

# Simple sequence repeat DNA markers in alfalfa and perennial and annual *Medicago* species

Noa Diwan, Arvind A. Bhagwat, Gary B. Bauchan, and Perry B. Cregan

**Abstract:** Simple sequence repeat (SSR) or microsatellite DNA markers have been shown to function well in plant and mammalian species for genetic map construction and genotype identification. The objectives of the work reported here were to search GenBank for the presence of SSR-containing sequences from the genus *Medicago*, to assess the presence and frequency of SSR DNA in the alfalfa (*Medicago sativa* (L.) L. & L.) genome, and to examine the function of selected markers in a spectrum of perennial and annual *Medicago* species. The screening of an alfalfa genomic DNA library and sequencing of clones putatively containing SSRs indicated approximately 19 000 (AT)<sub>n</sub> + (CT)<sub>n</sub> + (CA)<sub>n</sub> + (ATT)<sub>n</sub> SSRs in the tetraploid genome. Inheritance was consistent with Mendelian expectations at four selected SSR loci with different core motifs. Additionally, genotypes of a range of *Medicago* species, including 10 perennial subspecies of the *M. sativa* complex and other perennial and annual *Medicago* species, were analyzed at each of the loci to ascertain the presence, number, and size of SSR alleles at each locus in each genotype. These studies indicate that SSR markers can function in alfalfa for the construction of genetic maps and will also be useful in a range of *Medicago* species for purposes of assessing genetic relatedness and taxonomic relationships, and for genotype identification.

**Key words:** microsatellites, SSR markers, simple sequence repeats, alfalfa, annual medics.

**Résumé :** Les marqueurs de type microsatellite se sont avérés très utiles chez les espèces végétales et les mammifères pour la construction de cartes génétiques et l'identification génotypique. Les objectifs de ce travail étaient d'examiner GenBank pour la présence de séquences contenant des microsatellites chez les espèces du genre *Medicago*, de déterminer la présence et la fréquence de microsatellites dans le génome de la luzerne (*Medicago sativa* (L.) L. & L.) et d'examiner l'utilité de certains marqueurs chez une gamme d'espèces, tant annuelles que pérennes, du genre *Medicago*. Le criblage d'une banque génomique de la luzerne et le séquençage de clones contenant potentiellement des microsatellites ont indiqué qu'environ 19 000 microsatellites (AT)<sub>n</sub> + (CT)<sub>n</sub> + (CA)<sub>n</sub> + (ATT)<sub>n</sub> sont présents dans le génome tétraploïde. L'hérédité de quatre microsatellites choisis présentant des motifs répétés distincts s'est avérée mendélienne. De plus, les génotypes de toute une gamme d'espèces du genre *Medicago* comprenant 10 sous-espèces pérennes du complexe *M. sativa* et d'autres espèces pérennes et annuelles ont été analysés à chacun des loci pour déterminer la présence, le nombre et la taille des allèles à chaque locus dans chaque génotype. Ces études montrent que les microsatellites peuvent être employés chez la luzerne pour la construction de cartes génétiques et qu'ils seront utiles également chez toute une gamme d'espèces du genre *Medicago* pour des travaux d'identification génotypique ainsi que pour des études de parenté génétique et de relations phylogénétiques.

**Mots clés :** microsatellites, marqueurs SSR, répétition simple de séquences, la luzerne, medics annuelles.

[Traduit par la Rédaction]

## Introduction

Alfalfa (*Medicago sativa* (L.) L. & L.), the most important forage legume in North America (Barnes et al. 1988), belongs to

the genus *Medicago*, which includes both perennial and annual species (Lesins and Lesins 1979). Most of the perennial species of *Medicago*, including alfalfa, are tetraploid  $2n = 4x = 32$  outcrossing species; some are diploid ( $2n = 2x = 16$ ); and two, *Medicago cancellata* M.B. and *Medicago saxatillis* M.B., are hexaploid ( $2n = 6x = 48$ ) (Quiros and Bauchan 1988). The majority of the annual *Medicago* species ( $2n = 2x = 16$ ), with the exception of *Medicago constricta* Dur., *Medicago murex* Willd., *Medicago polymorpha* L., *Medicago praecox* DC, and *Medicago rigidula* Desr., which are  $2n = 2x = 14$ , are inbreeding species. Bauchan and Elgin (1984) determined that *Medicago scutellata* Mill. and *Medicago rugosa* Desr. are the only tetraploid annual species with  $2n = 4x = 30$ . Some of the annual *Medicago* species, also known as medics, such as *M. scutellata*, *M. polymorpha* L., *M. rugosa*, and *Medicago truncatula* Gaertn., are widely grown in Australia for forage and as green manure (Crawford et al. 1989).

Corresponding Editor: J.P. Gustafson.

Received March 12, 1997. Accepted September 2, 1997.

N. Diwan,<sup>1</sup> A.A. Bhagwat, G.B. Bauchan, and P.B. Cregan.<sup>2</sup>  
Soybean and Alfalfa Research Laboratory, Building 011, HH-19,  
United States Department of Agriculture, Agricultural Research  
Service, Beltsville, MD 20705, U.S.A.

<sup>1</sup> Present address: The National Plant Genome Center for Israel,  
Department of Plant Genetics, Weizmann Institute of Science,  
Rehovot 76100, Israel.

<sup>2</sup> Author to whom all correspondence should be addressed:  
(e-mail: pcregan@gig.usda.gov).

**Table 1.** Sequences of SSR (microsatellite) alfalfa PCR primer pairs; all primer sequences are given 5' to 3'.

Locus	Core motif	5'-End primer (sense)	3'-End primer (antisense)
AFca11	(CA) <sub>11</sub>	cttgagggaactattgttgagt	aacgtttcccaaacatactt
AFct32	(CT) <sub>14</sub>	ttttgtcccacctcattag	ttggtagattcaagggttac
AFct1	(CTT) <sub>9</sub> (CAA) <sub>3</sub>	cccatcaacaacatttca	ttgtgattggaacgagt
MTLEC2A	(AT) <sub>19</sub>	cggaaagattcttgaatagatg	tggttcgctgttctcatg

Simple sequence repeat (SSR) or microsatellite DNA markers are PCR-based multiallelic codominant genetic markers. SSRs consist of 2–5 nucleotide core units, such as (AT), (CTT), and (ATGT), that are tandemly repeated. The regions flanking the SSR are generally conserved among genotypes of the same species. PCR primers to the flanking regions are used to amplify the SSR-containing DNA fragment. Length polymorphism is created when PCR products from different alleles vary in length as a result of variation in the number of repeat units in the SSR. This length polymorphism can be visualized via electrophoresis of the PCR products (Weber and May 1989; Litt and Luty 1989; Tautz 1989). In plant species, SSR markers are being used to construct linkage maps in *Arabidopsis thaliana* (L.) Heynh. (Bell and Ecker 1994), soybean (*Glycine max* (L.) Merr.; Akkaya et al. 1995), and maize (*Zea mays* L.; Senior et al. 1996). SSRs are also being used to study allelic profiles of genotypes for purposes of genotype identification in crop species; a few examples are soybean (Akkaya et al. 1992; Rongwen et al. 1995; Diwan and Cregan 1997), maize (*Zea mays* L.; Senior and Heun 1993), sunflower (*Helianthus annuus* L.; Brunel 1994), avocado (*Persea americana*; Lavi et al. 1995), and rapeseed (*Brassica napus* L.; Kresovich et al. 1995). SSRs are reported to exhibit high levels of length polymorphism, with as many as 26 alleles at loci in soybean (Cregan et al. 1994a; Rongwen et al. 1995). This level of polymorphism is much higher than that found with RFLP (restriction fragment length polymorphism) markers using the same set of genotypes (Diwan and Cregan 1997).

Several studies have been conducted using molecular markers to assess the level of variation among perennial *Medicago* species and populations (Brummer et al. 1991; Echt et al. 1992; Yu and Pauls 1993; Kidwell et al. 1994), and one study used RAPD (random amplified polymorphic DNA) markers to study variation among annual species (Brummer et al. 1995). None of these studies utilized SSR markers.

The objectives of the work reported here were to assess the presence and frequency of SSR DNA sequences in the alfalfa genome and to study SSR markers with various core motifs across the genus *Medicago*.

## Materials and methods

### Development of SSR markers

#### *Selection of SSR-containing sequences from GenBank*

All available DNA sequence data pertaining to *Medicago* species available in GenBank were searched for the presence of all possible di- and tri-nucleotide SSRs. Only microsatellites with a minimum length of 20 bp were selected.

#### *Selection of SSR-containing clones from a genomic alfalfa*

##### *DNA library*

Alfalfa DNA from bulked plants of population W10 AC3 (Elgin and Ostazeski 1984) was digested with the restriction enzyme *Sau3A*, and fragments of 400–600 bp in length were cloned into pBluescript II KS+ (Stratagene, LaJolla, Calif.). The procedure of Cregan et al. (1994b) was followed. The ligation products were used to transform *Escherichia coli* XL1-Blue, followed by blue–white color selection. The library was screened by colony hybridization (Sambrook et al. 1989). Filters were hybridized to <sup>32</sup>P-labeled oligonucleotide probes (oligo-GA, oligo-AT, oligo-CA, and oligo-ATT), which were labeled using a reaction mixture containing an oligonucleotide template, an appropriate octamer primer, Klenow polymerase, [ $\alpha$ -<sup>32</sup>P]dATP, and dNTPs. Membranes were washed for 2 h at temperatures ranging from 37 to 50°C (depending on the specific probe) for 45 min in 2× SSC (1× SSC: 0.15 M NaCl plus 0.015 M sodium citrate), before being exposed to x-ray film. Positive clones identified in the first screening were purified and screened again with the appropriate SSR-containing oligonucleotide probe to verify the presence of the microsatellite.

To verify the presence of SSRs in positive clones and to allow the design of PCR primers to the regions flanking the SSR, the sequence of clones producing positive hybridization signals was determined. Plasmid DNA of sequencing quality was obtained using the QIAwell-8 Plasmid Kit (Qiagen Inc., Chatsworth, CA 91311). Sequencing was performed using a Perkin Elmer Prism 373A DNA Sequencer. Sequencing was from both KS and SK primer sites of pBluescript, using the Taq DyeDeoxy Terminator Cycle Sequencing Kit (Perkin Elmer, Applied Biosystems, Foster City, CA 94404).

#### *Selection and preparation of PCR primers to microsatellite loci*

Primers were selected using Designer PCR (Research Genetics, Huntsville, AL 35801). In all cases, the  $T_m$  of primers was 55.0 ± 1.0°C. The selection of nearly identical  $T_m$  values for all primers allowed the use of a standard set of PCR amplification conditions. Primers (Table 1) were synthesized on a 391 DNA Synthesizer PCR-MATE (Perkin Elmer, Applied Biosystems).

### Alfalfa genetic material and DNA isolation

#### *Diploid alfalfa cross*

A cross was made between a CADL (cultivated alfalfa at the diploid level; Bingham and McCoy 1979) plant (W2xiso) and a *M. sativa* ssp. *caerulea* Schmalh. plant (PI 440501) to produce an F<sub>1</sub> plant. The F<sub>1</sub> was selfed, producing the mapping population described by Brummer et al. (1993). The two parent plants and the F<sub>1</sub> plant were included in this study.

#### *Tetraploid alfalfa cross*

Two alfalfa plants, MN White-MA (MA-4), a plant with white flower color, and A-620, a plant with purple flower color, were crossed by Dr. Joseph H. Bouton (University of Georgia, Athens, Ga.) to

**Table 2.** Perennial and annual *Medicago* species analyzed using primers to SSR loci AFct32, AFctt1, AFca11, and MTLEC2A.

<b>Accessions representing the <i>M. sativa</i> complex</b>	
Sect. <i>Falcago</i> subsect. <i>Falcatae</i>	<i>M. sativa</i> ssp. <i>falcata</i> UGA108602 (2x) <i>M. sativa</i> ssp. <i>falcata</i> , Very Fall Dormant germplasm source (4x), PI 536532 <i>M. sativa</i> ssp. <i>caerulea</i> UAG214 (2x) <i>M. sativa</i> ssp. <i>glutinosa</i> UAG85 (4x) <i>M. sativa</i> ssp. <i>sativa</i> (4x), from Turkistan, PI 536537 <i>M. sativa</i> ssp. <i>sativa</i> (4x), from India, PI 536536 <i>M. sativa</i> ssp. <i>sativa</i> (4x), from Peru, PI 536535 <i>M. hemicycla</i> PI 315481 (2x) <i>M. glomerata</i> UAG1529 (2x) <i>M. prostrata</i> UAG465 (2x)
<b>Accessions from other perennial <i>Medicago</i> species</b>	
Sect. <i>Falcago</i> subsect. <i>Rupestres</i>	<i>M. cancellata</i> UAG43 (6x)
Sect. <i>Falcago</i> subsect. <i>Daghnestanicae</i>	<i>M. pironae</i> Vis. UAG1971 (4x)
Sect. <i>Falcago</i> subsect. <i>Papillosae</i>	<i>M. papillosa</i> Boiss. UAG96 (4n)
Sect. <i>Arborea</i>	<i>M. arborea</i> PI 307390
Sect. <i>Marinae</i>	<i>M. marina</i> UAG1209
Sect. <i>Suffruticosae</i>	<i>M. suffruticosa</i> Raymond UAG1544
Subgenus <i>Orbicularia</i> sect. <i>Carstiensis</i>	<i>M. carstiensis</i> Wulf. PI 212936
<b>Accessions from annual <i>Medicago</i> species</b>	
Subgenus <i>Lupularia</i>	<i>M. lupulina</i> cv. George
Subgenus <i>Orbicularis</i> sect. <i>Orbiculares</i>	<i>M. orbicularis</i> PI 459206
Subgenus <i>Spirocarpos</i> sect. <i>Rotatae</i>	<i>M. rugosa</i> cv. Paraponto <i>M. scutellata</i> cv. Sava
Subgenus <i>Spirocarpos</i> sect. <i>Pachyspirae</i>	<i>M. littoralis</i> cv. Harbinger AR <i>M. truncatula</i> cv. Caliph
Subgenus <i>Spirocarpos</i> sect. <i>Leptospirae</i>	<i>M. minima</i> PI 499002
Subgenus <i>Spirocarpos</i> sect. <i>Intertextae</i>	<i>M. polymorpha</i> cv. Santiago <i>M. intertexta</i> PI 498832

**Note:** UGA, University of Georgia; UAG designates genotypes from the Karl Lesins collection, Genetics Department, University of Alberta, Edmonton, Alberta, Canada.

produce an F<sub>1</sub> plant. MA-4 is a cytoplasmic male sterile maintainer selected from Turkistan germplasm and was developed by a cooperative breeding program between the United States Department of Agriculture, Agricultural Research Service (USDA-ARS), St. Paul, Minnesota and The Institute of Plant Breeding and Acclimatization, Radzikow, Poland. MA-4 was obtained from Dr. Donald Barnes (USDA-ARS and Department of Agronomy and Plant Genetics, University of Minnesota, St. Paul, Minn.) A-620 is a selection from the cultivar Apollo, and was obtained from Sue Croughan, Louisiana State University, Rice Research Station, Crowley, La. The two parent plants and the F<sub>1</sub> plant were included in this study.

#### *Perennial Medicago species of the M. sativa complex*

Accessions were chosen to represent all sections and subsections of the subgenus *Medicago*, including all species in Section *Falcago* subsection *Falcatae* that includes the *M. sativa* complex (Quiros and Bauchan 1988). The *M. sativa* complex includes both diploid and tetraploid species and subspecies. *Medicago sativa* ssp. *sativa* (tetraploid), *M. sativa* ssp. *falcata* Arcengeli (both diploid and tetraploid forms exist), *M. sativa* ssp. *caerulea* (diploid), and *M. sativa* ssp. *glutinosa* M.B. (tetraploid) are all included in the complex. These taxa are all interfertile, share the same karyotype, and have been given the taxonomic rank of species by some authors (Lesins and

Lesins 1979) and of subspecies by others (Gunn et al. 1978). Accessions representing three other species in the *M. sativa* complex were also chosen for the study (Table 2). Seeds were obtained from the Western Regional Plant Introduction Station, Pullman, Wash., U.S.A.

#### *Other perennial Medicago species*

Additional perennial *Medicago* species were selected to represent the *Falcago* subsections *Rupestres*, *Daghnestanicae*, and *Papillosae*, as well as the sections *Arborea*, *Marinae*, and *Suffruticosae* and the subgenus *Orbicularia* section *Carstiensis* (Table 2). Seeds of the Plant Introductions (PI) were obtained from the Western Regional Plant Introduction Station, Pullman, Wash., U.S.A. Accessions with a UAG designation were obtained from the Karl Lesins Collection, University of Alberta, Edmonton, Alberta, Canada.

#### *Annual Medicago species*

Accessions of agronomically important annual medics were chosen to represent the three subgenera of the genus *Medicago* that include annual species: *Lupularia*, *Orbicularia*, and *Spirocarpos* (Quiros and Bauchan 1988). Seeds of the cultivars George, Paraponto, Sava, Harbinger AR, Caliph, and Santiago were obtained from the Peterson Seed Co., Savage, Minn., U.S.A.

**Table 3.** SSR DNA in the alfalfa genomic library based upon sequence analysis.

Probe	No. of clones containing SSRs <sup>a</sup>	Frequency per Mbp <sup>b</sup>
poly(AT/TA)	17	5.4
poly(CA/GT)	10	3.2
poly(CT/GA)	15	4.7
poly(ATT/TAA)	10	3.2
Compound repeats	8	2.5
Total	60	19.0

<sup>a</sup>Includes only clones with (AT)<sub>n</sub>, (CA)<sub>n</sub>, or (CT)<sub>n</sub> core motifs containing 10 or more repeat units, and (ATT)<sub>n</sub>, or compound dinucleotide-trinucleotide core motifs with 7 or more repeat units.

<sup>b</sup>Library size = 3.17 Mbp.

#### DNA isolation

DNA was isolated from the diploid and tetraploid parents and their F<sub>1</sub>s and from single plants of the accessions indicated in Table 2, using a standard CTAB (hexadecyltrimethylammonium bromide) procedure (Brummer et al. 1991).

#### PCR conditions and PCR product separation

The forward primers to SSR loci Afct32, Afca11, and Afctt1 were 5'-end labeled using 10 pmoles of the 5'-end primer, 5 µL 10× T4 polynucleotide kinase buffer (containing 0.5 M Tris-HCl (pH 7.6), 100 mM MgCl<sub>2</sub>, and 100 mM 2-mercaptoethanol (United States Biochemical, Cleveland, Ohio)), 20 pmole of [<sup>32</sup>P]ATP at 3000 Ci/mmol (1 Ci = 37 GBq), and 3 U of T4 polynucleotide kinase (United States Biochemical, Cleveland, Ohio) in a total volume of 50 µL.

The Afct32, Afca11, and Afctt1 reaction mixtures contained 30 ng of genomic DNA, 1.5 mM Mg<sup>2+</sup>, 0.15 µM of the end-labeled forward and reverse primers, 100 µM of each nucleotide, 1× PCR buffer (containing 50 mM KCl plus 10 mM Tris-HCl (pH 9.0)), 0.1% Triton X-100, and 1 U Taq DNA polymerase in a total volume of 10 µL. Reaction mixtures for the MTLEC2A locus contained 30 ng of genomic DNA, 1.5 mM Mg<sup>2+</sup>, 0.15 µM of forward and reverse primers, 100 µM of each nucleotide, 0.1 µL of 3000 Ci/mmol [<sup>32</sup>P]dATP (10 FCi/µL), 1× PCR buffer (containing 50 mM KCl plus 10 mM Tris-HCl (pH 9.0)), 0.1% Triton X-100, and 1 U Taq DNA polymerase in a total volume of 10 µL. Cycling consisted of 30 s denaturation at 94°C, 25 s annealing at 51 or 52°C, and 25 s extension at 68°C for 32 cycles on a MJ Research model PTC-100 thermocycler (MJ Research, Inc., Watertown, Mass.). PCR products (3 µL/lane) were separated on a DNA sequencing gel containing 6% polyacrylamide, 30% formamide, 5.6 M urea, and 1× TBE (90 mM Tris-borate plus 2 mM EDTA, pH 8.0) at 60 W constant power for 2–3 h.

## Results and discussion

#### Frequency of SSRs in the alfalfa genome

The frequency of SSRs in the alfalfa genomic library is given in Table 3. SSR DNA is abundant in the alfalfa genome, with an estimated 19 000 (AT)<sub>n</sub> + (CT)<sub>n</sub> + (CA)<sub>n</sub> + (ATT)<sub>n</sub> repeats in the tetraploid genome. Therefore, it appears likely that SSR markers can be used to construct high density linkage maps of alfalfa. The most frequent dinucleotide repeat in alfalfa is AT/TA followed by CT/GA. These results are in agreement with frequencies of SSR DNA in other plants surveyed using

GenBank sequences (Morgante and Olivieri 1993; Lagercrantz et al. 1993; Wang et al. 1994). CA/GT repeats, which are the most common repeat in mammalian genomes, were abundant in alfalfa relative to other plant genomes, where these repeats are relatively rare.

#### Medicago SSRs in GenBank

A search for all possible di- and tri-nucleotide repeats in sequences from all *Medicago* species available in GenBank produced three SSR loci: MTLEC2A, the lec2 gene for lectin from *Medicago truncatula* with a (AT)<sub>20</sub> core (Bauchrowitz et al. 1992); ALFLEGHEMA, the alfalfa leghemoglobin gene from *M. sativa* with a (TAA)<sub>11</sub> core (Davidowitz et al. 1990; unpublished results); and ALFUBIQUIT, *M. sativa* ubiquitin carrier protein mRNA with a (CT)<sub>3</sub>CC(CT)<sub>8</sub> core sequence (Pramanik and Bewley 1993; unpublished results). PCR primers were selected for these three sequences. However, only the MTLEC2A locus gave products in the predicted product length range and was used in the study.

#### SSRs markers in the tetraploid and diploid alfalfa crosses

Observation of the alleles present in the diploid and tetraploid parents and their F<sub>1</sub> progeny permits an assessment of the phenotype of SSR alleles at each locus. In this study, only one F<sub>1</sub> plant was examined at each ploidy level. Therefore, the F<sub>1</sub> plants in the study do not represent the range of genotypes possible in the F<sub>1</sub> generation. Also, until definitive data are obtained from mapping these loci, the fragments referred to as alleles would better be referred to as "putative alleles."

At the Afct32 locus (Fig. 1a), an allele generally appears to have two bands differing in length by 1 bp. Additional bands that are one and two bases shorter in length than the two primary bands are sometimes present, as is shown in Fig. 1a, lanes 3 and 4, which are the products amplified from W2xiso and the F<sub>1</sub> of W2xiso × *M. sativa* spp. *caerulea* (PI 440501), respectively. The size of the upper band was used to define the size of an allele (Table 4). It is relatively simple to follow the inheritance of alleles from parents to progeny, as in the case of the tetraploid parents and F<sub>1</sub> in Fig. 1a, lanes 6–8. The MA-4 and A-620 parents are both triallelic, with alleles of approximately 102, 122, and 128 and 102, 106, and 156 nucleotides, respectively. The F<sub>1</sub> is tetra-allelic, with alleles of 102, 106, 121, and 127 nucleotides.

At the Afctt1 locus, an allele appears to have a wider range of phenotypes (Fig. 1b). In the diploid and tetraploid alfalfa parents and F<sub>1</sub>s (Fig. 1b, lanes 3–8), alleles appear to possess two bands that differ in length by three nucleotides. The lower band is generally more intense and better resolved and was therefore used to define allele size. The W2xiso diploid parent has alleles of 106 and 107 nucleotides (Fig. 1b, lane 3), the *M. sativa* spp. *caerulea* (PI 440501) parent has alleles of 97 and 118 nucleotides (Fig. 1b, lane 5), and the F<sub>1</sub> has alleles of 106 and 118 nucleotides (Fig. 1b, lane 4). Likewise, in the cross of the tetraploid alfalfa parents, while difficult to determine, MA-4 appears diallelic with alleles of approximately 116 and 122 nucleotides (Fig. 1b, lane 6), the A-620 parent has alleles of 97, 106, and 113 nucleotides (Fig. 1b, lane 8), and the F<sub>1</sub> is triallelic, with alleles of 97, 106, and 122 nucleotides (Fig. 1b, lane 7).

Exact allele sizing at the Afca11 locus (Fig. 1c) was difficult. In the diploid and tetraploid alfalfa parents and F<sub>1</sub>s in

**Table 4.** Number and size (in nucleotides (nt)) of alleles present at four SSR loci in the parents and F<sub>1</sub> of the diploid and tetraploid alfalfa mapping populations.

	AFct32		AFctt1		AFca11		MTLEC2A	
	No.	Size (nt)	No.	Size (nt)	No.	Size (nt)	No.	Size (nt)
Diploid population								
W2xiso (♀ parent)	2	138, 142	2	106, 107	1	146	1	185
<i>M. sativa</i> (♂ parent)	2	116, 148	2	97, 118	2	138, 154	2	186, 190
F <sub>1</sub>	2	142, 148	2	106, 118	2	146, 154	2	185, 190
Tetraploid population								
MA-4 (♀ parent)	3	102, 121, 127	2	116, 122	3	138, 148, 160	3	176, 178, 186
A-620 (♂ parent)	3	102, 106, 157	3	97, 106, 113	3	146, 162, 164	2	176, 186
F <sub>1</sub>	4	102, 106, 121, 127	3	97, 106, 122	3	148, 160, 164	3	176, 178, 186

**Table 5.** Number and size range of alleles at four SSR loci in 26 perennial and annual *Medicago* spp.

	AFct32		AFctt1		AFca11		MTLEC2A	
	No.	Size range (bp)	No.	Size range (bp)	No.	Size range (bp)	No.	Size range (bp)
All genotypes	21	100–188	13	90–135	12	83–177	12	136–304
<i>M. sativa</i> complex (10 genotypes)	14	100–188	11	102–135	10	133–177	7	136–236
Other perennial <i>Medicago</i> spp. (7 genotypes)	5	102–136	7	99–120	6 <sup>a</sup>	141–177	6 <sup>b</sup>	180–304
Annual <i>Medicago</i> spp. (9 genotypes)	3	104–128	6	90–114	4	83–147	1 <sup>c</sup>	166

Note: See Table 2 for listing of genotypes.

<sup>a</sup>Nonspecific PCR products were present for *M. suffruticosa*.

<sup>b</sup>No PCR products were present for *M. suffruticosa* and *M. pironae*.

<sup>c</sup>PCR products were present only in *M. littoralis* and *M. truncatula*.

Fig. 1c, lanes 3–8, an allele generally appeared to consist of three bands differing in length by one nucleotide. The upper band was the most intense and was therefore used to estimate allele size. In lane 3, the W2xiso parent appeared to be a homozygote and, based upon the genotype of the F<sub>1</sub> in Fig. 1c, lane 4, carries an allele of 146 nucleotides. However, the phenotype of the W2xiso parent was not similar to the other genotypes, in that the higher molecular weight band was much less intense than the lower band. The *M. sativa* spp. *caerulea* (PI 440501) parent had alleles of 138 and 154 nucleotides and the F<sub>1</sub> is a heterozygote with alleles of 146 and 154 nucleotides. The estimated allele sizes of the tetraploid alfalfa parents and F<sub>1</sub> are given in Table 4.

While sometimes difficult to interpret clearly, allelic phenotypes at the AFct32, AFctt1, AFca11, and MTLEC2A loci of the parents and F<sub>1</sub> of the diploid and tetraploid alfalfa populations were consistent with Mendelian inheritance of SSR alleles in both ploidy levels. This suggests that development of both diploid and tetraploid genetic linkage maps based on the segregation of SSR alleles is feasible in alfalfa. There are three published alfalfa linkage maps, all of diploid alfalfa (Brummer et al. 1993; Echt et al. 1993; Kiss et al. 1993). Using RAPD markers, Yu and Pauls (1993) were the first to attempt to map alfalfa at the tetraploid level. They concluded that mapping alfalfa at the tetraploid level is feasible; however, because of the dominant nature of RAPD markers, only markers that were either in simplex or duplex conditions could be mapped. SSRs

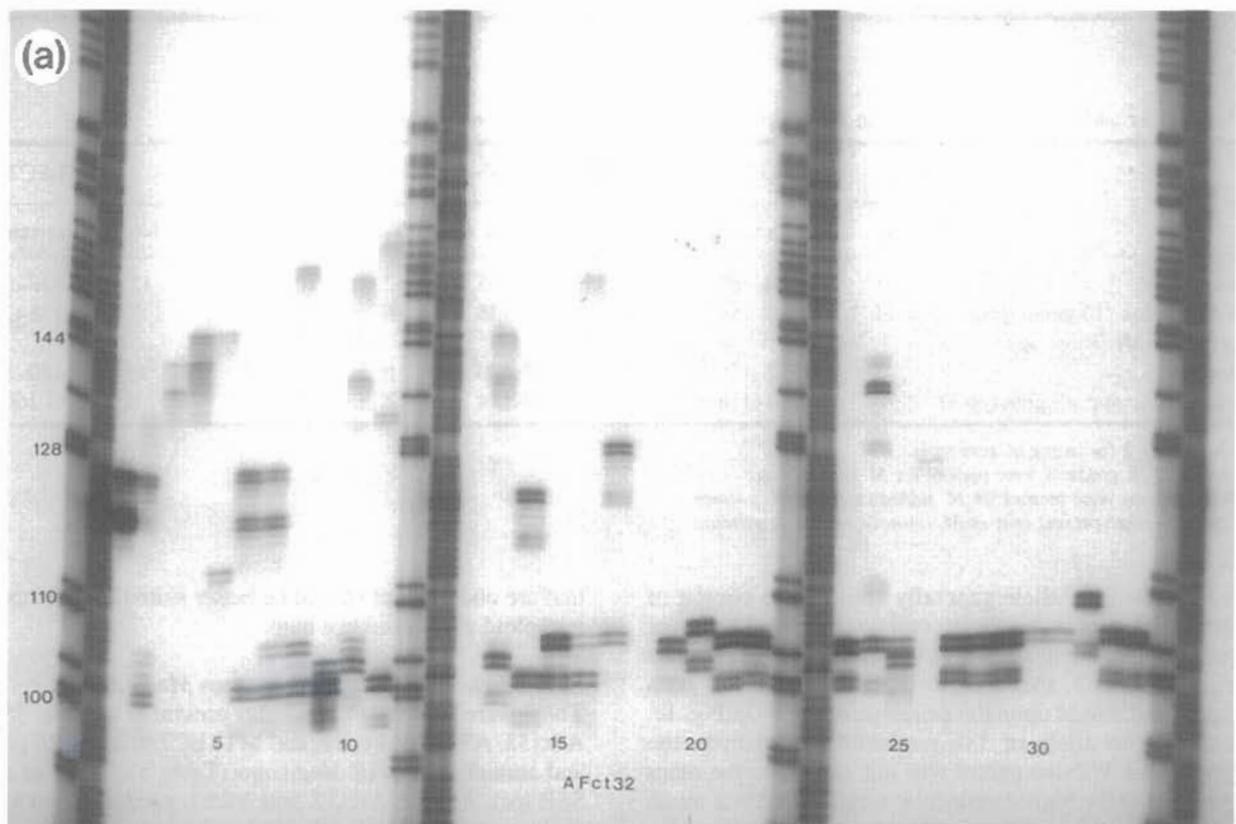
that are codominant should be better suited for constructing a tetraploid alfalfa linkage map.

#### SSR markers present in the genus *Medicago*

There were from 12 to 21 alleles present at the four SSR loci, AFct32, AFctt1, AFca11, and MTLEC2A, in the 26 perennial and annual species of *Medicago* (Table 5). Three of the four SSR loci, AFca11, AFct32, and AFctt1, produced 1–4 alleles in each genotype tested. Allele sizes were within the expected range for products at each locus. The number of PCR products amplified from the genomic DNA of each of the genotypes corresponded to the ploidy level of the species, i.e., diploid genotypes produced up to two alleles and tetraploid genotypes up to four. One exception was the AFca11 locus in *Medicago suffruticosa*, where multiple bands were present, indicating annealing of the PCR primers at numerous places in the genome. The fourth locus, MTLEC2A, which was the only locus from GenBank and which originated from *M. truncatula*, produced PCR products in only two of the annual species, *M. truncatula* and *Medicago littoralis*. No amplification occurred when *Medicago prostrata*, a species closely related to the *M. sativa* complex, was used as template. A similar lack of PCR products resulted with two other perennial species, *M. suffruticosa* and *Medicago pironae*. These amplification failures are probably due to a sequence mismatch at one or both priming sites.

For all four loci, the largest number of alleles was present

**Fig. 1.** SSR alleles at three loci across perennial and annual *Medicago* species. (a) Locus AFct32. (b) Locus AFct11. (c) Locus Afca11. Lane 1, the pBluescript II KS + (Stratagene, LaJolla, Calif.) plasmid containing the alfalfa genomic DNA that was sequenced to determine the PCR primers. Lane 2, W10, the alfalfa population from which the DNA library was constructed (a bulk of 50 plants). Lanes 3–5: parents and  $F_1$  of the diploid alfalfa cross: lane 3, the CADL parent (W2xiso); lane 4, the  $F_1$  plant; and lane 5, the *M. sativa* ssp. *caerulea* parent (PI 440501). Lanes 6–8: parents and  $F_1$  of the tetraploid alfalfa cross: lane 6, the MN White-MA parent (MA-4); lane 7, the  $F_1$  plant; and lane 8, the A-620 parent. Lanes 9–11: alfalfa germplasm from: lane 9, Turkistan; lane 10, India; and lane 11, Peru. Lanes 12–17 and 23: species from the *M. sativa* complex: lane 12, *M. sativa* ssp. *falcata* UGA108602 (2x); lane 13, *M. sativa* ssp. *falcata* (4x); lane 14, *M. sativa* ssp. *caerulea* UAG214 (2x); lane 15, *M. sativa* ssp. *glutinosa* UAG85 (4x); lane 16, *M. glomerata* UAG1529 (2x); lane 17, *M. hemicycla* PI 315481 (2x); and lane 23, *M. prostrata* UAG465 (2x). Lanes 18–25 (excluding lane 23): perennial *Medicago* species: lane 18, *M. suffruticosa* UAG1544; lane 19, *M. papillosa* UAG96; lane 20, *M. pironae* UAG1971; lane 21, *M. carstiensis*, PI 212936; lane 22, *M. cancellata* UAG43; lane 24, *M. marina* UAG1209; and lane 25, *M. arborea* PI 307390. Lanes 26–34: annual *Medicago* species: lane 26, *M. lupulina* cv. George; lane 27, *M. orbicularis* PI 459206; lane 28, *M. rugosa* cv. Paraponto; lane 29, *M. scutellata* cv. Sava; lane 30, *M. littoralis* cv. Harbinger AR; lane 31, *M. truncatula* cv. Caliph; lane 32, *M. minima* PI 499002; lane 33, *M. polymorpha* cv. Santiago; and lane 34, *M. intertexta* PI 498832. Note that at the AFct32 locus (a), lanes 18 and 23 were switched and are lane 18, *M. prostrata* UAG465; and lane 23, *M. suffruticosa* UAG1544.

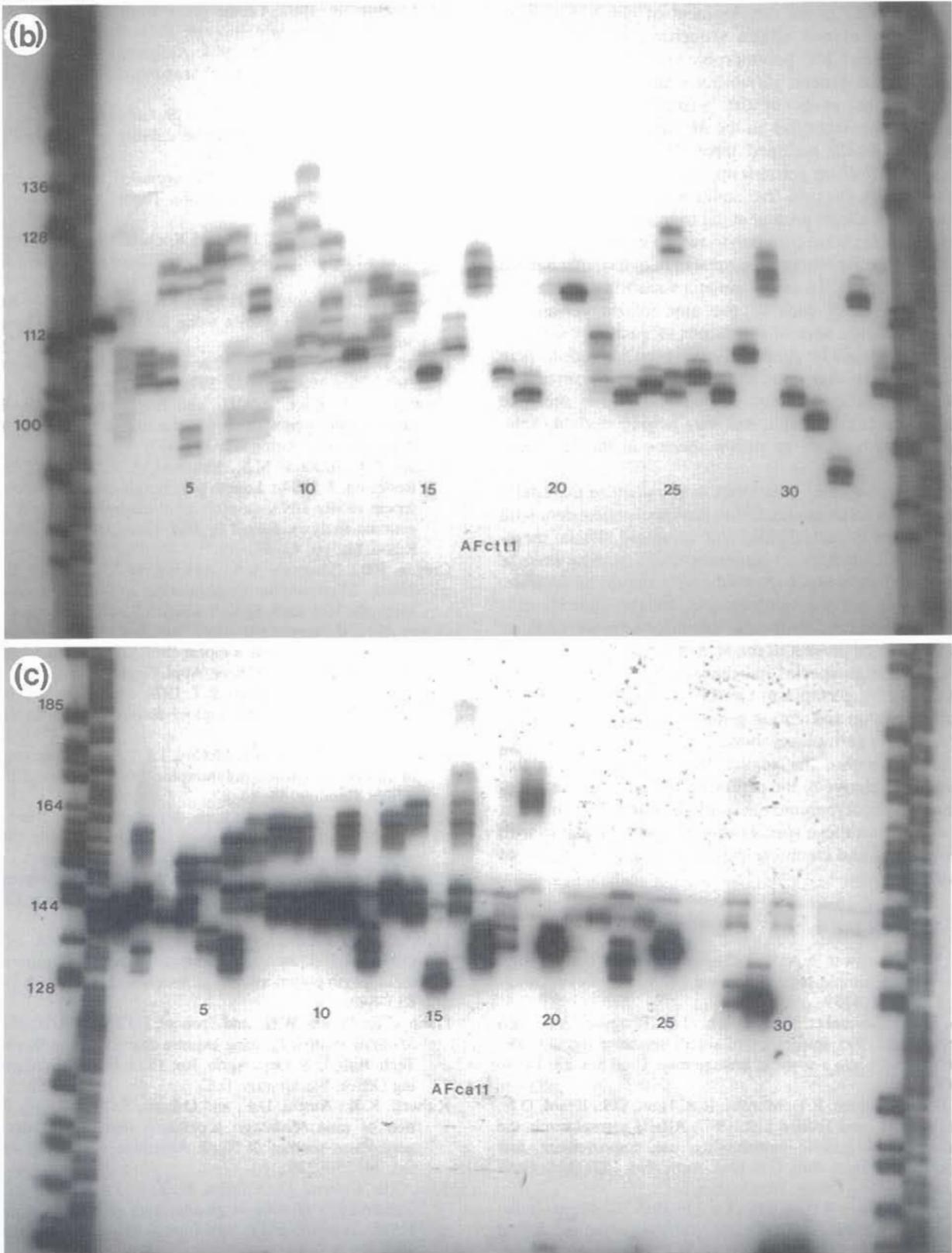


in species of the *M. sativa* complex, followed by perennial species that are more distantly related to alfalfa, and are not part of the *M. sativa* complex. Several of the perennial species, especially ones in the *M. sativa* complex, are autotetraploids and all are outcrossers, resulting in the possibility of up to four alleles per genotype per locus (Fig. 1). Dunbier and Bingham (1974) have proposed the theory of maximum heterozygosity in alfalfa, which emphasizes the importance of intralocus interactions and associates tri- and tetra-allelic loci with maximum performance, as measured by forage yield, fertility, and seed mass. Large genetic variation within and among the diploid subspecies of the *M. sativa* complex and tetraploid alfalfa cultivars was also shown by Brummer et al. (1991) using RFLP markers. This large genetic variation is not surprising and has also been observed in other outcrossing crops, such as *Brassica* spp. (Figdore et al. 1988) and maize

and melon (Shattuck-Eidens et al. 1990). In contrast, relatively limited allelic variation was present in the annual *Medicago* species included in this study. This could be due to the facts (i) that most annual *Medicago* species are diploids ( $2n = 2x = 14$  or  $16$ ) and potentially can only have up to two alleles at any given locus and (ii) that all annual species in the genus *Medicago* are self-pollinating (Lesins and Lesins 1979). Therefore, in most of the annual *Medicago* species only one band (or allele) was present per genotype. The only two exceptions were *M. scutellata* and *M. rugosa*, the only known annual tetraploids (Bauchan and Elgin 1984), in which two alleles were present at the AFct11 locus.

All cultivated alfalfa in North America has been produced from nine germplasm sources that are pure *M. sativa* ssp. *sativa*, pure *M. sativa* ssp. *falcata*, or hybrids of these two subspecies (Barnes et al. 1977; Quiros 1983). Kidwell et al.

Fig. 1 (concluded).



(1994) investigated genetic diversity among the nine germplasm sources used in the development of North American alfalfa using 35 cloned alfalfa sequences as probes. The 35 clones produced 180 polymorphic restriction fragments. However, because genetic variability within was as great as among accessions, seven of the germplasm sources were indistinguishable. Genotypes in the *M. sativa* complex in the current study, which included three *M. sativa* germplasm sources, one *M. falcata* germplasm source, a diploid form of *M. sativa* ssp. *falcata*, and five additional *Medicago* species, had from 7 to 14 alleles present at the four SSR loci (Table 5). Some of the alleles were common to two or more of the germplasm sources, while others were present in one source but not in the others (Fig. 1). To assess genetic variability within and among accessions of each of the nine alfalfa germplasm sources using SSRs, several accessions of each of the germplasm sources should be studied. However, observations here indicate that SSRs may contribute additional information to the study of genetic variation among and within the nine germplasm sources of alfalfa, and may help in studying relationships among as well as within species of the *M. sativa* complex.

This study shows that SSR DNA is abundant in the alfalfa genome and that inheritance of SSR markers is consistent with Mendelian rules in both diploid and tetraploid alfalfa; therefore, using SSR markers to construct alfalfa genetic linkage maps at both diploid and tetraploid levels should be feasible. SSR markers should also be useful tools for genotype identification within the genus *Medicago*. The large number of alleles demonstrated to be present in the *M. sativa* complex and other perennial *Medicago* species indicates the possibility of using SSRs markers for germplasm identification and for studying relationships within and among genotypes from the nine North American alfalfa germplasm sources. Although the number of alleles present within the annual *Medicago* species in this study was low relative to the perennial species, the identification of annual genotypes and the study of genetic relationships among and within these species should also be feasible with the development and careful selection of an appropriate set of SSR markers.

## References

- Akkaya, M.S., Bhagwat, A.A., and Cregan, P.B. 1992. Length polymorphism of simple sequence repeat DNA in soybean. *Genetics*, **132**: 1131–1139.
- Akkaya, M.S., Shoemaker, R.C., Specht, J.E., Bhagwat, A.A., and Cregan, P.B. 1995. Integration of simple sequence repeat (SSR) DNA markers into a soybean linkage map. *Crop Sci.* **35**: 1439–1445.
- Barnes, D.K., Bingham, E.T., Murphy, R.P., Hunt, O.J., Beard, D.F., Skrdla, W.H., and Teuber, L.R. 1977. Alfalfa germplasm in the United States: genetic vulnerability, use, improvement, and maintenance. *Tech. Bull. U.S. Dep. Agric. No. 1571*. U.S. Government Printing Office, Washington, D.C.
- Barnes, D.K., Goplen, B.P., and Baylor, J.E. 1988. Highlights in the U.S.A. and Canada. In *Alfalfa and alfalfa improvement*. Edited by A.A. Hanson, D.K. Barnes, and R.R. Hill. Agronomy Monograph No. 29. American Society of Agronomy – Crop Science Society of America – Soil Science Society of America, Madison, Wis. pp. 1–24.
- Bauchan, G.R., and Elgin, J.H., Jr. 1984. A new chromosome number for the genus *Medicago*. *Crop Sci.* **24**: 193–195.
- Bauchrowitz, M.A., Barker, D.G., Nadaud, I., Rouge, P., and Lescure, B. 1992. Lectin genes from the legume *Medicago truncatula*. *Plant Mol. Biol.* **19**: 1011–1017.
- Bell, C.J., and Ecker, J.R. 1994. Assignment of 30 microsatellite loci to the linkage map of *Arabidopsis*. *Genomics*, **19**: 137–144.
- Bingham, E.T., and McCoy, T.J. 1979. Cultivated alfalfa at the diploid level: origin, reproductive stability, and yield of seed and forage. *Crop Sci.* **19**: 97–100.
- Brummer, E.C., Kochert, G., and Bouton, J.H. 1991. RFLP variation in diploid and tetraploid alfalfa. *Theor. Appl. Genet.* **83**: 89–96.
- Brummer, E.C., Bouton, J.H., and Kochert, G. 1993. Development of an RFLP map in diploid alfalfa. *Theor. Appl. Genet.* **86**: 329–332.
- Brummer, E.C., Bouton, J.H., and Kochert, G. 1995. Analysis of annual *Medicago* species using RAPD markers. *Genome*, **38**: 362–367.
- Brunel, D. 1994. A microsatellite marker in *Helianthus annuus* L. 1994. *Plant Mol. Biol.* **24**: 397–400.
- Crawford, E.J., Lake, A.W.H., and Boyce, K.G. 1989. Breeding annual *Medicago* species for semiarid conditions in southern Australia. *Adv. Agron.* **42**: 399–437.
- Cregan, P.B., Akkaya, M.S., Bhagwat, A.A., Lavi, U., and Jiang Rongwen, J. 1994a. Length polymorphisms of simple sequence repeat (SSR) DNA as molecular markers in plants. In *Plant genome analysis*. Edited by P.M. Gresshoff. CRC Press, Boca Raton, Fla. pp. 43–49.
- Cregan, P.B., Bhagwat, A.A., Akkaya, M.S., and Jiang Rongwen, J. 1994b. Microsatellite fingerprinting and mapping of soybean. *Methods. Mol. Cell. Biol.* **5**: 49–61.
- Diwan, N., and Cregan, P.B. 1997. Automated sizing of fluorescent labeled simple sequence repeat (SSR) markers to assay genetic variation in soybean. *Theor. Appl. Genet.* In press.
- Dunbier, M.W., and Bingham, E.T. 1974. Maximum heterozygosity in alfalfa: results using haploid-derived autotetraploids. *Crop Sci.* **15**: 527–531.
- Echt, C.S., Erdahl, L.A., and McCoy, T.J. 1992. Genetic segregation of random amplified polymorphic DNA in diploid cultivated alfalfa. *Genome*, **35**: 84–87.
- Echt, C.S., Kidwell, K.K., Knapp, S.J., Osborn, T.C., and McCoy, T.J. 1993. Linkage mapping in diploid alfalfa (*Medicago sativa*). *Genome*, **37**: 61–71.
- Elgin, J.H., and Ostazeski, S.A. 1984. Registration of eleven Race 1 and Race 2 anthracnose resistant alfalfa germplasm lines. *Crop Sci.* **24**: 623–624.
- Figdore, S.S., Kennard, W.C., Song, K.M., Slocum, M.K., and Osborn, T.C. 1988. Assessment of the degree of restriction fragment length polymorphism in *Brassica*. *Theor. Appl. Genet.* **75**: 833–840.
- Gunn, C.R., Skrdla, W.H., and Spencer, H.C. 1978. Classification of *Medicago sativa* L. using legume characters and flower colors. *Tech. Bull. U.S. Dep. Agric. No. 1574*. U.S. Government Printing Office, Washington, D.C.
- Kidwell, K.K., Austin, D.F., and Osborn, T.C. 1994. RFLP evaluation of nine *Medicago* accessions representing the original germplasm sources of North American alfalfa cultivars. *Crop Sci.* **34**: 230–236.
- Kiss, G.B., Csanadi, G., Kalman, K., Kalo, P., and Okresz, L. 1993. Construction of a basic genetic map for alfalfa using RFLP, RAPD, isozyme and morphological markers. *Mol. Gen. Genet.* **238**: 129–137.
- Kresovich, S.K., Szewc-McFadden, A.K., and Bliet, S.M. 1995. Abundance and characterization of simple-sequence repeats (SSRs) isolated from a size-fractionated genomic library of *Brassica napus* L. (rapeseed). *Theor. Appl. Genet.* **91**: 206–211.

- Lagercrantz, U., Ellegren, H., and Andersson, L. 1993. The abundance of various polymorphic microsatellite motifs differs between plants and vertebrates. *Nucleic Acids Res.* **21**: 1111–1115.
- Lavi, U., Akkaya, M.S., Bhagwat, A.A., Lahav, E., and Cregan, P.B. 1995. Generation of simple sequence repeat DNA markers in avocado (*Persea americana* M.). *Euphytica*, **80**: 171–177.
- Lesins, K., and Lesins, I. 1979. Genus *Medicago* (Leguminosae). A taxogenetic study. Junk, The Hague, The Netherlands.
- Litt, M., and Luty, J.A. 1989. A hypervariable microsatellite revealed by *in vitro* amplification of a dinucleotide repeat within the cardiac muscle actin gene. *Am. J. Hum. Genet.* **44**: 397–401.
- Morgante, M., and Olivieri, A.M. 1993. PCR-amplified microsatellites as markers in plant genetics. *Plant J.* **3**: 175–182.
- Pramanik, S.K., and Bewley, J.D. 1993. A ubiquitin carrier protein cDNA from developing alfalfa embryos. *Plant Physiol.* **102**: 1049–1050.
- Quiros, C.F. 1983. Alfalfa, luzerne (*Medicago sativa* L.) In *Isozymes in plant genetics and breeding*, Part B. Edited by S.D. Tanksley, and T.J. Orton. Elsevier Scientific Publishing Co., Amsterdam. pp. 253–294.
- Quiros, C.F., and Bauchan, G.R. 1988. The genus *Medicago* and the origin of *Medicago sativa* complex. In *Alfalfa and alfalfa improvement*. Edited by A.A. Hanson, D.K. Barnes, and R.R. Hill. Agronomy Monograph No. 29. American Society of Agronomy – Crop Science Society of America – Soil Science Society of America, Madison, Wis. pp. 93–124.
- Rongwen, J., Akkaya, M.S., Bhagwat, A.A., Lavi, U., and Cregan, P.B. 1995. The use of microsatellite DNA markers for soybean genotype identification. *Theor. Appl. Genet.* **90**: 43–48.
- Sambrook, J., Fritsch, E.F., and Maniatis, T. 1989. *Molecular cloning: a laboratory manual*. 2nd ed. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
- Senior, M.L., and Heun, M. 1993. Mapping maize microsatellites and polymerase chain reaction confirmation of the targeted repeats using a CT primer. *Genome*, **36**: 884–889.
- Senior, M.L., Chin, E.C.L., Lee, M., Stuber, C.W., and Smith, J.S.C. 1996. Simple sequence repeat markers developed from maize sequences found in the Genbank database: map construction. *Crop Sci.* **36**: 1676–1683.
- Shattuck-Eidens, D.M., Bell, R.N., Neuhausen, S.L., and Helentjaris, T. 1990. DNA sequence variation within maize and melon: observations from polymerase chain reaction amplification and direct sequencing. *Genetics*, **126**: 207–217.
- Tautz, D. 1989. Hypervariability of simple sequences as a general source of polymorphic DNA markers. *Nucleic Acids Res.* **17**: 6463–6471.
- Wang, Z., Weber, J.L., Zhong, G., and Tanksley, S.D. 1994. Survey of plant short tandem DNA repeats. *Theor. Appl. Genet.* **88**: 1–6.
- Weber, J.K., and May, P.E. 1989. Abundant class of human DNA polymorphisms which can be typed using the polymerase chain reaction. *Am. J. Hum. Genet.* **44**: 388–397.
- Yu, K.F., and Pauls, K.P. 1993. Rapid estimation of genetic relatedness among heterogeneous populations of alfalfa by random amplification of bulked genomic DNA samples. *Theor. Appl. Genet.* **86**: 788–794.