

Contrasting patterns of genetic diversity in neutral markers and agromorphological traits in wild and cultivated populations of *Medicago sativa* L. from Spain

Eric Jenczewski, Michel Angevain, André Charrier,
G rard G nier, Jo lle Ronfort, Jean-Marie Prosperi*

Laboratoire de ressources g n tiques et d'am lioration
des luzernes m diterran ennes, Station de g n tique et am lioration des plantes,
Inra, Domaine de Melgueil, 34130 Mauguio, France

Abstract – Crop/weed gene flow is often of great concern for a better understanding of germplasm diversity. This paper provides an overview of the levels and partition of genetic diversity assessed between wild and cultivated populations of *Medicago sativa* L. in Spain for both molecular markers and agromorphological traits. We report that most wild populations preserve a great agronomic and morphological originality despite their occasional parapatry with cultivated forms and the occurrence of crop/weed gene flow in many locations. However, populations that display either a continuous range of intermediate plants between the typical wild and cultivated forms, or a mosaic of individuals that can be assigned to each gene pool respectively, were also identified. Providing that reproductive isolation is not strict, we hypothesise that contrasted selective pressure between wild and cultivated populations mainly accounted for the maintenance of their phenotypic integrity.   Inra/Elsevier, Paris
gene flow / genetic diversity / *Medicago sativa* / genetic markers

R sum  – Comparaison de la diversit  g n tique mesur e sur la base de marqueurs neutres et de caract res agro-morphologiques dans des populations sauvages et cultiv es de luzerne (*Medicago sativa* L.) en Espagne. L'existence d'un flux g nique entre des populations cultiv es et sauvages joue souvent un r le pr pond rant dans l' volution de la diversit  g n tique de ces derni res. Nous pr sentons ici une vue g n rale des niveaux, et de l'organisation de la diversit  g n tique mesur e sur la base de marqueurs neutres et de caract res agro-morphologiques dans des populations sauvages et cultiv es de luzerne (*M. sativa* L.)

* Correspondence and reprints

en Espagne. La plupart des populations sauvages maintiennent une claire originalité agronomique et morphologique bien qu'elles soient parapatriques de populations cultivées et qu'elles fassent parfois l'objet d'introgessions régulières de leur part. Néanmoins notre étude a permis de mettre en évidence un ensemble de populations sauvages moins typiques, présentant un continuum de formes intermédiaires ou une mosaïque d'individus caractéristiques de chaque forme. Notre étude suggère que l'isolement reproducteur entre populations sauvages et cultivées n'est pas total et qu'il est vraisemblable que de fortes pressions de sélection engendrées par le milieu jouent un rôle déterminant dans le maintien de l'intégrité phénotypique des formes sauvages. © Inra/Elsevier, Paris

diversité génétique / flux de gènes / marqueurs génétiques / *Medicago sativa*

1. INTRODUCTION

Optimal strategies for sampling genetic resources, for their use and conservation, require an accurate understanding of the way genetic diversity is created or maintained in a crop's wild relatives (Frankel et al., 1995). The amount and organisation of genetic diversity depends on breeding structure (Levin, 1988; Schoen and Brown, 1991), population dynamics (Harrison and Hastings, 1996) and the extent of gene flow between populations (Slatkin, 1987). Crop/weed gene flow could consequently be of great concern for a better understanding of germplasm diversity. Natural hybridisations have been shown to promote the evolution of several weed species into more aggressive and invasive forms (reviewed in Barrett, 1983). They now represent an avenue of escape for engineered genes (Ellstrand and Hoffman, 1990). On the other hand, crop/weed gene flow also provides a great opportunity to make introgressions of wild genetic resources in plant improvement easier. However, genetic isolation cannot be restricted to reproductive barriers (Harrison, 1990) and the direct achievement of hybrids is only a prerequisite step in the introgressive process (see McCoy et al., 1991; Rieseberg et al., 1996). Bearing in mind Harlan and de Wet's (1971) concept of crop gene pools, crop/weed hybridisations will be more likely to occur and to play a preeminent role when the wild and cultivated germplasms are more closely related.

Crop/weed gene flow is mostly pollen-mediated (Ellstrand and Hoffman, 1990) and thus requires cross-compatibility and flowering period overlaps within a crop/weed complex (Klinger et al., 1992). Self-incompatibility, high outcrossing rates and generalist pollination mechanisms significantly increase the likelihood of gene exchanges (Arias and Rieseberg, 1994). It has long been considered that distance alone could prevent crop/weed hybridisations but Klinger et al. (1992) also emphasised that crop/weed mating is not necessarily restricted to narrow zones of close parapatry; long-distance gene dispersal can be effective as well (Manasse, 1992; Arias and Rieseberg, 1994), but critically depends on the combined effects of the population sizes and the distance of the recipient wild populations from the cultivated source (Klinger et al., 1992).

These prerequisites are satisfied in the wild and cultivated populations of *Medicago sativa* L. from Spain: both forms contain autotetraploid, outcrossing and partially autoincompatible plants (so that outcrossing rates classically vary from 85 to 100 %) which are called 'Mielgas' and 'Alfalfas', respectively. Wild

populations occur along roadsides, orchards and fields, and grow in parapatry with cultivated plants in many locations. Their flowering periods overlap and both forms share the same generalist pollinators (honey bees, bumblebees, megachiles). Cross-compatibility has been verified experimentally (Prosperi et al., 1996). Nonetheless, wild populations seem to maintain morphological originality despite their parapatry with cultivated forms of the same species. Although a wide range of variability has been observed, most of the wild plants usually display a crawling growth habit and rhizomes, and thus clearly differ from the cultivated erect ones. Ubiquitous natural hybridisations (Langevin et al., 1990 and references therein; Klinger et al., 1992; Arias and Rieseberg, 1994) even at the interspecific level (Doebley, 1990; Till-Bottraud et al., 1992; Wilson and Manhart, 1993) suggests that striking morphological differences are often irrelevant to reproductive affinity (Doebley, 1990). The nature and dynamics of gene flow between the wild and cultivated populations of *M. sativa* can, however, be questioned since gene flow is expected to produce homogenising effects: such cases of morphological convergence have been documented in wild and cultivated populations of rice (Langevin et al., 1990), whereas a soft reproductive isolation maintains the phenotypic integrity of both forms of pearl millet (Robert et al., 1991).

This paper provides an overview of the levels and partition of genetic diversity assessed between wild and cultivated populations of *M. sativa* L. in Spain for both molecular markers and agromorphological traits. Using allozyme and random amplified polymorphic DNA (RAPD) markers, we estimated the neutral genetic differentiation between these two forms and assessed its relative extent as compared to the spatial genetic structure within each gene pool. Theoretically, this structure results from a balance between gene flow and genetic drift; it provides a null model to assess which forces govern the differentiation achieved for agromorphological characters, which potentially mirror selection effects as well. The results will be discussed in the context of crop/weed gene flow and its balance with random genetic drift and heterogeneous selection.

2. MATERIALS AND METHODS

2.1. Plant material and experimental design

One hundred and three wild populations of *M. sativa* were sampled in Spain (figure 1), over the 190 sites where *Medicago* species were collected. An ecological description of the sites where these populations were collected is given in Prosperi et al. (1989). They encompassed the widest range of variability observed *in natura*. Introgressions from *M. falcata*, a related species of the same complex (Falcatae), were sometimes identified through the occurrence of variegated flowers and C-shaped pods. In each population, mature pods were randomly collected from a great number of plants (more than 50), then mixed and threshed in the laboratory.

Seeds were put in Petri dishes for 2 days to germinate and afterwards transplanted in a randomised four-blocks design with 12 control cultivars (including Spanish landraces, French Flemish and Provence cultivars and two North American cultivars selected for grazing); four replications of 25 plants, bedded out in 1 m² plot, were implemented for each population (one replicate

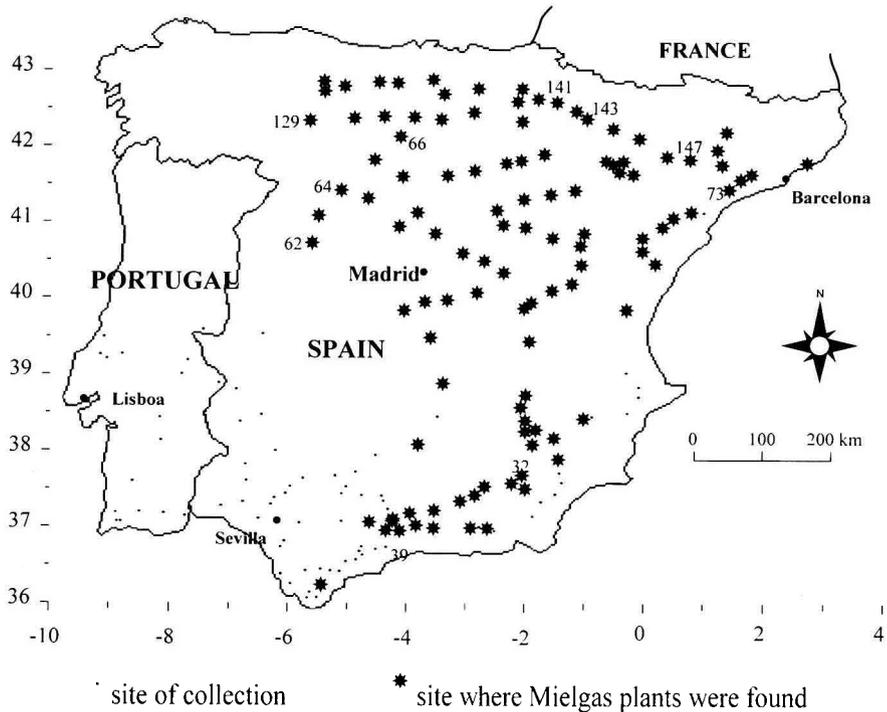


Figure 1. Geographic distribution of wild populations of *Medicago sativa* in Spain. Populations analysed with genetic markers are indicated by their site number.

per block). Evaluation was conducted over 5 years at Montpellier (France) with restrained irrigation during the first 2 years and dried conditions afterwards. Apart from the 1st and 3rd sowing years, plots were harvested in June, July, September and November. Seed production was measured the 3rd year after sowing, and plots were thus only harvested in September and November.

Allozyme and RAPD surveys were implemented on a subset of nine wild populations; they were randomly chosen within each group (two to three populations per group) resulting from the survey of their agronomic and morphological features (see Section 3.1). Four Spanish cultivated populations of *M. sativa* and a French wild population of *M. falcata* (referred to as Malzeville hereafter) were also examined. Two newly collected wild (Es141) and cultivated (Cult141) populations were included in this marker-based analysis because they grew in close parapatry: wild plants were collected all around the cultivated field. Given that they were prospected in 1995, these two populations were not included in the agromorphological survey. Thirty-six plants per population were used for these analyses.

2.2. Agronomic and morphological traits

Sixty-two agronomic and morphological traits were initially scored, but statistical analysis was only performed with a synthetic data set that had been purged of the most obvious redundancies. This was essentially achieved

by averaging all measurements of a given trait over time as far as they were highly correlated. Thirty-one variables were finally analysed. Most of them are quantitative measurements of vegetative dry matter produced between two successive cuts or in a year, seed production and proportion of plants with rhizomes or variegated flowers. Growth habit, plant volume, winter growth and autumn regrowth (as documented by the length of the growing stem), colonisation ability, perenniality and leaf size were evaluated on a one-to-nine scale. To account for within-population variations, we surveyed differences between the maximum and the minimum values scored within a given plot for growth habit, volume and winter growth.

2.3. Allozymes survey

Fresh pieces of actively growing young leaf tissue were ground in cold 0.1 M Tris-HCl (pH = 7.2) extraction buffer. The supernatant was absorbed onto filter paper wicks, kept frozen at -80°C overnight, and loaded afterwards onto 13 % starch gels. In this study, we only focused on the glucose-6-phosphate dehydrogenase (*G6PD*), the isocitrate dehydrogenase (*IDH*) and the leucine aminopeptidase (*LAP*) enzyme systems. *G6PD* and *IDH* were resolved with a Tris-citrate (pH = 7) buffer system whereas *LAP* was stained with a lithium borate (pH = 8.3) one. Banding intensity was used to identify unbalanced heterozygotes.

2.4. RAPD survey

Total deoxyribonucleic acid (DNA) was extracted according to the procedure described by Tai and Tanksley (1990). RAPD assays were performed in a final volume of 25 μL containing 10 mM Tris-HCl pH 8.0, 50 mM KCl, 0.2 mg/mL gelatine, 0.1 % Triton X-100, 1.5 mM MgCl_2 , 200 μM of each dNTP (Boehringer, Mannheim, Germany), 50 μM single Bioprobe (Montreuil sous Bois, France) primer, 0.4 U Taq[®] (Appligene-Oncor, France) and approximately 30 ng total DNA template. Each reaction mixture was overlaid with mineral oil (Sigma). Amplifications were carried out in a Biometra-trio thermal cycler, programmed for 40 cycles with the following temperature profile: 30 s at 94°C , 45 s at 40°C and 75 s at 72°C . The initial and final steps were at 94°C for 3 min and 72°C for 6 min, respectively. RAPD products were resolved through a 1.8 % agarose gel, run at 8.5 V/cm for 5 h in 0.5x Tris-borate (0.045 M) EDTA (0.000 02 M), pH 8.0. A 1 kb DNA ladder (BRL) was run as a molecular size control. DNA was stained with ethidium bromide (3 mg/L) and photographed in ultraviolet light ($\lambda = 260\text{ nm}$). The presence or absence of bands was scored from negatives.

2.5. Data analysis

Statistical analysis was performed on mean population values, averaged over replications for each trait using the Proc MEAN procedure (SAS[®] Institute, 1989). Principal component analysis (PCA) was conducted with the Proc PRINCOMP procedure. Cultivated control populations were not included in the analysis but were plotted afterwards on the PCA axes. Mean values over accessions within each cluster were estimated with the Proc MEAN procedure.

Allozymes allow accurate estimations of both genotypic and allelic frequencies; F -statistics (Wright, 1951) could thus be used to estimate population differentiation. F_{st} were computed using Weir and Cockerham's (1984) framework adapted for autotetraploid organisms. Significance was checked by a Fischer's exact test for population differentiation (Raymond and Rousset, 1995) expanded to autotetraploid organisms. Gene pool differentiations were assessed through an hierarchical analysis (Weir, 1996). Confidence intervals for the among-groups differentiation were built through permutations (Sokal and Rohlf, 1995).

RAPD markers are dominant (Williams et al., 1991; Yu and Pauls, 1993). Fragments were scored as discrete character state (presence versus absence). Provided that populations do not depart from panmictic equilibrium, allelic frequencies can then be estimated from the proportion of individuals sharing null amplification. PCA was performed with these estimates after an angular transformation. To assess the overall variability within and among populations, Nei's index (1973) of gene diversity was estimated for each RAPD locus and over all loci. The total observed diversity (H_T) was partitioned into within- and among-population components (H_S and D_S , respectively). Estimates of population differentiation were expressed as the extent of among-population diversity relative to total-population diversity (G_{st}).

3. RESULTS

3.1. Agromorphological traits

The first three axes of the PCA accounted for 74 % of the overall variability. Principal component 1 (PC1) (58 % of the overall variation) was mostly associated with growth habit, dormancy, vegetative vigour and dry matter production. Principal component 2 (PC2) (9 % of the overall variation) accounted for perenniality, dry matter production for the first cuts and the proportion of rhizomes. Variables that depicted population heterogeneity for growth habit or volume strongly contributed to principal component 3 (PC3) (7 % of the overall variation). Plots of the first two principal components (*figure 2*) suggested that three major groups could be identified.

The partition of a hierarchical classification of populations, based on Euclidean distances calculated from their coordinates on the first three principal components, confirmed that 73 % of the total variation was between these three groups. *Table I* summarises their main features. Wild typical populations (group 1) exhibited crawling growth habit and rhizomes, which allowed an increased soil colonisation. They globally displayed a long winter vegetative dormancy and a weak ability for regrowth after cutting (notably in autumn) as compared to control cultivars. This explains why most of their annual vegetative biomass was harvested at the first cut. However, a great variability was still observed for growth rhythm, and the first cut accounted for 51–75 % of the total annual dry matter production (*table I*). These populations were also characterised by a greater perenniality and a seed production that was only 30 % lower than that of the control cultivars. Populations from group 2 did not substantially differ from the typical wild ones; they were, however, separated by PC3 (*figure 2*). These populations displayed significant differences between

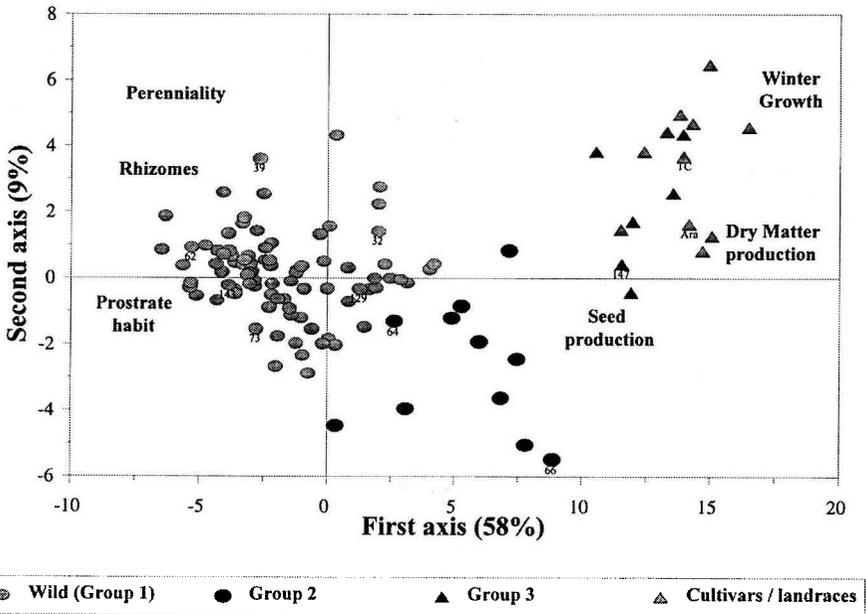


Figure 2. Plots of the first and second axis of a principal component analysis based on 31 agronomic and morphological traits and 103 wild populations of *Medicago sativa*. Cultivated populations were plotted despite their noninclusion in the analysis.

extreme individuals for their growth habit or volume. Populations from group 3 appeared to be closely associated with control cultivars even though they had been sampled in the wild. They displayed agronomic potentials barely inferior to those of the control cultivars and appeared to be homogeneous, apart from their winter growth.

3.2. Allozyme survey

Three enzyme systems were examined in this study; they resolved five putative loci (*IDH*, *PGD1*, *PGD2*, *LAP1*, *LAP2*). None of them was monomorphic but, in most populations and for most loci, one allele was far more common than the others (figure 3). Polymorphism was not very high despite the fact that four or five alleles were detected at each locus over all populations. We did not find any 'diagnostic' allele (i.e. common in a population but absent from the other) either between *M. falcata* (which was represented by one single population) and *M. sativa*, or between wild and cultivated populations (figure 3). However, *LAP1* and *LAP2* displayed large variance in allele frequencies between Malzeville and all the other populations and the highest differentiation was thus measured at the interspecific level; F_{st} values varied from 0.194 to 0.301 between Malzeville and the other populations.

Genetic differentiation over all populations of *M. sativa* was low ($F_{st} = 0.053$) but statistically significant ($P < 0.0001$). Cultivated populations were not significantly differentiated from one another ($F_{st} = 0.0034$; $P = 0.273$).

Table I. Comparative mean group values for main agronomic and morphological traits. The three clusters result from principal component analysis (PCA).

	Group 1		Group 2		Group 3		Control cultivars	
	Mean	(Min-Max)	Mean	(Min-Max)	Mean	(Min-Max)	Mean	(Min-Max)
Winter growth ¹	1.91	(1.17-3.08)	3.15	(2.58-4.17)	5.18	(4.42-7.33)	5.32	(2.94-8.58)
Growth habit ²	6.28	(5.25-7.17)	5.40	(4.83-6.00)	4.22	(3.92-4.67)	3.94	(2.92-4.44)
Autumn regrowth ³	1.90	(1.08-3.33)	3.60	(2.17-5.42)	6.30	(5.50-7.75)	6.77	(5.25-7.92)
Rhizomes ⁴	6.15	(4.25-8.75)	4.55	(3.25-6.33)	2.75	(2.00-3.75)	2.28	(1.13-4.17)
Leaf size ⁵	3.15	(1.88-4.50)	4.20	(3.50-5.00)	5.92	(5.38-6.63)	6.16	(5.67-6.63)
Dry matter first cut ⁶	1.65	(1.26-2.23)	1.78	(1.35-2.04)	1.70	(1.48-1.89)	1.82	(1.21-2.11)
Dry matter second cut ⁶	0.61	(0.40-0.97)	0.90	(0.52-1.14)	1.12	(1.06-1.18)	1.27	(1.03-1.37)
Dry matter third cut ⁶	0.31	(0.12-0.59)	0.54	(0.27-0.89)	0.76	(0.69-0.83)	0.94	(0.7-1.05)
Dry matter fourth cut ⁶	0.07	(0.00-0.30)	0.27	(0.08-0.51)	0.40	(0.36-0.45)	0.49	(0.37-0.60)
Total dry matter ⁶	2.65	(1.95-3.74)	3.47	(2.22-4.57)	3.95	(3.80-4.13)	4.52	(3.61-4.92)
Seed production ⁷	104.30	(33.1-162.7)	130.90	(66.7-210.8)	140.70	(100.7-179.9)	137.30	(94.2-192.0)
1 000 seeds weight ⁷	1.51	(1.00-1.76)	1.68	(1.54-1.85)	1.85	(1.76-2.03)	1.97	(1.88-2.09)
Persistence ⁸	72.3	(38.0-94.0)	57.8	(45.0-68.0)	53.8	(44.0-60.0)	48.5	(34.3-61.0)
Colonisation ⁹	79.5	(35.0-97.0)	60.6	(41.2-80.0)	51.5	(43.7-66.2)	43.2	(24.3-60.0)
% Production first cut	63.2	(50.7-74.9)	52.1	(44.9-60.9)	42.5	(38.8-46.3)	40.2	(33.5-45.2)
% Production fourth cut	2.30	(0.0-4.0)	7.5	(3.4-11.1)	10.1	(8.6-11.6)	11.0	(7.7-14.5)
Winter growth dif. ¹⁰	0.71	(0.00-4.00)	3.55	(1.25-6.00)	4.25	(2.00-6.00)	1.81	(0.33-3.71)
Growth habit dif. ¹⁰	2.00	(0.75-3.67)	2.95	(1.50-4.00)	1.15	(1.00-1.50)	1.16	(0.00-1.69)
Volume dif. ¹⁰	1.61	(0.13-3.50)	4.33	(3.38-5.75)	3.80	(3.13-4.88)	3.51	(2.88-4.36)

Variables were scored as follows: ¹ Winter growth: from 1 = dormant to 9 = nondormant plant; ² growth habit: from 1 = erect to 9 = crawling plant; ³ autumn regrowth: from 1 = no regrowth to 9 = fast regrowth; ⁴ rhizomes: from 1 = no rhizome to 9 = rhizomatous plant; ⁵ leaf size: the length of the fourth subapical leaf varies from < 5 mm = 1 to > 40 mm = 9; ⁶ dry matter production is expressed in kg - cuts were performed in 1988 on 2 June, 11 July, 23 August and 9 November, respectively; ⁷ seed production is expressed in g; ⁸ persistence: % of remaining plants after 5 years; ⁹ colonisation: % of the soil surface covered; ¹⁰ expressed differences between the most extreme plants of a given plot.

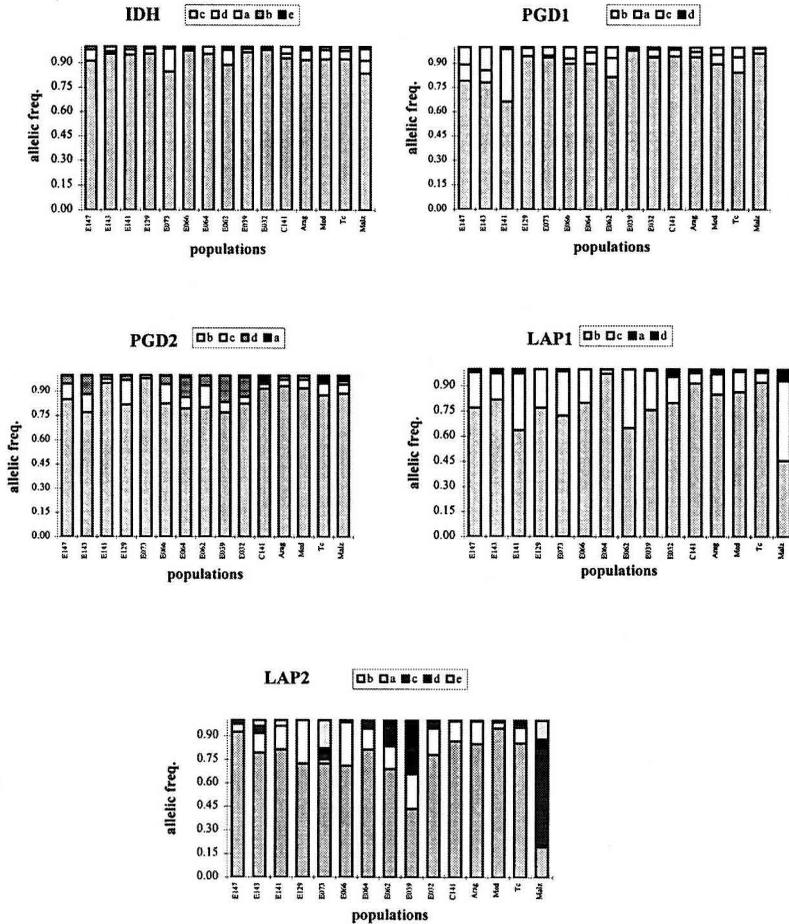


Figure 3. Allele frequencies distribution for five allozymic loci, 14 wild and cultivated populations of *Medicago sativa* and one population of *M. falcata* (Malzeville [Malz]). Wild populations are designated by their site number and the cultivated ones by their abbreviated name (Arag for Aragon, Med for Mediterraneo, TC for Tierra de Campos and C141 for Cult141).

Wild/cultivated differentiation did not stand out preeminently since wild populations were as differentiated from one another ($F_{st} = 0.05$; $P < 0.0001$) as they were from cultivated ones ($F_{st} = 0.05$; $P < 0.0001$). However, populations that displayed intermediate agronomic (and morphological) features between the two cultivated and wild ideotypes seemed to be slightly less differentiated from cultivated populations (F_{st} between groups = 0.0155) than typical wild populations (F_{st} between groups = 0.0295). Interestingly, parapatric populations Es141 and Cult141 were significantly differentiated from each other ($F_{st} = 0.116$; $P < 0.0001$) whereas population Es032 was barely differentiated from Aragon ($F_{st} = 0.008$; $P = 0.005$), Cult141 ($F_{st} = 0.018$; $P = 0.001$) or Tierra de Campos ($F_{st} = 0.021$; $P = 0.0001$).

3.3. RAPD survey

An initial survey had previously been carried out using 40 individuals of a wild population from the south of Spain and 40 individuals of a Spanish landrace (Totana) which is generally cultivated in the same area; 32 RAPD primers were used individually to screen for polymorphism. Of these, 16 did not allow any efficient amplification, nor did they lead to unambiguous and repeatable products. Furthermore, three efficient primers only resolved monomorphic fragments. The remaining 13 primers uncovered 61 polymorphic markers (e.g. 4.8 polymorphic markers per primer) out of the 65 unambiguous and fully reproducible fragments we scored (e.g. 5.0 fragments per primer). None of them could be regarded as diagnostic and only four markers strongly discriminate these two populations (unpublished results).

According to these preliminary results, five primers were selected and used to analyse all the 540 genotypes of this study. They resolved a total of 30 clear major products; each of them was assumed to identify a strictly biallelic locus with complete dominance. Three loci (6 %) appeared to be monomorphic over all populations but the number of polymorphic loci per population was highly variable (*table II*). We did not find any diagnostic band between wild and cultivated populations and only some rare fragments discriminated *M. falcata* from *M. sativa* (their frequencies were significantly different). This is emphasised by PCA. PC1 accounts for 19 % of the overall diversity and essentially separates population Malzeville (*M. falcata*) from all populations of *M. sativa* (data not shown). Plots of PC2 (18 %), PC3 (15 %) and PC4 (12 %) are presented in *figure 4*. The overall variability detected with RAPD markers is mostly due to typical wild populations that surround the cultivated ones; PC3 discriminates populations Es073 and Es129 from all the other wild populations and these two populations are separated from each other by PC2. The most striking result, however, is that populations from group 2 and Es147 (group 3), which display intermediate agronomic and morphological characters, tend to be closely grouped together with the cultivated ones (*figure 4*). This is in complete agreement with their genetic proximity with cultivated populations as revealed with allozymic markers.

Estimates of within-population genetic diversity ranged continuously from 0.130 (Malzeville) to 0.224 (Es147) (*table II*). Partitioning this diversity within and among populations indicated that on average 20 % of the overall diversity was partitioned among populations (G_{st}). However, this value is reduced to 12.4 % when partitioning is performed on *M. sativa* populations only, and to 5.9 % when cultivated populations are excluded from the analysis. In a previous study covering a broader range of population diversity, Crochemore et al. (1996) showed that up to 50.6 % of the total variance was among populations.

4. DISCUSSION

The survey of allozymes, RAPD and agronomic traits depicted a contrasting picture of the way genetic diversity is partitioned within and among wild and cultivated populations of *M. sativa* in Spain. Spatial patterns of genetic diversity usually result from the antagonistic effects of gene flow, heterogeneous natural selection and genetic drift (Slatkin, 1987; Harrison, 1990). We will thus discuss our results in relation to the balance between these evolutionary forces.

Table II. Within-population diversity on RAPD markers as estimated by the proportion of polymorphic fragments and Nei's index of gene diversity (H_s).

Population	Proportion of polym. fragments	H_s
Es032	0.82	0.181
Es039	0.77	0.173
Es062	0.88	0.213
Es064	0.77	0.168
Es066	0.74	0.174
Es073	0.63	0.144
Es129	0.77	0.204
Es141	0.74	0.130
Es143	0.77	0.180
Es147	0.88	0.224
Aragon	0.82	0.203
Mediterraneo	0.82	0.168
Tierra de C.	0.74	0.150
Cult141	0.77	0.189
Malzeville	0.85	0.129
H_T		0.219

H_T : Total observed diversity.

4.1. Agronomic distinction between wild and cultivated forms

This study suggests that most wild populations preserve their agronomic and morphological integrity despite occasional parapatry with cultivated forms. This originality has long been recognised since vernacular names contrasted wild 'Mielgas' with cultivated 'Alfalfas'. The differences observed in situ were maintained ex situ, suggesting that they are genetically based. Most agronomic and morphological traits were strongly correlated with each other (phenotypic correlation computed over population means; data not shown) and allowed a clear distinction between wild and cultivated populations of *M. sativa* (figure 2). However, atypical populations have also been identified because they gathered individuals with different volume, growth habit or winter growth features (i.e. group 2; table I), or because they contained such a high proportion of cultivated-like plants that their agronomic potentials are barely inferior to those of the cultivated ones (i.e. group 3; table I). This result shows that extreme morphotypes live together within some populations but does not provide any accurate information on the within-population variability for agronomic and morphological traits. Consequently, individual phenotypic evaluations have been started for each single plant that has been electrophoretically examined. Preliminary results confirm that some populations display a mosaic of individuals that can be roughly assigned to either the wild or to the cultivated gene pool; however, a continuous range of intermediate forms, from grazing rhizomatous plants to erect ones, has also been observed in some other populations (data not shown).

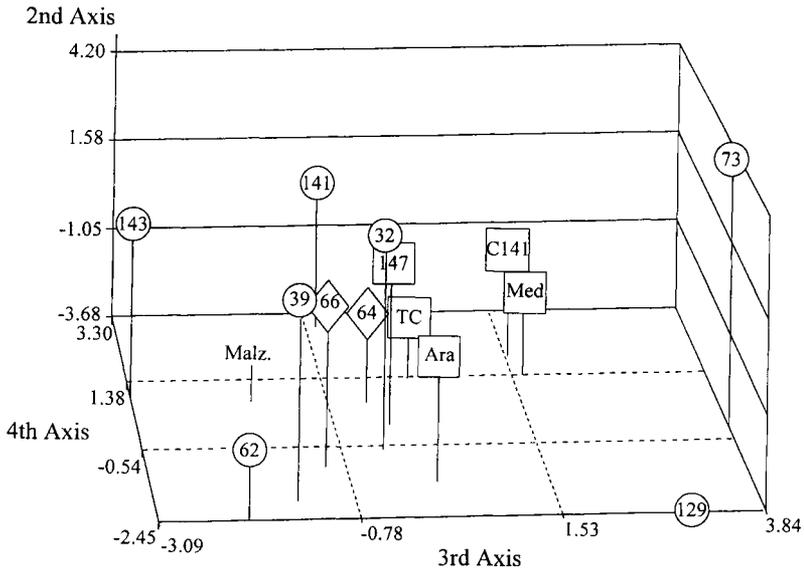


Figure 4. Plots of the second, third and fourth axis of a principal component analysis (PCA) based on random amplified polymorphic DNA (RAPD) frequencies.

Although intermediacy is far from being conclusive evidence of hybridisation (Rieseberg and Ellstrand, 1993), all these results question the nature and extent of gene flow between wild and cultivated populations of *M. sativa* in Spain. Gene flow is often regarded as an homogenising force (Slatkin, 1987) and should thus result in morphological convergence, as documented in red rice wild populations (Langevin et al., 1990). Nonetheless, agronomic and morphological traits can be misleading elements because they are potentially subject to selection. Strong selective differential can generate and maintain steep morphological clines at a few selected loci, whereas neutral alleles keep on flowing (e.g. Antonovics and Bradshaw, 1970; Barton and Hewitt, 1985 and references therein). The survey of population differentiation with neutral genetic markers is thus required to disentangle the relative effects of gene flow, genetic drift and selection (e.g. Bonnin et al., 1996).

4.2. Tracking crop/weed gene flow

The demonstration that crop/weed interbreeding occurs usually comes from the identification and the tracking of diagnostic alleles in the populations they do not initially stem from (Doebley, 1990; Klinger et al., 1992; Wilson and Manhart, 1993; Arias and Rieseberg, 1994). Yet, this approach can be limited by close genetic proximity between most crops and wild relatives (Doebley, 1990); distinguishing introgressive molecular markers from jointly inherited ones is not straightforward (Heiser, 1973). In this study, no diagnostic allele (either with allozymes or with RAPD) was identified between wild and cultivated populations (*figure 3*), but both current gene flow and recent evolutionary divergence could account for this result.

Similarly, Slatkin (1987) emphasised that population differentiation only mirrors the level of gene flow which must have been occurring in order to produce the observed patterns of allele frequencies and does not necessarily reveal how much gene flow is occurring now. In other words, the low levels of differentiation estimated in this study could just as well result from the persistence of a common origin as for the occurrence of current gene flow.

4.3. The persistence of an ancestral genetic background

Interpreting the low level of differentiation in terms of residual genetic background firstly requires that genetic drift should be very slow. This is likely since polysomic inheritance has a conservative effect on genetic diversity. Polyploidy (and allogamy) effectively induces larger effective population size (Moody et al., 1993) than in diploid (selfing) organisms. Moreover, forage crops are known to be poorly domesticated. Domestication can be viewed as a kind of founder effect promoting genetic drift due to the derivation of small samples of individuals from their wild progenitors (van Raamsdonk, 1993; Frankel et al., 1995). As a consequence, the less the domestication, the closer a crop is expected to be to its wild relatives. Genetic drift does exist, however, and is notably exemplified in population Es073; this population is geographically isolated and strongly decreasing in size as observed over 10 years (less than 15 individuals in 1995). This population is more strongly differentiated from all the other populations than they are differentiated from one another (data not shown). Es073 also presents one of the smallest amounts of genetic diversity (33 % of the markers are monomorphic, $H_s = 0.144$; table II).

Allozymes can be reasonably suspected to account partly for the persistence of an ancestral genetic background. Using these markers, cultivated populations do not appear significantly differentiated from one another although Mediterraneo is supposed to stem from a different origin (Delgado, 1996). However, allozymic and RAPD estimates of population differentiation run mostly in parallel in this study and systematically agree about the way diversity is partitioned. It is noteworthy, however, that comparing these estimates of population structure (F_{st} versus G_{st}) is not straightforward and caution is advised until the significance of the levels of differentiation documented with RAPD is tested. To address this issue, an extension of Excoffier's analysis of molecular variance (AMOVA) procedure (Excoffier et al., 1992; Stewart and Excoffier, 1996) for analysing autotetraploid population structure with dominant markers is in progress.

4.4. Evidence for the occurrence of current gene flow

Even a low genetic drift cannot satisfactorily explain why population Es032 appeared to be barely differentiated from cultivated populations, whereas it only contains typical wild individuals. This result strongly suggests that current gene flow can occur between wild and cultivated populations and that clear agronomic and morphological differences may sometimes be misleading concerning the extent of gene exchanges. This result strongly contrasts with the significant differentiation documented between populations Es141 and Cult141, which were about 200 m apart. This distance does not ensure reproductive isolation in *M. sativa* since long-distance gene dispersal (up to 1 000 m) has been

confirmed from alfalfa fields or among escaped plants in a pollen trap plot experiment (Saint Amand et al., 1996). This outward inconsistency can, however, be resolved by considering the manner in which wild and cultivated populations meet and mate. Actually only wild populations of *M. sativa* can persist over time because cultivated ones participate in a rotation with cereals and thus suffer from 'rapid' turnover (every 3 or 4 years). Consequently, preexistent wild/cultivated differentiation may still be observed between parapatric populations when their coming together is recent, and it will last unless subsequent hybridisations have occurred. Nevertheless, alternative hypotheses can also invoke the occurrence, or even the deployment, of reproductive barriers in some locations. However, although some populations contained individuals with obvious morphological differences, neither significant departure from panmixia nor positive assortative mating (e.g. homogamy) could be found (unpublished results). Conversely, population Es147 satisfies some of the expectations that are experimentally or theoretically born out for hybrid populations and notably the presence of rare alleles (RAPD) undetected in all the other populations (Golding and Strobeck, 1983), as well as its greatest intrapopulation diversity (Rieseberg and Wendel, 1993). Populations Es064 and Es147 not only displayed intermediate characters but also appeared to be more closely related to cultivated populations than typical wild populations (*figure 4*).

These preliminary results suggest that, if reproductive barriers between wild and cultivated populations of *M. sativa* exist, they do not result in a strict reproductive isolation. Further analysis is required and should notably focus on different parapatric populations to track clues for reproductive isolation, if it does exist. An accurate survey of populations from group 2 using both molecular markers and morphological traits should be significant, notably to draw definite conclusions about the 'hybrid' status of these populations.

4.5. The strength of natural selection

Providing reproductive isolation is not preminent, the clear distinction which has been documented for agronomic and morphological traits between wild and cultivated populations is likely to result from heterogeneous natural selection. Actually, the efficiency of introgressions mostly depends on the balance between antagonistic selective pressures acting on wild or cultivated populations in a contrasting way: as crop integrity is artificially maintained by man, the expression of gene flow should arise mainly in the wild forms (Langevin et al., 1990) which may shift and adapt to man-made habitats (Harlan, 1965; de Wet and Harlan, 1975). Conversely, most of the agronomic selected traits (dwarfing, nonshattering infrutescence, etc.) are detrimental or maladaptive in the wild (Ellstrand and Hoffman, 1990) so that natural selection should thus impede the crop-to-weed gene flow.

Though this study does not provide any accurate estimate of the strength of natural selection, it does suggest that selective pressure must be strong enough to account for the maintenance of the overall originality of population Es032, regardless of its close genetic proximity with cultivated populations as estimated with neutral markers. Moreover, the strength of selection is likely to vary from place to place. The overall Mediterranean climate of Spain is modulated by the balance between oceanic versus continental influences or the altitude. Average

annual rainfall varies from 150 to 1 000 mm and the north is more humid than the south. In this connection, it is noteworthy that populations from groups 2 and 3 are not randomly distributed but gathered in the north of Spain. The grazing pressure and the agricultural landscape are also very variable (Prosperi et al., 1989). In this scenario, gene flow should be offset by heterogeneous selection in the sites of parapatry between cultivated and wild populations of *M. sativa*. Future prospects should thus try to understand the exact role selection can play in maintaining germplasm diversity (hybrid performance in contrasted controlled situations, covariation between morphological character and neutral markers within natural populations). In parallel, it is crucial to determine whether selection is due to (exogenous) environmental pressures or to (endogenous) interactions within hybrid genomes (see Barton and Hewitt, 1985; Harrison, 1990).

5. CONCLUSION

Wild genetic resources have often attracted attention because of the wealth of genes for resistance they have acquired over long periods of coevolution with pathogens (Frankel et al., 1995). However, they can also be used as a source for new adaptations to less favourable environments. In this connection, 'Mielgas' populations appeared to be a promising germplasm for the development of fodder and hardy varieties adapted to rearing in marginal Mediterranean lands. It intrinsically carries attractive traits such as crawling growth habit, persistence and drought tolerance. However, it also displays good levels of seed production for a nonselected wild material. This study suggests that gene flow from the cultivated populations of *M. sativa* into the wild ones is likely to occur in many locations. This may be directly used in further introgressive breeding schemes but must also be taken into account before transgenic plants are rashly used in Spain.

ACKNOWLEDGEMENTS

We thank I. Delgado who has kindly provided Spanish cultivar seed sets and for his help during collection. Many thanks are due to O. Soudiere, D. Tauzin and M. Vabre for technical assistance in field evaluation and to S. Roussel, N. Courtes and F. Meziane for performing allozymic analysis. We are most grateful to M. Raymond for his tremendous help in performing Fischer's exact test on autotetraploid populations. This work has been supported by grants from the Bureau des Ressources Génétiques and the French Ministère de l'Agriculture.

REFERENCES

- Antonovics J., Bradshaw A.D., Evolution of closely adjacent plant populations. VIII. Clinal patterns at a mine boundary, *Heredity* 25 (1970) 349–362.
- Arias D.M., Rieseberg L.H., Gene flow between cultivated and wild sunflowers, *Theor. Appl. Genet.* 89 (1994) 655–660.
- Barrett S.C.H., Crop mimicry in weeds, *Econ. Bot.* 37 (1983) 255–282.
- Barton N.H., Hewitt G.M., Analysis of hybrid zones, *Ann. Rev. Ecol. Syst.* 16 (1985) 113–148.

Bonnin I., Prosperi J.M., Olivieri I., Genetic markers and quantitative genetic variation in *Medicago truncatula* (Leguminosae): a comparative analysis of population structure, *Genetics* 143 (1996) 1795–1805.

Crochemore M.L., Huyghe C., Kerlan M.C., Durand F., Julier B., Partitioning and distribution of RAPD variation in a set of populations of the *Medicago sativa* complex, *Agronomie* 16 (1996) 421–432.

de Wet J.M.J., Harlan J.R., Weeds and domesticates: evolution in man-made habitat, *Econ. Bot.* 29 (1975) 99–107.

Delgado Enguita I., The lucerne in Spain. Characterization of the cultivated and spontaneous ecotypes, in: *The Genus Medicago in the Mediterranean Region: Current Situation and Prospects in Research*, Hammamet, 19–22 October 1995, 1996, pp. 65–70.

Doebley J., Molecular evidence for gene flow among *Zea* species, *Bioscience* 40 (1990) 443–448.

Ellstrand N.C., Hoffman C.A., Hybridization as an avenue of escape for engineered genes, *Bioscience* 40 (1990) 438–442.

Excoffier L., Smouse P.E., Quattro J.M., Analysis of Molecular Variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data, *Genetics* 131 (1992) 479–491.

Frankel O.H., Brown A.H.D., Burdon J.J., *The Conservation of Plant Biodiversity*, Cambridge University Press, Cambridge, 1995.

Golding G.B., Strobeck C., Increased number of alleles found in hybrid populations due to intragenic recombination, *Evolution* 37 (1983) 17–29.

Harlan J.R., The possible role of weed races in the evolution of cultivated plants, *Euphytica* 14 (1965) 173–176.

Harlan J.R., de Wet J.M.J., Toward a rational classification of cultivated plants, *Taxon* 20 (1971) 509–517.

Harrison R.G., Hybrid zones: windows on evolutionary process, in: Futuyma D., Antonovics J. (Eds.), *Oxford Surveys in Evolutionary Biology*, vol. 7, Oxford University Press, New York, 1990, pp. 69–128.

Harrison S., Hastings A., Genetic and evolutionary consequences of metapopulation structure, *Trends Ecol. Evol.* 11 (1996) 180–183.

Heiser C., Introgression re-examined, *Bot. Rev.* 39 (1973) 347–366.

Klinger T., Arriola P.E., Ellstrand N.C., Crop–weed hybridization in radish (*Raphanus sativus*): effects of distance and population size, *Am. J. Bot.* 79 (1992) 1431–1435.

Langevin S.A., Clay K., Grace J.B., The incidence and effects of hybridization between cultivated rice and its related weed red rice (*Oryza sativa* L.), *Evolution* 44 (1990) 1000–1008.

Levin D.A., Local differentiation and the breeding structure of plant populations, in: Gottlieb L.D., Jain S.K. (Eds.), *Plant Evolutionary Biology*, Chapman and Hall, New York, 1988, pp. 305–329.

Manasse R.S., Ecological risks of transgenic plants: effects of spatial dispersion on gene flow, *Ecol. Appl.* 2 (1992) 431–438.

McCoy T.J., Echt C.S., Mancino L.C., Segregation of molecular markers supports an allotetraploid structure for *Medicago sativa* × *Medicago papillosa* interspecific hybrid, *Genome* 34 (1991) 574–578.

Moody M.E., Mueller L.D., Soltis D.E., Genetic variation and random drift in autotetraploid populations, *Genetics* 134 (1993) 649–657.

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

Prosperi J.M., Delgado Enguita I., Angevain M., Prospection du genre *Medicago* en Espagne et au Portugal, *Plant Genetic Resources Newsletter* 78/79 (1989) 27–29.

Prosperi J.M., Angevain M., Bonnin I., Chaulet E., Génier G., Jenczewski E., Olivieri I., Ronfort J., Genetic diversity, preservation and use of genetic resources of Mediterranean legumes: alfalfa and medics, in: *The Genus Medicago in the Mediterranean Region: Current Situation and Prospects in Research*, Hammamet, 19–22 October 1995, 1996, pp. 71–89.

Raymond M., Rousset F., An exact test for population differentiation, *Evolution* 49 (1995) 1280–1283.

Rieseberg L.H., Ellstrand N.C., What can molecular and morphological markers tell us about plant hybridization? *Crit. Rev. Plant Sci.* 12 (1993) 213–241.

Rieseberg L.H., Wendel J.F., Introgression and its consequences in plants, in: *Hybrid Zones and the Evolutionary Process*, Oxford University Press, Oxford, 1993, pp. 70–109.

Rieseberg L.H., Arias D.M., Ungerer M.C., Linder C.R., Sinervo B., The effects of mating design on introgression between chromosomally divergent sunflower species, *Theor. Appl. Genet.* 93 (1996) 633–644.

Robert T., Lespinasse R., Pernès J., Gametophytic competition as influencing gene flow between wild and cultivated forms of pearl millet (*Pennisetum typhoides*), *Genome* 34 (1991) 195–200.

Saint Amand P.C., Skinner D.Z., Peadar R.N., Risk of alfalfa transgene dissemination and scale dependent effects, in: *Report of the Thirty-Fifth North American Alfalfa Improvement Conference*, Oklahoma City, 16–20 June 1996, 1996.

SAS[®] Institute Inc., SAS\STAT User's Guide, Version 6, 4th ed., SAS Institute, Cary, NC, 1989.

Schoen D., Brown A.H.D., Intraspecific variation in population gene diversity and effective population size correlates with the mating system in plants, *Proc. Natl. Acad. Sci. USA* 88 (1991) 4494–4497.

Slatkin M., Gene flow and the geographic structure of populations, *Science* 236 (1987) 787–792.

Sokal R.R., Rohlf F.J., *Biometry*, 3rd ed., W.H. Freeman and Co., New York, 1995.

Stewart C.N. Jr., Excoffier L., Assessing population genetic structure and variability with RAPD data: application to *Vaccinium macrocarpon* (American Cranberry), *J. Evol. Biol.* 9 (1996) 153–171.

Tai T.H., Tanksley S.D., A rapid inexpensive method for isolation of total DNA from dehydrated plant tissue, *Plant Mol. Biol. Rep.* 8 (1990) 297–303.

Till-Bottraud I., Reboud X., Brabant P., Lefranc M., Rherissi B., Vedel F., Darmency H., Outcrossing and hybridization in wild and cultivated foxtail millets: consequences for the release of transgenic plants, *Theor. Appl. Genet.* 83 (1992) 940–946.

van Raamsdonk L.W.D., Wild and cultivated plants: the parallelism between evolution and domestication, *Evol. Trends Plants* 7 (1993) 73–84.

Weir B.S., *Genetic Data Analysis II*, Sinauer Associates, Inc., Sunderland, MA, 1996.

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

Williams J.G.K., Kubelik A.R., Livak K.J., DNA polymorphisms amplified by arbitrary primers are useful as genetic markers, *Nucleic Acids Res.* 18 (1991) 6531–6535.

Wilson H., Manhart J., Crop/weed gene flow: *Chenopodium quinoa* Willd. and *C. berlandieri* Moq., *Theor. Appl. Genet.* 86 (1993) 642–648.

Wilson H.D., Gene flow in squash species, *Bioscience* 40 (1990) 449–455.

Wright S., The genetical structure of populations, *Ann. Eugen.* 15 (1951) 323–354.

Yu K., Pauls K.P., Segregation of random amplified polymorphic DNA markers and strategies for molecular mapping in tetraploid alfalfa, *Genome* 36 (1993) 844–851.