Theor Appl Genet (1989) 78:480–488

Variation in outcrossing rate and population genetic structure of *Clarkia tembloriensis* (Onagraceae)

T. P. Holtsford*, ** and N. C. Ellstrand

Department of Botany and Plant Sciences, University of California, Riverside, CA 92521-0124, USA

Received February 22, 1989; Accepted April 12, 1989
Communicated by P. M. A. Tigerstedt

Summary. Outcrossing rate estimates for eight accessions of *Clarkia tembloriensis* indicate that this annual plant species has a wide interpopulational range of outcrossing rate ($\hat{t}=0.03-0.87$). Populations' t estimates were significantly correlated with observed heterozygosity and mean number of alleles per locus. Estimated fixation indices, $\hat{f}$, for most populations were very close to their expected values, F_{eq}, for a given $\hat{t}$. Nei's gene diversity statistics showed that the group of outcrossing populations have more total genetic variation and less differentiation among populations than does the group of selfing populations. These results indicate that the breeding system of *C. tembloriensis* has had a strong influence on the amount and distribution of genetic variation within and among its populations.

Key words: Outcrossing rate – Population structure – Inbreeding coefficients – Gene diversity – Isozymes

Introduction

The breeding system has long been thought to be a major determinant of population genetic structure (Wright 1921). Outbreeding populations are expected to have greater alleleic diversity, higher levels of heterozygosity, and show less differentiation among populations than self-fertilizing populations (Wright 1921; Allard et al. 1968). Most of the empirical data concerning this issue have come from studies of flowering plants. Population structure statistics of species with different breeding systems have been compared and the predicted patterns generally hold (Hamrick et al. 1979; Loveless and Hamrick 1984). However, very few studies have examined the relationship between outcrossing rate and population genetic structure within single species with wide ranges of outcrossing rates (Schoen 1982b; Glover and Barrett 1987). Intraspecific comparisons are desirable because interpopulation variation in the genetic background of the populations being studied is minimized. The more genetically similar the populations being compared, the more confidence one may have that the estimated values of the population structure statistics are the result of the breeding system and not of other unmeasured genetic or historical factors.

In this paper we present data on the relationship between outcrossing rate and population genetic structure in *Clarkia tembloriensis* Vasek (Onagraceae). This annual herb grows on grassy hillsides of the Temblor range to the west, and in the Tejon hills to the southeast of the San Joaquin Valley in central California. Vasek and Harding (1976) estimated outcrossing rates of four populations of *C. tembloriensis,* using single-locus petal color and shape polymorphisms. Their estimates ranged from $\hat{t}=0.08$ in a small-flowered non-dichogamous population to $\hat{t}=0.83$ in a large-flowered protandrous population. In the current study we use polymorphisms at isozyme-encoding loci as the genetic markers for estimating outcrossing rates and population structure statistics. These loci have the advantage that their phenotype is expressed codominantly in both maternal and in seedling tissue. Thus, progenies need not be grown to flowering to evaluate their phenotype nor is progeny testing required to detect heterozygotes. Also, while petal coloration might influence mating patterns (Brown and Clegg 1984; Schoen and Clegg 1985), a plant's genotype at isozyme loci is not likely to have any influence on outcrossing rate

* To whom correspondence should be addressed
** Present address: W. K. Kellogg Biological Station, Michigan State University, Hickory Corners, MI 49060, USA

Table 1. Electrophoresis systems summary

Gel buffer system	Current (mA)	Volts (max)	Run time (h)	Enzymes assayed
Histidine-citrate pH 6	25	150	6.5	PGM, 6PGD
Morpholine-citrate pH 6	30	150	6.5	SKDH, MDH
LiOH-boric acid pH 8.3	75	200	4.0	GDH, AAT, TPI
LiOH-boric acid pH 8.3	75–50	200–300	7.0	PGI

[a] A double-long (27 cm) gel was used to separate the allozymes of PGI

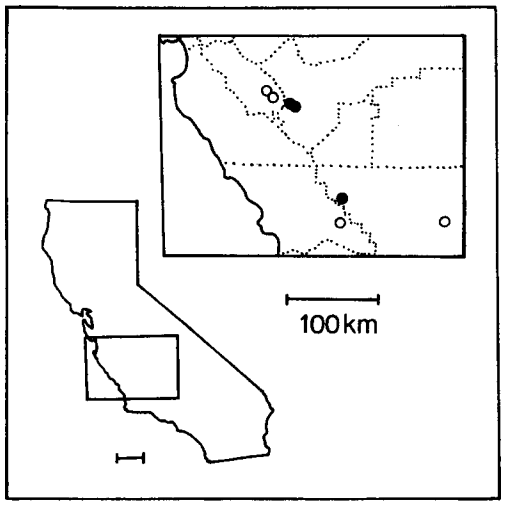

Fig. 1. Collection sites. *Dotted lines* are California county boundaries. *Closed circles* represent primarily selfing populations and *open circles* represent primarily outcrossing populations. From north to south, the population designations are: I1, I2, CC1, CC2, CR, MK (southwest), and HC (southeast)

by affecting pollinator behavior. We used isozyme polymorphisms to estimate single- and multi-locus measures of outcrossing rate and several population structure statistics for seven populations of *C. tembloriensis* which span the range of flower morphology and phenology of the species.

Materials and methods

Seven populations were sampled in 1986 and the MK population was sampled again in 1987 (Fig. 1). *Clarkia tembloriensis* is most often found on north-facing slopes in association with *Haplopappus linearifolius, Isomeris arborea* and several annual rangeland grasses. Fruits were collected from randomly-chosen plants with three exceptions. Two accessions (I2-86 and CC2-86) were collected along linear transects, which were placed in the approximate center of the populations laterally and ran from the populations' lower to their upper boundaries. All plants which were within 10 cm of the tape were sampled. One accession, HC-85, was provided by Dr. F. C. Vasek by sampling 20 haphazardly-chosen plants throughout the population.

Progenies of six to eight individuals were assembled by drawing one seed each from a separate fruit, in order to minimize the possibility of correlated mating events which would violate an assumption of the mixed mating model used to estimate outcrossing rates (Clegg 1980; Schoen and Clegg 1984, 1986; Schoen 1985). If a maternal plant produced less than six fruits the progeny were drawn randomly from all the seeds collected from that plant.

The procedures we used for extraction of plant material and starch gel electrophoresis were those of Gottlieb (1981a, 1984; Table 1). The following enzyme systems were assayed: phosphoglucomutase (PGM, two loci scored, two other putative loci unresolved), 6-phosphogluconate dehydrogenase (PGD, four loci), shikimate dehydrogenase (SKDH, one locus), malate dehydrogenase (MDH, four loci), glutamate dehydrogenase GDH, (one locus), aspartate/alpha-keto-glutarate transaminase (AAT, two loci), triose-phosphate isomerase (TPI, four loci), and phosphogluconate isomerase (PGI, three loci). Loci were numbered sequentially with the most anodally-migrating locus designated as locus 1. The formal genetic analyses of the loci encoding PGI, TPI, and PGD in *Clarkia* have been published (Gottlieb 1984). Genetic interpretations of the banding patterns at the other putative loci could be made, because the number of loci encoding these isozymes and their inheritance in *Clarkia* are generally understood (Gottlieb 1981b, 1982, 1984; Soltis and Bloom 1986). Inspection of progeny groups during the course of this study revealed segregation at the putative loci which was consistent with a Mendelian interpretation. The marker loci are not likely to be tightly linked, since alleles from all pairs of electrophoretically detectable loci in *Clarkia* segregate independently (L. D. Gottlieb, personal communication). The independent segregation of alleles at the marker loci is an important assumption for the unbiased estimation of multi-locus outcrossing rates (Ritland and Jain 1981), sample variances of observed heterozygosities (Archie 1985), and sample variances of multilocus coancestry coefficients (Weir and Cockerham 1984).

Two loci, SKDH and PGM-4, in three accessions, CC1-86, MK-86, and I1-86, were assayed from leaf tissue of plants whose seeds were collected for outcrossing rate estimation. All other maternal genotypes were inferred from the genotypes of six to eight their progeny (Brown and Allard 1970; Schoen 1982b). If 30 complete maternal genotypes could not be constructed for a population because fewer than six seeds per maternal plant could be germinated, then one progeny was randomly chosen from the unrepresented maternal plant and its genotype added to the data set. Mean observed heterozygosities, $\bar{H}_o$, mean number of alleles per locus, $\bar{k}$, and the percent of loci which were polymorphic, PLP, were calculated for the set of maternal genotypes from each population. A locus was considered polymorphic if the frequency of the most common allele was ≤ 0.95. Pearson's and Spearman's rank correlation coefficients were calculated among $\hat{t}$, $\bar{k}$, $\bar{H}_o$, and PLP. The Shapiro-Wilk statistic was calculated to test the normality of the distributions of the residuals of these variables when regressed on each other (Shapiro and Wilk 1965).

Single- and multi-locus outcrossing rate estimates, $\hat{t}$, were calculated by the methods of Ritland and Jain (1981). We estimated t for every locus which was polymorphic, with three exceptions. The locus, PGI-2, had a null allele which co-migrated with a very common active allele in all populations except CR-86 and CC2-86. Since null heterozygotes could not be distinguished from homozygotes for the active allele, these loci were not used for any calculations except in the populations noted above. GDH activity was too low to give readable zymograms except in HC-85. Standard errors of the t estimates were calculated by inverting the information matrix of the maximum-likelihood estimation procedure. The goodness-of-fit of the data to the assumptions of the mixed-mating model was evaluated by a Chi-square test (Ritland and Jain 1981). Those loci whose χ^2 value was significant at the $p<0.05$ level were excluded from the multi-locus t estimates.

The coancestry coefficient, f, was calculated as $\hat{f}=1-\hat{h}_o/\hat{h}_e$, where $\hat{h}_o$ and $\hat{h}_e$ are, respectively, the frequencies of heterozygotes observed and expected under random mating. Expected heterozygosity was rescaled to account for the number of genotypes in the sample, n, following Nei and Chesser (1983); corrected $\hat{h}_e=(n/n-1)\ (\hat{h}_e-\hat{h}_o/2n)$. Separate f estimates were calculated for every polymorphic locus in each population for the same maternal genotypes used to estimate t. The value of f was also estimated for a subsample of the progeny used for t estimation. One progeny genotype was selected randomly from the six to eight progeny electrophoresed per maternal plant. Population f values were estimated separately, i.e., no hierarchical F statistics for the whole assemblage of populations were calculated. We did this because the populations were found to have widely varying outcrossing rates. Hence, the values of $\hat{F}_{IS}$ and $\hat{F}_{ST}$ would mask interesting variation since these statistics are means of values calculated for each population, which in this case are inbred to widely different degrees. However, $\hat{G}_{ST}$ was calculated for the whole assemblage of populations, and separately for the groups of predominantly selfing and predominantly outcrossing populations. Estimates of f were combined over loci using the weighting advocated by Reynolds et al. (1983) and Weir and Cockerham (1984). The f estimates were then jackknifed over loci by eliminating one locus at a time and recalculating $\hat{f}$. This procedure yields a less biased estimate of f and allows the variance of $\hat{f}$ to be estimated numerically (Weir and Cockerham 1984).

Nei's gene diversity statistics were calculated by the methods of Nei and Chesser (1983) for unequal sample sizes using the program Genestat (Whitkus 1985). Only the MK-86 accession represented the MK population in this analysis; we did not want to bias the gene diversity estimations by including two collections of the same population. In addition to estimating H_T, H_S, D_{ST}, G_{ST} and R_{ST} for the assemblage of all seven populations, we have grouped the populations into selfing and outcrossing groups, based on several morphological and phenological traits (Vasek and Weng 1988; T. P. Holtsford and N. C. Ellstrand, unpublished results). The intermediate outcrossing MK-86 accession was included with HC-85, I1-86, and I2-86 in the outcrossing group because its glasshouse autofertility is low and its average stigma exertion, degree of protandry, and pollen/ovule ratio indicate a preponderance toward outcrossing. CR-86, CC1-86, and CC2-86 comprise the selfing group, based on these same traits and their low outcrossing rate estimates. All loci assayed (18–20 per population, Table 3) were included in this analysis, since a large sample of both monomorphic and polymorphic loci are required to give an unbiased estimate of gene diversity and differentiation (Nei 1987, p. 190). We also estimated gene diversity statistics using only the polymorphic loci so that our results may be compared to other published estimates (e.g., Loveless and Hamrick 1984).

Results

Outcrossing rates vary widely among populations of *C. tembloriensis* (Table 2). The range in single-locus outcrossing rate estimates is $\hat{t}=0-0.88$, while the range in multi-locus estimates is $\hat{t}=0.03-0.87$. Although there is some heterogeneity among the single-locus estimates within populations (e.g., MDH-1 in CC1-86 and PGM-4 in MK-86), there is general agreement between the t estimates for a given population. Chi-square tests of the goodness-of-fit of the observed frequencies of progeny genotypes to those expected, based on the maternal genotype frequencies, $\hat{t}$, and pollen allele frequencies, are significant in 7 of 36 tests (denoted by * in Table 2). Most of the data fit the assumptions of the mixed-mating model (Clegg 1980) sufficiently to give reasonable t estimates (Ritland and Jain 1981).

The values of the mean number of alleles per locus, $\bar{k}$, and the mean percentage of loci at which individuals were heterozygous, $\bar{H}_o$, for the reconstructed maternal genotypes are positively correlated with the multi-locus t estimates (Tables 3 and 4). The residuals of $\bar{k}$ and PLP, when these variables were regressed on $\hat{t}$, were satisfactorily normally distributed. Only the distribution of residuals of $\bar{H}_o$ was significantly non-normal, as determined by the value of the Shapiro-Wilk statistic. Hence, the Spearman correlation coefficients are more valid for $\bar{H}_o$, while the more powerful Pearson's r estimates may be used for correlations among the other variables. There was a non-significant positive association between the percentage of loci which were polymorphic (PLP) and $\hat{t}$.

Coancestry coefficient estimates for the maternal plants, $\hat{f}_M$, are indistinguishable from the F_{eq} value expected given the populations' outcrossing rates in five of

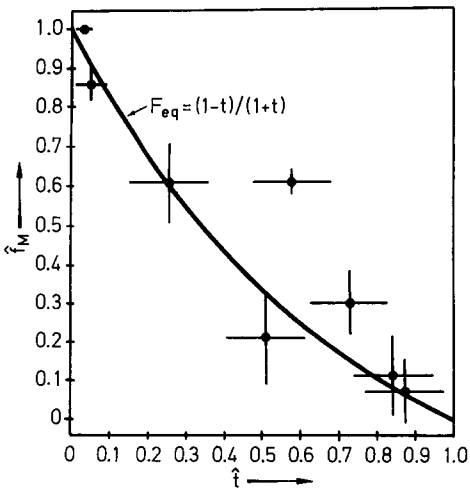

Fig. 2. The coancestry coefficient of maternal plants, $\hat{f}_M$, plotted as a function of multi-locus outcrossing rate, $\hat{t}$. The estimates are plotted ± 2 standard errors of the statistic. The value of f expected at inbreeding equilibrium, F_{eq}, is also plotted

Table 2. Single and multi-locus estimates of outcrossing rate $\hat{t}$, and fixation index $\hat{f}$, and their standard errors (SE). Loci whose single-locus $\hat{t}$ is asterisked (*) were not included in the multi-locus $\hat{t}$. All loci were included in the jackknife estimate of multi-locus $\hat{f}_P$. Allele frequencies are for the total sample of progeny genotypes while $\hat{f}_P$ values are for a sample of one randomly chosen progeny per maternal plant

Pop.-yr.	No. Families, no. progeny	Locus	Allele freq.	$\hat{t}$ (SE)	$\hat{f}_P$ (SE)
CR-86	26, 156	PGM-4	0.92/0.08	0.01 (0.01)	1.0
		PGI-2	0.76/0.16/0.08	0.02 (0.02)	1.0
		TPI-1	0.08/0.92	0.01 (0.01)	1.0
		PGD-2	0.96/0.04	0.00 (0.02)	1.0
		Multi-locus		0.03 (0.01)	1.0 (0.0)
CC2-86	40, 240	PGI-2	0.55/0.45	0.05 (0.02)	0.90
		PGD-2	0.17/0.83	0.01 (0.01)	1.0
		Multi-locus		0.05 (0.02)	0.92 (0.04)
CC1-86	31, 186	SKDH	0.11/0.18/0.71	0.28 (0.06)	0.71
		MDH-1	0.11/0.89	0.03 (0.02)*	1.0
		PGD-3	0.07/0.93	0.18 (0.16)	1.0
		Multi-locus		0.26 (0.05)	0.77 (0.1)
MK-86	26, 194	SKDH	0.04/0.95/0.01	0.46 (0.22)	0.55
		PGM-4	0.44/0.56	0.18 (0.04)*	0.68
		TPI-1	0.78/0.22	0.61 (0.10)	0.61
		Multi-locus		0.58 (0.05)	0.65 (0.02)
MK-87	38, 228	PGD-2	0.93/0.07	0.52 (0.17)	0.54
		PGM-4	0.60/0.40	0.32 (0.06)*	0.20
		TPI-1	0.81/0.19	0.51 (0.11)	0.66
		Multi-locus		0.51 (0.05)	0.36 (0.11)
HC-85	15, 117	PGM-4	0.36/0.32/0.32	0.57 (0.08)	0.28
		AAT-1	0.37/0.63	0.36 (0.08)	0.54
		PGD-4	0.74/0.26	0.36 (0.09)*	0.31
		PGD-2	0.87/0.13	0.47 (0.16)	0.65
		GDH	0.19/0.81	0.86 (0.15)	0.33
		SKDH	0.78/0.22	0.85 (0.15)	−0.27
		Multi-locus		0.73 (0.05)	0.29 (0.04)
I1-86	27, 164	SKDH	0.37/0.21/0.42	0.77 (0.07)*	0.22
		PGM-4	0.68/0.24/0.08	0.72 (0.09)	−0.07
		AAT-1	0.52/0.48	0.88 (0.10)	0.07
		Multi-locus		0.84 (0.05)	0.10 (0.05)
I2-86	26, 156	SKDH	0.16/0.30/0.54	0.54 (0.07)*	0.11
		PGM-4	0.71/0.27/0.02	0.84 (0.10)	0.23
		AAT-1	0.53/0.47	0.68 (0.08)	0.01
		PGD-2	0.89/0.11	0.74 (0.19)	−0.06
		Multi-locus		0.87 (0.05)	0.09 (0.02)

Table 3. Measures of allelic diversity. $\bar{k}$: mean number of alleles per locus. PLP: percent of loci which were polymorphic. $\bar{H}_0$: mean observed percentage of loci which were heterozygous. Standard errors of the means are in parentheses

Accession	$\hat{t}$	No. loci	$\bar{k}$ (SE)	PLP	$\bar{H}_0$ (SE)
CR-86	0.03	20	1.25 (0.10)	25.00	0.97 (0.67)
CC2-86	0.05	20	1.10 (0.07)	10.00	0.50 (0.37)
CC1-86	0.26	18	1.28 (0.18)	16.67	1.11 (0.43)
MK-87	0.51	18	1.29 (0.11)	16.67	2.47 (0.70)
MK-86	0.58	18	1.11 (0.07)	16.67	1.77 (0.59)
HC-85	0.73	19	1.67 (0.20)	31.58	14.90 (1.39)
I1-86	0.84	18	1.55 (0.22)	22.22	9.50 (1.16)
I2-86	0.87	19	1.50 (0.20)	26.31	9.12 (0.80)

Table 4. Pearson's (upper triangle) and Spearman's rank (lower triangle) correlation coefficients among multi-locus outcrossing rates: $\hat{t}$, mean observed heterozygosities: $\bar{H}_0$, mean number of alleles/locus: $\bar{k}$, and percent of loci polymorphic: PLP. Significance levels in parentheses are denoted by *: $p<0.05$, **: $p<0.01$, ***: $p<0.001$

	$\hat{t}$	$\bar{H}_0$	$\bar{k}$	PLP
$\hat{t}$	–	0.78 (*)	0.72 (*)	0.52 ($p=0.18$)
$\bar{H}_0$	0.85 (**)	–	0.94 (***)	0.80 (*)
$\bar{k}$	0.71 (*)	0.93 (**)	–	0.91 (**)
PLP	0.51 ($p=0.19$)	0.85 (**)	0.91 (**)	–

Table 5. Interpopulation f statistics. Multilocus outcrossing rate: $\hat{t}$, F_{eq}: expected $\hat{f}$ at inbreeding equilibrium, $\hat{f}_M$: $\hat{f}$ of maternal plants, $\Delta F = \hat{f}_M - F_{eq}$, $\hat{f}_P$: $\hat{f}$ of progeny. Significant differences denoted by **: $p<0.01$, ***: $p<0.001$

Accession	$\hat{t}$	F_{eq}	$\hat{f}_M$ (SE)	ΔF	$\hat{f}_P$ (SE)	$\hat{f}_P - \hat{f}_M$
CR-86	0.03	0.94	1.0 (0.0)	0.06	1.0 (0.0)	0.0
CC2-86	0.05	0.90	0.86 (0.02)	−0.04	0.92 (0.04)	0.06
CC1-86	0.26	0.59	0.61 (0.04)	0.02	0.77 (0.10)	0.16
MK-87	0.51	0.32	0.21 (0.06)	−0.11	0.36 (0.11)	0.15
MK-86	0.58	0.27	0.65 (0.02)	0.38***	0.66 (0.02)	0.01
HC-85	0.73	0.16	0.33 (0.04)	0.17**	0.29 (0.04)	−0.04
I1-86	0.84	0.09	0.11 (0.05)	0.02	0.10 (0.05)	−0.01
I2-86	0.87	0.07	0.07 (0.04)	0.0	0.09 (0.02)	0.02

Table 6. Nei's statistics of gene diversity and differentiation for three groups: all 7 populations (tot), 4 outcrossing populations (ox), and 3 selfing populations (sf). Estimates are means over all 18 loci (all) or only those loci which were polymorphic in each group (poly). See 'Results' for descriptions of statistics

	tot		ox		sf	
	all	poly	all	poly	all	poly
$\hat{H}_T$	0.19	0.39	0.20	0.45	0.13	0.30
$\hat{H}_S$	0.08	0.17	0.12	0.26	0.04	0.09
$\hat{D}_{ST}$	0.11	0.22	0.08	0.19	0.10	0.22
$\hat{G}_{ST}$	0.57	0.57	0.42	0.42	0.71	0.71
$\hat{R}_{ST}$	1.54	1.53	0.96	1.66	3.64	2.15

eight cases (Table 5, Fig. 2). The MK-86 and HC-85 accessions are significantly more inbred than expected (Student's $t=34.0$ and 4.25, $df=2$ and 5, respectively, where $df=$ no. loci-1). There was no variation in $\hat{f}$ among loci in CR-86; all 30 maternal genotypes were homozygous at all four polymorphic loci. The genotype at one of these loci, PGM-4, was known by assaying maternal leaf tissue. Maternal genotypes at the other three loci were inferred from an array of six progeny genotypes. Since there is no variation in $\hat{f}_M$ in this accession, no statistical test can be done.

Inbreeding coefficients for the maternal ($\hat{f}_M$) and progeny ($\hat{f}_P$) generations are generally similar (Table 5). In none of the individual populations are offspring significantly more inbred than their parents (by Student's t-test, $df=2$ (no. loci-1)). A Wicoxon signed-rank test shows that the overall difference between $\hat{f}_P$ and $\hat{f}_M$, considering all seven non-zero differences, is positive and marginally significant ($T_s=4$, $n=7$, $p=0.055$).

The group of outcrossing populations has a greater average total genetic diversity, $\hat{H}_T$, greater gene diversity within populations, $\hat{H}_S$, and lower gene diversity between populations, $\hat{D}_{ST}=\hat{H}_T-\hat{H}_S$, than the selfing group (Table 6). The selfing populations have less genetic diversity and are more differentiated from each other than the outcrossing populations. The coefficient of gene differentiation among populations, $\hat{G}_{ST}=\hat{D}_{ST}/\hat{H}_T$, is greater for the selfing populations than for the outcrossers, as is the ratio of inter- to intra-population gene diversity, $\hat{R}_{ST}=\hat{D}_m/\hat{H}_S$, ($\hat{D}_m=s\hat{D}_{ST}/s-1$), where s is the number of populations in the group). The relative magnitudes of the gene diversity estimates for the group of selfing populations versus the group of outcrossing populations are changed very little when the monomorphic loci are left out of the analysis.

Discussion

Clarkia tembloriensis exhibits a remarkable range of outcrossing rates among its populations. Eighty-four percent of the possible range of $\hat{t}$ (0.03–0.87) was represented among the sampled populations. Few other species show as much mating system variation: *Gilia achilleifolia* (81%, Schoen 1982a), *Lupinus succulentus* (83%, Harding and Barnes 1977), *Lupinus nanus* (100% range of $\hat{t}$ among subspp., 94% within subsp. *apricus*, Harding et al. 1974), *Collinsia sparsiflora* (84%, Allard et al. 1977), and *Eichhornia paniculata* (69%, Glover and Barrett 1986).

The t estimate from I1-86 is very similar to that of Vasek and Harding (1976) for the same population using seeds collected in 1965. The Idria Road populations no. 1 and no. 2 of Vasek and Harding (1976) were 20 m apart and correspond to the I1-86 accession of this paper. Vasek and Harding (1976) estimated outcrossing rates of 0.67 and 0.83 for Idria Road nos. 1 and 2, respectively, while in our study single locus $\hat{t}$ ranged from 0.72 to 0.88 and the multi-locus estimate was 0.84 (Table 2).

The progeny genotype distributions of seven of the loci did not fit the mixed mating model as judged by a χ^2 test. In all but one case (SKDH in I1-86, Table 2), the loci that do not conform to the model yield lower t estimates than other loci in those populations. The low values of these t estimates and the poor fit of the data from these loci to the mixed-mating model may reflect biparental inbreeding (Ritland and Jain 1984). Outcrossing to related individuals occurs if alleles are spatially clumped and matings are more likely among neighbors. Distance-de-

pendent matings in spatially-structured populations will bias t downward (Ennos and Clegg 1982).

Our data suggest that the mating system has probably had a major effect on the distribution of genetic variation within and among populations of *C. tembloriensis*. The most direct population-level effect of increased self-fertilization is decreased frequency of heterozygotes in the population (Wright 1921). This expectation is borne out by our data, whether heterozygosity is expressed as $\bar{H}_O$ (Tables 4 and 5) or as $\hat{f}$ (Fig. 2).

Inbreeding is expected to promote genetic drift by reducing effective population size N_e (Crow 1954). Since selfing affects heterozygosity directly, one would expect the relationship between $\hat{t}$ and $\bar{H}_o$ to be stronger than that of $\hat{t}$ with $\bar{k}$. If polyallelic loci lose diversity to genetic drift one allele at a time, then $\bar{k}$ should be influenced by outcrossing rate more directly than PLP. This expectation is borne out in *C. tembloriensis* – the magnitude of the Spearman's and Pearson's correlation coefficients are greatest between $\hat{t}$ and $\bar{H}_o$ and lowest between $\hat{t}$ and PLP (Table 4). This ranking of correlation coefficients, according to the directness of the effect of outcrossing rate on the statistic in question, was also found in *E. paniculata* (Glover and Barrett 1987) and *Lycopersicon pimpinellifolium*, a species exhibiting a 40% range of outcrossing rates among populations (Rick et al. 1977). In *G. achilleifolia*, however, $\bar{H}_o$ was not as strongly correlated with $\hat{t}$ as were $\bar{k}$ and PLP (Schoen 1982b).

Although the direct effect of outcrossing rate on average heterozygosity is clear, one should interpret the correlations between $\hat{t}$ and $\bar{k}$ with caution (Jain 1975, 1976). Effective population size, N_e may be influenced by factors other than inbreeding. Population bottlenecks – small numbers of founding, surviving, or reproducing plants – may bring about reductions in N_e and promote drift (Crow 1954). The distribution of the selfing populations on the ecological and geographic margin of *C. tembloriensis'* range suggests that the lower genetic diversity in the selfing populations, relative to the outcrossing populations, may be due in part to population bottlenecks. Higher rates of self-fertilization are often found in geographically marginal (Rick et al. 1977; Schoen 1982 a) or ecologically marginal populations (Allard et al. 1977; Brown et al. 1978). The more selfing populations of *Clarkia tembloriensis* are consistently found at lower elevations on the eastern margin of the north-south band that describes the species range (Vasek and Harding 1976; personal observation). Colonization of geographically marginal sites seems more likely if the colonizers are self-pollinating, since they do not require other individuals to be present for successful mating (Baker 1955; Jain 1976). The habitats of the selfing *C. tembloriensis* populations are on the ecological boundary of the species distribution as well as on its geographic boundary. All annual species in these sites flower and die earlier than their conspecifics growing sympatrically with the more outcrossing *C. tembloriensis* populations, presumably in response to the earlier drying out of the soil in these lower, hotter habitats. Hence, bottlenecks due to poor plant survivorship in very dry years seem more likely in the marginal selfing populations. Plants which happen to be growing in a good microclimate may survive, as will plants with drought-resistant genotypes.

There is no good evidence for selection being the primary evolutionary force shaping the distribution of alleles at the marker loci examined in *C. tembloriensis*. If the restricted allelic diversity in selfing populations in *C. tembloriensis* was primarily due to strong selection for adaptation to xeric conditions, then one might expect the same set of alleles at some of the marker loci to be found in all xeric sites, as has been found in several species along various environmental gradients (Allard et al. 1972; Hamrick and Allard 1972; Antlfinger 1981; Nevo et al. 1986, 1988). However, the matrix of Nei's genetic distances between populations of *C. tembloriensis* shows that the selfing populations are no more similar to each other than they are to the outcrossing populations (T. P. Holtsford and F. C. Vasek, unpublished results). If strong selection for xeric adaptation did play a role in reducing the allelic diversity in this species, then the marker loci were not directly selected upon and were not in linkage disequilibrium with the selected loci. This is also true for *Eichhornia paniculata;* the most highly selfing populations in Jamaica and Brazil have the smallest genetic distances to nearby highly outcrossing populations in those regions, not to the other selfing populations (Glover and Barrett 1987). In *G. achilleifolia*, the selfing populations are all clustered at the northern end of the species range, so that the effects of selection and founder effects are confounded in the values of the genetic distances among populations (Schoen 1982a). Whether by selection or drift, reduced numbers of individuals surviving to successfully reproduce in dry years may have played a large role in the loss of allelic diversity in the selfing populations of *C. tembloriensis*. Additionally, if pollinator abundance is also reduced in dry years, self-pollinating genotypes may leave more offspring than potentially pollination-limited outcrossers. Thus, the same ecologically harsh conditions which serve to diminish allelic diversity by selection or drift may also select for self-pollinating traits. Reduced allelic diversity then may be correlated with, but not necessarily caused by, evolution of a self-pollinating mating system.

The estimated values of Wright's (1965) inbreeding coefficient, $\hat{f}$, generally correspond to their expectations based on the t estimates (Table 5, Fig. 2). There are two cases of deviations of the f estimate for the sample of maternal genotypes, $\hat{f}_M$, from $F_{eq}=(1-t)/(1+t)$, the expected equilibrium value of F if inbreeding was due only to self-fertilization. The difference between F_{eq} and $\hat{f}_M$ is

known as ΔF (Brown 1979). ΔF is positive for the MK-86 and HC-85 accessions; these samples are more inbred than expected at inbreeding equilibrium (Table 5, Fig. 2). Selfing species are often found to have higher levels of heterozygosity than expected (negative ΔF values), and many outcrossing populations have lower heterozygosities than expected (Brown 1979; Schoen 1982b). The ΔF values in this study do not show such a pattern, as was also the case in *E. paniculata* (Glover and Barrett 1987).

The MK-86 accession is also significantly more inbred than the sample from the same population in the next year MK-87. In the 1986 sample, the SKDH and PGM-4 genotypes are known from having assayed maternal leaf tissue. The TPI-1 genotypes from 1986 and all genotypes from 1987 were inferred from the array of six to eight progeny genotypes assayed from each mother. The gene frequencies from the two collections are significantly different by a Chi-square test of heterogeneity for the PGM-4 and PGD-2 loci ($\chi^2 = 10.5$, 5.8; $p < 0.001$, 0.01, respectively). SKDH was monomorphic in the 1987 sample and PGD-3 was monomorphic in the 1986 sample, but these loci gave similar t and f estimates in the years they were polymorphic (Table 3). Although PGM-4 is the locus whose f and t estimates are most different between years, if this locus is excluded from the multilocus f estimates, the values of $\hat{f}_M$ are still significantly different from each other ($\hat{f}_M = 0.64$ and 0.34 for 1986 and 1987, respectively). The spring season of 1987 was considerably drier than that of 1986 and all populations of *C. tembloriensis* observed had many fewer plants in them. In particular, the MK population was much smaller and less dense in 1987 than in 1986. It may be that the harsh conditions of the 1987 growing season imposed selection on this population such that plants with more heterozygous loci were more likely to survive.

Heterozygote advantage, as measured by the difference between $\hat{f}$ of maternal and progeny generations, is not significant in any single population surveyed in this study (Table 5). MK-87 and CC1-86 are the only accessions where the parental generation appears substantially less inbred than their progeny. However, the large variance in $\hat{f}_P$ in these samples rendered the differences between $\hat{f}_P$ and $\hat{f}_M$ non-significant. The nearly significant Wilcoxon signed rank statistic indicates that there may be an overall trend for the maternal plants to be less inbred than their progeny. In several studies, adult plants were significantly less inbred than their seeds, implying heterotic selection during the course of development (e.g., Ellstrand et al. 1978; Moran and Brown 1980; Schoen 1982b; Kesseli and Jain 1985).

Nei's gene diversity statistics (Table 6) confirm the nature of the relationship between genetic structure and outcrossing rate that the previous discussion has suggested. The average gene diversity, H_T, in the group of four outcrossing populations is larger than the $\hat{H}_T$ in the group of three selfing populations. Decomposing these gene diversities into the gene diversities within and among the populations in each group shows that the group of outcrossing populations sampled here has a higher mean *within* population diversity, $\hat{H}_S$, than the selfing group. The mean *between* population diversity, $\hat{D}_{ST}$, is slightly greater among the selfing populations than among the outcrossing populations: 0.10 versus 0.08 when all loci are included. The average coefficient of gene differentiation among populations within the two groups, $\hat{G}_{ST} = \hat{D}_{ST}/\hat{H}_T$, is higher for the three selfing populations, 0.71 versus 0.42 for the four outcrossing populations. $\hat{G}_{ST}$ is unaffected by the inclusion of monomorphic loci.

The $\hat{G}_{ST}$ of the outcrossing group is significantly greater than the $\hat{G}_{ST}$ average of 0.12 (SE = 0.04) measured for 76 predominantly outcrossed species (Loveless and Hamrick 1984). The $\hat{G}_{ST}$ for the selfing *C. tembloriensis* populations is not significantly different from the average $\hat{G}_{ST} = 0.52$ (SE = 0.21) value of 39 autogamous species (Loveless and Hamrick 1984). When only polymorphic loci are considered, the estimate of $\hat{H}_T$ of 0.45 for the outcrossing *C. tembloriensis* populations is substantially greater than the average $\hat{H}_T = 0.24$ (SE = 0.06) for predominantly outcrossing species presented by Loveless and Hamrick (1984). The $\hat{H}_T$ of 0.30 for our selfing populations is very close to the average for autogamous species ($\hat{H}_T = 0.29$, Loveless and Hamrick 1984). Our $\hat{H}_S$ values indicate that the genetic diversities within populations of the selfing and outcrossing groups of *C. tembloriensis* are comparable to the averages for populations of selfing and primarily outcrossing species [$\hat{H}_S = 0.13$ (SE = 0.03) and 0.21 (SE = 0.03), respectively, Loveless and Hamrick 1984].

Gene diversity statistics can be rescaled to be independent of the number of populations in each group ($\hat{G}'_{ST} = \hat{D}_m/\hat{H}_T$, Nei 1973, 1987 p. 191). The average (over all loci) minimum genetic distance between populations in the selfing and outcrossing groups is $\hat{D}_m = 0.10$ for the outcrossers and 0.14 for the selfers. $\hat{G}'_{ST} = 0.40$ for the outcrossers and 0.78 for the selfers. These rescaled statistics have the limitation of not being simple additive components of the total gene diversity (Nei 1973, 1987). However, they show that the general pattern of genetic organization within and among the outcrossing and selfing populations of *C. tembloriensis* does not hinge on the sample containing more outcrossing than selfing populations. When all loci are considered, the mean values of $\hat{R}_{ST} (= \hat{D}_m/\hat{H}_S$, the ratio of inter- to intra-populational gene diversity) shows that the group of selfing populations have more than three times the genetic diversity spread between their populations than is represented within an average selfing population. There is about the same amount of genetic diversity within an average outcrossing population as occurs among them.

Acknowledgements. This paper is part of the dissertation of T. H. Dr. F. C. Vasek was invaluble to this project; he supplied collection sites, seeds, and a wealth of information about *Clarkia* biology and evolution. The authors thank J. Lee, V. Weng, S. Rogers, M. Bartholomew, L. Patty, B. LaCerra, C. La Place, J. Nason, and D. Glover for their help with field, lab, or greenhouse work. F. C. Vasek, M. T. Clegg, D. E. Glover, and J. Kohn made helpful comments on the manuscript. T. and D. Twisselman, P. and L. Martin, and Mr. and Mrs. E. Garcia graciously allowed T. H. to collect seeds and conduct experiments on their ranches. This study was funded by NSF Dissertation Research Grant BSR8612338 and a Chancellor's Patent Fund Award from University of California, Riverside.

References

Allard RW, Jain SK, Workman PL (1968) The genetics of inbreeding species. Adv Genet 14:55–131

Allard RW, Babbel GR, Clegg MT, Kahler AL (1972) Evidence for coadaptation in *Avena barbata*. Proc Natl Acad Sci USA 69:3043–3048

Allard RW, Kahler AL, Clegg MT (1977) Estimation of mating cycle components of selection in plants. In: Christiansen FB, Fenchel TM (eds) Measuring selection in natural populations. Springer, Berlin Heidelberg New York, pp 1–20 (Lecture notes in biomathematics 19)

Antlfinger AE (1981) The genetic basis of microdifferentiation in natural and experimental populations of *Borrichia frutescens* in relation to salinity. Evolution 35:1056–1068

Archie JW (1985) Statistical analysis of heterozygosity data: independent sample comparisons. Evolution 39:623–637

Baker HG (1955) Self-compatibility and establishment after "long-distance" dispersal. Evolution 9:347–348

Brown AHD (1979) Enzyme polymorphism in plant populations. Theor Pop Biol 15:1–42

Brown AHD, Allard RW (1970) Estimation of the mating system in open-pollinated maize populations using isozyme polymorphisms. Genetics 66:133–145

Brown BA, Clegg MT (1984) Influence of flower color polymorphism on genetic transmission in a natural population of the common morning glory *Ipomoea purpurea.* Evolution 38:796–803

Brown AHD, Zohary D, Nevo E (1978) Outcrossing rates and heterozygosity in natural populations of *Hordeum spontaneum* Koch in Israel. Heredity 41:49–62

Clegg MT (1980) Measuring plant mating systems. Bioscience 30:814–818

Crow JF (1954) Breeding structure of populations. II. Effective populations number. In: Bancroft TA, Gown JW, Lush JL (eds) Statistics and mathematics in biology. Iowa State University Press, Ames/IA, pp 543–556

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

Ennos RA, Clegg MT (1982) Effect of population substructuring on estimates of outcrossing rate in plant populations. Heredity 48:283–292

Glover DE, Barrett SCH (1986) Variation in the mating system of *Eichhornia paniculata* (Spreng.) Solms. (Pontederiaceae). Evolution 40:1122–1131

Glover DE, Barrett SCH (1987) Genetic variation in continental and island populations of *Eichhornia paniculata* (Pontederiaceae). Heredity 59:7–17

Gottlieb LD (1981a) Gene numbers in species of Asteraceae that have different chromosome numbers. Proc Nat Acad Sci USA 78:3726–3729

Gottlieb LD (1981b) Electrophoretic evidence and plant populations. Progr Phytochem 7:1–46

Gottlieb LD (1982) Conservation and duplication of isozymes in plants. Science 216:373–380

Gottlieb LD (1984) Electrophoretic analysis of the phylogeny of the self-pollinating populations of *Clarkia xantiana.* Plant Syst Evol 147:91–102

Hamrick JL, Allard RW (1972) Microgeographical variation in allozyme frequencies in *Avena barbata.* Proc Natl Acad Sci USA 69:2100–2104

Hamrick JL, Linhart YB, Mitton JB (1979) Relationship between life history characteristics and electrophoretically detectable genetic variation in plants. Annu Rev Ecol Syst 10:173–200

Harding J, Barnes B (1977) Genetics of *Lupinus.* X. Genetic variability, heterozygosity, and outcrossing in colonial populations of *Lupinus succulentus.* Evolution 31:247–255

Harding J, Mankinen CB, Elliott MH (1974) Genetics of *Lupinus.* VII. Outcrossing, autofertility, and variability in natural populations of the *nanus* group. Taxon 23:729–738

Jain SK (1975) Population structure and the effects of breeding systems. In: Frankel OH, Hawkes JG (eds) Crop genetic resources for today and tomorrow. Cambridge University Press, London, pp 15–36

Jain SK (1976) The evolution of inbreeding in plants. Annu Rev Ecol Syst 7:469–495

Kesseli RV, Jain SK (1985) Breeding systems and population structure in *Limnanthes.* Theor Appl Genet 71:292–299

Levene H (1949) On a matching problem arising in genetics. Ann Math Stat 20:91–94

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

Moran GF, Brown AHD (1980) Temporal heterogeneity of outcrossing rates in alpine ash (*Eucalyptus delegatensis* R. T. Bak.). Theor Appl Genet 57:101–105

Nei M (1973) Analysis of gene diversity in subdivided populations. Proc Nat Acad Sci USA 70:3321–3323

Nei M (1987) Molecular evolutionary genetics. Columbia University Press, New York

Nei M, Chesser RK (1983) Estimation of fixation indices and gene diversities. Ann Human Genet 47:253–259

Nevo E, Beiles A, Zohary D (1986) Genetics resources of wild barley in the Near East: structure, evolution, and application in breeding. Biol J Linn Soc 27:355–380

Nevo E, Beiles A, Krugman T (1988) Natural selection of allozyme polymorphisms: a microgeographic climatic differentiation in wild emmer wheat (*Triticum dicicoides*). Theor Appl Genet 75:529–538

Reynolds J, Weir BS, Cockerham CC (1983) Estimation of the coancestry coefficient: basis for a short-term genetic distance. Genetics 105:767–779

Rick CM, Fobes JF, Holle M (1977) Genetic variation in *Lycopersicon pimpinellifolium:* evidence of evolutionary change in mating systems. Plant Syst Evol 127:139–170

Ritland K, Jain S (1981) A model for the estimation of outcrossing rate and gene frequencis using *n* independent loci. Heredity 47:35–52

Schoen DJ (1982a) The breeding system of *Gilia achilleifolia:* variation in floral characteristics and outcrossing rate. Evolution 36:352–360

Schoen DJ (1982b) Genetic variation and the breeding system of *Gilia achilleifolia.* Evolution 36:361–370

Schoen DJ (1985) Correlation between classes of mating events in two experimental plant populations. Heredity 55:381–385

Schoen DJ, Clegg MT (1984) Estimation of mating system parameters when outcrossing events are correlated. Proc Natl Acad Sci USA 81:5258–5262

Schoen DJ, Clegg MT (1985) The influence of flower color on outcrossing rate and male reproductive success in *Ipomoea purpurea*. Evolution 39:1242–1249

Schoen DJ, Clegg MT (1986) Monte Carlo studies of plant mating system estimation models: the one-pollen parent and mixed mating models. Genetics 112:927–945

Shapiro SS, Wilk MB (1965) An analysis of variance test for normality (complete samples). Biometrika 52:591–611

Soltis PS, Bloom WL (1986) Genetic variation and estimates of gene flow in *Clarkia speciosa* subsp. *polvantha* (Onagraceae). Am J Bot 73:1677–1682

Vasek FC, Harding J (1976) Outcrossing in natural populations. V. Analysis of outcrossing, inbreeding, and selection in *Clarkia exilis* and *Clarkia tembloriensis*. Evolution 30:403–411

Vasek FC, Weng V (1988). Breeding systems of *Clarkia* sect. *Phaeostoma* (Onagraceae) I. Pollen-ovule ratios. Syst Bot 13:336–350

Weir BS, Cockerham CC (1984) Estimating F-statistics for the analysis of population structure. Evolution 38:1358–1370

Whitkus R (1985) FORTRAN program for computing genetic statistics from allelic frequency data. J Hered 76:142

Wright S (1921) Systems of mating. Genetics 6:111–178

Wright S (1965) The interpretation of population structure by F-statistics with special regard to systems of mating. Evolution 19:395–420