Theor Appl Genet (1994) 89: 419–425

L. Gentzbittel · Y.-X. Zhang · F. Vear · B. Griveau
P. Nicolas

RFLP studies of genetic relationships among inbred lines of the cultivated sunflower, *Helianthus annuus* L.: evidence for distinct restorer and maintainer germplasm pools

Received: 20 September 1993 / Accepted: 2 February 1994

Abstract One-hundred-and-eighty-one nuclear DNA probes were used to examine restriction-fragment length polymorphism in inbred lines of the cultivated sunflower (*Helianthus annuus* L.). The probes were from six libraries: two genomic libraries – one made with *Pst*I and the other with *Hind*III, and four cDNA libraries – from etiolated plantlets, green leaves, ovaries, petals and anthers. Total DNA from 17 inbred lines representing an overview of the genetic stocks of sunflower, including restorer and maintainer lines of the classical cytoplasmic male sterility, was digested with four different restriction enzymes and probed in 331 probe-enzyme combinations. Of 181 clones analysed, 73 probes were found to be polymorphic. Genetic distances between inbreds were calculated from the resultant proportion of shared bands and submitted to principal component analysis and the UPGMA 'tree-making' method. The RFLP analysis allowed a clear differentiation between restorer and maintainer lines of the cytoplasmic male sterility, together with a grouping of some of the genotypes from the same origin. The analysis of the accuracy of distance estimation as a function of the number of probe-enzyme combinations used, indicates that 40–50 combinations ensure a confidence level of near 95%. Considering the inbreds as representatives of the range of cultivated inbreds, estimates of gene diversity, as well as estimates of average gene diversity between and within the sets of restorer and maintainer lines, were calculated. Estimation of gene diversity showed that the available genetic variability in cultivated sunflower, based on allelic frequencies, is lower than that of other plants (H=0.20). Moreover, we show that the proportion of genetic variability due to the difference between maintainer and restorer lines (D_m) is about 2%.

Key words Sunflower · RFLP · Genetic diversity
Cytoplasmic male sterility

Communicated by J. W. Snape

L. Gentzbittel[1] (✉)
GIE CARTISOL, 12 Avenue George V,
75008 Paris, France

Y.-X. Zhang
GIP GEVES, La Minière,
78285 Guyancourt Cedex, France

F. Vear
INRA, Station de Génétique et d'Amélioration des Plantes,
12 avenue du Brezet,
63039 Clermont-Ferrand Cedex 01, France

B. Griveau
INRA, Station de Génétique et d'Amélioration des Plantes,
Service Tournesol, Domaine de Melgueil, 34130 Maugio, France

P. Nicolas
Université Blaise Pascal-Clermont II,
Laboratoire Organisation et Variabilité des Génomes Végétaux,
4 rue Ledru, 63038 Clermont-Ferrand Cedex 01, France

Present address:
[1] GIE CARTISOL, 4 rue Ledru, 63038 Clermont-Ferrand Cedex 01, France

Introduction

The sunflower (*Helianthus annuus* L.) is one of the most important oil-seed crops in the world. It was introduced into Spain from North America for floral decoration in the 16th century (Schilling and Heiser 1981). Selection for oil-content was first carried out in Russia at the beginning of this century and, through continuous breeding efforts, inbred lines with a 50% oil content were obtained. With the discovery in France of a cytoplasmic male sterility system (Leclercq 1969), hybrid sunflowers have become predominant (except in a few countries including the Ukraine, Russia and Spain). To-date, a large number of breeding programs, using similar genetic resources derived from Russian populations, have dramatically increased the number of inbred lines, with over 400 entries described in the French reference catalogue alone.

Sunflower breeding programs have three major aims: the improvement of seed and oil yields; the improvement of yield stability, in particular by the introduction of resistance to diseases or extreme climatic conditions; and the continued research for new sources of cytoplasmic male

sterility (Serieys and Vincourt 1987). As in maize, a commercial hybrid of sunflower must be male fertile. It is produced by crossing a female parent (with cytoplasmic male sterility, called the 'A line' – the line is maintained by crossing with an isogenic male-fertile maintainer, called the 'B line' and named the 'M' line in this study) with a male parent possessing the appropriate fertility-restoration system (an 'R' line). Breeding programs appear to have kept R- and M-germplasm pools separated, in order to maximise heterosis in commercial hybrids.

Despite the relatively short history of sunflower breeding, little is known about the genetic relationships of the cultivated sunflower. As the number of inbreds described increases and the discrimination of lines by physiological and morphological characteristics becomes more difficult, breeders need more accurate identification criteria. Currently, inbred lines are only discriminated on the basis of morphological characters, although the use of isozyme markers for the same purpose has recently been described (Quillet et al 1992). DNA markers such as restriction-fragment length polymorphism (RFLPs) have been applied to interspecific studies of the genus *Helianthus* (Choumane and Heizmann 1988; Rieseberg et al. 1988, 1990; Gentzbittel et al. 1992). We present here an RFLP comparison of a representative set of sunflower inbred lines, and the estimated genetic similarities among them. Estimates of the overall genetic variability, together with the definition of gene diversity between the subset of R and M inbreds, are also presented, with the long-term goal of establishing a linkage map of sunflower which will combine molecular markers with genes for qualitative morpho-agronomic traits.

Materials and methods

Plant materials

Seedlings from 17 cultivated *H. annuus* inbred lines were grown in the field for DNA extraction. These inbreds represent the main genetic resources available in cultivated sunflower (Table 1). They originate from western and eastern Europe, North America, North Africa and Argentina and are restorer (R-line) or maintainer (M-line) inbreds of the classical male sterility (Leclercq 1969). All lines were maintained at the INRA Plant Breeding stations at Montpellier and Clermont-Ferrand (France).

DNA extractions

All leaf tissue was frozen at –80 °C and ground to a fine powder in liquid nitrogen. Total DNA was extracted from the frozen tissues with CTAB extraction buffer, as described previously (Gentzbittel et al. 1992). The precipitated DNA was hooked out with a glass pipette, redisolved in TE, and quantified in a spectrophotometer.

Construction of partial genomic and cDNA libraries

Nuclear DNA of *H. annuus* (inbred line HA89) was isolated following the method of Kiss and Solimosy (1987), and restricted with *Pst*I and *Hin*dIII. The fragments were size-selected on low-melting agarose gel (1–2 kb) and ligated into the plasmid pUC18. DH5-alpha bacterial cells were then transformed with the ligated plasmids and subsequently plated on LB containing ampicillin, X-gal and IPTG. White colonies carrying insert DNA were selected, grown overnight in liquid media, and stored at –20 °C in 50% glycerol.

Table 1 *Helianthus annuus* inbreds used for RFLP analysis

Inbred	R/M[a]	Origin[b]	Known gene pool[c]
PAT4	R	France	× *H. tuberosus*
PX7	R	Argentina	USA
RHA274	R	USA	USA
RHA345	R	USA	USA
RHA266	R	USA	USA
PAC2	R	France	× *H. petiolaris*
PB3	R	France	× *H. petiolaris*
CANP3	M	Russia	Russia
H52	M	South Africa	
2603	M	Morocco	
89B1	M		
GH	M	Romania	Romania
CX	M	France	Russia, USA, France
HA124	M	USA	Russia
HA89	M	USA	Russia
SD	M	France	Russia
Z2736.2	M	Romania	Romania

[a] R or M refers to the restorer or maintainer phenotype of the male sterility, respectively
[b] Country from which the inbreds were originally distributed
[c] Known genetic pool from which the inbred was created, with introgression source

Partial cDNA libraries were constructed from expressed sequences of green leaves, etiolated plantlets, ovaries, petals and anthers, using the Superscript kit (Gibco-BRL). The cDNAs were size-selected by gel filtration, ligated into the pSPORT plasmid vector (Gibco-BRL) and transformed into E. coli DH5-alpha as described above.

Restriction analysis and Southern hybridizations

Total DNA from each of the sunflower inbred lines was routinely digested for 5 h with four restriction endonucleases (*Eco*RI, *Hin*dIII, *Bgl*II and *Eco*RV) and then subjected to electrophoresis in TAE on 0.8% agarose horizontal slab gels. Gels were Southern blotted on Hybond-N^+ (Amersham) and hybridized following conventional methods with randomly-primed ^{32}P-labelled DNA probes, in a hybridization oven (Appligene). Filters were washed with 0.5 × SSC at 65 °C for final stringency, and exposed to X-ray films (Amersham Hyperfilm MP) with two intensifying screens for 3–5 days. The films were scored for the presence or absence of each individual band, by using the BioImage analysis system of Millipore.

Genetic and statistic analyses

For each probe-enzyme combination, the different hybridized restriction fragments were recorded for all accessions, and a matrix (X) of DNA fragments versus inbreds was constructed. The matrix of the number of common signals in all pairwise comparisons between inbred lines (N) was computed as $N = X'.X$ where X' is the transposed matrix of X. Nei's F matrix and the nucleotide substitution matrix (Nei and Li 1979) were computed with a program developed in our laboratory (Gentzbittel and Nicolas 1990). The former was subjected to principal component analysis, being considered equivalent to a correlation matrix, whereas a phylogenetic tree was produced from the latter by the UPGMA clustering method. Genetic diversity statistics were calculated assuming that each probe was located at one locus and that each restriction-fragment banding pattern of a probe was equivalent to an allele at this locus. Gene diversity (H) was computed, as well as intralocus [$V_S(h)$] and interlocus [V(h)] variances. For the latter purpose, the sunflower lines were divided into two subpopulations consisting of R and M lines. Under this subdivided-population model, the total gene diversity (H_T) was partitioned into within-population diversity (HS) and between-population diversity (D_{ST}) (Nei 1987). Estimates of the coefficient of gene differentiation ($G_{ST} = D_{ST}/H_T$) and absolute gene differentiation [$D_m = sD_{ST}/(s-1)$, where s is the number of subpopulations corrected for small num-

bers of subpopulations] were computed in order to compare the gene diversities in cultivated sunflower with previously-described data.

The accuracy of distance estimation was computed for all six pairwise distance estimates among two restorer inbred lines (RHA266 and RHA274) and two maintainer lines (HA89 and SD). A previously-described method (Gentzbittel et al. 1992) allows the computation of the dispersion of the distance on both sides of the mean value (coefficient of dispersion), at each step, for any number of probe-enzyme combinations used. It is thus possible to plot the coefficient of dispersion for a pair of samples as a function of the number of probe-enzyme combinations used to estimate the F value.

Results and discussion

Level of polymorphism

RFLPs of the 17 inbred lines were studied with 181 probes and with four enzymes giving a total of 331 probe-enzyme combinations and 1 113 hybridizing fragments. Seventy-three probes were polymorphic (40% of the total probes), with 614 banding patterns in 101 probe-enzyme combinations. No significant differences were observed in the percentage of polymorphic probes with the different cDNA libraries (except for etiolated plantlets, where the high level of redundancy can be explained by an amplification step in the construction of the library). However, the *Pst*I and *Hin*dIII genomic libraries gave noticeably fewer polymorphisms (Table 2). For polymorphic probe-enzyme combinations, the number of band patterns ranged from 2 to 17, with an average of about five patterns (Fig. 1); and the number of bands per pattern ranged from 2 to 23 with an average of three bands (Table 2). For our set of unselected probes, we found a level of polymorphism that is somewhat lower than that described for other crops. Nei's F index (Table 3) ranges from 0.88 to 0.94, whereas it has been reported as low as 0.10 for wild oat (Goffreda et al. 1992). In a previous study (Gentzbittel et al. 1992), we found an average pairwise interspecific distance (in the mean number of nucleotide substitutions per nucleotide site) over the *Helianthus* genus of about 0.039. This value was in the range of the genetic distances described, for example, in *Solanum* (Miller and Tanksley 1990). In contrast, the estimated distances between inbreds of sunflower (Table 3) range from 0.0032 (inbred H52 versus inbred 2603) to 0.00681(inbred PAT4 versus RHA266), indicating a five- to tenfold reduction of variability in comparison with inter-

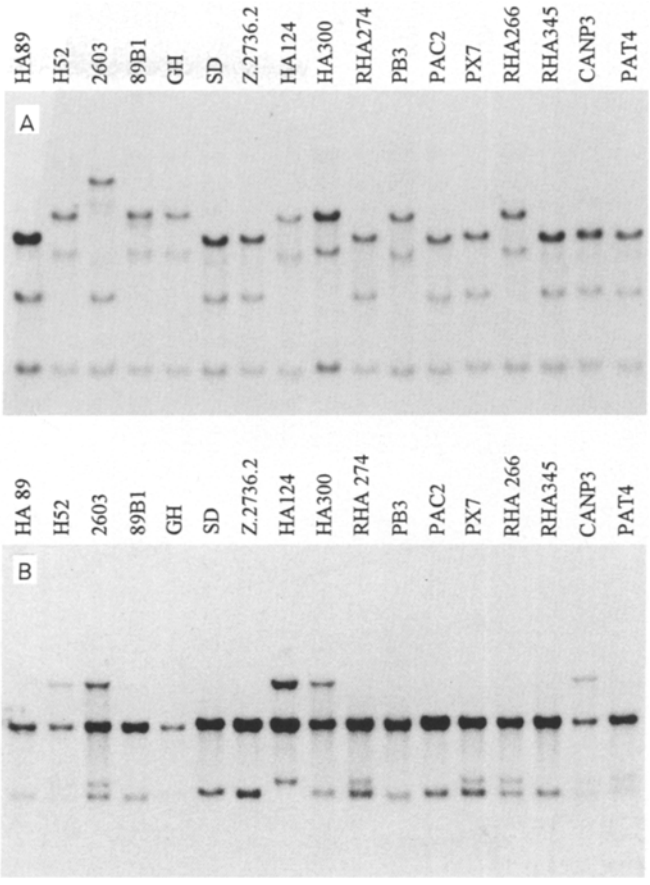

Fig. 1 A, B Autoradiogram from Southern hybridization of 17 inbred lines of cultivated sunflower. **A** *Eco*RI-digested DNA hybridized with cDNA probe K017. **B** *Hin*dIII-digested DNA hybridized with cDNA probe K398

specific distances. The same results were obtained by Quillet et al. (1992) by using eight isozyme systems. They were able to discriminate 52 inbred lines of cultivated sunflower, but state that an important reduction of variability had occurred in the breeding process, resulting in a noteworthy homogenization of biochemical characteristics, as very few rare alleles were found. However, comparisons of our distance data with previously published data are difficult because probes are usually pre-selected for high rates of polymorphism.

Table 2 Summary of the polymorphism detected by the various libraries

Source	Number of probes screened	Number of polymorphic probes	Percentage of polymorphic probes	Average number of patterns/polymorphic probe	Average number of bands per pattern
*Pst*I genomic	13	4	31	5.2	2.3
*Hin*dIII genomic	69	8	12	5.4	3.5
Plantlets cDNA	29	9	31	5.1	2.5
Leaves cDNA	27	19	70	5	3
Ovaries cDNA	22	19	86	5.3	2.2
Anther cDNA	21	14	67	5.3	3.1
Total (Average)	181	73	40	(5.2)	(2.8)

Table 3 Nei's F index (below diagonal) and estimates of nucleotide substitutions per nucleotide site ($\times 10^5$; above diagonal) for the 17 inbred lines of sunflower listed in Table 1

Line	89B1	CANP3	GH	H52	HA124	HA300	HA89	PAC2	PAT4	PB3	PX7	RHA266	RHA274	RHA345	SD	Z2736.2	2603
89B1	–	512	410	468	476	415	499	581	593	588	527	633	587	621	499	509	384
CANP3	0.91	–	610	493	538	483	569	518	597	600	582	497	479	580	554	607	538
GH	0.93	0.90	–	470	536	468	560	680	709	673	596	672	626	698	406	460	428
H52	0.92	0.92	0.92	–	428	410	384	592	620	570	538	585	561	603	465	461	321
HA124	0.92	0.91	0.91	0.93	–	455	576	584	641	577	538	629	553	661	451	476	480
HA300	0.93	0.92	0.92	0.93	0.92	–	462	649	632	605	542	612	573	652	477	495	469
HA89	0.91	0.90	0.90	0.93	0.90	0.92	–	631	622	609	576	624	615	634	488	455	532
PAC2	0.90	0.91	0.89	0.90	0.90	0.89	0.89	–	636	463	439	705	446	604	677	587	643
PAT4	0.90	0.90	0.88	0.90	0.89	0.89	0.90	0.89	–	562	680	681	628	693	644	651	671
PB3	0.91	0.90	0.89	0.90	0.90	0.90	0.90	0.92	0.90	–	489	617	417	560	617	551	636
PX7	0.91	0.90	0.90	0.91	0.91	0.91	0.90	0.92	0.89	0.92	–	541	343	463	487	555	560
RHA266	0.89	0.91	0.89	0.90	0.89	0.90	0.89	0.88	0.89	0.90	0.91	–	520	510	550	595	658
RHA274	0.90	0.92	0.89	0.90	0.91	0.90	0.90	0.92	0.89	0.93	0.94	0.91	–	356	495	542	605
RHA345	0.89	0.90	0.88	0.90	0.89	0.89	0.89	0.90	0.88	0.90	0.92	0.91	0.94	–	582	642	669
SD	0.91	0.91	0.93	0.92	0.92	0.92	0.92	0.89	0.89	0.90	0.92	0.91	0.92	0.90	–	404	444
Z2736.2	0.91	0.90	0.92	0.92	0.92	0.92	0.92	0.90	0.89	0.91	0.91	0.90	0.91	0.89	0.93	–	512
2603	0.93	0.91	0.93	0.94	0.92	0.92	0.91	0.89	0.89	0.89	0.90	0.89	0.90	0.89	0.92	0.91	–

Genetic relationships

The genetic relationships between the inbred lines are represented by principal component analysis (Fig. 2) and UPGMA cluster analysis (Fig. 3), based on Nei's F matrix and estimates of nucleotide substitutions (Table 3). These two methods gave similar results: the inbred PAT4, which was obtained from an interspecific cross with *H. tuberosus* (2n=6x=102) (Table 1), was placed outside the pool of the lines ; inbreds PAC2 and PB3, which contain restoration genes from *H. petiolaris* (2n=34) for cytoplasmic male sterility, are grouped together. All the inbreds of east-European origin (GH, CX, SD, Z2736.2, HA124 and HA89) are on the same branch of the UPGMA tree. These results are in accordance with the pedigree map of sunflower described by Korell et al. (1992).

The major division among the set of inbreds is defined by the restorer or maintainer genotype of an inbred. This grouping is independent of the origin of the restoration system: American inbreds (PX7, RHA274, RHA 345, RHA 266) and French inbreds (PAC2 and PB3) are grouped together, although these lines were created independently and probably do not have the same restorer genes. American inbreds came from crosses with the wild *H. annuus* genotype HA61, whereas the French inbreds came from an interspecific cross with *H. tuberosus*. The only exception

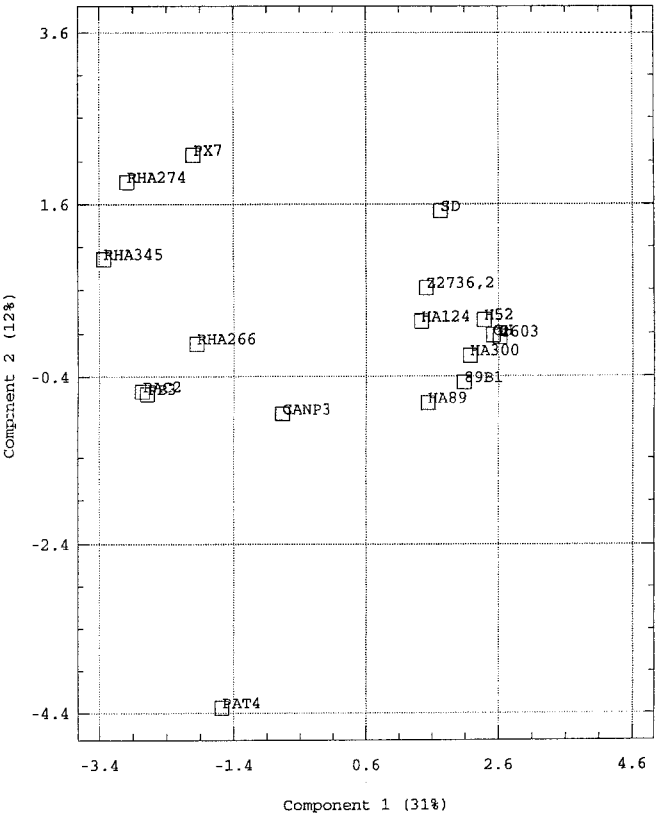

Fig. 2 Plot of the first (31% of total information) versus the second (12% of total information) principal coordinate scores based on Nei's F matrix, for 17 inbred lines

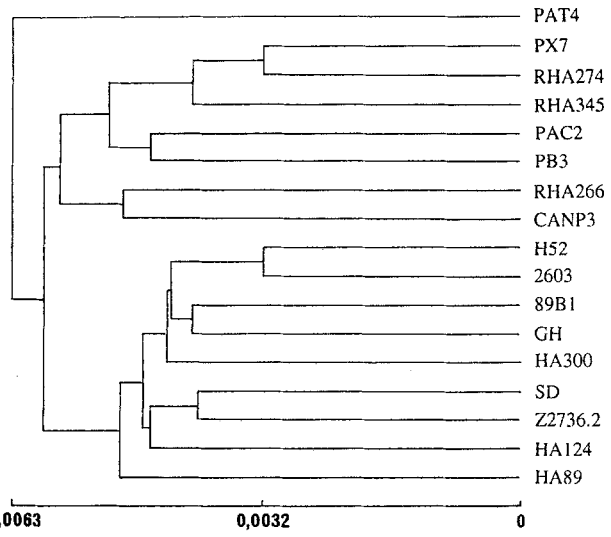

Fig. 3 UPGMA dendrogram constructed using Nei's nucleotide-substitution matrix among 17 inbreds

is the presence of the male-sterile maintainer CANP3, of Russian origin, close to the restorer inbred lines. However, as CANP3 is known to come from a VNMINK population, from which RHA266 was also derived, the grouping of the two lines seems not to be artifactual. The fact that restorer inbreds are grouped, although they do not belong to the same system of restoration nor to the same gene pool, could be explained by selection constraints (e.g., all the 'R' lines in this study are mildew-resistant). The same type of bottleneck effect can be postulated for the grouping of maintainer lines, as Russian selections were made on oil-content. This dual classification by cluster analysis is also a consequence of sunflower breeders having separate breeding programs for R and M inbreds to maximise the potential for heterosis in commercial hybrids. Our results are similar to those of Berry et al. (1994), based on RFLP patterns, where a different set of sunflower inbreds grouped into R- and M-germplasms pools.

Confidence levels

The accuracy of distance estimates between samples, with particular reference to the number of probe-enzyme combinations used, is a general problem. For inbred lines of sunflower, the study of the accuracy of distance estimation as a function of the number of polymorphic probe-enzymes combinations used indicated that 101 polymorphic combinations produce a general dispersion of $< 2\%$ on both sides of the mean value, and that 50 combinations could ensure a dispersion distance of $< 5\%$. Interestingly, the analysis of the correlation of all the pairwise distances estimated with the same polymorphic probe, but with different restriction endonucleases, showed that the majority of the probes (60%) exhibited little or no correlation (R^2 less than 0.3) between the estimated distances, and only 10% of the probes showed a complete correlation between the estimated distances. We hypothesize that this result arises because the frequency of insertions/deletions between restriction sites is considerably lower than that of point mutations within restriction sites. This is in agreement with previous findings concerning wild species of the genus *Helianthus* (Gentzbittel et al. 1992).

Estimation of available variability

By considering this set of inbreds as representative of the population of sunflower lines, estimates of general gene diversity were computed on the assumption that each probe is situated at a single locus, and that each restriction fragment pattern can be assumed to be an allele at this locus. As the majority of probes were hybridized on several restriction endonucleases digests, and to overcome redundancy in the data, we considered two groups of loci: one constitued by the probes that were studied with endonucleases *Eco*RI or *Bgl*II, the other formed by the probes that were analysed with *Hin*dIII or *Eco*RV. The results are presented in Table 4 and provide information about the degree of genetic variability that can be used for breeding schemes. The gene diversity estimated from the two groups of data are the same, and we found an overall gene diversity (H) of 0.207. This value is similar to that described for *Avena* ($H_k = 0.35$, Goffreda et al. 1992), but considerably lower than that described for *Brassica* ($H = 0.65$, McGrath and Quiros 1992) or lentil ($H = 0.521$, Havey and Muehlbauer 1989). The spread of average gene diversity per locus (Fig. 4) presents an L-shaped curve, as described by Nei (1987) under a neutral allele model. The intralocus variance is tenfold lower than the interlocus variance (Table 4), indicating that the sampling error of inbreds at each locus is negligible and need not be reduced by increasing the number of lines sampled. Hence, considering the number of loci studied (158) and the intralocus variances, the results obtained seem neither to be a sampling effect of the inbreds nor of the loci studied. These conclusions are also in agreement with previously-decribed data (Quillet et al. 1992), based on isozyme diversity, showing a great homogenization of biochemical characteristics.

Table 4 Gene diversity statistics for 17 sunflower inbred lines

Loci[a] analysed with:	Gene diversity (H)	SE	Intra-locus variance	Inter-locus variance
*Eco*RI or *Bgl*II[b]	0.204	0.0261	0.00117	0.1074
*Hin*dIII or *Eco*RV[c]	0.211	0.0255	0.00114	0.1019[a]

[a] A locus is defined as a probe-enzyme combination, an allele at this locus as a restriction fragment pattern
[b] Computed on the basis of 58 polymorphic loci and 108 monomorphic loci
[c] Computed on the basis of 57 polymorphic loci and 108 monomorphic loci

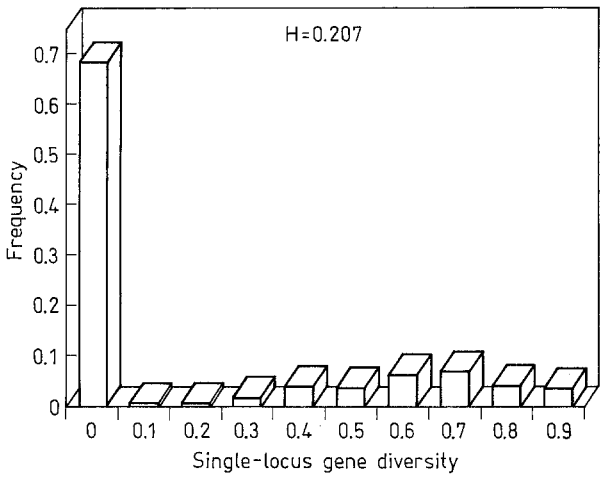

Fig. 4 Distributions of single-locus gene diversity for 17 inbreds. H stands for average diversity among the set of inbreds

As RFLP analysis seems to be able to define groups that are used to obtain hybrids showing heterosis (putative 'heterotic' groups), we can use allele frequencies to ascertain the proportion of overall genetic variability available between R- and M-germplasm pools. We have thus estimated the average gene diversity within (H_S) and between 'R' and 'M' populations (D_{ST}), as well as the proportion of this variability which is explained by the difference between these two subsets of inbreds (G_{ST} and D_m). Most of the gene diversity in the total population (H_T) is explained by the within-subsets gene diversity (Table 5) while the absolute difference between R and M lines is about 2% (D_m=0.023 or 0.024). Interestingly, the average variability within each subset (H_S) is almost the same, indicating a consistent effort for preserving the variability in each group. This low genetic difference between the R- and M-germplasm pools is significant, as tested by Berry et al. (1994) by discriminant analysis. These results are, however, based on the assumption that each probe is situated at a single locus and they need to be corrected as preliminary linkage data indicates that 21% of the probes are located at two or more loci (data not shown).

Conclusions

Our results are the first description of an extensive RFLP analysis of the cultivated sunflower, whereas numerous species have already been subjected to this type of analysis (for example *Arachis*, Kochert et al. 1991 ; bean, Nodari et al. 1992 ; *Vicia*, Van de Ven et al. 1993 ; *Vigna*, Fatokun et al. 1993; as well as maize, rice and tomato). Our analysis of gene diversity shows that the available genetic variability for sunflower is one of the lowest described; indeed, such a low gene diversity has only been reported with protein-loci data for natural populations of man (Nei and Roychoudhury 1982). The simplest explanation for this is based on the particularly recent origin of the cultivated sunflower, and is probably also influenced by a bottleneck effect in the early stages of selection for oil-content. In fact, almost all the actual genotypes are directly or indirectly derived from Russian populations, with the addition of only 30 years of divergent selection. This could also explain the relatively minor variability (2%) between the putative 'heterotic' groups of R- and M-germplasm pools (Table 5). However, the frequency of polymorphic probes (40%) allows easy RFLP studies, and we demonstrate that 50 probe-enzyme combinations ensure a < 5% error on distance estimation. We postulate also that sunflower genome evolution has been mainly due to point mutations rather insertions/deletions.

Our results have significance in terms of sunflower breeding, which seems to be concentrated on relatively narrow germplasm sources. In view of the apparent low genetic variability available in such populations, using RFLPs to determine genetic relatedness among genotypes would allow sunflower breeders to choose parental germplasm material with a view to maximizing variability in their breeding programs, both within and between R and M inbred subsets. For example, in maize, 20% of the genetic distance between lines is explained by differences between flint and dent (Melchinger et al. 1992), whereas in sunflower, although based on different computations, this quantity is estimated at about 2% between R and M lines. Helentjaris et al. (1985) suggested a number of factors that could account for differences in genomic variability; these include differences in mating systems, breeding

Table 5 Gene-diversity statistics for the two subsets of restorer and maintainer inbred lines of sunflower

Loci[a] analysed with	Gene diversity				
	Total H_T[d]	Within H_S	Between D_{ST}	G_{ST}	Absolute D_m[e]
*Eco*RI or *Bgl*II[b]	0.195	0.184	0.011	0.058	0.023
*Hin*dIII or *Eco*RV[c]	0.199	0.187	0.012	0.061	0.024

[a] A locus is defined as a probe-enzyme combination, an allele at this locus as a restriction fragment pattern
[b] Computed on the basis of 58 polymorphic loci and 108 monomorphic loci
[c] Computed on the basis of 57 polymorphic loci and 108 monomorphic loci
[d] To compare with H in Table 2
[e] On the basis of s=2 subpopulations (restorer and maintainer inbred lines)

systems, and domestication events. The former may not play a role in sunflower variability, but the latter two are particularly pertinent to the history of sunflower breeding. This suggests that interspecific or intraspecific crosses with wild *H. annuus*, or unstudied populations, will be a useful way for the creation of new source populations. As an example, PAT4, which comes from an interspecific cross (Table 1), is clearly outside the pool of the inbreds. We believe that sunflower breeders may need to direct additional efforts toward introgression from wild ecotypes or species.

Acknowledgements We thank V. de Conto and S. Haby (GIE CARTISOL) and A. Perrault and J. Phillipon (INRA) for their excellent technical assistance. A. Berville (INRA) is also acknowledged for interesting discussions about molecular markers. This work was financed with private funds from the GIE CARTISOL, a grant from the Ministère de l'Agriculture et de la Forêt of France, and a State - Region contract 'Semences et Plants'.

References

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

Choumane W, Heizmann P (1988) Structure and variability of nuclear ribosomal genes in the genus *Helianthus*. Theor Appl Genet 76:481–489

Fatokun CA, Danesh D, Young ND, Stewart EL (1993) Molecular taxonomic relationships in the genus *Vigna* based on RFLP analysis. Theor Appl Genet 86:97–104

Gentzbittel L, Nicolas P (1990) Improvement of 'a BASIC program to construct evolutionary trees from restriction endonuclease data' with the use of Pascal. J Hered 81:491–492

Gentzbittel L, Perrault A, Nicolas P (1992) Molecular phylogeny of the *Helianthus* genus, based on nuclear restriction-fragment length polymorphism (RFLP). Mol Biol Evol 9:872–892

Goffreda JC, Burnquist WB, Beer SC, Tanksley SD, Sorrells ME (1992) Application of molecular markers to assess genetic relationships among accessions of wild oat, *Avena sterilis*. Theor Appl Genet 85:146–151

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

Helentjaris T, King G, Slocum M, Siedenstrang C, Wegman S (1985) Restriction-fragment length polymorphisms as probes for plant diversity and their development as tools for applied plant breeding. Plant Mol Biol 5:109–118

Kiss T, Solymosy F (1987) Isolation of high-molecular-weight plant nuclear DNA suitable for use in recombinant DNA technology. Acta Biochim Biophys Hung 22:1–5

Kochert G, Halward T, Branch WD, Simpson CE (1991) RFLP variability in peanut (*Arachis hypogaea* L.) cultivars and wild species. Theor Appl Genet 81:565–570

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

Leclercq P (1969) Une stérilité mâle cytoplasmique chez le Tournesol. Ann Amélior Plant 19:99–106

McGrath JM, Quiros CF (1992) Genetic diversity at isozyme and RFLP loci in *Brassica campestris* as related to crop type and geographical origin. Theor Appl Genet 83:783–790

Melchinger AE, Boppenmaier J, Dhillon BS, Pollmer WG, Herrmann RG (1992). Genetic diversity for RFLPs in European maize inbreds. II. Relation to performance of hybrids within, versus between, heterotic groups for forage traits. Theor Appl Genet 84:672–681

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

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

Nei M, Li W-H (1979) Mathematical model for studying genetic variation in terms of restriction endonucleases. Proc Natl Acad Sci USA 76:5269–5273

Nei M, Roychoudhury AK (1982) Genetic relationship and evolution of human races. Evol Biol 14:1–59

Nodari RO, Koinange EMK, Kelly JD, Gepts P (1992) Towards an integrated linkage map of common bean. 1. Development of genomic DNA probes and levels of restriction-fragment length polymorphism. Theor Appl Genet 84:186–192

Quillet MC, Vear F, Branlard G (1992) The use of isozyme polymorphism for identification of sunflower (*Helianthus annuus*) inbred lines. J Genet Breed 46:295–304

Rieseberg LH, Soltis DE, Palmer JD (1988) A molecular reexamination of introgression between *Helianthus annuus* and *H bolanderi* (*Compositae*). Evolution 42:227–238

Rieseberg LH, Beckstrom-Sternberg S, Doan K (1990) *Helianthus annuus* ssp. *texanus* has chloroplast DNA and nuclear ribosomal RNA genes of *Helianthus debilis* ssp *cucumerifolius*. Proc Natl Acad Sci USA 87:593–597

Schilling EE, Heiser CB (1981) Infrageneric classification of *Helianthus* (compositae). Taxon 30:393–403

Serieys H, Vincourt P (1987) Caractérisation de nouvelles sources de stérilité mâle chez le Tournesol. Les colloques de l'INRA no 45, INRA, Paris, pp 53–64

Van de Ven WTG, Duncan N, Ramsay G, Phillips M, Powell W, Waugh R (1993) Taxonomic relationships between *V. faba* and its relatives based on nuclear and mitochondrial RFLPs and PCR analysis. Theor Appl Genet 86:71–80