Theor Appl Genet (1997) 95:400–407

Vipa Hongtrakul · Gordon M. Huestis
Steven J. Knapp

Amplified fragment length polymorphisms as a tool for DNA fingerprinting sunflower germplasm: genetic diversity among oilseed inbred lines

Received: 10 June 1996 / Accepted: 4 April 1997

Abstract Amplified fragment length polymorphism (AFLP) analysis is a rapid and efficient method for producing DNA fingerprints. The AFLP diversity of sunflower has not been described, and much of the public germ plasm of sunflower has not yet been fingerprinted. Our objectives were to: (1) estimate genetic similarities, polymorphism rates, and polymorphic information contents (PICs) for AFLP markers among elite public oilseed inbred lines, and (2) assess the genetic diversity of inbred lines using genetic similarities estimated from AFLP fingerprints. We produced fingerprints for 24 public inbred lines of sunflower (*Helianthus annuus* L.) using six AFLP primer combinations. These primers produced a total of 359 AFLP markers or about 60 markers per primer combination. Genetic similarities ranged from 0.70 to 0.91, polymorphism rates ranged from 7 to 24%, and PICs ranged from 0.0 to 0.5. Genetic similarities were lower overall for maintainer (B) × restorer (R) crosses than for B × B or R × R crosses. Principal-coordinate and cluster analyses separated lines into two groups, one for B-lines and another for R-lines. These groupings illustrate the breeding history and basic heterotic pattern (B × R) of sunflower and the widespread practice of using B × B and R × R crosses to develop new lines. There were, nevertheless, distinct subgroups within these groups. These subgroups may represent unique heterotic groups and create a basis for formally describing heterotic patterns in sunflower.

Key words DNA fingerprinting · AFLPs · RFLPs · *Helianthus annuus* L.

Communicated by G. Wenzel

V. Hongtrakul · G. M. Huestis · S. J. Knapp (✉)
Department of Crop and Soil Science, Crop Science 451C, Oregon State University, Corvallis, OR 97331–3002, USA
Fax: +1 541 737 1334
E-mail: knapp@css.orst.edu

Introduction

The genetic diversity of germ plasm collections can be assessed through the analysis of pedigree records and DNA fingerprints. Both methods have been widely used in crop plants to identify breeding bottlenecks, reconstruct breeding histories, classify germ plasms, and describe heterotic groups and patterns (Smith et al. 1990, 1991, 1993; Smith and Smith 1992; Arias and Rieseberg 1995; Mumm and Dudley 1994; Mumm et al. 1994; Sneller 1994). Pedigree or co-ancestry analysis has been a powerful tool for describing the genetic diversity of elite soybean (*Glycine max* L.) (Sneller 1994), barley (*Hordeum vulgare* L.) (Graner et al. 1994; Melchinger et al. 1994; Tinker et al. 1993), and maize (*Zea mays* L.) (Smith et al. 1990; Smith and Smith 1992; Mumm et al. 1994) germ plasm. The aim of genetic diversity analysis is to discover patterns of genetic diversity obscured by the complexities of pedigree records. Such an analysis has not been done in sunflower, a crop with complex pedigree records (Korell et al. 1992).

The ancestral relationships between sunflower (*Helianthus annuus* L.) inbred lines, wild populations, and land races have been studied using random amplified polymorphic DNAs (RAPDs) (Arias and Rieseberg 1995) and restriction fragment length polymorphisms (RFLPs) (Berry et al. 1995; Genztbittel et al. 1995). Using two sets of elite oilseed inbred lines, Berry et al. (1995) and Genztbittel et al. (1995) found lines to be strongly separated into maintainer (B) and restorer (R) groups. These groups reflect the breeding history of sunflower. Elite sunflower germ plasm has been funneled through at least three major breeding bottlenecks, one for oilseed traits (e.g., achene morphology and kernel oil percentage), one for self-compatibility and self-pollination, and another for hybrid seed-production traits (e.g., fertility restoration and maintenance and branching) (Miller 1987; Korell et al.

1992). New inbred lines are most often produced from R × R or B × B crosses, as opposed to B × R or elite × exotic crosses, to maintain heterosis and traits essential for hybrid seed production (Miller 1987).

One of the practical uses of genetic-diversity analysis in maize has been to describe heterotic groups and patterns (Lee et al. 1989; Melchinger et al. 1990; Smith et al. 1990, 1991; Dudley et al. 1991; Bernardo 1992; Messmer et al. 1992). A heterotic group is a collection of closely related inbred lines. The co-ancestries within a heterotic group are usually high, whereas the co-ancestries between two heterotic groups comprising a heterotic pattern are usually low. The classification of heterotic groups and patterns in maize has been done using DNA fingerprint and pedigree analysis, experience, and single-cross hybrid performance. Formal heterotic groups, apart from maintainer and restorer groups *per se* (Berry et al. 1995; Genztbittel et al. 1995), have not been described in sunflower.

Genetic markers and maps have, until recently, been lacking in sunflower. Berry et al. (1994) and Gentzbittel et al. (1994) were the first to describe RFLP markers for sunflower. They reported polymorphism rates ranging from 20 to 45% among 41 *H. annuus* inbred lines. These lines were mostly elite B- and R-lines. This work laid the foundation for the first RFLP map of cultivated sunflower (Berry et al. 1995).

Amplified fragment length polymorphisms (AFLPs) have emerged as powerful tools for DNA fingerprinting and genetic mapping (Zabeau 1993). AFLP marker polymorphisms are produced when restriction-site differences exist between two DNA sources. One of the strengths of AFLPs is the sheer number of markers produced per assay (Thomas et al. 1995; Vos et al. 1995). The complexity (number of fragments) of AFLP fingerprints can be manipulated by increasing or decreasing the number of selective bases and changing base composition (Zabeau 1993). The AFLP diversity of sunflower has not been described. We completed a DNA fingerprinting study in sunflower using a random sample of AFLP markers. Our objectives were to: (1) estimate genetic similarities, polymorphism rates, and polymorphic information contents for AFLP markers among inbred lines, and (2) assess the genetic diversity of inbred lines using genetic similarities estimated from AFLP fingerprints.

Materials and methods

Plant materials, DNA extraction, and AFLP assays

Twelve restorer (R) and 10 maintainer (B) inbred lines, one genetic stock (P-21), and one cytoplasmic-genic male-sterile (A) line, cmsHA822, were used in this study (Table 1). Young leaves were harvested from green house-grown plants and frozen at −80°C. Leaf tissue was ground by hand to a fine powder in liquid nitrogen using quartz sand as an abrasive. Total DNA was extracted using a modified CTAB procedure (Webb and Knapp 1990).

AFLP fingerprints were produced for each line using protocols described by Zabeau (1993) and Vos et al. (1995). One-half Microgram of genomic DNA was digested with *Eco*RI and *Mse*I in

Table 1 Types, pedigrees, and sources of cytoplasmic-genic male-sterile (A), maintainer (B), and restorer (R) inbred lines and genetic stocks (G) of sunflower (*H. annuus* L.) used in the AFLP fingerprinting study

Line	Type	Pedigree	Source
RHA265	R	2*Peredovik/Texas Wild (953-102-1-1-41)	Korell et al. (1992)
RHA271	R	PI343765/HA119//HA62-4-5/2/T-66006-2-1-31-1	Korell et al. (1992)
RHA273	R	PI343765/HA119//HA62-4-5/2/T-66006-2	Fick et al. (1975)
RHA274	R	PI343765/HA119//HA62-4-5/2/T-66006-2	Fick et al. (1975)
RHA294	R	Multiple Source Open-Pollinated Population	Miller et al. (1983)
RHA358	R	RHA274*3/DDR	Miller and Gulya (1989)
RHA365	R	Select	Miller and Gulya (1990)
RHA373	R	RHA274/82-ROM-R31	Miller (1992)
RHA374	R	ARG-R43	Miller (1992)
RHA377	R	RHA299//SOREM-HT-58/RHA801	Miller (1992)
RHA801	R	Multiple Source R-Line Population	Roath et al. (1981)
RHA858	R	P1161/RHA298	Roath et al. (1987)
HA89	B	VNIIMK 8931	Korell et al. (1992)
HA124	B	VNIIMK 8883	Korell et al. (1992)
HA234	B	2*SMENA//HA6/HA8	Korell et al. (1992)
HA369	B	ARG-8018	Miller and Gulya (1990)
HA370	B	RK-74-198	Miller and Gulya (1990)
HA371	B	H-52	Miller and Gulya (1990)
HA372	B	H-55	Miller and Gulya (1990)
HA821	B	HA300	Roath et al. (1986)
HA822	B	HA400	Roath et al. (1986)
cmsHA822	A	HA400	Roath et al. (1986)
P21	G	2*Peredovick/CMS953-102-1-1-41	Korell et al. (1992)

One-Phor-All buffer (10 mM Tris-acetate, pH 7.5, 10 mM Mg acetate, 50 mM K acetate, and 5 mM DDT) (Pharmacia, Upsala, Sweden). *Eco*RI and *Mse*I adapters were subsequently ligated to the digested DNA fragments (the *Eco*RI adapter was 5′ biotinylated). Fragments containing *Eco*RI biotinylated ends were subtracted from the reaction mixture using streptavidin beads, thereby reducing the number of fragments. Sequences of the adapters and adjacent restriction sites served as primer binding sites for amplifying the selected fragments.

A specific population of fragments was amplified from the reaction mixture by adding nucleotides to the 3 ends of the primers in two steps (Zabeau 1993). One nucleotide (+1 primers) was added for the first amplification step, while three nucleotides (+3 primers) were added for the second amplification step. DNA was PCR-amplified for 30 cycles using 5 µl of template DNA and +1 primers (*Eco*RI +1 and *Mse*I +1) (Table 2). The 30 cycles were run at 94°C for 30 s, 60°C for 30 s, and 72°C for 60 s. The second amplification step used +3 primers (*Eco*RI +3 and *Mse*I +3) (Table 2). The DNA template for this step was the PCR product produced by the first step. The *Eco*RI +3 primers were end-labelled with ^{33}P using T4 polynucleotide kinase. The *Mse*I +3 primers were unlabelled. DNA was amplified for one cycle at 94°C for 30 s, 65°C for 30 s, and 72°C for 60 s, then for 12 cycles with a 0.7°C annealing temperature decrease per cycle, and finally for 24 cycles at 94°C for 30 s, 56°C for 30 s, and 72°C for 60 s (Zabeau 1993; Vos et al. 1995).

The PCR products produced by the second amplification step were mixed with an equal volume of loading buffer (98% formamide, 10 mM EDTA, 0.025% xylene cyanol, and 0.025% bromophenol blue) and heated for 5 min. at 90°C. Eight-Microliter samples were loaded into pre-warmed 4.5% acrylamide gels with 7.5 M urea (standard sequencing gel) and 0.5 × TBE running buffer (0.045 M Tris borate and 0.001 M EDTA, pH 8.0). Gels were run at 50 V/cm using a constant wattage until the forward running dye (bromophenol blue) reached the end of the gel. The gels were dried and exposed to X-ray film for 1–3 days. AFLP bands ranging in length from 50 to 350 bases were scored as present (1) or absent (0).

The reproducibility of AFLP fingerprints was assessed by comparing the marker phenotypes from duplicate and replicate assays of all 24 lines based on the *Mse*I + 3-AAG primer (Table 2). Duplicate AFLP fingerprints were produced using two aliquots of one AFLP-PCR product (running separate lanes of the same AFLP-PCR product). Replicate AFLP fingerprints were produced by repeating the assays using separate aliquots of the original DNA samples.

Statistical analyses

Polymorphism rates were estimated for all possible pairs of lines by dividing the number of polymorphic bands by the total number of bands. Matches between missing bands were included in the total. The probability of a polymorphism between two randomly drawn lines (the polymorphic information content or PIC) was estimated using $PIC = \frac{\sum_{k=1}^{b} [(1 - P_p^2) + (1 - P_A^2)]_k}{b}$, where P_p is the frequency of lines in which the kth fragment was present (frequency of the amplified allele), p_A is the frequency of lines in which the kth fragment was absent (frequency of the null allele), $k = 1, 2, \ldots, b$, and $b = 359$ is the total number of AFLP bands scored. This parameter is sometimes called heterozygosity.

The genetic similarity between two inbreds was estimated using Gower's coefficient of similarity (Gower 1971). The genetic similarity between inbred i and j was estimated using

$$S_{ij} = \sum_{k=1}^{b} (w_{ijk} S_{ijk}) / \sum_{k=1}^{b} w_{ijk}$$

Table 2 Oligonucleotide adapter and primer names and sequences for six selective amplified fragment length polymorphism primer combinations (marker assays).

Name	Sequence
*Eco*RI Adapter	5′-BIO-CTCGTAGACTGCGTACC CTGACGCATGGTTAA-5′
*Mse*I Adapter	5′-GACGATGAGTCCTGAG TACTCAGGACTCAT-5′
*Eco*RI + 1	5′-AGACTGCGTACCAATTC/C-3′
*Mse*I + 1	5′-GACGATGAGTCCTGAGTAA/A-3′
*Eco*RI + 3	5′-GACTGCGTACCAATTC/CAG-3′
*Mse*I + 3-AAG	5′-GATGAGTCCTGAGTAA/AAG-3′
*Mse*I + 3-AAT	5′-GATGAGTCCTGAGTAA/AAT-3′
*Mse*I + 3-ACG	5′-GATGAGTCCTGAGTAA/ACG-3′
*Mse*I + 3-AGA	5′-GATGAGTCCTGAGTAA/AGA-3′
*Mse*I + 3-AGG	5′-GATGAGTCCTGAGTAA/AGG-3′
*Mse*I + 3-ATA	5′-GATGAGTCCTGAGTAA/ATA-3′

where w_{ijk} is a weight for inbred i and j and AFLP fragment k, S_{ijk} is the marker phenotype or score (present or absent) for inbred i and j and fragment k, $i = j = 1, 2, \ldots, n$, and $n = 23$ is the number of inbreds. The similiarity between two inbreds was estimated by ignoring null matches: (1) if inbred i and j shared a band, then $S_{ij} = 1$; (2) if inbred i and j did not share a band, then $S_{ij} = 0$; (3) if inbred i or j or i and j shared a band, then $w_{ij} = 1$; and (4) if inbred i and j lacked a band, then $w_{ij} = 0$.

Principal-coordinate analysis was done using the PROC PRINCOMP procedure of SAS (1992) and the genetic-similarity matrix. Cluster analysis was done using the average linkage algorithm of PHYLIP (Felsenstein 1993) and the genetic distance matrix ($D_{ij} = \sqrt{1 - S_{ij}}$). A phenogram was produced from the output of PHYLIP using TREETOOL (Maciukenas et al. 1991).

Results

Six AFLP primer combinations (Table 2) produced 359 scoreable AFLP markers (fragments) (Fig. 1). Each primer combination produced about 60 strongly amplified and scoreable fragments between 50 and 350 bp in size (Fig. 1). There were no scoring discrepancies between duplicate and replicate AFLP fingerprints produced using one primer combination (+ 3-AAG primer) (Table 2) or between AFLP fingerprints for a pair of isogenic lines (HA822 and cmsHA822).

The number of polymorphic fragments per fingerprint (primer combination) ranged from 4.2 for RHA274 × RHA358 to 14.3 for HA124 × RHA801 (Table 3). Polymorphism rates ranged from 7.0% for RHA274 × RHA358 to 23.8% for HA124 × RHA801. B × B and R × R crosses had similar minimum and maximum polymorphism rates. B × R crosses had greater minimum and maximum polymorphism rates than B × B and R × R crosses.

Roughly half (47.9%) of the AFLP fragments were polymorphic in at least one pair of lines (187 AFLP fragments were monomorphic) (Fig. 2). The PIC scores for AFLPs ranged from 0.0 to 0.5 (Fig. 2). Mean PIC scores were 0.115 for R-lines, 0.151 for B-lines, and 0.14 overall. The distribution of PIC scores was nearly

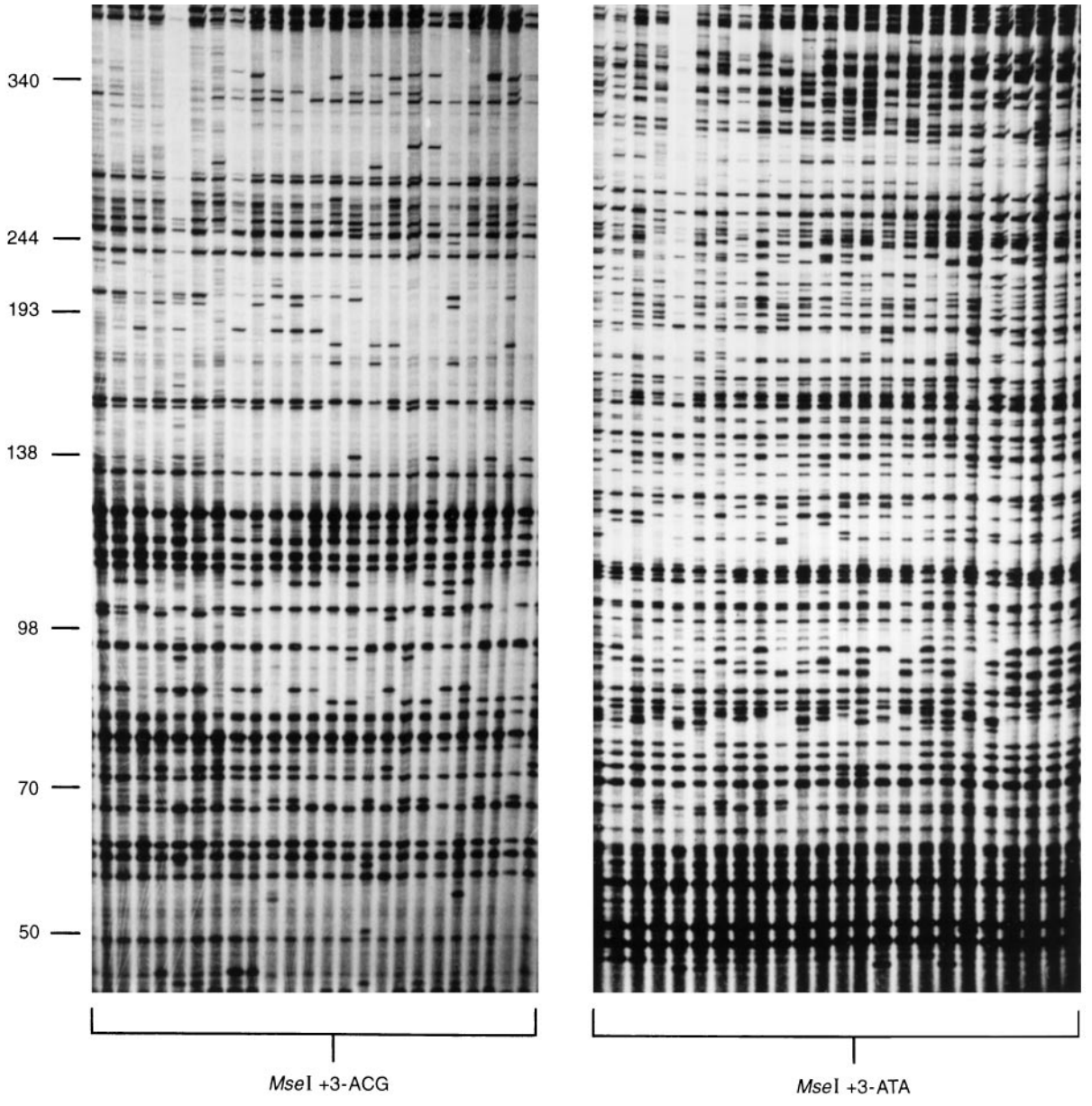

Fig. 1 AFLP fingerprints produced by primer pairs *Eco*RI + 3/*Mse*I + 3-AAG (left) and *Eco*RI + 3/*Mse*I + 3-AAT (right) for 24 inbred lines of sunflower. There is one lane per inbred line for each primer pair

uniform (random) for the 172 polymorphic AFLP markers (Fig. 2). PIC scores were maximum (0.5) for 5.6% of the AFLP markers.

Genetic similarities between lines ranged from 0.70 to 0.91 (Fig. 3). The similarities between B × R crosses tended to be lower than between B × B and R × R crosses. Principal-coordinate and cluster analyses separated lines into two major groups, one comprised of B-lines and one comprised of R-lines (Figs. 4–6). The first three principal coordinates accounted for 34% of the genetic-similarity variance. The phenogram (Fig. 4) and principal-coordinate maps (Figs. 5, 6) show the groups found with both methods.

Lines within these two groups were not completely homogeneous (Figs. 4–6). There were at least two distinct B-line subgroups: HA852 and HA89 formed one subgroup (subgroup B_1), while the remaining B-lines formed a second more-dispersed subgroup (subgroup B_2). There were four distinct R-line subgroups: (1) RHA858, RHA271, RHA273, RHA801, RHA265, RHA374, RHA377, and RHA858 (subgroup R_1); (2) RHA274, RHA373; and RHA358 (subgroup R_2); (3) RHA294 (Subgroup R_3); and (4) RHA365 (subgroup R_4). RHA365 was more similar to certain B-lines (e.g.,

Table 3 Minimum and maximum polymorphism rates (rate) and number of polymorphic markers per AFLP primer combination (number) for 359 AFLP markers among 22 maintainer (B) or restorer (R) inbred lines of sunflower

Group	Statistic	Cross	Rate (%)	Number
B × B	Minimum	HA89 × HA852	7.5	9.5
	Maximum	HA234 × HA369	21.8	13.1
R × R	Minimum	RHA274 × RHA358	7.0	4.2
	Maximum	RHA358 × RHA365	19.8	11.8
B × R	Minimum	HA370 × RHA265	13.4	7.8
	Maximum	HA124 × RHA801	23.8	14.3

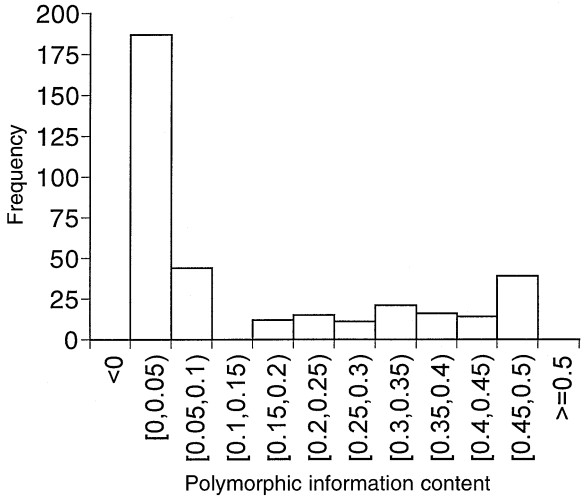

Fig. 2 Distribution of polymorphic information content scores for 359 AFLP markers among 23 inbred lines of sunflower

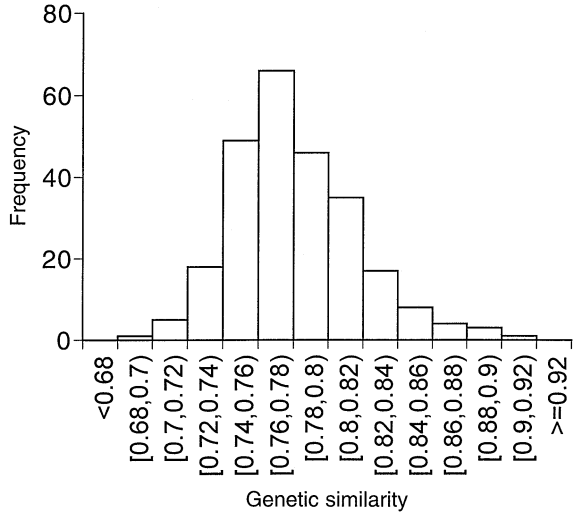

Fig. 3 Distribution of genetic similarities among 23 inbred lines of sunflower estimated from 359 AFLP markers

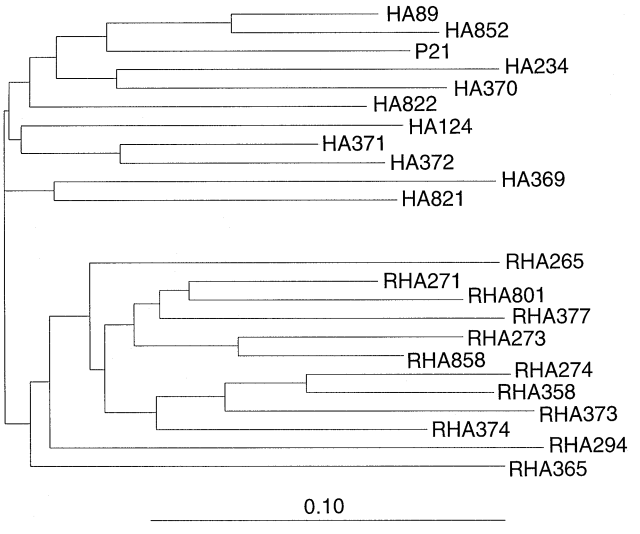

Fig. 4 Phenogram produced by cluster analysis of the genetic distance matrix estimated using 359 AFLP markers and 23 inbred lines of sunflower

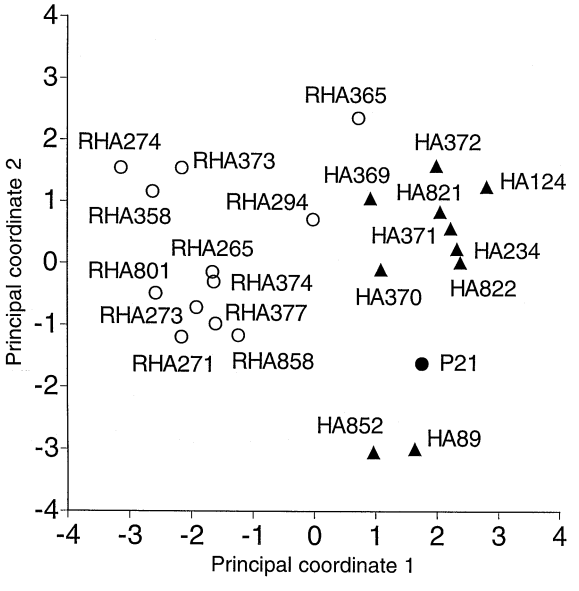

Fig. 5 Principal-coordinate map for the first and second principal coordinates estimated for 359 AFLP markers using the genetic similarity matrix for 23 inbred lines of sunflower

HA369 and HA372), than to most of the other R-lines (Figs. 5, 6). Subgroups R_1, R_2, and R_3 were nearly equally separated. The greatest separation was between subgroups R_1 and R_4 (Figs. 5, 6).

RHA274 ranked as the first or second most-polymorphic line for 70% of the B-lines, while RHA365 was the first or second most polymorphic line for 40% of the B-lines (Table 4). RHA274 was most polymorphic with HA234, HA371, HA372 and HA822, while

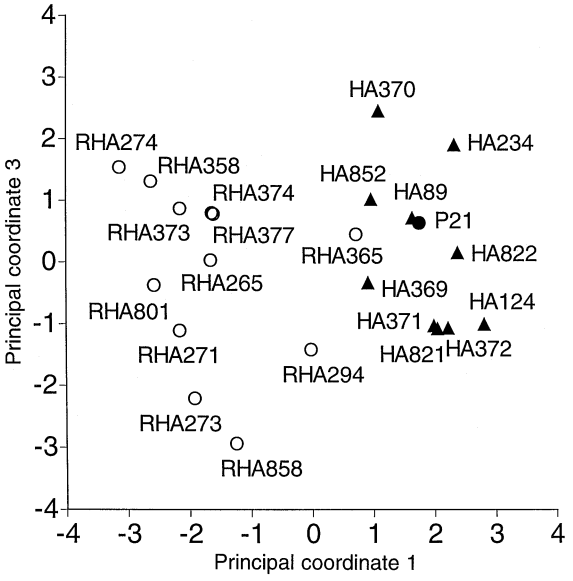

Fig. 6 Principal-coordinate map for the first and third principal coordinates estimated for 359 AFLP markers using the genetic similarity matrix for 23 inbred lines of sunflower

Table 4 The five most-polymorphic crosses and the number of polymorphic fragments (shown in parentheses) per cross for 359 AFLP markers and ten B-lines of sunflower

	Rank of Cross				
Line	1	2	3	4	5
HA89	RHA365 (72)	RHA274 (69)	RHA373 (66)	RHA294 (64)	RHA358 (64)
HA124	RHA801 (86)	RHA274 (84)	RHA373 (81)	RHA377 (78)	RHA358 (77)
HA234	RHA274 (83)	RHA365 (79)	HA369 (79)	RHA373 (77)	RHA273 (76)
HA370	HA369 (73)	HA372 (72)	HA124 (71)	RHA294 (70)	RHA365 (70)
HA371	RHA274 (67)	RHA373 (65)	RHA377 (65)	HA369 (64)	RHA801 (64)
HA372	RHA274 (74)	HA370 (72)	RHA374 (72)	HA852 (71)	HA369 (69)
HA369	RHA365 (85)	RHA377 (82)	HA234 (79)	RHA801 (74)	HA370 (73)
HA821	RHA377 (79)	RHA274 (74)	RHA358 (68)	RHA801 (68)	RHA365 (67)
HA822	RHA274 (73)	HA369 (72)	RHA377 (71)	RHA373 (71)	RHA801 (66)
HA852	RHA365 (77)	HA124 (72)	HA372 (71)	RHA274 (70)	RHA373 (67)

RHA365 was most polymorphic with HA89, HA369 and HA852 (Table 4). RHA377 was most polymorphic with HA821, while RHA801 was most polymorphic with HA124 (Table 4). HA370 was unusual among B-lines: three B-lines (HA369, HA372, and HA124) were most polymorphic with this line (Table 4).

Discussion

AFLPs are a powerful tool for fingerprinting inbred lines, producing genetic maps, and marker-assisted selection in crop plants (Vos et al. 1995). They have virtually eliminated the DNA-marker bottleneck in sunflower. This bottleneck persisted for many years, impeded the use of markers in sunflower breeding, and also impeded the development of the genetic map of sunflower. The percentage of polymorphic AFLP fragments in our study was similar to the percentage of polymorphic RFLP probes reported by Berry et al. (1994), and Gentzbittel et al. (1994). Fourteen percent of the genomic clones and 61.1% of the cDNA clones tested by Gentzbittel et al. (1994), and 48.1% of the genomic clones (ignoring redundant clones) and 47.2% of the cDNA clones tested by Berry et al. (1994) were polymorphic in at least one pair of lines, whereas 47.9% of the AFLP fragments were polymorphic in at least one pair of lines in our study (Fig. 2). AFLPs, however, had lower PIC scores than RFLPs. Berry et al. (1994) reported a mean PIC of 0.49 for a selected set of 57 RFLP probes (185 RFLP bands). This is significantly greater than the mean PIC we estimated for AFLPs (0.14) (Fig. 2).

PIC-score differences between AFLPs and RFLPs have two primary causes (both markers detect DNA polymorphisms caused by restriction-site mutations, insertions, or deletions). The maximum PIC score for an AFLP marker (or any bi-allelic marker) is 0.5, whereas the maximum PIC score for an RFLP marker is 1.0. When an AFLP fragment is present in half and missing in half of the lines, the PIC score is 0.5. Roughly 5% of the AFLP fragments in our study had maximum PIC scores. When each line has a unique RFLP allele, the PIC score is 1.0. PIC scores this high have not been reported for RFLP markers in sunflower. The PIC score for an RFLP marker can often be increased by testing additional restriction enzymes (increasing the number of probe-enzyme combinations).

Although AFLPs have lower PIC scores than RFLPs in sunflower, they produce more polymorphic markers per assay than RFLPs and other genetic markers, require limited pre-screening, no prior development, and are sufficiently polymorphic and abundant to produce genetic maps using virtually any cross between elite inbred lines (Tables 1 and 3). AFLPs are dominant when visually scored. This is a drawback for some applications; however, densitometry can be used to estimated allele doses and the densitometric differences between one or two doses are often great enough to distinguish between homozygotes and heterozygotes (Vos et al. 1995). AFLP fragments of a specific length do not necessarily represent specific loci across genetic backgrounds, but many do. This affects some applications, but is offset by the sheer number of DNA targets accessed by AFLP technology (Vos et al. 1995).

Three genetic diversity studies have been done in sunflower using 62 elite oilseed B- or R-lines and RFLPs (Berry et al. 1995; Gentzbittel et al. 1995) or AFLPs (Figs. 4–6). B- and R-lines were strongly separated into groups in each study. These groups, as previously stated, reflect the fundamental heterotic pattern of sunflower ($B \times R$) and the widespread practice of producing new lines using $B \times B$ or $R \times R$ crosses. The more important question raised by these analyses is whether or not they highlight unique heterotic groups of sunflower and whether or not these groups reflect more than one heterotic pattern. Work on this problem in sunflower has been insufficient. Novel heterotic groups and patterns undoubtedly exist in sunflower, but have not been described.

Miller and Gulya (1990) and Miller (1992) introgressed diversity from unique germ plasm sources into elite genetic backgrounds. RHA365, which is an outlier among R-lines (Figs. 4–6), formed a unique R-line subgroup, presumably because this line was developed from a Romanian single-cross hybrid (Select) (Miller and Gulya 1990). The genetic background of this hybrid seems to be unique (Figs. 4–6). Some lines developed using apparently unique germ plasm sources, however, fell into groups with other widely used inbreds. RHA373 was developed from RHA274/82ROM-R31 and fell in the RHA274 subgroup (R_2), whereas RHA377 was developed from RHA299//Sorem/RHA801 and fell in the RHA801 subgroup (R_1) (Miller 1992).

The B_2 subgroup may be too diverse to be classified as a single heterotic group. Some of the members of the B_2 subgroup (e.g., HA370, HA371 and HA372) were developed from crosses between other members of the group and elite South African germ plasm (RK-74-198, H-52, and H-55, respectively) (Miller and Gulya 1990) and may represent unique heterotic groups. HA371 and H-52 seem to be closely related to HA821 (Figs. 4–6). HA372 and H-55 also seem to be closely related to HA821, but less so than HA371 and H-52. HA370 and HA369, a line developed from Argentinian germ plasm (ARG-8018) (Miller and Gulya 1992), seem to be unique among these lines (Figs. 4–6). Both were on the fringes of the B_2 subgroup (Figs. 4–6).

CANP3 (a B-line) was clearly separated from other B-lines in the Gentzbittel et al. (1994) study, while HA89 and HA852 (subgroup B_1) were clearly separated from the bulk of the other B-lines in our study (Figs. 4–6). Gentzbittel et al. (1994) found that PAT4 (an R-line) was clearly separated from the other R-lines they tested.

The six subgroups we proposed (B_1, B_2, R_1, R_2, R_3, and R_4) create a working model for describing heterotic groups in sunflower analogous to those found in maize (Lee et al. 1989; Melchinger et al. 1990; Smith et al. 1990; Dudley et al. 1991; Bernardo 1992; Messmer et al. 1992). The proposed subgroups can be used as a basis for relating single-cross hybrid performance to genetic similarities within and between groups; e.g., (1) are all the lines in a proposed heterotic group closely related, (2) are some $B \times R$ heterotic patterns ($B_1 \times R_1$, $B_2 \times R_1$, $B_1 \times R_2$, …) superior to others or unique in some important way, and (3) are lines from different groups comprising a heterotic pattern distantly related or unrelated? Sunflower breeders can undoubtedly describe heterotic patterns from experience; however, associations between genetic similarities and hybrid performance warrants study and should shed light on heterotic groups and patterns in sunflower.

Acknowledgements This work was funded by grants from the USDA-NRICGP (95-37300-1573) and Pioneer Hi-Bred International (Johnston, Iowa). We are grateful to Alex Kahler (Biogenetic Services, Brookings, South Dakota) for supplying DNA samples, Jerry Miller (USDA-ARS, Fargo, North Dakota) for supplying seed samples, and Mark Zabeau (Keygene, Wageningen, the Netherlands) and Richard Michelmore (U.C.-Davis, California) for supplying unpublished AFLP protocols. Oregon Agric. Exp. Stat. Tech. Rep. No. 11041.

References

Arias DM, Rieseberg LH (1995) Genetic relationships among domesticated and wild sunflower (*Helianthus annuus*, Asteraceae). Econ Bot 49:239–248

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

Berry ST, Allen RJ, Barnes SR, Caligari PDS (1994) Molecular-marker analysis of *Helianthus annuus* L. 1. Restriction fragment length polymorphism between inbred lines of cultivated sunflower. Theor Appl Genet 89:435–441

Berry ST, Leon AJ, Hanfrey CC, Challis P, Burkholz A, Barnes SR, Rufener GK, Lee M, Caligari PDS (1995) Molecular-marker analysis of *Helianthus annuus* L. 2. Construction of an RFLP map for cultivated sunflower. Theor Appl Genet 91:195–199

Dudley JW, Saghai-Maroof MA, Rufener GK (1991) Molecular markers and grouping of parents in a maize breeding program. Crop Sci 31:718–723

Felsenstein J (1993) Documentation for PHYLIP (Phylogeny Inference Package) Version 3.5C. University of Washington, Seattle

Fick GN, Kinman ML, Zimmer DE (1975) Registration of RHA 273 and RHA 274 sunflower parental lines. Crop Sci 15:106

Gentzbittel L, Zhang Y-X, Vear F, Griveau B (1994) RFLP studies of genetic relationships among inbred lines of the cultivated sunflower, *Helianthus annuus* L.: evidence for distinct restorer and maintainer germ plasm pools. Theor Appl Genet 89: 419–425

Gower JC (1971) A general coefficient of similarity and some of its properties. Biometrics 27:857–874

Graner A, Ludwig WF, Melchinger AE (1994) Relationships among European barley germ plasm. II. Comparison of RFLP and pedigree data. Crop Sci 34:1199–1205

Korell G, Mosges F, Friedt W (1992) Construction of a sunflower pedigree map. Helia 15–17:7–16

Lee M, Godshalk EB, Lamkey KR, Woodman WW (1989) Association of restriction fragment length polymorphisms among maize inbreds with the agronomic performance of their crosses. Crop Sci 29:1067–1071

Maciukenas M, Olsen G, Larsen N, Woese C (1991) Treetool – a phylogenetic data visualizer, editor, and tree formatter. Version 1.0, Ribosomal RNA Database Project, University of Illinois, Champaign-Urbana

Melchinger AE, Lee M, Lamkey KR, Hallauer AR, Woodman WL (1990) Genetic diversity for restriction fragment length polymorphisms and heterosis for two diallel sets of maize inbreds. Theor Appl Genet 80:488–496

Melchinger AE, Graner A, Singh M, Messmer MM (1994) Relationships among European barley germ plasm: I. Genetic diversity among winter and spring cultivars revealed by RFLPs. Crop Sci 34:1191–1199

Messmer MM, Melchinger AE, Lee M, Woodman WL, Lee EA, Lamkey KR (1991) Genetic diversity among progenitors and elite lines from the Iowa stiff stalk synthetic (BSSS) maize population: comparison of allozyme and RFLP data. Theor Appl Genet 83:99–107

Miller JF (1987) Sunflower. In: Fehr WR (ed) Principles of cultivar development. Vol. 2, crop species. Wiley, New York, pp. 626–668

Miller JF (1992) Registration of five oilseed sunflower germ plasm restorer lines (RHA 373 to 377) and two nuclear male-sterile populations (NMS 274 and 801). Crop Sci 32:1298

Miller JF, Gulya T (1989) Registration of seven reduced-height sunflower restorer germ plasm lines. Crop Sci 29:1332–1333

Miller JF, Gulya T (1990) Registration of ten oilseed sunflower germ plasm lines. Crop Sci. 30:430–431

Miller JF, Gulya T, Roath WW (1983) Registration of RHA 309, RHA 310, and RHA 311 sunflower germ plasm. Crop Sci 29: 1332–1333

Mumm RH, Dudley JW (1994) A classification of 148 U.S. maize inbreds. I. Cluster analysis based on RFLPs. Crop Sci 34:842–851

Mumm RH, Hubert LJ, Dudley JW (1994) A classification of 148 U.S. maize inbreds: II. Validation of cluster analysis based on RFLPS. Crop Sci 34:852–865

Roath WW, Miller JF, Gulya T (1981) Registration of HA801 sunflower germ plasm. Crop Sci 21:479

Roath WW, Miller JF, Gulya T (1986) Registration of sunflower parental lines HA821 and HA822. Crop Sci 26:217

Roath WW, Miller JF, Gulya T (1987) Registration of six oilseed sunflower parental lines. Crop Sci 27:825–826

SAS (1992) SAS-STAT Users guide vol. 2 (4th edn.). GLM-VARCOMP. Cary, North Carolina

Smith JSC, Smith OS, Bowen SL, Tenborg RA, Wall SJ (1991) The description and assessment of distances between inbred lines of maize. III. A revised scheme for the testing of distinctiveness between inbred lines utilizing DNA RFLPs. Maydica 36: 213–226

Smith JSC, Smith OS, Wright S, Wall SJ, Walton M (1993) Diversity of U.S. hybrid maize germplasm as revealed by restriction fragment length polymorphisms. Crop Sci. 32:598–604

Smith OS, Smith JSC (1992) Measurement of genetic diversity among maize hybrids: a comparison of isozymic, RFLP, pedigree, and heterosis data. Maydica 37:53–60

Smith OS, Smith JSC, Bowen SL, Tenborg RA, Wall SJ (1990) Similarities among a group of elite maize inbred as measured by pedigree, F_1 heterosis, and RFLPs. Theor. Appl. Genet. 80: 833–840

Sneller CH (1994) Pedigree analysis of elite soybean lines. Crop Sci 34:1515–1522

Thomas CM, Vos P, Zabeau M, Jones DA, Norcott KA, Chadwick BP, Jones JDG (1995) Identification of amplified fragment length polymorphism (AFLP) markers tightly linked to the tomato *Cf-9* gene for resistance to *Claosporium fulvum.* Plant J 8:785–794

Tinker NA, Fortin MG, Mather DE (1993) Random amplified polymorphic DNA and pedigree relationships in spring barley. Theor Appl Genet 85:976–984

Vos P, Hogers R, Bleeker M, Reijans M, van de Lee T, Hornes M, Frijters A, Pot J, Peleman J, Kuiper M, Zabeau M (1995) AFLP: a new technique for DNA fingerprinting. Nucleic Acids Res 23: 4407–4414

Webb DM, Knapp SJ (1990) DNA extraction from a previously recalcitrant plant genus. Mol Biol Rep 8:180–185

Zabeau M (1993) Selective restriction fragment amplification: a general method for DNA fingerprinting. European Patent Application No. 0-534-858-A1