Partitioning and distribution of RAPD variation in a set of populations of Medicago sativa complex

M.L. Crochemore, Christian C. Huyghe, Marie-Claire Kerlan, F. Durand, Bernadette Julier

agronomie (1996) 16, 421-432

Partitioning and distribution of RAPD variation in a set of populations of the *Medicago sativa* complex

ML Crochemore, C Huyghe *, MC Kerlan, F Durand, B Julier

Station d'amélioration des plantes fourragères, Inra, F-86600 Lusignan, France

(Received 21 December 1995; accepted 29 May 1996)

Summary — Lucerne *(Medicago sativa)* is a major perennial forage legume and includes two main sub-species. The variation available within a group of 26 tetraploid populations of this complex was investigated with random amplified polymorphic DNA (RAPD) markers. Thirty seedlings per population were analysed. Twenty-nine reproducible markers, 24 being polymorphic, were obtained with four primers. The variation partitioning was studied using the AMOVA technique. The genetic variation proved to be nearly equally distributed within and between populations. The between-population variation was further partitioned between groups (*falcata*, Flemish, mediterranean) and between populations within groups. The latter source of variation was the major one. Although this study included a sub-species level, the within-population variation was very large probably due to the outcrossing reproduction and the tetraploidy. A similar approach was used to distinguish varieties and proved to be very efficient. The within-population dissimilarity indices were very variable according to the populations; the *falcata* and Flemish-type materials showed on average a larger within-population dissimilarity. The between-population dissimilarities were calculated and a dendrogram was drawn. This made it possible to separate the populations belonging to the two sub-species and the populations of subspecies *sativa* largely introgressed by *falcata*. The relationships among the *sativa* populations partly fitted with the known origin of the material or with their agronomic behaviour.

***Medicago sativa* = lucerne / RAPD marker / variation pattern**

Résumé — Variation intra- et inter-population et relations entre populations du complexe *Medicago sativa* mises en évidence à l'aide de marqueurs RAPD. *La luzerne* (Medicago sativa) *est une légumineuse fourragère majeure comprenant deux sous-espèces principales,* sativa *et* falcata. *La variation disponible au sein d'un ensemble de 26 populations tétraploïdes de ce complexe a été analysée à l'aide de marqueurs RAPD* (random amplified polymorphic DNA). *Trente plantules par population ont été étudiées. Vingt-neuf marqueurs reproductibles dont 24 marqueurs polymorphes ont été mis en évidence à l'aide de quatre amorces. La variation génétique qui a été étudiée à l'aide de la technique Amova est distribuée de façon presque égale au sein des populations et entre les populations. La variation inter-population a été répartie en variation inter-groupes (*falcata, *flamand et méditerranéen) et entre populations au sein des groupes. Cette dernière source de variation est la plus importante. Bien que cette étude ait pris en compte la variation d'origine sous-spécifique, l'importante variation intra-population est vraisemblablement liée à la reproduction allogame et à l'autotétraploïdie de la luzerne pérenne. La même technique a été utilisée pour étudier la distinction entre les variétés et s'avère très performante pour cela. La dissimilarité intrapopulation varie selon la population, les populations dormantes et de type flamand, montrant en moyenne une plus grande dissimilarité intrapopulation. Les dissimilarités interpopulation ont été calculées et un dendrogramme a été tracé pour cet indice de distance. Il*

* Correspondence and reprints

permet de séparer les populations des deux sous-espèces ainsi que les populations sativa *fortement introgressées de* falcata. *Les relations entre les populations* sativa *correspondent partiellement à l'origine du matériel et à son comportement agronomique.*

Medicago sativa = *luzerne / RAPD / variation*

INTRODUCTION

Lucerne (*Medicago sativa* L), one of the world's most important forage legumes, is an autotetraploid, outcrossing and perennial species. The *M sativa* complex includes nine subspecies (Gunn et al, 1978), which are either diploid or tetraploid and include ssp *sativa* which has purple flowers, tap roots, erect plants with coiled pods and ssp *falcata* which has yellow flowers, fasciculate roots, prostate plants with sickle-shaped pods and a strong winter dormancy. At the tetraploid level, the crosses between *sativa* and *falcata* give fully fertile hybrids (Lesins and Lesins, 1979). The intermediate forms originating from such crosses are widely used in breeding and cultivation in northern Europe and northern America because of their frost and winter hardiness (Michaud et al, 1988).

Information about germplasm diversity, diversity distribution and genetic relationships among material under selection is of fundamental importance in breeding. Ideally, methods for understanding genetic relationships between populations and variation partitioning within and between populations in alfalfa should be based on comparison of plants using monogenic traits whose expression is not affected by plant development. Although the random amplified polymorphic DNA (RAPD) markers are dominant, they can usefully contribute to such an investigation.

RAPDs are dominant markers which were first developed by Welsh and McClelland (1990) and Williams et al (1990). They are random pieces of DNA amplified from the genome by a technique based on the polymerase chain reaction. RAPDs have been used for genome mapping (Carlson et al, 1991; Kazan et al, 1993), to describe phylogenetic relationships (Puterka et al, 1993), to identify cultivars (Hu and Quiros, 1991; Yu and Nguyen, 1994; Nienhuis et al, 1995) and to assess levels and patterns of genetic variation (Chalmers et al, 1992; Huff et al, 1993; Castagnone-Sereno et al, 1994; Baruffi et al, 1995; Nesbitt et al, 1995). Within the *Medicago* genus, RAPD markers have been used to estimate genetic relatedness in cultivated lucernes (Yu and Pauls, 1993a), to develop a genomic map (Echt et al, 1992; Kiss et al, 1993; Yu and Pauls, 1993b), to analyse the genetic variability in annual diploid species (Brummer et al, 1995; Bonnin et al, 1996), to characterize woody species of *Medicago* (Chebbi et al, 1995) and to tag genes (Yu and Pauls, 1993c).

The genetic material sampled in this study comprises tetraploid populations and varieties from the *M sativa* complex and covers a wide range of the genetic variability which forms the base for breeding populations in most countries (Michaud et al, 1988; Julier et al, 1995). The present study presents a methodological approach using RAPDs to investigate the partitioning of variation within and between populations and between groups of a set of 26 *M sativa* populations. This approach will be used to discriminate lucerne varieties. The relationships between populations, either natural or cultivated, based on the genetic distances, will be analysed.

MATERIALS AND METHODS

Plant material

Twenty-six tetraploid populations of *M sativa* ssp *sativa* and ssp *falcata* were involved in the study. Table I presents the origin of these populations and their genetic status. Monte Oscuro, Maron and Malzeville are wild populations. Monte Oscuro is a population from northern Spain and seeds were collected in 1981 (Delgado, 1989). Maron and Malzeville were collected in the east of France as seeds and cuttings respectively (C Ecalle, personal communication). The other 23 populations are cultivated, either local landraces or varieties showing a large range of winter dormancy. The cultivars and experimental genotypes are synthetics with different numbers of constituents. This number is not available for all the varieties. For those available, it ranges from 4 for 63-28P to 28 for Maya. The seeds of the varieties were kindly provided by the breeders. The seeds of the local landraces were multiplied at Lusignan (France) by intercrossing 80 plants grown from the seeds originally received. The number of progenitors of these landraces or the number of seeds or plants originally collected is not available. As a source of additional information, the percentage of variegated flowers is also given in table I. This percentage was scored on 90 spaced plants grown in the field at Lusignan from the same seed batches as those used

Table I. Genetic material used in this study.

Name	*Country*	*Sub-species*	*Type*	*Percentage of variegated flowers* [a] *(%)*	*Group* [b]	*Number of seedlings studied*
Maron	France	*falcata*	Wild	8	1	21
Malzeville	France	*falcata*	Wild	0	1	28
Flamande	France	*sativa*	Landrace	20	2	30
Luçon	France	*sativa*	Landrace	20	2	28
Maya	France	*sativa*	Variety	14	2	28
Europe	France	*sativa*	Variety	10	2	28
Rival	France	*sativa*	Variety	16	2	28
Orca	France	*sativa*	Variety	11	2	30
63-28P	France	*sativa*	Experimental	34	2	30
Luzelle	France	*sativa*	Variety	38	2	29
Magali	France	*sativa*	Variety	15	3	29
Coussouls	France	*sativa*	Variety	33	3	27
Sabre	Canada	*sativa*	Variety	53	2	28
Victory	Canada	*sativa*	Variety	28	2	29
Monte Oscuro	Spain	*sativa*	Wild	2	3	29
Mediterraneo	Spain	*sativa*	Variety	0	3	29
Lodi	Italy	*sativa*	Variety	2	3	27
Julus	Sweden	*sativa*	Variety	20	2	29
Dem3	Morocco	*sativa*	Landrace	0	3	30
Pool5	Morocco	*sativa*	Pool	1	3	30
Gabes	Tunisia	*sativa*	Landrace	0	3	28
Sewa	Egypt	*sativa*	Landrace	3	3	30
Higazi	Sudan	*sativa*	Landrace	0	3	27
CUF101	USA	*sativa*	Variety	0	3	29
WL514	USA	*sativa*	Variety	3	3	28
Nabuwakaba	Japan	*sativa*	Variety	0	2	29

[a] The percentage of variegated flowers was scored on spaced plants in the field; [b] group 1: *falcata* populations; group 2: Flemish-type populations; group 3: Provence and Mediterranean-type populations.

for the RAPD analysis. For the analysis of the data, the populations were gathered into three groups: *falcata* populations; Flemish type populations; and Provence and Mediterranean-type populations (table I).

Thirty seedlings per population were grown in a growth cabinet set to a continuous light at a constant temperature of 18 °C. The last trifoliolate leaves were sampled on each plant to yield 30–40 mg of fresh material. The leaves were oven-dried overnight at 60 °C before DNA extraction. The number of seedlings with a successful DNA extraction is given for each population in table I.

DNA extraction

Genomic DNA was extracted from each plant by homogenizing the dry leaves in 800 µL of extraction buffer (made by mixing 100 mM Tris HCl pH 8.0, 5.0 mM EDTA pH 8.0, 500 mM NaCl, 1.25% SDS (w/v) and 0.38 g of sodium bisulfite) in a sterile Eppendorf tube (Tai and Tanksley, 1991). The proteins were precipitated by adding 250 µL, potassium acetate 5 M and incubated on ice for 30 min. The samples were centrifuged at 12 000 rpm for 3 min. The supernatants were removed to new Eppendorf tubes. The DNA was precipitated by adding one volume of pure isopropanol at –20 °C for 1 h. After two washes in 75% ethanol, the DNA pellets were dissolved in 100 µL of TE buffer. RNAse treatment was performed by adding 2.5 µL RNAse (2.5 µg/mL) for 15 min. The DNA was precipitated by sodium acetate 3 M (1/10 vol) and 3 vol of ethanol 95% and incubated at –20 °C for 1 h. The samples were centrifuged at 15 000 rpm for 5 min. The pellets were washed with ethanol 75% and dissolved in 50 µL of sterile distilled water without agitation. The DNA was quantified using a Hoefer Scientific TK0100 fluorometer.

DNA amplification and separation

The DNA amplification protocol was adapted from Williams et al (1990). Each 25 µL reaction contained 25 ng of DNA, 10 mM Tris HCl pH 8.3, 50 mM KCl, 1.5 mM $MgCl_2$, 0.01% gelatine, 0.2 mM each of dATP,

dCTP, dGTP, dTTP (Pharmacia), 0.2 µM of random primer and 0.4 units of *Taq* polymerase (Stehelin and Co, Basel, Switzerland). The reactions were overlaid with mineral oil. Amplification was performed in a Perkin Elmer thermocycler starting with 4 min of denaturation at 94 °C followed by 37 cycles of 1 min at 93 °C, 1 min at 45 °C and 1 min at 72 °C. The last cycle ended with 6 min at 72 °C. The RAPD fragments were separated by 1.8% agarose-gel electrophoresis with TAE 1x buffer (0.04 M Tris acetate pH 8.0 and 10 mM EDTA) and visualized with ethidium bromide.

Four ten-mer Operon primers (Operon Technologies Inc, Alameda, CA, USA), set B, were used. They were known to reveal polymorphism on some populations of *M sativa* (Huguet, personal communication). Banding patterns were obtained on a total of 737 individual plants for each primer. The markers are identified in the tables and figures by the name of the primer and by the length of the fragment.

Statistical analysis

Mean population analysis

For each population, the percentage of presence of the different bands was calculated. The frequencies of occurrence of the polymorphic loci were analysed through principal component analysis.

Individual analysis

The dissimilarity index *(d)* was calculated by comparing banding patterns between pairs of individuals according to the following formula:

$$d = N_{AB} / (N_A + N_B)$$

where N_A and N_B are the number of fragments in individuals A and B respectively and N_{AB} is the number of fragments that differ between the two individuals (Gilbert et al, 1990; Yukhi and O'Brien, 1990). A matrix of genetic distances between all individuals based on dissimilarity indices was obtained.

Variation partitioning

Components of variance of the genetic distance attributable to differences between groups, between populations within groups, and between individuals within populations were estimated from this matrix using AMOVA (analysis of molecular variance; Excoffier et al, 1992). AMOVA variance components were used as estimates of the genetic diversity, partitioned into within and between populations within and between groups. The number of permutations between all individuals for testing the significance of the variances was set at 100. The null distribution of the within-population variance was tested by allocating each individual to a randomly chosen population. The null distribution of the between-populations within-groups variance was tested by permuting individuals within groups without regard to population. The null distribution of the between-group variance was tested by permuting whole populations across groups. With this last system of permutations, the size of the groups was variable.

Variety distinction

The 15 varieties of our study were pairwise analysed using the AMOVA approach with two variance components: within varieties, between varieties. The null distribution of the between-varieties variance was tested by 200 random permutations of the individuals without regard to variety. This approach was only performed on the varieties which are expected to have similar constitutions even if the number of constituents in the original polycross may be different.

Within-population and between-population dissimilarities

An estimate of within-population dissimilarity was calculated. If n individuals from the same population were compared, the average of the $(n(n-1)/2)$ dissimilarity values gives an estimate of the within-population dissimilarity. Similarly, the dissimilarity value, d_{xy}, between two populations x and y is given by the average of the $(n_x\ n_y)$ dissmilarity values between the n_x and n_y individuals of the two populations.

The between-population dissimilarities were used to draw a dendrogram with the unweighted pair group mathematical average method of aggregation.

To test differences between populations for the within-population dissimilarity value, a bootstrap routine (Efron, 1979) was performed. As proposed by Van Dongen (1995), resampling was performed across individuals to reflect the variation due to the sampling of individuals within the populations. Thirty random samples were created for each population. Similarly, by bootstraping simultaneously in two populations, a series of 30 distance matrices of dissimilarities was created. The dendrogram with the UPGMA method was drawn for each of these matrices. A majority-rule consensus tree was built using the CONSENSE procedure of PHYLLIP 3.5c (Felsenstein, 1985). To run the majority-rule consensus procedure of PHYLLIP 3.5c, which was first defined to bootstrap over loci, the dendrogram obtained from each of the 30 iterations was converted into a data format compatible with the CONSENSE procedure.

RESULTS

Frequencies among populations

In the material involved in this study, banding patterns were observed on 737 plants. Twenty-nine

bands were observed and selected as they proved to be reproducible. Twenty-four of these bands were polymorphic. The mean frequencies of the different polymorphic bands are presented in table II alongside the standard deviation and the minimum and maximum values among populations.

A principal component analysis performed on the frequencies illustrates the distribution of the variation for the different bands among the populations. Figure 1 shows the distribution of the bands (fig 1a) and of the populations (fig 1b) on the first two axes which represent only 25% of the total variation. Figure 1A illustrates that there is no association of bands among the populations under study. The analysis of the correlations between the frequencies of the bands in the different populations showed that out of the 276 correlations, only 11 were significant at the $P < 0.05$ level. This means low levels of links between bands. Figure 1B shows that most of the variation observed for the frequencies calculated for each band and taken into account in the first two axes was due to a limited number of populations, Maron, Malzeville, Maya, 63-28P, Luzelle and Rival, ie, mainly *falcata* populations and Flemish-type populations.

Variance partitioning

When the AMOVA was first performed with only within-population and between-population effects, there was a significant contribution ($P < 0.01$) of the between-population variance to the total variance. A new analysis with a between-group effect and a between-population within-group effect was performed. The results of the AMOVA partitioning are given in table III. Although the populations under study cover a wide range of the genetic variability available for breeding, including a sub-species variation, the within-population variation still accounts for 50.6% of the total variance. It is consequently necessary to analyse both the within-population and between-population variations. Both the between-population within-group and between-group variance significantly contributed to the total variance. The between-group effect accounted for 7.7% of the variation. This important contribution to the total variance is mainly due to the presence of *falcata* populations. Indeed, when testing the null distribution for the between-group variance, the highest values of the estimate of the between-group variance were

Table II. Mean frequency over the whole population of the different RAPD polymorphic loci; and standard deviation minimum and maximum values among the 26 lucerne populations.

Variable	*Mean*	*Standard deviation*	*Minimum*	*Maximum*
B01-400	0.8511	0.3220	0	1.0000
B01-440	0.9920	0.0230	0.8929	1.0000
B01-570	0.0449	0.1975	0	1.0000
B01-600	0.0771	0.1257	0	0.4286
B01-700	0.8782	0.2823	0	1.0000
B01-710	0.0299	0.0835	0	0.3667
B01-730	0.0082	0.0308	0	0.1429
B01-740	0.0596	0.2139	0	1.0000
B01-940	0.0380	0.1429	0	0.6667
B01-960	0.4334	0.2580	0	0.8929
B01-980	0.4827	0.4919	0	1.0000
B06-470	0.0320	0.1235	0	0.6207
B06-830	0.8672	0.3162	0	1.0000
B06-1020	0.8459	0.2290	0	1.0000
B06-1350	0.2455	0.3046	0	1.0000
B07-385	0.9481	0.1115	0.6429	1.0000
B07-580	0.0318	0.0882	0	0.3448
B07-740	0.9908	0.0247	0.9000	1.0000
B07-980	0.6614	0.2867	0	1.0000
B10-480	0.9444	0.0650	0.7500	1.0000
B10-550	0.2916	0.2434	0	1.0000
B10-580	0.0208	0.0755	0	0.3333
B10-700	0.2571	0.1657	0.0345	0.7857
B10-1250	0.6224	0.3146	0	1.0000

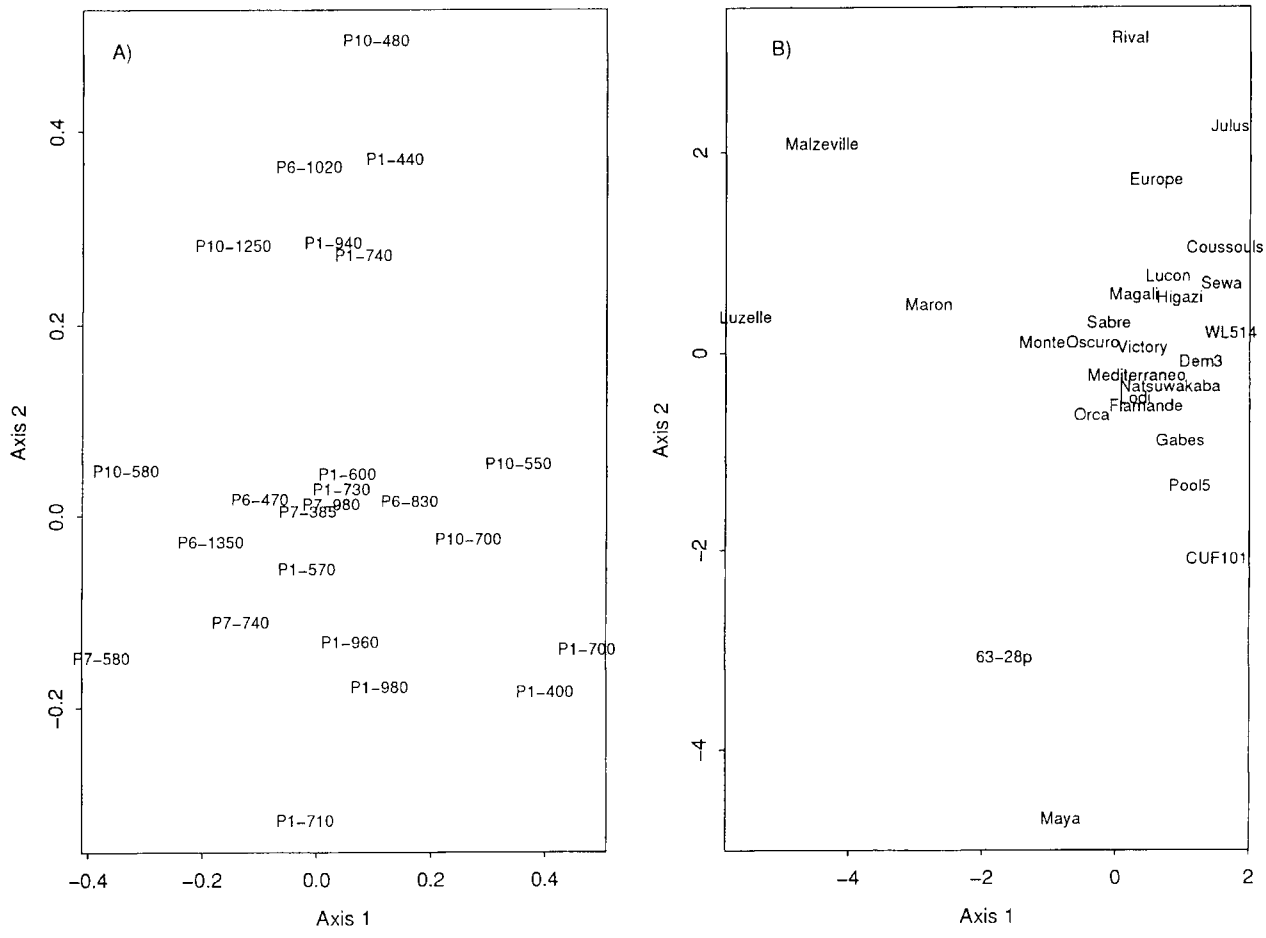

Fig 1. Graph of the first two axes of the principal component analysis performed on the frequencies of presences of the bands in the different populations. **A**) The contribution of the different markers to the axes; **B**) the distribution of the populations.

obtained when Maron and Malzeville were randomly allocated to the first group. The lowest values were obtained when they were randomly allocated to different groups. The between-population within-group effect contributed more than 40% to the total variance. This is likely to be due to the very broad basis of the different groups, the Flemish group, for instance, gathering both the very dormant north American variety Sabre and the landrace Luçon, which probably originates from a different genetic background.

Variety distinction

The varieties involved in this study were compared pairwise. Table IV shows the contribution of the between-variety variance to the total variance in each of the possible comparisons and the probability of obtaining a more extreme estimate of the between variance by random permutations of individuals across varieties. In four combinations only, the between-variety variance is not significant at the $P < 0.05$ level. These four

Table III. Analysis of molecular variance (AMOVA) for 29 RAPD markers in 26 populations of *M sativa*.

Source of variation	df	*MSD*	*Variance component*	*Total (%)*	P
Between group	2	232.7×10^{-2}	0.67×10^{-2}	7.7	< 0.01
Between population within group	23	96.5×10^{-2}	3.64×10^{-2}	41.7	< 0.01
Within population	711	4.42×10^{-2}	4.42×10^{-2}	50.6	< 0.01

Mean squared deviations (MSD), variance component estimates, percentage of total variation (total %) contributed by each component and the probability *(P)* of obtaining a more extreme estimate by chance alone are presented.

combinations are Natsuwakaba–Europe, and the three possible combinations between Lodi, CUF101 and WL514. These last three varieties differ only by slight differences in the frequencies of three bands. The AMOVA approach shows a high potential to discriminate lucerne varieties with only four primers and 24 polymorphic loci. The addition of an extra adequate primer to the set of the four primers used in the present experiment is likely to enable the distinction of all the varieties.

Within-population dissimilarity

The populations were compared for the value of the within-population dissimilarity. An analysis of

Table IV. Percentage of the between-varieties variance in the total variance and the probability of obtaining a more extreme estimate of the between varieties by chance alone (in italics) using an AMOVA approach.

	1	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*	*11*	*12*	*13*	*14*
1) Natsuwakaba	–													
2) Magali	26.5 *0.000*	–												
3) Europe	**20.3** ***0.200***	27.6 *0.000*	–											
4) Rival	48.6 *0.000*	50.2 *0.000*	43.4 *0.000*	–										
5) Maya	24.1 *0.035*	30.4 *0.000*	38.4 *0.000*	49.1 *0.000*	–									
6) Orca	42.5 *0.000*	41.1 *0.000*	39.9 *0.000*	60.6 *0.000*	42.3 *0.000*	–								
7) Luzelle	56.6 *0.000*	54.0 *0.000*	44.3 *0.000*	54.7 *0.000*	58.2 *0.000*	59.0 *0.000*	–							
8) Julus	53.2 *0.000*	57.3 *0.000*	50.0 *0.000*	39.1 *0.000*	54.4 *0.000*	63.8 *0.000*	58.8 *0.000*	–						
9) Lodi	37.6 *0.000*	40.0 *0.000*	34.8 *0.030*	43.3 *0.000*	45.5 *0.000*	55.5 *0.000*	46.8 *0.000*	46.2 *0.000*	–					
10) Mediterraneo	48.9 *0.000*	57.3 *0.000*	44.8 *0.000*	55.7 *0.000*	52.4 *0.000*	62.6 *0.000*	59.0 *0.000*	54.9 *0.000*	37.8 *0.000*	–				
11) Coussouls	49.5 *0.000*	46.2 *0.000*	47.4 *0.000*	64.9 *0.000*	47.3 *0.000*	56.0 *0.000*	67.7 *0.000*	62.0 *0.000*	64.8 *0.000*	67.6 *0.000*	–			
12) WL514	44.9 *0.000*	41.6 *0.000*	40.0 *0.000*	47.6 *0.000*	49.2 *0.000*	53.0 *0.000*	51.2 *0.000*	46.7 *0.000*	**18.5** ***0.990***	43.9 *0.000*	55.5 *0.000*	–		
13) CUF101	36.7 *0.000*	43.2 *0.000*	40.3 *0.000*	46.3 *0.000*	40.1 *0.000*	53.3 *0.000*	50.3 *0.000*	43.1 *0.000*	**14.6** ***0.985***	40.3 *0.000*	55.2 *0.000*	**12.5** ***0.930***	–	
14) Sabre	38.5 *0.000*	46.8 *0.000*	31.1 *0.000*	42.3 *0.000*	43.0 *0.000*	50.8 *0.000*	40.1 *0.000*	47.4 *0.000*	37.4 *0.000*	46.6 *0.000*	51.2 *0.000*	33.3 *0.000*	31.1 *0.000*	–
15) Victory	32.7 *0.000*	42.7 *0.000*	35.4 *0.000*	**27.9** ***0.050***	42.3 *0.000*	54.4 *0.000*	49.5 *0.000*	39.7 *0.000*	**28.2** ***0.050***	40.9 *0.000*	53.4 *0.000*	34.0 *0.000*	30.4 *0.000*	24.4 *0.000*

The non-significant values ($P \geq 0.05$) are in bold.

variance of the average within-population dissimilarity was performed using the iterations of the bootstrap as replicates. Significant differences between populations were found ($P < 0.01$). The comparisons of the means, using an SNK test which controls the Type I experiment-wise error rate under the complete null hypothesis, are presented in table V. Sabre, a Canadian variety, was four times more variable than Coussouls, a French variety of mediterranean type.

Even if the dormant material, ie, the Flemish-type populations and the *falcata* populations, shows on average a larger within-population dissimilarity than the mediterranean-type material (data not shown), there are contrasted situations within each group. For instance, Gabes, a Tunisian landrace, shows a within-population dissimilarity as wide as that of Sabre. Dem3 and Pool5, which have different origins and constitutions, were the more variable mediterranean materials; Dem3 was a landrace collected from the valley of Demnate region in Morocco in 1981–1982 (Birouk et al, 1989), while Pool5 was built by pooling five Moroccan populations from Saharian oasis and mountain regions (Birouk and Guy, 1986).

Table V. Comparisons of the populations for the within-population dissimilarity.

Population	*Mean*	*SNK grouping*	
Sabre	0.129	a	
Flamande	0.119	b	
Gabes	0.112	c	
Maron	0.110	c	
Maya	0.108	c	
Europe	0.099	d	
Luzelle	0.093	d	e
Luçon	0.093	d	e
CUF101	0.092	d	e
63-28P	0.084	f	e
Dem3	0.084	f	e
Sewa	0.081	f	g
Victory	0.080	f	g
Pool5	0.080	f	g
Orca	0.073	h	g
WL514	0.071	h	i
Naztsuwakaba	0.065	j	i
Monte Oscuro	0.065	j	i
Rival	0.064	j	k
Lodi	0.060	j	k
Mediterraneo	0.057	l	k
Magali	0.054	l	k
Higazi	0.054	l	k
Julus	0.049	l	
Coussouls	0.035	m	
Malzeville	0.034	m	

Values followed by the same letter do not differ at $P < 0.05$ using a SNK test controlling the experimentwise error rate.

The *falcata* populations, Maron and Malzeville, showed very contrasted structures with a wide within-population variation for Maron and a very narrow one for Malzeville. The way these populations were collected may explain this difference. While 4 800 seeds were collected to build the population Maron, cuttings were collected on the site of Malzeville. Some 122 cuttings were sampled out of which only 64 plants proved to be true *falcata* type (yellow flowers and sickled-shaped pods). The other plants were discarded before producing seeds. Out of the 64 plants, 50 produced seeds and were defined as the original Malzeville population. The mode of collection and the subsequent selection of true *falcata* types may explain the narrow genetic variability within this population.

The fact that the dormant material shows a larger within-population variation may be due to two factors. The first is that most of the dormant materials are varieties or experimental genotypes coming from breeding programs. The varieties are bound to precise methods of maintenance which avoid a rapid modification of the genetic basis. The natural populations or landraces used in this study had been multiplied several times and this could have led to a progressive narrowing of their genetic diversity. Moreover, the original genetic basis is not always known and the collection itself could have only taken into account a part of the genetic variability of the original population. The case of the *falcata* population Malzeville is a good example of this feature. A second factor could be the fact that most of the dormant populations show large introgressions of *M sativa* by *M falcata* and consequently are at the intersection of two gene pools.

Between-population dissimilarity

Figure 2 shows the dendrogram based on the between-population dissimilarities using the UPGMA method. The values at the different nodes of the dendrogram indicate the frequencies of occurrence of the corresponding nodes in the 30 different samples of bootstrap.

The dendrogram separates Gabes, Malzeville, Luzelle, Sabre and Maron from the other populations. The other populations are distributed in two sub-groups with three French Flemish-type vari-

eties, Maya, Orca and Europe, in an intermediate position. In each of the two sub-groups, there is a cluster of Mediterranean populations, Mediterraneo, CUF101, WL514, Pool5 and Lodi in one sub-group and Coussouls, Magali, Higazi, Dem3 and Sewa in the other.

Apart from the position of the intermediate group of Maya, Orca and Europe, the groups of populations appear fairly stable against the bootstrap on the individuals in each population. The relatively low frequencies of occurrence of the intermediate group are due to the small distances between each of these populations and the rest of the sub-groups. Small changes in the distances as a consequence of resampling on the individuals lead to modification of the position in the tree and reduce the reliability of the nodes.

DISCUSSION

This study only used four primers which showed 29 bands, out of which 24 were polymorphic. This limited number of bands reduces the possibility of drawing phylogenetic relationships between populations and the dendrogram must be analysed with caution. In most studies of relationships between populations and cultivars a much larger number of polymorphic loci are included. In *Brassica oleracea*, Kresovich et al (1992) used 117 polymorphic fragments. In *Phaseolus lunatus*, Nienhuis et al (1995) used 125 markers. For discrimination between samples or varieties, the number of markers are usually smaller. Yu and Pauls (1993a) studying the genetic relatedness on bulked DNA of lucerne varieties and Van De Ven and McNicol (1995) on

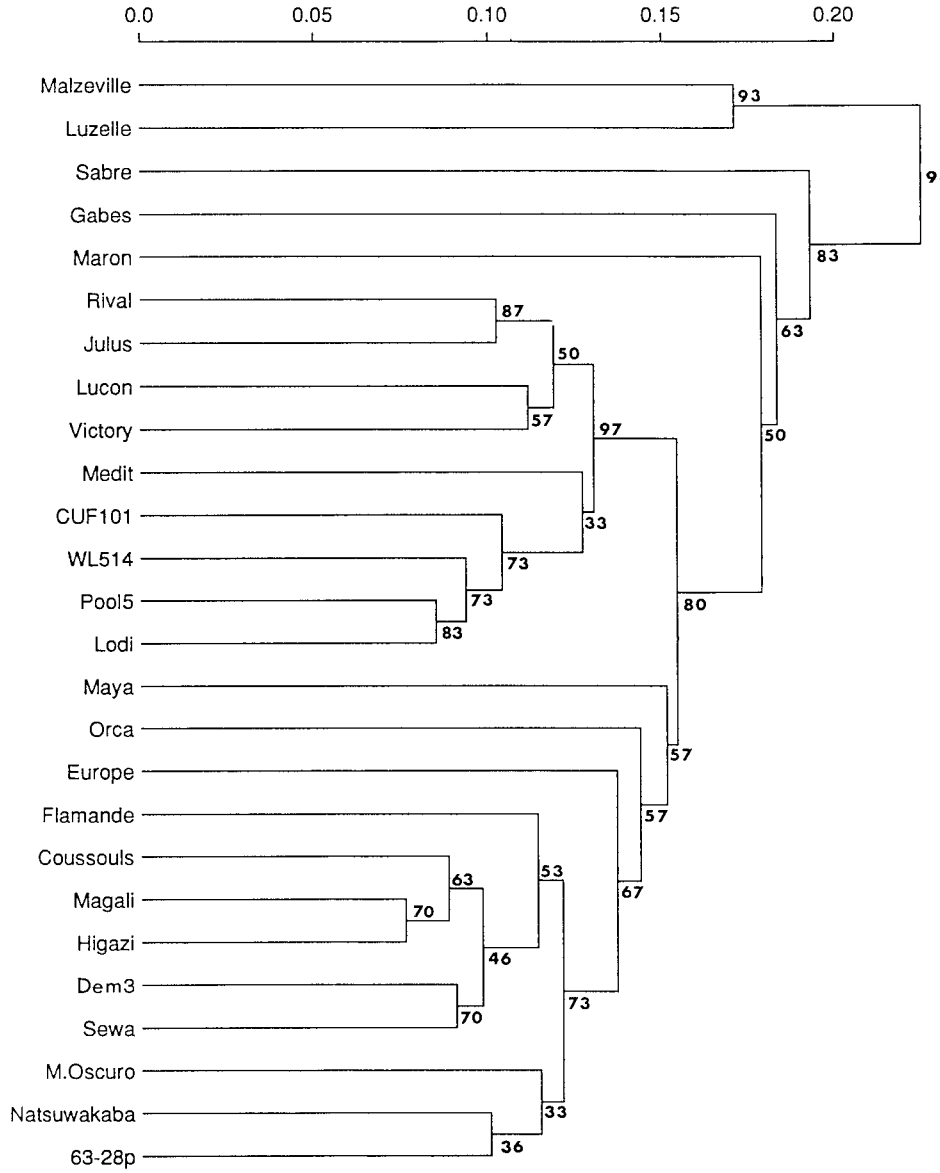

Fig 2. Dendrograms of the 26 populations of *M sativa* derived from RAPD dissimilarities values. Numbers are bootstrapping indices of the level of support for individual nodes calculated with Phyllip 3.5c package on 30 resamplings over individuals within population.

Sitka spruce clones used ten primers with on average five products per primer and six primers yielding a total of 24 polymorphic bands respectively. When the partitioning of variation was studied, the number of markers was variable and related to the number of DNA samples. Chalmers et al (1992) in *Gliciridia* scored 63 polymorphic bands on 50 samples. Nesbitt et al (1995) scored 147 polymorphic bands on 52 samples of *Eucalyptus*. In *Picea abies*, Bucci and Menozzi (1995) scored 20 markers on 288 DNA samples. At the interspecific level of *Meloidogyne*, Castagnone-Sereno et al (1994) used five primers yielding 74 polymorphic bands on 12 samples. Barasubyie et al (1995) scored 39 markers to partition the variation among 35 *Verticillium albo-atrum* isolates. In our study, where we had 737 DNA samples from 26 lucerne populations, we chose to reduce the number of primers to analyse all the DNA samples and get an optimal estimate of the variance partitioning.

The RAPD variation in *M sativa* is equally distributed within and between populations. To our knowledge, this is the first report of the within- and between-population variation in *M sativa* using RAPD markers. The level of partitioning of RAPD variation has been examined in both outcrossing and selfing species. The results of these studies have been variable depending on the range of material under study and on the breeding systems of the species involved. In buffalograss *(Buchloë dactyloides)*, a cross-pollinating species, Huff et al (1993) found that the within-population variation accounted for 72.9 to 80.5% of the variation according to the geographic origin. Similarly, in *Eucalyptus globulus* which is also a cross-pollinating species, Nesbitt et al (1995) found a within-population variation ranging from 73.8 to 94.9% according to the geographic origin. On a leguminous tree, *Gliciridia sepium,* which is an obligate outbreeder, Chalmers et al (1992) found that the within-population variation accounted for 40.1% of the variation. This low percentage of variation is similar to the variation found by Dawson et al (1993) on *Hordeum spontaneum* (43%) which has high selfing levels. In *M truncatula*, a self-pollinated legume, Bonnin et al (1996) found that the within-population variance component, in a set of four populations collected in the south of France, accounted for 55% of the total variance.

M sativa is an allogamous species and so the intermediate level of the within-population variation found in our study (50%) is likely to be due to the material involved. The populations were deliberately chosen to cover a wide range of variation and a wide range of origin. The sub-species variation of *M sativa* was taken into account in the present study as some *falcata* populations were included. This significantly contributed to the between-population variance, as the between-group variance accounted for 7.7% of the total variance and was greatly due to the presence of two *falcata* populations. Despite the wide range of populations involved in the study, the within-population variation was very large. This clearly indicates that the within-population variation may be a major source of variation for characters which have not yet been selected for. The prevalence of the within-population variation was also demonstrated on tetraploid populations of *Trifolium pratense*, another cross-pollinating perennial forage legume (Kongkiatngam et al, 1995). *M sativa* is a long-lived, outcrossing species and so it is the type of species in which Hamrick and Godt (1989) found the highest levels of allozyme variation within populations. However, the codominant status of the isozyme systems is not directly relevant for the dominant markers such as RAPD. The bias due to the dominance on a tetraploid species such as lucerne could be investigated.

The AMOVA technique made it possible to partition the variation for RAPD markers within and between populations. This approach was first developed on haplotypes and can be directly used on haploid organisms such as fungi (Barasubiye et al, 1995) or on selfing species. It has been successfully used for RAPD markers on outcrossing species and heterozygous specimens (Nesbitt et al, 1995). In lucerne, it is possible to use it to compare varieties in a pairwise test and assess them for distinction. In our study, in 101 of the 105 possible pairwise comparisons, the varieties proved to be different at a probability level of 5%. Consequently, with a reasonable amount of work (DNA extraction and PCR amplification with four primers) and with a reasonable level of reproducibility (major bands and stability against resampling), the strategy of describing a variety by 30 of its individuals is reliable. When increasing the number of varieties involved in the study, the number of primers will need to be larger but only in a reasonable proportion.

The dendrogram of the 26 populations derived from RAPD dissimilarity values does not completely fit with the known origin of the material and this may be clearly explained by the small number of polymorphic loci. Gabes from Tunisia belongs to a separate germplasm group and

does not connect with the other populations from northern or eastern Africa. The two *falcata* populations and the varieties Luzelle and Sabre are away from the rest of the material. Luzelle and Sabre are the only two populations of this study with plants with yellow flowers in the original polycross and with plants with variegated flowers in the commercial seed, both indicating the presence of *falcata*-type progenitors.

While the two wild *falcata* populations and the wild *sativa* Monte Oscuro show very similar patterns of growth (Julier et al, 1995), they are very different for the RAPDs. This shows that they belong to different gene pools but evolved similarly on the morphological and agronomical points of view in relation to similar natural selection pressures. This dendrogram must be cautiously considered because of the limited number of loci included in the present study. As a consequence, bootstrapping over the bands would yield low levels of reliability because the bands provide different information.

Assuming that the initial sample was representative of the populations, the technique of bootstrap over individuals as suggested by Van Dongen (1995) proved to be successful in testing the differences between populations for the extent of the within-population dissimilarity. This strategy also made it possible to investigate how robust the relationships between populations are against resamples of the original data and consequently against resampling new seedlings in the same populations. As shown by the results of the consensus tree, the nodes of the dendrograms are reasonably stable and would be found again if 30 extra seedlings were analysed for each population. It also suggests that 30 individuals per population seems to be a reliable sample size even for lucerne populations with a large within-population dissimilarity. This sample size should be checked in a new experiment. Indeed, the bootstrap method does not generate any new data and assumes that the initial sample is a 'true' one. The possibility of a wrong initial sample cannot be avoided by resampling in the same set.

Using the methodological approach based on AMOVA and presented in this paper on a set of populations and varieties, plant breeders can use RAPD markers to distinguish between varieties and to compare new ones from those previously registered. Such an approach could be successfully used in the DUS (distinction–uniformity–stability) test. Similarly, it could be useful to distinguish and structure genetic resources.

ACKNOWLEDGMENT

ML Crochemore was supported by a fellowship of the CNPq (Conselho Nacional de Desenvolvimento Cientifico E Tecnologico, Brazil).

REFERENCES

Barasubiye T, Parent JG, Hamelin RC, Laberge S, Richard C, Dostaler D (1995) Discrimination between alfalfa and potato isolates of *Verticillium albo-atrum* using RAPD markers. *Mycol Res* 99, 1507-1512

Baruffi L, Damiani G, Guglielmino CR, Bandi C, Malacrida AR, Gasperi G (1995) Polymorphism within and between populations of *Ceratitis capitata*: comparison between RAPD and multilocus enzyme electrophoresis. *Heredity* 74, 425-432

Birouk A, Dattee Y, Sadiki M, Roumet P (1989) Évaluation agronomique et adaptation de populations marocaines de luzerne (*Medicago sativa* L). *agronomie* 9, 363-376

Birouk A, Guy P (1986) Partition de populations marocaines de luzerne locale en vue de leur constitution en réservoir massal. In: *Distance, divergence et variabilité génétique* (Inra Éditions), Séminaire du département de génétique et d'amélioration des plantes, Méribel, 196-203

Bonnin I, Huguet T, Gherardi M, Prosperi JM, Olivieri I (1996) High level of polymorphism and spatial structure in a selfing plant species, *Medicago truncatula* (Leguminosae), using RAPD markers. *Am J Bot* (in press)

Bucci G, Menozzi P (1995) Genetic variation of RAPD markers in a *Picea abies* Karst population. *Heredity* 75, 188-197

Brummer EC, Bouton JH, Kochert G (1995) Analysis of annual *Medicago* species using RAPD markers. *Genome* 38, 365-367

Carlson JE, Tulsieram LK, Glaubitz JC, Luk VWK, Kauffeldt C, Rutledge R (1991) Segregation of random amplified DNA markers in F1 progeny of conifers. *Theor Appl Genet* 83, 194-200

Castagnone-Sereno P, Vanlerberghe-Masutti F, Leroy F (1994) Genetic polymorphism between and within *Meloidogyne* species detected with RAPD markers. *Genome* 37, 904-909

Chalmers KJ, Waugh R, Sprent JI, Simons AJ, Powell W (1992) Detection of variation between and within populations of *Gliricidia sepium* and *G maculata* using RAPD markers. *Heredity* 69, 465-472

Chebbi H, Pascual-Villalobos MJ, Cenis JL, Correal E (1995) Caractérisation morphologique et moléculaire des espèces ligneuses du genre *Medicago*. *Fourrages* 142, 191-206

Dawson IK, Chalmers KJ, Waugh R, Powell W (1993) Detection and analysis of genetic variation in

Hordeum spontaneum populations from Israel using RAPD markers. *Mol Ecol* 2, 151-159

Delgado Enguita I (1989) Estudio de la variabilidad de las mielgas aragonesas en areas de precipitation annual inferior a 600 mm. PhD thesis, Universidad Politecnica de Madrid, Madrid

Echt CS, Erdahl LA, McCoy TJ (1992) Genetic segregation of random amplified polymorphic DNA in diploid cultivated alfalfa. *Genome* 35, 84-87

Efron B (1979) Bootstrap methods: another look at the jacknife. *Ann Stat* 7, 1-26

Excoffier L, Smouse PE, Quattro JM (1992) Analysis of molecular variance inferred from metric distances among haplotypes: application to human mitochondrial DNA restriction data. *Genetics* 131, 479-491

Felsenstein J (1985) Confidence limits on phylogenies: an approach using the bootstrap. *Evolution* (Lawrence, KS) 39, 783-791

Gilbert DA, Lehman N, O'Brien SJ, Wayne RK (1990) Genetic fingerprinting reflects population differentiation in the Californian Channel Island fox. *Nature* 344, 764-767

Gunn CR, Skrdla WH, Spencer HC (1978) Classification of *Medicago sativa* L using legume characters and flower colors. USDA-ARS Tech Bull 1574 US Government Printing Office, Washington, DC, USA

Hamrick JL, Godt MJW (1989) Allozyme diversity in plant species. In: *Plant Population Genetics Breeding, and Genetic Resources* (AHD Brown, MT Clegg, AL Kahler, BS Weir, eds), Sinauer, Sunderland, MA, USA, 43-63

Hu J, Quiros CF (1991) Identification of broccoli and cauliflower cultivars with RAPD markers. *Plant Cell Rep* 10, 505-511

Huff DR, Peakall R, Smouse PE (1993) RAPD variation within and among natural populations of outcrossing buffalograss [*Buchloë dactyloides* (Nutt) Engelm]. *Theor Appl Genet* 86, 927-934

Julier B, Porcheron A, Écalle C, Guy P (1995) Genetic variability for morphology growth and forage yield among perennial diploid and tetraploid lucerne populations (*Medicago sativa* L). *agronomie* 15, 295-304

Kazan K, Manners JM, Cameron DF (1993) Inheritance of random amplified polymorphic DNA markers in an interspecific cross in the genus *Stylosanthes*. *Genome* 36, 50-56

Kiss GB, Csanadi G, Kalman K, Kalo P, Okresz L (1993) Construction of a basic genetic map for alfalfa using RFLP, RAPD, isozyme and morphological markers. *Mol Gent Genet* 238, 129-137

Kongkiatngam P, Waterway MJ, Fortin MG, Coulman BE (1995) Genetic variation within and between two cultivars of red clover (*Trifolium pratense* L): comparisons of morphological, isozyme, and RAPD markers. *Euphytica* 84, 237-246

Kresovich S, Williams JGK, McFerson JR, Routman EJ, Schaal BA (1992) Characterization of genetic identities and relationships of *Brassica oleracea* L via a random amplified polymorphic DNA assay. *Theor Appl Genet* 85, 190-196

Lesins KA, Lesins I (1979) *Genus Medicago* (Leguminosae): *A Taxogenetic Study*. Junk bv Publishers, The Hague

Michaud R, Lehman WF, Rumbaugh MD (1988) World distribution and historical development. In: *Alfalfa and Alfalfa Improvement* (AA Hanson, ed), ASA-CSSA-SSSA Publishers, Agronomy monograph N° 29, Madison, WI, USA, 25-91

Nesbitt KA, Potts BM, Vaillancourt RE, West AK, Reid JB (1995) Partitioning and distribution of RAPD variation in a forest tree species, *Eucalyptus globulus* (Myrtaceae). *Heredity* 74, 628-637

Nienhuis J, Tivang J, Skroch P, DosSantos JB (1995) Genetic relationships among cultivars and landraces of lima bean (*Phaseolus lunatus* L) as measured by RAPD markers. *J Am Soc Hort Sci* 120, 300-306

Puterka GJ, Black WC, Steiner WM, Burton RL (1993) Genetic variation and phylogenetic relationships among worldwide collections of Russian wheat aphid, *Diuraphis noxia* (Mordvilko) inferred from allozyme and RAPD-PCR markers. *Heredity* 70, 604-618

Tai TH, Tanksley SD (1991) A rapid and inexpensive method for isolation of total DNA from dehydrated plant tissue. *Plant Mol Biol Rep* 8, 297-303

Van De Ven WTG, McNicol RJ (1995) The use of RAPD markers for the identification of Sitka spruce *(Picea sitchensis)* clones. *Heredity* 75, 126-132

Van Dongen S (1995) How should we bootstrap allozyme data? *Heredity* 74, 445-447

Welsh J, McClelland M (1990) Fingerprinting genomes using PCR with arbitrary primers. *Nucleic Acids Res* 18, 7213-7218

Williams JGK, Kubelick AR, Livak KJ, Rafalski JA, Tingey SV (1990) DNA polymorphisms amplified by arbitrary primers are useful as genetic markers. *Nucleic Acids Res* 18, 6531-6535

Yu K, Pauls KP (1993a) Rapid estimation of genetic relatedness among heterogeneous populations of alfalfa by random amplification of bulked genomic DNA samples. *Theor Appl Genet* 86, 788-794

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

Yu K, Pauls KP (1993c) Identification of an RAPD marker associated with somatic embryogenesis in alfalfa. *Plant Mol Biol* 22, 269-277

Yu LX, NGuyen HT (1994) Genetic variation detected with RAPD markers among upland and lowland rice cultivars (*Oryza sativa* L). *Theor Appl Genet* 87, 668-672

Yukhi N, O'Brien SJ (1990) DNA variation of the mammalian major histocompatibility complex reflects genomic diversity and population history. *Proc Natl Acad Sci USA* 87, 836-840