Theor Appl Genet (1998) 96 : 406–412

M. Ghérardi · B. Mangin · B. Goffinet
D. Bonnet · T. Huguet

A method to measure genetic distance between allogamous populations of alfalfa (*Medicago sativa*) using RAPD molecular markers

Received: 7 July 1997 / Accepted: 28 October 1997

Abstract Alfalfa (*Medicago sativa* L.) is a forage legume of world-wide importance whose both allogamous and autotetraploid nature maximizes the genetic diversity within natural and cultivated populations. This genetic diversity makes difficult the discrimination between two related populations. We analyzed this genetic diversity by screening DNA from individual plants of eight cultivated and natural populations of *M. sativa* and *M. falcata* using the RAPD method. A high level of genetic variation was found within and between populations. Using five primers, 64 intense bands were scored as present or absent across all populations. Most of the loci were revealed to be highly polymorphic whereas very few population-specific polymorphisms were identified. From these observations, we adopted a method based on the Roger's genetic distance between populations using the observed frequency of bands to discriminate populations pairwise. Except for one case, the between-population distances were all significantly different from zero. We have also determined the minimal number of bands and individuals required to test for the significance of between-population distances.

Key words *Medicago sativa* · RAPD · Cultivars · Genetic distance

Communicated by G. Wenzel

M. Ghérardi · D. Bonnet[1] · T. Huguet (✉)
Laboratoire de Biologie Moléculaire des Relations Plantes-Microorganismes, CNRS-INRA, BP27, Centre de Recherches INRA de Toulouse, 31326 Castanet-Tolosan Cedex, France
Fax: 33-(0)5 61 28 50 61
E-mail: thuguet@toulouse.inra.fr

B. Mangin · B. Goffinet
Station de Biométrie et d'Intelligence Artificielle, BP27, Centre de Recherches INRA de Toulouse, 31326 Castanet-Tolosan Cedex, France

Present address:
[1] Novartis Seeds S.A., 12 chemin de l'Hobit, B.P. 27, 31790 Saint Sauveur, France

Introduction

Alfalfa (*Medicago sativa*) is a forage crop of considerable world-wide importance (Barnes et al. 1988). However, despite its agronomical importance, breeding improvements are complicated by the high genetic complexity of this species and its cultivated forms. Cultivated alfalfa is an outcrossing autotetraploid plant developed by combining different *M. sativa* and *M. falcata* germplasm sources in order to maximize heterosis and to secure multiple pest resistance. In addition, cultivated alfalfa varieties exist as synthetic varieties developed by randomly intermating selected parents from the nine original germplasm sources. As a consequence, the genetic structure of alfalfa cultivars is imprecise compared to fixed cultivated forms like inbred lines or hybrids.

An accurate genetic characterization of an alfalfa cultivar is thus difficult, but is nevertheless necessary since plant breeders need to distinguish between two cultivars, even closely related ones, and to compare new ones from those previously registered. Methods for detecting isoenzyme polymorphisms were first used for genetic diversity analyses in order to trace the ancestors of alfalfa cultivars or to quantify the within and between population components of genetic diversity in alfalfa germplasms (Quiros 1983). However, the level of polymorphism detected with such markers in insufficient for cultivar identification. Neutral DNA-based molecular markers allows a more precise and non-environment-dependant identification of individual plant genotypes but the first described technique (RFLP), although useful and efficient, is very laborious when used routinely on a large number of plants (Brummer et al. 1991; Kidwell et al. 1994). The report that a single 10-mer oligonucleotide, of arbitrary nucleotide sequence, will direct the amplification of discrete loci has given rise to RAPD (random amplified polymorphic DNA; Williams et al. 1990) markers. These PCR-based

dominant markers considerably speed up processes like the construction of a genetic map (Reiter et al. 1992; Yu and Pauls, 1993) or the determination of the genetic diversity within species or populations (Hadrys et al. 1992). Since RAPD markers allow a rapid analysis of the polymorphism of many individuals, they were chosen to estimate genetic distances between cultivated alfalfa varieties. In order to take into account the genetic complexity of alfalfa cultivars, each individual was studied separately with several arbitrary primers. The within-population frequency of each marker could thus be calculated. We propose a mathematical method to estimate this genetic distance between alfalfa populations.

The objectives of the present study were to use the RAPD method: (1) to estimate the genetic distance between two alfalfa populations, and (2) to evaluate the minimal number of RAPD bands, and/or individuals, to be analyzed for this estimation.

Materials and methods

Plant material

The following cultivars representative of alfalfa germplasm origins (Julier et al. 1995) were employed: Oro (Provence type), Rival, Europe, Gemini and Euver (Flemish type), Medalfa (intermediate type) together with two tetraploid wild populations, Monte Oscuro (*M. sativa*) and Maron (*M. falcata*). The alfalfa varieties were provided by GEVES [Groupe d'Etudes et de Contrôle des Variétés et des Semences], Le Magneraud (France). The two tetraploid non-cultivated wild populations, Monte Oscuro (*M. sativa*) and Maron (*M. falcata*), were provided by Dr. P. Guy (Julier et al. 1995). Alfalfa seeds were surface-sterilized and germinated as previously described (Truchet et al. 1989) and grown in a growth chamber for 3 weeks.

DNA isolation

For each plant, about 20–40 mg of young leaves were taken and dried in an oven at 60°C overnight. DNA was then purified as described by Tai and Tanksley (1990). The yield was about 500 ng of DNA per mg of leaves (fresh weight).

RAPD-marker analysis

The experimental conditions employed were very similar to those described by Williams et al. (1990): 20 ng of DNA were amplified using 50 ng of 10-mer primers (Operon Technologies, Alameda, USA) in the presence of 1.5 mM of $MgCl_2$ and 0.4 *Taq* Polymerase units (Stehelin, Basel, Switzerland) during 40 cycles of PCR amplification. These conditions allowed a good reproducibility of the bands and of their relative intensity. Amplification products were identified using 2.4% agarose-gel electrophoresis. Analyses were done manually by an examination of the photograph of the gel after staining in 1 mg/l of Ethidium Bromide. For each DNA sample, amplification gave rise to a characteristic profile which was represented by a number of bands, each band being denoted as 0 (absence) or 1 (presence). The intensity of the bands was not taken into account, even if only the intense ones were considered.

Results

Can RAPD bands be found which discriminate between populations?

The first objective was to detect if some population-specific bands could be found. From each population, we arbitrarily selected 16–50 plants and analysed them with the RAPD technique. All the individuals were tested using five arbitrary 10-mer primers (Table 1): B1, B6, B7, B8 and B10. Each primer gave clear banding profiles showing various degrees of polymorphism. By pooling the five primers, a similar number of clear and reproducible bands (45–52) could be detected after amplification in each population (Table 2). The total number of bands observed was 64. Most of the bands (67–75%) were polymorphic, not only within each population but also between populations (Table 3, as an example with a B1 primer). Using population pairwise comparisons, it was shown that the number of bands which discriminated between two populations or cultivars ranged from 6 to 22 (data not shown). However, only a limited number of bands allowed a clear identification (Table 4) and only one band (out of 64) appeared population-specific: the 600-bp (B10 primer) band which was present in 32% of the plants of the *M. falcata* Maron population but never in the other populations. For other bands, only high frequency differences could be used to differentiate the populations: for instance with primer B1, the 650-bp band is present

Table 1 Nucleotide sequence of the primers used in RAPD analysis

Primer	5′ → 3′ Sequence
B1	GTTTCGCTCC
B6	TGCTCTGCCC
B7	GGTGACGCAG
B8	GTCCACACGG
B10	CTGCTGGGAC

Table 2 Within-population polymorphism. N refers to the number of informative individuals tested. The number of bands is the total number of RAPD bands observed in each population using the five primers. The % of polymorphic bands refers to bands which are not observed in all the tested individuals of a population

Population	N	Number of bands	% Polymorphic bands
Europe	45	52	73
Rival	47	46	70
Medalfa	47	49	71
Oro	35	45	75
Gemini	28	51	72
Euver	19	46	70
Monte-Oscuro	20	46	67
Maron	21	51	69

Table 3 Banding pattern polymorphism with the B1 primer. The data refer to the frequency (%) of individuals within each population showing the amplification band. Band length is expressed in base pairs. N is the number of individuals analysed in each population with a primer

Band length	Europe	Rival	Medalfa	Oro	Gemini	Euver	M.-Oscuro	Maron
1000	21	12	15	7	50	37	16	26
980	100	92	100	100	96	95	100	84
960	58	62	39	43	29	58	27	31
850	100	100	100	100	100	100	100	100
740	2	44	0	0	12	0	0	53
680	100	90	100	97	100	100	94	52
650	2	6	0	0	0	5	0	26
600	15	12	15	35	12	16	22	94
530	2	16	6	23	25	5	0	0
500	4	12	21	5	12	21	0	0
492	10	0	20	7	29	5	0	0
430	97	100	87	95	100	100	100	100
N	46	50	46	39	24	19	18	19

Table 4 Frequency (%) of ten selected amplified bands. Band length is expressed in base pairs

Primer	Band length	Europe	Rival	Medalfa	Oro	Gemini	Euver	M.-Oscuro	Maron
B1	650	2	6	0	0	0	5	0	26
B1	740	2	44	0	0	12	0	0	53
B6	480	10	0	4	8	0	0	65	22
B6	1350	4	54	6	2	0	0	4	0
B7	430	4	18	0	0	8	10	25	56
B7	1018	24	4	0	37	75	89	94	100
B8	1050	26	0	41	26	41	26	35	4
B10	600	0	0	0	0	0	0	0	32
B10	920	0	0	2	0	4	0	0	64
B10	1476	30	64	34	30	41	26	79	0

at 26% in Maron, the 740-bp band at 53% in Maron and 44% in Rival cultivars, but both bands were also observed in other populations with a lower frequency. By contrast, the 1050-bp band (obtained with B8) was lacking only in the Rival cultivar.

Reflecting the high level of genetic polymorphism found in alfalfa with these markers, it was not possible to find two individuals having the same genotype band pattern over all the primers studied, even within a single population (data not shown).

Estimation of the genetic distance between complex allogamous populations using RAPD markers

Estimation of genetic distance

Due to the lack of population-specific bands, a discrimination between cultivars could only be achieved by using the population frequencies of bands.

The initial approach was to use bands that presented the largest difference between two populations This method can be efficient but, as the bands are precisely choosen for their discriminatory power, it is difficult to determine the correct threshold for a test of comparison. Additionally, such a method cannot be used in a global way for comparing all the varieties.

Therefore, in what follows we define the following distance between the populations. The re-sampling method proposed by Excoffier et al. (1992) will be used to test for a null value of this distance. For each band b, $(b = 1, \ldots, B)$, y_{ij}^b denotes the variable value y for individual j $(j = 1, \ldots, N_i)$ belonging to cultivar; this measure is 1 or 0 depending on the presence or the absence of the band b. We assume that the individuals studied represent a random sample of the cultivar. The mean per cultivar value for each band b is $E(y_{ij}^b) = p_i^b$.

The distance between two populations i and i' is defined as in Rogers (1972) by

$$d_{ii'}^2 = \frac{1}{B}\sum_{b=1}^{B}(p_i^b - p_{i'}^b)^2.$$

From now on this "squared" distance will simply be called a distance. The unbiased estimator of $d_{i,i'}^2$ is

$$\hat{d}_{ii'}^2 = \frac{1}{B}\left\{\sum_{b=1}^{B}(\hat{p}_i^b - \hat{p}_{i'}^b)^2 - \sum_{b=1}^{B}\left[\frac{\hat{p}_i^b(1-\hat{p}_i^b)}{N_i - 1} + \frac{\hat{p}_{i'}^b(1-\hat{p}_{i'}^b)}{N_{i'} - 1}\right]\right\},$$

Table 5 Estimated distance (to be multiplied by 10^3) between populations. *ns* denotes an estimated distance not significantly different from 0

Population	Gemini	Euver	Europe	Rival	Medalfa	Oro	Monte-Oscuro
Euver	5.4 *ns*						
Europe	22.4	21.2					
Rival	32.8	38.2	19.2				
Medalfa	35.0	28.6	15.4	32.4			
Oro	26.2	25.6	14.4	34.4	24.6		
Monte-Oscuro	25.6	32.6	32.4	41.2	46.0	21.6	
Maron	51.8	50.6	63.4	77.4	77.4	50.0	41.8

where $\hat{p}_i^b = \sum_{j=1}^{N_i} y_{ij}^b/N_i$ and $\hat{p}_{i'}^b \sum_{j=1}^{N_{i'}} y_{i'j}^b/N_{i'}$. The negative estimates of the distance are set at zero.

To obtain the distribution of this estimator under the null hypothesis $d = 0$, we proceed by re-sampling. The individuals are mixed and associated at random in one or the other cultivar. We thus obtain a value of the estimator $\hat{d}_{ii'}^2$ under the null hypothesis. We repeat this procedure 500 times and obtain the distribution of $\hat{d}_{ii'}^2$ under the null hypothesis. A $\hat{d}_{ii'}^2$ value is then considered different from 0, with a α% type-I error ($\alpha = 5\%$), if $\hat{d}_{ii'}^2$ is greater than the α quantile of the distribution.

The estimation and test for the distance between sets uses the same procedure. Table 5 gives the estimates $\hat{d}_{ii'}^2$, multiplied by 10^3, of the distance between populations. The analysis is done on the total number of bands (B = 64) revealed by the primers. We have used all the bands because, in the genetic distance measurement, each locus, whether polymorphic or not, must be taken into account in order to avoid over-estimation of the distance. It appears that all the distances are significantly different from 0, except for the distance between Gemini and Euver.

Minimal number of bands and individuals required to detect a difference between two varieties

It is important to identify the number of bands and individuals that are necessary to detect a significant difference between two populations. To this end, we need to consider that the observed bands are a sample of a number of possible bands in the genome. One can then calculate the variance of the estimator of the distance $\hat{d}_{ii'}^2$ between two populations, i and i', in order to derive the minimum number of individuals and bands needed to control a fixed value of the type-I and type-II errors. This variance, denoted $v(\hat{d}_{ii'}^2, B, N)$, will depend on the number B of bands, N the number of measured individuals, and on the value $\hat{d}_{ii'}^2$.

Items (1), (2), (3) and (4), below describe the method used to calculate this variance. Several results are obtained using different numbers, N_i and N_i, of measurements for the two populations, but the final formula is given in the simplified situation where $N_i = N_{i'} = \mathrm{N}$. Additional hypotheses are necessary compared to the previous section; in particular, we assume that all the bands are independant of each other.

(1) Let us denote $d_{ii'}^2(b) = (p_i^b - p_{i'}^b)^2$, the distance between the populations i and i' for the band b. Its estimator is

$$\hat{d}_{ii'}^2(b) = (\hat{p}_i^b - \hat{p}_{i'}^b)^2 - \frac{\hat{p}_i^b(1-\hat{p}_i^b)}{N_i - 1} - \frac{\hat{p}_{i'}^b(1-\hat{p}_{i'}^b)}{N_{i'} - 1}.$$

We thus have $E[\hat{d}_{ii'}^2(b)] = d_{ii'}^2(b)$ and the variance of $\hat{d}_{ii'}^2(b)$ will be denoted by $v_{ii'}(b)$.

(2) We consider that there are N_b bands in the genome. One of these bands can be drawn at random with an equi-probability of $\frac{1}{N_b}$. This band is denoted as b_1^*. We therefore have

$$E[(d_{ii'}^2(b_1^*)] = \frac{1}{N_b}\sum_{b}^{N_b} d_{ii'}^2(b) = d_{ii'}^2.$$

The variance $\mathrm{var}[d_{ii'}^2(b_1^*)]$, denoted as $w_{ii'}$, represents the variability of $d_{ii'}^2(b)$ among the bands.

(3) If B bands $\underline{b}^* = (b_1^*, \ldots, b_j^*, \ldots, b_B^*)$ are drawn independantly, each with the same distribution as b_1^*, and conditional to the value taken by the vector $\underline{b}^*$, we get

$$E[\hat{d}_{ii'}^2(\underline{b}^*)/\underline{b}^*] = E\left[\frac{\sum \hat{d}_{ii'}^2(b_j^*)}{B}\Big/\underline{b}^*\right] = d_{ii'}^2(\underline{b}^*) = \frac{\sum d_{ii'}^2(b_j^*)}{B},$$

and supposing that the different bands are in linkage equilibrium $\mathrm{var}[d_{ii'}^2\ (\underline{b}^*)/\underline{b}^*] = \sum_j \frac{v_{ii'}(b_j^*)}{B^2}$.

(4) Globally, when B is sufficiently large, the unconditional distribution of $\hat{d}_{ii'}^2(\underline{b}^*)$ can be approximated by the normal distribution with an expectation $d_{ii'}^2$ and a variance

$$\begin{aligned}\mathrm{var}[\hat{d}_{ii'}^2(\underline{b}^*)] &= E\{\mathrm{var}[\hat{d}_{ii'}^2(\underline{b}^*)/\underline{b}^*]\} \\ &\quad + \mathrm{var}\{E[\hat{d}_{ii'}^2(\underline{b}^*)/\underline{b}^*]\} \\ &= E\left[\frac{\sum_j v_{ii'}(b_j^*)}{B^2}\right] + \mathrm{var}\left[\frac{\sum_j d_{ii'}^2(b_j^*)}{B^2}\right] \\ &= E\left[\frac{\sum_j v_{ii'}(b_j^*)}{B^2}\right] + \frac{w_{ii'}}{B}.\end{aligned}$$

Some simplifications are necessary to obtain a usable formula. For sufficiently large values of N_i and $N_{i'}$, we have $v_{ii'}(b) = 2\sigma^4_{ii'}(b)\left[1 + 2\frac{d^2_{ii'}(b)}{\sigma^2_{ii'}(b)}\right]$ with $\sigma^2_{ii'}(b) = \left[\frac{p^b_i(1-p^b_i)}{N_i} + \frac{p^b_{i'}(1-p^b_{i'})}{N_{i'}}\right]$. To simplify the calculation, we will consider $N_i = N_{i'} = N$ and $p^b_i(1-p^b_i) = 0.25$, is approximately correct given a large interval for p^b_i.

Therefore we obtain $v_{ii'}(b) = \frac{1}{N}\left[\frac{0.5}{N} + 2d^2_{ii'}(b)\right]$ and

$$\mathrm{var}[\hat{d}^2_{ii'}(b^*)] = v(d^2, B, N) = \frac{1}{NB}\left(\frac{0.5}{N} + 2d^2_{ii'}\right) + \frac{w_{ii'}}{B}.$$

For $d^2_{ii'} = 0$, the variance of $\hat{d}^2_{ii'}(\underline{b}^*)$ is $v(0, B, N) = \frac{0.5}{N^2B}$.

We can now deduce a formula to derive the minimum number B of bands and N of individuals that are needed to detect a difference between two populations, with a type-I error of 5%, a power of 90%, and an actual distance of d^2.

The usual test statistic to detect a difference between two varieties is $T = \hat{d}^2/\sqrt{v(0, B, N)}$. This test statistic follows a standard normal distribution under the null hypothesis $d^2 = 0$; the 5% threshold is therefore 1.645. Under the alternative hypothesis for d^2, $\hat{d}^2$, is normally distributed with an expectation d^2 and a variance $v(d^2, B, N)$. We thus have

$$\begin{aligned}
&\Pr_{d^2}(T \geq 1.645) \\
&= \Pr_{d^2}\left(\hat{d}^2 \geq 1.645 \times \sqrt{\frac{0.5}{N^2B}}\right) \\
&= \Pr_{d^2}\left(\frac{\hat{d}^2 - d^2}{\sqrt{v(d^2, B, N)}} \geq \frac{1.645 \times \sqrt{\frac{0.5}{N^2B}} - d^2}{\sqrt{v(d^2, B, N)}}\right) \\
&= 0.90.
\end{aligned}$$

The formula is therefore

$$d^2 = 1.645\sqrt{\frac{0.5}{N^2B}} + 1.28\sqrt{v(d^2, B, N)}.$$

Using the hypothesis of independance between the estimates $\hat{d}^2_{ii'}(b)$ for different bands, it is possible to estimate $w_{ii'}$ with $v[\hat{d}^2_{ii'}(b)]$ the empiric variance between the bands, namely:

$$\hat{w}_{ii'} = [\hat{d}^2_{ii'}(b)] - \frac{1}{2}\left(\frac{1}{N_i} + \frac{1}{N_{i'}}\right) \times \left[\frac{1}{2}\left(\frac{0.5}{N_i} + \frac{0.5}{N_{i'}}\right) + 2\hat{d}^2_{ii'}\right].$$

Figure 1 gives the estimate of $w_{ii'}$ (y axis) and the estimates of $d_{ii'}$ (x axis) for all the populations. From this plot, we define four hypothetical pairs of populations for which we study the necessary number of bands and individuals to detect a difference. For these four

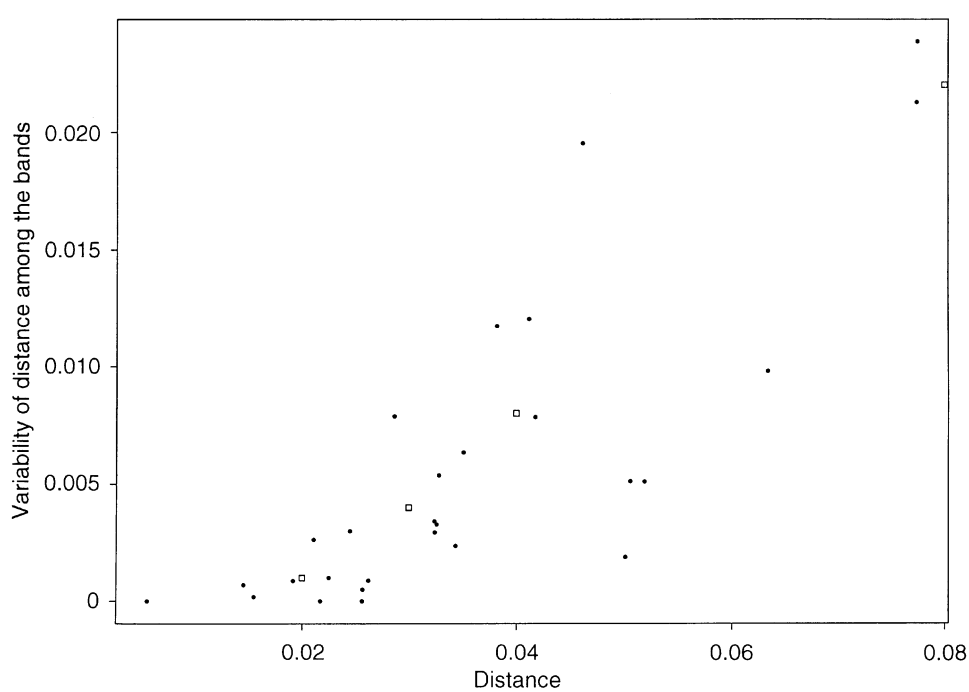

Fig. 1 Plot of the variability among the bands versus distance for all the couples of varieties. *Squared plots* are the hypothetical couples of varieties used in Fig. 2

chosen points, the function that links the minimal number of individuals to the minimal number of bands to detect a difference with a power of 90% and a type-I error of 5% is shown in Fig. 2. For a pair of varieties that is similar to Europe and Rival ($d = 0.01$, $w = 0.0003$) at least 25 bands from 40 individuals are sufficient to detect a difference.

Analysis of commercial sets

We tested if the multiplication of synthetic alfalfa varieties can increase the within-population variability in such a way that seed lots, from the same cultivar, could be discriminated. Twenty individuals from standard cultivars of Europe and Rival, together with two of their commercial seed lots resulting from two different multiplications, were analysed using the same methodology. Table 6 shows that the estimates $\hat{d}^2_{ii'}$ of the distance between a standard set and the two commercial sets for the cultivars Europe and Rival are not significantly different from zero. Moreover, the cultivar standards used in this experiment were not significantly different from the populations analysed above (data not shown).

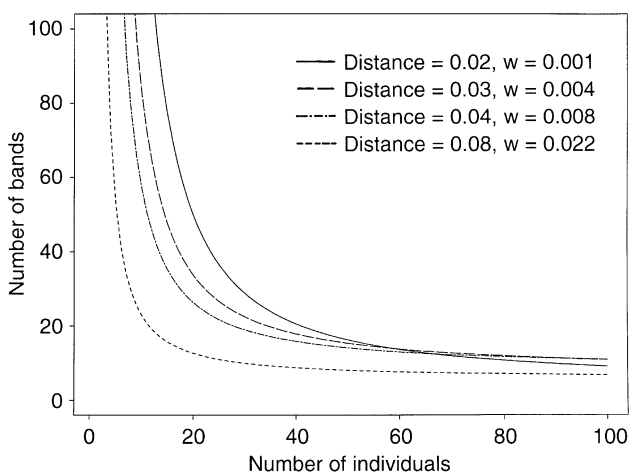

Fig. 2 Link between the number of individuals and the number of bands to detect a difference with a power of 90%

Table 6 Estimated distance (to be multiplied by 10^3) between the standards of Europe and Rival varieties and two samples of their commercial lots (20 individuals). *ns* denotes an estimated distance not significantly different from 0

Commercial lots	Standards	
	Europe	Rival
# 1	0*ns*	0*ns*
# 2	0*ns*	2.66*ns*

Discussion

The primary objective of this study was to find a method to evaluate the genetic distance between alfalfa populations. Alfalfa is an allogamous, polyploid species resulting in highly heterozygous, heterogeneous populations. Thus, much of the genetic diversity should be apportioned among individuals as observed by Kidwell et al. (1994), Quiros and Bauchan, (1988), Brummer et al. (1991) and Crochemore et al. (1996). For this reason, we analysed a number of individuals within each population using the RAPD method which rapidly generates many polymorphic bands, each, *a priori*, representative of a genetic locus.

Our results show clearly the great genetic diversity existing within alfalfa populations. About the 2/3rds of the 64 RAPD bands scored were polymorphic displaying a large range of within-population variation; in none of the eight populations analysed, were we able to find two identical genotypes. Most of the bands were present in all the populations and could be observed with different frequencies varying according to the population. In only one case, a band (B10-600 bp) appeared to be specific to a single population: the *M. falcata* Maron population; however this band was present in only 32% of the individuals of this population. Conversely, the B7-1018 bp and the B8-1050 bp bands were present in all the populations except Medalfa and Rival, respectively. It is noteworthy that these bands are present with a very low frequency in other varieties suggesting that their absence could simply be due to the number of individuals tested.

It is clear from our results that the pairwise comparison of populations allowed us to detect, specific bands that are present in only one of the populations. From this point of view, Maron (*M. falcata*) shows the largest number of discriminating bands. However, when three populations at a time were compared there is a rapid shift-down in the number of population-specific bands. Nevertheless, Table 4 points to some interesting features: at least, 10 bands out of 64 show striking variations of frequency among populations; for instance, the high frequencies of B1-740 bp in Rival and Maron or of B6-1350 bp in Rival only.

The discovery of specific bands could be of interest in order to track the introgression of genetically distant material, such as a *M. falcata* origin, within cultivated alfalfas. However, it is clear that the mixing of the germplasms which occured during the breeding of alfalfa is not only likely to prevent this identification of specific RAPD bands but should be of little value in the discrimination between two cultivars because they would not be representative of their global genome.

As a consequence, it is clear that populations can be distinguished only by an estimation of their genetic distance which must also take into account the diversity existing between individuals within populations.

We propose a mathematical method to define the distance between populations from the polymorphism frequency of individuals. This method is based on the pairwise comparison of the RAPD band frequency of a number of individuals taken at random within populations.

From Table 5, we can see that all except Euver and Gemini can be discriminated when comparing populations pairwise. The Maron population gave the greatest distance with other populations, which is in agreement with the results of Kidwell et al. (1994), who used RFLP to measure the genetic distance between nine *Medicago* accessions, Julier et al. (1995), using phenotypical traits, and Crochemore et al. (1996), with RAPD markers. The close distance between the Monte-Oscurro and Oro cultivars can be readily explained since Oro is a Provence-type of alfalfa, predominantly an *M. sativa*-type, and Monte-Oscurro is a Mielga population originating from spontaneous alfalfa *M. sativa* populations.

It should be noted that due to the tetraploidy of alfalfa and the dominant nature of RAPD markers (Williams et al. 1990), only one allele out of four is sufficient to give rise to an amplification product, that is to a band (Yu and Pauls 1993). Consequently, this method does not allow one to measure the real genetic distance between populations but only to discriminate populations according to their RAPD fingerprints. Therefore, we can ask if these statistically significant estimates of distance are really biologically significant. Nevertheless, this method could be successfully used in the D.U.S. (distinction, unifomity and stability) test to compare new cultivars from those previously registered. For instance, we have compared samples of yearly multiplication sets from the Rival and Europe cultivars. Table 6 shows that this method could also be used as an additional test for the conformity of cultivar sets.

Future experimentation to estimate the distance between two alfalfa cultivars could be done using our study on the minimal number of bands and individuals. The minimization of the cost of experimentation, by reducing the number of bands and the number of individuals, could be resolved using the shape of function of Fig. 2: only 25 bands and 40 individuals are sufficient to detect a difference. However, the genetic hypotheses made to obtain these functions have to be carefully kept in mind, in particular the linkage equilibrium of the bands. Moreover, these functions have been obtained for one estimated distance; if a new variety is to be compared with more than one known variety, the type-1 error of 5% could be too liberal. Nevertheless, this methodology could be used without difficulty with another type-1 error or another *a priori* power. Furthermore , this method can be used directly with other dominant markers, such as AFLP markers, or, with some modifications, with co-dominant markers.

Acknowledgements This study was supported by a grant from the Ministère de l'Agriculture et de la Pêche. We thank C. Baril, B. Julier and J. Ronfort for critical reading of the manuscript. We also thank Dr. Y. Dattée for her constant interest and support.

References

Barnes DK, Goplen BP, Baylor JE (1988) Highlights in the USA and Canada. In: Hanson AA, Barnes DK, and Hill RR, Jr. (eds) Alfalfa and alfalfa improvement. American Society of Agronomy, Madison, Wisconsin. Monograph 29, pp 1–24

Brummer EC, Kochert G, Bouton JH (1991) RFLP variation in diploid and tetraploid alfalfa. Theor Appl Genet 83:89–96

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

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

Hadrys H, Balick M, Schierwater B (1992) Applications of random amplified polymorphic DNA (RAPD) in molecular ecology. Mol Ecol 1:55–63

Julier B, Porcheron A, Ecalle C, Guy P (1995) Genetic variability for morphology, growth and forage yield among perennial diploid and tetraploid lucerne populations (*Medicago sativa* L.) Agronomie 15:1–10

Kidwell KK, Austin DF, Osborn TC (1994) RFLP evaluation of nine *Medicago* accessions representing the original germplasm sources for North American Alfalfa Cultivars. Crop Sci 34:230–236

Quiros CF (1983) Alfalfa (*Medicago sativa* L.). In: Tanksley SD, Orton TJ (eds.) Isozymes in plant genetics and breeding. Part B. Elsevier Scientific Publ. Co., Amsterdam, pp 253–294

Quiros CF, Bauchan GR (1988) The genus *Medicago* and the origin of the *Medicago sativa* complex. In: Hanson AA, et al. (eds) Alfalfa and alfalfa improvement. Agron Monogr 29. ASA, CSSA and SSSA, Madison, Wiscohsin, USA, pp 93–124

Reiter RS, Williams JGK, Feldmann KA, Rafalski JA, Tingey SV, Scolnik PA (1992) Global and local genome mapping in *Arabidopsis thaliana* by using recombinant inbred lines and random amplified polymorphic DNAs. Proc Natl Acad Sci USA 89:1477–1481

Rogers JS (1972) Measures of similarity and genetic distance. Studies in genetics VII. Univ Texas Publ 7213:145–153

Tai TH, Tanksley SD (1990) A rapid and inexpensive method for isolation of total DNA from dehydrated plant tissue. Plant Mol Biol Rep 8:297–303

Truchet G, Ccmut S, Debilly F, Odorico R, Vasse J (1989) The *Rhizobium*-legume symbiosis: two methods to discriminate between nodules and other root-derived structures. Protoplasma 149:82–88

Williams JC, Kubelik AR, Litvak KJ, Rafalski JA, Tingey SV (1990) DNA polymorphism amplified by arbitrary primers are useful as genetic markers. Nucleic Acids Res 18:6531–6535

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