



## A novel statistical method for assessing SSR variation in autotetraploid alfalfa (*Medicago sativa* L.)

Zhi-Peng Liu<sup>1,2,3</sup>, Gong-She Liu<sup>1</sup> and Qing-Chuan Yang<sup>3</sup>

<sup>1</sup>Key Laboratory of Photosynthesis and Environmental Molecular Physiology, Institute of Botany, Chinese Academy of Sciences, Beijing, China.

<sup>2</sup>Graduate School of the Chinese Academy of Sciences, China.

<sup>3</sup>Turf and Forage Science, Institute of Animal Science, Chinese Academy of Agricultural Sciences, Beijing, China.

### Abstract

The level of variation of simple sequence repeat (SSR) markers in cultivated alfalfa from American, Australian and Chinese sources was evaluated using a novel autotetraploid statistical method to calculate the effective number of alleles, the allele frequencies and heterozygosity. We used 19 SSR primers to screen seven polymorphic SSR loci in 320 plants from eight populations. The genetic distance and phylogenetic analysis (DISPAN) program was used to calculate the inter- and intra-population genetic relationships using the conventional binary absence/presence (0/1) method and our novel autotetraploid method. The autotetraploid method resulted in significantly higher heterozygosity ( $p < 0.01$ ), average effective number ( $p < 0.01$ ) and lower standard genetic distance ( $p < 0.01$ ) than the binary method. Our results suggest that our new autotetraploid method is a very useful tool for assessing genetic variation and genetic relationships in all autotetraploid plant species.

**Key words:** autotetraploid, genetic variation, *Medicago sativa* L., SSR, statistical genetics.

Received: July 3, 2006; Accepted: September 29, 2006.

### Introduction

Alfalfa, *Medicago sativa* L., is the most cultivated forage legume due to its ability to fix atmospheric dinitrogen and its high protein content. The analysis of the genetic variability within and among populations of cultivated alfalfa can assess future risk of genetic erosion and help in the development of sustainable conservation and genetic improvement strategies. However, since alfalfa is autotetraploid ( $2n = 4x = 32$ ), allogamous and seed-propagated successful assessment of the genetic diversity of alfalfa has been hampered by the statistical methods available (Stanford 1951, Flajoulot *et al.* 2005).

Compared with other DNA-based markers such as restriction fragment length polymorphisms (RFLP), random amplified polymorphic DNA (RAPD), single nucleotide polymorphisms (SNPs) and amplified fragment length polymorphisms (AFLP), simple sequence repeat (SSR) markers occur frequently in plants and are multiallelic, co-dominant, highly reproducible and can function with

low-quality DNA (Morgante and Olivieri 1993; Wang *et al.* 1994). Many SSR markers have been developed and are widely used in plants for genetic mapping, genetic diversity assessment, population genetics and marker-assisted selection (Gupta and Varshney 2000).

With the rapid development of expressed sequence tags (ESTs), a large number of SSR markers have been developed from the ESTs library of *Medicago truncatula* (Baquerizo *et al.* 2001; Eujayl *et al.* 2004), a model diploid organism with  $2n = 2x = 16$ . Eujayl *et al.* (2004) searched 147,000 *M. truncatula* ESTs and identified 455 SSR primer pairs which produced characteristic SSR bands of the expected length in *Medicago* species. Ellwood *et al.* (2006) used six simple sequence repeat (SSR) loci to analyze the genetic diversity and relationships between randomly selected specimens from 192 accessions in the core *M. truncatula* collection.

It thus seems that SSR markers would be a powerful molecular approach for assessing genetic diversity and germplasm characterization in tetraploid alfalfa. Bernadette *et al.* (2003) used 87 SSR primer pairs, most from *M. truncatula* ESTs, for genotyping and mapping tetraploid alfalfa populations and SSR markers have been applied to alfalfa in several other studies (Diwan *et al.* 1997, 2000;

Send correspondence to Qing-Chuan Yang. Turf and Forage Science, Institute of Animal Science, Chinese Academy of Agricultural Sciences, Beijing 100094, China. E-mail: qchyang66@yahoo.com.cn.

Mengoni, 2000a, b; Baquerizo *et al.* 2001; Eujayl *et al.* 2004; Flajoulot *et al.* 2005; Sledge *et al.* 2005; Ellwood *et al.* 2006). Even so, relatively little is known about the utility of SSR markers in elucidating the genetic relationships within and among populations of cultivated alfalfa because of the autotetraploid inheritance of alfalfa and codominant characteristic of SSR marker. A method for analyzing the segregation patterns of SSR markers in autotetraploid plants has been developed which can predict the genotypes of a pair of autotetraploid specimens at a genetic locus from their gel-pattern phenotypes and the phenotypes of their offspring (Luo *et al.* 2000, 2001; Hackett and Luo 2002). This methodology has greatly facilitated the construction of genetic linkage maps in tetraploid plants and has resulted in the application of SSR methodology to the analysis of genetic diversity within and among tetraploid plant populations.

There are large differences between the genotypes and the phenotypes of autotetraploid plants due to the possibilities of different dosages of alleles and the presence of null alleles (Table 1), which means the failure of the PCR primers to anneal to the DNA template and the consequent lack of the corresponding electrophoretic band (Luo *et al.* 2000). The null allele often occurs in cultivated alfalfa because there is more SSR variation in autotetraploid than diploid plants (Callen *et al.* 1993) and it has also been pointed out that when *M. truncatula* SSR primers are used in alfalfa the number of null alleles should be higher (Bernadette *et al.* 2003).

Jenczewski *et al.* (1999) has pointed out that the frequencies of the null alleles, the amount of self-fertilization within population ( $s$ ) and the proportion of double reduction at each loci ( $\alpha$ ) should be evaluated when describing genetic relationships in autotetraploid alfalfa accessions. The allele frequencies, including the null alleles, shown in Table 1 form the basis of our hypothesis that in natural populations of cultivated alfalfa all genotype frequencies within the same set of banding patterns (gel) are equivalent, assuming that that the populations were mating at random ( $s = 0$ ) and that in autotetraploid alfalfa loci did not experience double reduction ( $\alpha = 0$ ). These hypotheses are supported by the previous work of Jenczewski *et al.* (1999).

Based on allele frequencies, the formulas of heterozygosity ( $H$ ) and the average effective number of alleles ( $N_e$ ) were put forward in our study.

In the present paper we describe the use of SSR markers and a novel autotetraploid statistical method to investigate the genetic variability within and among cultivated alfalfa accessions. This methodology can be used to gain a better understanding of biodiversity conservation and the utilization of genetic resources for use in designing breeding programs for cultivated alfalfa.

## Material and Methods

### Plant materials

All alfalfa (*Medicago sativa* L.) accessions investigated were supplied by the Institute of Animal Science, Chinese Academy of Agricultural Sciences (IAS-CAAS), Beijing. The eight accessions studied were: two native Chinese accessions, Weinan, a landrace from Shaanxi Province, and Zhongmu N. 1, an improved variety at IAS-CAAS; and six accessions Alfanafa, Atlantic, AZ-Germ-Salt, AZ-88NDC, AZ-90NDC-ST and Salado introduced from Australia and America. Except for Weinan, all the accessions were improved varieties. Seeds were germinated and grown in a greenhouse for 30 days before collecting 0.5 g of fresh young leaves from each of 40 randomly selected seedlings per accession.

### DNA preparation and SSR amplification

For each accession, total genomic DNA was extracted from individual seedlings using a mini-extraction kit according to the manufacturer's protocol (BioDev-Tech, Beijing, China) and stored at  $-20^{\circ}\text{C}$  until needed. To determine the optimum polymerase chain reaction (PCR) conditions we carried out trials using the PCR protocol described below but varying the concentration of  $\text{MgCl}_2$ , dNTPs, Taq DNA polymerase and genomic DNA and other parameters, *e.g.* melting point ( $T_m$ ,  $^{\circ}\text{C}$ ), annealing time etc. The number of 19 SSR primer pairs was screened using ten DNA samples from different accessions and on the basis of this preliminary data seven SSR primers (AFca16, AFca1, AFca11, AFct32,

**Table 1** - Relationship between marker banding phenotypes and genotypes at a single locus. Different letters represent different alleles, with 'O' denoting the null allele. The expected allele frequencies were calculated on the assumption that the frequencies of each genotype of the same banding phenotype (*i.e.* number of bands present on the electrophoresis gel after staining) are equivalent in natural populations of cultivated alfalfa. For example, for the 'One band' banding phenotype all corresponding genotypes (AAAA, AAAO, AAOO and AOOO) share the same genotype frequency.

Banding phenotype	Genotype	Expected allele frequencies				
		A	B	C	D	O
One band	AAAA, AAAO, AAOO and AOOO	5/2	0	0	0	3/2
Two bands	AAAB, AABB and ABAB AABO, ABBO) and ABOO	5/3	5/3	0	0	2/3
Three bands	AABC, ABBC, ABCC and ABCO	5/4	5/4	5/4	0	1/4
Four bands	ABCD	1	1	1	1	0

AFct11, MTLEC2A, and B14B03) (Diwan *et al.* 1997; Bernadette *et al.* 2003) were chosen. The PCR amplification was carried out in a 12 µL final volume containing 15 ng of genomic DNA as template, 0.2 mM of each dNTP (Sangon Co., Shanghai, China), 0.25 µM of each primer, 1.0 unit of Taq DNA polymerase (Sangon Co., Shanghai, China) in 1 x PCR reaction buffer (Sangon Co., Shanghai, China). The optimal concentration of the Mg<sup>2+</sup> ions was different in each primer pairs (Table 2). The PCR conditions consisted of an initial denaturation at 95 °C for 5 min, followed by 33 cycles of denaturation at 94 °C for 45 s, annealing at 56 °C (depended on primers, Table 2) for 45 s and extension at 72 °C for 1 min, with a final 72 °C extension for 10 min.

The PCR products (3 µL per lane) were resolved on 8% or 12% (w/v) polyacrylamide gel, using an acrylamide/bisacrylamide ratio of 29:1, and run at 175 V constant voltage for 10 h at 4 °C. The PBR322/*Msp I* marker (Huamei, Beijing, China) was also run on each gel as a molecular weight standard. Bands were stained using a silver-nitrate method (Tixier *et al.* 1997) and scanned and analyzed using the Gel Logic 200 system and the associated 1D Image Analysis Software (Kodak, USA).

**SSR analysis**

The SSR banding phenotypes on the visualized gel were scored by two different protocols, the conventional band absence (0) or presence (1) binary method (Botstein *et al.* 1980; Anderson *et al.* 1993) and our novel tetraploid method.

In the conventional binary 0/1 method, the formulas for the binary allele frequencies (*P*), heterozygosity (*h*) (Nei 1987), effective number of alleles (*Ne*) and the polymorphism information content (*PIC*) (Botstein *et al.* 1980; Anderson *et al.* 1993) were calculated as follows:

$$P_i = \frac{1}{N} \sum_{i=1}^n f_i \quad h = 1 - \sum_{i=1}^n P_i^2$$

where *P<sub>i</sub>* is the frequency of an allele for locus *i*, *f<sub>i</sub>* (0/1) is the presence or absence of a specific band in each plant, *n* is the number of plants in the population and *N* is the total number of locus *i* bands in a population. So, this leads to the equation

$$Ne = \sum_{i=1}^n \frac{Ne \times i}{n} = \frac{\sum_{i=1}^n \left( \frac{1}{\sum_{j=1}^m P_{ij}^2} \right)}{n}$$

$$PIC = 1 - \sum_{i=1}^n P_{ij}^2 - \sum_{i=1}^{n-1} \sum_{j=i+1}^n 2P_i^2 P_j^2$$

where *P<sub>ij</sub>* is the frequency of the *j* allele of locus *i*, *m* is the total number of locus *i*, *Ne* is the effective number of alleles for locus *i* and *n* is the total number of loci.

In our new autotetraploid method the autotetraploid heterozygosity (*h'*) was calculated using the same formula as for the binary heterozygosity (*h*), and the allele frequencies (*p<sub>i</sub>*) and the effective number of alleles (*Ne'*) were calculated as follows:

$$\bar{P}_i = \frac{P'_i}{(P_e + P_0)} \quad P'_i = \frac{5}{2} \sum_{i=1}^n \alpha_i + \frac{5}{3} \sum_{i=1}^n \beta_i + \frac{5}{4} \sum_{i=1}^n \chi_i + \sum_{i=1}^n \epsilon_i$$

$$P_E = \sum_{i=1}^n P_i P_0 = \frac{3}{2} \sum_{i=1}^n \alpha_i + \frac{2}{3} \sum_{i=1}^n \beta_i + \frac{1}{4} \sum_{i=1}^n \chi_i$$

$$Ne = \sum_{j=1}^n \frac{1 / \left( \sum_{i=1}^j \bar{P}_i^4 + 4P_0 \sum_{i=1}^j \bar{P}_i^3 + 6P_0^2 \sum_{i=1}^j \bar{P}_i^2 + 4P_0^3 \sum_{i=1}^j \bar{P}_i + P_0^4 \right)}{n}$$

where *P<sub>i</sub>'* is the expected number of the *j* allele (expect for the null allele) of a locus, *P<sub>E</sub>* is the sum of *N<sub>j</sub>* in locus *i*, *P<sub>0</sub>* is the expected number of null alleles for a locus. The banding

**Table 2** - The seven pairs of alfalfa polymorphic simple sequence repeat (SSR) primers used in this study. Note that the number of alleles does not include the null allele.

Loci	Forward (F) and reverse (R) primers	Core motif	Melting point (T <sub>m</sub> , °C)	Amount of Mg <sup>2+</sup> (µL)	Allele size (bp)	Number of alleles	Linkage group
AFca16	(F) GGTCGAACCAAGCATGT (R) TAAAAAACATTACATGACCTCAAA	(CA) <sub>12</sub>	51	0.90	92-120	6	8
AFca1	(F) CGTATCAATATCGGGCAG (R) TGTTATCAGAGAGAGAAAGCG	(CT) <sub>4</sub> (CA) <sub>10</sub>	60	0.75	147-183	9	4
AFca11	(F) CTTGAGGGAACACTATTGTTGAGT (R) AACGTTTCCCAAAACATACTT	(CA) <sub>11</sub>	56	1.80	147-189	14	6
AFct32	(F) TTTTTGTCCACCTCATTAG (R) TTGGTTAGATTCAAAGGGTTAC	(CT) <sub>14</sub>	56	0.75	110-142	8	3
AFct11	(F) GGACAGAGCAAAGAACAAT (R) TTGTGTGAAAGAATAGGAA	(CT) <sub>12</sub>	56	1.75	192-256	7	7
MTLEC2A	(F) CGGAAAGATTCTTGAATAGATG (R) TGGTTCGCTGTTCTCATG	(AT) <sub>19</sub>	56	1.80	152-220	6	3
B14B03	(F) GCTTGTCTTCTTCAAGCTC (R) ACCTGACTTGTGTTTTATGC	(CA) <sub>9</sub>	54	0.75	152-180	9	5

phenotypes are given by  $\alpha$  = one band,  $\beta$  = two bands,  $\chi$  = three bands and  $\varepsilon$  = four bands.

For both methods, genetic diversity was estimated using Nei's standard genetic distance ( $D$ , Nei 1972) computed using the genetic distance and phylogenetic analysis DISPAN program (Ota 1993) and unweighted pair group-method with arithmetic mean (UPGMA) dendrograms (Sneath and Sokal 1973) were drawn from genetic distance ( $D$ ) matrices using the DISPAN clustering routine.

## Results

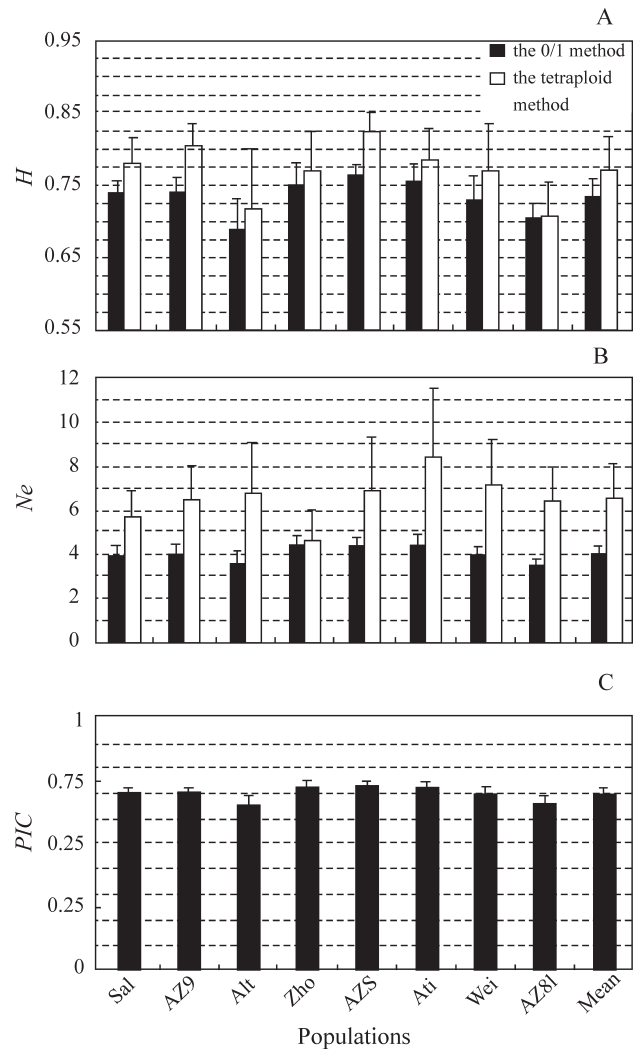
### Genetic diversity within and among accessions

By optimizing the PCR conditions for each primer pair, especially the  $Mg^{2+}$  concentration and the melting temperature ( $T_m$ , °C) (Table 2), we were able to visualize the SSR banding patterns of the accessions, an example being the SSR variation at the AFca16 locus for the Alfanafa accession shown in Figure 1. We used the allele frequencies to calculate the heterozygosity ( $h$ ), the number of effective alleles ( $N_e$ ) and the polymorphism information content ( $PIC$ ) by both the binary method and our novel autotetraploid method and found that both statistical methods indicated that there was substantial variability between the eight accessions studied. In the eight accessions we found 59 bands (alleles) at seven SSR loci, with 27 (45.8%) occurring in all the accessions. The number of bands detected in each accession was 51 in AZ-GermSalt, 49 in Atlantic, 48 in Salado and Zhongmu N. 1, 47 in AZ-90NDC-ST, 46 in Alfanafa, 44 in AZ-88NDC and 41 in Weinan.

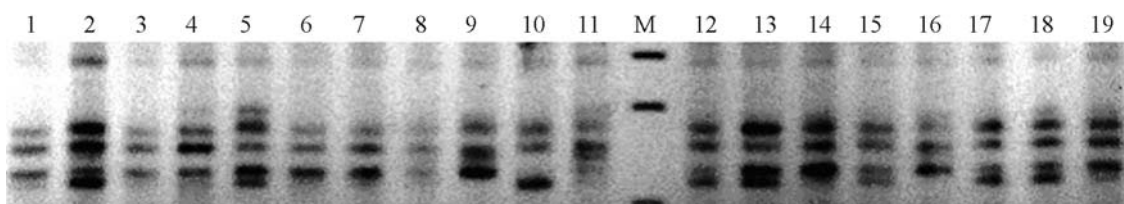
### Comparison of the binary and autotetraploid statistical methods

With the binary 0/1 method, within population SSR variability was high, with the heterozygosity by the binary method ( $h$ ) ranging from 0.6877 for the Alfanafa accession to 0.7640 for the AZ-GermSalt accession (Figure 2A) and the mean effective number of alleles by the binary method ( $N_e$ ) ranging from the 3.4 for Alfanafa and AZ-88NDC to 4.4 for Zhongmu N. 1' (Figure 2B). The most powerful index for evaluating loci polymorphism is the  $PIC$  value, which ranged from 0.6276 for the B14B03 locus to 0.7989 for the AFca11 locus. The  $PIC$  values for the AFca16, AFca11, AFct11, MTLEC2A and B14B03 loci were all higher than 0.6500 and the mean  $PIC$  for all seven loci was

0.6977 (Figure 3). The regression coefficient between  $h$  and  $PIC$  was 0.9917 ( $|t| = 7.697 > t_{0.01}$ ,  $p < 0.01$ ).



**Figure 2** - The effective number of alleles ( $N_e$ ) and heterozygosity ( $h$ ) and their respective standard errors of eight alfalfa accessions as calculated using two statistical methods. Figure A shows the heterozygosity of eight alfalfa accessions by the binary method and our novel autotetraploid statistical method. Figure B shows the number of effective alleles by these two methods. All the values calculated using the autotetraploid method are significantly higher than those using the binary method because the autotetraploid method takes the null allele into account and uses the expected allele frequencies (See Table 1). Figure C shows the polymorphic information content ( $PIC$ ) +SE of eight alfalfa populations as calculated using the binary method. Accessions: Alfanafa (Alf), Atlantic (Atl), AZ-GermSalt (AZS), AZ-88NDC (AZ8), AZ-90NDC-ST (AZ9), Salado (Sal), Weinan (Wei), Zhongmu N. 1 (Zho).



**Figure 1** - Simple sequence repeat variation at the AFca16 locus of some individual Alfanafa plants. The letter 'M' denotes the molecular marker, 147 bp, 110 bp and 90 bp from top to bottom of the lane.



Our novel autotetraploid method showed greater within population SSR variability, with the maximum autotetraploid heterozygosity ( $h'$ ) being 0.8234 for the AZ-GermSalt accession. The mean  $h'$  value for the eight accessions was 0.7707, significantly higher than the mean heterozygosity ( $h = 0.7338$ ,  $p < 0.05$ ) as calculated by the binary method (Figure 2A). The mean effective number of alleles by our autotetraploid method ( $Ne'$ ) was also significantly higher ( $p < 0.01$ ) than the effective number of alleles as calculated by the binary method ( $Ne$ ) (Figure 2B).

Standard genetic distance and dendrogram

The allele frequencies (Table 3) and DISPAN program were used to calculate the genetic distance ( $D$ ) for the results produced by the binary method ( $D$ ) and the novel autotetraploid method ( $D'$ ), and two UPGMA dendrograms constructed (Figure 4). With the binary method the minimum  $D$  value was 0.0173 between the Salado and Zhongmu N. 1 accessions and the maximum  $D$  value was 0.2217 between the Alfanafa and Weinan accessions. These findings were supported by similar findings for the autotetraploid statistical method, which found a minimum  $D'$  value of 0.0119 between the Salado and Zhongmu N. 1

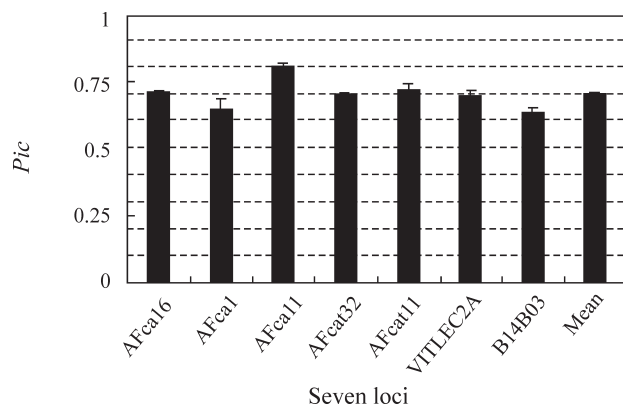


Figure 3 - The polymorphism information content (PIC) and corresponding standard error of seven simple sequence repeat (SSR) loci in eight alfalfa accessions calculated using the conventional binary method. See Figure 2 for the abbreviations.

accessions and a maximum  $D'$  value of 0.1473 between the Alfanafa and Weinan accessions. However, there were significant differences ( $p < 0.01$ ) between the  $D$  and  $D'$  values.

When plotted as a dendrogram the data from the binary analysis clustered the eight alfalfa populations were into two groups (Figures 4A), Weinan and Atlantic being combined into one group and the other six accessions constituting the other group. Data from our novel autotetraploid method produced similar results, except that there were differences in the AZ-GermSalt and AZ-90DNC-ST branches (Figure 4B).

Discussion

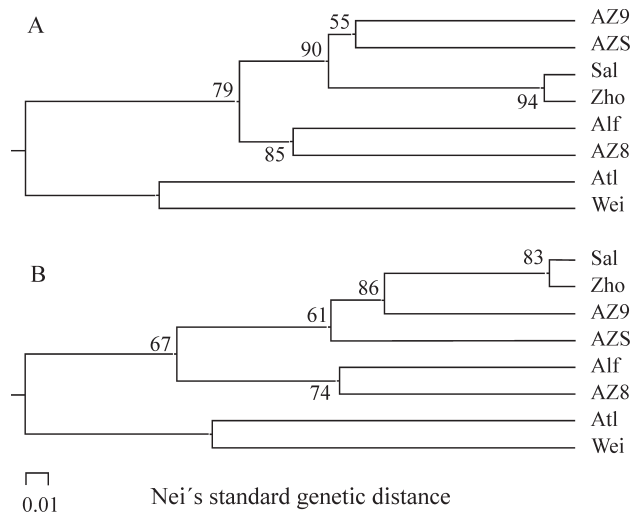
This study is the first attempt to evaluate genetic variation patterns in autotetraploid plant using a new statistical method. The data from the two different statistical methods provide evidence of inter- and intra-populational genetic polymorphisms which could be interpreted as being due to allogamy and the tetraploid nature of alfalfa. The PIC index describes diversity within accessions (intra-populational diversity) and characterizes the degree of polymorphism in each locus, a PIC value of less than 0.25 indicating low polymorphism, a value between 0.25 and 0.5 average polymorphism and a value higher than 0.5 a highly polymorphic locus (Botstein *et al.*, 1980). In our present study we found that seven loci all had PIC values exceeding 0.5 and could be considered highly polymorphic (Figure 3). The highest PIC value was 0.7302 for AZ-GermSalt accession, indicating that this accession had the greatest genetic diversity of all the accessions investigated (Figure 2C). However, all the eight alfalfa accessions studied presented high heterozygosity as measured by both  $h$  and  $h'$ , supporting previous reports of high heterozygosity in alfalfa as detected using RFLP (Brummer *et al.*, 1991) and SSR and RAPD (Mengoni *et al.*, 2000a).

In our study we found that both statistical methods clustered the eight accessions into two groups, which may be related to their salt-tolerance traits because Alfanafa, AZ-GermSalt, AZ-90NDC-ST, RFLPAZ-88NDC, Salado

Table 3 - Standard genetic distances ( $D$ ) between populations calculated using two different statistical methods, the 0/1 method (upper diagonal) and the tetraploid method (lower diagonal).

Accession	Sal	AZ9	Alf	AZS	Atl	Wei	AZ8	Zho
Salado (Sal)		0.0486	0.0839	0.0490	0.1071	0.1480	0.0701	0.0173
AZ-90DNC-ST (AZ9)	0.0277		0.0816	0.0548	0.1065	0.1317	0.0603	0.0726
Alfanafa (Alf)	0.0826	0.0651		0.1048	0.1826	0.2217	0.0713	0.1303
AZ-GermSalt (AZS)	0.0406	0.0333	0.0890		0.1610	0.1601	0.0744	0.0771
Atlantic (Atl)	0.0848	0.0610	0.1278	0.1073		0.1045	0.1138	0.1002
Weinan (Wei)	0.1218	0.0721	0.1473	0.1133	0.0628		0.1258	0.1230
AZ-88NDC (AZ8)	0.0642	0.0342	0.0398	0.0545	0.0787	0.0686		0.0823
Zhongmu N.1 (Zho)	0.0119	0.0361	0.1025	0.0503	0.0688	0.0837	0.0538	

The standard genetic distances computed by the 0/1 method and the tetraploid method are significantly different ( $|d| = 3.018 > t_{0.01}$ ,  $p < 0.01$ ).



**Figure 4** - Unweighted pair group-method with arithmetic mean (UPGMA) dendrogram of eight autotetraploid alfalfa accessions based on Nei's standard genetic distance. Figure A shows the results for the conventional binary method and Figure B those for our novel autotetraploid statistical method. The numbers at nodes are the percent occurrence in 1,000 bootstrap replicates. See Figure 2 for the abbreviations.

and Zhongmu N. 1 clustered together in one group and, except for accession AZ-88NDC, these accessions are salt-tolerant, with AZ-GermSal and AZ-90NDC-ST being the most classic salt-tolerant alfalfa varieties (unpublished data from our laboratory) while Alfanafa, Salado and Zhongmu N. 1 are also salt-tolerant (Yang and Sun 2001). A separate cluster was formed by the Atlantic and Weinan accessions, known to carry the salt-sensitive trait (Yang and Sun 2001). The differences between the conventional binary and our novel autotetraploid dendrograms (Figures 4a and 4b) are probably due to the different statistical theory on which the two methods are based.

The usefulness of nuclear SSR in the analysis of alfalfa germplasm has been reported previously (Mengoni *et al.*, 2000a) but our novel autotetraploid method allows a clearer insight into the genetic variability of cultivated tetraploid alfalfa using SSR markers because when calculating the allele frequencies our method considered not only the differences between marker phenotypes and genotypes but also the frequency of null alleles. Compared with the binary analysis, our new statistical method not only produced a significantly higher number of effective alleles and heterozygosity values and lower standard genetic distances but also a different UPGMA dendrogram. Because the autotetraploid method emphasizes the complex mechanisms of autotetraploid population genetics and codominant genetic SSR markers the results of this type of analysis should be closer to true genetic characteristics of autotetraploid alfalfa.

In our tetraploid method, we did not present a new formula for calculating the *PIC* because this is very complex for autotetraploid plants. In the traditional binary method, we found that the correlation coefficient between

the heterozygosity (*h*) and *PIC* was 0.9917 ( $p < 0.01$ ), the similar correlation also being reported by Qu *et al.* (2004), and that the populations showing the maximum and the minimum *h* (Figure 2B) and *PIC* (Figure 2C) values were AZ-GermSalt ( $h = 0.7640$ ,  $PIC = 0.7302$ ) and Alfanafa ( $h = 0.6877$ ,  $PIC = 0.6474$ ), respectively. This suggests that the autotetraploid *PIC* value (*PIC'*) could be estimated from the autotetraploid *h'* value and used to evaluate the genetic diversity within an autotetraploid population.

Polyploidy is very prevalent in plants, with at least 50% and perhaps up to 95% of angiosperms having experienced one or more episodes of chromosome doubling during their evolutionary history (Leitch and Bennett 1997). A large proportion of cultivated crops are allopolyploids, prominent among them being wheat, cotton, tobacco and many of the forage grasses. Autotetraploids are less common, but include some important plants such as alfalfa, birdsfoot trefoil, potato, tea and rose. During the work described in this paper we employed cultivated alfalfa as a model plant to establish a novel autotetraploid statistical method for studying the inter- and intra-population genetic relationship between autotetraploid species. The results of our study indicate that the tetraploid statistical method will be useful in a range of autotetraploid species for gene localization, construction of linkage maps and the planning of breeding programs.

## Acknowledgements

The authors are grateful to Professor P.Z. Guo (Capital Normal University, Beijing, China) for useful advice on analyzing the data and the Chinese Academy of Sciences Foundation for grant number NK10.5-A-05.

## References

- Anderson JA, Churchill GA, Sutcliffe JE, Tanksley SD and Sorrells ME (1993) Optimizing parental selection for genetic linkage maps. *Genome* 36:181-186.
- Baquerizo EB, Desplanque B, Prosperi JM and Santoni S (2001) Characterization of microsatellite loci in the diploid legume *Medicago truncatula* (barrel medic). *Mol Ecol Notes* 1:1-3.
- Bernadette J, Sandrine V, Philippe B, Gaele C, Sylvain S, Thierry H and Christian H (2003) Construction of two genetic linkage maps in cultivated tetraploid alfalfa (*Medicago sativa*) using microsatellite and AFLP markers. *BMC Plant Biol* 3:1-19.
- Botstein D, White RL, Skolnick M and Davis RW (1980) Construction of genetic linkage map in man using restriction fragment length polymorphisms. *Am J Hum Genet* 32:314-331.
- Callen DF, Thompson AD, Shen Y, Phillips HA, Richards RI, Mulley JC and Sutherland GR (1993) Incidence and origin of "null" alleles in the (AC)<sub>n</sub> microsatellite markers. *Am J Hum Genet* 52:922-927.
- Diwan N, Bhagwat AA, Baughan GR and Cregan PB (1997) Simple sequence repeat (SSR) DNA markers in alfalfa and perennial and annual *Medicago* species. *Genome* 40:887-895.

- Diwan N, Bouton JH, Kochert G and Cregan PB (2000) Mapping of simple sequence repeat (SSR) DNA markers in diploid and tetraploid alfalfa. *Theor Appl Genet* 101:165-172.
- Ellwood SR, Souza NKD, Kamphuis LG, Burgess TI, Nair RM and Oliver RP (2006) SSR analysis of the *Medicago truncatula* SARDI core collection reveals substantial diversity and unusual genotype dispersal throughout the Mediterranean basin. *Theor Appl Genet* 112:977-983.
- Eujayl I, Sledge MK, Wang L, May GD, Chekhovskiy K, Zwonitzer JC and Mian MA (2004) *Medicago truncatula* EST-SSR reveal cross-species genetic markers for *Medicago* spp. *Theor Appl Genet* 108:414-422.
- Flajoulot S, Ronfort J, Baudouin P, Barre P, Huguet T, Huyghe C and Julier B (2005) Genetic diversity among alfalfa (*Medicago sativa*) cultivars coming from a breeding program, using SSR markers. *Theor Appl Genet* 111:1420-1429.
- Gupta PK and Varshney RK (2000) The development and use of microsatellite markers for genetic analysis and plant breeding with emphasis on bread wheat. *Euphytica* 113:163-185.
- Hackett CA and Luo ZW (2002) TetraploidMap, software suite for calculating calculating linkage maps for autotetraploid populations. <ftp://ftp.bioss.sari.ac.uk/pub/cah/>.
- Jenczewski E, Prospero JM and Ronfort J (1999) 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* 8:1317-1330.
- Leitch IJ and Bennett MD (1997) Polyploidy in angiosperms. *Trends Plant Sci* 2:470-476.
- Luo ZW, Hackett CA, Bradshaw JE, McNicol JW and Milbourne D (2000) Predicting parental genotypes and gene segregation for tetrasomic inheritance. *Theor Appl Genet* 100:1067-1073.
- Luo ZW, Hackett CA, Bradshaw JE, McNicol JW and Milbourne D (2001) Construction of a genetic linkage map in tetraploid species using molecular markers. *Genetics* 157:1369-1385.
- Mengoni A, Gori A and Bazzicalupo (2000a) Use of RAPD and microsatellite (SSR) variation to assess genetic relationships among populations of tetraploid alfalfa, *Medicago sativa*. *Plant Breeding* 119:311-317.
- Mengoni A, Ruggini C, Vendramin G and Bazzicalupo (2000b) Chloroplast microsatellite variation in tetraploid alfalfa. *Plant Breeding* 119:509-512.
- Morgante M and Olivieri AM (1993) PCR-amplified microsatellites as markers in plant genetics. *Plant J* 3:175-182.
- Nei M (1972) Genetic distance between populations. *Amer Nat* 106:283-292.
- Nei M (1987) *Molecular Evolutionary Genetics*. Columbia University Press, New York, 512 pp.
- Ota T (1993) *Dispan: Genetic Distance and Phylogenetic Analysis*. Pennsylvania State University, University Park.
- Qu LJ, Wu GQ, Li XY and Yang N (2004) Conservation efficiency of local chicken breeds in different farms as revealed by microsatellite markers. *Acta Genetica Sinica* 31:591-595.
- Sledge MK, Ray IM and Jiang G (2005) An expressed sequence tag SSR map of tetraploid alfalfa (*Medicago sativa* L.). *Theor Appl Genet* 111:980-992.
- Sneath PH and Sokal RR (1973) *Numerical Taxonomy*. Freeman, San Francisco, 573 pp.
- Stanford EH (1951) Tetrasomic inheritance in alfalfa. *Agron J* 43:222-225.
- Tixier MH, Sourdilie RM, Leroy P and Bernard M (1997) Detection of wheat microsatellite using a non radioactive silver-nitrate staining method. *J Genet Breed* 51:175-177.
- Wang Z, Weber JL, Zhang G and Tanksley SD (1994) Survey of plant short tandem DNA repeats. *Theor Appl Genet* 88:1-6.
- Yang QC and Sun J (2001) Studies on RAPD polymorphic in salt-tolerant alfalfa and salt-sensitive alfalfa. *Acta Agrestia Sinica* 9:83-86.

Associate Editor: Márcio de Castro Silva Filho