CELL BIOLOGY & MOLECULAR GENETICS

Forage Yields of Alfalfa Populations Derived from Parents Selected on the Basis of Molecular Marker Diversity

K. K. Kidwell, L. M. Hartweck, B. S. Yandell, P. M. Crump, J. E. Brummer, J. Moutray, and T. C. Osborn*

ABSTRACT

It is important to know if forage yields of alfalfa (*Medicago sativa* L.) cultivars could be improved by using a subset of genetically diverse parents selected from a larger population of parents for a synthetic cultivar. We used restriction fragment length polymorphisms to assess the level of genetic diversity among 93 alfalfa genotypes previously selected as parents for a potential commercial synthetic cultivar. On the basis of average pairwise genetic distances between genotypes, four synthetic populations (samples) were developed with 2, 4, 8, 12, 16, or 24 parents selected for genetic dissimilarity (DIS) or similarity (SIM). Resulting synthetic populations were evaluated for forage yield in replicated field trials at three locations for 2 yr. Forage yields of populations with low numbers of parents were highly variable among samples, and as number of parents increased, genetic diversity levels and yield variation decreased. No significant difference ($P > 0.05$) was detected between forage yields of DIS and SIM groups averaged over parent number and samples; however, yield variation among samples within parent number of DIS and SIM groups was highly significant ($P < 0.0001$). The SIM group had more sample variation than did the DIS group, and more SIM than DIS populations were among the lower yielding populations. The lack of significant differences between the forage yields of DIS and SIM populations may be due to linkage equilibrium in the population used for selecting parents and the inability to target heterozygosity to specific genome regions affecting yield.

Most alfalfa cultivars are genetically broad-based synthetics developed by randomly intermating selected parents and advancing their offspring through several generations of open pollination (Busbice, 1969; Hill and Elgin, 1981; Hill et al., 1988). Several studies have been conducted to determine the optimum number of parents to intermate to maximize forage yields in a synthetic. In general, the highest yielding experimental synthetic cultivars had greater than four but fewer than 16 parents (Busbice and Gurgis, 1976; Hill and Elgin, 1981). Fewer parents are required when they are genetically diverse rather than highly related, but at least four parents are needed to avoid excessive inbreeding (Busbice, 1970).

Difficulties associated with distinguishing genetically diverse individuals have forced most breeders to include larger numbers of parents in their synthetic cultivars to minimize inbreeding and conserve genetic variability for other traits. Over half of the cultivars released between 1973 and 1982 were synthesized from more than 40 parental genotypes (Hill et al., 1988). The large amount of genetic variability for forage yield observed within most alfalfa cultivars suggests that a subset of the original parents could be intermated to produce a higher yielding synthetic cultivar (Hill and Elgin, 1981).

In previous studies, we used molecular markers to estimate the genetic distance of pairs of alfalfa genotypes and the heterozygosity of their single-cross progenies (Kidwell et al., 1994a,b). These studies included four diploid genotypes with a range of genetic diversity and their isogenic tetraploids derived by chromosome doubling. We found that for the tetraploid genotypes the genetic distance and heterozygosity estimates had significant, positive correlations with each other and with forage yields of the single-cross progenies. The results suggest that molecular markers may be useful for identifying pairs of genetically diverse parents for producing highly heterozygous and high yielding, single-cross hybrid progenies.

Molecular markers also could be used to identify more than two genetically diverse individuals. The resulting synthetic populations should have a higher number of different alleles and more heterozygosity at marker loci than unselected populations or populations derived from genetically similar parents on the basis of molecular markers. If diversity at marker loci reflects diversity at linked loci, this selection would increase the chance of having at least one favorable dominant allele at linked loci affecting forage yield and greater overall complementary gene interaction. In contrast, selecting genetically similar parents by molecular markers should result in fewer different alleles, lower heterozygosity, a lower chance of favorable dominant alleles and less complementary gene interaction. This reasoning is analogous to that used in explaining the roll of linkats and complementary gene interaction in progressive heterosis of tetraploid alfalfa (Bingham, 1983) and in the in-

K.K. Kidwell, Dep. of Crop and Soil Sciences, Washington State Univ., Pullman, WA 99164-6420; L.M. Hartweck, Dep. of Plant Biology, Univ. of Minnesota, St. Paul, MN 55108; T.C. Osborn, Dep. of Agronomy, 1575 Linden Dr., Univ. of Wisconsin, Madison, WI 53706; B.S. Yandell, Dep. of Horticulture and Statistics, 1575 Linden Dr., Univ. of Wisconsin, Madison, WI 53706; P.M. Crump, Dep. of Computing and Biometry, 1675 Observatory Dr., Univ. of Wisconsin, Madison, WI 53706; and J.E. Brummer and J. Moutray, ABI - Alfalfa, 2316 259th St., Ames, IA, 50014. Contribution from the Dep. of Agronomy, Univ. of Wisconsin, Madison. Received 30 Aug. 1997. *Corresponding author (tcosborn@facstaff.wisc.edu).

Published in Crop Sci. 39:223–227 (1999).

Abbreviations: GD, genetic dissimilarity; DIS, population from genetically dissimilar parents; SIM, population from genetically similar parents; RFLP, restriction fragment length polymorphism; ANOVA, analysis of variance; GLM, general linear model.

creased yield associated with tetraploid versus diploid alfalfa (Bingham et al., 1994).

In the study reported here, we used molecular markers to select sets of genetically similar and dissimilar parents from a group of elite alfalfa genotypes. Different numbers of parents, ranging from two to 40, were selected and progenies from intermating these parents were evaluated to determine the effects of selecting different numbers of genetically similar or dissimilar parents on forage yields.

MATERIALS AND METHODS

Plant Materials, RFLP Analysis and Data Collection

Tetraploid alfalfa genotypes evaluated in this study were selected from populations developed for commercial exploitation by ABI-Alfalfa. The origin of these populations were nine cultivars (Vernal, Weevlchek, Titan, Anchor, Saranac, Tempo, Kanza, Dawson, and Cody) that had been intermated and undergone 12 cycles of phenotypic recurrent selection for disease resistance, leaf hopper resistance, and winter survival. The resulting populations included germplasm from seven of the nine (excluding African and Peruvian) germplasm sources for North American alfalfa cultivars (Barnes et al., 1977). Several hundred genotypes from these populations were initially screened in growth chambers by standard procedures for acceptable levels of resistance to bacterial wilt, Fusarium wilt, Verticillium wilt, Anthracnose, Phytophthora root rot and Aphanomyces. The surviving genotypes were vegetatively propagated and evaluated in fields at locations throughout the Midwest in 1990 and 1991 to confirm original disease resistance results and to select for vigor, winter survival, potato leaf hopper resistance, and resistances to other leaf diseases (data not shown). Ninety-three genotypes with superior phenotypic characteristics and multiple disease resistances were selected as parents for a synthetic population to be evaluated for commercial release.

The 93 genotypes were analyzed for RFLPs by previously described procedures for total genomic DNA isolation, DNA quantification, restriction enzyme digestion, gel electrophoresis, Southern blotting, probe labeling by random-hexamer priming, hybridization, and autoradiography (Kidwell et al., 1994b). *Eco*RI or *Hin*dIII digested plant genomic DNAs were hybridized to 61 recombinant DNA clones, 36 of which were used previously to locate 2 to 7 RFLP loci on each of the eight linkage groups of diploid alfalfa (Echt et al., 1994; Tavoletti et al., 1996). The remaining clones were from the same genomic DNA and cDNA libraries used previously to select clones for mapping alfalfa (Echt et al., 1994; Kidwell and Osborn, 1993; Tavoletti et al., 1996). Banding profiles were recorded for each RFLP probe by assigning a number to every unique band detected on autoradiographs based on migration distance compared with λ DNA digested with restriction enzymes. Presence/absence data were combined for all RFLP probes, and these data were used to measure genetic diversity among genotypes.

Genetic Diversity Estimates, Parental Selection, and Population Synthesis

Genetic dissimilarities between all pairs of individuals were calculated as the complement of the Dice Coefficient (Dice, 1945) by the computer program NTSYS-pc (Rohlf, 1992). Genetic dissimilarities were assumed to represent the level of genetic diversity between two individuals, with a value of zero indicating no RFLPs detected between a pair of genotypes for the set of probes used and a value of 1.0 indicating a pair of individuals differed for all RFLPs detected. Groups of genetically dissimilar (DIS) parents were identified by initially selecting two parents at random from the highest 2% ranking of GDs. SAS/IML (SAS, 1989) procedures were developed (programs available upon request) to search the 93 × 93 GD matrix to identify a third parent that would maximize the average of the pairwise GDs with the two parents that had been selected previously. The search was repeated until 24 parents had been identified, and the samples with 2, 4, 8, 12, 16, and 24 parents were selected to represent a series of DIS parents. The entire process was repeated three times starting with new pairs of parents selected at random from the 2% tails of the GD frequency distribution. Thus, we obtained four samples of DIS parents for each parent number level. This procedure does not guarantee that parents with the maximum average GD were selected, but it is a fast computational method of selecting multiple samples of parents with near maximum GD. Four samples of genetically similar (SIM) parents at the same parent number levels were identified in a similar fashion by selecting four two-parent combinations from the lowest 2% GD ranking and then adding parents that minimized the average of the pairwise GDs. In addition, one sample each of 40 DIS and 40 SIM parents was selected. This resulted in a total of 50 selected groups of parents plus the 93 original genotypes for creating 51 synthetic populations.

Selected parental genotypes were vegetatively propagated and hand pollinations without emasculation were initiated in the greenhouse in October 1992. Groups of plants to be intercrossed were distributed in a random order throughout the greenhouse, and pots of genotypes within populations were rotated frequently to ensure that each single cross within a population occurred at a similar frequency. Approximately equal quantities of seed from each female parent were combined to represent a synthetic population for planting in the field.

Field Evaluation and Data Analyses

Forage yields of the 51 synthetic populations and three check cultivars (ABI-9017, Vernal, and Apollo Supreme) were evaluated at three locations: two sites at Ames, IA; one at Livingston, WI. A second site at Livingston was lost because of flooding in the establishment year, and a trial in Nampa, ID, was dropped from the analysis because the location was irrigated and had a significantly different variance than the other locations. Entries were seeded in single row plots 1.8 m long at a seeding rate of 0.6 g per row in the fall of 1993. A row of ABI-9017 was planted between each entry to ensure even competition for entry rows. Four replications of each population were evaluated at each location using a randomized complete block design. Top growth was hand harvested from 0.9 m per row when the entire field reached approximately 10% bloom and the same plots were harvested four times during each of the 1994 and 1995 growing season. Forage samples were dried, and weights were recorded and combined over the eight harvests for each field trial.

Two ANOVAs were performed with data for all populations derived from the 2, 4, 8, 12, 16, and 24 parent numbers of the DIS and SIM diversity groups. In the first ANOVA, a split-split plot model was used with diversity group (fixed effect) as the main plot, parent number within diversity group (fixed effect) as the subplot, and sample within parent number and diversity group (random effect) as the sub-subplot. We performed a second ANOVA to examine variation between diversity groups at each location and parent number level and variation among samples at each location, parent number level

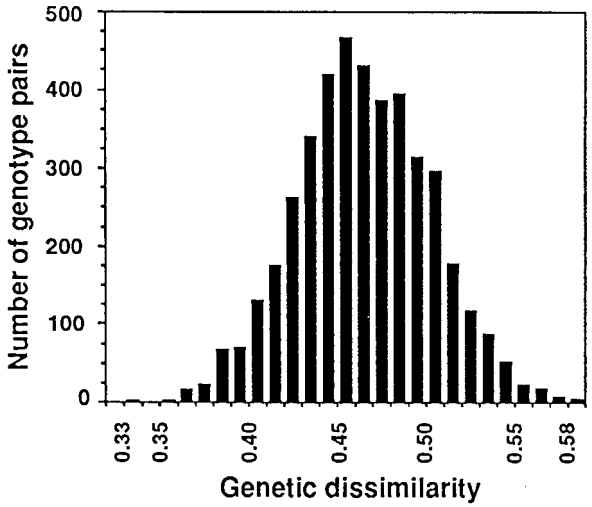

Fig. 1. Frequency distribution of the 4278 pairwise genetic dissimilarity values for 93 tetraploid alfalfa genotypes (mean = 0.46). Values were calculated as the complement of the Dice coefficient with data for 399 RFLPs.

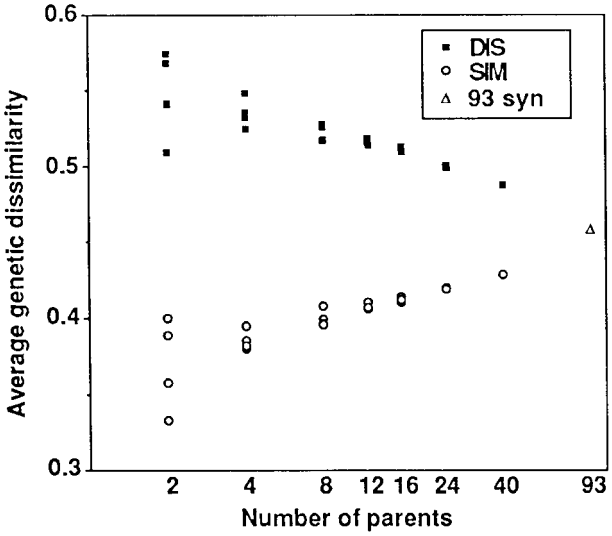

Fig. 2. Plot of number of parents (log scale) versus average genetic dissimilarity of alfalfa parents selected for genetic dissimilarity (DIS) or for genetic similarity (SIM) from a group of 93 parents (93 syn).

and diversity group, using the appropriate error terms to calculate *F* values. All models were analyzed by SAS GLM procedures (SAS, 1989).

RESULTS

Each of the 61 RFLP probes used in this study detected polymorphism among the 93 alfalfa genotypes. Presence or absence data for 399 polymorphic restriction fragments were scored for each genotype with the number of fragments scored per probe ranging from 2 to 13 (avg. of 6.5 fragments per probe). Genetic dissimilarity values for the 4278 possible pairwise comparisons among the 93 plants were normally distributed (Fig. 1). Groups of parents were selected and their average GD values were compared (Fig. 2). All groups of DIS parents had larger average GD values and all groups of SIM parents had smaller values than the average GD value for the 93 plants. As parent number increased, the average GD decreased for DIS populations and increased for SIM populations and the among sample variation decreased for both DIS and SIM populations (Fig. 2).

In the ANOVA for forage yield, location, block within location, sample and location × sample effects were significant sources of variation; however, diversity group (DIS vs. SIM) and parent number within diversity group were non-significant (Table 1). Because of the significant sample effects in the main model, we analyzed the sample variation for each parent number within diversity group at each location (Table 2). In general, as parent number increased, the among sample variation for average GD (Fig. 2) and forage yield (Fig. 3) decreased simultaneously. Both DIS and SIM populations derived from two parents had significant sample variation, but the SIM populations were more variable than the DIS populations for populations having four to 16 parents (Table 2; Fig. 3).

The mean forage yields of DIS and SIM populations at each parent number level were compared at the three locations. Only one comparison was significant (24 DIS

Table 1. ANOVA of forage yield (grams dry matter/row combined over eight harvests) at three locations for 48 alfalfa synthetic populations created by selecting four samples each of 2, 4, 8, 12, 16, or 24 parents on the basis of genetic similarity or dissimilarity (diversity group).

Source of variation†	df‡	MS§	F ratio	*F*	*P* > *F*¶
(1) loc	2	425 852	1/2	11.8	0.0031
(2) block(loc)	9	36 120	2/8	3.2	0.0023
(3) diversity	1	87 216	3/5	1.1	0.3060
(4) parent(diversity)	10	35 270	4/5	0.4	0.9187
(5) sample[parent(diversity)]	36	80 875	5/10	10.7	0.0001
(6) loc × diversity	2	5 021	6/8	0.5	0.6383
(7) loc × parent(diversity)	20	11 601	7/8	1.0	0.4259
(8) loc × sample[parent(diversity)]	72	11 115	8/10	1.5	0.0107
(9) model sum	152	36 648	9/10	4.9	0.0001
(10) residual	423	7 529			
Total	575				

† loc = location; (block(loc)) = block within location; diversity = diversity group; parent(group) = parent number within diversity group; sample(parent(diversity)) = random samples of parents selected within parent number and diversity group.
‡ df = degrees of freedom.
§ MS = mean square.
¶ *P* > *F* = probability of a larger F value.

Table 2. Mean forage yields (averaged over replications and samples) for each parent number (2, 4, 8, 12, 16, 24, 40, and 93) and diversity group (DIS and SIM) at each of three locations, and significance of sample variation for means based on four samples.

Parent no. and diversity group	No. of samples	Location		
		Ames 1	Ames 2	Livingston
		g DM/row		
2 DIS	4	822**†	761**	671**
2 SIM	4	741**	699**	710**
4 DIS	4	813	803**	717
4 SIM	4	779**	811*	714**
8 DIS	4	817	825	768
8 SIM	4	851**	811**	723
12 DIS	4	807	766	709
12 SIM	4	747*	722*	666
16 DIS	4	842*	767	715
16 SIM	4	822	768	701**
24 DIS	4	839	769	788
24 SIM	4	804	801	688
40 DIS	1	899	849	717
40 SIM	1	793	853	718
93 synthetic	1	823	828	788
ABI9017	1	757	833	749
Vernal	1	777	740	599
Apollo Supreme	1	793	770	625

† Means followed by * and ** had significant sample variation at $P < 0.05$ and $P < 0.01$, respectively.

> 24 SIM at Livingston, $P < 0.05$, sample and overall means shown in Fig. 2 and Table 2). However, for two-thirds of all DIS-SIM comparisons, the DIS populations had higher forage yields than the SIM populations (Table 2). One factor contributing to this observation was the greater number of low yielding SIM populations (compared with DIS) at each location (Fig. 3).

DISCUSSION

The purpose of this study was to determine if selection of parents on the basis of molecular marker diversity would affect forage yields of synthetic populations derived from the parents. We tested this hypothesis by selecting subsets of parents that had either larger (DIS) or smaller (SIM) average GD values compared with the value for 93 genotypes and evaluated populations derived from intermating these parents for forage yield. Populations derived from smaller numbers of parents varied more for forage yield than those derived from larger numbers of parents, and some of these populations yielded significantly more than the 93 parent synthetic population. However, selection based on molecular marker diversity had no consistent effect on forage yield.

Our results could have been effected by limitation in the range of genetic diversity among the 93 genotypes that we used for selection. To address this issue, the average GD values from this study can be compared with values from a previous study on genetic diversity within and between nine accessions representing the original germplasm sources for North American alfalfa cultivars (Kidwell et al., 1994c). The genetic diversity of genotypes used in that study should reflect the potential range of diversity in cultivated alfalfa, with intra-accession diversity representing a minimum diversity level and inter-accession diversity representing a maximum

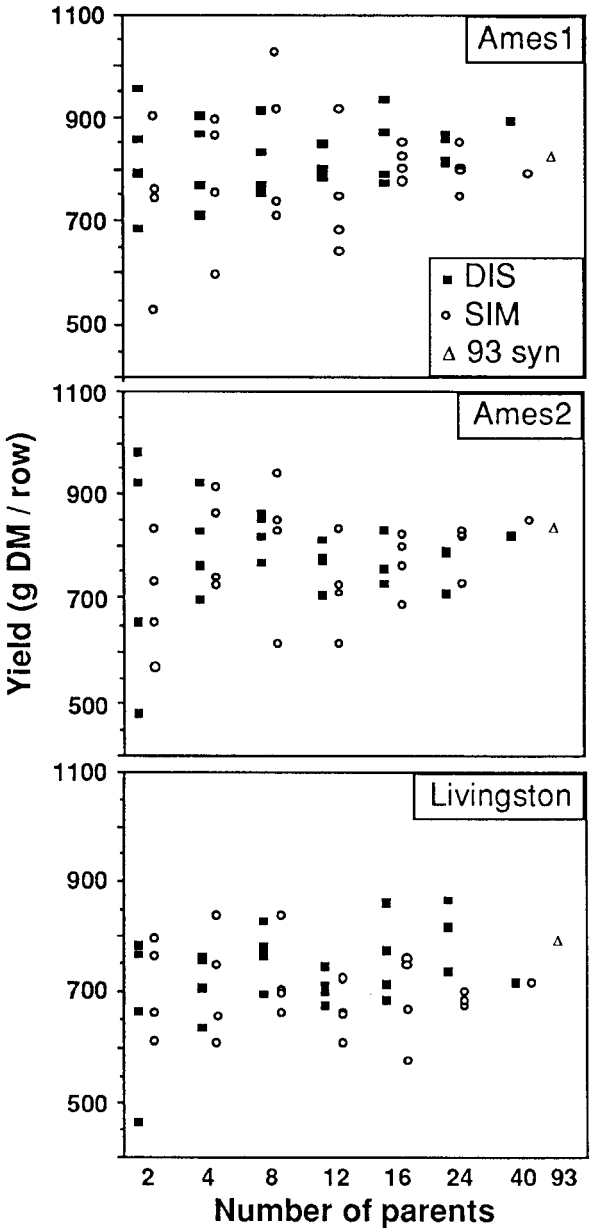

Fig. 3. Plot of the number of parents versus forage yield (grams dry matter/row combined over harvests) of alfalfa synthetic populations grown at three locations. Genetic diversity of the parents based on RFLP assessment are indicated (DIS: genetically dissimilar; SIM: genetically similar; 93 syn: 93 parent synthetic population).

diversity level. The largest average GD values from that study (one minus the average Dice coefficients in Table 2 of Kidwell et al., 1994c) were between the 12 genotypes of *M. falcata* L. and the 12 genotypes of other accessions (0.50–0.54). These values are similar to those of our 12 DIS parents (0.52) even though they were calculated with only the comparisons of genotypes between accessions. The smallest values were among the 12 genotypes within accessions (0.24–0.37). These are considerably lower than those of our 12 SIM parents (0.41), and may reflect the fact that the 93 genotypes used in our study had been derived from the original germplasm sources after many generations of intermating.

Repeated cycles of intermating moves the resulting

population toward linkage equilibrium, breaking up linkages among alleles that were previously associated because of common ancestry. Thus, diversity at molecular marker loci among the parents we selected may not have accurately reflected diversity and heterozygosity at linked trait loci. In fact, it may be very difficult to achieve substantial differences in heterozygosity throughout the genome by selecting parents from populations near linkage equilibrium. Also, regions in which heterozygosity would affect yield may occur in only a portion of the genome. Bernardo (1992) showed that these regions should be emphasized in using markers to select parents of a hybrid; however, his models were for inbred parents and required prior knowledge of quantitative trait locus effects. This information will be difficult to obtain and implement for selection of multiple parents from populations of tetraploids in linkage equilibrium.

Although molecular markers may not be useful for selecting genotypes to be used directly as parents of synthetic cultivars, they could be used to help identify and develop genetically unique germplasm that complements existing cultivar germplasm. We recognize that additive gene effects, as well as dominance effects and complementary gene interactions, are important for improving alfalfa forage yields. Woodfield and Bingham (1995) demonstrated this by improving forage yields of closed, two-allele populations for which dominance and dominance interactions could only decrease. Thus, parental development and selection strategies that capitalize on both additive and non-additive forms of genetic variation will be most successful in improving alfalfa forage yield.

ACKNOWLEDGMENTS

We thank Robert Vogelzang and Rebecca Burkhamer for their assistance with the Livingston field trial, and E.T. Bingham for helpful discussions. Funding was provided by ABI-Alfalfa and the College of Agricultural and Life Sciences, Univ. of Wisconsin-Madison.

REFERENCES

Barnes, D.K., E.T. Bingham, J.D. Axtell, and W.H. Davis. 1972. The flower, sterility mechanisms, and pollination control. p. 123–141. *In* Alfalfa science and technology. Agron. Monogr. 15. ASA, Madison, WI.

Bernardo, R. 1992. Relationship between single-cross performance and molecular marker heterozygousity. Theor. Appl. Genet. 83: 628–634.

Bingham, E.T. 1983. Maximizing hybrid vigour in autotetraploid alfalfa. Ciba Found. Symp. 97:130–141.

Bingham, E.T., R.W. Groose, D.R. Woodfield, and K.K. Kidwell. 1994. Complementary gene interactions in alfalfa are greater in autotetraploids then diploids. Crop Sci. 34:823–829.

Busbice, T.H. 1969. Inbreeding in synthetic varieties. Crop Sci. 9:601–604.

Busbice, T.H. 1970. Predicting yield of synthetic varieties. Crop Sci. 10:265–269.

Busbice, T.H., and R.Y. Gurgis. 1976. Evaluating parents and predicting performance of synthetic alfalfa varieties. USDA ARS-S-130, U.S. Gov. Print. Office, Washington, DC.

Dice, L.R. 1945. Measures of the amount of ecological association between species. Ecology 26:297–302.

Echt, C.S., K.K. Kidwell, T.C. Osborn, S.J. Knapp, and T.J. McCoy. 1994. Linkage mapping in diploid alfalfa (*Medicago sativa*). Genome 37:61–71.

Hill, Jr., R.R., and J. H. Elgin, Jr. 1981. Effect of number of parents on performance of alfalfa synthetics. Crop Sci. 21:298–300.

Hill, Jr., R.R., J.S. Shenk, and R.F Barnes. 1988. Breeding for yield and quality. p. 809–825. *In* A.A. Hanson et al. (ed.) Alfalfa and alfalfa improvement. ASA, Madison, WI.

Kidwell, K.K., and T.C. Osborn. 1993. Variation among alfalfa somaclones in copy number of repeated DNA sequences. Genome 36:906–912

Kidwell, K.K., E.T. Bingham, D.R. Woodfield, and T.C. Osborn. 1994a. Relationships among genetic distance, forage yield and heterozygosity of isogenic diploid and tetraploid alfalfa populations. Theor. Appl. Genet. 89:323–328.

Kidwell, K.K., D.R. Woodfield, E.T. Bingham, and T.C. Osborn. 1994b. Molecular marker diversity and yield of isogenic 2× and 4× single-crosses of alfalfa. Crop Sci. 34:784–788.

Kidwell, K.K., D. Austin, and T.C. Osborn. 1994c. RFLP evaluation of nine *Medicago* accessions representing original germplasm sources for North American alfalfa. Crop Sci. 34:230–236.

Rohlf, F.J. 1992. NTSYS-pc: Numerical taxonomy and multivariate analysis system. Version 1.70. State Univ. of New York Stony Brook, NY.

SAS Inst. 1989. SAS language guide for personal computers. Release 6.04 (ed.) SAS Inst., Cary, NC.

Tavoletti, S., F. Veronesi, and T.C. Osborn. 1996. RFLP linkage map of an alfalfa meiotic mutant based on an F1 population. J. Hered. 87:167–170.

Woodfield, D.R., and E.T. Bingham. 1995. Improvement in two-allele autotetraploid populations of alfalfa explained by accumulation of favorable alleles. Crop Sci 35:988–994.