)	0.756 (0.029)	0.519 (0.067)	0.813 (0.029)	0.759 (0.061)	0.243 (0.082)	0.278 (0.123)
	c	–	–	–	–	0.006 (0.005)	0.034 (0.002)	–	–
Fum-1	*a*	0.014 (0.008)	0.024 (0.002)	–	–	–	–	–	–
	b	0.078 (0.018)	0.073 (0.042)	–	–	–	–	–	–
	c	0.908 (0.019)	0.902 (0.043)	–	–	–	–	–	–
Got-2	*a*	0.172 (0.028)	0.125 (0.047)	0.206 (0.019)	0.167 (0.043)	0.563 (0.043)	0.500 (0.112)	0.873 (0.054)	0.667 (0.102)
	b	0.827 (0.028)	0.850 (0.050)	0.793 (0.019)	0.815 (0.044)	0.418 (0.047)	0.464 (0.105)	0.127 (0.054)	0.333 (0.102)
	c	0.002 (0.001)	0.025 (0.018)	0.001 (0.000)	0.019 (0.013)	0.019 (0.010)	0.036 (0.028)	–	–
Idh-1	a	0.217 (0.031)	0.200 (0.068)	0.173 (0.018)	0.222 (0.062)	0.078 (0.014)	0.036 (0.017)	0.005 (0.005)	0.053 (0.006)
	b	0.506 (0.037)	0.625 (0.070)	0.578 (0.021)	0.574 (0.065)	0.527 (0.043)	0.464 (0.060)	0.823 (0.030)	0.632 (0.137)
	c	–	–	–	–	–	–	0.167 (0.028)	0.263 (0.121)
	d	0.277 (0.035)	0.175 (0.053)	0.248 (0.023)	0.204 (0.056)	0.395 (0.037)	0.500 (0.062)	0.005 (0.004)	0.053 (0.034)
Lap-1	*a*	0.991 (0.004)	0.975 (0.016)	0.990 (0.003)	0.964 (0.003)	0.796 (0.036)	0.759 (0.081)	0.306 (0.075)	0.389 (0.097)
	b	0.009 (0.004)	0.025 (0.018)	0.009 (0.003)	0.018 (0.002)	0.179 (0.033)	0.207 (0.081)	0.538 (0.090)	0.556 (0.096)
	c	–	–	0.001 (0.001)	0.018 (0.002)	0.025 (0.008)	0.034 (0.002)	0.155 (0.050)	0.056 (0.038)
Pgi-1	*a*	0.032 (0.009)	0.049 (0.027)	0.081 (0.014)	0.074 (0.039)	0.052 (0.011)	0.034 (0.002)	0.230 (0.032)	0.111 (0.062)
	b	0.956 (0.011)	0.927 (0.027)	0.919 (0.014)	0.926 (0.039)	0.948 (0.011)	0.966 (0.002)	0.765 (0.032)	0.833 (0.069)
	c	0.012 (0.007)	0.024 (0.002)	–	–	–	–	0.005 (0.000)	0.056 (0.046)
Pgm-1	*a*	0.285 (0.036)	0.425 (0.088)	0.324 (0.027)	0.315 (0.071)	0.012 (0.006)	0.033 (0.002)	0.050 (0.026)	0.053 (0.005)
	b	0.715 (0.036)	0.575 (0.088)	0.676 (0.027)	0.685 (0.071)	0.970 (0.012)	0.933 (0.004)	0.900 (0.052)	0.895 (0.044)
	c	–	–	–	–	0.018 (0.009)	0.033 (0.002)	0.050 (0.028)	0.053 (0.044)
Tpi-2	*a*	0.217 (0.031)	0.200 (0.068)	0.303 (0.025)	0.259 (0.062)	0.175 (0.033)	0.214 (0.057)	–	–
	b	0.506 (0.037)	0.625 (0.070)	0.697 (0.025)	0.741 (0.062)	0.825 (0.033)	0.786 (0.057)	–	–

Table 2 Family size (n), maximum-likelihood outcrossing estimate (t_m) and standard error, apparent outcrossing estimate (t_a) and standard error, and detectable outcrossing rate (t_d) for families in three populations of *A. incarnata* in northern Virginia. Families numbered 1–15 in the Clarke County population represent the same individuals in both years

Population	n	t_m	SE	t_a	SE	t_d
Clarke 1995						
1	36	0.94	0.05	0.864	0.000	0.857
2	35	1.00	0.00	0.945	0.003	0.941
3	45	1.00	0.00	0.932	0.000	0.932
4	23	0.97	0.08	0.930	0.000	0.818
5	23	1.00	0.00	0.876	0.004	0.864
6	31	0.95	0.04	0.957	0.002	0.933
7	27	0.78	0.09	0.911	0.002	0.731
8	25	0.92	0.16	0.610	0.005	0.583
9	34	1.00	0.00	0.737	0.004	0.394
10	15	0.64	0.15	0.874	0.005	0.643
11	32	0.75	0.08	0.882	0.002	0.710
12	25	1.00	0.00	0.889	0.003	0.667
13	37	1.00	0.01	0.919	0.002	0.722
14	20	0.76	0.11	0.702	0.008	0.684
15	67	0.88	0.05	0.744	0.002	0.742
16	23	1.00	0.00	0.959	0.005	0.955
17	35	0.99	0.03	0.703	0.004	0.676
18	15	1.00	0.05	0.877	0.009	0.857
19	25	0.95	0.05	0.923	0.000	0.917
20	31	0.83	0.18	0.847	0.003	0.533
Clarke 1996						
1	51	0.97	0.03	0.904	0.003	0.900
2	35	0.96	0.05	0.719	0.005	0.735
3	24	1.00	0.00	0.963	0.004	0.957
4	54	0.76	0.13	0.800	0.002	0.528
5	23	0.96	0.05	0.869	0.005	0.864
6	21	0.91	0.06	0.900	0.000	0.900
7	26	0.73	0.18	0.436	0.006	0.440
8	41	1.00	0.00	0.731	0.003	0.725
9	35	1.00	0.00	0.732	0.005	0.706
10	17	1.00	0.00	0.887	0.006	0.875
11	17	0.90	0.08	0.884	0.006	0.875
12	12	0.92	0.07	0.976	0.003	0.909
13	13	0.35	0.16	0.778	0.011	0.250
14	14	0.63	0.15	0.585	0.013	0.615
15	37	0.99	0.03	0.778	0.004	0.750
16	26	0.98	0.07	0.939	0.002	0.800
17	22	0.99	0.05	0.870	0.004	0.571
18	16	1.00	0.00	0.880	0.009	0.867
19	19	0.86	0.11	0.793	0.007	0.833
20	14	1.00	0.00	0.980	0.002	0.923
21	17	1.00	0.00	0.881	0.006	0.875
22	18	1.00	0.00	0.457	0.010	0.412
23	53	1.00	0.00	0.816	0.003	0.808
24	16	1.00	0.00	1.000	0.000	1.000
25	44	0.98	0.03	0.933	0.000	0.930
26	23	0.69	0.16	0.572	0.007	0.545
27	27	0.90	0.06	0.892	0.005	0.885
Frederick						
1	20	0.86	0.14	0.684	0.000	0.684
2	24	0.57	0.15	0.452	0.007	0.435
3	33	0.97	0.04	0.944	0.004	0.938
4	30	1.00	0.01	0.766	0.005	0.759
5	30	0.95	0.09	0.697	0.005	0.690
6	22	1.00	0.00	1.000	0.000	1.000
7	33	0.76	0.09	0.694	0.004	0.688
8	11	0.99	0.11	0.820	0.013	0.800
9	30	0.96	0.05	0.869	0.005	0.862
10	15	1.00	0.08	0.864	0.007	0.857
11	33	0.77	0.11	0.638	0.005	0.625
12	32	0.88	0.08	0.817	0.004	0.839
13	20	0.38	0.14	0.353	0.008	0.316
14	33	0.72	0.09	0.697	0.005	0.688
Fauquier						
1	24	1.00	0.00	1.000	0.000	1.000
2	17	1.00	0.00	1.000	0.000	1.000
3	23	0.81	0.14	0.615	0.008	0.636
4	33	0.60	0.14	0.405	0.003	0.406
5	14	1.00	0.00	0.003	0.002	0.000
6	32	0.67	0.12	0.571	0.005	0.548
7	32	0.97	0.03	0.971	0.003	0.968
8	14	0.68	0.18	0.578	0.013	0.538
9	18	0.93	0.07	0.721	0.008	0.471

Table 3 Spearman rank-sum correlation coefficients among estimates for family level outcrossing rate and floral neighborhood size in three populations of *A. incarnata* from northern Virginia. See text for description of neighborhood calculation. See Table 2 for definitions of abbreviated parameters

Population	t_a	t_d	Floral neighborhood
Clarke 1995			
t_m	0.40	0.39	0.34
t_a		0.76****	0.46*
t_d			0.16
Clarke 1996			
t_m	0.31	0.36	−0.15
t_a		0.88****	0.03
t_d			0.02
Frederick			
t_m	0.83***	0.83***	−0.24
t_a		0.99****	0.06
t_d			−0.08
Fauquier			
t_m	0.53	0.41	−0.37
t_a		0.88**	−0.31
t_d			−0.51

* $P < 0.05$; ** $P < 0.01$; *** $P < 0.001$; **** $P < 0.0001$

the outcrossing rate in 15 of the same individuals from the Clarke County population in both 1995 and 1996. There was no relationship between years in the estimates for t_m ($r_s = 0.08$, $P = 0.78$) or t_a ($r_s = 0.28$, $P = 0.30$) for these individuals. The neighborhood index (N), however, was strongly correlated between years ($r_s = 0.84$, $P = 0.0001$) for these individuals. N was significantly positively correlated with t_a for the Clarke County population in 1995. There was no relationship between floral neighborhood and outcrossing rate in any of the other populations.

The population-level estimates of mating-system parameters revealed that all populations were largely outcrossing (Table 4), which agrees with earlier

Table 4 Mating-system parameters estimated from natural populations of *A. incarnata* located in northern Virginia. These parameter estimates were used to generate simulated data sets (see Materials and methods). Estimates of the standard error (in parentheses) were obtained from 100 bootstraps of the data. F = inbreeding coefficient, t_m = multilocus outcrossing rate, r_t = correlation between progeny pairs for outcrossing, r_p = correlation between progeny pairs for outcrossed paternity. * = parameter estimates that were assigned values of zero for the simulations because estimates were not significantly different from zero

Population	F	t_m	r_t	r_p
Clarke 1995	− 0.030*	0.960	− 0.004*	0.196
	(0.096)	(0.024)	(0.198)	(0.054)
Clarke 1996	− 0.035*	0.961	0.029*	0.121
	(0.087)	(0.018)	(0.164)	(0.054)
Frederick	0.003*	0.902	0.133*	0.092
	− (0.086)	(0.033)	(0.088)	(0.027)
Fauquier	0.088	0.942	0.215*	0.482
	(0.167)	(0.044)	(0.540)	(0.180)

observations in these populations (Ivey et al. 1998). Estimates of Wright's (1922) fixation index and of correlations for outcrossing between progeny pairs were not significantly different from zero in any of the populations; thus, they were assigned a value of zero for the simulated data sets. The one exception was the Fauquier County population; the estimate for the fixation index was marginally significant for this population, so we retained this value for the simulations. Estimates of correlations for outcrossed paternity between progeny pairs were significantly positive in all populations (Table 4). This indicates variation among plants in paternal success, which has been reported previously (Ivey et al. 1998).

The Monte-Carlo simulations revealed that a large amount of variation in family outcrossing rate estimates, using any of the estimation procedures, is expected simply by chance alone (Table 5). For the Frederick and Fauquier County populations, the observed variance in family outcrossing estimates using any of the three procedures was greater than the expected variance determined from the simulations (Table 5). This was also true for the Clarke County population in 1996, except when the t_m procedure was used. For the Clarke County population in 1995, only when the t_d procedure was used did the observed variance in outcrossing exceed the expected variance. Nonetheless, the observed variance exceeded the 95th percentile for the expected variance in outcrossing in only three cases: in the Fauquier County population when the t_a and t_d procedures were used, and in the Clarke County population in 1996 when the t_d procedure was used (Table 5). Thus, for the Fauquier County population, based on estimates of t_a and t_d, and for the Clarke County population in 1996, based on estimates of t_d, it appears that there is significant variation among plants in outcrossing rate. For the remaining cases, all of the variation observed for family outcrossing rate could be explained by the statistical variation associated with using these procedures.

Table 5 Results from a Monte-Carlo simulation estimating the expected variance in family outcrossing estimates using three estimation procedures. t_m = multilocus maximum-likelihood estimation, t_a = apparent outcrossing estimation, t_d = detectable outcrossing rate. See text for details

Population	Observed variance, t_m	Expected variance, t_m	95th Percentile expected variance, t_m
Clarke 1995	0.012	0.018	0.042
Clarke 1996	0.023	0.014	0.039
Frederick	0.035	0.020	0.055
Fauquier	0.027	0.016	0.064
	Observed variance, t_a	Expected variance, t_a	95th Percentile expected variance, t_a
Clarke 1995	0.010	0.021	0.038
Clarke 1996	0.023	0.022	0.032
Frederick	0.031	0.022	0.042
Fauquier	0.106	0.017	0.042
	Observed variance, t_d	Expected variance, t_d	95th Percentile expected variance, t_d
Clarke 1995	0.023	0.021	0.037
Clarke 1996	0.036	0.023	0.034
Frederick	0.034	0.022	0.044
Fauquier	0.109	0.017	0.045

Table 6 Spearman's rank-sum correlations of observed floral neighborhood size with family outcrossing estimates from 40 simulated data sets using multilocus maximum-likelihood estimation (t_m), apparent outcrossing estimation (t_a), or detectable outcrossing rate (t_d). Only statistically significant correlations are shown. See text for details

Population	t_m	t_a	t_d
Clarke, 1995	− 0.54**	None	None
	− 0.47*		
	0.46*		
Clarke, 1996	0.46**	− 0.48**	− 0.51**
	0.42*	− 0.55**	− 0.53**
	− 0.39*	− 0.40*	− 0.39*
		− 0.46*	− 0.45*
		− 0.49**	− 0.49**
Frederick	None	0.53*	0.60*
		0.81***	0.72**
		− 0.53*	0.53*
		0.56*	
Fauquier	− 0.71*	− 0.91***	− 0.81**
		0.67*	0.67*
		0.88**	0.90***
		− 0.74*	− 0.71*
			− 0.70*

* $P < 0.05$; ** $P < 0.01$; *** $P < 0.001$

For each population, the 95th percentile of expected variance for t_m was larger than that for t_a (Table 5).

Similarly, for all populations except Fauquier County, the observed variance in family estimates was greater for the t_m estimation procedure than for the t_a procedure.

The correlations between observed floral neighborhood size and outcrossing estimates from the simulated data sets revealed that spurious correlations can occur using any of the procedures (Table 6). Overall, nearly 7% of the correlations we calculated were statistically significant. We found that, overall, 4.4% of the simulated data sets estimating t_m, 7.5% of the data sets estimating t_a, and 8.1% of the data sets estimating t_d were significantly correlated with observed floral neighborhood size. In three of the four populations, significant correlations with floral neighborhood size were observed in 10% of the data sets estimating t_a. In two populations, significant correlations with floral neighborhood size were observed in 7.5% of the simulated data sets estimating t_m. In two populations, floral display size was significantly correlated with 12.5% of the data sets estimating t_d.

Discussion

We found that all three of the procedures used to estimate family outcrossing rates had considerable statistical variation associated with them. This was reflected in the large expected variances defined through Monte-Carlo simulations. To understand patterns of mating system evolution, accurate estimates of family outcrossing from nature are essential. It appears, however, that the estimates generated by the statistical tools currently available are not very accurate (Morgan and Barrett 1990; Cruzan et al. 1994). Nonetheless, numerous studies have relied on such potentially suspect estimates of family outcrossing for interpreting the biology of their systems (e.g., Shaw and Allard 1982; Perry and Dancik 1986; Morgan and Barrett 1990; Perry and Knowles 1990; El-Kassaby et al. 1993; Cruzan et al. 1994; Burgess et al. 1996; El-Kassaby and Jaquish 1996).

As reported by Cruzan et al. (1994), we found that the procedure used to estimate t_a tends to generate lower estimates of outcrossing than the procedure used to estimate t_m. Even so, we found that the range of expected variance in outcrossing defined by the simulations is smaller when estimating t_a than when estimating t_m. Furthermore, the estimates of standard error around t_a were smaller than those around estimates of t_m in the original data. These observations might suggest that estimates of outcrossing using the t_a procedure are more accurate and precise than those generated using the t_m procedure. We observed a considerable range, however, in expected variance for outcrossing using both procedures. Any accuracy gained through use of the t_a procedure appears to be slight. Moreover, the differences in standard errors between the two procedures is probably just a computational artifact (Cruzan et al. 1994).

Milkweeds characteristically have low levels of fruitset and, whereas *A. incarnata* has somewhat higher levels (Wilbur 1976), we still were not able to sample more than an average of 30 progeny per maternal plant from any population. Thus, one might suspect that the wide variation we observed in our estimates of outcrossing reflects sampling error because relatively few progeny were sampled. Indeed, the breadth of the confidence intervals for the expected variance narrowed as the average family size increased. Furthermore, standard errors of individual estimates of t_a decreased with sample size ($r_s = -0.41$, $P = 0.0004$), but there was no statistically significant relationship between the standard errors of individual estimates of t_m and sample size ($r_s = -0.15$, $P = 0.22$). Nonetheless, previous studies have revealed that increasing the number of progeny sampled per family only slightly narrows the confidence interval for expected variance in outcrossing rate (Morgan and Barrett 1990). Furthermore, the sample sizes used in our data are larger than those reported in some earlier studies estimating open-pollinated family outcrossing rates (Ritland and Ganders 1985; Morgan and Barrett 1990; Cruzan et al. 1994).

It is possible that the poor accuracy in outcrossing estimates that we observed was related to violations of model assumptions due to the structure of our data. For example, similarity between the maternal genotype and the pollen pool allele frequencies can diminish the power to detect outcross events in her progeny (Shaw et al. 1981; Morgan and Barrett 1990). We found a significant relationship between the probability of outcrossing being undetected (Shaw et al. 1981) and t_d in only one population (Clarke 1995: $r_s = -0.65$, Clarke 1995 refers to a population studied in the article, not an article under reference $P = 0.002$). This suggests that, for this population only, variation in the maternal genotype may have contributed to the variation in our estimates of outcrossing. Correlations among pollen genotypes in the outcrossed pollen pool may have contributed to inaccuracies in estimates of t_m, but this problem should not have affected the other estimation procedures (Cruzan et al. 1994) and we observed inaccuracy across all procedures. Indeed, there was little relationship among the different estimates of outcrossing within populations. We also found no relationship between years for estimates collected from the same plants. Similar observations have been interpreted by others as evidence for environmental influences on outcrossing (e.g., El-Kassaby et al. 1993), but it seems likely that chance statistical variation contributed significantly to this pattern as well.

Some workers have resolved the problem of estimating outcrossing rates in natural populations by placing plants of predetermined genotypes into experimental arrays (e.g., Motten and Antonovics 1992; Karron et al. 1995). Using such techniques, the statistical variation

associated with family outcrossing estimates can be significantly reduced, since outcrossing rate can be estimated without using these models. Because of the constraints of an experimental system, however, it is more difficult to extrapolate conclusions from these results to a natural setting.

Although estimates of family outcrossing from nature may be somewhat inaccurate, it may be possible to interpret the basis of the observed variation by investigating the relationship between the estimates and environmental or other parameters (e.g., El-Kassaby et al. 1993; Cruzan et al. 1994; Burgess et al. 1996; El-Kassaby and Jaquish 1996). Nonetheless, for our data, we found that a statistically significant relationship could be expected to arise by chance on an average of about once every 15 times such a correlation is calculated. When t_d was the method employed, that figure rose to 1 in 8 times for some populations. Thus, it may be difficult to interpret the significance of variation in outcrossing rate estimates using observed variation in environmental factors unless the frequency of expected spurious correlations is understood.

We failed to detect significant variation among plants in estimates of t_m or t_a except in one population. Since these plants varied considerably in their floral neighborhood size and, as a consequence, probably varied in success at pollination and in the frequency of self-pollination, it is possible that post-pollination factors (e.g., selective fruit abortion: Bookman 1984) governed the success of self-pollinated ovaries. Variation in post-pollination success has been previously documented in milkweeds and is sometimes invoked to explain low levels of fruit-set in milkweeds (Wyatt and Broyles 1994).

Why did we observe such a high variance among plants for estimates of t_a and t_d in the Fauquier County population? The wet, abandoned pasture habitat of all three populations appeared similar; thus, this seems unlikely to have played a role. Taxonomic differences may have contributed to this pattern; the Fauquier County population represents ssp. *pulchra*, which appears to have lower outcrossing rates than populations of ssp. *incarnata* (Ivey et al. 1998). The higher variance may have been an analytical artifact; perhaps we observed a higher variance in these estimates due to some factor not controlled in our simulations. Alternatively, two of the individual estimates of t_a from this population were considerably lower than the corresponding estimates of t_m (families 4 and 5), which increased the variance in t_a relative to that of t_m. Furthermore, one of these maternal plants (5) had an isozyme genotype similar to the outcross pollen-pool allele frequencies; thus, it may have been more difficult to detect true outcross events in her progeny. Overall, however, there was no significant relationship between the probability of detecting outcrossing events (Shaw et al. 1981) and estimates of t_a ($r_s = -0.18$, $P = 0.64$) or t_d ($r_s = -0.35$, $P = 0.35$) in this population.

Although it seems plausible to suspect that environmental or genetic factors influence variation in the outcrossing rates of individual mixed-mating, open-pollinated plants, the tools currently available appear to be unsatisfactory for describing that relationship. Experimental populations remain the most advantageous settings for investigations related to such questions. The development of functional statistical tools for such investigations in natural settings will provide fertile grounds for future research.

Acknowledgments We thank D. E. Carr, P. M. Dixon, J. L. Hamrick, M. T. Morgan, and R. R. Sharitz for providing comments on earlier versions of this paper. Laboratory assistance was provided by G. Gaston. Support for this research was provided in part by grants to C.T.I. from the American Museum of Natural History's Theodore Roosevelt Memorial Fund, Sigma Xi, the University of Georgia Botany Department's Palfrey Fund, and the University of Virginia's Blandy Experimental Farm.

References

Barrett SCH, Harder LD (1996) Ecology and evolution of plant mating. Trends Ecol Evol 11:73–79

Bookman SS (1984) Evidence for selective fruit production in *Asclepias*. Evolution 38:72–86

Brown AHD, Barrett SCH, Moran GF (1985) Mating-system estimation in forest trees: models, methods and meanings. In: Gregorius HR (ed) Population genetics in forestry, Springer, Berlin Heidelberg New York, pp 32–49

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

Burgess IP, Williams ER, Bell JC, Harwood CE, Owen JV (1996) The effect of outcrossing rate on the growth of selected families of *Eucalyptus grandis*. Silvae Genet 45:97–100

Clegg M (1980) Measuring plant mating systems. BioScience 30:814–818

Cruzan MB (1998) Genetic markers in plant evolutionary ecology. Ecology 79:400–412

Cruzan MB, Arnold ML (1994) Assortative mating and natural selection in an *Iris* hybrid zone. Evolution 48:1946–1958

Cruzan MB, Hamrick JL, Arnold ML, Bennett BD (1994) Mating-system variation in hybridizing irises: effects of phenology and floral densities on family outcrossing rates. Heredity 72:95–105

El-Kassaby YA, Jaquish B (1996) Population density and mating pattern in western larch. J Hered 87:438–443

El-Kassaby YA, Meagher MD, Davidson R (1993) Temporal variation in the outcrossing rate in a natural stand of western white pine. Silvae Genet 42:131–135

Ellstrand NC, Torres AM, Levin DA (1978) Density and the rate of apparent outcrossing in *Helianthus annuus* (Asteraceae). Systematic Bot 3:403–407

Fyfe JL, Bailey NTJ (1951) Plant breeding studies in leguminous forage crops. I. Natural crossbreeding in winter beans. J Agric Sci 41:371–378

Ivey CT, Lipow SR, Wyatt R (1998) Mating systems and interfertility of swamp milkweed (*Asclepias incarnata* ssp. *incarnata* and ssp. *pulchra*). Heredity (in press)

Jarne P, Charlesworth D (1993) The evolution of the selfing rate in functionally hermaphrodite plants and animals. Annu Rev Ecol Syst 24:441–466

Karron JD, Thumser NN, Tucker R, Hessenauer AJ (1995) The influence of population density on outcrossing rates in *Mimulus ringens*. Heredity 75:175–180

Kephart SR (1981) Breeding systems in *Asclepias incarnata* L., *A. syriaca* L., and *A. verticillata* L. Am J Bot 68 : 226–232
Klinkhamer PGL, deJong TJ (1993) Attractiveness to pollinators: a plant's dilemma. Oikos 66 : 180–184
Kunin WE (1993) Sex and the single mustard: population density and pollinator behavior effects on seed-set. Ecology 74 : 2145–2160
Morgan MT, Barrett SCH (1990) Outcrossing rates and correlated mating within a population of *Eichhornia paniculata* (Pontederiaceae). Heredity 64 : 271–280
Motten AF, Antonovics J (1992) Determinants of outcrossing rate in a predominantly self-fertilizing weed, *Datura stramonium* (Solanaceae). Am J Bot 79 : 419–427
O'Malley D, Wheeler NC, Guries NP (1980) A manual for starch-gel electrophoresis. Staff Paper Series ♯11, Coll Agric Life Sci, Univ Wisconsin, Madison
Perry DJ, Dancik BP (1986) Mating system dynamics of lodgepole pine in Alberta, Canada. Silvae Genet 35 : 190–195
Perry DJ, Knowles P (1990) Evidence of high self-fertilization in natural populations of eastern white cedar (*Thuja occidentalis*). Can J Bot 68 : 663–668
Ritland K (1989) Correlated matings in the partial selfer *Mimulus guttatus*. Evolution 43 : 848–859
Ritland K, Ganders FR (1985) Variation in the mating system of *Bidens menziesii* (Asteraceae) in relation to population substructure. Heredity 55 : 235–244
Ritland K, Jain S (1981) A model for the estimation of outcrossing rate and gene frequencies using *n* independent loci. Heredity 47 : 35–52
Schoen DJ, Clegg MT (1984) Estimation of mating system parameters when outcrossing events are correlated. Proc Natl Acad Sci USA 81 : 5258–5262
Shaw DV, Allard RW (1982) Estimation of outcrossing rates in Douglas-fir using isozyme markers. Theor Appl Genet 62 : 113–120
Shaw DV, Kahler AL, Allard AW (1981) A multilocus estimator of mating-system parameters in plant populations. Proc Natl Acad Sci USA 78 : 1298–1302
Smyth CA, Hamrick JL (1984) Variation in estimates of outcrossing in musk thistle populations. J Hered 75 : 303–307
Stuber CW, Goodman MM, Johnson FM (1977) Genetic control and racial variation of β-glucosidase isozymes in maize (*Zea mays* L.). Biochem Genet 15 : 383–394
Watkins L, Levin DA (1990) Outcrossing rates as related to plant density in *Phlox drummondii*. Heredity 65 : 81–90
Werth CR (1985) Implementing an isozyme laboratory at a field station. Virginia J Sci 36 : 53–76
Wilbur HM (1976) Life history evolution in seven milkweeds of the genus *Asclepias*. J Ecol 64 : 223–240
Woodson RE (1954) The North American species of *Asclepias* L. Ann Mo Bot Gard 41 : 1–211
Wright S (1922) Coefficients of inbreeding and relationship. Am Nat 56 : 330–338
Wyatt R, Broyles SB (1994) Ecology and evolution of reproduction in milkweeds. Annu Rev Ecol Syst 25 : 423–441