COMPARATIVE ANALYSIS OF GENETIC RELATIONSHIPS IN SUNFLOWER INBRED LINES, BASED ON ISOZYMIC, RAPD AND PEDIGREE DATA.

Guillermo G. Pizarro, Alicia D. Carrera, Mónica Poverene
Universidad Nacional del Sur, Departamento de Agronomía,
San Andrés 800, 8000 Bahía Blanca, Argentina
Fax: 54-291-4595127; e-mail: gpizarro@criba.edu.ar

Raúl H. Rodríguez, Mercedes M.Echeverría, María T. Salaberry
Unidad Integrada Balcarce, INTA/FCA-UNMdP
CC 276, 7620 Balcarce, Argentina.
e-mail: rhrodriguez@balcarce.inta.gov.ar

Abstract

Genetic relationships have seldom been analysed with different types of markers in order to compare the information provided by each marker class. The objectives of this study were: (i) to use isozymic and RAPD markers to determine estimates of genetic similarity among 21 selected sunflower accessions, (ii) to compare the utility of this two markers systems for evaluating similarity, (iii) to examine the agreement between genetic similarity coefficients and the coancestry coefficient (*C*).
The usefulness of each marker system was examined in terms of the amount of polymorphism detected (arithmetic mean heterozygosity or Hav) and the number of loci revealed (effective multiplex ratio or E). The overall utility of both markers was evaluated by the marker index (MI) which is the product of Hav and E. Hav values calculated for isozymes and RAPD were not significantly different. Marker index for RAPD was almost two-fold the value for isozymes.
Mean Jaccard similarity coefficients (JSC) were calculated from 97 polymorphic RAPD bands and six polymorphic isozyme loci, and were similar for both markers. The correlation between RAPD and isozymes similarity matrices was low (r = 0,22). Across all 210 possible inbred lines combinations, there was not correlation between *C* and JCS based in RAPD nor isozymes data. However when considering only related lines (*C* >0,1) the correlation coefficient between *C* and JCS-RAPD became significant ($p<0,05$) and increased to r = 0,31.

Introduction

Sunflower is a very important oil-seed production crop in Argentina. Cropped on 3 million hectares, it constitutes a very important economic resource and nowadays our country is the world´s first largest producer.

A detailed knowledge of the genetic relationships among accessions is an important factor for the success of plant breeding programs and for efficient sampling and more informed utilisation of available germplasm. The use of morphological traits to study such relationships is subject to the influence of the environment, and is limited by the paucity of the traits showing polymorphism, specially when dealing with elite genotypes. In comparison with morphological markers, molecular markers offer significant advantages with respect to increased number of detectable loci, overall phenotypic neutrality, and the ability to score plants at any developmental stage.

Isozymes have been used for the characterization and identification of inbred lines and varieties in sunflower (Quillet et al., 1992; Carrera & Poverene, 1995). More recently RAPDs have been developed for assessing genetic relationships in sunflower (Teulat et al., 1994; Lawson et al., 1994).

Genetic relationships have seldom been analysed with different types of markers in order to compare the information provided by each marker class. A comparison of different marker techniques is timely, principally taking into account that there is no report of comparison of information provided by isozymes and RAPD data in sunflower.

The objectives of this study were: (i) to use isozymic and RAPD markers to determine estimates of genetic similarity among 21 selected sunflower accessions, (ii) to compare the utility of this two markers systems for evaluating similarity, (iii) to examine the agreement between genetic similarity coefficients and the coancestry coefficient. Results would also provide information for parental selection in a breeding program.

Materials and methods

Plant material

Twenty one inbred lines (Table 1), obtained from INTA Balcarce (Argentina) were chosen to explore the diversity of sunflower germplasm. All these inbreds have been extensively used in sunflower breeding programs.

Coefficient of coancentry (*C*) from each entry were obtained from the SUNflower GENomE (SUNGENE) database (http://www.css.orst.edu/knapp-lab/sungene)(Cheres and Knapp, 1998) or calculated as described by Falconer (1970).

Isozymes Assays

Samples were prepared from seeds soaked for 24 hours (48 hrs for PGM) using a 0.1M Tris-HCl-mercaptoethanol buffer (pH 7.5). The following enzymes were assayed: acid phosphatase (ACP), esterase (EST), glutamate dehydrogenase (GDH), phosphoglucoisomerase (PGI), 6-phosphogluconate dehydrogenase (PGD) and phosphoglucomutase (PGM). Allozymes were resolved on 12% horizontal starch gel. The buffer systems and staining methods are described in Carrera and Poverene (1995), after Soltis et al. (1983). The number of loci and alleles were interpreted according to Torres (1983), Kahler and Lay (1985), Rieseberg and Soltis (1989), and Carrera and Poverene (1991, 1995). Loci were designated with the most anodally migrating isozyme given number 1, and additional loci numbered sequentially in order of decreasing electrophoretic mobility. The most anodally migrating allozyme was designated by the letter ‘a’.

RAPD Assays

Sunflower DNA was prepared and handled as described by Fütterer et al. (1995).

Reactions volumes of 25 μl were prepared containing 50 ng of genomic DNA, 0,3 μM primer, 0,2 mM of each dNTP, 1U Taq polymerase (Gibco), 2 mM $MgCl_2$, 50 mM KCl, 10 mM Tris-HCl (pH 8,4). Mixture was gently mixed and overlaid with one drop of heavy mineral oil. Decamer oligonucleotide primers (sets G and P) were purchased from Operon Technologies (Alameda, CA). The RAPD PCR amplifications were performed using a MJ Thermal Cycler (Model PTC100). The thermocycling program consisted of an initial 6 min hold at 94 ºC, 40 cycles of 15 s at 94 ºC, 45 s at 40 ºC and 1 min at 72 ºC, ending with a 6-min hold at 72 ºC. The reactions were then held at 4 ºC until used.

Amplified products (18 μl) were electrophoresed in a 1,5 % (w/v) agarose gel in 1x TAE. In each gel, a 100-bp ladder was included as a molecular weight standard. Amplification products were separated according to size by electrophoresis at 95 V for 2 h. Gels were stained with EtBr and photographed over UV light.

Data analysis

Polymorphism was scored on a presence or absence basis and data were analyzed using the program NTSYS-pc version 1.40 (Applied Biostatistic Inc. 1988). Monomorphic markers were excluded from the analysis. Genetic similarity values used for cluster analyses were calculated between all 210 possible pairs of lines based on Jaccard coefficient (Jaccard, 1908).

Similarity matrices obtained with isozymes and RAPD data and the coefficient of coancestry matrix were compared using the Mantel matrix-correspondence test (Mantel 1967). Dendrograms were obtained by the UPGMA (unweighted pair group method using arithmetic averages) cluster analysis performed with NTSYS-pc program.

The expected heterozygosity (Hn), the arithmetic mean heterozygosity (Hav), the fraction of polymorphic loci (β), the effective multiplex ratio (E), and the marker index (MI) were calculated as reported by Powell et al. (1996). The sum of effective number of alleles (SENA) was calculated by determining the effective number of alleles for each locus ($n_e = 1/\Sigma p_i^2$), reducing it by 1, and summing over all loci: $SENA = \Sigma[(1/\Sigma p_i^2)-1]$

For isoenzymes standard measures of genetic variation were also calculated, including the proportion of polymorphic loci (P), the mean number of alleles across all loci (A), the mean number of alleles per polymorphic locus (Ap).

Results

Levels of polymorphism

RAPD preliminary experiments were carried out to optimise reaction conditions. The components considered were the concentration of template DNA (from 0,4 to 2 ng/μl), primer (from 0,2 to 0,6 μM), Taq polymerase (from 1 to 2 U/reaction) and Mg (from 0,5 to 2,5 mM).

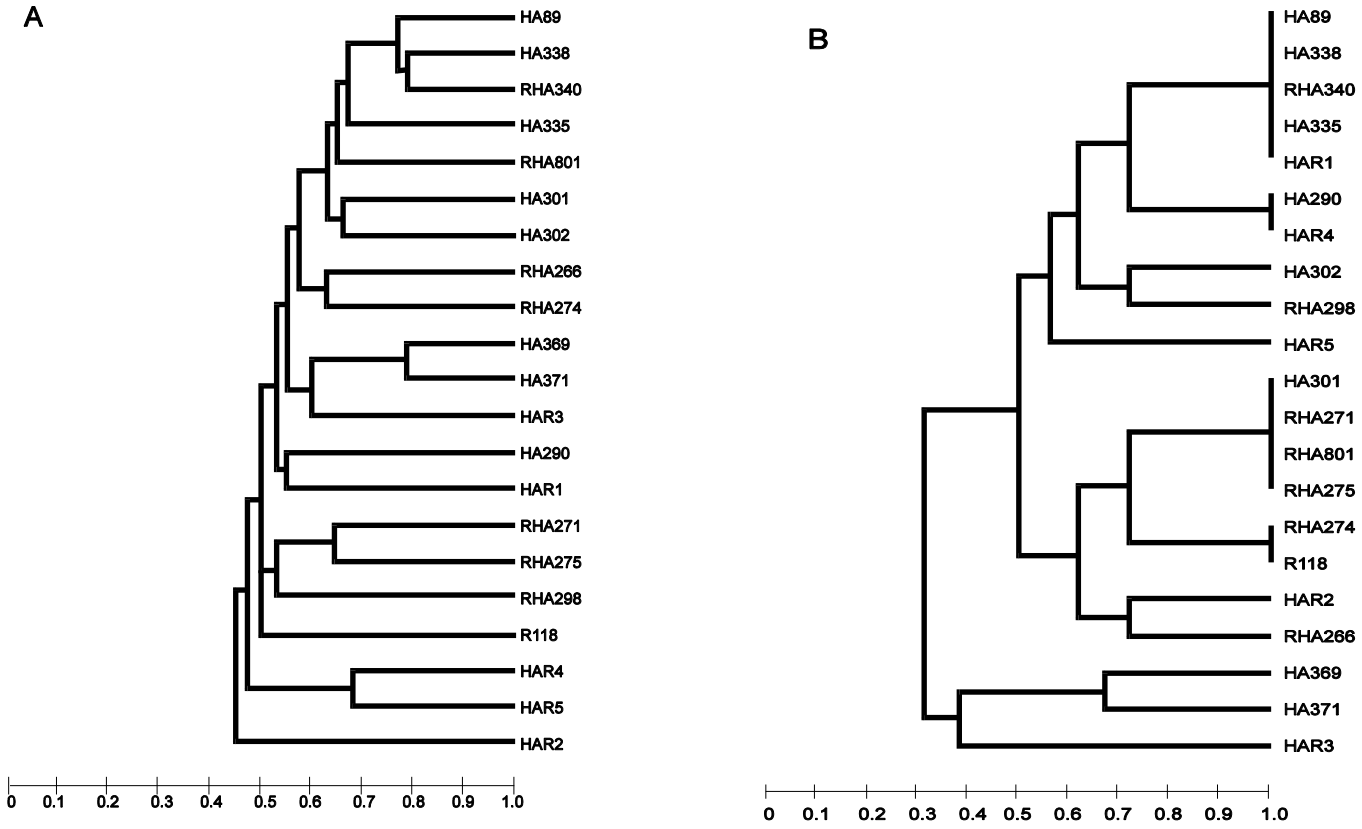
Among the 33 primers used, 3 (9 %) produced a monomorphic banding pattern and 5 (15 %) produced amplification products that were too faint to score or could not be consistently reproduced. The remaining 25 primers (75,8 %), used to evaluate all 21 accessions, detected one or more polymorphic bands. A total of 136 scorable RAPD bands were amplified, 97 (71 %) of which were polymorphic (3,88 bands/primer). No single RAPD primer differentiated all 21 accessions, but on the basis of the 97 RAPD polymorphisms each inbred line was uniquely fingerprinted.

The six enzyme systems assayed revealed 13 loci, with a total of 20 alleles (A = 1.53). The average number of alleles per polymorphic locus (Ap) was 2.166. Six loci were found to be polymorphic: *Acp-1*, *Est-1*, *Gdh-2*, *Pgd-3*, *Pgi-2* and *Pgm-1*. Therefore, the proportion of polymorphic loci (P) was 0.46. Five of these loci were localised onto four linkage groups of the

Table 1. Pedigree of sunflower inbred lines analyzed by isoenzymes and RAPD markers

Line	Pedigree
HA 290	4*P-21—VR1/HA 60
HA 301	Peredovik 301
HA 302	Peredovik 304
HA 89	VNIIMK 8931 Sel
HA 371	H52 Sel
HA 369	ARG 8018
HA 338	HA 89*3/H. praecox 419
HA 335	HA 89*3/H. annuus 423
HA-R1	Pergamino 71/538 Selection
HA-R2	Impira INTA Selection
HA-R3	Charata Selection
HA-R4	Saenz Peña 74-1-2 Selection
HA-R5	Guayacan INTA Selection
RHA 266	2*Peredovik/953-102-1-1-41
RHA 271	CMS PI 343765/HA 119//HA 62-4-5/2/T 66006-2-1-31-1
RHA 274	CMS PI 343765/HA 119//HA 62-4-5/2/T 66006-2
RHA 275	CMS PI 343765/HA 119//HA 62-4-5/2/T 66006-2-2-11-3-2
RHA 340	HA89*3/H. argophyllus 415
RHA 801	Derived from a Restorer Composite
RHA 298	CMS HA 89/RHA 273
R 118	Derived from a Restorer Composite

Figure 1. Relationships among 21 sunflower inbred lines revealed by cluster analysis of Jaccard genetic similarity coefficients calculated from (A) RAPD data, and (B) isozymes data.



public RFLP map (Carrera et al., 1998). There were 12 different genotypes, 8 (38%) of which were unique within this set of material.
Arithmetic mean heterozygosity (Hav) values calculated for isozymes and RAPD (Table 2) were not significantly different (standard two-sample t-test). Marker Index (MI) for RAPD was almost two-fold the value for isozymes. According to the small number of isozyme genes available to be assayed, the sum of effective numbers of alleles (SENA) found to be lowest for this marker.

Correspondence of similarity measured between markers systems
Jaccard coefficient of similarity (JCS) ranged from 0,365 to 0,792 (mean 0,533; SD = 0,0061) for RAPD and from 0,090 to 1 (mean 0,532; SD = 0,512) for isozymes. The coefficient of coancestry (*C*) ranged from 0 to 0,875 (mean 0,122; SD = 0,198).
Figure 1 shows isozyme and RAPD dendrograms for the 21 sunflower genotypes. The goodness of fit of the UPGMA dendrograms for isozymes (r = 0,84) and RAPD (r = 0,80) were highly significant (p<0,001), and the high r values indicate that the trees gave a reliable representation of the genetic similarity between genotypes.
The comparison of isozyme and RAPD similarity matrices, was significant (p< 0,05), however, the correlation coefficient was low (r = 0,22). Across all 210 possible inbred lines combination, there was not correlation between *C* and JCS based in RAPD nor isozyme data. However when considering only related lines (*C* >0,1) the correlation coefficient between *C* and JCS-RAPD became significant (p<0,05) and increased to r = 0,31.

Table 2. Comparison of the average expected heterozygosity for polymorphic markers, Hav and its standard deviation, of the fraction of polymorphic markers β, multiplex ratio n, the effective multiplex ratio E, the marker index MI, and the sum of effective numbers of alleles SENA for each marker class, calculated on the basis of experimental data obtained from 21 sunflower lines.

Marker system	Number of loci	Hav	SD of Hav	β	n	E	MI	SENA
Isozymes	6	0,335	0,136	0,46	4,33	1,99	0,67	3,45
RAPD	97	0,304	0,150	0,71	5,44	3,86	1,17	46,64

Discussion
To our knowledge this is the first report of a comparison involving isozymes and RAPD markers.
The coincidence of isozymes and RAPD Hav values are according with a similar number of alleles (all but one isozyme gene, Est-1, showed two alleles like RAPD) and similar allelic frequencies. The higher MI value for RAPD markers can be explained by a greater number of independent loci analysed per assay and a higher fraction of polymorphic loci.
RAPD and isozymes presented similar mean JSC values, but isozymes showed a broader range, and the highest SD value.
A possible explanation for the low correlation value between JSC-isozymes / JSC-RAPD and not correlation between JSC-isozymes / *C*, could be the low number of loci and poor genome coverage by isozyme markers. The different capacity to detect DNA mutations (many variation at the DNA level can remain hidden at the protein level because several single mutations do not change the amino-acid composition or the global charge of proteins) could also account for the low correlation value between JSC-isozymes / JSC-RAPD. In addition, the disparity between genetic diversity levels detected by pedigree and molecular marker-based JSC, and the lack of correlation between the respective matrices may result from fundamental differences in the

concepts underlying both measures. The coefficient of coancestry is an indirect measure based on the probability that two alleles of a locus taken at random, one of each parent, are "identical by descent". It assumes that the original ancestors are unrelated and each progeny received half of its genes from each parent. Thus, it ignores the effects of selection, mutation and genetic drift. For markers, genetic similarity is a direct measure of their resemblance in the DNA and should reflect the proportion of "genes alike in state" irrespective of whether the identity is caused by alleles identical by descent or those "only alike in state" (Bohn et al., 1999).

The increased correlation between *C* and JCS-RAPD when considering only related lines (*C*>0,1) may be attributed to minor effect of selection and drift in the germplasm and that pedigree relationships may be erroneous or imprecise for distantly related material.

We had used marker index and its components to examine the efficiency of the two marker systems. In conclusion this study has demonstrated that isozymes and RAPD had similar Hav values, but the overall efficiency is higher for RAPD due to its greater value of the effective multiplex ratio. Moreover, while isozymes could differentiated only eight lines, RAPD could uniquely discriminate between all lines. What is more JSC-RAPD presented correlation with *C* for related (*C* >0,1) pairs of sunflower inbreds, suggesting that RAPD data may help quantify the degree of relatedness in sunflower germplasm.

References

Bohn M, Utz H, Melchinger A. (1999). Genetic similarities among winter wheat cultivars determined on the basis of RFLPs, AFLPs, and SSRs and their use for predicting progeny variance. Crop Science 39: 228-237.

Carrera, A & Poverene, M. (1991).Variabilidad genética en girasol mediante marcadores isoenzimáticos. Invest. Agr. Prot. Veg. 3:281-293.

Carrera A, and Poverene M. (1995). Isozyme variation in Helianthus petiolaris and sunflower, H. annuus. Euphytica 81: 251-257.

Carrera A, Pizarro G, Poverene M, Feingold F, Berry S, León A. (1998) Mapping sunflower isozymes in an RFLP linkage map. Plant & Animal Genome VI Conference. San Diego CA. P335

Cheres M, and Knapp S. (1998). Ancestral origins and genetic diversity of cultivated sunflower: coancestry analysis of public germplasm. Crop Science 38: 1476-1482.

Falconer DS. (1970). Introduction to quantitative genetics. 1st ed. Compañía Editorial Continental. México

Fütterer J, Gisel A, Iglesias V, Kloti A, Kost B, Mittelsten Scheid O, Neuhaus G, Neuhaus-Url G, Schrott M, Shillito R, Spangenberg G, Wang ZY. (1995). Standard Molecular Techniques for the Analysis of Transgenic Plants. En Gene Transfer to Plants. Potrykus Y, Spangenberg G (eds.) Springer Lab Manual. Cap 25: 215.

Jaccard P. (1908) Nouvelles recherches sur la distribution florale. Bull Soc. Vaud. Sci. Nat. 44:223-270.

Kahler A & Lay L. (1985) Genetics of electrophoretic variants in the annual sunflower. J. Hered. 76: 335-340.

Lawson WR, Henry RJ, Kochman JK, Kong GA. (1994). Genetic diversity in sunflower (*Helianthus annuus* L.) as revealed by random amplified polymorphic DNA analysis. Aust. J. Agric. Res. 45(7):1319-1327.

Mantel N. (1967) The detection of disease clustering and a generalized regression approach. Cancer Res 27:209-220.

Powell W, Morgante M, Andre C, Hanafey M, Vogel J, Tingey S, Rafalski A. (1996). The comparison of RFLP, RAPD, AFLP and SSR (microsatellite) markers for germplasm analysis. Mol. Breed. 2: 225-238.

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 L & Soltis D. (1989). Assessing the utility of isozyme number for determining ploidal level: Evidence from Helianthus and Heliomeris (Asteraceae). Aliso 12: 277-286.

Soltis D, Haufler C, Darrow D and Gastony G. (1983) Starch gel electrophoresis of ferns: A compilation of grinding buffers, gel and electrode buffers, and staining schedules. American Fern Journal 73: 9-27.

Teulat B, Zhang YX, Nicolas P. (1994). Characteristics of random amplified polymorphic DNA markers discriminating *Helianthus annuus* inbred lines. Agronomie 14: 497-502.

Torres A. (1983). Sunflower (Helianthus annuus L.). In "Isozymes in Plant Gentics and Breeding, Part B". Eds S.D. Tanksley and T.J. Orton, T.J., Elsevier Science Publishers B.V., Amsterdam, pp. 329-338

Torres A.M. and Diedenhoffen (1976) The genetic control of sunflower seed acid phosphatase. Canadian J. Genet. Cytol. 18: 709-716.