Genetic Mapping of Factors Affecting Quantitative Variation for Flowering in Sunflower

A. J. Leon, F. H. Andrade, and M. Lee*

ABSTRACT

The number of days from seedling emergence to flowering (DTF) is a major consideration in sunflower (*Helianthus annuus* L.) breeding because the maximum yield of the crop can only be achieved if the cultivars are phenologically adapted to the production environment. Identification of genetic factors that affect flowering could create opportunities for improved breeding methods and for more fundamental investigations of this important trait and its interactions with the environment. The objectives of this study were to locate quantitative trait loci (QTL) for DTF in an elite sunflower population evaluated in four environments. Two hundred thirty-five F_2-generation plants and their F_3 progeny of a single-cross population of two divergent inbred lines were evaluated in four environments (Fargo, ND and Venado Tuerto, Daireaux, and Balcarce in Argentina). Detection of QTL was facilitated with a genetic linkage map of 205 loci defined by restriction fragment length polymorphism (RFLP) and composite interval mapping. Five QTL of five linkage groups accounted for 89% of the genetic variation for DTF. Gene action was additive at four QTL and dominant at the other locus. Three QTL were detected in all environments and generations. The parental effects and the relative magnitudes of the genetic effects of those QTL were consistent across generations and environments.

The genetic and environmental controls of flowering in sunflower are certainly complex and mostly undefined. Our abilities to investigate and manipulate the phenotype in selection programs could be enhanced with improved resolution of genetic factors that influence flowering and the rate at which genotypes proceed from seedlings to anthesis. Most genetic studies of flowering in sunflower have assessed the phenotype as the number of days from seedling emergence to anthesis (DTF). Polygenic inheritance patterns have been reported in most studies (Stoenescu, 1974; Machacek, 1979), although there is some evidence of genetic factors with major, qualitative effects (Jan, 1986). Additive gene action has the greatest influence on flowering (Miller et al., 1980; Roath et al., 1982; Reid, 1992; Alvarez et al., 1992; El-Hity, 1992), but dominant effects have been noted (Jan, 1986). Estimates of broad-sense heritability have ranged from 0.62 to 0.95 (Shabana, 1974; Alvarez et al., 1992; Berretta de Berger and Miller, 1984; Miller and Fick, 1997).

The genetic components of flowering in sunflower have not been described within the context of contemporary genetic analysis and molecular linkage maps. Thus, there is very limited information on genetic factors affecting flowering or DTF, their locations in the genome, and their linkage and interaction with other genes, traits, and environmental cues. To our knowledge, genetic analysis of DTF in sunflower with a comprehensive genetic map based on DNA marker loci has not been done. Therefore, our understanding of this complex trait would be advanced through genetic mapping of QTL with DNA markers. Ultimately, such information could facilitate marker-assisted selection in breeding programs and other more fundamental inquiry. The objectives of this study were (i) to locate QTL for DTF using replicated progeny evaluated in four environments and (ii) to compare detection of QTL for DTF using individual plants in the F_2 generation and their F_3 families.

MATERIALS AND METHODS

Germplasm and Field Design

A cross was made between nonrestorer inbred lines (B lines) ZENB8 (female) and HA89 (male). A single plant of the F_1 generation was self-pollinated to create seed of the F_2 generation. ZENB8, a proprietary inbred line, flowers ≈75 d after planting at photoperiods (15–16 h) and temperatures typical of growing seasons of the locations used in this study (Fargo, ND and Venado Tuerto, Daireaux, and Balcarce in Argentina). HA89, released by the USDA, flowers ≈65 d after planting under the same conditions at these locations.

The F_2 generation was planted at Fargo on 14 May 1992. Two seeds per hill were sown with a hand planter and thinned to one plant per hill. The space between rows was 75 cm and the distance between hills within a row was 30 cm. Five rows of each parent and the F_1 were planted at different periods (−10, −5, 0, +5, +12 d relative to the F_2 planting date) to estimate the within-row error variance (Leon et al., 1995). Before anthesis, individual heads of the F_2-generation plants were covered with pollination bags to ensure self-pollination and production of F_3 generation seed. Two hundred thirty-five $F_{2:3}$ families were planted with a hand planter at Daireaux, Venado Tuerto, and Balcarce on 17, 18, and 20 Nov. 1992. One row per family was planted at each location. Fifteen replicates of each parent and the F_1 hybrid were included to provide an estimate of the error variance within and across locations. Rows were 3 m long and contained ten hills. The space between rows was 70 cm. Three seeds per hill were planted and seedlings were thinned to one plant per hill. The families, parents, and F_1 were randomly assigned to plots at each location.

The number of days from emergence (VE) to 50% flowering (R5.5) was recorded for individual F_2 plants and their corresponding progenies (growth stages as defined by Schneiter and Miller, 1981). The day of flowering of an F_3 progeny row was the day when 50% of the plants reached the R5.5 stage.

The RFLP map and segregation data have been described

A.J. Leon, Zeneca Semillas, Balcarce Research Station, CC: 30, (7620) Balcarce, Pcia. de Bs. As., Argentina; F.H. Andrade, Mar del Plata University, (7620) Balcarce, Pcia. de Bs. As., Argentina; M. Lee, Dep. of Agronomy, Iowa State University, Ames, IA 50011-1010. Received 4 Dec. 1998. *Corresponding author (mlee@iastate.edu).

Published in Crop Sci. 40:404–407 (2000).

Abbreviations: cM, centimorgans; DTF, days to flowering; LOD, $\log_{10}$ of the odds; QTL, quantitative trait loci; RFLP, restriction fragment length polymorphism.

previously (Berry et al., 1995; Leon et al., 1995, 1996). The 205 RFLP loci covered 1380 centimorgans (cM) and were arranged in 17 linkage groups, the haploid number of chromosomes in this species. The average interval size was 5.9 cM. The genetic map was constructed using MAPMAKER version 3.0 (Lander et al., 1987). Genotypic classes at 23 loci deviated significantly from the expected ratios. Those loci exhibited a deficiency in the ZENB8 homozygous class. The majority of the loci with deviant ratios (18 of 23 loci) mapped to four regions, representing Linkage Groups G, L, and P (Berry et al., 1995).

Statistical Analysis

To estimate the total phenotypic variability due to genetic effects, the broad-sense heritability was estimated according to Allard (1966, p. 96) for individual plants in the F_2 generation (Leon et al., 1995). The within-row variance in the F_2 generation was estimated by pooling within-row variances of the parent and F_1 rows. The error variance among rows was estimated in the F_2 generation. Genetic variation was then estimated by subtracting the within- and among-row variances from the phenotypic variance (Leon et al., 1995). For the F_3 families, broad-sense heritabilities were estimated using variance components according to Fehr (1987, p. 96). The heritability on a plot basis is given by (for each location) $h^2 = \sigma_g^2/(\sigma_e^2 + \sigma_{G\times E}^2)$ and heritability on an entry-mean basis (across locations) by $h^2 = \sigma_g^2/(\sigma_e^2/rt + \sigma_{ge}^2/t + \sigma_g^2)$, where σ_g^2, σ_e^2, and $\sigma_{G\times E}^2$ are genotypic variance, experimental error variance, and genotype × environment variance, respectively, with t and r the number of environments and replications within environments, respectively. Estimates of σ_e^2 within and across locations were obtained from the parents and F_1, according to Hallauer and Miranda (1988, p. 113). The significance of the genotype × environment interaction was tested according to Hallauer and Miranda (1988, p. 113), using the σ_e^2 estimated from the parents and F_1 across locations.

Composite interval mapping (Zeng, 1994) was used for mapping QTL. Phenotypic data consisted of trait values for each F_2 plant or F_3 family evaluated at each location and the average value of the F_3 families across locations (the mean environment). The use of single replicates of each family in each environment has been described previously for QTL mapping in maize for grain yield (Stuber et al., 1992; Beavis et al., 1994) and plant height (Beavis et al., 1991). Computations were performed with PLABQTL Version 1.1 (Utz and Melchinger, 1996) as described in detail by Bohn et al. (1996) and Austin and Lee (1998). The initial analysis was made with the "first" statement to check the database for errors and outliers. A second analysis was conducted to select cofactors using the "model D" and "scan" statement with a $\log_{10}$ of the odds (LOD) threshold value of 2.5. The third analysis was done adding the preselected cofactors in the "cov" statement and the "smodel" statement for detection of digenic epistatic interactions between QTL that had significant main effects. The coefficient of determination (R^2) of the model for the mean environment (the average of the other environments) was compared with the broad-sense heritability to calculate the amount of genetic variation associated with the RFLP loci.

The QTL and their positions were used in simultaneous multiple regression to estimate the additive (a) and dominance (d) effects for the F_2 and F_3 generations. The DTF data identified the median of a row (when 50% of the plants in a row have flowered) in the F_3 generation and not the mean value of the trait because the DTF of the later plants of the row were not recorded (per all other studies of QTL for flowering in crop species). Therefore, dominance effects for DTF were not multiplied by a factor of two. Under the assumption of $|d| \leq |a|$, when we measure the median DTF of an F_3 family derived from an F_2 plant heterozygous at a specific locus (QTL), we are directly assessing the DTF of heterozygous genotypes in the row. For F_3 progeny from a heterozygous F_2 plant for a specific locus, the expected mean is $E(\bar{Y}_{F3,H}) = \mu + 0.5d$ (Falconer, 1981) and $d = 2\{E(\bar{Y}_{F3,H}) - [E(\bar{Y}_{F3,A}) + E(\bar{Y}_{F3,B})/2]\}$, where $E(\bar{Y}_{F3,H})$, $E(\bar{Y}_{F3,A})$, $E(\bar{Y}_{F3,B})$ are the expected average values of an F_3 progeny from self-pollinating an F_2 plant heterozygous at a given locus, homozygous for the allele from parent A at a given locus, and homozygous for the allele from parent B at a given locus, respectively; μ is the population average; and d is the dominance deviation. Thus the difference between the heterozygous and the midparent mean values needs to be multiplied by a factor of two. When a median value is used instead of a mean value, the previous approach yields upwardly biased estimates of dominance because $E(\tilde{Y}_{F3,H}) = \mu + d(1 - 2p)$, where $\tilde{Y}_{F3,H}$ is the median value, p is the probability of having at least 50% of the n F_3 plants in a row (from a heterozygous F_2 plant at a given locus) homozygous (AA) at that locus. For an n value of 10 plants in a row, p is equal to 0.02; then $E(\tilde{Y}_{F3,H}) = \mu + 0.96d$ ($\cong \mu + d$), and is a direct estimate of dominance effects in the F_2 generation. The above expression can be extended to other generations and different numbers of plants in a progeny by adjusting the p value.

The d/a (dominant/additive) ratio scale described by Edwards et al. (1987) was used to classify gene action [A = additive or partial dominance ($0 < |d/a| < 0.55$); D = partial dominance or dominance ($0.55 < |d/a| < 1.20$); OD = overdominance ($|d/a| > 1.20$)].

Table 1. Means, variance components, and broad-sense heritabilities for days to flowering (DTF) for the ZENB8 × HA89 sunflower population.

	Venado Tuerto	Daireaux	Balcarce	Fargo	Mean environment
	DTF means†				
ZENB8	69 ± 2.0†	72 ± 1.2	74 ± 1.0	86 ± 2.0	76 ± 0.6
HA89	59 ± 2.0	67 ± 1.2	62 ± 1.0	80 ± 2.0	68 ± 0.6
F_1	60 ± 2.0	66 ± 1.2	64 ± 1.0	78 ± 2.0	67 ± 0.6
F_2				81 ± 2.0	
F_3	63 ± 2.0	68 ± 1.2	67 ± 1.0		
	Variance components‡				
σ_e^2	1.01	4.12	1.57	9.16	4.21
σ_g^2	10.89	10.04	13.80	13.98	5.60
$\sigma_{g\times e}^2$					0.71
σ_{ph}^2	11.90	14.16	15.37	23.14	
H	0.92	0.71	0.90	0.60	0.82

† Mean number of days from seedling emergence to 50% flowering (R5.5 stage) ± 2 standard errors of mean.

‡ σ_e^2 = experimental error variance, σ_g^2 = genotypic variance, $\sigma_{g\times e}^2$ = genotype × environment interaction variance, σ_{ph}^2 = phenotypic variance, H = Broad-sense heritability.

Table 2. Analysis of variance for days to flowering (DTF) for 235 F_3 families of the ZENB8 × HA89 sunflower population evaluated at four environments.

Source of variation	MSE†	F test
Environment (location)	14503.7	2959.9***
Family (genotype)	49.7	10.1***
Family × environment	4.9	1.2
Error‡	4.2	

** Significant at the 0.001 probability level.

† Mean square error.

‡ Variance error was estimated from the parents and F_1 that were replicated 15 times in each environment.

Table 3. Summary of quantitative trait loci (QTL) associated with days to flowering (DTF) in the mean environment for the ZENB8 × HA89 sunflower population.

Linkage group	Position†	RFLP loci‡	LOD§	R^2¶	*a*#	*d*††	|*d*/*a*|‡‡	Gene action§§
	cm							
A	38	C0266–C0341	10.8	19.1	−2.22	0.04	0.02	A
B	64	C1735–C0741	38.4	52.9	3.19	−2.69	0.84	D
H	70	C0523–C0515	2.7	5.2	1.04	0.26	0.25	A
I	56	C1891–C0851	4.8	9.1	1.11	−0.58	0.52	A
L	54	C0230–C0628	7.8	14.2	1.54	−0.42	0.27	A
Total¶¶				72.9				

† Position of likelihood peak (highest LOD score on a given linkage group).
‡ Restriction fragment length polymorphism loci flanking the likelihood peak of the QTL.
§ LOD is $\log_{10}$ of the odds.
¶ Coefficient of determination (percentage of phenotypic variance explained by the QTL).
# Additive (*a*) value. Negative sign (−) indicates an increase of the mean value of the trait due to HA89 alleles. A positive sign (+) indicates an increase of the mean value of the trait due to ZENB8 alleles.
†† Dominant (*d*) values. A positive sign means dominance for higher values of the trait. A negative value means dominance for lower values of the trait.
‡‡ Absolute ratio of the average dominance and additive effects at a QTL.
§§ A = additive or partial dominance ($0 < |d/a| < 0.55$); D = partial dominance or dominance ($0.55 < |d/a| > 1.20$). Based on the scale of the $|d/a|$ ratio.
¶¶ Estimate of total variance obtained from the simultaneous fit of all QTL detected for DTF.

RESULTS AND DISCUSSION

ZENB8 flowered later than HA89 at each location by 5 to 12 d. The average difference was 8 d. Directional dominance for earliness was indicated, as the F_1 had similar values to HA89 and the means of the F_2 and F_3 generations were between the midparent value and HA89 (Table 1). Coefficients of skewness were positive: 0.61, 0.85, 0.73, 2.14, and 1.17 for Venado, Daireaux, Balcarce, Fargo, and the mean environment, respectively. Broad-sense heritabilities ranged from 0.60 in the F_2 generation at Fargo to 0.92 for the F_3 families at Venado. The heritability estimated on an entry basis in the mean environment was 0.82 (Table 1). These values are similar to those obtained with other populations in other environments (Shabana, 1974; Alvarez et al., 1992; Miller and Fick, 1997). The genotype × environment interaction was not significant (Table 2).

Five QTL were associated with DTF in Linkage Groups A, B, H, I, and L (Table 3). Those QTL accounted for 73% of the phenotypic and 89% of the genotypic variation in the mean environment. Quantitative trait loci in Linkage Groups A and B had the highest LOD scores in each environment and in the mean environment (LOD 10.8 and 38.4, respectively), and they accounted for 84% of the genetic variation associated with RFLP loci. Evidence of additive × dominance digenic epistasis was found between QTL in Linkage Groups A and H. That interaction accounted for 2% of the genetic variation attributable to marker loci. The genetic positions and parental effects of the QTL were very consistent across environments and generations. With the exception of QTL of Linkage Groups H and I, all QTL were detected in every environment.

The genetic effects for higher values of DTF at four QTL were derived from the late-flowering parent, ZENB8. The only exception was the QTL on Linkage Group A. With the exception of the dominant effects at the QTL of Linkage Group B, gene action was additive and in accordance with previous reports (Miller et al., 1980; Roath et al., 1982; Reid, 1992; Alvarez et al., 1992).

In sunflower, DTF is controlled primarily by the genotype, photoperiod, and temperature (Goyne et al., 1977; Marc and Palmer, 1981; Goyne and Hammer, 1982). The lack of a genotype × environment interaction could be explained by the similarity of photoperiods among environments. With the exception of the cool temperatures toward the end of the growing season at Fargo, temperatures among the locations were also very similar throughout the growing season.

The QTL identified in this study could be used for marker-assisted selection for these and related environments. Since it is known that the inbred lines ZENB8 and HA89 are photoperiod sensitive (A. Leon, 1991, unpublished data), further research is being conducted to genetically resolve that component of flowering. Further understanding of the components of DTF and the interaction with the environment will refine the use of marker-assisted selection for modifying DTF for a wider range of environmental conditions and for understanding the influence of DTF on the expression and perception of other traits such as grain quality.

ACKNOWLEDGMENTS

The authors thank Dr. Florin Stoenescu, Nora Costa, and Abelardo de la Vega for supervising the experiments at their research stations; Dr. Simon Berry for reviewing the manuscript; Dr. Martin Grondona for statistical support; and Dr. Ian Bridges and Zeneca Semillas (ADVANTA group) for making this research possible.

REFERENCES

Allard, R.W. 1966. Principles of plant breeding. John Wiley and Sons, New York.

Alvarez, D., P. Ludueña, and E. Frutos. 1992. Variability and genetic advance in sunflower. p. 963–968. *In* Proc. of the 13th Int. Sunflower Conf. Pisa, Italy. Int. Sunflowers Assoc., Paris.

Austin, D.F., and M. Lee. 1998. Detection of quantitative trait loci for grain yield and yield components in maize across generations in stress and nonstress environments. Crop Sci. 38:1296–1308.

Beavis, W.D., D. Grant, M. Albertsen, and R. Fincher. 1991. Quantitative trait loci for plant height in four maize populations and their associations with qualitative genetic loci. Theor. Appl. Genet. 83:141–145.

Beavis, W.D., O.S. Smith, D. Grant, and R. Fincher. 1994. Identification of quantitative trait loci using a small sample of topcrossed and F_4 progeny from maize. Crop Sci. 34:882–896.

Berretta de Berger, A., and J.F. Miller. 1984. Genetic study of two sources of reduced height in sunflower. p. 11–12. *In* Proc. of the

6th Sunflower Res. Workshop. Bismarck, ND. Feb. 1984. Natl. Sunflower Assoc., Bismarck, ND.

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

Bohn, M., M.M. Khairallah, D. Gonzalez-de-Leon, D.A. Hoisington, H.F. Utz, J.A. Deutsch, D.C. Jewell, J.A. Mihm, and A.E. Melchinger. 1996. QTL mapping in tropical maize: I. Genomic regions affecting leaf feeding resistance to sugarcane borer and other traits. Crop Sci. 36:1352–1361.

Edwards, M.D., C.W. Stuber., and J.F. Wendel. 1987. Molecular-marker-facilitated investigations of quantitative-trait loci in Maize. I. Numbers, genomic distribution and types of gene action. Genetics 116:113–125.

El-Hity, M.A. 1992. Genetical analysis of some agronomic characters in sunflower (*Helianthus annuus* L.). p. 1118–1128. *In* Proc. of the 13th Int. Sunflower Conf. Pisa, Italy. Int. Sunflower Assoc., Paris.

Falconer, D.S. 1981. Introduction to quantitative genetics. 2nd ed. Longman, New York.

Fehr, W. 1987. Principles of cultivar development. Vol. 1. McGraw-Hill, New York.

Goyne, P.J., and G.J. Hammer. 1982. Phenology of sunflower cultivars. II. Controlled environments studies of temperature and photoperiod effects. Aust. J. Agric. Res. 33:251–261.

Goyne, P.J., D.R. Woodruff, and J.D. Churchett. 1977. Prediction of flowering in sunflowers. Aust. J. Exp. Agric. Anim. Husb. 17: 475–481.

Hallauer, A.R., and J.B. Miranda. 1988. Quantitative genetics in maize breeding. 2nd ed. Iowa State Univ. Press, Ames.

Jan, C.C. 1986. The inheritance of early maturity and short-stature of *H. annuus* line. *In* Proc. of the 9th Sunflower Research Workshop. Fargo, ND. Natl. Sunflower Assoc., Bismarck, ND.

Lander, E.S., P. Green, J. Abrahamson, A. Batlow, M.J. Daly, S.E. Lincoln, and L. Newburg. 1987. MAPMAKER: An interactive computer package for constructing primary genetic linkage maps of experimental and natural populations. Genomics 1:174–181.

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

Leon, A.J., M. Lee, G.K. Rufener, S.T. Berry, and R.P. Mowers. 1996. Genetic mapping of a locus (*Hyp*) affecting seed hypodermis color in sunflower. Crop Sci. 36:1666–1668.

Machacek, C. 1979. Study of the inheritance of earliness in sunflower (*Helianthus annuus* L.). Genet. Slechteni 15(3):225–232.

Marc, J., and J.H. Palmer. 1981. Photoperiodic sensitivity of inflorescence initiation and development in sunflower. Field Crops Res. 4:155–164.

Miller, J.F., and G.N. Fick, 1997. The genetics of sunflower. p. 441–495. *In* A.A. Schneiter (ed). Sunflower technology and production. Agron. Monogr. 35. ASA, CSSA, and SSSA, Madison, WI.

Miller, J.F., J.J. Hammond, and W.W. Roath. 1980. Comparison of inbred vs. single-cross testers and estimation of genetic effects in sunflower. Crop Sci. 20:703–706.

Reid, R.J. 1992. Combining ability of sunflower inbred lines derived from the United States and Argentina. M.S. thesis. North Dakota State Univ., Fargo.

Roath, W.W., J.J. Hammond, and J.F. Miller. 1982. Genetic effects of days to flowering in sunflower (*Helianthus annuus* L.) under short day regime. p. 247–249. *In* Proc. of the 10th Int. Sunflower Conf. Surfers Paradise, Australia. 14–18 Mar. 1982. Int. Sunflower Assoc., Paris.

Schneiter, A.A., and J.F. Miller. 1981. Description of sunflower growth stages. Crop Sci. 21:901–903.

Shabana, R. 1974. Genetic variability of sunflower varieties and inbred lines. p. 263–269. *In* Proc. of the 6th Int. Sunflower Conf. Bucharest, Romania. 22–24 July 1974. Int. Sunflower Assoc., Paris.

Stoenescu, F. 1974. Genetics. p. 92–125. *In* A.V. Vranceanu (ed.) Floarea-soarelui. Editura Academiei Republicii Socialiste, Bucharest, Romania.

Stuber, C.W., S.E. Lincoln, D.W. Wolff, T. Helentjaris, and E.S. Lander. 1992. Identification of genetic factors contributing to heterosis in a hybrid from two elite maize inbred lines using molecular markers. Genetics 132:823–839.

Utz, H.F., and A.E. Melchinger. 1996. PLABQTL: A program for composite interval mapping of QTL. [Online.] Available at http://probe.nalusda.gov:8000/otherdocs/jqtl/ (verified 29 Sept. 1999). J. Quant. Trait Loci 2:1.

Zeng, Z.B. 1994. Precision mapping of quantitative trait loci. Genetics. 136:1457–1468.