Theor Appl Genet (1992) 85:146–151

THEORETICAL AND APPLIED GENETICS

Application of molecular markers to assess genetic relationships among accessions of wild oat, *Avena sterilis*

J.C. Goffreda*, W.B. Burnquist, S.C. Beer, S.D. Tanksley, and M.E. Sorrells

Department of Plant Breeding, Cornell University, Ithaca, NY 14853, USA

Received November 1, 1991; Accepted April 7, 1992
Communicated by P.L. Pfahler

Summary. The *Avena sterilis* collection in the National Small Grains Collection (NSGC) is an invaluable source of genetic variation to be exploited by oat breeding programs. Prior knowledge of the structure and distribution of genetic variation within the *A. sterilis* collection would be useful to efficiently screen the collection for valuable traits. To determine genetic structure within a subset of the collection, restriction fragment length polymorphisms were analyzed in a stratified sample of 173 accessions originating in eight countries of Africa and Southwest Asia. Of the 48 probes used for this study 43 detected polymorphism among accessions. The average number of RFLP patterns per probe ranged from 2.9 among Ethiopian accessions to 3.7 among those from Iran. Genetic variation, as measured by genetic distances and polymorphic indexes, was highest in Iran and lowest in Ethiopia. The probability of drawing a genotype from Iran or Iraq that is not present in the more western regions was high, indicating large genetic divergence of the Iran-Iraq accessions from the other regional collections surveyed. Cluster analysis of genetic distances and probabilities of unique genotypes clearly differentiated the eastern region (Iran and Iraq) from the western region (Algeria, Ethiopia, Israel, Lebanon, Morocco, and Syria). The western region could be further subdivided into two clusters, an African cluster (Algeria, Ethiopia, and Morocco) and a southwestern Asia cluster (Israel, Lebanon, and Syria). Genetic distances were generally related to but not proportional to geographical distances.

Key words: *Avena sativa* – Genetic diversity – Genetic structure – Restriction fragment length polymorphisms (RFLPs)

* *Present address:* Department of Horticulture, Cook College, Rutgers University, New Brunswick, NJ 08903, USA
Correspondence to: J.C. Goffreda

Introduction

Avena sterilis, a wild hexaploid progenitor of cultivated oats, *Avena sativa*, constitutes a large reservoir of genetic diversity that is readily accessible through conventional breeding (Rajhathy and Thomas 1974). This gene pool remains under-utilized because it is difficult to identify desirable traits and evaluate the agronomic potential of individual accessions in wild species. Screening the nearly 7,000 accessions of *A. sterilis* in the National Small Grains Collections (NSGC) for agronomically useful traits would be a bewildering task. Nevertheless, *A. sterilis* is known to possess several useful traits, including high growth rates (Takeda and Frey 1977), resistance to crown rust (Martens et al. 1980; Wahl 1958, 1970), resistance to nematodes (Clamot and Rivoal 1984), herbicide resistance (Somody et al. 1984), and high protein grain (Takeda and Frey 1985).

Information on the geographic structure of genetic diversity within progenitor species is important to optimize the identification of useful alleles. A systematic grouping of *A. sterilis* accessions in the NSGC would facilitate the use of this gene pool for oat improvement by reducing the numbers of genotypes the breeder would need to evaluate and by increasing the probability of selection of desirable genotypes for crossing. Recently, restriction fragment length polymorphisms (RFLPs) have been employed to determine the genetic relationships of organellar DNA in *Avena* species (Murai and Tsunewaki 1986; Rines et al. 1988) and nuclear DNA in other plant species (e.g., Song et al. 1988; Havey and Muehlbauer 1989; Keim et al. 1989; Wang and Tanksley 1989; Miller and Tanksley 1990; Furnier et al. 1990; Lubbers et al. 1991). However, there has been no comprehensive study to determine the nature of nuclear genetic diversity within *A. sterilis*.

The primary objective of the research presented here was to determine the genetic relationships among *A. sterilis* accessions from a stratified random sample of 173 accessions in the NSGC by analysis of nuclear RFLPs.

Materials and methods

Plant materials

Seedlings from 173 *Avena sterilis* accessions in the National Small Grains Collection (NSGC) were grown in the greenhouse for DNA extraction. These accessions represent a stratified sample from eight countries, Morocco, Algeria, Ethiopia, Israel, Lebanon, Syria, Iraq, and Iran. All accessions were originally obtained by the NSGC from the Canadian Department of Agriculture's collection of wild species in the genus *Avena* (CAV collection).

DNA extractions

Approximately 15 g of fresh seedling leaf tissue was ground to a fine powder in a chilled mortar and pestle using liquid nitrogen with a trace of glass beads. Ground tissue was stored at −80 °C. DNA was extracted from the frozen tissue with urea-phenol extraction buffer (Shure et al. 1983) using the method described by McCouch et al. (1988). The DNA samples were RNased (to 33 ng/ml) and re-precipitated with 0.1 volume 3 *M* sodium acetate and 2 volumes of chilled 95% ethanol. The precipitated DNA was immediately hooked out with a glass pipette, redissolved in TE, and quantified on a spectrophotometer.

Construction of partial genomic library

Total plant DNA of *Avena sativa* (cv 'Brooks') was restricted with *Pst*I and size selected (1–2 kb) on a 10–40% sucrose gradient. The selected fragments were ligated into dephosphorylated PGem4z plasmid. DH5-alpha bacterial cells were then transformed with the ligated plasmids and subsequently plated on LB media containing X-gal and IPTG. White colonies carrying insert DNA were selected, and plasmid DNA was isolated using the miniprep procedure decribed by Wilimzig (1985). Plasmid DNAs were digested with *Pst*I, electrophoresed, and Southern blotted. Low-copy insert DNAs were identified by hybridization of the plasmid filters with labelled 'Brooks' total DNA and rice chloroplast (cp) DNA. Clones containing inserts which did not produce any signal in the autoradiographs were selected. These low-copy clones were further selected by labelling the entire plasmid and hybridizing it to small survey filters. Filters were washed to 0.5 × SSC at 65 °C. Out of 200 clones carrying insert DNA, 48 low-copy clones which produced well-defined bands in the autoradiographs were selected for this study.

Restriction analysis and Southern hybridizations

Approximately 7–10 µg of total DNA from each accession was digested with *Eco*RI and electrophoresed in 0.9% agarose horizontal slab gels. Gels were Southern blotted and hybridized as described by Bernatzky and Tanksley (1986). The filters were hybridized with the selected random-hexamer labelled plasmids overnight, washed to 0.5 × SSC at 65 °C, and exposed to X-ray film (Kodak XAR-5) with a single intensifying screen for 5–7 days. The films were scored for the presence or absence of each individual band.

Genetic analysis

An estimate of mean genetic diversity within each geographic region was obtained using a modification of Nei's adjusted polymorphic index (H_{ik}) (Nei and Roychoudhury 1974). Because *Avena sterilis* is hexaploid, many probes hybridized to loci on sets of homoeologous chromosomes, and thus allelic RFLPs could not be determined. Therefore, the adjusted polymorphic index, H_{ik}, was computed for each probe as $H_{ik}=[n_k/(n_k-1)](1-\Sigma_j P_{ijk}^2)$, where n_k is the number of accessions from region k, and P_{ijk} is the frequency of restriction fragment pattern j with probe i in region k. The mean adjusted polymorphic index, H_k, was defined as the sum of the H_{ik}'s for each probe ($i=1,2,\ldots,L$) averged over the number of probes ($H_k=(1/L)\Sigma_i H_{ik}$).

Another estimate of genetic diversity was calculated based on Nei and Li's DNA divergence (Nei 1987; *d*, eq. 5.55). This is an estimate of the mean number of nucleotide substitutions per nucleotide site. DNA divergence was calculated with the program HyperRFLP (ZW Wang and SD Tanksley personal communication) for each pairwise combination of the 173 accessions.

Hedrick's probability of drawing a unique genotype (Hedrick 1971), defined as restriction fragment banding pattern for a particular probe, was used to describe genetic differences between two geographic regions. This probability is defined as $U_{km}=(1/L)\Sigma_i\Sigma_j P_{ijk}$ when $P_{ijm}=0$, where L is the number of probes, P_{ijk} is the frequency of restriction fragment pattern j with probe i in region k, and P_{ijm}, the frequency of restriction fragment pattern j with probe i in region m, is zero. U_{km} is the probability of drawing a genotype in region k not present in region m, averaged over all probes.

Cluster analyses were performed on a matrix of Nei's DNA divergence averaged within countries (8 × 8) and on the matrix of Hedrick's probability of drawing a unique genotype, using the average linkage option of SAS's Cluster procedure (SAS 1985). Principal coordinate analysis based on Nei's DNA divergence (173 × 173 matrix) was performed using SAS's Princomp procedure.

Results

Of the 48 probes examined in this study, 43 probes were polymorphic, producing between 1 and 11 distinct fragments and between 2 and 36 restriction fragment banding patterns. A total of 230 fragments were scored with an average of 4.8 fragments produced per probe. A total of 379 different restriction fragment patterns were observed over the 48 clones with an average of 7.9 patterns per probe.

Genetic diversity, as measured by the mean polymorphic index (H_k), ranged from 0.31 in Ethiopia to 0.45 in Iran and, as measured by Nei's mean genetic distance for all pairs of accessions from a single country, ranged from 0.10 in Ethiopia to 0.16 in Iran (Table 1). These two diversity parameters were highly correlated ($R^2=0.98$). The average number of RFLP patterns per probe ranged from 2.9 in Ethiopia to 3.7 in Iran. Accessions from Iran possessed the greatest genetic diversity as measured by these three parameters. Iraq and Israel also showed higher levels of genetic variation than the other countries. Ethiopia possessed the least genetic variation.

Iran and Iraq were the countries in which there were the greatest probabilities of drawing a genotype that was

absent from some other country. These mean probabilities, averaged over probes and countries, were 0.223 and 0.189, respectively (Table 2). However, the probabilities of drawing a genotype from Iran that was not present in Iraq and vice-versa were low (0.151 and 0.080, respectively), indicating that accessions from Iran and Iraq are genetically closely related. There was a high probability of drawing a genotype from other countries that was not present in Iran (mean of 0.199) or Iraq (mean of 0.163), also suggesting considerable genetic divergence of the Iran-Iraq region from the more western regions surveyed (Algeria, Ethiopia, Israel, Lebanon, Morocco, Syria). The large genetic divergence of the Iran-Iraq region from the western regions produces an upward bias in the mean Hedrick's probability when averaged over countries.

Iran also had the greatest proportion of accessions with unique fragment banding patterns (0.075, averaged over probes) (Fig. 1). Algeria had the lowest proportion of unique genotypes (0.008), but this could be due in part to the smaller sample from this country (16 versus 24 accessions). The probability that a country possesses a restriction fragment pattern present in another country, averaged over probes and countries, was greatest for Israel (0.901) and lowest for Iran (0.801) (Fig. 2). Israeli accessions had the greatest probability of possessing restriction fragment patterns common to other regions. When Rezai (1977) clustered *A. sterilis* accessions by a combination of disease resistance, chemical composition, and agronomic traits, Israeli accessions occurred in all major clusters.

Dendrograms constructed from the cluster analyses of the matrix of Nei's DNA divergence averaged within countries (8 × 8) and the matrix of Hedrick's probability of drawing a unique genotype are presented in Figs. 3 and 4. Both dendrograms illustrate that Iran and Iraq (eastern region) are genetically divergent from the six western countries examined. Two clusters can be differentiated within the western region, a cluster of African countries (Algeria, Ethiopia, and Morocco) and a cluster of Southwest Asian countries (Israel, Lebanon, and Syria). Principal coordinate analysis of Nei's DNA divergence matrix (173 × 173) indicated that 60% of the variation could be explained by the first principal axis, which separated Iran and Iraq from the more western regions (Fig. 5). The second principal coordinate, which accounted for almost one-fourth of the remaining variation, loosely divided the western region into a cluster of African countries and a cluster of Southwest Asian countries. The remaining variation was relatively homogeneously distributed.

Table 1. Number of accessions, average number of RFLP patterns, mean polymorphic index, and mean Nei's genetic distance for each country

Region	Number of accessions	Average number of RFLP patterns/probe	Mean polymorphic index (H_k)	Mean Nei's genetic distance
Algeria	16	2.7	0.34	0.11
Ethiopia	25	2.9	0.31	0.10
Iran	24	3.7	0.45	0.16
Iraq	18	3.2	0.39	0.13
Israel	23	3.4	0.39	0.13
Lebanon	24	2.9	0.33	0.11
Morocco	24	3.4	0.34	0.12
Syria	19	3.0	0.35	0.11

Discussion

Many surveys of intraspecific diversity have been conducted using morphological and protein traits, but to

Table 2. Hedrick's probability of drawing a unique genotype between pairwise combinations of the eight countries averaged over 48 probes[a]

Country in which RFLP pattern is present	Country in which RFLP pattern is absent									
	Algeria	Ethiopia	Iran	Iraq	Israel	Lebanon	Morocco	Syria	Other region	Mean[b]
Algeria	–	0.113	0.216	0.152	0.059	0.092	0.089	0.111	0.008	0.119
Ethiopia	0.128	–	0.271	0.174	0.061	0.087	0.107	0.101	0.032	0.133
Iran	0.289	0.232	–	0.150	0.181	0.234	0.251	0.221	0.075	0.223
Iraq	0.237	0.230	0.080	–	0.176	0.207	0.194	0.201	0.030	0.189
Israel	0.139	0.146	0.193	0.194	–	0.092	0.138	0.091	0.020	0.142
Lebanon	0.124	0.108	0.210	0.159	0.056	–	0.115	0.078	0.014	0.122
Morocco	0.117	0.121	0.210	0.157	0.085	0.128	–	0.099	0.042	0.131
Syria	0.127	0.111	0.213	0.158	0.076	0.093	0.107	–	0.027	0.126

[a] Genotype is defined as restriction fragment pattern for a particular probe
[b] Mean calculation excludes the probabilities of a country with itself

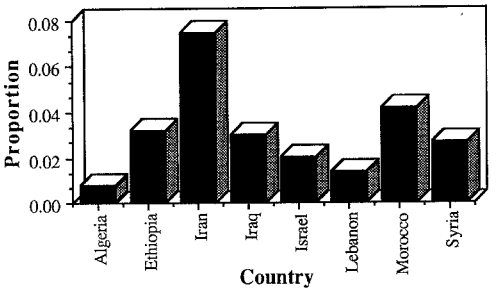

Fig. 1. Proportion of accessions with fragment banding patterns unique to that country averaged over probes

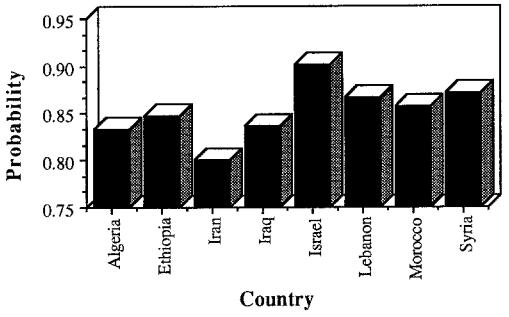

Fig. 2. Probability that a country possesses at least 1 accession with a restriction fragment pattern drawn from another country averaged over probes and countries

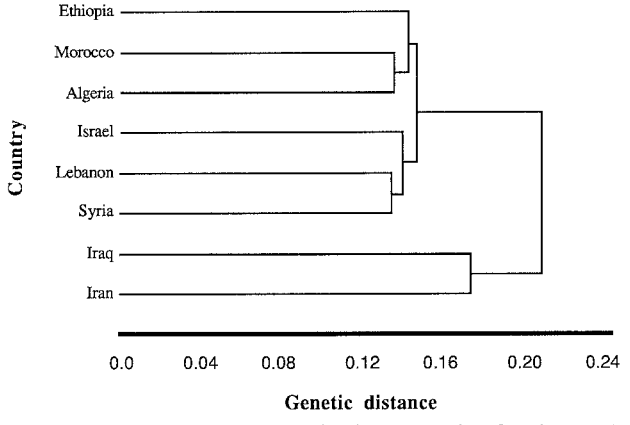

Fig. 3. Dendrogram constructed using a matrix of Nei's DNA divergence among the eight countries

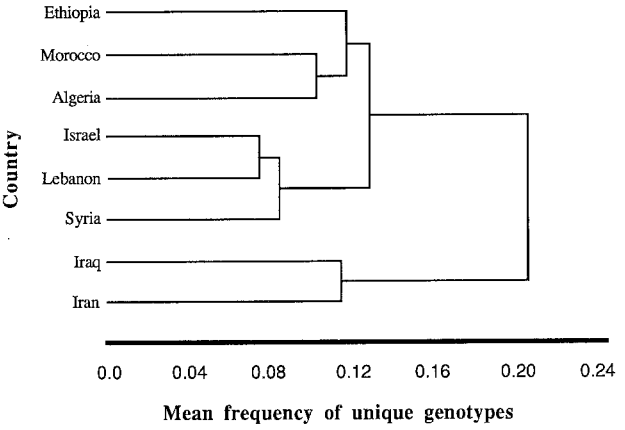

Fig. 4. Dendrogram constructed using a matrix of Hedrick's probabilities of drawing a unique genotype among the eight countries

date few have employed RFLPs. While intraspecific polymorphism has been detected in mitochondrial and chloroplast DNAs (e.g., Chowdhury and Smith 1988; Baum and Bailey 1989), the level has generally been inadequate to discriminate patterns of diversity, as shown in this survey. Holwerda et al. (1986) did find sufficient mtDNA and cpDNA polymorphism within two species of *Hordeum* to discern geographic patterns of diversity. Rines et al. (1988) included only 2 accessions of *A. sterilis* in their survey of mtDNA variability in *Avena*, but these showed different banding patterns in three of the five digests. One limitation of plastid DNA in studies of intraspecific relationships is its inability to reveal introgression (e.g., Furnier et al. 1990).

In the present study of nuclear RFLPs, 88% of the probes detected polymorphism with one restriction enzyme. Among 102 accessions of *Triticum tauschii*, Lubbers et al. (1991) found polymorphism at 80% of the loci using a single restriction enzyme. Among 70 *Oryza sativa* cultivars, Wang and Tanksley (1989) detected polymorphism with all 10 probes by restricting the DNA with five different enzymes. Considerably lower levels of nuclear RFLPs were found in *Glycine* subgenus *Soja* (Keim et al. 1989).

In digests of nuclear DNA from inbred diploids, or of cpDNA or mtDNA, the mutational events underlying RFLPs can sometimes be inferred, especially if different digests are compared. In the present study of a hexaploid species, which employed a single restriction enzyme, the number or kind of change(s) which had produced the observed variants were ambiguous. Therefore, it is uncertain how well Nei's measure of DNA divergence estimates the number of base-pair substitutions per nucleotide site. The second approach that we used, condensing all of the fragments hybridizing to a probe to a single banding pattern, may result in considerable loss of information. Nevertheless, these two approaches produced similar results, which increases our confidence that either is useful for elucidating patterns of diversity. Surveying the diploid species, *Triticum tauschii*, Lubbers et al. (1991) found that clustering and principal coordinate analysis, both based on fragment banding patterns and Hedrick's measure of genotype identity, produced patterns of diversity in general agreement with those found in morphological and isozyme studies. Song et al. (1988) investigated relationships within and among *Brassica* species, selecting phylogenetic trees based on parsimony

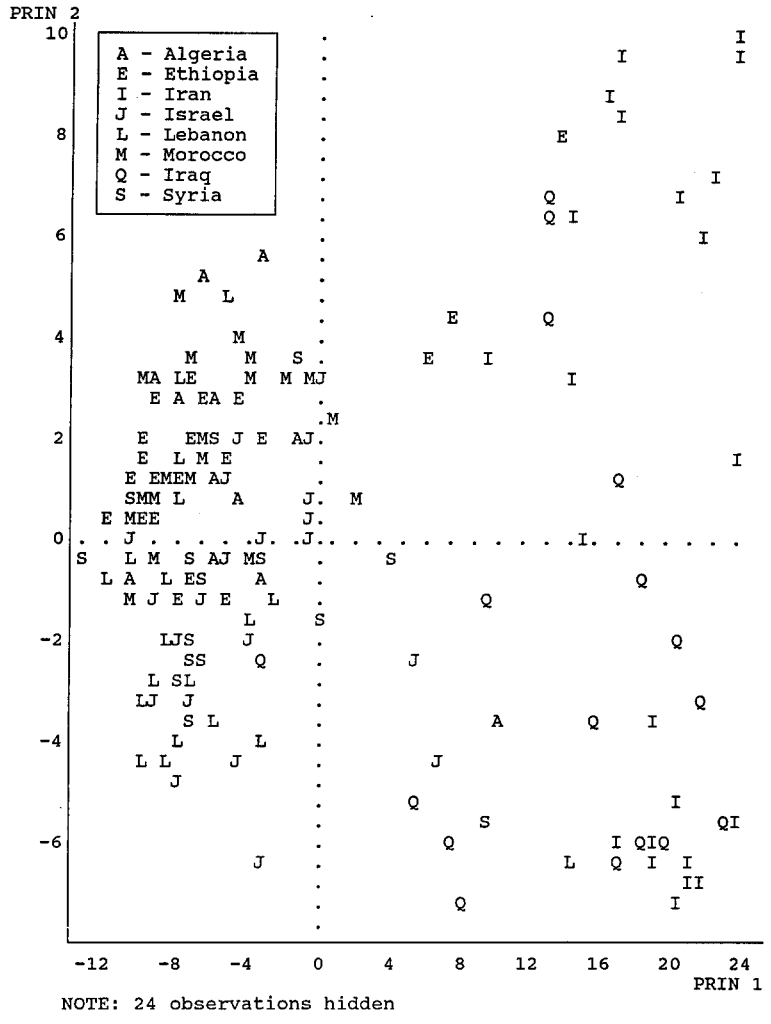

Fig. 5. Plot of the first versus the second principal coordinate scores based on the analysis of Nei's DNA divergence matrix

analysis of restriction fragments. These trees did not correspond closely to the total number of dissimilar fragments between pairs of accessions, but they did correspond to morphological and geographic groupings. Those authors attempted to use fragment patterns, but they found that those generated by different probe-enzyme combinations clustered accessions in different ways.

Based upon cytological evidence, Rajhathy and Thomas (1974) proposed that hexaploid oats evolved from the *magna-murphyi* complex of tetraploids. Since the *magna-murphyi* complex is unique to the Western Mediterranean region, Rajhathy and Thomas (1974) considered this region to be the best candidate for the center of origin of the hexaploid oats. Accessions from Morocco and Algeria, representative of that region in the present study, were not particularly diverse by any of the measures we employed. For this species, the probable center of origin is not a center of diversity. It is possible that greater variation in other parts of its range (i.e., Iran and Israel) resulted from evolution under environmentally heterogeneous conditions. Nevo and Beiles (1989) found that Israel was a center of diversity in *Triticum dicoccoides* and attributed this to Israel's range of climatic and edaphic conditions. Rezai (1977) stated that four clusters of *A. sterilis* accessions from Iran corresponded to different climatic zones. Of the eight countries sampled in the present study, the sample from Iran was collected from a larger area than that from any other country.

Rezai (1977) and others have noted geographic differences in mean values for particular traits in *A. sterilis*, but also considerable overlap in ranges of values for those traits. Breeders desiring to introgress genes for large seed size, for example, might restrict their attention to accessions from Western Mediterranean countries in order to maximize the percentage of accessions with acceptably large seed. However, the divergence of germ plasm from Iraq and Iran, compared to that from Mediterranean countries, argues for inclusion of the former in surveys for any quantitative trait. Even if the percentage of acceptable accessions from that region would be small,

alleles for that trait are likely to be different from, and perhaps complementary to, those obtained from other regions. The relationship of these divergent germ plasm pools within *A. sterilis* to that of *A. sativa* remains an interesting question.

Acknowledgements. We would like to thank N. Vorsa, R. Ortiz, and P. Smouse for their valuable comments on this manuscript.

References

Baum BR, Bailey LB (1989) Species relationships in the *Hordeum murinum* aggregate viewed from chloroplast DNA restriction fragment patterns. Theor Appl Genet 78:311–317

Bernatzky R, Tanksley SD (1986) Majority of random cDNA clones correspond to single loci in the tomato genome. Mol Gen Genet 203:8–14

Chowdhury MKU, Smith RL (1988) Mitochondrial DNA variation in pearl millet and related species. Theor Appl Genet 76:25–32

Clamot G, Rivoal R (1984) Genetic resistance to cereal cyst nematode *Heterodera avenae* Woll. in wild oat *Avena sterilis* I. 376. Euphytica 33:37–32

Furnier GR, Cummings MP, Clegg MT (1990) Evolution of the avocados as revealed by DNA restriction fragment variation. J Hered 81:183–188

Havey MJ, Muehlbauer FJ (1989) Variability for restriction fragment lengths and phylogenies in lentil. Theor Appl Genet 77:839–843

Hedrick PW (1971) A new approach to measuring genetic similarity. Evolution 25:276–280

Holwerda BC, Jana S, Crosby WL (1986) Chloroplast and mitochondrial DNA variation in *Hordeum vulgare* and *Hordeum spontaneum*. Genetics 114:1271–1291

Keim P, Shoemaker RC, Palmer RG (1989) Restriction fragment length polymorphism diversity in soybean. Theor Appl Genet 77:786–792

Lubbers EL, Gill KS, Cox TS, Gill BS (1991) Variation of molecular markers among geographically diverse accessions of *Triticum tauschii*. Genome 34:354–361

Martens JW, McKenzie RIH, Harder DE (1980) Resistance to *Puccinia gramminis avenae* and *P. coronata avenae* in the wild and cultivated *Avena* populations of Iran, Iraq, and Turkey. Can J Genet Cytol 22:641–649

McCouch SR, Kochert G, Yu ZH, Wang ZY, Khush GH, Coffman WR, Tanksley SD (1988) Molecular mapping of rice chromosomes. Theor Appl Genet 76:815–829

Miller JC, Tanksley SD (1990) Analysis of phylogenetic relationships and genetic variation in the genus *Lycopersicon*. Theor Appl Genet 80:437–448

Murai K, Tsunewaki K (1986) Phylogenetic relationships between *Avena* species revealed by the restriction endonuclease analysis of chloroplast and mitochondrial DNAs. In: Lawes DA, Thomas H (eds) Proc 2nd Int Oats Conf. Martinus Nijhoff Publ, Dordrecht, The Netherlands, pp 34–38

Nei M (1987) Molecular evolutionary genetics. Columbia University Press, New York

Nei M, Roychoudhury, AK (1974) Sampling variance of heterozygosity and genetic distance. Genetics 76:379–390

Nevo E, Beiles A (1989) Genetic diversity of wild emmer wheat in Israel and Turkey: structure, evolution, and application to breeding. Theor Appl Genet 77:421–455

Rajhathy T, Thomas H (1974) Cytogenetics of oats (*Avena* L.). The Genetics Society of Canada, Misc Publ No 2

Rezai A (1977) Variation for some agronomic traits in the world collection of wild oat (*Avena sterilis* L.). PhD thesis, Iowa State University, Ames, Iowa

Rines HW, Gengenbach BG, Boylan KL, Storey KK (1988) Mitochondrial DNA diversity in oat cultivars and species. Crop Sci 28:171–176

SAS Institute Inc (1985) SAS users' guide: statistics, version 5. SAS Institute Inc, Cary, N.C.

Shure M, Wessler S, Federoff N (1983) Molecular identification of the *Waxy* locus in maize. Cell 35:225–233

Somody CN, Nalewaja JD, Miller SD (1984) Wild oat (*Avena fatua*) and *Avena sterilis* morphological characteristics and response to herbicides. Weed Sci 32:353–359

Song KM, Osborn TC, Williams PH (1988) *Brassica* taxonomy based on nuclear restriction fragment length polymorphisms (RFLPs) 2. Preliminary analysis of subspecies within *B. rapa* (syn. *campestris*) and *B. oleracea*. Theor Appl Genet 76:593–600

Takeda K, Frey KJ (1977) Growth rate inheritance and association with other traits in backcross populations of *Avena sativa* × *A. sterilis*. Euphytica 26:309–317

Takeda K, Frey KJ (1985) Simultaneous selection for grain yield and protein percentage in backcross populations from *Avena sterilis* × *A. sativa* matings by using the independent culling levels procedure. Theor Appl Genet 69:375–382

Wahl I (1958) Studies on crown rust and stem rust on oats in Israel. Bull Res Counc Isr Sec D 6:145–166

Wahl I (1970) Prevalence and geographic distribution of resistance to crown rust in *Avena sterilis*. Phytopathology 60:746–749

Wang ZY, Tanksley SD (1989) Restriction fragment length polymorphism in *Oryza sativa* L. Genome 32:1113–1118

Wilimzig M (1985) LiCl-boiling method for plasmid minipreps. Trends Genet 1:158