B2-coperez

# PROGRESS IN THE DEVELOPMENT OF DNA-BASED MARKERS FOR HIGH STEARIC ACID CONTENT IN SUNFLOWER

**Begoña Pérez-Vich**, SES-Europe NV/SA, Advanta Biotechnology Laboratory, Industriepark, Soldatenplein Z2 nr. 15, B-3300 Tienen, Belgium
Fax: +32 16 808263; e-mail: begonia.perez@ses-europe.be

**Jose María Fernández-Martínez**, Instituto de Agricultura Sostenible (CSIC), Apartado 4084, E-14080 Córdoba, Spain
Fax: +34 957 499252; e-mail: cs9femaj@uco.es

**Juan Muñoz-Ruz**, Instituto de Agricultura Sostenible (CSIC), Apartado 4084, E-14080 Córdoba, Spain
Fax: +34 957 499252; e-mail: jmruz@cica.es

**Steven J. Knapp**, Department of Crop and Soil Science, CRS451C, Oregon State University, Corvallis, OR 97331, USA
Fax: +541 737 1334; e-mail: knapps@css.orst.edu

**Simon T. Berry**, SES-Europe NV/SA, Advanta Biotechnology Laboratory, Industriepark, Soldatenplein Z2 nr. 15, B-3300 Tienen, Belgium
Fax: +32 16 808263; e-mail: simon.berry@ses-europe.be

Summary: Sunflower seed oil with higher stearic acid (C18:0) content could be useful for some food manufacturing process. A high C18:0 line (CAS-3) and a low C18:0 line (HA-89) were crossed in order to obtain an $F_2$ population segregating for C18:0. RFLP and AFLP markers were used to construct a linkage map of this population for the identification of QTL associated with high C18:0. A major QTL controlling the level of C18:0 was found on linkage group 1, which explained 86.8% of the phenotypic variation observed for C18:0 in this $F_2$ population. Moreover, candidate genes from the fatty acid biosynthetic pathway were also mapped. The major loci detected by the Δ9-stearate desaturase (U91340) gene probe coincided with the QTL peak for C18:0 on linkage group 1. A FatB thioesterase gene (AF036565) was also mapped, but no association between this marker and the levels of C18:0 was found.

## INTRODUCTION

Sunflower (*Helianthus annuus* L.) seed oil contains mostly unsaturated fatty acids (oleic, C18:1, and linoleic, C18:2). For some food manufacturing processes, a higher content of saturated fatty acids is desirable (Kinney, 1999). Recently a new EMS generated mutant, designated CAS-3, has been developed, which has high levels of C18:0 (>22%) (Osorio et al., 1995). Classical genetic studies indicated that the C18:0 concentration in this line was controlled by partially recessive alleles at two independent loci (*Es1*, *Es2*) (Pérez-Vich et al., 1999). The effect of the *Es1* locus on the C18:0 concentration is greater than that of the *Es2* locus. It is possible to speculate about the function of these two loci based on our knowledge of the fatty acid biosynthetic pathway. The amount of stearate available for incorporation into seed storage lipids is likely to be controlled by the combined regulation of the following fatty acid biosynthetic enzymes: β-ketoacyl-acyl carrier protein (ACP) synthase II, which forms stearoyl-ACP (C18:0-ACP), and stearoyl-ACP desaturase, acyl-ACP thioesterases and acyl-ACP transferases, all of which use stearoyl-ACP as a substrate (reviewed in Harwood, 1996) (Fig. 1). Altered regulation of any one of these enzymes could potentially affect the level of C18:0 accumulated in the seed oil.

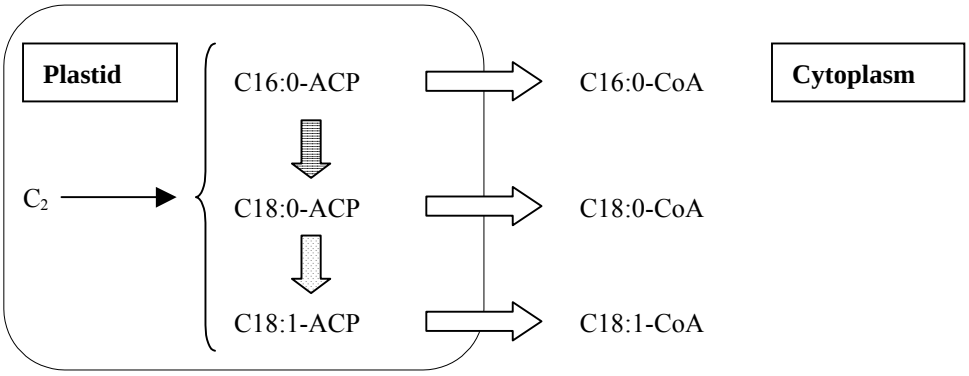


**Figure 1.-** Schematic representation of the biochemical pathway involved in the synthesis of fatty acids in the plastid (Harwood, 1996). ⇛ βKetoacyl-ACP synthase II, ⇨ stearoyl-ACP desaturase, and ⇨ acyl-ACP thioesterase.

Although sunflower has a tremendous economic and biological significance, genetic marker and map development emerged in this crop nearly a decade after the advent of RFLP mapping in other important crop plant species (Knapp et al., in press). The application of DNA markers to sunflower breeding has been focused mainly on the development of markers for disease resistance and some important agronomic traits (Knapp et al., in press). To date very few molecular studies, relating to sunflower seed oil content and quality, have been published (Leon et al., 1995; Dehmer and Friedt, 1998; Kabbaj et al., 1996; Hongtrakul et al., 1998a, 1998b; Lacombe et al., 1999). The objective of this research was to construct a genetic linkage map comprising both RFLP (Restriction Fragment Length Polymorphism) and AFLP (Amplified Fragment Length Polymorphism) markers in order to identify QTL controlling the high levels of C18:0 in the mutant line CAS-3. In addition to these anonymous DNA markers, cDNA clones of a Δ9-stearate desaturase (Hongtrakul et al., 1998a), and a FatB thioesterase were also mapped as potential candidate genes for high C18:0 QTL.

## MATERIAL AND METHODS

The $F_2$ population used in this study was derived from a cross between the high C18:0 line CAS-3, and the standard-low C18:0 line HA-89 (Pérez-Vich et al,. 1999). Fatty acid composition was determined for each of 149 $F_2$ half-seeds by gas-liquid chromatography (Pérez-Vich et al., 1999). $F_2$ half-seeds were grown to obtain the $F_3$ seeds. $F_3$ families from each individual $F_2$ plant were obtained in the field at Venado Tuerto (Argentina).

Tissue from 15 plants of each $F_3$ family was bulked prior to DNA extraction in order to recreate the $F_2$ genotype. The procedures of DNA isolation, digestion, Southern blotting, and hybridization were carried out as reported by Berry et al. (1994, 1995). Additionally, cDNA clones for the fatty acid biosynthesis enzymes Δ9-stearate desaturase [SAD-17 (U91340), Hongtrakul et al., 1998a], and FatB thioesterase (AF036565) were used as RFLP probes. AFLP analyses were carried out by Keygene N.V., Wageningen, NL. The segregation of alleles at the RFLP marker loci was checked against the expected ratios for codominant (1:2:1) or dominant (3:1) markers in an $F_2$ population using a Chi-squared test. An RFLP-AFLP linkage map was constructed with MAPMAKER V 3.0 (Lander et al., 1987), using the Kosambi centiMorgan function. For QTL analysis by interval mapping, using the computer program MAPMAKER/QTL V 1.1 (Lander and Botstein, 1989), a sub-set of codominant marker loci were selected to avoid problems associated with marker clustering. A LOD threshold of 3.0 was used to identify QTL regions affecting the concentrations of all the fatty acids measured. The data set and the unlinked markers were also analysed using a one-way ANOVA.

## RESULTS AND DISCUSSION

*Linkage map*

From the 167 RFLP probes screened against the parental lines, HA-89 and CAS-3, 72 were polymorphic (43%). These 72 polymorphic probes detected 85 loci, with only 15 of these scored as dominant markers. Keygene ran a twelve *Eco*RI/*Mse*I primer combinations on 134 individuals, which revealed 226 loci. All the AFLP loci were scored codominantly based on signal intensity differences between homozygotes and heterozygotes, using Keygene's proprietary software system; however most loci contained a variable number of dominant scores (i.e. C's and D's). A total of 298 AFLP+RFLP loci were arranged into 16 linkage groups (LGs) with a minimum LOD score of 3.0 and in comparison with our existing maps. The linkage group nomenclature is according to Berry et al. (1997). Three markers were unlinked. LG 6 had no polymorphic AFLP or RFLP markers.

For the QTL analyses, a sub-set of 156 loci was selected, which included the loci detected by the SAD-17, and the FatB cDNA probes. These loci covered 1559 cM of the sunflower genome. The average distance between adjacent markers was 10.0 cM, although on LG 2, 13, 14, and 16 there were intervals greater than 30 cM due to the lack of polymorphic markers within these regions. The mean length of a linkage group was 97.4 cM, but the range in length was from 73.4 to 145.1 cM. No clear regions of distorted segregation were detected, although 7 individual RFLP markers segregated significantly different (5% level) from the expected ratios of 1:2:1 and 3:1.

*QTL analysis*

One major QTL controlling the C18:0 concentration (peak LOD 53.9) was identified on LG 1 (Fig. 2). This QTL alone explained 86.8% of the phenotypic variation for the high C18:0 trait and its gene action was partially recessive (Table 1). MAPMAKER/QTL gave the most likely position of the QTL as 4.3 cM to the right of the main SAD-17 locus (Fig. 2), and ANOVA analysis showed that SAD-17 locus on LG 1 was the most significant marker (Table 1). Significant associations between the C18:0 concentration and other markers linked to SAD-17 were also found (Fig. 2). A FatB thioesterase gene (AF036565) was also mapped. It was located in LG 7 and no significant association between this FatB locus and the C18:0 concentration was found.

**Table 1.-** ANOVA statistics for the SAD-17 locus showing its significant affect ($P$<0.001) on the levels of C16:0, C18:0 and C18:1 in the $F_2$ of the cross HA-89 x CAS-3.

| Fatty acid | SAD-17 marker loci on LG 1 | | | | | | |
|---|---|---|---|---|---|---|---|
| | Mean (m) fatty acid (%) and F value[a] | | | | Number of individuals[a] | | |
| | m(A) | m(H) | m(B) | F | n(A) | n (H) | n(B) |
| C16:0 | 8.21 | 7.16 | 7.07 | 9.53 | 28 | 68 | 34 |
| C18:0 | 19.56 | 9.59 | 6.16 | 86.25 | 28 | 68 | 34 |
| C18:1 | 19.01 | 35.24 | 40.13 | 22.83 | 28 | 68 | 34 |

[a] The A allele corresponds to the line CAS-3, and the B allele to the line HA-89

Since C18:0 is desaturated to C18:1 by the Δ9-stearate desaturase (reviewed in Harwood, 1996), it is not surprising to find this major C18:0 QTL very close to a SAD-17 locus. This QTL also significantly affects the levels of C16:0 and C18:1 (Table 1 and Fig. 2). Again this is not unexpected because, due to the stepwise biosynthetic formation of seed oil fatty acids (reviewed in Harwood, 1996), the increase in one fatty acid can affect the level of the other fatty acids both upstream and downstream in the pathway.

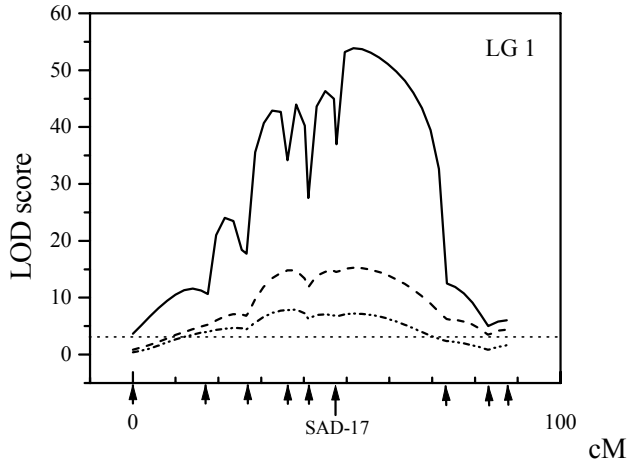


**Figure 2.-** LOD profiles for loci controlling stearic acid ( — ), oleic acid ( --- ) and palmitic acid (-..-..-) content on the LG 1 of sunflower. The X-axis shows the marker loci, represented by *black arrows*. The *dashed line* indicates a LOD score threshold of 3.0.

From the LOD plot in Fig. 2 it appeared that there were multiple LOD peaks which could imply the presence of more than one QTL on LG 1. However, the LOD score from a multi-locus model, which included the three highest LOD peaks, was similar to that of a single QTL model (i.e. 55.06 versus 53.86). Moreover, when the QTL to the right of SAD-17 was fixed, and LG 1 was scanned again to look for a second QTL, no significant increase in the LOD score was observed. Therefore, we consider the hypothesis of multiple high C18:0 QTLs in this region unlikely. The multiple LOD peaks are due to a decrease in the LOD score at the marker loci (Fig. 2), which maybe caused by a mixture of codominant and dominant scoring and/or genotyping errors.

Pérez-Vich et al. (1999) reported one major gene, designated *Es1*, controlling the high C18:0 levels in CAS-3, which was the only one mutated in the original parental line of CAS-3, RDF-1-532. Taking into account these results, the Δ9-stearate desaturase locus on LG 1 probably corresponds to the *Es1* gene. Moreover, the high C18:0 mutation in CAS-3 is only expressed in seeds (Cantisán et al., 1999), which supports the alteration of a gene mainly expressed in this tissue. Hongtrakul et al (1998a) have shown that SAD-17 is more strongly expressed in developing seeds than in other tissues or developmental stages.

Apart from the main codominant locus on LG 1, the SAD-17 gene probe also detected a second dominant locus on LG 4 (Fig. 3). Hongtrakul et al. (1998a) reported the isolation of two Δ9-stearate desaturases that were highly expressed in seeds: SAD-17 and SAD-6 (U91339). These two genes have a nucleotide identity of 84.1% (Hongtrakul et al., 1998a) and they do in fact cross-hybridise to one another and to this locus on LG 4 (unpublished data). The main SAD-6 locus was not polymorphic in this cross, but it actually maps to LG 11 (unpublished data). Lacombe et al. (1999) have shown that a Δ9-stearate desaturase gene can detect a high level of polymorphism between inbred lines of sunflower, when used as an RFLP probe. This is probably due to the fact that the probe is detecting three loci simultaneously on three different linkage groups. No high C18:0 QTL were detected on LGs 4 and 11 in this population.

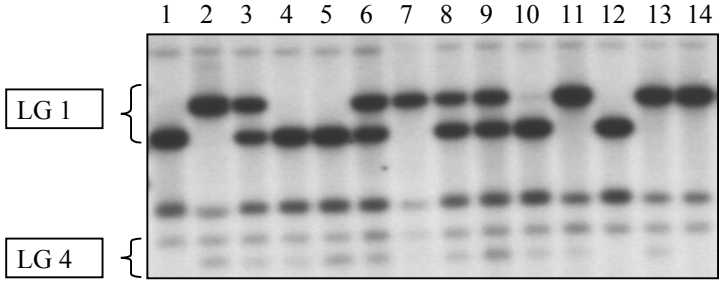


**Figure 3.-** Southern analysis of the $^{32}$P-labelled SAD-17 probe hybridized to *Dra*I digested DNA from the high C18:0 line CAS-3 (lane 1), the low C18:0 line HA-89 (lane 2), and twelve $F_2$ individuals from the cross HA-89 x CAS-3 (lanes 3 to 14).

ACKNOWLEDGEMENTS

This work was funded by a postdoctoral grant from the Spanish Ministerio de Educación y Cultura. We are especially grateful to Alberto Leon and Martín Grondona for their statistical advice.

## REFERENCES


Berry, S.T., Allen, R.J., Barnes, S.R., and Caligari, P.D.S. (1994) Molecular analysis of *Helianthus annuus* L. 1. Restriction length polymorphism between inbred lines of cultivated sunflower. Theor. Appl. Genet. 89: 435-441.

Berry, S.T., Leon, A.J., Hanfrey, C.C., Challis, P., Burkholz, A., Barnes, S.R., Rufener, G. K., Lee, M., and Caligari, P.D.S. (1995) Molecular marker analysis of *Helianthus annuus* L. 2. Construction of an RFLP linkage map for cultivated sunflower. Theor. Appl. Genet. 91: 195-199.

Berry, S.T., Leon, A.J., Peerbolte, R., Challis, C., Livini, C., Jones, R., and Feingold, S. (1997) Presentation of the Advanta sunflower RFLP linkage map for public research. In "Proc. 19TH Sunflower Res. Workshop", Fargo, USA, January 9-10, pp 113-118.

Cantisán, S., Martínez-Force, E., Álvarez-Ortega, R., and Garcés, R. (1999) Lipid characterization in vegetative tissues of high saturated sunflower mutants. Journal of agricultural and food chemistry 47: 78-82.

Dehmer, K. J., and Friedt, W. (1998) Development of molecular makers for high oleic acid content in sunflower. Industrial Crops and Products 7: 311-315.

Harwood, J.L. (1996) Recent advances in the biosynthesis of plant fatty acids. Biochim. Biophys. Acta 1301: 7-56.

Hongtrakul, V., Slabaugh, M.B., and Knapp, S. (1998a) DFLP, SSCP, and SSR markers for -9-stearoyl-acyl carrier protein desaturases strongly expressed in developing seeds of sunflower: intron lengths are polymorphic among elite inbred lines. Molecular Breeding 4: 195- 203.

Hongtrakul, V., Slabaugh, M.B., and Knapp, S. (1998b) A seed specific -12 oleate desaturase gene is duplicated, rearranged, and weakly expressed in high oleic acid sunflower lines. Crop Sci. 38: 1245-1249.

Kabbaj, A., Vervoort, V., Abbott, A.G., Tersac, M., and Bervillé, A. (1996) Expression d'une stéarate et d'une oléate désaturases chez le tournesol normal et à haute teneur en acide oléique, clonage de fragments génomiques et variabilité chez quelques *Helianthus* OCL 3: 452-458.

Kinney, A.J. (1999) New and improved oils from genetically-modified oilseed plants. Lipid Technol. Newsletter, April 1999, pp 36-39.

Knapp, S.J., Berry, S.T., and Rieseberg L.H. (In press) Genetic mapping in sunflowers. In "DNA-Based Markers in Plants", Volume 1, Second Edition, Eds. R.L. Philips and I.K. Vasil, Kluwer Academic Publishers.

Lacombe, S., Lambert, P., Cellier, F., Casse F., and Bervillé, A. (1999) RFLP profiles in low oleic sunflower using *SDI*-, a stearoyl-ACP, and an oleoyl-PC desaturases c DNAs. Helia 22: 19-28.

Lander, E., Green, P., Abrahanson, J., Barlow, A., Daley, M., Lincoln, S., and Newburg, L. (1987) MAPMAKER: an interactive computer package for constructing primary genetic linkage maps of experimental and natural populations. Genomics 1: 174-181.

Lander, E. S., and Botstein, D. (1989) Mapping Mendelian factors underlying quantitative traits using RFLP linkage maps. Genetics 121: 185-199.

Leon, A.J., Lee, M., Rufener, G.K., Berry, S.T., and Mowers, R.P. (1995) Use of RFLP markers for genetic linkage analysis of oil percentage in sunflower seed. Crop Sci. 35: 558-564.

Osorio, J., Fernández-Martínez, J.M., Mancha, M., and Garcés, R. (1995) Mutant sunflowers with high concentration of saturated fatty acids in the oil. Crop Sci. 35: 739-742.

Pérez-Vich, B., Garcés R., and Fernández-Martínez, J.M. (1999) Genetic control of high stearic acid content in the seed oil of sunflower mutant CAS-3. Theor Appl Genet: 99: 663-669.