A Hierarchical View of Genetic Structure in the Rare Annual Plant *Clarkia springvillensis*

KIMBERLIE A. McCUE, EDWARD S. BUCKLER, AND TIMOTHY P. HOLTSFORD

Division of Biological Sciences, 105 Tucker Hall, University of Missouri, Columbia, MO 65211, U.S.A.

Abstract: *Genetic structure at several spatial scales was examined in the rare California annual,* Clarkia springvillensis. *Using seven isozyme-encoding loci as genetic markers, we assessed the amount and distribution of genetic variation among three populations and eight subpopulations. Total genetic variation was lower than in species with similar life history traits but equivalent to that of other endemic plants. Spatial autocorrelation showed some evidence for very limited differentiation within subpopulations at a scale of 1-2 m. The subpopulations, separated by tens of meters, were found to be more differentiated from each other (F_{sp} = 0.084) on average than were populations (F_{pt} = 0.017). This local genetic differentiation was not correlated with physical distance between subpopulations. The low F_{pt} estimates suggest that substantial gene flow is occurring among populations. However, the lack of correlation between genetic and geographic distances and the significant differentiation of subpopulations suggest that genetic drift is occurring* within *populations. Therefore, we believe the apparent homogeneity of populations is due to each population's gene frequencies' being an average of several divergent subpopulations. If drift is causing differentiation within populations, it may eventually cause differentiation between populations. The importance of using a hierarchical approach to evaluating genetic structure is clear. Patterns occurring at one spatial scale may not be evident at others. One should not necessarily conclude that gene flow is substantial and that the risk of genetic erosion via drift is negligible just because differentiation between populations is small; the system may not be at equilibrium. This lesson is particularly important when recent changes in climate or land use are apparent.*

Una visión jerárquica de la estructura genética de la planta anual rara *Clarkia springvillensis*

Resumen: *Se examinó la estructura genética de la planta anual californiana rara* Clarkia springvillensis *en varias escalas espaciales. Utilizando siete loci codificadores de isozimas como marcadores genéticos estimamos la cantidad y distribución de variación genética entre tres poblaciones y ocho subpoblaciones. La variación genética total fue menor en especies con tablas de vida similares pero equivalente a la de otras plantas endémicas. La autocorrelación espacial mostró alguna evidencia de diferenciación muy limitada entre subpoblaciones en una escala de 1-2 m. Se encontró que las subpoblaciones, separadas por decenas de metros, en promedio estaban más diferenciadas entre sí (F_{pt} = 0.084) que las poblaciones (F_{pt} = 0.017). Esta diferenciación genética local so se correlacionó con la distancia física entre las subpoblaciones. Las estimaciones bajas de F_{pt} sugieren que esta ocurriendo una sustancial deriva génica entre las poblaciones. Sin embargo, la falta de correlación entre las distancias genéticas y geográficas y la diferenciación significativa de las subpoblaciones sugiere que la deriva génica esta ocurriendo* dentro *de las poblaciones. Por tanto, pensamos que la homogeneidad aparente de las poblaciones se debe a que sus frecuencias génicas son promedio de subpoblaciones divergentes. Si la deriva esta causando diferenciación dentro de las poblaciones eventualmente puede producir diferenciación entre poblaciones. Es clara la importancia de utilizar un enfoque jerárquico para evaluar la estructura genética. Los patrones que ocurren en una escala espacial pueden no ser evidentes en otra. No se debería concluir que la deriva génica es considerable y que el riesgo de erosión genética vía deriva génica es despreciable solo porque la diferenciación entre poblaciones es pequeña; el sistema puede no estar en equilibrio. Esta lección es particularmente importante cuando son aparentes los cambios climáticos o de uso de suelo recientes.*

Address correspondence to T. P. Holtsford (email holtsford@biosci.mbp.missouri.edu).
Paper submitted August 7, 1995; revised manuscript accepted November 28, 1995.

Introduction

Knowledge of the amount and spatial distribution of genetic diversity in natural populations is critical for developing successful management strategies for populations of rare species (Loveless & Hamrick 1984; Millar & Libby 1991). The amount of genetic diversity present may serve as an indication of the long-term survival potential of a species. High levels of genetic variation may enable a population to adapt to changing environments (Beardmore 1983; Huenneke 1991; Van Treuren et al. 1991) or perhaps to exploit new environments, thus increasing the species' range. Knowledge of the genetic structure within and among populations can allow conservationists to manage natural populations to best conserve maximum diversity and to sample populations adequately if the establishment of an ex situ population is a goal.

The spatial distribution of genetic variability is the product of environmental influences, life history traits, and a species' demographic past. Therefore, genetic structure can be used to infer past patterns of gene flow (Dewey & Heywood 1988; Schoen & Latta 1989), drift (Barbujani 1987), or local adaptation due to natural selection (Epperson & Allard 1989; Epperson 1990). Genetic structure will also influence and be influenced by the mating system (Holtsford & Ellstrand 1989; Coates & Sokolowski 1992). Because drift, local adaptation, and the mating system can create nonrandom distributions of genotypes, biologists and managers should evaluate the genetic structure in populations before implementing management strategies. For example, some species will have high variation within populations but little differentiation among them. In this case protecting or managing a few populations may adequately protect the species' gene pool. Conversely, some species will have high variation both within and among populations. Conservation of the total gene pools of such species would require a priori knowledge of the distribution of variation.

Given the significance of understanding genetic structure for conservation, serious consideration should be given to determining the scale at which structure is to be evaluated. Genetic structure can be detected at the level of the population, subpopulation, or continuously among neighboring individuals. Although methods have been applied to reveal patterns among populations or subpopulations (Dole & Sun 1992; Richter et al. 1994) and small-scale structure (Dewey & Heywood 1988), rarely are all levels examined simultaneously. If patterns are not consistent across spatial scales, then the use of a single-level approach could miss important information. Therefore, we employed a hierarchical approach to examine genetic diversity and structure in an annual plant species. Our sampling strategy allowed us to assess the amount and distribution of genetic diversity among individuals, subpopulations, and populations. This step-down process, compared with single-level analyses, should provide a clearer picture of the distances over which differentiation is occurring (Loveless & Hamrick 1984).

We studied genetic structure in *Clarkia springvillensis*, an endangered annual plant endemic to California that exhibits several characteristics that could lead to spatial structuring of genetic variation. *Clarkia springvillensis* is patchily distributed throughout its 15.54 km^2 range, primarily in small populations that fluctuate in size yearly (California Department of Fish and Game 1991). Seed and pollen dispersal are probably limited because they are accomplished by gravity and bees, respectively. We pose hypotheses regarding the processes giving rise to the genetic structure and discuss the possible implications of our results for genetic assessments of rare species and for the long-term management of *C. springvillensis*.

Methods

Study System and Sampling

C. springvillensis Vasek (Onagraceae) is an annual plant that is narrowly distributed within Tulare County, near the town of Springville, California. It is listed as endangered under the California Endangered Species Act. Eleven extant populations are known, nine along the North Fork of the Tule River and two along the Middle Fork of the Tule River. Populations generally range in size from zero to several hundred plants in any given year. However, in some years, a few populations support thousands of individuals (California Department of Fish and Game 1991). Within populations the plants are patchily distributed. Aggregations of individuals most often occur on unstable slopes where grass densities and shade cover are low.

During April of 1993 we sampled three populations to determine the amount of variation present at isozyme encoding loci. We sampled two populations along the North Fork of the Tule River: Bear Creek (BC) and Springville *Clarkia* Ecological Reserve (SCER). These sites are separated by approximately 300 m. The third site, Gauging Station (GS), is located along the river's Middle Fork and is approximately 8 km from the BC and SCER sites. The only other population known along the Middle Fork contained so few plants that we did not sample it for fear of negatively affecting its persistence. In 1993 the SCER and BC sites contained over 1000 plants each. The GS site had several hundred plants.

At each site we ran a series of transects through discrete subpopulations. Three of five subpopulations were sampled at SCER, three of three at BC, and two of two at GS. We chose starting points for transects with a blind toss and chose a compass bearing with a random number generator. We mapped each transect with respect to the others and sampled leaf tissue from individuals at randomly chosen points along each transect. Sam-

Table 1. Summary of electrophoresis procedures.*

Gel buffer system	*Current (mA)*	*Volts (max)*	*Run time (hr)*	*Enzymes assayed*
LiOH-boric/citric acid pH 8.3	75	13/cm	4.5	AAT, GDH
LiOH-boric/citric acid pH 8.0	75	13/cm	4.5	ACP
Histidine-citrate pH 7.0	35	10/cm	7.5	MDH, PGM, SKDH

**Gel buffer systems 7, 8, and 11, respectively, from Soltis et al. (1983).*

ple sizes were 30 plants per subpopulation, 90 plants total for SCER and BC and 60 plants for GS.

Isozyme Analysis

The tissue samples of two leaves per plant were packed on ice in the field and express mailed to the University of Missouri for immediate extraction. Crude extractions were performed by freezing tissue with liquid nitrogen, grinding the tissue with mortar and pestle, and adding approximately 1 mL of extraction buffer (Gottlieb 1981). The extracts were absorbed onto filter paper wicks and stored at −80°C.

The following enzyme systems were assayed after starch gel electrophoresis of the crude extracts: acid phosphatase (*Acp*, one locus scored), aspartate-α-ketoglutarate transaminase (*Aat*, one locus scored, one putative locus unresolved), glutamate dehydrogenase (*Gdh*, one locus), malate dehydrogenase (*Mdh*, two loci), shikimate dehydrogenase (*Skdh*, one locus), and phosphoglucomutase (*Pgm*, 3 loci, Table 1). The most anodally migrating bands were designated "1", with slower loci and alleles numbered sequentially thereafter. Genetic interpretation of the banding patterns was based on previous studies elucidating the number of loci encoding these enzymes and their inheritance in *Clarkia* (Gottlieb & Weeden 1979; Gottlieb 1984; Soltis et al. 1987). Examination of progeny groups in *C. springvillensis* indicates Mendelian segregation of electromorphs.

Diversity Statistics and *F* Statistics

We calculated Nei's genetic diversity statistics to estimate diversity present within and among subpopulations and populations (Table 2). We used the methods of Nei (1973) to estimate gene diversity or expected het-

Table 2. Frequencies of isozyme variants and Nei's gene diversity in subpopulations (H_s), populations (H_p), and total (H_t) of *C. springvillensis*.[a]

Locus	*Allele*	*BC 1*	*BC 2*	*BC 3*	*GS 1*	*GS 2*	*SCER 1*	*SCER 2*	*SCER 3*
Mdh-f	1	0.000	0.000	0.033	0.017	0.034	0.000	0.036	0.034
	2	1.000	1.000	0.967	0.983	0.845	1.000	0.964	0.966
	3	0.000	0.000	0.000	0.000	0.121	0.000	0.000	0.000
Mdh-s	1	1.000	1.000	1.000	1.000	0.967	1.000	1.000	1.000
	2	0.000	0.000	0.000	0.000	0.033	0.000	0.000	0.000
Skdh	1	0.000	0.050	0.000	0.220	0.033	0.000	0.017	0.017
	2	0.217	0.125	0.052	0.040	0.217	0.091	0.083	0.190
	3	0.533	0.350	0.448	0.740	0.433	0.705	0.767	0.500
	4	0.250	0.325	0.310	0.000	0.317	0.159	0.083	0.241
	5	0.000	0.150	0.190	0.000	0.000	0.045	0.050	0.052
Pgm-1	1	0.050	0.000	0.000	0.000	0.000	0.017	0.000	0.017
	2	0.950	0.875	0.850	1.000	0.950	0.983	0.833	0.983
	3	0.000	0.125	0.150	0.000	0.050	0.000	0.167	0.000
Pgm-2	1	0.067	0.018	0.000	0.000	0.000	0.000	0.000	0.138
	2	0.933	0.929	1.000	0.614	1.000	1.000	1.000	0.845
	3	0.000	0.054	0.000	0.386	0.000	0.000	0.000	0.017
Pgm-3	1	0.367	0.241	0.150	0.000	0.067	0.077	0.308	0.207
	2	0.617	0.672	0.817	1.000	0.933	0.846	0.654	0.724
	3	0.017	0.086	0.033	0.000	0.000	0.077	0.038	0.069
Acp	1	0.767	0.583	0.567	0.907	0.783	0.414	0.759	0.650
	2	0.233	0.417	0.433	0.093	0.217	0.586	0.241	0.350
$H_s = 0.230$	H_{si}[b] =	0.238	0.293	0.255	0.154	0.223	0.180	0.226	0.273
$H_p = 0.240$	H_{pi}[b] =		0.268		0.211			0.240	
$H_t = 0.260$									

[a] *Gene diversity statistics do not include monomorphic loci.*
[b] i = *estimates for individual subpopulations or populations.*

Table 3. *F* statistics (±2 SE jackknifed over loci) for three populations of *C. springvillensis*.*

Site subpopulation	f_j	F_{isk}	F_{ipk}	F_{spk}
Bear Creek (BC)	—	0.191 ± 0.032	0.214 ± 0.342	0.028 ± 0.004
1	0.237 ± 0.204	—	—	—
2	0.273 ± 0.254	—	—	—
3	0.055 ± 0.112	—	—	—
Gauging Station (GS)	—	0.175 ± 0.054	0.322 ± 0.268	0.178 ± 0.050
1	−0.128 ± 0.244	—	—	—
2	0.348 ± 0.154	—	—	—
Springville Clarkia Ecological Reserve (SCER)	—	0.061 ± 0.072	0.122 ± 0.102	0.065 ± 0.016
1	0.167 ± 0.302	—	—	—
2	−0.061 ± 0.072	—	—	—
3	0.101 ± 0.126	—	—	—
Totals (hierarchical values)			F_{it} = 0.201 ± 0.022	F_{sp} = 0.084 ± 0.007 F_{pt} = 0.017 + 0.005

*i = *individuals;* s = *subpopulations;* p = *populations;* t = *total;* j = *individual subpopulations;* k = *individual populations.*

erozygosity (H) and coefficient of genetic differentiation between sites (G_{st}). Gene diversity in each (i^{th}) subpopulation was estimated as

$$H_{s_i} = \sum_{k=1}^{nloci} \sum_{j=1}^{nalleles} 1 - p_{jk}^2,$$

where p_{jk} is the frequency of the j^{th} allele at the k^{th} locus. The average gene diversity of all subpopulations was then estimated as the average of the H_{s_i}. To estimate diversity in each population, H_{p_i}, p_{jk} was the average (over subpopulations) allele frequency for each population. The H_p was then estimated as the average of the H_{p_i}. The total gene diversity present in the sample, H_t, was estimated from the grand average (over populations) allele frequencies. All H estimates included only polymorphic loci so that the estimates could be compared with Table 4 of Hamrick and Godt (1990). Comparisons of our H estimates with those compiled by those authors (Hamrick & Godt 1990) were made using Student's t tests.

The F statistics were estimated using the methods of Weir and Cockerham (1984) for populations and subpopulations (Table 3). Hierarchical F statistics were estimated using the methods of Weir (1990) with the following correction for the equation R_2 on page 158: the last term should be (MSG + MSI)/2 so that it has an expectation of $p_A(1 - p_A)$, (B. S. Weir, personal communication). Variances of the F statistics were estimated by jackknifing across loci (Weir & Cockerham 1984). Standard errors of the jackknife statistics were estimated following Sokal and Rohlf (1981).

Genetic Distances

We chose to use the Cavalli-Svorza and Edwards (1967) chord distance for estimating genetic distance instead of the commonly used methods of Nei (1972) (Table 4). The modified chord distance (Eq. 9.16 of Nei 1987) is a more appropriate method for characterizing divergence due to drift alone, whereas Nei's genetic distance models longer-term divergence due to mutation and drift (chapter 9 of Nei 1987). Cavalli-Svorza's distance statistic is also more sensitive than Nei's D to small changes in gene frequencies (Felsenstein 1985), and with small differences in gene frequencies the chord distance increases linearly with time (Nei 1987). The distance esti-

Table 4. Cavalli-Svorza genetic distances (above diagonal) and geographic distances (m; below diagonal) between subpopulations* of *C. springvillensis*.

	BC 1	*BC 2*	*BC 3*	*GS 1*	*GS 2*	*SCER 1*	*SCER 2*	*SCER 3*
BC 1	—	0.080	0.095	0.225	0.082	0.065	0.071	0.029
BC 2	16.70	—	0.030	0.213	0.095	0.068	0.052	0.045
BC 3	95.86	56.70	—	0.244	0.082	0.051	0.037	0.066
GS 1	8024	8000	7988	—	0.179	0.196	0.184	0.187
GS 2	8036	8012	8000	58.30	—	0.093	0.079	0.085
SCER 1	361.4	313.2	265.1	7988	8000	—	0.069	0.050
SCER 2	385.5	337.3	265.1	7964	7976	26.70	—	0.068
SCER 3	409.6	385.5	313.2	7951	7964	72.50	16.70	—

**Subpopulation names supplied in Table 3.*

mates and UPGMA clustering of those distances were done using the "Gendist" and "Neighbor" routines of Phylip version 3.5c (Felsenstein 1989).

To determine whether the genetic distance among subpopulations is correlated with the physical distance between them we used Kendall's and Spearman's tests for association between the matrix of genetic distances and the matrix of geographic distances. These tests are generally more useful than Mantel's test because their power does not depend on the scale of the distance measures used (Dietz 1983). The significance of Spearman's r and Kendall's τ was estimated by permutation tests, following Dietz (1983).

Migration Rates and Spatial Autocorrelation

For estimating rates of migration (Nm) we used Crow and Aoki's (1984) equation:

$$F_{st} = \frac{1}{1 + 4Nm\alpha},$$

where $\alpha = [n/(n-1)]^2$ and corrects for the number of populations or subpopulations being evaluated. This is modified from Wright's (1951) original infinite island model for estimating gene flow using F_{st}.

Spatial autocorrelation (Sokal & Oden 1978) was used to assess genetic structure within subpopulations. Moran's I was used as a measure of similarity at isozyme-encoding loci of plants separated from each other by 1 m, 2 m, . . . , 10 m. A positive estimate of I indicates a positive association between the genotypes of plants separated by a specified distance. We estimated I for the most common allele at each locus, provided that the frequency of that allele was ≤0.95 in each subpopulation. The estimates of I were then plotted as a function of the distance class to produce a correlogram.

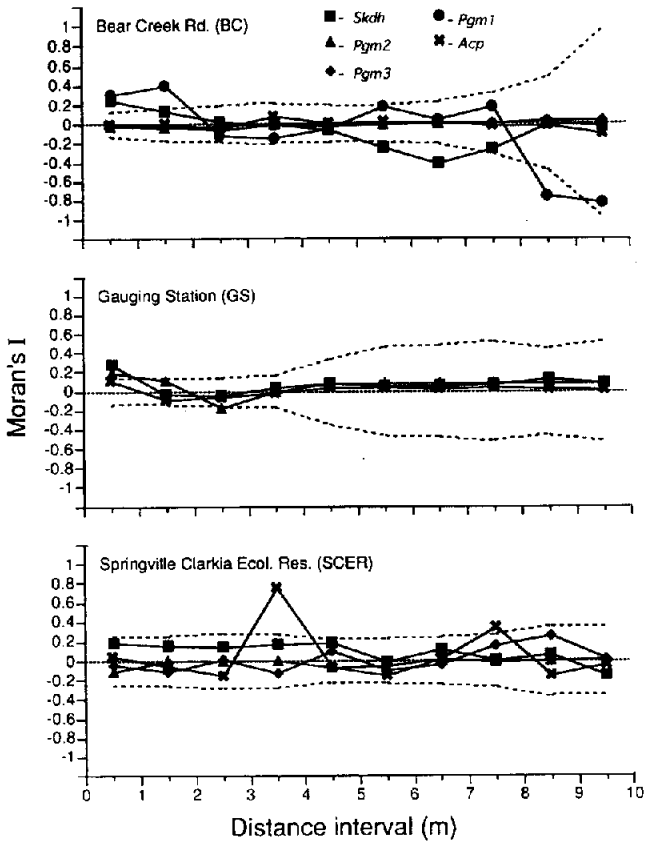

Figure 1. Correlograms showing Moran's I *statistic of association as a function of interplant distance for three populations of* Clarkia springvillensis. *Data are plotted at the midpoint of discrete 1-m intervals (i.e., the points plotted at 1.5 m are the estimates of* I *for all pairs of plants separated by >1 m and ≤2 m). The* I *estimates are for the most common allele at each locus. Dashed lines are ±2 standard normal deviates of the null expectation of* I *given the sample size of pairs of plants in each distance class.*

Results

Two of the nine isozyme-encoding loci assayed were monomorphic in the sample (*Aat-1; Gdh*). All other loci showed polymorphism in at least one population (Table 2). The GS 2 was the only subpopulation to have unique alleles, which were found at the *Mdh-s* and *Mdh-f* loci.

Spatial autocorrelation revealed signs of genetic structure within subpopulations at GS and BC. Plants separated by less than 1 m in GS and by less than 2 m in BC were significantly more similar to each other at some loci than would be expected if alleles were randomly distributed (Fig. 1). Two loci showed negative autocorrelation in BC for the 6- to 7-m and 8- to 9-m distance intervals, suggesting that plants 6–9 m apart may belong to different genetic neighborhoods. Another way to quantify genetic structure within patches is the inbreeding coefficient f. Although subpopulations BC 1, BC 2, and GS 2 had f estimates >0.2, the large variation in f among loci rendered these estimates not significantly greater than zero (by Student's t-tests, Table 3). The SCER population did not show any genetic structure within subpopulations—either by F statistics or by spatial autocorrelation—with one anomalous exception. Plants at SCER that were separated by 3–4 m and by 7–8 m were significantly more similar to each other at *Acp* than randomly chosen plants.

Nei's gene diversity statistics showed considerable variation among subpopulations and populations (Table 2). *Clarkia springvillensis* has levels of genetic diversity within populations that are comparable to an average mixed-mating, animal-pollinated species ($H_s = 0.221$) and species with narrow geographic range ($H_s = 0.215$) (Student's $t = 1.108$, $df = 59$; $t = 1.911$, $df = 82$; both, $p > 0.05$; note that within-population diversity in Hamrick & Godt [1990] is denoted H_s, whereas our estimates

of the same parameter are denoted H_p). However, *C. springvillensis* had significantly *more* genetic variation within populations (H_p), than the average annual plant species (H_s = 0.200), greater H_p than the average species with gravity dispersed seeds (H_s = 0.207), greater H_p than other mid-successional species (H_s = 0.205), and more within-population diversity than the average endemic plant species (H_s = 0.163) (t = 3.322, df = 145, $p < 0.001$; t = 2.991, df = 160, $p < 0.01$; t = 3.169, df = 120, $p < 0.01$; and t = 4.767, df = 51, $p < 0.001$; respectively). Considering the total genetic diversity of the sample (H_t), *C. springvillensis* is comparable to mixed-mating, animal-pollinated species (H_t = 0.304, t = 1.984, df = 59, $p > 0.05$) and to species with endemic ranges (H_t = 0.263, t = 0.129, df = 51, $p > 0.05$). However, *C. springvillensis* has lower total genetic diversity than the average species with an annual life history (H_t = 0.330), gravity-dispersed seeds (H_t = 0.306), mid-successional associations (H_t = 0.287), and narrow geographic range (H_t = 0.300) (t = 5.813, df = 145, $p < 0.001$; t = 4.586, df = 160, $p < 0.001$; t = 2.241, df = 120, $p < 0.05$; and t = 2.651, df = 81, $p < 0.01$; respectively).

Differentiation of subpopulations varied significantly among populations (F_{sp_k}, Table 3). Additionally, the hierarchical F statistics indicate that the average differentiation between subpopulations within populations F_{sp} is five times greater than differentiation between populations, F_{pt} (t = 24.8, df = 12, $p < 0.05$, Table 3). Nei's G_{st} was estimated to be 0.036 at the population level. This statistic is another estimate of the differentiation among populations (and designated G_{pt} following the above notation) and is not significantly different from our estimate of F_{pt} (t = 1.9, $p > 0.05$). We present the G_{st} estimate so that our results can be more directly compared with published estimates of differentiation among populations (Table 4 of Hamrick & Godt 1990). *C. springvillensis* shows much less differentiation among populations than the average annual species (G_{st} = 0.357), animal-pollinated mixed-mating species (G_{st} = 0.216), species with gravity dispersed seeds (G_{st} = 0.277), mid-successional species (G_{st} = 0.259), and species with either endemic (G_{st} = 0.248) or narrow geographic ranges (G_{st} = 0.242) (t = 13.329, df = 145, $p < 0.001$; t = 7.438, df = 59, $p < 0.001$; t = 11.441, df = 160, $p < 0.001$; t = 10.095, df = 120, $p < 0.001$; t = 5.675, df = 51, $p < 0.001$; t = 8.531, df = 81, $p < 0.001$; respectively).

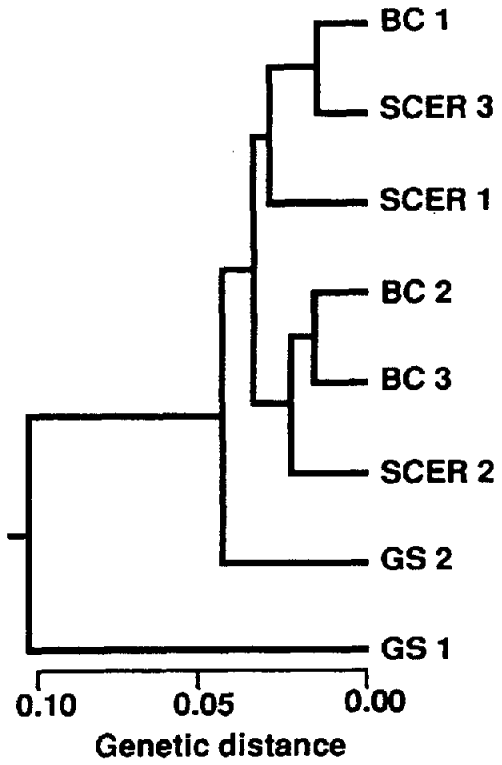

Figure 2. UPGMA clustering of the Cavalli-Svorza genetic distances (Table 4) among all subpopulations sampled within three populations of C. springvillensis. *Subpopulation names supplied in Table 3.*

At the population level, genetic distances are proportional to geographic distance. Populations SCER and BC are only approximately 300 m apart and are the least diverged from each other (D = 0.011). The GS population is approximately 8 km from the other populations and is correspondingly more genetically distant: D = 0.081 and 0.089 from SCER and BC, respectively. When all eight subpopulations are included, the matrix of genetic distances is significantly correlated with the matrix of geographic distance (Spearman's r = 6883.5, Kendall's τ = 48, $p < 0.05$ for both). However, we were concerned that the significance of this relationship was primarily the result of the comparisons of subpopulations along the Middle Fork of the Tule River (GS) with the subpopulations on the North Fork watershed (SCER and BC), because GS was most distant from SCER and BC both genetically and geographically. Within the North Fork watershed BC subpopulations 2 and 3 had little genetic distance between them, but they were not geographically proximal. The BC 1 and SCER1 populations did not cluster with their respective nearest neighbors (Fig. 2). When the matrix association tests were run on the six North Fork subpopulations alone there was no significant association between geographic and genetic distance (Spearman's r = 866, Kendall's $\tau = -13$, $p > 0.05$ for both).

We used F_{pt} for three populations to estimate that N_m (number of migrants per generation) is 6.6. However, when we used F_{sp} for eight subpopulations N_m was 2.1.

Discussion

Intrasubpopulation Structure

C. springvillensis exhibited little evidence of genetic structure within subpopulations. Three of eight subpop-

ulations had relatively high inbreeding coefficients, although the f estimates were not significantly greater than zero (GS 2 and BC 1 and 2, Table 3). The correlogram for BC shows a typical isolation by distance pattern for two loci: positive genetic associations for nearby pairs of plants (<1 or 2 m) and negative genetic associations for more distant pairs of plants (>6 or 8 m, Fig. 1). Therefore, the Wahlund effect may be contributing to the high f in BC. However, despite the high f in GS 2, the GS correlogram shows little evidence for small-scale spatial structure. We believe the geometry of the subpopulations may help explain why BC subpopulations exhibit evidence of substructure and the other populations do not. The subpopulations of BC are roughly linear, narrow strips (<1 m wide and about 15 m long) distributed along a steep bank and bounded by dense woody vegetation above and gravel below. The subpopulations in SCER and GS are amorphous two-dimensional patches on moderate slopes. The linear arrangement of the plants in BC may restrict the average dispersal distance because pollen or seeds dispersed >1 m above or below the population will not encounter suitable habitat to germinate or receptive stigmas to pollinate. Restricted dispersal distance will lead to greater genetic spatial autocorrelation. This result supports the theoretical prediction that differentiation proceeds much more rapidly in one-dimensional systems than in two- or three-dimensional systems (Kimura & Weiss 1964).

We suspect a small amount of self-fertilization may also be increasing homozygosity over random mating expectations. *C. springvillensis* flowers are roughly 3 days protandrous, and style length is about 1.6 cm (Vasek 1977). If the relationship between those floral traits and the selfing rate is similar to that of the closely related *C. temblorіensis*, we would predict a self-fertilization rate of about 10% (Holtsford & Ellstrand 1992).

Population and Subpopulation Structure

Genetic diversity at polymorphic loci within populations of *C. springvillensis* is comparable to or greater than that of other species with similar life-history traits. However, total diversity over all populations sampled is less than that of other annual, mid-successional species but comparable to that of other endemics with a mixed-mating system. Interestingly, genetic differentiation among populations (F_{pt}) was much lower than is typical for populations of mixed-mating animal-pollinated, gravity-dispersed, mid-successional narrow-range, or endemic annuals (Hamrick & Godt 1990).

Three scenarios might explain the low differentiation among populations. First, selection on the allozyme variants (or genes linked to them) has prevented the populations from diverging. A second possibility is that substantial gene flow could be maintaining the homogeneity of populations. The third explanation for small divergence among populations is that these populations have only recently been isolated from each other and therefore have not had time to diverge substantially. Distinguishing among these explanations using gene frequency data is difficult if not impossible (Felsenstein 1982). However, we believe that the hierarchical data, along with other natural history data and empirical evidence, support the hypothesis of recent divergence over the other two possibilities.

Natural selection could be expected to maintain similarities between populations if the habitats occupied were homogeneous. At each locus the same allele is the most common in every population. However, if selection was the force maintaining the genetic similarity we see, then that similarity should be strongest at the subpopulation level. Yet, the subpopulations are more genetically distinct from each other than are the populations (F_{sp} versus F_{pt}, Table 3). Therefore, we think selection is an unlikely explanation for the lack of differentiation between populations.

Gene flow is often invoked as the force preventing the divergence of populations. A migration rate (Nm) of 1 or greater should be sufficient to impede divergence. If we use F_{pt} for three populations then $Nm = 6.6$, indicating substantial gene flow. However, if we use F_{sp} for eight subpopulations then $Nm = 2.1$. It seems unlikely that gene flow would be greater among populations separated by hundreds or thousands of meters than among patches separated by tens of meters Rather, we think the low F_{pt} estimate is due to the effect of averaging gene frequencies over many moderately diverged subpopulations. Adding to the argument against interpopulation gene flow is the fact that *Clarkia* has no obvious long-distance seed dispersal features and is bee-pollinated (MacSwain et al. 1973). In *Clarkia concinna* bee pollinators move only short distances and are more likely to visit all flowers in one patch rather than travel between patches. Isolation distances under 100 m were sufficient to lower reproductive success in small patches of *C. concinna* (Groom 1995). Slatkin (1987) reviewed several examples where indirect methods of evaluating gene flow gave higher estimates than the estimates obtained through direct methods. He suggests that in these cases there may have been higher levels of gene flow in the recent past but that these species have probably undergone large-scale demographic change that now restricts migration.

We hypothesize that *C. springvillensis* may have recently had a more continuous distribution and that some environmental change caused the plants to retreat into more isolated populations with more patchy distributions of individuals within them. We can estimate the time since isolation given our estimate of genetic divergence and a rough estimate of the effective population sizes. Genetic divergence due to drift is related to the time of divergence and the effective population size as

$F_{st} = 1 - e^{-t/2N_e}$, where N_e is the effective population size and t is the number of non-overlapping generations since the populations were separated (Nei 1987). Using effective subpopulation size estimates of 200, 300, and 1000 plants for GS, BC, and SCER, respectively, and $F_{sp} = 0.084$, we estimate that the subpopulations of *C. springvillensis* have been isolated from each other for approximately 60 years. This should be considered an "order of magnitude estimate" owing to the uncertainty of the N_e estimates (precise estimation would require population size counts for every year since isolation), and the possibility that seed-bank carryover lengthens generation time beyond 1 year in this otherwise annual species.

We suggest that genetic drift is responsible for the divergence of subpopulations. The populations have not diverged substantially because they are aggregates of those subpopulations. The fluctuation in population and subpopulation size observed in *C. springvillensis* (California Department of Fish and Game 1991) would reduce effective population sizes and promote drift. The dominance of drift relative to gene flow is strongly suggested by the genetic distance analysis. Geographical distance was not a significant predictor of genetic distance in the North Fork populations (SCER and BC). Restricted seed and pollen flow among subpopulations should lead to isolation by distance—neighboring subpopulations would exchange more genes and hence be less divergent from each other than would more widely separated groups. The lack of correspondence between genetic and physical distance suggests they are diverging randomly with respect to each other.

Several ecological factors could have contributed to recent changes in the distribution of *C. springvillensis*. The Little Ice Age caused important climatic fluctuations throughout the world between AD 1500 and 1920 (Denton & Karlén 1973). Tree ring data from Visalia (45 km northwest of the populations examined here) indicate that precipitation averages were 10.5% lower annually and 20.6% lower during the winter rainy season (when *C. springvillensis* seeds germinate) between AD 1602 — 1904 compared with 1960—1991 (Fritts 1991; NOAA 1992). In the semi-arid valley grasslands and foothill woodland environment of *C. springvillensis* (approximately 258 mm of rain annually, NOAA 1992), changes in the amount and seasonality of rainfall could favor one congener over another. *Clarkia springvillensis* is a xerically adapted derivative of the widespread species *C. unguiculata* (Vasek 1977). *C. springvillensis* may never have proliferated beyond the Tule River drainage, however, its populations may have been more continuous during periods of lower rainfall. Perhaps only since the climate began to return to present day norms did *C. springvillensis* recede to its small enclaves.

The distribution of *C. springvillensis* has probably also been influenced by grazing with the introduction of cattle ranching in the 1860s. Ranchers also introduced several European grass species into California during the mid-1800s (e.g., *Bromus, Avena, and Festuca* spp. [Burcham 1981; Bartolome 1987]). The extant populations of *C. springvillensis* all occupy roadcuts and other unstable slopes with obviously lower grass density than adjacent sites. Hence, competitive interactions with annual grasses might play a critical role in subdividing the range of *C. springvillensis*. Decreased fire frequency during this century probably also reduced the availability of open habitats that *C. springvillensis* seems to require.

Efforts directed at maintaining the vigor of the extant populations should begin with experiments aimed at determining whether competition with annual grasses or shading by shrubs and trees limits *C. springvillensis'* survivorship or reproduction. If competition is important, then it will be critical to find out what levels of fire frequency and grazing intensity provide the appropriate amount of disturbance.

The results of this study have several important implications for conservation biologists as well as ecologists and evolutionary biologists. The use of several different measures of genetic variation on a hierarchical scale is valuable for drawing conclusions about the source of genetic structure in a species or a population. For instance, based solely on the low genetic differentiation among populations, we would have probably concluded that there was gene flow between populations. However, the smaller-scale genetic differentiation among subpopulations (which is not correlated with the geographical distance), suggests that gene flow is low and that differentiation is occurring via drift *at the subpopulation level*. The utility of F_{st} (or G_{st}) estimates at only one level of spatial scale should be considered. These estimates may be informative if the level examined is the level on which structure occurs. If this is not the case, incorrect conclusions concerning the mechanisms operating within species could be reached. Of practical consideration is the possibility of inaccurate or inadequate sampling of the species if a germplasm collection is established. Additionally, the patterns seen in *C. springvillensis* suggest that recent ecological history, rather than currently observable population processes, probably played the decisive role in the large-scale genetic structure of this species. Consideration of nonequilibrium hypotheses should be made for all species—especially when there is reason to believe that important climate or land use changes have occurred less than 100 generations ago.

Because *C. springvillensis* is a protected, endangered plant, the information gathered in this study should be useful in guiding management. The genetic diversity within populations is comparable to other species with similar life-history attributes. However, the subpopulations have diverged significantly, as have the populations along the North versus Middle Forks of the Tule River. The data suggest it is only a matter of time before

genetic drift further fragments the populations within these watersheds. Therefore, efforts to preserve several sites along each of the drainages of the Tule River is recommended. In particular, the GS population is important because it contains the most divergent subpopulation (GS 1) and the only subpopulation (GS 2) with unique alleles at the marker loci.

Acknowledgments

We appreciate the logistical help we received from D. Steeck, S. Morey, R. Goss, G. Martynn, T. Bowers, R. Galloway, T. Studley, and especially Rick Mitchell. We thank T. Cornine and C. Jungbluth for help with electrophoresis and plant care, and Becky Sherry and Charlie Hoyt for field assistance. We thank A. Ippolito, C. Howell, J. Kohn, and two anonymous reviewers for commenting on earlier drafts of the manuscript. Bruce Weir shared his expertise on the estimation and interpretation of hierarchical *F* statistics. J. Dietz provided the FORTRAN code to perform the matrix association tests, and M. McPeek calculated the hierarchical *F* statistics. Funding was provided by The Nature Conservancy's California Natural Heritage Division (thanks especially to Michael Golden) and by the University of Missouri Research Board Grant 93-060.

Literature Cited

Barbujani, G. 1987. Autocorrelation of gene frequencies under isolation by distance. Genetics **117**:777–782.

Bartolome, J. A. 1987. California annual grassland and oak savannah. Rangelands **9**:122–125.

Beardmore, J. A. 1983. Extinction, survival, and genetic variation. Pages 125–151 in C. M. Schonewald-Cox, S. M. Chamber, B. Macbryde, and L. Thomas, editors. Genetics and conservation: a reference for managing wild animal and plant populations. Benjamin/Cummings, Menlo Park, California.

Burcham, L. T. 1981. California rangelands in historical perspective. Rangelands **3**:95–104.

California Department of Fish and Game, Region 4. 1991. The fifth annual population studies of Kaweah Brodiaiea (*Brodiaea insignis*) and Springville Clarkia (*Clarkia springvillensis*). California Department of Fish and Game, Fresno, California.

Cavalli-Svorza, L. L., and A. W. F. Edwards. 1967. Phylogenetic analysis: models and estimation procedures. Evolution **32**:550–570.

Coates, D. J., and R. E. S. Sokolowski. 1992. The mating system and patterns of genetic variation in *Banksia cuneata* A.S. George (Proteaceae). Heredity **69**:11–20.

Crow, J. F., and K. Aoki. 1984. Group selection for a polygenic behavioral trait: estimating the degree of population subdivision. Proceedings of the National Academy of Science of the United States of America **81**:6073–6077.

Denton, G. H., and W. Karlèn. 1973. Holocene climatic variations—their pattern and possible cause. Quaternary Research **3**:155–205.

Dewey, S. E., and J. S. Heywood. 1988. Spatial genetic structure in a population of *Psychotria nervosa* I. Distribution of genotypes. Evolution **42**:834–838.

Dietz, E. J. 1983. Permutation tests for association between two distance matrices. Systematic Zoology **32**:21–26.

Dole, J. A., and M. Sun. 1992. Field and genetic survey of the endangered Butte County meadowfoam—*Limnanthes floccosa* subsp. *californica* (Limnanthaceae). Conservation Biology **6**:549–558.

Epperson, B. K. 1990. Spatial autocorrelation of genotypes under directional selection. Genetics **124**:757–771.

Epperson, B. K., and R. W. Allard. 1989. Spatial autocorrelation analysis of the distribution of genotypes within populations of lodgepole pine. Genetics **121**:369–377.

Felsenstein, J. 1982. How can we infer geography and history from gene frequencies. Journal of Theoretical Biology **96**:9–20.

Felsenstein, J. 1985. Phylogenies from gene frequencies: a statistical problem. Systematic Zoology **34**:300–311.

Felsenstein, J. 1989. PHYLIP—phylogeny inference package, version 3.2. Cladistics **5**:164–166.

Fritts, H. C. 1991. Reconstructing large-scale climatic patterns from tree ring data: a diagnostic analysis. University of Arizona Press, Tucson.

Gottlieb, L. D. 1981. Gene numbers in species of Asteraceae that have different chromosome numbers. Proceedings of the National Academy of Science of the United States of America **78**:3726–3729.

Gottlieb, L. D. 1984. Electrophoretic analysis of the phylogeny of the self-pollinating populations of *Clarkia xantiana*. Plant Systematics and Evolution **147**:91–102.

Gottlieb, L. D., and N. F. Weeden. 1979. Gene duplication and phylogeny in *Clarkia*. Evolution **33**:1024–1039.

Groom, M. J. 1995. Effects of isolation and patch size on the population dynamics of an annual herb, *Clarkia concinna*. Ph.D. dissertation. University of Washington, Seattle.

Hamrick, J. L., and M. J. W. Godt. 1990. Allozyme diversity in plant species. Pages 43–63 in A. H. D. Brown, M. T. Clegg, A. L. Kahler, and B. S. Weir, editors. Plant population genetics, breeding, and genetic resources. Sinauer Associates, Sunderland, Massachusetts.

Holtsford, T. P., and N. C. Ellstrand. 1989. Variation in outcrossing rate and population genetic structure of *Clarkia tembloriensis* (Onagraceae). Theoretical and Applied Genetics **78**:480–488.

Holtsford, T. P., and N. C. Ellstrand. 1992. Genetic and environmental variation in floral traits affecting outcrossing rate in *Clarkia tembloriensis* (Onagraceae). Evolution **46**:216–225.

Huenneke, L. F. 1991. Ecological implications of genetic variation in plant populations. Pages 31–44 in D. A. Falk and K. E. Holsinger, editors. Genetics and conservation of rare plants. Oxford University Press, New York.

Kimura, M., and G. H. Weiss. 1964. The stepping stone model of population structure and the decrease of genetic correlation with distance. Genetics **49**:561–576.

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

MacSwain, J. W., P. H. Raven, and R. W. Thorpe. 1973. Comparative behavior of bees and Onagraceae. IV. Clarkia bees of the western United States. University of California Publications in Entomology **70**:1–80.

Millar, C. I., and W. J. Libby. 1991. Strategies for conserving clinal, ecotypic, and disjunct population diversity in widespread species. Pages 149–170 in D. A. Falk and K. E. Holsinger, editors. Genetics and conservation of rare plants. Oxford University Press, New York.

Nei, M. 1972. Genetic distance between populations. American Naturalist **106**:283–291.

Nei, M. 1973. Analysis of gene diversity in subdivided populations. Proceedings of the National Academy of Sciences of the United States of America **70**:3321–3323.

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

National Oceanic and Atmospheric Administration. 1992. Climatography of the United states, no. 81. Monthly station normals of tem-

perature, precipitation, and heating and cooling degree days, 1961-1990. National Climatic Data Center, Ashville, North Carolina.

Richter, T. S., P. S. Soltis, and E. E. Soltis. 1994. Genetic variation within and among populations of the narrow endemic, *Delphinium viridescens* (Ranunculaceae). American Journal of Botany **81:** 1070-1076.

Schaal, B. A., and W. G. Smith. 1980. The apportionment of genetic variation within and among populations of *Desmodium nudiflorum*. Evolution **44:**214-221.

Schoen, D. J., and R. G. Latta. 1989. Spatial autocorrelation of genotypes in populations of *Impatiens pallida* and *Impatiens capensis*. Heredity **63:**181-189.

Slatkin, M. 1987. Gene flow and the geographic structure of natural populations. Science **236:**787-792.

Sokal, R. R., and F. J. Rohlf. 1981. Biometry. 2nd edition. Freeman, San Francisco.

Sokal, R. R., and N. L. Oden. 1978. Spatial autocorrelation in biology 1. Methodology. Biological Journal of the Linnean Society **10:**199-228.

Soltis, D. E., C. H. Haufler, D. C. Darrow, and G. J. Gastony. 1983. Starch gel electrophoresis of ferns: a compilation of grinding buffers, gel and electrode buffers, and staining schedules. American Fern Journal **73:**9-27.

Soltis, P. S., D. E. Soltis, and L. D. Gottlieb. 1987. Phosphoglucomutase gene duplications in *Clarkia* (Onagraceae) and their phylogenetic implications. Evolution **41:**667-671.

Van Treuren, R., R. Bijlsma, W. V. Delden, and N. J. Ouberg. 1991. The significance of genetic erosion in the process of extinction. I. Genetic differentiation in *Salvia pratensis* and *Scabiosa columbaria* in relation to population size. Heredity **66:**181-189.

Vasek, F. C. 1977. Phenotypic variation and adaptation in *Clarkia* Section *Phaeostoma*. Systematic Botany **2:**251-279.

Weir, B. S. 1990. Genetic data analysis. Sinauer Associates, Sunderland, Massachusetts.

Weir, B. S., and C. C. Cockerham. 1984. Estimating F-statistics for the analysis of population structure. Evolution **38:**1358-1370.

Wright, S. 1951. The genetical structure of populations. Annals of Eugenics **15:**323-354.