

EVOLUTION
INTERNATIONAL JOURNAL OF ORGANIC EVOLUTION
PUBLISHED BY
THE SOCIETY FOR THE STUDY OF EVOLUTION

Vol. 55

July 2001

No. 7

Evolution, 55(7), 2001, pp. 1269–1282

PERSPECTIVE:
EVOLUTION OF FLOWER COLOR IN THE DESERT ANNUAL *LINANTHUS PARRYAE*:
WRIGHT REVISITED

DOUGLAS W. SCHEMSKE^{1,2} AND PAULETTE BIERZYCHUDEK³

¹Department of Botany, Box 355325, University of Washington, Seattle, Washington 98195-5325
E-mail: schem@u.washington.edu

³Department of Biology, Box 53, Lewis and Clark College, Portland, Oregon 97219-7899
E-mail: bierzch@lclark.edu

Abstract.—*Linanthus parryae*, a diminutive desert annual with white or blue flowers, has been the focus of a long-standing debate among evolutionary biologists. At issue is whether the flower color polymorphism in this species is the product of random genetic drift, as Sewall Wright argued, or of natural selection, as proposed by Carl Epling and his colleagues. Our long-term studies of three polymorphic populations in the Mojave Desert demonstrate that flower color is subject to selection that varies in both time and space in its direction and magnitude. For all sites taken together, blue-flowered plants produced more seeds than white-flowered plants in years of relatively low seed production, whereas white-flowered plants had higher fitness in years of high seed production. Evidence of selection on flower color was found in two of the three study sites. Differences in fitness between the color morphs were sometimes large, with selection coefficients as high as 0.60 in some years. Our longest period of observations was at Pearblossom site 1, where plants reached appreciable densities in seven of the 11 years of study. Here we found significant differences in the seed production of the color morphs in six years, with three years of blue advantage and three years of white advantage. For all sites taken together, total spring precipitation (March and April) was positively correlated with both absolute and relative seed production of the color morphs. At Pearblossom site 1, blue-flowered plants typically had a fitness advantage in years of low spring precipitation, whereas white-flowered plants had a fitness advantage in years of high spring precipitation. This temporal variation in selection may contribute to the maintenance of the flower-color polymorphism at Pearblossom site 1, whereas gene flow from neighboring populations is proposed as the principal factor maintaining the polymorphism at the other study sites. We found no significant differences between the color morphs in pollinator visitation rate or in their carbon isotope ratio, a measure of water-use efficiency. Although the mechanism of selection remains elusive, our results refute Wright's conclusion that the flower color polymorphism in *L. parryae* is an example of isolation by distance, a key component of his shifting balance theory of evolution.

Key words.—Flower color, genetic drift, isolation by distance, natural selection, polymorphism, shifting balance theory, Sewall Wright.

Received August 9, 2000. Accepted April 4, 2001.

Linanthus had great significance for Wright as his first and best example of isolation by distance, a theory that Wright considered to be one of his most important contributions to evolutionary biology.

Provine (1986, p. 485)

The relative contribution of natural selection and random genetic drift to the evolution of adaptations has been the

subject of debate for decades (Endler 1986; Provine 1986; Coyne et al. 1997). Many of the early architects of the Modern Synthesis maintained that adaptive evolution occurred principally by natural selection within populations (Fisher 1930; Haldane 1932; Mayr 1942; Fisher and Ford 1947). This view was perhaps best exemplified by Fisher (1922), who believed that adaptation resulted from the fixation of many individually favorable mutations, each with small phenotypic effects. In contrast, Sewall Wright suggested that adaptation may often involve the stochastic process of genetic drift in addition to natural selection.

In Wright's shifting balance theory of evolution (Wright 1931, 1978, 1982), populations acquire favorable gene com-

² Present address: Department of Botany and Plant Pathology, Michigan State University, East Lansing, Michigan 48824-1312.

binations by a three-phase process involving both genetic drift and selection. In phase 1, genetic drift within demes decreases mean fitness, allowing them to descend adaptive peaks and to move across adaptive valleys. In phase 2, mass selection elevates demes to new, higher adaptive peaks; and in phase 3, interdeme selection allows demes with higher fitness to spread across the adaptive landscape. Phase 3 of the shifting balance has received the most attention (Crow et al. 1990; Wade and Goodnight 1991; Barton 1992; Kondrashov 1992), but recent theoretical work has focused on phase 1, and the potential role of drift in the formation of adaptive gene complexes (Goodnight 1995; Coyne et al. 1997).

Wright (1931) suggested that phase 1 of the shifting balance would occur most effectively in species subdivided into small, partially isolated demes. He proposed that restricted gene flow in continuously distributed populations was the principal means of attaining such population structure. Wright found support for this theory, which he termed "isolation by distance" (Wright 1943a), from field studies of a flower color polymorphism in the annual plant *Linanthus parryae*. These studies were initiated by Theodosius Dobzhansky and Carl Epling in 1941, and were continued by Epling and his colleagues over a period of more than 20 years (Epling and Dobzhansky 1942; Epling et al. 1960). According to Wright, this was "probably the most thoroughly studied case of a simple Mendelian polymorphism, within and among populations" (Wright 1978, p. 194).

Reporting the results from a single year of field observations, Epling and Dobzhansky (1942) at first concluded that the spatial pattern of flower color in *L. parryae* was best explained by random genetic drift, and this was supported by Wright's analysis of their data published soon thereafter (Wright 1943b). However, following many years of detailed field observations and experiments, Epling et al. (1960) rejected this earlier conclusion, and instead proposed that natural selection, not genetic drift, maintained the polymorphism. Wright again reanalyzed the data collected by Epling and his colleagues, and disagreed with their new interpretation, concluding once more that genetic drift was of primary importance (Wright 1978).

This wide disparity in opinions is reflected in the varied interpretations of the *L. parryae* studies presented over the years by other workers. Stebbins (1950) cited *L. parryae* as an example of isolation by distance. In contrast, Mayr (1965) concluded that the evidence collected by Epling et al. (1960) showed that the flower color polymorphism in *L. parryae* was maintained by selection. Grant (1981, p. 25) concluded that the spatial distribution of the color morphs "may well have developed as a result of restricted gene flow and neighborhood size, and without any role of environmental selection, although this is by no means certain." Waser and Price (1985) questioned Wright's conclusion that spatial patterns such as those exemplified by *L. parryae* are due to isolation by distance and suggested the alternative hypothesis of local selection. Finally, Levin (1988, p. 314) suggested that "one highly probable case for local random differentiation is described by the spatial distribution of flower polymorphism in the outcrossing annual *Linanthus parryae*."

Our research was motivated by this long-standing debate

over the importance of genetic drift and natural selection in *L. parryae* and by the pivotal role this species has played in the development of evolutionary theory (Provine 1986). We find that natural selection on flower color in *L. parryae* can be strong, but that the direction and intensity of selection vary in both time and space. Our studies suggest that the frequencies of blue- and white-flowered *L. parryae* are largely a product of temporal and spatial heterogeneity in local selection pressures—not of isolation by distance, as concluded by Wright (Wright 1943a, 1978).

We first describe the biology of *L. parryae*, then provide a chronological account of the nearly 40 years of research conducted on this species. This background allows us to place our approaches and interpretations in a historical context. We then outline the research objectives of our study.

Natural History of Linanthus parryae

Linanthus parryae (Polemoniaceae) is a diminutive (< 3 cm tall) annual native to the southern and western edges of the Mojave Desert in California (Patterson 1993). It has large flowers (1.0–1.5 cm long, 2 cm wide) and a few tiny, needlelike leaves, with calyces comprising most of the photosynthetic area. Epling and Dobzhansky (1942) described two flower morphs, blue and white, with the white form most common throughout the range. Crossing studies carried out by Epling and his colleagues provided evidence that flower color is controlled by a single gene, with blue dominant to white (Epling et al. 1960; Wright 1978), although the details of these experiments were never published (H. Lewis, pers. comm.). The blue flowers in some populations of *L. parryae* display considerable variation in intensity (e.g., Antelope Valley, see below), suggesting the presence of modifier genes.

Seed germination occurs after winter rains, plants flower in early to late April, and seeds mature in late May to early June. The seeds are passively dispersed, falling from the capsule as it dries. There is no germination in years with low rainfall (Epling et al. 1960; pers. obs.), and ungerminated seeds remain dormant in the soil. In wet years plants can be exceptionally dense, hence the common name "desert snow." Flower and seed number per plant are highly variable, with plants in dry years producing an average of one or two flowers and 10–30 seeds, whereas in good years reproductive output may increase by nearly 10-fold. The sole pollinator is the beetle *Trichochorus* sp. (Melyridae), and the flowers are self-incompatible (Epling et al. 1960; D. W. Schemske and P. Bierzychudek, pers. obs.).

Background to the Debate

The first research carried out on *L. parryae* was a 1941 census of flowering plants by Epling and Dobzhansky (1942). They recorded the numbers of blue- and white-flowered plants at four different stations distributed every 0.5 mi along a 200-mi grid of roads traversing the Mojave Desert. The white morph was more abundant overall; 78% of their samples consisted only of white-flowered individuals (data from table 1 of Epling and Dobzhansky 1942). For this reason, they designated certain geographical regions as "variable areas" and focused the rest of their analysis on these sam-

pling stations. Most of the samples in these variable areas contained at least some blue plants, and about 10% were monomorphic for blue. Based on the data for the variable areas, they found that the frequency distribution of the percentage blue plants was U-shaped, and concluded that this "resembles Wright's curves for the distribution of gene frequencies in effectively small populations" (Epling and Dobzhansky 1942, p. 331). This was consistent with Wright's (1931) expectation that in the absence of selection, mutation, and migration, alternate alleles in different populations will become fixed by random genetic drift.

Wright published his own analysis of Epling and Dobzhansky's data (Wright 1943b). He calculated F , "the amount of differentiation among groups taken at random from the whole" (Wright 1943b, p. 123), then estimated the effective population size that would be required to produce the observed F . Wright's classic paper "Isolation by Distance" (Wright 1943a) provided the theoretical development required for this analysis, and was published in tandem with his *L. parryae* paper (Wright 1943b). For the case of complete outcrossing, with the blue allele dominant to the white, Wright estimated an effective population size on the order of 14–25 individuals. He also calculated F for a hierarchy of spatial subdivisions, ranging from square feet to hundreds of square miles, and plotted F as a function of the size of the subdivision. Wright concluded that "The distribution of blue and white can be accounted for most easily by supposing that most of the differentiation of the smaller categories [subdivisions] is random in character" (Wright 1943b, p. 155).

Carl Epling maintained his interest in *L. parryae*, and in collaboration with Harlan Lewis and Francis Ball, initiated a long-term study of *L. parryae* in 1944 that resulted in the third publication on the biology of this species (Epling et al. 1960). Their work provided two major findings that challenged Wright's conclusions. First, they demonstrated that *L. parryae* seeds could remain viable in the soil for at least seven years, and probably much longer, thereby producing a large "effective breeding group" (Epling et al. 1960, p. 254). This result was based on an extraordinary field experiment in which flowering plants were counted, then removed from marked plots for seven successive years. Over this time period more than 16,000 plants were observed, all of which had emerged from the seed bank. Second, by establishing permanent census transects, they demonstrated that morph frequencies were far more stable than was expected from Wright's estimate of the effective population size. Their transect was 10 ft wide and 0.5 m long, divided into 260, 10 ft \times 10 ft quadrats. In each quadrat they recorded the number of blue- and white-flowered plants in the single square-foot sample with the highest plant density. They monitored this transect every year from 1944 to 1958, except for years when *L. parryae* was absent (two years) or rare (three years). They found considerable spatial variation in morph frequencies, but very little temporal variation, and concluded that "if genetic drift has played a role, it has been of only local consequence and not persistent in its effects" (Epling et al. 1960, p. 254). Instead, they hypothesized "an intense local selection because the blues are concentrated in certain areas and because persisting clines of blue and white frequencies have been found" (p. 254). Epling et al. acknowledged, how-

ever, that they saw no evidence that such selection was actually taking place. They concluded that the distribution of blue- and white-flowered plants was due to "selection operating at relatively low intensity or intermittently" (Epling et al. 1960, p. 254).

In 1978 Wright returned to the *L. parryae* problem, summarizing earlier results and adding some new analyses (Wright 1978). He reanalyzed the census data in Epling et al. (1960) and subsequent unpublished data collected by Epling and his colleagues. Wright created 52 groups, each comprised of five quadrats, and conducted χ^2 -tests of the change in morph frequency across five different time periods spanning 22 years. He showed that there were few significant changes, and that these were "widely scattered," leading him to the conclusion that "In the main, the changes were clearly of the nature of random genetic drift" (Wright 1978, p. 222).

Wright again calculated F -statistics from the data collected by Epling and Dobzhansky (1942), and plotted F_{IS} and F_{ST} as a function of the size of the region sampled. He concluded that "these curves are fairly typical of those expected where diversification at all levels is built up merely by isolation by distance" (Wright 1978, p. 203). Wright conceded that genetic drift could not explain the unexpectedly high degree of differentiation at the largest spatial scales: "It is barely possible that the profound differentiation of the four primary subdivisions is built up from random drift of neighborhoods. This area, containing millions of neighborhoods, is probably so great, however, that such a buildup would almost certainly be prevented by recurrent mutation" (Wright 1978, p. 223). Instead, he hypothesized that white-flowered plants had "a slight, general selective advantage," with selection coefficients in the range of $s = 0.01$ – 0.0001 , and that this was more likely due to "significant environmental differences among broad regions than among the many thousands of small areas that differ apparently at random in blue frequency" (Wright 1978, p. 223).

In summary, Wright believed that random genetic drift prevailed at the level of demes, and at all but the largest spatial scales, where selection played a modest role. This is at odds with Epling et al. (1960), who concluded that even small-scale differences in morph frequency were due to natural selection, not to genetic drift.

Research Objectives

We have conducted long-term field studies to investigate two questions concerning flower color evolution in *L. parryae*: (1) Is the polymorphism observed within populations selectively neutral? (2) Are the differences in flower color observed between populations due to spatially varying natural selection? Here we report on the evidence for selection in polymorphic populations. Our investigations of flower color evolution in nearly monomorphic populations will be presented in a later paper.

Whereas both Wright and Epling et al. (1960) relied on indirect approaches to reach their conclusions, we estimated selection on flower color directly, by measuring seed production of the two morphs. We made this comparison over a series of 11 years, because the maintenance of a polymorphism by directional selection requires, at the very least, that

the direction of selection fluctuate over time (Turelli et al. 2001). Thus, we determined whether the relative fitness of the two morphs varies from one year to another, and if so, what mechanisms might be responsible for any such variation.

In addition to measuring fitnesses of blue- and white-flowered morphs of *L. parryae* and how these values may change over time, we also investigated some possible agents of selection. We first asked if *Linanthus*'s pollinators might preferentially visit one flower color over another, and if so, whether the preferred morph changes over time. Evidence for pollinator preference in polymorphic species exists from other studies (Kay 1978; Brown and Clegg 1984; Stanton 1987; Stanton et al. 1989).

We also asked if flower color might have indirect effects on plant fitness. In several other California *Linanthus* spp. that produce both white and pigmented flowers, the pigmented morph is typically associated with stressful environments, for example serpentine soils or habitats with a rapid decline in soil moisture through the growing season (D. W. Schemske, pers. obs.). Pleiotropic effects of flower color are known from other species (Burdon et al. 1983; Schoen et al. 1984; Ernst 1987; Taylor and Jorgensen 1992; Rausher and Fry 1993). Of particular interest is the finding by Levin and Brack (1995) that white-flowered individuals of *Phlox drummondii*, a species that is in the same family (Polemoniaceae) as *Linanthus*, had lower fitness than pink-flowered morphs and that this fitness difference was due to pleiotropic effects. In a desert environment, the ability to tolerate soil moisture deficits is of obvious survival value. Thus, we asked whether plants with different colored flowers may also differ in their water use efficiency, such that soil moisture levels are experienced differently by the two morphs. To test this possibility, we analyzed carbon isotope ratios of dried plant tissue to infer the water use efficiency of the two morphs (Ehleringer 1988).

MATERIALS AND METHODS

Study Sites

Our work was conducted in two different localities north of the San Gabriel Mountains in Los Angeles County, California. The two Pearblossom sites, which are about 1.5 km apart, are both located along Avenue V, 5 km west of Pearblossom (elevation 900 m), between the Pearblossom Highway (CA 138) and 96th Street East; this is about midway between the towns of Pearblossom and Littlerock. This creosote bush desert has *L. parryae* populations that differ widely in the frequency of the two color morphs, including nearly monomorphic populations as well as the two polymorphic ones that we report on here. Pearblossom 1, the site further to the east, overlaps with the study sites used by Carl Epling and his colleagues in their studies (H. Lewis, pers. comm.; Epling and Dobzhansky 1942; Epling et al. 1960). This site maintains its variable nature for at least 0.5 km in all directions. In contrast, Pearblossom 2 is closely bounded by nearly monomorphic populations, with a mostly blue-flowered population 50 m from its western edge and a mostly white-flowered population 50 m from its eastern edge.

The Antelope Valley locality (elevation 900 m) is 3 km

south of the intersection of Lancaster Road and 200th Street West, west of the town of Lancaster, California, and 55 km WNW of the Pearblossom locality. This small, mixed population of *L. parryae* is not more than 500 m in diameter and is confined to the base of a narrow valley; the nearest monomorphic *L. parryae* population is at least 0.5 km away.

For most of the data we have collected, we provide three levels of analysis, at increasingly more restricted spatial scales. First, we combine all three populations and all available years to provide the largest possible sample and most general result; second, we present the combined data from the two Pearblossom populations, a smaller and more localized sample; and third, we provide an analysis for Pearblossom 1 only, the site that we have studied for the longest time period and the one that overlaps the area studied by Epling and his colleagues.

Morph Frequency, Density, and Reproduction

To measure morph frequency and density and to measure fitness of the two morphs, at each site we established three 50m × 1 m permanent transects placed roughly parallel to one another and approximately 7 m apart. In each year of the study, during the peak of flowering (mid–late April), we censused the numbers of flowering individuals of each flower color within each 1-m quadrat of every transect, 150 m² at each site. In years when plants were present, we placed individually numbered plastic stakes next to a subset of individuals from within these transects and collected their seeds. We marked plants in pairs, one blue and one white, that were nearest neighbors. Blue plants were always less frequent than white, so we typically marked every blue plant in each quadrat as well as that plant's nearest white neighbor. In years of low plant density, we occasionally marked plants just outside the quadrats to increase the sample size, but the distance between paired plants was never greater than 1 m and typically was on the order of a few centimeters. In years when plant density was particularly high and the number of blue plants per subquadrat (0.25 m²) was greater than five, we marked only five randomly selected pairs per subquadrat. We marked an average of 340 (range 88–521) plant pairs at a site each year.

We collected each marked plant once its seeds had matured (late May, early June) into an individual paper envelope. Each plant was examined under a dissecting microscope in the laboratory, where we counted the number of flowers and seeds produced. We then returned the seeds from each plant to the exact position where they were collected, still marked by its plastic stake. Seeds were returned as soon as possible after counting, usually in July of that year. In total, we counted 41,718 plants in our census transects and 713,364 seeds on marked plants.

We used the seed number of flowering plants as our measure of fitness in this outcrossing, hermaphroditic annual. Our failure to include the male component of fitness will doubtless cause some error in our estimates of the intensity of selection on flower color. We also compared flower number of the two morphs to attempt to understand the causes of any differences in seed production between them.

Although it is clearly possible that selection may operate

prior to flowering, our field observations support the findings of Epling et al. (1960) that most seedlings survive to flower. Because all withered plants remain standing, dry and brittle, for many months, evidence of mortality was not easy to overlook. Furthermore, the seeds from monomorphic populations of each morph that we germinated in the greenhouse showed no obvious differences in their germination rates (P. Bierzychudek and D. W. Schemske, pers. obs.). We attempted to compare the longevity of seeds of the two morphs in the seed bank, but this experiment was vandalized and yielded no useful information.

We compared flower number and seed number of blue- and white-flowered plants by Wilcoxon paired two-sample tests. Sequential Bonferroni tests (Rice 1989) were carried out separately for each site, with a tablewide α level of 0.05. We provide the statistical results for individual tests of site-year combinations both with and without the Bonferroni correction. Flower number could not be determined for plants that had experienced severe herbivory, so these plants were omitted from the analyses of flower number, but not from the seed number analysis. Plants that had flowered but produced no seeds, due either to herbivory or to fruiting failure, were included in the analyses of seed number. The seeds usually fell from the fruits once the plants were placed in the collecting envelope, so it was not possible to estimate the seed production of individual fruits within a plant. This prevented us from estimating the contributions of flower number per plant and of seed number per fruit to the variation observed in total seed number per plant. Sample sizes varied among years and sites, due primarily to temporal variation in plant density.

We used Fisher's combined probability test (Fisher 1954) to assess the overall significance of the statistical analyses conducted on flower and seed number. This method is used when the combined data are unlikely to meet the assumptions of more complex, parametric analyses such as two-way ANOVA (Sokal and Rohlf 1994). For each site, we calculated $-2 \sum \ln P_k$, where P_k is the probability level of the k th independent test of significance. The combined probability is distributed as χ^2 with $2k$ degrees of freedom (Sokal and Rohlf 1994).

Precipitation and Plant Performance

To determine whether there was an association between plant performance and precipitation levels, we identified the times of year during which precipitation was most strongly correlated with plant density and seed production. We did this by calculating the Spearman rank correlation coefficient between these variables and precipitation for each month of the growing season (January–April) separately, as well as for successive combinations of months (e.g., February + March, February + March + April, etc.). We used data for the period 1988–1998, but for the analysis of seed production, we excluded years in which flowering plants were absent or nearly so (1990, 1994, 1996, and 1997).

We obtained precipitation data from the Palmdale, California, recording station, located 18 km NW of the Pearblossom locality and 43 km ESE of the Antelope Valley locality. Although the Pearblossom recording station is closer

to our Pearblossom sites than the Palmdale station, its operation has been inconsistent, resulting in missing data for some months of the study. The precipitation recorded during the period of study at Palmdale and Pearblossom was highly correlated for the entire growing season (January–March, $r = 0.95$, $n = 9$, $P < 0.0001$) and for each individual month (January, $r = 0.96$, $n = 11$, $P < 0.0001$; February, $r = 0.99$, $n = 10$, $P < 0.0001$; March, $r = 0.90$, $n = 10$, $P < 0.0001$; April, $r = 0.83$, $n = 10$, $P < 0.01$).

Finally, desert precipitation is notoriously variable, so the 11 years of our study may not be representative of longer-term trends in precipitation. We compared the rainfall pattern observed during our study to that for the period from 1941 to 1987, that is, from the first year of Epling and Dobzhansky's original study until the year before our study. Comparisons were made for those monthly combinations of rainfall found to best explain the observed variation in plant density and plant performance.

Possible Mechanisms of Selection

Pollinator preferences

To determine whether pollinators preferentially visited blue- or white-flowered *L. parryae*, observers watched a randomly chosen plant for 30 sec, counting the number of beetle visits that took place during that time. At the end of the period, the observer moved to the nearest plant of the other color for an additional 30 sec. This sequence of alternating-color observations continued throughout the period that pollinator activity was most intense, between 1000 and 1500 h. Beetles did not appear to be disturbed by the presence of observers. It was usually impossible to follow beetles as they moved between plants, so we could not estimate the degree of flower-color preference of individual pollinators. All observations were made near the Pearblossom 2 site, because pollinator densities were consistently higher there than at the other sites. Pollinator observations were carried out during the peak of the blooming season, for 33.7 observer-hours in 1991 and 74.0 observer-hours in 1992. Because we observed no evidence of heterogeneity in the visitation pattern based on date or time of day, we pooled the observations from each year. We used a *G*-test with one degree of freedom to determine if the number of visits observed was different from the null expectation of equal visitation to the two morphs.

Water-use efficiency

A number of studies have demonstrated that the relative ratio of $^{13}\text{C}/^{12}\text{C}$, typically expressed as $\delta^{13}\text{C}$, provides an indirect measure of plant water use efficiency (WUE; Ehleringer et al. 1987; Farquhar et al. 1988; Martin and Thorstenson 1988). Plants with high WUE possess higher (more positive) carbon isotope ratios, a relationship that results from the strong discrimination by the initial carboxylating enzyme ribulose bisphosphate carboxylase against heavier carbon isotopes (Ehleringer 1988). To determine if flower color in *L. parryae* has a pleiotropic effect on WUE, we compared the carbon isotope ratios ($\delta^{13}\text{C}$) for blue- and white-flowered plants collected from Pearblossom 1 in 1998. The expense of these analyses required that we create pooled samples of

TABLE 1. Density and frequency of blue- and white-flowered *Linanthus parryae* at the three study sites in the Mojave Desert. Standard deviation for percentage blue calculated according to Snedecor and Cochran (1967).

Site	Year	Number of plants ¹			Density (plants/m ²)	% blue (SD)
		Blue	White	Total		
Pearblossom 1	1988	371	2383	2754	18.4	13.5 (0.65)
	1989	32	201	233	1.6	13.7 (2.25)
	1990	0	0	0	0	
	1991	138	732	870	5.8	15.9 (1.24)
	1992	479	2789	3268	21.8	14.7 (0.62)
	1993	200	1726	1926	12.8	10.4 (0.70)
	1994	2	5	7	0.047	28.6 (17.08)
	1995	1153	8818	9971	66.5	11.6 (0.32)
	1996	~0	~0	~0	~0	
	1997	0	0	0	0	
	1998	308	3124	3432	22.9	9.0 (0.49)
Pearblossom 2	1991	596	902	1498	10.0	39.8 (1.26)
	1992	603	1074	1677	11.2	36.0 (1.17)
	1993	184	424	608	4.1	30.3 (1.86)
	1994	0	0	0	0	
Antelope Valley	1990	336	1244	1580	10.5	21.3 (1.03)
	1991	520	2735	3255	21.7	16.0 (0.64)
	1992	490	4925	5415	36.1	9.0 (0.39)
	1993	1185	4039	5224	34.8	22.7 (0.56)

¹ Pooled date from three, 1 m × 50 m transects at each site.

individuals having the same flower color. We paired a pooled sample of 10 blue individuals with another pooled sample of 10 white individuals. Each of the 10 white individuals was the nearest neighbor of one of the 10 blue individuals in its paired sample. The 10 pairs of blue- and white-flowered individuals were collected from within a few meters of one another within one of the transects. This process was repeated six times from each of the three transects, for a total of 18 pairs of pooled samples. By pooling tissue from 10 different plants to construct each sample, we reduced the variation among samples and increased the precision in estimating the mean. Each sample was prepared for analysis by grinding the roots, stems, and leaves of individual plants to a fine powder. Flowers and fruits were not included. The carbon isotope ratio of plant tissue was determined by the Stable Isotope Ratio Facility for Environmental Research at the University of Utah (<http://ecophys.biology.utah.edu/sirfer>). We used paired *t*-tests to compare the carbon isotope ratios of the two morphs.

RESULTS

Plant Density and Morph Frequency

Each of our three study sites displayed striking temporal variation in plant density (Table 1). For example, at Pearblossom 1, mean annual plant density ranged from zero to > 60 plants/m² over the 11 years of observation. Mean densities at Antelope Valley varied less from year to year than at the Pearblossom sites, and moderate flowering plant densities were found at this site in 1990, a year when Pearblossom 1 did not produce any plants.

The frequencies of the blue and white morphs, in contrast, displayed much less temporal variation. At Pearblossom 1, including only the seven years with a significant density of plants, the frequency of the blue morph varied from 9.0 ± 0.5% (SD) in 1998 to 15.9 ± 1.2% in 1991 (Table 1). At

Antelope Valley, morph frequency fluctuated somewhat more; here the frequency of the blue morph varied from 9.0 ± 0.4% in 1992 to 22.7 ± 0.6% in 1993 (Table 1). At both Pearblossom sites the frequency of blue-flowered plants declined steadily from 1991 to 1993.

Plant Fitness

Mean flower production for all sites and years is given in Table 2. Individual tests of significance revealed that blue-flowered plants produced significantly more flowers than whites at Pearblossom 1 in 1988 and at Antelope Valley in 1990, whereas white-flowered plants produced significantly more flowers than blues at Pearblossom 1 in 1998 and at Pearblossom 2 in 1991. All but the 1990 comparison at Antelope Valley remain significant after the Bonferroni correction (Table 2). There was considerable temporal variation in flower number at all sites, although the range of mean values at Pearblossom (1.2–15.0 flowers/plant) was much greater than that at Antelope Valley (1.5–4.9 flowers/plant). Fisher's combined probability test revealed a significant overall effect of flower color on flower number at Pearblossom 1 ($\chi^2 = 41.2$, df = 14, $P < 0.0001$), but not at Pearblossom 2 ($\chi^2 = 11.0$, df = 6, $P = 0.090$) or at Antelope Valley ($\chi^2 = 14.6$, df = 8, $P = 0.071$).

Individual tests of significance revealed that the seed production of the two color morphs was significantly different in six of the seven years that plants flowered at Pearblossom 1 (Table 3; Fig. 1), and in one of three years at Pearblossom 2 (Table 3; Fig. 2A). At Pearblossom 1, blue-flowered plants produced significantly more seeds than whites in 1988, 1992, and 1993, whereas whites produced significantly more seeds than blues in 1991, 1995, and 1998 (Table 3; Fig. 1). After the Bonferroni correction, three of the comparisons of mean seed number at Pearblossom 1 remain significant (1988, 1993, and 1995; Table 3). White-flowered plants at Pear-

TABLE 2. Flower production by blue- and white-flowered plants of *Linanthus parryae* in the Mojave Desert. The periods of observation at each site were: Pearblossom 1, 1988–1998; Pearblossom 2, 1991–1994; Antelope Valley, 1990–1993 (Table 1). Within these periods, plants were not sampled in years of low density (Table 1). Probability levels in bold face identify comparisons that are significant after sequential Bonferroni tests conducted for each site (tablewide α -level of 0.05).

Site	Year	Overall mean ²	Number of flowers per plant				Statistical comparisons of the two morphs ¹		
			Blue		White		Z	N ³	P
			Mean	SD	Mean	SD			
Pearblossom 1	1988	2.08	2.53	1.67	2.01	1.37	5.62	430	<0.0001
	1989	1.23	1.28	0.55	1.22	0.50	1.05	100	0.29
	1991	14.97	14.23	10.78	15.11	10.79	1.12	322	0.26
	1992	2.96	3.15	2.09	2.93	2.08	1.21	138	0.23
	1993	1.28	1.40	0.76	1.27	0.75	1.12	62	0.26
	1995	4.38	4.41	2.93	4.37	2.67	0.03	416	0.98
	1998	7.93	7.06	4.37	8.01	4.91	3.02	334	0.0026
Pearblossom 2	1991	8.35	7.92	4.61	8.64	5.17	2.55	451	0.0107
	1992	3.96	4.07	3.46	3.90	2.93	0.40	145	0.69
	1993	1.66	1.69	1.31	1.65	0.97	0.60	68	0.55
Antelope Valley	1990	3.23	3.80	3.18	3.08	2.13	2.25	88	0.0244
	1991	4.87	4.88	4.48	4.87	5.27	0.87	374	0.39
	1992	4.53	5.03	4.87	4.48	3.41	1.35	175	0.18
	1993	1.49	1.46	0.90	1.50	0.79	0.86	294	0.39

¹ Wilcoxon paired one-sample tests.

² Calculated by weighting mean flower number of each morph by its frequency (Table 1).

³ Number of sample pairs: one blue-flowered and one white-flowered plant per pair.

blossom 2 produced significantly more seeds than blues in 1991. In 1992 the difference in seed production at this site was marginally significant, with whites producing more seeds than blues (Table 2; Fig. 2A). The magnitude of these differences was often quite large (Figs. 1, 2), with coefficients of selection against the disadvantaged morph as high as 0.60 in some years (Table 3). By contrast, at Antelope Valley there were no significant differences in seed number between morphs (Table 3; Fig. 2B). Fisher's combined probability test revealed a significant overall effect of flower color on seed number at Pearblossom 1 ($\chi^2 = 62.2$, df = 14, $P < 0.0001$) and at Pearblossom 2 ($\chi^2 = 25.3$, df = 6, $P < 0.001$), but not at Antelope Valley ($\chi^2 = 4.9$, df = 8, $P = 0.77$).

To determine if there was significant temporal variation in relative seed production, for each site-year combination we counted the number of sample pairs in which blue-flowered plants produced more seeds than whites and in which white-flowered plants produced more seeds than blues. We then conducted a χ^2 -analysis for each site to determine if the two categories of relative seed production (blue > white, white > blue) were independent of year. At Pearblossom 1, analysis of the 1880 sample pairs (ties not included) censused over seven years revealed that the relative seed production of blue- and white-flowered plants varied significantly across years ($\chi^2 = 41.5$, df = 6, $P < 0.0001$). In contrast, there was no significant temporal variation in relative seed production at

TABLE 3. Seed production by blue- and white-flowered plants of *Linanthus parryae* in the Mojave Desert. Sampling periods are explained in Table 2. Selection coefficients, given for years/sites where there was a significant difference in seed production between morphs, are calculated as $1 - (w_1/w_2)$, where w_1 is the disadvantaged morph and w_2 is the morph having higher seed production. Probability levels in bold face identify comparisons that are significant after sequential Bonferroni tests conducted for each site (tablewide α -level of 0.05).

Site	Year	Overall mean ²	Number of seeds per plant				Selection coeff. (s)	Statistical comparisons of the two morphs ¹			
			Blue		White			Z	N ³	P	
			Mean	SD	Mean	SD					
Pearblossom 1	1988	21.12	30.45	38.17	19.67	26.69	0.35	4.74	430	<0.0001	
	1989	3.99	5.00	8.99	3.83	7.23		0.59	131	0.550	
	1991	295.71	260.13	235.91	302.42	259.88	0.14	2.29	322	0.022	
	1992	14.06	18.58	40.26	13.28	35.49	0.29	2.12	359	0.034	
	1993	1.29	2.77	11.80	1.12	6.04	0.60	2.70	270	0.007	
	1995	72.46	65.62	72.88	73.35	73.69	0.11	2.74	463	0.006	
	1998	124.29	108.87	100.12	125.81	123.97	0.14	2.37	465	0.018	
Pearblossom 2	1991	225.20	198.17	156.50	243.06	195.32	0.18	4.19	451	<0.0001	
	1992	19.32	16.58	38.72	20.86	39.10	0.21	1.85	406	0.064	
	1993	4.54	4.29	16.41	4.65	14.05		0.67	262	0.500	
Antelope Valley	1990	38.11	43.18	62.54	36.74	45.42		0.04	88	0.970	
	1991	91.77	88.79	116.32	92.34	121.36		0.30	374	0.760	
	1992	45.59	49.45	67.11	45.21	68.41		1.14	244	0.250	
	1993	9.07	8.99	17.99	9.09	15.58		0.70	521	0.480	

¹ Wilcoxon paired one-sample tests.

² Calculated by weighting mean seed number of each morph by its frequency (Table 1).

³ Number of sample pairs: one blue-flowered and one white-flowered plant per pair.

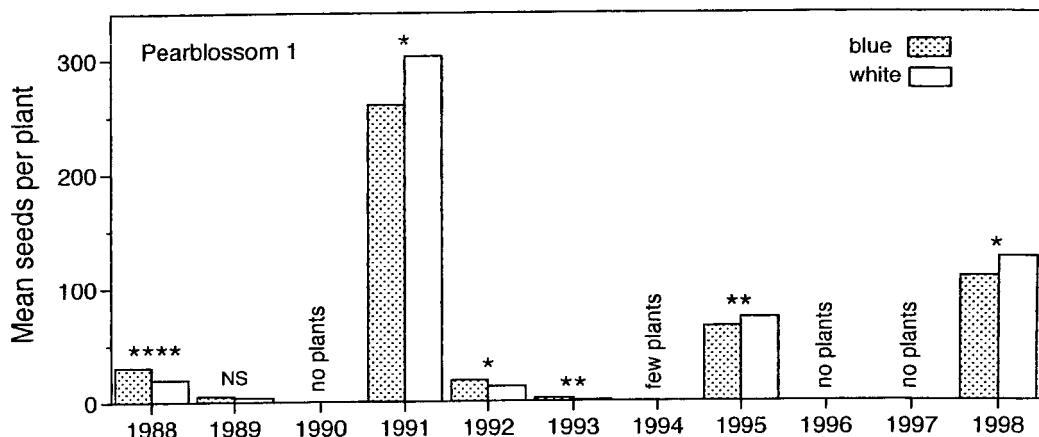


FIG. 1. Mean seeds per plant for white-flowered and blue-flowered *Linanthus parryae* at Pearblossom 1 over the years 1988–1998. Means, standard deviations, and sample sizes are given in Table 3. NS, $P > 0.05$; * $P < 0.05$; ** $P < 0.01$; *** $P < 0.0001$.

Pearblossom 2 ($\chi^2 = 3.27$, df = 3, $P = 0.20$) or at Antelope Valley ($\chi^2 = 3.00$, df = 4, $P = 0.39$).

There was considerable variation in mean seed production per plant at all sites, as calculated by weighting each morph's mean seed production by its frequency (Table 3; Figs. 1, 2A,

2B). For example, at Pearblossom 1 there was a 200-fold range of variation in seed production over the seven years with flowering plants, from a mean of 1.3 seeds/plant in 1993 to nearly 300 seeds/plant in 1991. These differences prompted us to ask if there was a relationship between the relative fitness of the two morphs (mean seeds per plant for blue-flowered plants/mean seeds per plant for white-flowered plants) and the quality of the year, as measured by the mean seed production of plants that year. We found significant negative correlations between relative fitness and mean seed production for all sites and years combined ($r_s = -0.62$, $N = 14$, $P < 0.05$; Fig. 3), for the combined data from the two Pearblossom sites ($r_s = -0.66$, $N = 10$, $P < 0.05$), and for Pearblossom 1 alone ($r_s = -0.86$, $N = 7$, $P < 0.05$). These results suggest that variation in precipitation among years might explain temporal variation in relative fitness of the two morphs.

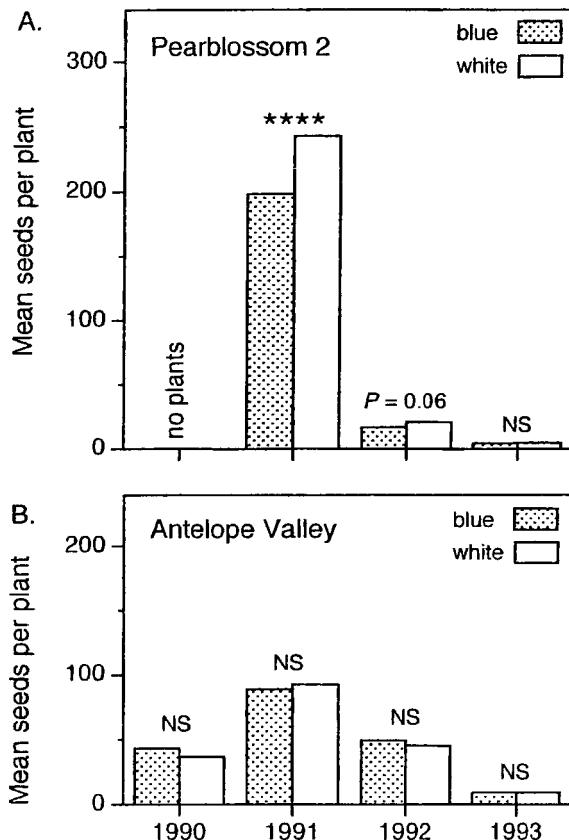


FIG. 2. Mean seeds per plant for white-flowered and blue-flowered *Linanthus parryae*. Means, standard deviations, and sample sizes are given in Table 3. NS, $P > 0.05$; *** $P < 0.0001$. Significance levels for marginally significant comparisons are given above the bars. (A) Pearblossom 2 over the years 1990–1993. (B) Antelope Valley over the years 1990–1993.

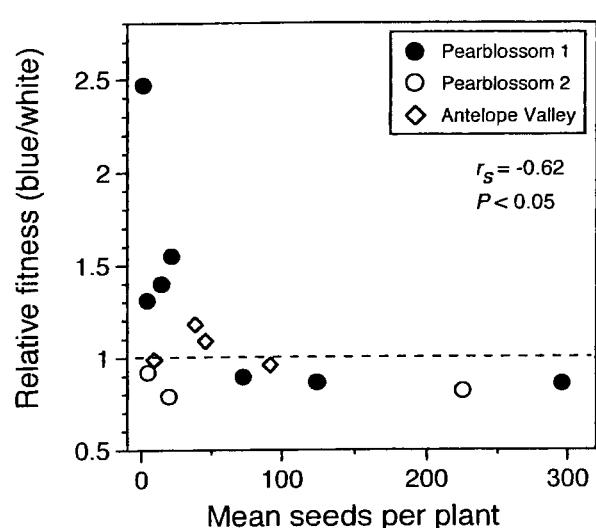


FIG. 3. Mean relative fitness (mean seeds per plant for blue-flowered plants/mean seeds per plant for white-flowered plants) as a function of mean seed number per flowering plant. Data from the same sites and years as Figures 1, 2A, B. The dashed line represents equal fitness for both morphs.

Association between Plant Density, Fitness, and Precipitation

Both in our study and that of Epling et al. (1960), there were some years when flowering plants were absent or rare. In these years, winter precipitation levels were probably too low to trigger germination. In the greenhouse, *Linanthus* seeds could only be induced to germinate by simulating protracted periods of rainfall; simply keeping seeds moist was insufficient (P. Bierzychudek, pers. obs.). Plant density in our census transects was most highly correlated with precipitation in January (for all sites and years combined, $r_s = 0.66$, $N = 19$, $P = 0.006$), typically the month when seeds germinate in the field. We observed the highest plant density at Pearblossom 1 in 1993, which had the wettest January in the last half century (Fig. 4A), with 13.6 times more precipitation (19.1 cm) than the mean for the four years with few or no plants ($\bar{x} = 1.4$ cm).

To investigate the possibility that temporal variation in precipitation is a cause of the significant negative relationship we observed between absolute and relative seed production (Fig. 3), we sought to identify the precipitation pattern most highly correlated with plant fitness. We found that the combined precipitation in March and April gave the highest correlation with mean seed production for all sites combined ($r_s = 0.80$, $N = 14$, $P = 0.004$; Fig. 5A), for the two Pearblossom sites ($r_s = 0.88$, $N = 10$, $P = 0.009$), and for Pearblossom 1 ($r_s = 0.86$, $N = 7$, $P = 0.036$). March–April rainfall was also significantly correlated with relative fitness (blue seed production/white seed production) for all sites combined ($r_s = -0.60$, $N = 14$, $P = 0.032$; Fig. 5B) and for the combined data for the two Pearblossom sites ($r_s = -0.69$, $N = 10$, $P = 0.038$). With increasing rainfall, the seed production of the blue morph decreased relative to that of the white morph. Even for Pearblossom 1 alone, the correlation between relative fitness and March–April rainfall was negative and marginally significant ($r_s = -0.71$, $N = 7$, $P = 0.080$).

The precipitation patterns we observed during our study were comparable to longer-term trends (Figs. 4A, B). For the period 1941–1987, the median January precipitation at Palmdale was 2.49 cm (mean = 3.76 cm, SD = 4.1 cm, range 0.00–15.9 cm), whereas the median for the period 1988 to 1998 was 2.79 cm (mean = 4.91 cm, SD = 5.9 cm, range 0.7–19.1 cm; Fig. 4A). These differences were not statistically significant (Mann-Whitney U -test, $Z = 0.73$, $df = 56$, $P = 0.463$). Plants were rare or absent in four of the 11 years of study at Pearblossom 1 (Table 1; Fig. 4A), a frequency (36%) nearly identical to that observed by Epling and colleagues (4/13 = 31%; Epling and Dobzhansky 1942; Epling et al. 1960). For the period 1941–1987, the median March–April precipitation at Palmdale was 3.86 cm (mean = 4.79 cm, SD = 4.34 cm, range 0.0–18.6 cm), and for 1988–1998 the median March–April precipitation was 2.52 cm (mean = 3.79 cm, SD = 3.50 cm, range 0.2–10.6 cm; Fig. 4B). Again, these differences were not statistically significant (Mann-Whitney U -test, $Z = 0.63$, $df = 55$, $P = 0.531$). March–April precipitation patterns during the two time periods are also similar when we exclude those years in which January precipitation was so low that germination was unlikely.

Mechanisms of Selection

On average, white-flowered plants had the advantage in good years, whereas blue-flowered plants had higher fitness in poor years. The mechanisms responsible for this temporal variation in the relative fitness of the two morphs could operate at a variety of life stages. For example, the morphs might vary in their ability to produce large numbers of flowers, in the percentage of their flowers that mature and produce seeds, or in their seed production per flower. These different possibilities suggest different likely mechanisms of selection.

We first investigated the relationship between seed production and flower production of the two morphs and found that variation among sites and years in mean seeds per plant was due largely to variation in mean flower number. The correlation between mean seed number per plant and mean flower number per plant was significant for all sites combined ($r_s = 0.95$, $N = 14$, $P < 0.001$), for the combined data for the two Pearblossom sites ($r_s = 0.95$, $N = 10$, $P < 0.01$), and for Pearblossom 1 ($r_s = 0.93$, $N = 7$, $P < 0.05$). We also found a significant, positive correlation ($P < 0.0001$) between flower and seed number per plant for each of the 14 site-year combinations.

In three of the four site-year combinations with a significant difference in flower production between the two morphs (Table 2) we also observed a significant difference in seed production (Table 3). For 10 of the 14 site-year combinations, the morph with the highest flower production had the highest seed production. We also found a significant positive correlation between the relative fitness of the two morphs and their relative flower production (mean flowers per plant for blue-flowered plants/mean flowers per plant for white-flowered plants) for all sites and years combined ($r_s = 0.75$, $N = 14$, $P < 0.01$), for the combined data from the two Pearblossom sites ($r_s = 0.77$, $N = 10$, $P < 0.01$), and for Pearblossom 1 ($r_s = 0.93$, $N = 7$, $P < 0.05$). These results suggest that the differences in seed production between color morphs are due largely to mechanisms that influence flower production per plant, rather than to factors influencing seed production per flower.

Pollination

The beetle *Trichochorus* sp. (Melyridae) was the only flower visitor. We observed 131 pollinator visits in 1991 and 271 visits in 1992. In each year pollinators exhibited slightly higher visitation to white-flowered plants (55.7% of visits in 1991; 51.7% of visits in 1992), but this was not significantly different from the null expectation in either year (1991: $G = 1.72$, $df = 1$, $P = 0.20$; 1992: $G = 0.30$, $df = 1$, $P = 0.59$).

Water-use efficiency

We found no significant differences between morphs in their carbon isotope ratios. The mean WUE of the pooled blue-flowered samples was $-27.24 \pm 0.32\%$, whereas the mean of the white-flowered samples was $-27.13 \pm 0.30\%$ (paired t -test, $n = 18$, $P = 0.19$).

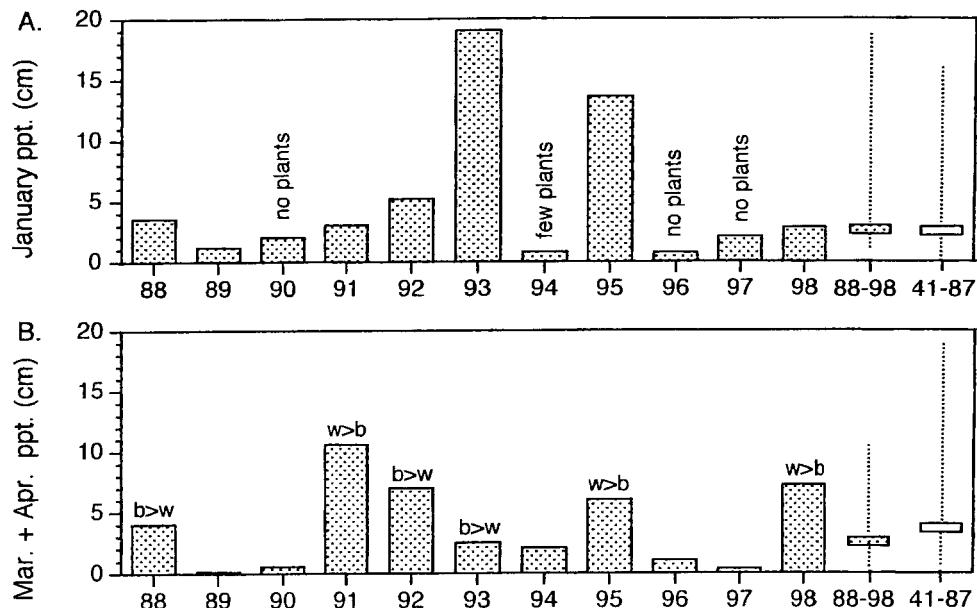


FIG. 4. Precipitation patterns for Pearblossom, California, during the course of the present study. (A) January precipitation levels in centimeters for each year of the present study. Years when flowering plants were absent or rare at Pearblossom 1 are indicated by text above the bars. Shown at the far right are the median and range for the 11 years of the present study (1988–1998) and for 1941–1987, a period that begins with the year of Epling and Dobzhansky's (1942) field study and ends in the year preceding our study. (B) Combined March–April precipitation levels in centimeters for each year of the present study. Shown at the far right are the median and range for these 11 years and the median and range for the period 1941–1987 (see above). A comparison of the mean seed production by the two morphs at Pearblossom 1 (b, blue; w, white) is summarized for each year with a significant density of flowering plants (data from Table 3). Years when one morph produced significantly more seeds per plant than the other are indicated by text above the bars.

DISCUSSION

Sewall Wright believed that flower color in *L. parryae* was a nearly neutral character (Wright 1978). He found evidence of only slight selective differences between the morphs ($s \leq 0.01$), and these were manifest only at the largest spatial scales. The effective population size that he estimated for this species was small enough for random genetic drift to override such a small selective difference between the morphs (Wright 1978). He proposed that the combination of weak selection and small effective population size in *L. parryae* would result in isolation by distance, that is, the loss of genetic variation within populations and the evolution of genetic differentiation among populations. The spatial pattern of flower color observed in *L. parryae* was thought to conform to these predictions (Provine 1986).

In contrast, our data demonstrate that flower color in *L. parryae* is sometimes subject to strong selection. At Pearblossom 1, the average selection coefficient (s) was 0.23 (including all years with flowering plants; Table 3). Although it is theoretically possible for genetic drift to override selection of this magnitude, it would require an extremely small effective population size. The conditions for genetic drift to override selection, $N_e s < 1$, are satisfied at Pearblossom 1 only if the effective population size (N_e) is less than 4.4. We do not have an estimate of the effective population size in *L. parryae*, so we cannot determine precisely if genetic drift could override the selection observed in our studies. Nevertheless, as suggested by Epling et al. (1960), it seems very unlikely that the effective population size of *L. parryae* is small enough to negate the effects of selection. Its long-lived

seed bank, exclusively outcrossing mating system, and high flowering plant densities (up to 67 plants/m² at Pearblossom 1; Table 1) are not characteristics typically associated with low effective population size (Crawford 1984).

Spatial and Temporal Variability in Selection

At Pearblossom 1, where from 1988 to 1998 we observed seven *Linanthus* flowering seasons, blue-flowered plants produced significantly more seeds than white-flowered plants in three years, and white-flowered plants outproduced blue plants in three years. At Pearblossom 2, we found that white-flowered plants produced significantly more seeds than blue-flowered plants in two of three years. At Antelope Valley, we never found a significant difference between morphs in seed production. Thus, only at Antelope Valley are we unable to reject Wright's contention that flower color is neutral and subject only to the stochastic process of genetic drift.

The *L. parryae* population at Antelope Valley behaved somewhat differently from those at Pearblossom, fluctuating less in density from year to year. In 1990, when no plants were observed at Pearblossom 1, plant density at Antelope Valley was reasonably high (Table 1). Plants at Antelope Valley were smaller and produced fewer flowers and seeds, and differences in seed production between the two morphs were not as great as at Pearblossom. We also observed that the intensity of color in blue flowers at Antelope Valley was more variable than at Pearblossom, where the blue and white morphs were strikingly different. These subtler differences in color were associated with correspondingly smaller differences in performance at Antelope Valley.

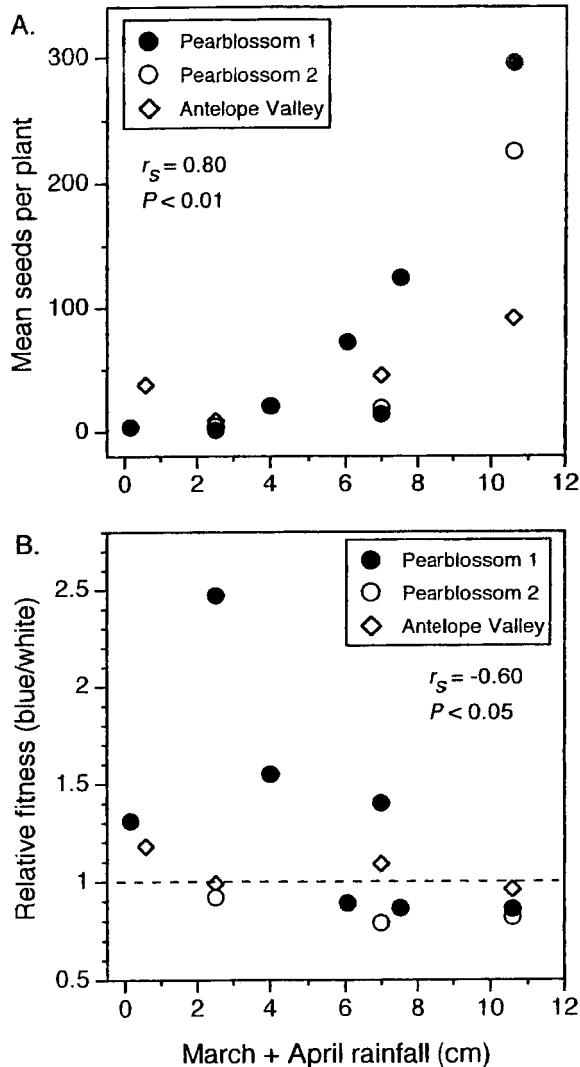


FIG. 5. (A) Mean seeds per plant, with both color morphs pooled, plotted as a function of combined March and April rainfall. (B) Relative fitness (mean seeds per plant for blue-flowered plants/mean seeds per plant for white-flowered plants) plotted as a function of combined March and April rainfall. The dashed line represents equal fitness for both morphs.

For all populations taken together, we observed a negative correlation between relative fitness and mean seed production of the morphs and a significant positive correlation between mean seed production and March–April precipitation. Thus, the performance of blue- and white-flowered plants appears to be determined by some temporally varying environmental factor, most likely precipitation. This is perhaps best illustrated by the seed production data from Pearblossom 1, where the mean March–April precipitation for the three years of white advantage was nearly twice that for the years of blue advantage (7.96 cm vs. 4.50 cm; Fig. 4B). Four years of our study (1991, 1993, 1995, and 1998) were classified as El Niño, the tropical climatic phenomenon that often results in a marked increase in winter precipitation along the western Pacific (Trenberth 1997; N. Mantua, pers. comm.). El Niño occurs in this region every three to 12 years (Quinn 1987),

and may represent a major cause of the temporal variation in selection we witnessed at Pearblossom 1. The 1993 El Niño produced high January precipitation but low March–April precipitation at Pearblossom 1, thus the conditions in this year favored blue-flowered plants (Figs. 4A, B). In contrast, each of the remaining El Niño years produced high March–April precipitation, and these were also the only years of white advantage at Pearblossom 1 (1991, 1995, and 1998; Fig. 1).

Mechanisms of Selection

Levels of spring rainfall appear to have very different effects on the fitness of the two color morphs in our study populations. In particular, at Pearblossom 1 the relative seed production of blue-flowered plants was highest in drier years, whereas that of white-flowered plants was highest in wetter years. This suggests an interaction between flower color genotype and environment.

Pollinators could cause selection on flower color if they visited one color morph more frequently than another, yet we found no significant color preference in either of our two years of pollinator observations. We measured pollinator visitation to the two morphs in 1991, a year of very high seed production and white advantage at Pearblossom 1, and in 1992, a year in which this site displayed much lower seed production and blue advantage. Pollinators in both years visited the two morphs at nearly equal rates; therefore it does not appear that the temporal variation in selection that we observed at Pearblossom 1 can be explained by changes in pollinator preference. Because the frequencies of the two morphs did not change appreciably between the two years in which we observed pollinators (1991, 1992; Table 1), we cannot rule out the possibility that pollinators might demonstrate frequency-dependent color preferences, a behavior that could contribute to the maintenance of the polymorphism. We suggest that pollinators probably do not contribute to selection on flower color in *L. parryae*, but caution that more detailed study is warranted. For example, it would be of interest to determine if pollinator foraging behavior differs between the color morphs, due perhaps to morph-specific variation in pollen production, and if this could cause differences between morphs in their male or female reproductive success.

Factors influencing plant growth are the most likely explanations for the observed temporal differences in plant size. We found no evidence that the two morphs differ in WUE, as indicated by their carbon isotope ratios. The substantial cost of these analyses required that we limit our investigation to only a single year (1998) in just one site (Pearblossom 1). This was a year of high spring precipitation, so it would be useful to have a comparative sample from a drier year. Because of the limited sampling and lack of replication across sites and years, the finding of no difference between the morphs in WUE should be viewed with caution.

There are countless other possible pleiotropic effects of flower color that may explain the fitness differences of the two morphs. Recent findings suggest that white-flowered plants possess a higher tissue concentration of potentially toxic cations such as magnesium and selenium and that pop-

ulation differentiation for flower color is associated with spatial variation in soil chemistry (D. W. Schemske, unpubl. data). These surprising results suggest a hypothesis to explain the relationship between relative fitness and precipitation that we observed at Pearblossom 1 (Fig. 1). In years with low soil moisture, the tissue concentrations of detrimental cations in white-flowered plants may become high enough to reduce their fitness, whereas blue-flowered plants are less adversely affected, by virtue of their reduced cation uptake. In years with high soil moisture, detrimental cations become diluted, and blue-flowered plants may display lower fitness because their mechanism of cation exclusion also reduces the uptake of favorable minerals such as calcium and phosphorus. A similar mechanism could explain the marked spatial variation in flower color observed by Epling and Dobzhansky (1942) and Epling et al. (1960). Experimental studies are now underway to test this hypothesis in *L. parryae* and in other congeners that display a similar flower-color polymorphism.

Maintenance of the Polymorphism

Wright believed that random genetic drift was the principle mode of evolution within *L. parryae* populations and that this process was a sufficient explanation for the spatial distribution of flower color. The geographical surveys conducted by Epling and Dobzhansky (1942) revealed that polymorphic populations of *L. parryae* were rather rare. According to this argument, polymorphic populations should quickly become fixed for alternate alleles, and this would result in spatial differentiation for flower color.

One possible alternative to this explanation is that temporally varying natural selection is actively maintaining polymorphic populations. The requirements for fluctuating selection to maintain genetic variation are rather restrictive (Haldane and Jayakar 1963; Hedrick et al. 1976; Hedrick 1986), but they become significantly less so with overlapping generations (Templeton and Levin 1979; Ellner and Hairston 1994; Ellner and Sasaki 1996). The long-lived seed bank of *L. parryae* provides such age structure.

For Pearblossom 1, where we have the longest record of plant performance, there is clear evidence that selection is both strong and temporally variable. Recent theoretical analysis of these data (Turelli et al. 2001) suggests that fluctuating selection, coupled with a long-lived seed bank, can maintain the polymorphism. Specifically, both morphs are maintained if the genotype whose arithmetic and geometric mean relative fitness are both less than one also has a relative fitness greater than one in years with its highest seed production. This is precisely what we observed at Pearblossom 1, where white-flowered plants have both the lowest arithmetic and geometric mean relative fitness, but are maintained in the population because of their fitness advantage in the years of highest seed production (Turelli et al. 2001). Precipitation patterns during our 11 years of observation at Pearblossom 1 were comparable to those over the last half century, thus the fluctuating selection we documented at this site may be sufficient to maintain the polymorphism. Nevertheless, we cannot rule out the possibility that other mechanisms, such as heterozygote superiority due to pleiotropy (Rausher and Fry 1993) or frequency-dependent selection mediated by pollinator behavior

(Rausher et al. 1993), might also contribute to the maintenance of the polymorphism. A complete discussion of the mechanisms contributing to the maintenance of the flower-color polymorphism at Pearblossom 1 is given in Turelli et al. (2001).

The morph frequency in the area surrounding Pearblossom 1 is very similar to that observed in our transects at this site, so it is unlikely that the flower color polymorphism there is due to gene flow from adjacent populations. This is not the case for Pearblossom 2, which has neighboring populations that are primarily blue- and primarily white-flowered. Here, with only three years of data on seed production, we have not observed fluctuating selection. Instead, we observed strong selection against white-flowered plants in one year and a marginally significant advantage for this morph in another year. If selection over longer time periods typically favors white-flowered plants, the flower color polymorphism at Pearblossom 2 could be maintained by gene flow from neighboring blue-flowered populations.

At Antelope Valley, in four years of observation there was never a significant advantage for one flower color morph over the other. However, flower color at this site was more variable, suggesting a possible role for modifier genes. At this site we observed flowering in years where there were no plants at the Pearblossom sites, and a range of seed production that was lower than that at Pearblossom. Although we cannot reject the hypothesis that flower color at Antelope Valley is a neutral character, further study is clearly warranted.

The Historical Debate

Epling and Dobzhansky (1942) first concluded that the spatial pattern of flower color in *L. parryae* was consistent with Wright's prediction that random genetic drift could cause the genetic differentiation of populations. Wright (1943) followed their paper with a detailed reanalysis of their data, concluding that *L. parryae* was an example of isolation by distance, the initial step in the shifting balance process. However, there was an early suggestion that this conclusion might be incorrect. William Hovanitz, a graduate student at Cal Tech who had worked with Dobzhansky on *Drosophila*, wrote to Wright in 1942: "it seems to me very difficult to imagine a gene existing in wild populations which has absolutely no different physiologic effect on individuals carrying it as compared with a type standardized as 'wildtype.'" Hovanitz went on to describe what he knew of *L. parryae* populations in the field and concluded: "there is fairly good reason to suppose, therefore, that the blue flowered plants are being selected for in these areas" (Provine 1986, p. 375).

After nearly 20 years of additional field studies, this was precisely the conclusion reached by Epling and his colleagues (Epling et al. 1960). They suggested that the earlier conclusions reached by Epling and Dobzhansky (1942) and Wright (1943) were wrong, and proposed the alternative explanation that the flower color polymorphism in *L. parryae* was the product of natural selection (Epling et al. 1960). As described by Provine (1986), Wright was dismayed by these new findings, but declined Epling's invitation to contribute a technical appendix to his paper. Wright reanalyzed the entire dataset

and published the results many years later, staunchly maintaining his view that natural selection on flower color was of only minor significance in *L. parryae* (Wright 1978).

One possible explanation for Wright's persistence is that Epling et al. (1960) failed to conduct detailed, direct observations, and instead focused more on the possible causes of selection than on selection itself. For example, they investigated whether there were differences in soils between all-blue and all-white areas and found none. Likewise, they compared the composition of the plant community in all-blue and all-white areas and again found no differences. Their conclusion that flower color in *L. parryae* is subject to selection was based on experimental evidence that this species has a large, long-lived seed bank and on long-term census data showing that morph frequencies were temporally stable. Even in their own minds, this was not strong evidence for selection: "the frequencies of blue and white flowered plants are in the long run the product of selection operating at an intensity we have been unable to measure" (Epling et al. 1960, p. 254).

Wright had another reason for doubting the conclusions of Epling et al. (1960). *Linanthus parryae* was Wright's "first and best example of isolation by distance" (Provine 1986, p. 485) and "an example of the whole shifting balance theory of evolution in nature in action" (Provine 1986, p. 378). With so much at stake, Wright was apparently unable to consider objectively the conclusions of Epling et al. (1960) that flower color in *L. parryae* was not the result of isolation by distance. It is hard to fathom how Wright could give so little merit to Epling's findings. Wright clearly valued the census data collected by Epling and his colleagues, as it formed the basis for all his own analyses, yet he apparently saw no inconsistency in ignoring virtually all of their other data.

Despite considerable controversy, the shifting balance theory remains one of the most celebrated of Wright's contributions to evolutionary biology (Provine 1986; Crow 1991; Wade and Goodnight 1991; Coyne et al. 1997). His detailed analysis of the flower color polymorphism in *L. parryae* was pivotal to the development of this theory (Provine 1986). We provide strong evidence that Wright greatly underestimated the role of natural selection in polymorphic populations of *L. parryae*. This conclusion is supported by additional studies showing that a sharp, local discontinuity in flower color is associated with a number of environmental differences and by transplant studies that demonstrate selection against the rare color morph (D. W. Schemske and P. Bierzychudek, unpubl. ms.). Our results for *L. parryae* are therefore consistent with Mayr's suggestion that "Selective neutrality can be excluded almost automatically wherever polymorphism or character clines are found in natural populations" (Mayr 1965, p. 207).

ACKNOWLEDGMENTS

Over the 11 years of this project, we received assistance from many individuals and institutions, only some of which we are able to mention here. H. Lewis showed us the location of the original study sites; C. Taylor allowed us to see C. Epling's maps and field notes; S. Curtis and E. Soulanielle provided expert field and laboratory assistance for several key years of the study; N. Mantua provided information on

El Niño; Pomona College students E. Chen, A. Scherer Desmarais, H. Foshee, B. Godsey, J. Astle Happy, S. Im, H. Merchant, K. Merriam, and Y. Smith together counted hundreds of thousands of *Linanthus* seeds; dozens more Pomona College students provided field and/or laboratory assistance; D. Boose, P. Pack, and B. Best helped with many aspects of the fieldwork; R. Levin provided us with a base camp in southern California; and The Pines provided an oasis of sanity in the biological and cultural desert of Palmdale. We thank J. Coyne, H. Lewis, M. Turelli, and R. Lande for useful discussion throughout the course of our studies; P. Miksová and N. R. Temkin for statistical advice; and two anonymous reviewers for helpful comments on the manuscript.

Financial support was provided by National Science Foundation BSR-8918246 to the University of Washington (DWS) and National Science Foundation RUI BSR-8919674 to Pomona College (PB). Other support was provided by Pomona College Schenck Funds and a Pomona College Seaver Grant to PB. DWS and PB thank their ancestors for the stress-tolerance traits that allowed them to tolerate the desert's prickliness as well as each other's.

LITERATURE CITED

- Barton, N. H. 1992. On the spread of new combinations in the third phase of Wright's shifting balance. *Evolution* 46:551–557.
- Brown, B. A., and M. T. Clegg. 1984. The influence of flower color polymorphisms on genetic transmission in a natural population of the common morning glory, *Ipomoea purpurea*. *Evolution* 38: 796–803.
- Burdon, J. J., D. R. Marshall, and A. H. D. Brown. 1983. Demographic and genetic changes in populations of *Echium plantagineum*. *J. Ecol.* 71:667–679.
- Coyne, J. A., N. H. Barton, and M. Turelli. 1997. Perspective: a critique of Sewall Wright's shifting balance theory of evolution. *Evolution* 51:643–671.
- Crawford, T. J. 1984. What is a population? Pp. 135–173 in B. Shorrocks, ed. *Evolutionary ecology*. Blackwell Scientific, Oxford, U.K.
- Crow, J. 1991. Was Wright right? *Science* 253:973.
- Crow, J., W. R. Engels, and C. Denniston. 1990. Phase three of Wright's shifting balance theory. *Evolution* 44:233–247.
- Ehleringer, J. R. 1988. Carbon isotope ratios and physiological processes in aridland plants. Pp. 41–54 in: P. W. Rundel, J. R. Ehleringer, and K. A. Nagy, eds. *Stable isotopes in ecological research*. Springer, New York.
- Ehleringer, J. R., J. P. Comstock, and T. A. Cooper. 1987. Leaf-twig carbon isotope ratio differences in photosynthetic-twigs desert shrubs. *Oecologia* 71:318–320.
- Ellner, S., and N. G. Hairston Jr. 1994. Role of overlapping generations in maintaining variation in a fluctuating environment. *Am. Nat.* 143:403–441.
- Ellner, S., and A. Sasaki. 1996. Patterns of genetic polymorphism maintained by fluctuating selection with overlapping generations. *Theor. Popul. Biol.* 50:31–36.
- Endler, J. A. 1986. *Natural selection in the wild*. Princeton Univ. Press, Princeton, NJ.
- Epling, C., and T. Dobzhansky. 1942. Genetics of natural populations. IV. Microgeographic races in *Linanthus parryae*. *Genetics* 27:317–332.
- Epling, C., H. Lewis, and F. M. Ball. 1960. The breeding group and seed storage: a study in population dynamics. *Evolution* 14: 238–255.
- Ernst, W. H. O. 1987. Scarcity of flower colour polymorphism in field populations of *Digitalis purpurea* L. *Flora* 179:231–239.
- Farquhar, G. D., K. T. Hubick, A. G. Condon, and R. A. Richards. 1988. Carbon isotope fractionation and plant water-use efficiency. Pp. 21–40 in: P. W. Rundel, J. R. Ehleringer, and K. A.

- Nagy, eds. Stable isotopes in ecological research. Springer, New York.
- Fisher, R. A. 1922. On the dominance ratio. Proc. R. Soc. Edinburgh. 42:321–341.
- . 1930. The genetical theory of natural selection. 2d ed. Oxford Univ. Press, Oxford, U.K.
- . 1954. Statistical methods for research workers. 12th ed. Oliver and Boyd, Edinburgh.
- Fisher, R. A., and E. B. Ford. 1947. The spread of a gene in natural conditions in a colony of the moth *Panaxia dominula*. Heredity 1:143–174.
- Goodnight, C. J. 1995. Epistasis and the increase in additive genetic variance: implications for phase 1 of Wright's shifting-balance process. Evolution 49:502–511.
- Grant, V. 1981. Plant speciation. 2d ed. Columbia Univ. Press, New York.
- Haldane, J. B. S. 1932. The causes of evolution. Longman, London.
- Haldane, J. B. S., and S. D. Jayakar. 1963. Polymorphism due to selection of varying direction. J. Genetics 58:237–242.
- Hedrick, P. W. 1986. Genetic polymorphism in heterogeneous environments: a decade later. Annu. Rev. Ecol. Syst. 7:535–566.
- Hedrick, P. W., M. E. Ginevan, and E. P. Ewing. 1976. Genetic polymorphism in heterogeneous environments. Annu. Rev. Ecol. Syst. 7:1–32.
- Kay, Q. O. N. 1978. The role of preferential and assortative pollination in the maintenance of flower colour polymorphisms. Pp. 175–190 in A. J. Richards, ed. Pollination and evolution. Royal Botanic Gardens, Sydney.
- Kondrashov, A. S. 1992. The third phase of Wright's shifting balance: a simple analysis of the extreme case. Evolution 46: 1972–1975.
- Levin, D. A. 1988. Local differentiation and the breeding structure of plant populations. Pp. 305–329 in L. D. Gottlieb and S. K. Jain, eds. Plant evolutionary biology. Chapman and Hall, London.
- Levin, D. A., and E. T. Brack. 1995. Natural selection against white petals in *Phlox*. Evolution 49:1017–1022.
- Martin, B., and Y. R. Thorstenson. 1988. Stable carbon isotope composition ($\delta^{13}\text{C}$), water use efficiency, and biomass productivity of *Lycopersicon esculentum*, *Lycopersicon pennellii*, and the F₁ hybrid. Plant Physiol. 88:213–217.
- Mayr, E. 1942. Systematics and the origin of species. Columbia Univ. Press, New York.
- . 1965. Animal species and evolution. Belknap Press of Harvard Univ. Press, Cambridge, MA.
- Patterson, R. 1993. *Linanthus*. Pp. 840–844 in J. C. Hickman, ed. The Jepson manual; higher plants of California. Univ. of California Press, Berkeley.
- Provine, W. B. 1986. Sewall Wright and evolutionary biology. Univ. of Chicago Press, Chicago.
- Quinn, W. H. 1987. El Niño. Pp. 411–414 in J. E. Oliver and R. W. Fairbridge, eds. The encyclopedia of climatology. Van Nostrand Reinhold, New York.
- Rausher, M. D., and J. D. Fry. 1993. Effects of a locus affecting floral pigmentation in *Ipomopsis purpurea* on female fitness components. Genetics 134:1237–1247.
- Rausher, M. D., D. Augustine, and A. VanderKool. 1993. Absence of pollen discounting in genotypes of *Ipomoea purpurea* exhibiting increased selfing. Evolution 47:1688–1695.
- Rice, W. R. 1989. Analyzing tables of statistical tests. Evolution 43:223–225.
- Schoen, D. J., D. E. Giannasi, R. A. Ennos, and M. T. Clegg. 1984. Stem color and pleiotropy of genes determining flower color in the common morning glory. J. Hered. 75:113–116.
- Snedecor, G. W., and W. G. Cochran. 1967. Statistical methods. 6th ed. Iowa State Univ. Press, Ames.
- Sokal, R. R., and F. J. Rohlf. 1994. Biometry. 3rd ed. W. H. Freeman and Co., San Francisco, CA.
- Stanton, M. L. 1987. Reproductive biology of petal color variants in wild populations of *Raphanus sativus*. I. Pollinator response to color morphs. Am. J. Bot. 74:178–187.
- Stanton, M. L., A. A. Snow, S. N. Handel, and J. Bereczky. 1989. The impact of a flower-color polymorphism on mating patterns in experimental populations of wild radish (*Raphanus raphanistrum* L.). Evolution 43:335–346.
- Stebbins, G. L., Jr. 1950. Variation and evolution in plants. Columbia Univ. Press, New York.
- Taylor, L. P., and R. Jorgensen. 1992. Conditional male fertility in chalcone synthase-deficient *Petunia*. J. Hered. 83:11–17.
- Templeton, A. R., and D. A. Levin. 1979. Evolutionary consequences of seed pools. Am. Nat. 114:232–249.
- Trenberth, K. E. 1997. The definition of El Niño. Bull. Am. Meteor. Soc. 78:2771–2777.
- Turelli, M., D. W. Schemske, and P. Bierzychudek. 2001. Stable two-allele polymorphisms maintained by fluctuating fitnesses and seed banks: protecting the blues in *Linanthus parryae*. Evolution 55:1283–1298.
- Wade, M. J., and C. J. Goodnight. 1991. Wright's shifting balance theory: an experimental study. Science 253:1015–1018.
- Waser, N. M., and M. V. Price. 1985. Reciprocal transplant experiments with *Delphinium nelsonii* (Ranunculaceae): evidence for local adaptation. Am. J. Bot. 72:1726–1732.
- Wright, S. 1931. Evolution in Mendelian populations. Genetics 16: 97–159.
- . 1943a. Isolation by distance. Genetics 28:114–138.
- . 1943b. An analysis of local variability of flower color in *Linanthus parryae*. Genetics 28:139–156.
- . 1978. Evolution and the genetics of populations. VI. Variability within and among natural populations. Univ. of Chicago Press, Chicago.
- . 1982. The shifting balance theory and macroevolution. Annu. Rev. Genet. 16:1–19.

Corresponding Editor: C. Eckert