

Mitochondrial DNA diversity and phenotypic variation in wild and cultivated populations of *Medicago sativa*: insights into the dynamics of contact zones between the two related forms

Marie-Hélène MULLER, Christian BALSERA, Gérard GÉNIER,
Jean-Marie PROSPERI, Sylvie ROUSSEL, Sylvain SANTONI,
Olivier SOUDIÈRE, Denis TAUZIN, Marin VABRE,
Joëlle RONFORT*

Laboratoire de ressources génétiques et d'amélioration des luzernes méditerranéennes, Station de génétique et amélioration des plantes, Institut national de la recherche agronomique, Domaine de Melgueil, 34130 Mauguio, France

Abstract – In order to understand the dynamics of contact zones between natural and cultivated populations of alfalfa (*Medicago sativa* L.) in Spain, we investigated the patterns of mitochondrial variation (characterized through restriction fragment length polymorphisms) for 173 individuals in four natural and four cultivated populations of this species. We also examined the relationship between mtDNA and phenotypic variation. This study shows that cytoplasmic variation occurs in both natural and cultivated populations of *Medicago sativa*. Among the five mitotypes identified in the species, two are specific from the natural populations. The distribution of shared mitotypes between wild and cultivated forms attested to the occurrence of gene flow through seed from cultivated towards natural populations. Comparisons between cytoplasmic and phenotypic patterns of variation give information on the patterns of gene exchange that occur in contact zones over successive generations.

mtDNA variation / phenotypic variation / seed dispersal / contact zones / *Medicago sativa*

Résumé – **Diversité mitochondriale et variation phénotypique dans des populations sauvages et cultivées de *Medicago sativa* : inférences sur la dynamique des zones de contact entre les deux formes apparentées.** Pour comprendre la dynamique de zones de contact entre des populations naturelles et cultivées de luzerne (*Medicago sativa* L.) en Espagne, nous avons étudié l'organisation de la diversité mitochondriale au sein de quatre populations naturelles et quatre populations cultivées, et examiné les relations entre variabilité mitochondriale et phénotypique. La diversité mitochondriale a été caractérisée par hybridation de séquences

* Correspondence and reprints
E-mail: ronfort@ensam.inra.fr

mitochondriales sur l'ADN total. Cette étude montre qu'il existe un polymorphisme cytoplasmique au sein des populations naturelles et cultivées de *Medicago sativa*. Cinq mitotypes ont été identifiés dont deux sont spécifiques des populations naturelles. Les trois mitotypes du compartiment cultivé, présents en fréquences variables dans le compartiment sauvage, attestent de l'existence de flux de graines des populations cultivées vers les populations sauvages. La comparaison de la variation mitochondriale et phénotypique nous permet de faire des inférences sur les échanges géniques en jeu dans les zones de contact au cours des générations successives.

diversité mitochondriale / variation phénotypique / flux de graines / zone de contact / *Medicago sativa*

1. INTRODUCTION

Secondary contact zones between sister species or subspecies are usually viewed as sites of active genetic and ecological interactions between differentiated populations. The study of such contact zones should provide insights into a range of problems in evolutionary biology: production of new recombinant types [28, 30], speciation [16], reinforcement [18]. Particular instances of such contact zones are those that involve wild and cultivated forms of a single species. Depending on the degree of domestication, the two forms are more or less differentiated and a strict reproductive isolation is rarely established [35]. Moreover, man artificially maintains the cultivated form, so that effective gene flow, if it occurs, is primarily unidirectional, from the cultivated form towards the wild relative [11, 29], although efficient gene flow from the wild to the cultivated populations may occur. Usually, these wild populations may constitute important genetic resources. Therefore, understanding the causes of such interactions and their potential consequences on the fate of the wild form should provide not only insights into evolutionary processes but also guidance in sampling of genetic diversity for future breeding purposes as well as *in situ* management strategies.

According to theoretical models, the fate of a contact zone generally depends on the balance between three forces: genetic drift, gene flow and selection, the strength of selection counteracting the otherwise inevitable erosion of group differences expected under recurrent gene flow. Patterns of genetic variation (as described by clinal variation or F-Statistics) and genetic association (among nuclear molecular markers, phenotypic traits, and environmental variation) are generally used to assess the relative importance of selection as compared to gene flow [3, 6]. When applied to contact zones between wild and cultivated relatives, shared ancestral polymorphisms present prior to domestication are likely to occur, even in the total absence of gene flow between the two entities. This reduces the availability of diagnostic molecular markers (*i.e.* markers that are specific and frequent in one population) and consequently hinders the use of classical model of population genetics to analyse the dynamics of contact zones [29]. Cytoplasmic markers could be helpful for that purpose. In

the case of mitochondrial DNA, despite a low rate of gene sequence evolution [37], a strong variability in size and gene arrangement exists [23], which makes it a powerful tool in detecting intraspecific diversity [38]. Moreover, due to their uniparental segregation pattern, cytoplasmic markers generally display increased population differentiation compared to nuclear markers [24]. Diagnostic markers are thus more likely to be found, especially for the cultivated form due to the bottleneck effect of domestication (see for example [9]). With this assumption, and for maternally transmitted organelles (as is generally the case in plants), cytoplasmic patterns of variation should provide information on the level of seed dispersal. Whereas if morphological traits distinguish natural and cultivated populations, associations between diagnostic cytoplasmic and phenotypic markers may help to understand the dynamics of the contact zone, especially the patterns of gene flow and mating within the introgressed populations.

In Spain, wild and cultivated populations of alfalfa (*Medicago sativa* L.) grow in parapatry in many locations: cultivated fields can be adjacent to natural populations. Both forms are autotetraploid, perennials and outcrossing. They overlap in flowering periods, share the same pollinators and are cross compatible as verified experimentally [25]. A crawling growth habit and rhizomes, differing clearly from the cultivated erect phenotype, characterize typical wild plants. Our general question is to highlight the processes (patterns of gene exchange, mode and intensity of selection) which determine the genetic structure of alfalfa natural populations and the maintenance of their morphological originality despite their frequent occurrence in parapatry with the cultivated form. A previous study of the overall differentiation among natural and cultivated populations of *Medicago sativa* collected throughout Spain has shown that wild and cultivated populations are weakly differentiated for nuclear molecular and biochemical markers [19–21]. Although these results suggest the occurrence of gene flow between the two entities, the lack of nuclear diagnostic markers between wild and cultivated populations makes it difficult to clearly understand the dynamics of these contact zones. In this paper, we report on a study of mitochondrial variability in natural and cultivated alfalfa populations collected in North-West Spain and examine the relationship between mitochondrial types and agro-morphological traits. In this species, mitochondria are only maternally transmitted [13]. By comparing patterns of mtDNA diversity among natural and cultivated populations to previous results on nuclear patterns of molecular variation, we discuss the relative importance of gene flow due to seed *versus* pollen dispersal. Results on the relationship between patterns of mtDNA diversity and phenotypic variation in natural populations are used to understand the patterns of gene exchange that occur within contact zones over successive generations.

2. MATERIALS AND METHODS

2.1. Studied populations

Four natural populations of *Medicago sativa* were sampled in the North-West of Spain in September 1997 (see map in Fig. 1). Two of these (Es066 and Es143) displayed a “pure” wild type, exhibiting crawling growth habit and rhizomes, although some typical cultivated forms could be found in the population Es143 (see section 3). The two other populations (Es127 and Es068) showed an intermediate phenotype, *i.e.* they contained a continuous range of plants with intermediate features between the wild and the cultivated forms, as expected with hybridization between the two entities. For all these populations, cultivated plants could be found in the vicinity of the sampled sites. In each population, mature pods were collected on 30 to 60 maternal plants, and each sampled maternal plant position was mapped. Pods were threshed in the laboratory and one seed per sampled maternal plant was put in Petri dishes for two days to germinate and afterwards transplanted into pots in the greenhouse. These plants were used for DNA extraction. Standard representatives from three Spanish cultivated landraces (*Tierra de Campos*, *Mediterraneo*, *Aragon*), and from a French “Flemish” cultivar (*Europe*) which is sometimes cultivated in Spain, were included in the molecular survey as cultivated controls (10 plants per population were analyzed). A wild population of *Medicago falcata* (a related species of the same complex) collected in the north of France (*Malzeville*) was used to calibrate the amount of differentiation.

2.2. Molecular survey

Total DNA was extracted according to [34] from frozen material and digested with individual restriction endonucleases (*EcoRI* and *HindIII*). Five micrograms of digested DNA was separated on 0.8% agarose gel and transferred onto charged Nylon membrane (Appligene-Oncor) using alkaline transfer. Probes were radiolabelled by random priming (20 μCi of ^{32}P αdCTP (3000 Ci/mmol), (1 Ci = 37 Gbq) per probe). Membranes were hybridized overnight at 65 °C in a 7% SDS, 0.5 M NaHPO₄ (pH 7.2) buffer [4] and then washed three times in 40 mM NaHPO₄ (pH 7.2), 0.1% SDS (20 min at 65 °C).

Eight mitochondrial probes kindly provided by P. Saumitou-Laprade (CNRS-Lille) and F. Vedel (Université Paris Sud) were used: the ATPase subunit six gene (*atp6*) of maize ATP synthetase [8], the ATPase subunit a gene (*atpa*) of sunflower ATP synthetase [33], four different fragments of the NADH dehydrogenase gene, subunit 1 (*nad1* and *nad1-d*) from Arabidopsis [36], subunit three from tobacco (*nad3*, [22]) and subunit six from wheat (*nad6*, [15]), and two fragments containing cytochrome oxidase subunits from *Oenothera* (*coxI* and *coxIII*, [17]). The restriction endonuclease/probe combinations (E/P) studied



Figure 1. Location of the 4 natural populations and sites where the three Spanish landraces originated (shaded areas).

were: *HindIII*/*nad1d*, *HindIII*/*atp6*, *HindIII*/*Atpa*, *HindIII*/*nad6*, *HindIII*/*coxI*, *HindIII*/*nad3*, and *EcoRI*/*coxIII*, *EcoRI*/*nad1*.

2.3. Morphological survey

Based on results of the molecular survey described above, we implemented an agro-morphological study to examine whether different mitotypes were associated with different phenotypic classes (namely cultivated *versus* wild phenotypes) within contact zones. For each population (Es066, Es068, Es127 and Es143), each mitotype (four were considered) was represented by at least 20 plants sampled in different maternal families (the number of plants per maternal family varies from 3 to 12, depending on the population cytoplasmic diversity as shown in Tab. I). For the 4 cultivated populations, 10 randomly chosen plants per population were included in the experimental design as cultivated control. Original seeds were put in Petri dishes for 2 days to germinate. The seedlings were then transplanted into plastic containers (4 plants / container of 0.5 m³) in a randomized 2-block design in September 1997. Plants were grown for two years under greenhouse conditions with continuous individual watering. Measurements were scored on each individual and were divided up into two broad categories: vegetative and reproductive characters. Vegetative characters included the production of rhizomes, the growth habit, the plant

Table I. Number of maternal families (and number of individual plants) from each population representing the different mitotypes in the agro-morphological survey.

Population	Mitotype				Total
	A	C	D	E	
Es066	6 (24)	3 (24)	2 (24)	–	11 (72)
Es068	6 (24)	3 (24)	6 (24)	3 (24)	18 (96)
Es127	17 (46)	11 (34)	20 (52)	11 (40)	59 (172)
Es143	7 (28)	–	5 (20)	3 (24)	15 (72)

volume and 3 characters of vegetative productivity: autumn regrowth, total vegetative productivity, and the ability to grow after the first cut in spring. The number of rhizomes was scored on 26/02/1999 from 1 = no rhizome to 5 = numerous rhizomes. The growth habit was measured as the ratio of the plant diameter to the plant height (both measured on 24/02/00). The product of these two values was used as a measure of plant's volume. Autumn regrowth and winter growth were assessed on the basis of the dry matter produced on 26/10/98 and on 01/03/99, respectively. The ability to grow after the first cut (performed on 01/03/99) was scored 15 days after cutting from 1 = weak re-growth to 9 = intense re-growth. Vegetative productivity was estimated as the total dry matter produced between the 26/10/98 and the 10/02/00 over all successive cuts.

Flowering traits included the date of flowering, inflorescence length, the number of flowers produced per inflorescence, the number of pods produced per pollinated flower, the number of seeds produced per pod and the mean seed weight. Flowering date was expressed as the number of days necessary for a plant to flower since the beginning of the year (01/01/99). At the date of first flowering in 1999, two stems were randomly chosen on each plant. On these two stems, the second inflorescences were used to measure the number of flowers produced per inflorescence and the inflorescence length. Two inflorescences were then hand-pollinated on each stem in order to assess traits related to seed production. Pollinations were made with a mixture of pollen from plants randomly chosen in the experimental design. After pollinations were complete, stems were covered and allowed to mature. Upon maturation, pods were harvested, scored and threshed in order to measure the number of seeds produced per pod and the mean seed's weight.

2.4. Data analysis

Due to the non-recombining nature of the mitochondrial genome on the spatial and temporal scale studied, the patterns detected with the different E/P (enzyme*probe) combinations were combined to define mitochondrial haplotypes. These will be referred to as mitotypes in what follows. To estimate population differentiation, pairwise F_{ST} estimates were computed by a "weighted"

analysis of variance on mitotype frequencies, using the Genepop software package [27] (option 3 and option 8 to adapt the haploid data set). Significance of population differentiation was assessed using exact tests [26] also available in the Genepop software.

Quantitative trait analyses were performed using the SAS/STAT software package [31]. Analyses of variance were computed on family mean values. Differences between population and cytoplasmic type were examined for each metric character using an analysis of variance for unequal sample sizes using type III SS (Proc GLM). The model included population, mitotypes, block, and population X mitotypes, as fixed factors. Duncan multiple range tests were performed to detect differences among populations and mitotypes for each trait. Phenotypic differences between mitotypes were also assessed within each population separately.

3. RESULTS

3.1. MtDNA polymorphism

Among the eight E/P combinations studied, only one was monomorphic (*HindIII/Cox1*); the number of patterns observed with the seven remaining E/P combinations ranged from 2 to 6 (6 variants for *HindIII/nad1d* and *HindIII/atpa*, 4 for *HindIII/atp6* and *EcoRI/coxIII*, 3 for *HindIII/nad6*, 2 for *HindIII/nad3* and *EcoRI/nad1*). Combining these patterns, 6 different mitotypes could be defined among the 173 plants surveyed. Two E/P combinations were sufficient to distinguish all the mitotypes, E/P combinations *HindIII/nad1d* and *HindIII/atpa* being redundant. Among these mitotypes, two (denoted A and C) were detected only in natural populations and one was specific to *Medicago falcata* (H). The three remaining were found in both wild and cultivated populations (B, D, E). The number of different mitotypes observed in a single population ranged from 2 to 5 which was large compared to the overall variability detected. As shown in Figure 2, the four cultivated populations shared the same mitochondrial types. Furthermore, they were not significantly differentiated from one another (Exact Test $P > 0.10$) in mitotype frequencies. The three mitotypes detected in these cultivars were all present in the natural populations. Interestingly, they occur more frequently in the two "intermediate" populations (Es068 and Es127) as compared to the typical wild populations (Es143 and Es066, see Fig. 3). Pairwise F_{ST} between either Es068 or Es127 and each of the cultivated populations were not significantly different from zero, except for the pair Es127-Aragon which was weakly differentiated (Tab. II). In contrast, Es143 and Es066 were significantly differentiated from both the cultivated populations and the intermediate natural populations, population Es066 showing the greater degree of differentiation

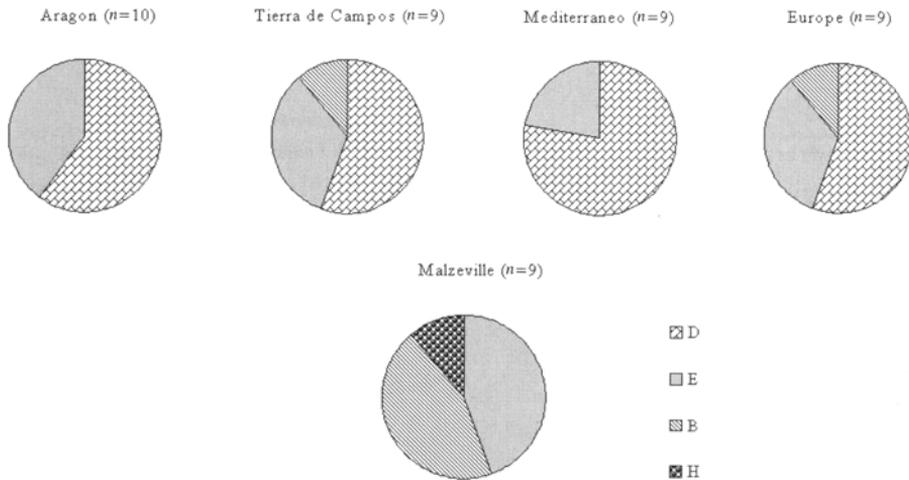


Figure 2. Patterns of mitotype frequencies within cultivated populations of *Medicago sativa*, and one population of *Medicago falcata* as revealed through Southern blot hybridization of total DNA with mtDNA probes (n = number of individuals surveyed).

compared to the cultivated populations (Tab. II). Pairwise F_{ST} between natural populations cannot be related to geographical position (as is evident when comparing Tab. II and the map in Fig. 1). Finally, *Medicago falcata* shared two mitotypes with the cultivated populations (B and E, Fig. 2) but was highly differentiated from all the studied populations (Tab. II).

3.2. Morphological variation among populations

For both vegetative and reproductive characters, univariate analysis of variance detected significant differences between natural and cultivated populations, except for one trait (total vegetative productivity). Compared to natural populations, the cultivated morphotype was characterized by a greater ability to grow in Spring, a larger volume, an erect growth habit, and the absence of rhizome. Cultivated plants also flowered earlier, and displayed larger inflorescences with more flowers and a better seed production (more seeds that were also heavier). Natural populations differed significantly from one another for autumn regrowth, plant diameter, and seed weight. When classifying the natural populations according to their phenotypic proximity as compared to the cultivated populations, population Es066 exhibited the most differentiated phenotype (autumn re-growth, plant diameter, see also growth habit, plant height and total vegetative productivity although not significant). In contrast, Es127 exhibited intermediate phenotypic values between Es066 and the mean cultivated phenotype for most characters (Tab. III).

Table II. Pairwise F_{ST} estimates among natural and cultivated populations of *Medicago sativa*.

	Es066	Es127	Es068	Es143	Aragon	Europe	Tierra de Campos	Mediterraneo
Es127	0.15**							
Es068	0.23**	0.01ns						
Es143	0.06*	0.04ns	0.05ns					
Aragon	0.45***	0.09*	0.04ns	0.18*				
Europe	0.41***	0.05ns	0.01ns	0.15*	-0.09ns			
TdeCampos	0.41***	0.05ns	0.01ns	0.15*	-0.09ns	-0.12ns		
Medit	0.50***	0.13ns	0.04ns	0.25*	-0.04ns	-0.04ns	-0.04ns	
Falcata	0.42***	0.16***	0.27***	0.25***	0.27**	0.18*	0.18*	0.41**

Levels of significance : *: $p < 0.05$; **: $p < 0.005$, ***: $p < 0.0005$ (following exact tests).

Table III. Mean phenotypic values in each population. For each trait, values followed by the same letter are not significantly different ($p>0.05$).

(a) *Vegetative traits*

Population	Autumn Regrowth	Spring regrowth	Total vegetative productivity	Plant height (H)	Plant Diameter (Di)	Volume (DixHM)	Growth Habit (Di/H)	Rhizome
Es066	13.8 b	2.2 b	207.3 a	14.6 b	34.3 b	527.3 b	2.7 a	6.5 a
Es068	24.1 ab	3.2 b	309.1 a	19.8 b	37.3 b	795.4 b	2.2 a	6.9 a
Es127	28.0 a	3.5 b	325.9 a	22.3 b	39.0 ab	969.1 b	2.1 a	5.8 a
Es143	25.8 a	3.4 b	304.8 a	27.0 b	38.7 ab	1257.8 b	2.0 a	4.4 a
Es143 without E	21.0 ab	2.5 b	246.8 a	16.7 b	35.6 b	673.0 b	2.5 a	5.7 a
Cultivated form	.	7.1 a	344.9 a	52.1 a	45.7 a	2529.3 a	0.9 b	0.4 b

(b) *Reproductive traits*

Population	Flowering date	Inflorescence Length	Number of flowers per inflorescence	Pods production	Seeds per pod	Seed weight (mg)
Es066	189 a	25.1 ab	14.1 ab	0.83 ab	4.9 b	2.0 bc
Es068	190 a	21.5 b	12.1 b	0.73 b	4.1 b	1.8 c
Es127	186 a	25.6 ab	14.2 ab	0.80 ab	4.5 b	2.1 b
Es143	187 a	23.7 b	14.3 b	0.76 ab	3.8 b	1.9 bc
Es143 without E	189 a	22.0 b	13.6 b	0.78 ab	3.9 b	1.9 bc
Cultivated form	175 b	29.5 a	17.5 a	0.86 a	6.8 a	2.7 a

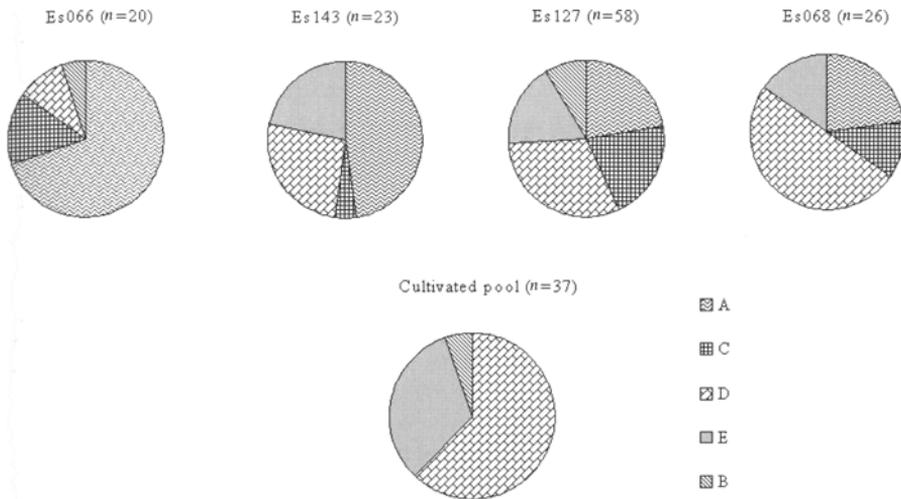


Figure 3. Patterns of mitotype frequencies within natural populations of *Medicago sativa*, and comparison with the patterns observed on cultivated populations when pooled together. Mitotypes were identified through Southern blot hybridization of total DNA on mtDNA sequences (n = number of individuals surveyed).

3.3. Association between mitotypes and morphological traits

Among the four natural populations studied, associations between phenotype and mitotype were found only in one population (Es143). In this population, the mitotype E was always associated with plants that exhibited a cultivated phenotype while plants associated with the two other mitotypes (A and D) displayed the typical wild morphological form (Tab. IV). This association was found for all the studied characters. Moreover, when removing plants bearing the mitotype E, mean phenotypic values of most characters in population Es143 were close to those observed in the “wild type” population Es066 (Tab. III). In contrast, no association was found in the three other populations, even in population Es127 in which each mitotype was represented by at least 11 maternal families.

4. DISCUSSION

4.1. What does cytoplasmic diversity tell us about cultivated populations?

Using RFLP analysis on mitochondrial DNA, we showed that cytoplasmic variation occurs in both natural and cultivated populations of *Medicago sativa*. The occurrence of intravarietal polymorphism, already shown for Japanese and

Table IV. Mean phenotypic values of the different mitotypes in each population. Within each population, values followed by the same letter are not significantly different ($P > 0.05$).

Population	Mitotype	Autumn regrowth	Winter growth	Spring regrowth	Total vegetative productivity	Plant Height (H)	Plant Diameter (Di)	Growth habit (Di × H)	Volume (Di × H)	Rhizome
Es066	A	12.4 a	5.1 a	2.2 a	192.0 a	14.7 a	34.3 a	2.7 a	521.7 a	6.1 a
	C	12.3 a	4.5 a	2.0 a	211.5 a	13.7 a	33.2 a	2.9 a	485.5 a	6.4 a
	D	20.7 a	6.0 a	2.3 a	243.7 a	15.4 a	35.9 a	2.7 a	575.1 a	7.0 a
Es068	A	27.7 a	17.1 a	3.7 a	330.1 a	20.6 a	39.6 a	2.3 a	921.5 a	6.8 a
	C	25.1 a	17.6 a	3.7 a	307.3 a	23.2 a	36.1 a	1.7 a	916.0 a	6.0 a
	D	21.3 a	10.7 a	2.7 a	296.1 a	16.5 a	36.9 a	2.4 a	632.5 a	7.2 a
	E	22.1 a	8.75 a	2.8 a	317.9 a	18.9 a	35.4 a	2.2 a	699.7 a	7.5 a
Es127	A	22.1 a	16.3 a	3.0 a	300.4 a	21.4 a	37.7 a	2.1 a	922.3 a	5.4 a
	C	31.0 a	16.0 a	3.7 a	324.9 a	24.4 a	41.4 a	2.1 a	1093.6 a	5.5 a
	D	28.2 a	16.8 a	3.4 a	330.1 a	20.7 a	38.0 a	2.1 a	882.2 a	5.9 a
	E	33.7 a	17.1 a	4.0 a	360.4 a	23.5 a	40.1 a	2.0 a	1045.9 a	6.3 a
Es143	A	20.7 a	8.3 a	2.4 a	249.5 a	16.5 a	35.4 a	2.4 a	627.3 a	6.3 a
	D	21.4 a	14.8 a	2.7 a	243.1 a	16.9 a	35.9 a	2.6 a	734.6 a	5.0 a
	E	46.6 b	40.6 b	6.8 b	529.7 b	48.1 b	50.2 a	1.2 b	2452.8 b	1.8 b
Cult	D	.	18.7 a	7.2 a	305.9 a	51.0 a	44.8 a	0.9 a	2361.1 a	0 a
	E	.	22.1 a	7 a	383.9 a	53.1 a	46.6 a	0.9 a	2484.9 a	0.2 a

American alfalfa varieties [1], was in accordance with the breeding methods involved for the production of alfalfa cultivars. As for most forage crops, alfalfa varieties are either landraces (*i.e.* *Tierra de Campos*, *Mediterraneo* and *Aragon*) or synthetic varieties (*i.e.* *Europe*). In contrast to inbred lines and hybrid varieties, these types of varieties imply the mixture of different components (components are either single genotypes, or populations); it was thus not surprising to find intravarietal mtDNA polymorphism. More surprisingly, the four cultivated populations studied were not differentiated from one another in their mitotype composition. The three Spanish populations and Europe are generally thought to result from different original genetic pools. *Tierra de Campos*, *Mediterraneo* and *Aragon* have been described as cultivated landraces issued from local selections on old Spanish cultivated populations. In contrast, “*Europe*” belongs to the Flemish pool, *i.e.* it originates from selection in cultivated populations from the North of Europe. The lack of mitotype differentiation between these populations, combined with the lack of differentiation observed on nuclear molecular markers [19–21] suggests either that these populations issued from the same domestication event or that earlier differences between ancient varieties have been erased through modern breeding.

Among the three mitotypes we detected in the cultivated populations, two were also present in *M. falcata*. Two non-exclusive phenomena could explain this result. First, these two mitotypes could be ancestral mitotypes, *i.e.* that occurred in the ancestral form of *M. sativa* and *M. falcata*. One or both of these mitotypes could also be specific mitotype(s) of *M. falcata*; their occurrence in cultivated *M. sativa* would then reflect “man-made” introgression events from *M. falcata* into *M. sativa*. Crosses between *M. falcata* and *M. sativa* are classically used in alfalfa breeding programs to improve the winter hardiness of the cultivated form. Such crosses are generally identified through the occurrence of variegated flowers in the cultivated populations. Examining the relationship between flower colors and mitotype variation could thus help to distinguish the two hypotheses above. An alternative explanation could be that the population *Malzeville* (used here as a reference population for *M. falcata*) has experienced gene flow from alfalfa populations cultivated in the North of France (in particular, Europe has been used extensively in this region). Hybrid swarms between domesticated *M. sativa* and wild *M. falcata* have already been reported [32]. This hypothesis is however unlikely since population *Malzeville* displays all the taxonomic characters of *M. falcata* and has been shown to be clearly differentiated from alfalfa varieties in a previous study (for both morphological and molecular markers, [19, 20]).

As compared to the cultivated pool, two supplementary mitotypes were detected among the four natural populations studied. Interestingly, both were found in each of these populations. This attests to the originality of the Spanish natural populations of *M. sativa* as compared to the cultivated form. This also suggests that neither *Tierra de Campos*, nor *Aragon*, nor *Mediterraneo* directly issued from local selection in these natural populations. An extension of this

mitochondrial survey to include other Spanish varieties and natural populations from the South of Spain is now in progress to confirm the originality of these natural populations. A further extension of this investigation to other cultivated and wild origins should provide information on the bottleneck effect that occurred during domestication.

4.2. Mitotype differentiation between natural and cultivated populations

Despite these two particular mitotypes, the four natural populations also contained the three mitotypes found in cultivated alfalfa. As for nuclear markers, this could be interpreted either as an ancestral polymorphism present prior to domestication or as an indication of the occurrence of gene flow from cultivated towards natural populations (or both). Interestingly, and in accordance with the gene flow hypothesis, the proportion of shared mitotypes between natural and cultivated populations was larger in the two intermediate populations as compared to the two populations that display the “typical” wild phenotype. In a previous investigation based on both nuclear markers and agromorphological variation, we already suggested that intermediate populations were likely to be hybrid swarms between natural and cultivated populations of *M. sativa* since they were not differentiated from cultivated populations for both types of diversity [19]. An allozyme survey of population Es127 (based on the 5 loci described in [19]) has confirmed that this population is also not differentiated from the cultivated pool for nuclear markers (data not shown). This further supports the assumption that gene flow is partly responsible for the occurrence of “cultivated” mitotypes within natural populations.

As compared to patterns of nuclear variation which reflect gene flow through both pollen and seed dispersal, the level of differentiation assessed on cytoplasmic variation reflects the level of gene flow that results from seed dispersal only. The lack of differentiation observed between intermediate and cultivated populations thus suggests that gene flow through seed plays an important role in contact zones between wild and cultivated populations of *Medicago sativa*. This contrasts with the classical view that gene flow between a crop and its wild or weedy relatives is mostly pollen-mediated [10], although recent studies focusing on GMO risk assessment have taken into account the potential occurrence of gene flow through seeds (for example see [7]). In the case of *M. sativa*, the frequent occurrence of irrigation canals around cultivated areas, but also the use of alfalfa to put roadsides under grass, probably increase the potential for cultivated seeds to establish in non-cultivated areas.

4.3. What does the relationship between cytoplasmic and phenotypic variation tell us about the dynamics of these contact zones?

Evidence for an association between a phenotypic trait and a particular mitotype can be the result of two different phenomena: (i) the involvement of cytoplasmic genes in the phenotypic expression of the related trait, and (ii) the occurrence of linkage disequilibrium between cytoplasmic and nuclear genomes. A common example of the former is the expression of sexual phenotypes in gynodioecious species, a breeding system characterized by the co-occurrence of male sterile (functionally female) and hermaphrodite plants within populations. In many gynodioecious species, the sexual phenotype is the result of an interaction between nuclear and cytoplasmic genes [5], and, in many instances, major rearrangements in mtDNA have been shown to be characteristic of male sterile lines [14]. In *Medicago sativa*, analyses of progenies resulting from crosses between wild and cultivated plants suggested the occurrence of maternal effects on different characters including growth habit and winter growth (Prosperi, unpublished results). Thus, evidence for a relationship between these characters and cytoplasmic variation could have suggested that those maternal effects result from a direct involvement of cytoplasmic genes in these characters. The involvement of organelle genomes in such morphological traits has, however, never been reported in plants. On the other hand, in contact zones between wild and cultivated populations, linkage disequilibrium between cytoplasmic and nuclear markers is expected when cultivated seeds have been established in a natural population but further gene exchange between the two entities has not occurred or has not been sufficient to erase that disequilibrium.

In the present study, a relationship between mitotype and phenotypic variation was only found in one population, Es143. For several reasons, this association is more likely to result from a linkage disequilibrium between cytoplasmic and nuclear genes and not from a direct involvement of cytoplasmic genes. First, if cytoplasmic genes were involved, then associations would have been evidenced in the other populations, at least in population Es127, in which similar levels of variation for both mitotypes and phenotypes occur as compared to population Es143. Second, in population Es143, mitotype/phenotype associations were detected for all the traits differentiating wild and cultivated populations, a pattern expected with linkage disequilibrium since the whole nuclear genome is in linkage disequilibrium with the cytoplasm (as long as the gene exchange between the two entities has not been sufficient). Finally, maternal families associated with mitotype E in population Es143 were not randomly distributed along the collection transect; all were sampled in the same area, *i.e.* at the end of the transect, so that a reproductive isolation between these plants and the wild plants could exist. No such distribution was observed for the other natural populations. Evidence for linkage disequilibrium between nuclear molecular or morphological markers and cytoplasmic variation in this population is now necessary to confirm our hypothesis. Unfortunately, as yet

we have failed to find nuclear diagnostic markers and consequently to detect nuclear introgression of cultivated genes into natural populations. The power to detect such a linkage disequilibrium, when the differentiation assessed with nuclear molecular markers is small, would thus be very low [2].

Assuming that linkage disequilibrium is responsible for a relationship between mitotype E and morphological traits in population Es143, the lack of such a relationship in the other populations gives information on their genetic status. In particular, a survey of allozymic variation in population Es127 has shown that genotypic frequencies in this population are not different from those expected with random mating (data not shown). This result associated with (i) the occurrence of 'cultivated' mitotypes in this population and (ii) the lack of mitotype/phenotype relationship suggests that no reproductive isolation between sympatric wild and cultivated plants occurs in this population. In such a population, it is thus reasonable to assume that recurrent seed flow from cultivated populations could result in the disappearance of the wild phenotype, assuming that selection against the cultivated phenotype is not too strong. In contrast, the linkage disequilibrium observed in population Es143 reflects that inter-crosses between wild and cultivated plants had not occurred in this population. It remains to be determined whether this reproductive isolation simply results from a recent event of (cultivated) seed migration or attests to the occurrence of selective processes that counteract further gene exchange between the two entities. Based on the results presented in this paper, examining the temporal evolution of both cytoplasmic and phenotypic variation in such populations could help to elucidate this question.

ACKNOWLEDGEMENTS

We thank J. David and E. Jenczewski for discussion of this work, and T. Bataillon for critical review of the manuscript. Many thanks are due to C. Gratraud for technical assistance in agromorphological evaluation. This work was funded by the *Bureau des ressources génétiques*.

REFERENCES

- [1] Ando S., Saeko M., Takahashi C., Shimizu T., Intravarietal differences in Mitochondrial DNAs of Alfalfa (*Medicago sativa* L.), *Breed. Sci.* 45 (1995) 227–228.
- [2] Arnold J., Cytonuclear disequilibria in hybrid zones, *Ann. Rev. Ecol. Syst.* 24 (1993) 521–554.
- [3] Barton N.H., Hewitt G.M., Analysis of hybrid zones, *Ann. Rev. Ecol. Syst.* 16 (1985) 113–148.
- [4] Church M.G., and Gilbert W., Genomic sequencing, *Proc. Natl. Acad. Sci. USA* 81 (1984) 1991–1995.

- [5] Couvet D., Atlan A., Belhassen E., Gliddon C.J., Gouyon P.H., Kjellberg F., Coevolution between two symbionts: the case of cytoplasmic male sterility in higher plants, *Oxf. Surv. Evol. Biol.* 7 (1991) 225–249.
- [6] Cruzan M.B., Arnold M.L., Ecological and genetic associations in an Iris Hybrid zone, *Evolution* 47 (1993) 1432–1445.
- [7] Desplanques B., Betteraves mauvaises herbes et rudérales : diversité génétique, traits d'histoire de vie et flux de gènes au sein du complexe d'espèces cultivées-sauvages *Beta vulgaris* ssp., thèse de l'Université Lille I (1999) 71 p.
- [8] Dewey R.E., Levings C.S. 3d, Timothy D.H., Nucleotide sequence of ATPase subunit 6 of Maize mitochondria, *Plant Physiol.* 79 (1985) 914–919.
- [9] Doebley J., Mapping the genes that made maize, *Trends Genet.* 8 (1992) 302–307.
- [10] Ellstrand N.C., Hoffman C.A., Hybridization as an avenue of escape for engineered genes, *Bioscience* 40 (1990) 438–442.
- [11] Ellstrand N.C., Prentice H.C., Hancock J.F., Gene flow and introgression from domesticated plants into their wild relatives, *Ann. Rev. Ecol. Syst.* 30 (1999) 539–563.
- [12] Ennos R.A., Estimating the relative rates of pollen and seed migration among plant populations, *Heredity* 72 (1994) 250–259.
- [13] Forsthoefel N.R., Bohnert H.J., Smith, S.E., Discordant inheritance of mitochondrial and plastid DNA in diverse alfalfa genotypes, *J. Hered.* 83 (1992) 342–345.
- [14] Hanson M.R., Plant mitochondrial mutations and male sterility, *Ann. Rev. Genet.* 25 (1991) 464–486.
- [15] Haouazine N., Pereira de Souza A., Jubier M.F., Lancelin D., Delcher E., Lejeune B., The wheat mitochondrial genome contains an ORF showing sequence homology to the gene encoding the subunit 6 of the NADH-ubiquinone oxidoreductase, *Plant. Mol. Biol.* 20 (1992) 395–404
- [16] Hewitt, G.M., Hybrid zones: natural laboratories for evolutionary studies, *Trends Ecol. Evolut.* 3 (1988) 158–167.
- [17] Hiesel R., Schobel W., Schuster W., Brennicke A., The cytochrome oxidase subunit I and subunit III genes in *Oenothera* mitochondria are transcribed from identical promoter sequences, *EMBO J.* 6 (1987) 29–34
- [18] Howard D.J., Reinforcement: Origins, dynamics and fate of an evolutionary hypothesis, in: Harrison, R.G. (ed.), *Hybrid zones and the Evolutionary process*, Oxford University Press, NY, 1993, pp. 46–69.
- [19] Jenczewski E., Angevain M., Charrier A., Génier G., Ronfort J., Prospero J-M., Contrasting patterns of genetic diversity in neutral markers and agromorphological traits in wild and cultivated populations of *Medicago sativa* L. from Spain, *Genet. Select. Evol.* 30 (1998) S103–S121.
- [20] Jenczewski E., Prospero J.M., Ronfort J., Evidence for gene flow between wild and cultivated *Medicago sativa* (Leguminosae) based on allozyme markers and quantitative traits, *Am. J. Bot.* 86 (1999) 677–687.
- [21] Jenczewski E., Prospero J.M., Ronfort J., Differentiation between natural and cultivated populations of *Medicago sativa* (Leguminosae) from Spain: analysis with random amplified polymorphic DNA (RAPD) markers and comparison to allozymes, *Mol. Ecol.* 8 (1999) 1317–1330.
- [22] Lelandais C., Guttieres S., Mathie, C., Vedel F., Remacle C., Marechal-Drouard L., Brennicke A., Binder S., Chetrit P., A promoter element active in run-off

- transcription controls the expression of two cistrons of *nad* and *rps* genes in *Nicotiana sylvestris* mitochondria, *Nucleic Acids Res.* 24 (1996) 4798–4804.
- [23] Palmer J.D., Mitochondrial DNA in plant systematics: applications and limitations, in: Soltis, P.S., Soltis, D.E., Doyle, J.-J. (eds.) *Molecular systematics of plants*, Chapman & Hall, London, 1992, pp. 36–39.
- [24] Petit R.J., Kremer A., Wagner D.B., Finite island model for organelle and nuclear genes in plants, *Heredity* 71 (1993) 630–641.
- [25] Prospero J.M., Angevain M., Bonnin I., Chaulet E., Génier G., Jenczewski E., Olivieri L., Ronfort J., Genetic diversity, preservation and use of genetic resources of mediterranean legumes: Alfalfa and Medics, in: Génier G., Prospero J.M. (eds.), *Proceedings of the meeting of the Mediterranean Working Group on Medicago of the FAO/CIHEAM Inter-Regional Research and Development Network on Pastures and Fodder Crops*, Hammamet (Tunisia), 19–22 October 1995, *Cahiers Options Méditerranéennes Vol. 8: The Genus Medicago in the Mediterranean Region: Current situation and prospects in research*, CIHEAM, Zaragoza, 1996 pp. 71–89.
- [26] Raymond M., Rousset F., An exact test for population differentiation, *Evolution* 49 (1995) 1280–1283.
- [27] Raymond M., Rousset F., GENEPOP (version 1.2): population genetics software for exact tests and ecumenism, *J. Hered.* 86 (1995) 248–249.
- [28] Rieseberg L.H., Hybrid origin of plant species, *Ann. Rev. Ecol. Syst.* 28 (1997) 359–389.
- [29] Rieseberg L.H., Wendel J.F., Introgression and its consequences in plants, in: Harrison, R.G. (ed.), *Hybrid zones and the Evolutionary process*, Oxford University Press, NY, 1993, pp. 70–110.
- [30] Rieseberg L.H., Whitton J., Gardner K., Hybrid zones and the genetic architecture of a barrier to gene flow between two sunflower species, *Genetics* 152 (1999) 713–727.
- [31] SAS® (Institute Inc., SAS/STAT User's guide, version 6, 4th ed., SAS Institute, Cary, NC, 1989.
- [32] Small E., Hybridization in the domesticated weed-wild complex, in Grant, W.F. (ed.), *Plant Biosystematics*, Academic press, Toronto, 1984, pp. 195–210.
- [33] Schuster W., Brennicke A., Pseudocopies of the ATPase α -subunit gene in *Oenothera* mitochondria are present on different circular molecules, *Mol. Gen. Genet.* 204 (1986) 29–35.
- [34] Tai T.H., Tanksley S.D., A rapid and inexpensive method for isolation of total DNA from dehydrated plant tissue, *Plant Mol. Biol. Rep.* 8 (1990) 297–303.
- [35] Van Raamsdonk L.W.D., Wild and cultivated plants: the parallelism between evolution and domestication, *Evol. Trends Pl* 7 (1995) 73–84.
- [36] Wissinger B., Schuster W., Brennicke A., Trans-splicing in *Oenothera* mitochondria: *nad1* mRNAs are edited in exon and trans-splicing group II intron sequences, *Cell* 65 (1991) 473–482
- [37] Wolfe K.H., Li W.H., Sharp P.M., Rates of nucleotide substitution vary greatly among plant mitochondrial, chloroplast, and nuclear DNAs, *Proc. Natl. Acad. Sci. USA* 84 (1987) 9054–9058.
- [38] Wu J., Krutovskii K.V., Strauss S.H., Abundant mitochondrial genome diversity, population differentiation and convergent evolution in pines, *Genetics* 150 (1998) 1605–1614