Multilocus genetic structure of ancestral Spanish and colonial Californian populations of *Avena barbata*

(allozyme variants/discrete log-linear multivariate analysis/multilocus epistatic interactions)

M. PÉREZ DE LA VEGA*, P. GARCÍA*, AND R. W. ALLARD†‡

*Departamento de Genética, Facultad de Biología, Universidad de León, 24071 León, Spain; and †Department of Genetics, University of California, Davis, CA 95616

Contributed by R. W. Allard, November 1, 1990

ABSTRACT We have applied a multivariate log-linear technique to the analysis of interlocus allelic associations among 14 allozyme loci in a sample of 4011 plants from 42 Spanish populations of *Avena barbata*. The loci fell into three natural groups of five, five, and four loci. The five loci of the first group are invariant, or nearly so, throughout the range of the species. The genetic organization of the loci of this set is defined by a single five-locus genotype; each allele of this predominant genotype is a “wild-type” allele that contributes favorably to adaptedness in all single-locus and multilocus configurations regardless of environment. Although allelic diversity is high in Spain for the nine loci of the second and third sets, log-linear analyses showed that these loci are tied together in Spanish populations through complex networks of overlapping lower-order interlocus interactions. The ancestral Spanish and colonial Californian gene pools are closely similar in allelic composition on a locus-by-locus basis; however, Spanish allelic configurations at two-locus and higher-order levels are usually different from and much less tightly organized than in Californian populations. We conclude that the major force involved in the evolution of the colonial populations was selection that led to reorganization, at the interlocus level, of the ancestral Spanish allelic ingredients into different multilocus genotypes adapted to Californian habitats.

The Slender Wild Oat, *Avena barbata* Pott ex Link., is a diploidized tetraploid grass that is distributed over a vast area from the Western Mediterranean Basin across the Middle East to Nepal (1). Historical records indicate that *A. barbata* was introduced to California from Spain in the period of exploration and colonization and that the species soon became a major component of a number of distinct habitats in California (2). Detailed studies of the ecogenetics of *A. barbata* in California (3) have established that the species is now differentiated into several ecotypes marked by specific multilocus combinations of alleles of discretely identifiable morphological, allozyme, and restriction-fragment variants and that these ecotypes are distributed in patterns that precisely overlay environmental heterogeneity, especially heterogeneity for available moisture and for temperature. *A. barbata* occurs in two major climatic zones in California (4), the Mediterranean hot summer zone, which includes the extensive semiarid grasslands and oak savannah of the foothills bordering the Central Valley, and the Mediterranean cool summer zone, which includes the intermontane regions of the Pacific coastal strip and the higher foothills of the Sierra Nevada mountains; the Mediterranean cool summer zone is characterized by mild wet winters and a diversity of habitats varying from mesic to semiarid. The great majority of populations in the hot summer zone are nearly monomorphic for a specific combination of alleles designated xeric,

whereas most populations in the cool summer zone are nearly monomorphic for a constellation of alleles designated mesic. In areas of interface between semiarid and mesic habitats, zones of polymorphism for xeric and mesic gametic types occur on a microgeographical scale, in which frequencies of the two gametic types are correlated with available moisture and temperature. Gametic types featuring mixtures of xeric and mesic alleles are rare (3). Also the size of xeric and mesic zones expands and contracts, and the frequencies of xeric and mesic gametic types shift rapidly in polymorphic zones in dry and wet years. A dozen or so additional ecotypes, each marked by a specific set of multilocus allelic complexes, have been identified in California; these ecotypes generally occupy small patches featuring unusual edaphic conditions.

It is apparent from a recent detailed study (5) of 42 Spanish populations of *A. barbata* that present-day Spanish and Californian populations are closely similar in allelic composition and allelic frequencies but that the arrays of multilocus genotypes found in the colonial gene pool are different from those of the ancestral Spanish gene pool. Main objectives of the present study were to determine whether ecotypes parallel to those marked by multilocus complexes of alleles in California also occur in Spain and whether rainfall and temperature are major environmental factors in Spain as in California. Genetic structure among loci has conventionally been assessed in terms of pairs of loci; however, two-locus measures are inadequate for analysis of multilocus structure, and we have consequently used a discrete multivariate log-linear technique (6), as adapted for study of associations among several loci with each other and with environmental factors (7), in our analyses of the Spanish multilocus *A. barbata* data. We conclude from our analyses that the present-day Spanish and Californian gene pools are differently structured on a multilocus basis and that the main force responsible for evolutionary change in the colonial populations was selection for particular combinations of alleles at many different loci that allowed the colonial populations to meet the different environmental challenges encountered in California.

MATERIALS AND METHODS

The materials of this study were the same 4011 plants collected from 42 Spanish populations located at sites illustrated in figure 1 of ref. 5. Each site was sampled by collecting a single panicle from ≈ 100 randomly chosen individuals. A single seed was sown from each panicle, and crude extracts from seedlings were assayed electrophoretically and scored for 14 allozyme loci; the 51 allozyme alleles and 80 homozygous single-locus allozyme genotypes observed in the 4011 individuals scored are described in ref. 5. The number of possible 14-locus genotypes is several orders of magnitude

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked “advertisement” in accordance with 18 U.S.C. §1734 solely to indicate this fact.

‡Present address: Department of Agronomy and Range Science, University of California, Davis, CA 95616.

larger than the number of plants (4011) classified, so that there was no possibility of analyzing all 14 loci simultaneously. Consequently, to reduce the number of classes to manageable proportions, we first divided the loci into three natural groups of five, five, and four loci and then reduced each locus to diallelic state by grouping the genotypes into two classes within each group. All five loci of the first group (*Got1*, *Got2*, *Pgml*, *Mdh1*, and *Mdh2*) are fixed or nearly fixed for the same alleles in Californian and in all Spanish and southwestern Asian populations. The genetic structure of these five loci is therefore defined by the single five-locus genotype that is predominant over the range of the species. Each of the alleles of this genotype is a "wild-type" allele that contributes favorably to superior adaptedness in all single-locus and multilocus configurations in all environments in which the species occurs.

The second group of loci included the five loci (*Pgd1*, *Lap1*, *Prx1*, *Acp1*, and *Est1*) that most clearly differentiate the xeric from the mesic ecotypes in California. These five loci are fixed or very nearly fixed in xeric habitats in California for a set of alleles designated allele "1" at each locus. The predominant five-locus genotype, designated 11111, precisely defines the genetic structure of these five loci in Californian xeric habitats; all of the alleles of this predominant genotype are wild-type alleles that interact favorably in all two- and higher-order allelic configurations in that habitat. Spanish materials in this group were classified into two classes respecting each locus: loci with the Californian xeric allele were designated 1 and loci with any other allele were designated 2; in total, 21 nonxeric 2 alleles, most of which were infrequent, were observed in Spain. In this notation 11222 indicates a plant with the Californian xeric allele at loci *Pgd1* and *Lap1* and the "synthetic" nonxeric allele at loci *Prx1*, *Acp1*, and *Est1*.

Loci of a third group (*Pgd2*, *Pgil*, *Mdh3*, and *Acp2*) are virtually invariant in California: the xeric and the mesic ecotypes, as well as most of the minor ecotypes, carry the same allele at each locus. Again the Spanish materials were classified into two groups, 1 vs. 2; thus 1111, 1122, and 2222 indicate, respectively, individuals with the predominant Californian allele at all four loci, at two loci (*Pgd2* and *Pgil*), and at none of the four loci. A total of 13 2 alleles, most of which were infrequent, was observed in Spain. After analysis of the above loci, a fourth set of five loci (*Pgd1*, *Pgd2*, *Prx1*, *Acp1*, and *Mdh3*), made up of a mixture of group 2 and group 3 loci, was formed and analyzed.

The specific multivariate analysis followed is a two-step process: (i) likelihood ratio tests were employed in a series to identify and eliminate terms that have statistically insignificant effects; (ii) log-linear models were constructed to fit the data to the remaining terms in the system. Models were fit in an hierarchical manner such that a higher-order term was included only when lower-order terms failed to fit the data; also, when a higher-order term was included, all of its lower-order relatives were also included. The best-fitting model was then chosen on the basis that it was maximally parsimonious (i.e., included a minimal number of terms) and that it also provided a statistically acceptable fit to the data. Two screening methods were used in model selection, stepwise selection (8) and marginal tests (9). In associating rainfall and temperature, which have important effects in determining multilocus associations in California (also in Israel, ref. 10), Spanish sites with < or > 500 mm of rainfall and with mean annual temperatures > or < 13°C were designated classes 1 and 2, respectively.

RESULTS

Loci of Group 2 (Californian Xeric Alleles). The five loci of this group—*Pgd1*, *Lap1*, *Prx1*, *Acp1*, and *Est1*—were coded

A, C, D, E, and F, respectively. The contingency table (not given) for these five loci has $2^5 = 32$ cells. Application of stepwise selection (8) and marginal tests (9) as screening procedures led to selection of the model [ACF] [ADE] [CDF] [CEF] as best fitting the data. The resulting likelihood-ratio statistic, G^2 (which is distributed approximately as χ^2) and χ^2 statistics, each with 12 df, took values of 14.9 and 15.7, $P = 0.245$ and 0.207, respectively, indicating a good fit to the model. Goodness-of-fit to the model was also tested by examining standardized residuals (*SR*) for each cell [$SR = (\text{observed} - \text{expected}) / \sqrt{\text{expected}}$] in the contingency table. *SR* values are distributed asymptotically following a normal distribution with mean zero and unit variance; thus, absolute values of 1.96, 2.58, and 3.29 indicate statistically significant departures from the model at probability (*P*) levels 0.05, 0.01, and 0.001, respectively. The largest $|SR|$ value was 1.6, also indicating that the above model fits the data well. The log-linear form of this model is as follows:

$$\begin{aligned} \ln m_{hijkl} = & u + u_{A(h)} + u_{C(i)} + u_{D(j)} + u_{E(k)} + u_{F(l)} + u_{AC(hi)} \\ & + u_{AD(hj)} + u_{AE(hk)} + u_{AF(hl)} + u_{CD(ij)} + u_{CE(ik)} \\ & + u_{CF(il)} + u_{DE(jk)} + u_{DF(jl)} + u_{EF(kl)} + u_{ACF(hil)} \\ & + u_{ADE(hjk)} + u_{CDF(ijl)} + u_{CEF(ikl)}, \end{aligned} \quad [1]$$

for all $h, i, j, k, l = 1, 2$, subject to the constraint that all of the marginal totals equal zero (6); m_{hijkl} is the expected number of individuals carrying the h th allele of locus A, the i th allele of locus C, and so on. On acceptance of the model the u terms were estimated. Terms such as $u_{A(h)}$ measure the main effect of each locus—i.e., provide comparisons of the allelic frequencies (1 vs. 2) at each locus, whereas terms such as $u_{AC(hi)}$ provide estimates of associations between alleles at pairs of loci. Additional information about associations can be obtained by estimation of standardized u values. There are two estimated u values for each single-locus effect—e.g., alleles 1 and 2 of locus A are related to each other by $u_{A(1)} = -u_{A(2)}$. There are four estimated u values for each pairwise association (e.g., locus A with locus C) related by $u_{AC(11)} = -u_{AC(12)} = -u_{AC(21)} = u_{AC(22)}$, because of the constraint that all marginal totals equal zero. The absolute values of the u values indicate the magnitude, and their algebraic signs show whether the association is positive or negative between alleles 1 and 2 in each term.

Standardized $|u|$ values for the terms in the model for locus set ACDEF are given in Table 1. The u values measuring the main effects of loci A, C, D, E, and F are all significant, indicating that the frequency of alleles 1 and 2 differed significantly in the Spanish materials for each locus. Numbers of individuals homozygous for 1 vs. 2 alleles for loci A, C, D, E, and F were 157:3854, 1771:2240, 648:3363, 3323:688, and 1105:2906, respectively. Thus, in our sample, Californian xeric alleles were more frequent only for locus E, and nonxeric alleles were more frequent for the other loci. Summed over all five loci, Californian xeric alleles were outnumbered by nonxeric alleles in our Spanish sample by 7004:13,051 (0.349:0.651), which suggests, in accord with rainfall data, that the sites we collected in Spain are more mesic, on the average, than the Californian xeric habitat.

According to the model [ACF] [ADE] [CDF] [CEF], loci A, C, D, E, and F are associated in Spain through a series of interlocking two- and three-locus associations. The pairwise terms are nested within the three-locus associations in this model, and they must consequently be interpreted with caution; however, inferences concerning the intensity and nature of the pairwise interactions can be made from the values (Table 1) and algebraic signs (not given) of the u

Table 1. Standardized absolute u values for the statistically significant terms included in the models for locus-set ACDEF and for locus-set ACDEF including environmental factors

ACDEF		ACDEF + variable		
Term	u	Term	T, u	R, u
A	15.2***	AV	4.4***	2.1*
C	2.0*	CV		2.3**
D	7.5***	EV	2.2*	
E	9.4***	ADV	2.7**	
F	7.8***	AEV		3.5***
AC	3.5***	CEV		3.0**
AF	3.6***	CFV		6.2***
CD	11.2***	DEV	2.1*	
CE	7.5***	DFV		4.3***
EF	12.1***	EFV	2.0*	
ACF	2.0*	ACFV	2.3*	
ADE	2.9**	ADEV		2.0*
CDF	11.9***	AEFV		3.2**
CEF	2.7**	CDFV	2.4*	
		DEFV	3.3***	
		ACDFV	3.0**	

Best-fitting models are ACDEF, [ACF] [ADE] [CDF] [CEF]; ACDEF, [ACE] [ADE] [DEFT] [ACDFT]; and ACDEF, [ACD] [ACE] [CDF] [CER] [CFR] [DFR] [ADER] [AEFR]. Variables were temperature (T) or rainfall (R) (see text). *, **, ***, P of 0.05, 0.01, and 0.001, respectively.

values. Taking the AC pair as an example, the observed numbers of gametic types 11, 12, 21, and 22 were 120:37:1651:2203, whereas, under the hypothesis of random association, the expected numbers are 69:88:1702:2152, respectively. A test of goodness-of-fit gives $\chi^2_{(1)} = 14.2$, $P < 0.001$, which suggests significant excesses and deficiencies in some gametic classes. Standardized residuals for gametic classes 11 and 12 are 6.1 and -5.4, which suggests that 11 is in significant excess and 12 is in significant deficiency ($P < 0.001$); standardized residuals for classes 21 and 22 are -1.2 and 1.1, respectively, suggesting that these two classes do not differ from expectations assuming randomness. SR values also suggest that significant associations occur between locus pairs AC, AF, CD, CE, and EF. Algebraic signs (not given) suggest that the following two-locus allelic configurations are in excess: AC(11, 22), AF(12, 21), CD(12, 21), CE(12, 21), and EF(12, 21). Thus only the AC association is in the same direction in Spain as in California.

The most intense three-locus association is CDF. The algebraic signs of the u values show that, even after allowing for all lower-order associations among loci C, D, and F, there are excesses in the triplets 121, 211, 222, and 112 and deficiencies in the corresponding triplets 212, 122, 111, and 221. Other three-locus terms showing significant excesses were ACF(211, 121, 112, 222), ADE(112, 222, 121, 211), and CEF(121, 211, 222, 112); there were corresponding deficiencies in the complementary terms. The existence of three-locus interactions implies that two-locus interactions are nested within the higher-order interactions; therefore the model [ACF] [ADE] [CDF] [CEF] includes all 10 of the second-order interactions, although as pointed out above, only five of the second-order interactions are statistically significant. In Californian xeric habitats the two-locus, three-locus, four-locus, and five-locus configurations are nearly always 11, 111, 1111, and 11111.

Loci of Group III (Californian Xeric-Mesic Alleles). The contingency table for the four loci of this group (B, G, H, and I) has 16 cells. The screening procedure identified the saturated model [BGHI] as the only one that fits the data. Under this model the observed and expected numbers in each cell are equal so the G^2 , χ^2 , and all SR values are zero. This model thus indicates that allelic state at any locus depends on allelic

state at each of the other three loci, so that associations can be explained only if all combinations are known. This strong interdependence indicates that caution is called for in interpreting individual terms within the [BGHI] model. Nevertheless, an idea of the relative importance of the individual terms can be obtained by examining the u values (Table 2). The values measuring the main effects of loci BGHI are all significant. Numbers of 1 vs. 2 alleles were 3239:772, 3485:526, 3280:731, and 2652:1359. Thus alleles that are predominant in California are also heavily favored in Spain (overall 12,656:3388 or 0.789:0.211). There are significant associations among locus pairs BG (21 and 12 in excess) and BI (21 and 12 in excess) and among locus triplets BGH (111, 221, 212, 122 in excess) and GHI (211, 121, 112, 222 in excess). The saturated [BGHI] set implies significant associations among all four loci (standardized $u = 2.6$) and algebraic signs of the standardized u values show that combinations 1111, 2211, 2121, 1221, 2112, 1212, 1122, and 2222 are in excess. Thus, in Spain, most of the excesses are in non-Californian combinations at the two-, three-, and four-locus levels.

Combined Analysis of the Two Subsets. To keep the number of gametic combinations within manageable proportions, a five-locus data set (ABDEH) was selected to represent the nine loci of locus groups 2 and 3. The selection was result guided. In the ACDEF subset, strong pairwise associations among loci A, C, D, and E indicate that allelic state at any one locus can be predicted in large part from allelic state at any of the other loci; also locus E is strongly associated with loci A and C. Accordingly, loci A, D, and E were chosen to represent the ACDEF set. Locus B was chosen to represent the BGHI set because it is strongly associated with G and I. *Mdh3* (H), which was not strongly associated with any other locus, was chosen to represent itself. The contingency table for these five loci has 32 cells. The best fitting model was [BDEH] [ABD] [ABE] [ADH] [AEH] with $G^2 = 6.33$, $P = 0.502$ and $\chi^2 = 10.78$, $P = 0.149$ for 7 df. The largest SR was 1.8. Thus the fit to the model is good. The main thing to note from the standardized u values (Table 3) is that the ABD, ABE, AEH, DEH, and BDEH associations are significant and that the BDEH term included two loci (DE) of the ACDEF set and two loci (BH) of the BGHI set. Thus, taking two-locus associations into account, the BDEH term can be extended, at the least, to a seventh-order term, BCDEFGH; this suggests that much of the genome of *A. barbata* has congealed into a multilocus complex in Spain. Again most of the excesses in allelic configurations were in non-Californian combinations.

Table 2. Standardized absolute u values for the statistically significant terms included in the models for locus-set BGHI and locus-set BGHI including environmental factors

BGHI		BGHI + variable		
Term	u	Term	T, u	R, u
B	10.2***	BV	2.0*	2.1*
G	11.1***	GV	4.9***	2.8**
H	7.4***	HV	2.3*	
I	5.2***	IV	2.0*	2.2*
BG	4.3***	BGV		3.7***
BI	4.2***	BHV	11.4***	
BGH	2.1*	GHV		2.5*
GHI	3.0**	HIV		2.4*
BGHI	2.6**	BGIV	4.2***	3.4***
		BHIV	9.8***	
		GHIV		3.8***

Best-fitting models are BGHI, [BGHI]; BGHIT, [BGH] [GHT] [BGIT] [BHIT]; and BGHIR, [BGHI] [BGIR] [GHIR]. Variables were temperature (T) or rainfall (R) (see text). *, **, ***, P of 0.05, 0.01, and 0.001, respectively.

Table 3. Standardized absolute u values for the statistically significant terms included in the models for the combined set of loci ABDEH excluding and including environment factors

ABDEH		ABDEH + variable		
Term	u	Term	T, u	R, u
ABD	2.2*	ABV		2.6**
ABE	3.5***	AHV	2.8**	
AEH	2.8**	BDV		2.9**
DEH	3.4***	EHV	5.0***	
BDEH	2.2*	ADHV		2.8**
		BDHV	4.0***	
		BEHV		3.3**

Best-fitting models are ABDEH, [ABD] [ABE] [ADH] [AEH] [BDEH]; ABDEHT, [ABE] [ADE] [ADT] [AEH] [AHT] [DET] [EHT] [BDHT]; and ABDEHR, [ABE] [BDH] [ABR] [AER] [BDR] [ADEH] [ADHR] [BEHR]. Variables were temperature (T) or rainfall (R) (see text). *, **, ***, P of 0.05, 0.01, and 0.001, respectively.

Relationships Among Loci and Environmental Factors. Relationships of the ACDEF, BGHI, and ABDEH sets of loci with temperature and with rainfall were also analyzed. Temperature and rainfall at the 42 individual collection sites in Spain were each classified into two disjoint classes based on whether mean annual temperature was $<$ or $>$ 13°C and whether annual was $<$ or $>$ 500 mm. Associations among loci and temperature or rainfall were assessed following the same procedures used in assessing associations among loci. The models obtained when temperature or rainfall were included are given in Tables 1, 2, and 3. In each of the three sets of loci both temperature and rainfall were significantly associated with specific allelic combinations; thus, temperature and rainfall affect multilocus structure in Spanish *A. barbata*. However, the allelic configurations that are in excess or deficiency often differ in the models that include rainfall or temperature. Thus, although the Spanish populations are structured on a multilocus basis, it is apparent that the patterns of structuring are much more complicated in Spain than in California.

DISCUSSION

Comparisons based on identity measures (5) have shown that the present-day Spanish and Californian gene pools are closely similar in overall allelic composition and in allelic frequencies. In the present study we analyzed the multilocus genetic structure of the Spanish gene pool in detail and related multilocus structure to rainfall and temperature. Our analyses show that the Spanish gene pool is structured on a multilocus basis and that rainfall and temperature have statistically significant effects on multilocus structure. However, allelic configurations of the Spanish populations at the two-locus and at higher-order levels are mostly different from and less tightly structured than those of the Californian populations. These differences in genetic organization appear associated with ecological differences in the habitats occupied by the species in Spain and California. In California *A. barbata* occurs in two climatic zones, one a semiarid zone characterized in many years by growth-limiting rainfall (250–500 mm) and mild temperatures (minima rarely $<-4^{\circ}\text{C}$) during the growing season (November–May) and the other a mesic zone characterized by generally adequate rainfall (500–1250 mm) and even milder temperatures (minima rarely $<0^{\circ}\text{C}$) during the growing season (October–June). Climatic data show that the 42 sites collected in Spain include a more complex array of temperature–rainfall combinations than occur in California; Spanish genetic data also show that most Spanish sites have unique multilocus structures. The mating system of *A. barbata* is favorable for the development and maintenance of a patchwork of interlocus combinations of

alleles that provide for locally high adaptedness (11). The one percent of outcrossing that occurs leads to recombination that produces numerous different interlocus allelic configurations within local populations, whereas the 99% of selfing restricts recombination sufficiently to protect favorable interallelic associations from being broken up by segregation before they can be tested by selection (11, 12). The tetraploidy of *A. barbata* allows the development and fixation of heteroallelic intralocus associations of alleles (13). Thus the mating system of *A. barbata*, in combination with differing modes of selection associated with local environmental heterogeneity, appears to have led to the development of a pathwork pattern of differing multilocus genetic organizations in Spain.

On introduction of *A. barbata* to California from Spain, the three sets of loci analyzed in this study each responded to the habitats encountered in a different way. The migrants that arrived in California were presumably fixed for the predominant five-locus genotype of the first set of loci. This constellation of alleles confers high adaptedness throughout the old world and evidently also in California, as it is the only genotype yet revealed by extensive sampling in California. The situation was apparently similar for the third set of loci. The Californian alleles of this set are also heavily favored in Spain (overall 0.789:0.211), and certain Californian lower-order configurations, as well as the Californian *IIII* four-locus configuration, were in excess in some Spanish populations. It is not difficult to imagine that these allelic configurations were introduced in migrants and that they increased rapidly in California, whereas the non-Californian alleles and allelic combinations were soon eliminated by selection in the colonial environments.

The migrants to California also presumably carried a more or less random sample of the 26 Spanish alleles and their combinations of the second set of five loci (5). The Californian xeric alleles and combinations were apparently highly favored in the Californian xeric habitat, and it seems likely that after some time these alleles and allelic combinations replaced nonxeric alleles and combinations of alleles in xeric habitats. The largest frequency change, from the Spanish frequency of 0.04 to the Californian frequency of 1.00, was for allele *I* of locus A, and the smallest frequency change, from the Spanish frequency of 0.83 to 1.00 in California, was for allele *I* of locus E. An opposite set of changes in the frequency of Californian xeric alleles took place in Californian mesic habitats; in these habitats the frequency of xeric alleles decreased to near zero, whereas one or another of the Spanish nonxeric alleles increased to a frequency near 1.00 during the same period. Note that the above allelic frequency changes were in opposite directions and balancing; thus, averaged over the Californian habitats, the evolutionary processes that occurred in California left the Californian gene pool virtually the same as the Spanish gene pool respecting allelic composition and allelic frequencies (5).

The outcome of evolutionary change respecting the multilocus structure of the Californian gene pool was very different: the evolutionary processes that took place in California led to extensive restructuring such that the multilocus genetic structure of the present-day colonial gene pool differs strikingly from that of the Spanish gene pool. Our results show that Spanish rainfall–temperature combinations are much more complex than those of California and that allelic associations that predominate in California are present in only a few places in Spain. Some of these Californian lower-order allelic associations were probably brought to California in migrants from Spain and contributed to the development of the higher-order associations that now predominate in California. There is, however, evidence from experimental populations that selection in new environments often rapidly breaks down existing multilocus allelic associ-

ations and produces new associations among favored alleles (11, 12). Thus, it is also likely that the differing modes of selection encountered in California quickly broke down the predominant non-Californian allelic configurations that came to California in the migrants and reorganized the basic allelic ingredients from Spain in one way in Californian xeric habitats, in another way in Californian mesic habitats, and in still other ways in the minor Californian habitats. We therefore conclude that multilocus selection, operating in different ways in the various colonial habitats encountered by the migrants, was the main force involved in the evolution of *A. barbata* in California.

This research was supported in part by grants from the U.S.-Spain Joint Committee for Scientific and Technological Cooperation (CCB8504-010), the Spanish Comision Interministerial de Ciencia y Tecnologia (PB85-0153), the U.S. National Science Foundation (BSR 83110869), the U.S. Public Health Service (National Institutes of Health Grant GM-32429), and the University of California (McDonald Endowment Fund).

1. Rajhathy, T. & Thomas, H. (1974) *Cytogenetics of Oats (Avena*

- L.*), Misc. Publ. Genet. Soc. Canada No. 2 (Genet. Soc. Canada, Ottawa).
2. Robbins, W. W. (1940) *Alien Plants Growing Without Cultivation in California*, Bull. 637 (Calif. Univ. Agric. Exp. Station, Berkeley, CA).
3. Allard, R. W., Babbel, G. R., Clegg, M. T. & Kahler, A. L. (1972) *Proc. Natl. Acad. Sci. USA* **69**, 3043-3048.
4. Durrenberger, R. W. (1960) *Patterns on the Land* (Roberts, Northridge, CA).
5. García, P., Vences, F. J., Pérez de la Vega, M. & Allard, R. W. (1989) *Genetics* **122**, 687-694.
6. Fienberg, S. E. (1980) *The Analysis of Cross-Classified Categorical Data* (MIT Press, Cambridge, MA).
7. Zhang, Q., Saghai Maroof, M. A. & Allard, R. W. (1990) *Theor. Appl. Genet.* **80**, 121-128.
8. Goodman, L. A. (1971) *Technometrics* **13**, 33-61.
9. Brown, M. B. (1976) *Appl. Stat.* **25**, 37-46.
10. Kahler, A. L., Allard, R. W., Krzakowa, M., Wehrhahn, C. F. & Nevo, E. (1980) *Theor. Appl. Genet.* **56**, 31-47.
11. Allard, R. W. (1975) *Genetics* **79**, 115-126.
12. Allard, R. W. (1988) *J. Hered.* **79**, 225-238.
13. García, P., Morris, M. I., Saénz-de-Miera, L. E., Allard, R. W., Pérez de la Vega, M. & Ladizinsky, G. (1991) *Proc. Natl. Acad. Sci. USA*, **88**, 1207-1211.