Allelic and Genotypic Composition of Ancestral Spanish and Colonial Californian Gene Pools of *Avena barbata*: Evolutionary Implications

P. Garcia,*[,1] F. J. Vences,* M. Pérez de la Vega* and R. W. Allard[†]

**Departamento de Genética, Facultad de Biologia, Universidad de León, 24071 León, Spain, and [†]Department of Genetics, University of California, Davis, California 95616*

Manuscript received November 16, 1988
Accepted for publication March 31, 1989

ABSTRACT

Spanish explorers and colonists inadvertently started a massive experiment in evolutionary genetics when they accidentally introduced *Avena barbata* to California from Spain during the seventeenth and eighteenth centuries. Assays of the Spanish and Californian gene pools of this species for 15 loci show that the present day Spanish gene pool, particularly that of Southwestern Spain, is identical or virtually identical to that of California for five loci and closely similar for nine loci. Despite their similar allelic and single-locus genotypic compositions, the present-day Spanish and Californian gene pools are differently structured on a multilocus genetic basis. Evolutionary implications of these results are discussed.

THE Slender Wild Oat, *Avena barbata* Pott ex Link is an annual, predominantly self-fertilizing, diploidized auto-allopolyploid ($2n = 4x = 48$ chromosomes) grass which is endemic to Southwestern Asia and the Mediterranean basin (MARSHALL and ALLARD 1970a; BAUM 1977; HUTCHINSON *et al.* 1983b). This species, which was unintentionally introduced to California from Spain during the period of exploration and colonization (ROBBINS 1940), has become a major component of grassland and grass-oak savannah habitats in California where it occurs in populations of millions of individuals. Detailed studies of the ecogenetics of *A. barbata* in California (MARSHALL and ALLARD 1969, 1970a,b; CLEGG and ALLARD 1972; HAMRICK and ALLARD 1972, 1975; ALLARD *et al.* 1972; MILLER 1977; ALLARD, MILLER and KAHLER 1978; HAKIM-ELAHI 1980; HUTCHINSON 1982; PINERO 1982; CLUSTER 1984) have established that: (1) the species is differentiated into a number of distinct ecotypes, each marked by specific combinations of alleles for morphological markers, allozyme loci, and quantitative characters and (2) that these ecotypes are distributed in patterns which are fine-scaled overlays of environmental heterogeneity, especially heterogeneity for available moisture. *A. barbata* occurs in two major climatic zones in California (DURRENBERGER 1960), the Mediterranean hot summer zone (mean temperature in the hottest month exceeds 22°) and the Mediterranean cool summer zone (mean temperature in the warmest month below 22°). The Mediterranean hot summer zone, which includes the extensive semiarid grasslands and oak savannah of the foothills bordering the Central Valley, is much more uniform environmentally than the Mediterranean cool summer zone; the latter zone, which includes the intermontane regions of the coastal strip and higher foothills of the Sierra Nevada mountains, features marked diversity in microhabitats varying from mesic to semiarid. The great majority of populations of the hot summer zone are monomorphic, or very nearly so, for a specific multilocus combination of alleles designated "xeric." In contrast most populations of the cool summer zone are monomorphic, or nearly so, for a complementary combination of alleles designated "mesic." Where semiarid and mesic habitats interface, patches of polymorphism for xeric and mesic genotypes occur on a microgeographical scale in which frequencies of the two genotypes are correlated with amounts of available moisture. However, about a dozen additional genotypes, each marked by its own multilocus set of alleles and occupying a habitat distinguishable from the standard mesic or xeric habitat, have been identified. Approximately 80% of the individual plants of *A. barbata* in California have the xeric multilocus set of alleles, approximately 10% of the mesic set, and the remaining 10% have one or another of the other dozen or so multilocus allelic complexes.

Studies of *A. barbata* from Israel (KAHLER *et al.* 1980) and Southwest Asia (R. W. ALLARD, unpublished data) have shown that allelic diversity is greater in *A. barbata* from these regions than in the California populations; however, as in the California populations, genetic variability in eastern Mediterranean and Southwest Asian populations is structured into multilocus complexes correlated with environment. Studies of small samples of *A. barbata* from the western Med-

[1] To whom reprint requests should be addressed.

iterranean basin (CLEGG and ALLARD 1972) indicate that the gene pool from that region is more like that of California than the gene pool of the eastern Mediterranean basin or Southwest Asia. This is consistent with the historical record which indicates that voyages of exploration and trade with California originated nearly exclusively in southwestern Spanish ports when *A. barbata* was unintentionally introduced to California in shipboard litter and as a contaminant in cereal crop seeds; hence it seems likely that *A. barbata* reached California primarily and perhaps exclusively from Spain, particularly southwestern Spain. However, very little is known about the Spanish gene pool; in the single report on Spanish *A. barbata* (JAIN and SINGH, 1979) only three enzyme systems were studied in 12 populations, sample sizes were very small and no geographical or ecological information was given. In this paper we show that the Spanish gene pool, particularly that of southwestern Spain, is very closely similar in allelic and genotypic content to that of California but that the Spanish and Californian populations differ in multilocus genotypic composition. The evolutionary implications of these findings are discussed in this paper and in a separate paper in which the differing multilocus structures of the ancestral Spanish and colonial Californian populations are analyzed in detail.

MATERIALS AND METHODS

The materials of this study were collected from 42 Spanish populations located at the sites shown in Figure 1. Each site was sampled by collecting a single panicle from approximately 100 randomly chosen plants. A single seed was sown from each panicle and crude extracts from the first and second leaves of 15-day-old seedlings were electrophoresed in 12% starch gels following procedures slightly modified from those of MARSHALL and ALLARD (1970b) and HUTCHINSON *et al.* (1983a). The enzyme systems assayed were acid phosphatase (ACPH, EC 3.1.3.2), esterase (EST, EC 3.1.1.2), glutamic oxaloacetic transaminase (GOT, EC 2.6.1.1), leucine aminopeptidase (LAP, EC 3.4.11.1), malic dehydrogenase (MDH, EC 1.1.1.37), peroxidase (PRX, EC 1.11.1.7), 6-phosphogluconate dehydrogenases (6-PGD EC 1.1.1.44), phosphoglucose isomerases (PGI, EC 5.3.1.9), and phosphoglucose mutases (PGM, EC 2.7.5.1). Staining procedures for ACPH, EST, LAP, PRX, and 6-PGD were those of HUTCHINSON *et al.* (1983a); procedures for PGI, GOT and MDH were similar to those of SHAW and PRASAD (1970). Seedlings of the California xeric and mesic genotypes were included in each gel as controls. Designations of loci and alleles follow those of HUTCHINSON *et al.* (1983a). The genetic control of the GOT, PGI and ACPH2 zones observed in the Spanish materials was worked out from segregation data from progenies of heterozygous plants found in the Spanish populations. Segregation patterns in families from heterozygous plants found in the Spanish populations showed that a duplicated single locus (*L, l*) governs the pubescent *vs.* glabrous leaf character in Spain, as in California.

Previous studies have shown that all populations of *A. barbata*, regardless of geographical location, are heavily self-

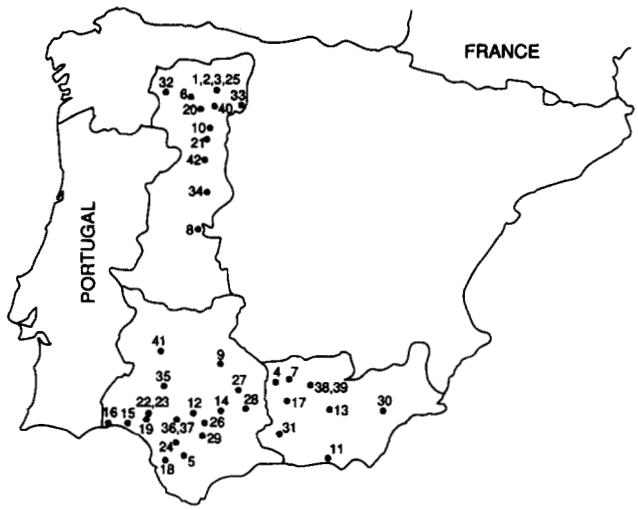

FIGURE 1.—Geographical locations of the 42 Spanish populations of *A. barbata*. 1, Agricolas 83; 2, Agricolas 84; 3, Agricolas 85; 4, Andújar; 5, Arcos de la Frontera; 6, Astorga; 7, Bailén; 8, Béjar; 9, Bélmez; 10, Benavente; 11, Berja; 12, Carmona; 13, C-323; 14, Ecija; 15, Huelva; 16, Isla Cristina; 17, Jaén; 18, Jerez de la Frontera; 19, KM 56; 20, La Bañeza; 21, La Encina; 22, La Palma del Condado 84; 23, La Palma del Condado 85; 24, Las Cabezas de San Juan; 25, León; 26, Marchena; 27, Medina-Azahara; 28, Montilla; 29, Morón de la Frontera; 30, N-342 A; 31, N-342B; 32, Ponferrada; 33, Sahagún; 34, Salamanca; 35, Santa Olalla; 36, Sevilla A; 37, Sevilla B; 38, Ubeda 84; 39, Ubeda 85; 40, Villamañán; 41, Zafra; 42, Zamora.

fertilized and that all individuals are homozygous for the great majority of loci. This was also found to be the case in the Spanish populations: mixed mating model estimates (SHAW, KAHLER and ALLARD 1981) of the amount of selfing, made in five of the populations, varied from 0.959 to 0.998 (mean = 0.985). The observed frequency of homozygotes in the 42 populations varied from 0.991 to 1.000 (mean = 0.998). When a plant is heterozygous at one locus it is also usually heterozygous at several other loci. Heterozygotes are thus either natural hybrids or descendants of recent natural hybrids between homozygotes which differed at several loci; because heterozygotes are infrequent in all populations, we have excluded them in our estimates of numbers of genotypes and genotypic frequencies. We have, however, included "fixed heterozygotes" in our estimates of numbers of genotypes and genotypic frequencies. When different alleles are present at duplicated loci in the two homologous genomes, both loci are nearly always homozygous and the two intragenomic homozygotes are "fixed heterozygotes" which breed true (ALLARD, MILLER and KAHLER 1978). As an example, the three-banded homozygote *96 96, 100 100* for the dimeric enzyme locus *6Pgd2* does not segregate on selfing; in contrast, however, the infrequent three-banded "true" single-heterozygotes *96 100, 100 100* and *100 100, 96 100*, and the "true" double heterozygote *96 100, 96 100* segregate in one-locus and two-locus patterns, respectively (HUTCHINSON *et al.* 1983b).

RESULTS

Allelic and genotypic content of the Spanish and Californian gene pools: Table 1 gives the alleles and single-locus homozygous genotypes, including fixed heterozygotes, which have been observed for the 15

TABLE 1

Alleles and single-locus homozygotes[a] observed in Spain and California[b]

Locus	Spain and California	Spain only	California only
Got1	*100 100*		
Got2	*100 100*	*100 106*[c]	
Pgm1	*100 100*		
Mdh1	*100 106*	*106 106*	
Mdh2	*100 100*	*100 106*[c]	
Mdh3	*100 107*	*95*[c] *103*[c], *95*[c] *108*[c], *100 100*, *107 107*	
Acp1	*100 101*	*98*[c] *98*[c], *98*[c] *100*, *98*[c] *101*, *100 100*, *101 101*	*96*[c] *100*
Acp2	*100 104*	*94*[c] *94*[c], *94*[c] *100*, *94*[c] *104*, *98*[c] *100*, *98*[c] *104*	
6Pgd1	*96 96*, *96 100*, *100 100*	*88*[c] *88*[c], *88*[c] *92*[c], *88*[c] *96*, *88*[c] *100*, *92*[c] *100*	
6Pgd2	*200 206*	*196*[c] *206*, *200 200*, *206 206*	*200 211*[c]
Pgi1	*100 100*	*85*[c] *100*, *85*[c] *105*[c], *90*[c] *90*[c], 90[c] 100, *100 115*[c], *105*[c] *105*[c]	
Lap1	*97 97*, *97 100*	*93*[c] *93*[c], *93*[c] *97*, *93*[c] *100*, *93*[c] *103*	
Lap1	*97 103*, *100 103*	*95*[c] *100*, *100 100*, *103 103*	
Prx1	*97 100*, *100 104*, *102 104*, *104 104*	*102 102*	
Est1	*88 100*, *93 100*, *93 105*, *96 96*, *96 100*, *96 105*, *96 106*, *100 100*, *100 102*, *100 105*, *100 106*, *105 105*, *105 106*, *106 106*	*85*[c] *100*, *88 88*, *88 105*, *93 93 105 110*[c], null	*93 106*
Ll	*Ll*, *LL Ll*, *ll*		
Total alleles	33	20	2
Total genotypes	38	45	3

The first of the two numbers in a column gives the allele for which one of the pair of homologous chromosomes is homozygous and the second number gives the allele for which the other pair of homologous chromosomes is homozygous, *e.g.*, *100 100* and *100 101* denote the genotypes *100 100*, *100 100* and *100 100*, *101 101*, respectively. Alleles listed in the first position are not necessarily located on the same member of a pair of homologous chromosomes.

[b] California data are from MILLER (1977), HAKIM-ELAHI (1980), PINERO (1982), HUTCHINSON (1982), HUTCHINSON *et al.* (1983a, b), and HAKIM-ELAHI and ALLARD (1983). Spanish data are partly from GARCIA (1988).

[c] Alleles observed in Spain alone or California alone. Addition "bands" which have been observed in locations other than California or Spain probably represent further alleles of the loci listed.

loci of this study in Spain and in California.

***Got1* and *Pgm1*:** Both of these loci are monomorphic for the same single-banded genotype (*100 100*, *100 100*) in all populations in both Spain and California. The ancestral and colonial gene pools are thus identical for these two loci.

***Got2*:** All California populations and 41 of the 42 Spanish populations are monomorphic for the single-banded genotype, *100 100*, *100 100*. An infrequent three-banded fixed heterozygote, *100 100*, *106 106* (banding intensity 1:2:1) was found along the genotype *100 100 100 100* in Spanish population 15. The ancestral and colonial gene pools are thus virtually identical for *Got2*.

***Mdh1*, *Mdh2* and *Mdh3*:** Only two genotypes have been found for *Mdh1*, an infrequent single-banded genotype (*106*, *106*, *106 106*) found only in Spain and a three-banded fixed heterozygote *100 100*, *106 106* (staining intensity 1:2:1) found in both Spain and California. Similarly only two genotypes have been found for *Mdh2*, the homozygote (*100 100*, *100 100*) found in both Spain and California and an infrequent fixed heterozygote (*100 100*, *106 106*) found in only one population in Spain (Population 15), but not in California. Thus the ancestral and colonial gene pools are also virtually identical for the *Mdh1* and *Mdh2*. *Mdh3* with five alleles (*95*, *100*, *103*, *107*, *108*) and five genotypes (Table 1) is variable in Spain. However, only one genotype (*100 100*, *107 107*) has been found in California; this genotype is also the most common

in Spain. Hybrid enzymes do not form between *Mdh1* and either *Mdh2* or *Mdh3*, but do form between *Mdh2* and *Mdh3*.

***Acp1* and *Acp2*:** Both of these loci are more variable in Spain than in California (Table 1). Allele 96 of *Acp1* occurs in the mesic genotype (*96 96, 100 100*) of California but has not been found in Spain, however, both alleles of the xeric genotype of California (*100 100, 101 101*) are frequent in Spain. Genotype *100 100, 104 104* of *Acp2* is part of both the xeric and mesic genotypes of California and it is also the most frequent genotype in Spain.

***6Pgd1* and *6Pgd2*:** Four alleles (*88, 92, 96, 100*) and eight genotypes of *6Pgd1* have been found in Spain but only two alleles (*96, 100*) and three genotypes have been found in California. Three alleles but only two genotypes of *6Pgd2* (*200, 206, 211*) have been found in California, whereas three alleles (*196, 200, 206*) and four genotypes have been found in Spain. Allele *211*, which is rare in California, has not been found in Spain. Genotype *200 200, 206 206*, which is part of both the xeric and mesic complexes, is very frequent in California and it is also present in most Spanish populations.

***Pgi1*:** Genotype *100 100, 100 100*, which is fixed in California, is also the most common genotype in Spain. Four additional alleles and six additional genotypes are found in low frequency in Spain.

***Prx1*:** This locus is variable (four alleles) in both Spain and California; however, one genotype (*102 102, 102 102*) not found in California occurs in Spain.

***Lap1*:** Three alleles (*97, 100, 103*) and four genotypes, *100 100, 103 103* (xeric), *97 97, 103 103* (mesic) *97 97, 97 97* and *97 97, 100 100* occur in California. The Spanish gene pool carries two alleles and seven genotypes additional to those observed in California. HAKIM-ELAHI and ALLARD (1983) have shown that allele *100* is present in only one genome and allele *103* in the other genome, whereas allele 97 is present in both genomes in California. These alleles occur in both genomes in Spain.

***Est1*:** This locus is highly variable in both Spain and California. Among the 10 alleles observed in Spain only three, all infrequent, have not been observed in California.

Pubescent *vs.* glabrous leaf (*Ll*): Xeric sites in California are usually fixed for glabrous leaf (*ll*) whereas mesic sites are often polymorphic for pubescent (*LL, Ll*) and glabrous leaf types. Both pubescent and glabrous types are widely distributed in Spain and approximately one half of Spanish populations are polymorphic for this locus.

The combined number of alleles observed for these 15 loci in Spain and California is 55, among which 53 (96%) were observed in Spain and 35 (64%) were observed in California. Among the alleles found in California only two (alleles *96* of *Acp1* and *211* of *6Pgd2*) have not been found in Spain. Nearly all of the 20 alleles which were present in Spain but not in California were infrequent or rare in Spain. The infrequency of these alleles in Spain and their absence in California suggests three possibilities: (1) that they failed to reach California due to sampling accidents associated with migration (founder effects); (2) that they reached California but did not become established in the colonial populations due to sampling accidents (genetic drift); or (3) that they were poorly adapted in California (as they appear to be in Spain) and were removed from the California populations by selection. The total number of homozygous single-locus genotypes (scored as duplicated homologs) found in Spain and California is 86, among which 83 (all except three) were found in Spain, but only 41 in California. As was the case with the alleles which were not found in California, the genotypes not found in California were rare or infrequent in Spain. Similarly genotypes which were frequent in Spain were usually also frequent in California. The Spanish and California gene pools are thus very closely similar in allelic and genotypic content respecting nearly all pairs of homologous loci. The genetic evidence thus suggests that the California gene pool was derived from a sample of the Spanish gene pool which included all of the common alleles and genotypes of Spain but only about one-half of the infrequent or rare alleles and genotypes for the 15 loci assayed. The same relationship appears to hold between the Spanish and eastern Mediterranean gene pools; the Spanish gene pool includes all of the common alleles but only a fraction of the infrequent or rare alleles of the more variable eastern Mediterranean populations.

Allelic and genotypic diversity within and among loci and populations: Allelic diversity per pair of homeologous loci is summarized in Table 2. The number of alleles per locus varied from one to 10 and the mean number of alleles per locus in the total sample of 4011 individuals was 2.06. Loci *Got1* and *Pgm1* were invariant in all of the 42 populations. *Got2, Mdh1,* and *Mdh2* were invariant in 41 of the 42 populations but each of these loci was polymorphic in a single population; the polymorphisms for *Got2* and *Mdh2* were due to presence of a single rare allele (in population 15 in each case) and for *Mdh1* due to presence of genotype *106 106, 106 106* in low frequency in population 9. The average number of alleles was 1.00 for *Got1* and *Pgm1*, 1.02 for *Got2* and *Mdh2*, and 2.00 for *Mdh1* (this locus was very nearly invariant for the fixed heterozygote *100 100, 106 106*).

Allelic diversity was much greater for the 10 remaining loci: the number of distinguishable alleles per locus varied from two to 10 and the mean number of alleles per locus, averaged over the 4011 individuals

TABLE 2

Number of different alleles per pair of homologous loci in the Spanish gene pool

Locus	No. of alleles/locus	Mean ± SEM
Got1	1	1.00 ± 0.00
Pgm1	1	1.00 ± 0.00
Got2	2	1.02 ± 0.15
Mdh2	2	1.02 ± 0.15
Mdh1	2	2.00[a] ± 0.00
Ll	2	1.40 ± 0.50
Acp1	3	2.14 ± 0.42
6Pgd2	3	2.31 ± 0.52
Prx1	4	2.62 ± 0.58
6Pgd1	4	2.69 ± 0.56
Acp2	4	3.02 ± 0.68
Pgi1	5	1.52 ± 0.71
Mdh3	5	2.62 ± 1.01
Lap1	5	3.07 ± 0.97
Est1	10	3.50 ± 1.13

[a] The fixed heterozygote *100 100 106 106* of *Mdh1* was present in very high frequency in all populations.

TABLE 3

Genotypic diversity (G_j) per locus

Locus	Range in G_j values for population	G_j values[a] ± SEM per locus	Polymorphic populations	
			No.	%
Got1	0	0 ± 0	0	0
Pgm1	0	0 ± 0	0	0
Got2	0–0.169	0.004 ± 0.026	1	2.4
Mdh2	0–0.169	0.004 ± 0.026	1	2.4
Mdh1	0–0.079	0.002 ± 0.012	1	2.4
Ll	0–0.490	0.049 ± 0.104	17	40.5
Acp1	0–0.698	0.186 ± 0.232	23	54.8
6Pgd2	0–0.491	0.173 ± 0.188	24	57.1
Prx1	0–0.700	0.271 ± 0.264	27	64.3
6Pgd1	0–0.755	0.216 ± 0.226	31	73.8
Acp2	0–0.666	0.242 ± 0.211	33	78.6
Pgi1	0–0.507	0.119 ± 0.176	18	42.9
Mdh3	0–0.662	0.162 ± 0.200	23	54.8
Lap1	0–0.756	0.397 ± 0.240	36	85.7
Est1	0–0.771	0.445 ± 0.253	38	90.5

[a] $\bar{G}_j$ = arithmetic mean of G_j values averaged over 42 populations.

assayed, varied from 1.40 to 3.50 (Table 2). Allelic and genotypic frequency data for the 15 loci in the 42 populations assayed are too extensive to be reported in full. Consequently, we have calculated a genetic diversity index, G_j, for each locus and population (MARSHALL and ALLARD 1970a; ZHANG and ALLARD 1986):

$$G_j = \sum_i^n p_i (1 - p_i) = 1 - \sum p_i^2)$$

in which p_i is the frequency of the ith genotype and n is the number of genotypes observed for the jth locus and population. The numerical value of this diversity index, which is zero when only one genotype is present, increases as the number of genotypes increases and approaches a maximum of unity when there are large numbers of equally frequent genotypes at a locus.

Table 3 gives the range and mean of estimated genetic diversity values, G_j, for the 15 loci assayed. G_j values were zero for *Got1* and *Pgm1* in all 42 populations. G_j values were also zero for *Got2*, *Mdh2*, and *Mdh1* in the 41 invariant populations of these loci and they were small, 0.169, 0.169, 0.079, respectively, in the populations for which these loci were polymorphic; mean G_j values for these three nearly invariant loci, averaged over all 42 populations, were also small (0.004, 0.004, 0.002, respectively). G_j values indicate that locus *Ll* was weakly polymorphic in our sample; the calculated values may, however, represent underestimates because differences in degree of pubescence, some of which may have been due to allelic diversity, were ignored (individuals were scored as either pubescent or glabrous). *Est1* was the most diverse locus: more populations (38/42 = 90.5%) were polymorphic for this locus than any other. The range in G_j values (0.000–0.771) and the mean G_j value, averaged over the 42 populations, were also larger for *Est1* than for any other locus. In general, G_j values increased in step with increasing numbers of alleles. There were, however, some irregularities in the increase, *e.g.*, although *Pgi1* was one of the most diverse among the variable loci when assessed in terms of number of alleles, it was one of the least diverse when assessed in terms of range in G_j values within populations (0.000–0.507), mean $\bar{G}_j$ averaged over all populations (0.119) and in number and percentage of populations in which it was polymorphic (18/42 = 47.9%). G_j values, which take both the number and frequency of genotypes into account, thus appear to be a more sensitive measure of diversity than numbers of alleles and/or genotypes alone.

Table 4 gives the range of G_j values observed in the 42 populations. Only one population (number 6) was monomorphic for all 15 loci (G_j = 0.000). None of the populations was polymorphic for all loci; 38 out of 42 of the populations were polymorphic for *Est1*, which was the most variable locus in this respect.

An important feature of variability in California populations is that genotypic state at a given locus often depends on genotypic state at other loci. We consequently tabulated the kinds and frequencies of 15-locus homozygous genotypes in each of the 42 Spanish populations as a basis for comparing the multilocus structures of the Spanish and Californian populations. The number of different 15-locus genotypic combinations observed in Spanish populations was 444; the single most frequent 15-locus genotype was found in 268 among the 4011 individuals scored (6.7%). Table 5 gives the numbers of populations in

TABLE 4

Genetic diversity (G_j) per population

Range in G_j values per population	No. of populations
0	1
0.000–0.049	8
0.050–0.099	8
0.100–0.149	5
0.150–0.199	5
0.200–0.249	5
0.250–0.299	9
0.310	1

which various numbers of 15-locus genotypes were observed. The number per population varied from only one to 39 (mean number per population was 11.9). To facilitate summarization of multilocus genotypic diversity within and among populations we calculated a multilocus genotypic diversity index for each population,

$$H_j = 1 - \sum g_i^2,$$

in which g_i is the frequency of the ith 15-locus genotype in the jth population. The properties of this index are parallel to those of G_j; the value of H_j is zero when only one 15-locus genotype is present in a population and values of H_j generally increase as the number of genotypes increases, approaching unity when there are numerous equally frequent genotypes in the population. Observed values of H_j (Table 5) ranged from 0.000 for population 15 (only one 15-locus genotype observed) to 0.999 for population 32 (39 15-locus genotypes observed). In about three-fourths of the Spanish populations (33 out of 42) there were six or more multilocus genotypes and H_j values were larger than 0.40. The great majority of populations in California (80%) are fixed or nearly fixed for the xeric genotype, about 10% are fixed or nearly fixed for the mesic genotype, and only a few populations (mostly nonxeric or nonmesic populations) were polymorphic for many loci. Genetic variability is therefore generally greater within Spanish than within Californian populations whereas differentiation among populations is more distinct among Californian than among Spanish populations.

HEDRICK's (1971) identity index (I) was calculated for each of the 861 possible pairwise combinations among the 42 Spanish populations. Observed I values ranged from 0.455 (population pair 12 and 40) to 0.988 (population pairs 1 and 2 and 1 and 3). Mean values of I ranged from 0.562 for population 12 to 0.794 for population 5, indicating that, on the average, population 12 was least and population 5 was most like all other populations genotypically. I values were also calculated for the California xeric-mesic pair of genotypes (I = 0.600 for the 15 loci of this study) and for each of the 42 Spanish populations with the

TABLE 5

Multilocus genotypic diversity (H_j) per population

No. of multilocus genotypes per population	No. of populations	H_j values	No. of populations
1	1	0	1
2	2	0–0.09	1
3	2	0.10–0.19	2
4	2	0.20–0.29	1
5	2	0.30–0.39	4
6–10	17	0.40–0.49	3
11–15	5	0.50–0.59	5
16–20	5	0.60–0.69	7
21–25	2	0.70–0.79	6
26–39	4	0.80–0.89	10
		0.90–1.00	2
Total	42		42

California xeric genotype ($\bar{I}$ = 0.662, range 0.543 (Population 12) to 0.775 (population 5) and with the California mesic genotype ($\bar{I}$ = 0.677, range 0.553 (population 6) to I = 0.772 (population 35). In comparisons of the 42 Spanish populations with the xeric or the mesic genotypes, I values were lower than 0.600 for only five populations (6, 12, 17, 21, 40) and I values were equal to or lower than 0.600 for both the xeric and mesic comparisons for only three populations (6, 12, 40). The above results indicate that the 42 Spanish populations are, on the average, slightly closer genetically to each other than they are to the California xeric and mesic genotypes and also that the Spanish populations are genetically intermediate between the xeric and mesic genotypes and differ from them almost equally.

Table 6 gives mean values for various diversity and identity parameters for populations from three geographical regions of Spain (Northern Plateau, Southeastern, Southwestern). The populations from the two southern regions were more variable, on the average, than the northern populations. It is particularly relevant to the present study that the populations from Southwestern Spain are more similar genetically to both the California xeric and California mesic populations than are the Southeastern and Northern Plateau populations.

DISCUSSION

Among the 15 loci of this study, two (*Got1* and *Pgm1*) were invariant and three (*Got2*, *Mdh1*, and *Mdh2*) were nearly invariant in the Spanish populations. All five of these loci appear to be invariant in California but at least four of them (no information is available for *Pgm1*) are somewhat more variable in most populations in Israel and Southwestern Asia than in the Spanish populations. The same five-locus genotype, which is fixed in California, was found in 99.7% of the 4011 individuals scored in Spain and it is also

TABLE 6

Geographical distribution of genetic variability in Spain

Measure	Northern plateau[a]	Southeastern[b]	Southwestern[c]
$\bar{A}^d$	1.81	2.12	2.14
$\bar{P}^e$	37	46	47
$\bar{MG}^f$	9.93	11.56	13.42
$\bar{G}_j^g$	0.147	0.199	0.183
$\bar{H}_j^h$	0.507	0.681	0.680
I_x^i	0.630	0.672	0.681
I_m^j	0.661	0.670	0.692

[a] Populations 1, 2, 3, 6, 8, 10, 20, 21, 25, 32, 33, 34, 40, 42.
[b] Populations 4, 7, 11, 13, 17, 30, 31, 38, 39.
[c] Populations 5, 9, 12, 14, 15, 16, 18, 19, 22, 23, 24, 26, 27, 28, 29, 35, 36, 37, 41.
[d] A = mean number of alleles per locus.
[e] $\bar{P}$ = mean percentage of polymorphic loci.
[f] $\bar{MG}$ = mean number of multilocus genotypes per population.
[g] $\bar{G}_j$ = mean genotypic diversity value per population.
[h] $\bar{H}_j$ = mean multilocus diversity value per population.
[i] I_x = mean identity with California xeric genotype.
[j] I_m = mean identity with California mesic genotype.

by far the most frequent genotype in Israel and Southwestern Asia. Thus the Spanish, Californian, and Eastern Mediterranean-Southwestern Asian gene pools are identical or virtually identical for these five loci. Four additional loci (*Mdh3*, *Acp2*, *6Pgd2* and *Pgi1*) are invariant in California, weakly variable in Spain and weakly to moderately variable in Israel and Southwestern Asia. The single four-locus genotype which predominates in California is the single most common one in Spain (29.4% of individuals scored) and also in Israel and Southwestern Asia. The gene pools of Spain and California are thus somewhat less similar to the Eastern Mediterranean gene pool than they are to each other for these four loci. Most populations are moderately to highly variable for the six remaining loci (*Acp1*, *6Pgd1*, *Prx1*, *Lap1*, *Est1*, *Ll*); levels of variability for these loci were generally lowest in California, intermediate in Spain and highest in Israel and Southwestern Asia. The same alleles tend to be moderately to highly frequent in each of the three gene pools. We interpret the above results to indicate that the gene pool of California was derived from a sample of the Spanish gene pool which included all of the common alleles but only about half of the infrequent or rare alleles of Spain. The data are less extensive and precise for the Eastern Mediterranean-Southwestern Asian gene pool; however, the Spanish gene pool appears to include all of the common alleles but only a fraction of the infrequent or rare alleles of the Eastern Mediterranean-Southwestern region. The loss of many of the infrequent or rare alleles may have resulted in part from sampling accidents (founder effects) during the spread of *A. barbata* westward to Spain from its probable Eastern Mediterranean-Southwestern Asian center of origin and also in part from founder effects during its subsequent migration from Spain to California; additional losses may also have resulted from sampling accidents (genetic drift) after the species reached Spain and then California and it also seems likely that selection contributed to the elimination of some of the less well adapted among the infrequent or rare alleles. It is also likely that selection played a role in the maintenance at moderate to high frequency of some of the better adapted among the common alleles in the gene pools of the three regions (ALLARD 1988).

Historical records indicate that *A. barbata* was brought to California as a contaminant in shipboard litter and in cereal seed shipments and that voyages to the New World originated almost exclusively in Southwestern Spanish ports during the period when the species was introduced to California; thus, the ancestors of the colonial populations probably came from Southwestern Spain. Furthermore, the climate of the areas of California which were colonized by *A. barbata* is similar to that of Southwestern Spain, leading to the hypothesis that the gene pool of California should be more similar genetically to the gene pool of that region than those of other regions of Spain. Genetic evidence supports this hypothesis: some of the alleles which are present in California are present in Southern but absent in Northern Spain (*e.g.*, allele *100* of *Prx1*); also genetic similarity between the California xeric and mesic genotypes and the Southwestern Spanish populations is higher than with the populations of Southeastern and the Northern Plateau regions of Spain.

The results of this study also show that, even though the present-day Spanish and Californian gene pools are closely similar to each other in allelic composition and allelic frequencies when they are compared on a locus-by-locus basis, the arrays of multilocus genotypes found in the present-day ancestral gene pool are very different from those found in the colonial gene pool. Two-locus and three-locus allelic associations are frequently the same in the Spanish and Californian populations; however, higher order allelic associations among loci are usually different, *e.g.*, neither the xeric nor mesic multilocus genotypes adapted to contrasting extreme habitats in California have been found in Spain. It therefore appears that changes in allelic frequencies, which are usually considered to be the elemental process of evolution, played a relatively minor role in the adaptive changes that are known to have occurred in the two centuries or more since the samples which founded the colonial populations were taken to the new world. Instead a different process, involving reorganization of the ancestral allelic ingredients of the Spanish gene pool into novel multilocus allelic combinations adapted to specific habitats, appears to have been the principal force involved in the evolution of adaptedness in the colonial Californian

populations. Ecogenetic aspects of the multilocus organization of the Spanish gene pool and its reorganization in the colonial populations will be analyzed and discussed in a separate paper in which we bring together genetic and environmental data.

This work was supported in part by grants from the U.S. Spain Joint Committee for Scientific and Technological Cooperation (CCB8504-101), the Comisión Interministerial de Ciencia y Tecnologiá (PB85-0153), the U.S. National Science Foundation (BSR 83110869), and the U.S. Public Health Service (NIH GM-32429).

LITERATURE CITED

ALLARD, R. W., 1988 Genetic changes associated with the evolution of adaptedness in cultivated plants and their wild progenitors. J. Hered. **79:** 225–238.

ALLARD, R. W., G. R. BABBEL, M. T. CLEGG and A. L. KAHLER, 1972 Evidence for coadaptation in *Avena barbata*. Proc. Natl. Acad. Sci. USA **69:** 3043–3048.

ALLARD, R. W., R. D. MILLER and A. L. KAHLER, 1978 The relationship between degree of environmental heterogeneity and genetic polymorphism, pp. 49–73 in *Structure and Functioning of Plant Populations*, edited by A. J. FREYSEN and J. W. WOLDENDORP. North-Holland, New York.

BAUM, B. R., 1977 *Oats: Wild and Cultivated* (Monograph No. 14). Canada Department Agriculture, Ottawa.

CLEGG, M. T., and R. W. ALLARD, 1972 Patterns of genetic differentiation in the Slender Wild Oat species, *Avena barbata*. Proc. Natl. Acad. Sci. USA **69:** 1820–1824.

CLUSTER, P. D., 1984 Correlation between genetic variation for allozyme markers and quantitative characters in *Avena barbata* Pott ex Link. Ph.D. dissertation, University of California, Davis.

DURRENBERGER, R. W., 1960 *Patterns on the Land.* Roberts, Northridge, Calif.

GARCIA, P., 1988 Analisis de la estructura genetica de poblaciones espanolas de *Avena barbata* Pott ex Link. Ph.D. dissertation, University of Leon, Spain.

HAKIM-ELAHI, A., 1980 Temporal changes in the population structure of the slender wild oat (*Avena barbata*) as measured by allozyme polymorphisms. Ph.D. dissertation, University of California, Davis.

HAKIM-ELAHI, A., and R. W. ALLARD, 1983 Distribution of homoeoalleles at two loci in a diploidized tetraploid: leucine aminopeptidase loci in *Avena barbata*. J. Hered. **74:** 379–380.

HAMRICK, J. L., and R. W. ALLARD, 1972 Microgeographical variation in allozyme frequencies in *Avena barbata*. Proc. Natl. Acad. Sci. USA **69:** 2100–2104.

HAMRICK, J. L., and R. W. ALLARD, 1975 Correlations between quantitative characters and enzyme genotypes in *Avena barbata*. Evolution **29:** 438–442.

HEDRICK, P. W., 1971 A new approach to measuring genetic similarity. Evolution **25:** 276–280.

HUTCHINSON, E. S., 1982 Genetic markers and ecotypic differentiation of *Avena barbata* Pott ex. Link. Ph.D. dissertation, University of California, Davis.

HUTCHINSON, E. S., A. HAKIM-ELAHI, R. D. MILLER and R. W. ALLARD, 1983a The genetics of the diploidized tetraploid *Avena barbata*. Acid phosphatase, esterase, leucine aminopeptidase, peroxidase, and 6-phosphogluconate dehydrogenase loci. J. Hered. **74:** 325–330.

HUTCHINSON, E. S., S. C. PRICE, A. L. KAHLER, M. I. MORRIS and R. W. ALLARD, 1983b An experimental verification of segregation theory in a diploidized tetraploid: esterase loci in *Avena barbata*. J. Hered. **74:** 381–383.

JAIN, S. K., and R. S. SINGH, 1979 Population biology of *Avena*. VII. Allozyme variation in relation to the genome analysis. Bot. Gaz. **140:** 356–362.

KAHLER, A. L., R. W. ALLARD, M. KRZAKOWA, C. F. WEHRHAHN and E. NEVO, 1980 Associations between isozyme phenotypes and environment in the slender wild oat (*Avena barbata*) in Israel. Theor. Appl. Genet. **56:** 31–47.

MARSHALL, D. R., and R. W. ALLARD, 1969 The genetics of electrophoretic variants in *Avena*. I. The esterase E_4, E_9, E_{10}, phosphatase P_5, and anodal peroxidase APX_5 loci in *A. barbata*. J. Hered. **60:** 17–19.

MARSHALL, D. R., and R. W. ALLARD, 1970a Isozyme polymorphisms in natural populations of *Avena fatua* and *A. barbata*. Heredity **25:** 373–382.

MARSHALL, D. R., and R. W. ALLARD, 1970b Maintenance of isozyme polymorphisms in natural populations of *Avena barbata*. Genetics **66:** 393–399.

MILLER, R. D., 1977 Genetic variability in the slender wild oat *Avena barbata* in California. Ph.D. dissertation, University of California, Davis.

PINERO, D., 1982 Correlations between enzyme phenotypes and physical environment in California populations of *Avena barbata* and *Avena fatua*. Ph.D. dissertation, University of California, Davis.

ROBBINS, W. W., 1940 Alien plants growing without cultivation in California. Calif. Univ. Agr. Expt. Sta. Bull. 637.

SHAW, C. R., and R. PRASAD, 1970 Starch gel electrophoresis of enzymes. A compilation of recipes. Biochem. Genet. **4:** 297–320.

SHAW, D. V., A. L. KAHLER and R. W. ALLARD, 1981 A multilocus estimator of mating system parameters in plant populations. Proc. Natl. Acad. Sci. USA **78:** 1298–1302.

ZHANG, Q., and R. W. ALLARD, 1986 The sampling variance of the genetic diversity index. J. Hered. **77:** 54–55.

Communicating editor: M. T. CLEGG