Theor Appl Genet (1995) 91: 195–199

S. T. Berry · A. J. Leon · C. C. Hanfrey · P. Challis
A. Burkholz · S. R. Barnes · G. K. Rufener · M. Lee
P. D. S. Caligari

Molecular marker analysis of *Helianthus annuus* L. 2. Construction of an RFLP linkage map for cultivated sunflower

Received: 10 November 1993 / Accepted: 18 April 1994

Abstract A detailed linkage map of *Helianthus annuus* was constructed based on segregation at 234 RFLP loci, detected by 213 probes, in an F_2 population of 289 individuals (derived from a cross between the inbred lines HA89 and ZENB8). The genetic markers covered 1380 centiMorgans (cM) of the sunflower genome and were arranged in 17 linkage groups, corresponding to the haploid number of chromosomes in this species. One locus was found to be unlinked. Although the average interval size was 5.9 cM, there were a number of regions larger than 20 cM that were devoid of markers. Genotypic classes at 23 loci deviated significantly from the expected ratios (1:2:1 or 3:1), all showing a reduction in the ZENB8 homozygous class. The majority of these loci were found to map to four regions on linkage groups G, L and P.

Communicated by J. W. Snape

S. T. Berry (✉) · C. C. Hanfrey · P. Challis
ZENECA Seeds, Plant Biotechnology,
Jealott's Hill Research Station, Bracknell,
Berkshire, RG12 6EY, UK

A. J. Leon
ZENECA Seeds-Argentina,
Research Department,
7620-Balcarce, Argentina

A. Burkholz · S. R. Barnes
ZENECA Seeds-SES,
Industriepark 15,
B-3300, Tienen, Belgium

G. K. Rufener
ICI Seeds, Research Department,
Slater,
IA 50244, USA

M. Lee
Department of Agronomy,
Iowa State University, Ames,
IA 50011, USA

P. D. S. Caligari
Department of Agricultural Botany, School of Plant Sciences,
University of Reading, Whiteknights,
P.O. Box 221,
Reading, Berkshire,
RG6 2AS, UK

Key words Restriction fragment length polymorphism (RFLP) · Linkage map · *Helianthus annuus* · Sunflower cDNAs

Introduction

The advent of restriction fragment length polymorphisms (RFLPs) as genetic markers (Botstein et al. 1980) has enabled the rapid development of linkage maps for a number of important crop species including maize (Helentjaris 1987), rice (McCouch et al. 1988), wheat (Liu and Tsunewaki 1991), barley (Heun et al. 1991), potato (Gebhardt et al. 1989), soybean (Shoemaker et al. 1992), *Phaseolus vulgaris* (Vallejos et al. 1992) and *Brassica napus* (Landry et al. 1991). The ubiquitous distribution of these markers in plant genomes has enabled the construction of highly saturated maps, thereby allowing the analysis of complex quantitative traits in species such as tomato (Paterson et al. 1988), maize (Edwards et al. 1987) and soybean (Diers et al. 1992).

Cultivated sunflower is a diploid species (2n=2x=34) and is second only to soybean in its importance as an annual oilseed crop. However, the genetics of only a small number of traits have been studied in detail (Miller 1992). Consequently, sunflower does not possess a classical genetic map. The few linkages reported to date include those between the nuclear genetic male sterility (ms) gene and the *T* gene coding for anthocyanin pigmentation in the plant (Leclerq 1966), and the isozyme loci peroxidase 3 (*Prx3*) and phosphoglucomutase 4 (*Pgm4*) (Kahler and Lay 1985). In addition, the use of molecular markers in sunflower has only focused on the identification of interspecific crosses (Krauter et al. 1991) and the taxonomy of the genus *Helianthus* (Choumane and Heizmann 1988; Gentzbittel et al. 1992). A recent fingerprinting study of elite inbred lines has shown that there are high levels of restriction fragment length polymorphism in cultivated sunflower (Berry et al. 1994). This level of polymorphism and the diploid nature of sunflower, together with the developed inbred lines and

suitable segregating populations, suggest that this species should be extremely amenable to RFLP mapping.

In the present paper we describe the construction of an RFLP map for sunflower which will be used to identify quantitative trait loci (QTL) affecting seed oil content and other agronomic traits.

Materials and methods

Plant material

An F_2 population (289 individuals) was made by selfing a single F_1 plant of the proprietary inbred line ZENB8 crossed to the public line HA89 (USDA). The F_2 population thus produced was also segregating for traits such as seed colour, seed oil content, seed length, days to flowering and head and stem diameter.

RFLP probes

A total of 322 cDNAs, isolated from an etiolated seedling library (coded with the prefix C), and 29 *Pst*I genomic clones (coded with the prefix H) had been selected as low-copy RFLP probes according to Berry et al. (1994). Thirty-nine low-copy cDNA clones were also obtained through GIE Cartisol, France, and all 390 clones were screened against the two parental lines, which had been digested with one of five restriction enzymes (*Dra*I, *Eco*RI, *Eco*RV, *Hin*dIII and *Xba*I) in order to identify polymorphic RFLP markers.

DNA isolation, digestion and Southern blotting

Three fully expanded leaves were cut from each F_2 plant and frozen on dry ice for transportation to the laboratory. The leaf tissue was then lyophilized, ground to a fine powder in a mill and DNA extracted using a modified version of the protocol described by Rogers and Bendich (1985). The DNA was quantified fluorimetrically and digested to completion using 4 U of enzyme per microgram of DNA according to the manufacturers instructions (Northumbria Biologicals). Per lane, 10 µg of DNA was loaded onto 1.0% TBE-agarose (Seakem) slab gels and electrophoresis performed overnight at 3 V/cm. The DNA was transferred onto Hybond-N membrane (Amersham) via Southern blotting with 20 × SSC and fixed by baking for 2 h at 80°C, followed by UV cross-linking at 60 mJ in a Stratalinker (Stratagene).

Hybridizations

Probe fragments were generated from recombinant plasmids using the polymerase chain reaction (PCR) and the products gel-purified prior to labelling with [α-^{32}P]dCTP (Amersham) via random priming (Feinberg and Vogelstein 1983). The unincorporated radioactivity was removed by spin-column chromatography using Sephadex G50 (Sigma). Hybridizations were performed at 60°C overnight using standard conditions, and the excess probe was removed by two, 30-min washes in 0.5 × SSC, 0.1% SDS at 60°C. Filters were exposed to X-OMAT AR X-ray film (Kodak) with two intensifying screens at –80°C. After autoradiography, the radioactive probes were removed from the filters by two 20-min washes in 0.1 × SSC, 0.5% SDS at 80°C.

Linkage analysis

Autoradiographs were independently scored twice and if conflicts in scoring arose and could not be resolved, the data were excluded from the analysis. If more than 5% of the RFLP scores for a given locus were missing, the hybridization was repeated. The segregation of the alleles at each locus was checked against the expected ratios for codominant (1:2:1) and dominant (3:1) markers using the chi-square test with a significance level of 5%.

The genetic map was constructed using MAPMAKER version 3.0 (Lander et al. 1987). Initially, a two point linkage analysis was conducted to determine the maximum likelihood recombination fraction and the LOD score for each of the possible pairs of loci. Linkage groups were formed using the "group" command on the two-point data with recombination values less than 0.375 and a constant LOD score of 3.0.

Three-point linkage analyses were conducted for the loci within each group, and these data were used in conjunction with the "order" command. The "order" function was repeated several times for each group to try and find the one containing the largest number of loci. These orders were then tested using the multipoint function "ripple", and the loci whose positions were ambiguous (i.e. those placed automatically at a LOD of 2.0) were noted.

Loci which had been excluded on the basis of the three point linkage data were placed using the "try" command, and loci which were initially unlinked were mapped using the "near" command by increasing the recombination default to 0.50. The Kosambi function was used to obtain the centiMorgan (cM) values (Kosambi 1944).

Multiple loci detected by a single probe were coded with the probe number plus the suffix A, B, C or D, etc. to indicate each duplicate locus.

Results

Probe selection and segregation

From the 390 RFLP clones screened against the parental inbred lines, 213 (55%) detected a polymorphism. The majority of these were revealed in digestions with the restriction enzymes *Eco*RI or *Eco*RV (Table 1). In general, probe/enzyme combinations were chosen on the basis of their ability to reveal clear and simple polymorphism in order to make the interpretation of the RFLP data more reliable. In total, the 213 polymorphic probes detected 235 loci, with only 25 of these scored as dominant markers (3:1).

Twenty-three loci [indicated by an asterisk (*) in Fig. 1] showed distorted segregation ratios ($P < 0.05$), all with a reduction in the ZENB8 homozygous class (Table 2). Of these loci, 18 mapped to four regions on linkage groups G, L and P.

If only the codominant markers are taken into account, within these regions, the HA89 and ZENB8 allele frequencies were 54% and 45%, respectively. In comparison, the allele frequency across the genome as a whole was 51% and 49% for HA89 and ZENB8, respectively.

Table 1 The numbers of probes mapped in conjunction with the five restriction enzymes used to reveal polymorphism between the parental inbreds HA89 and ZENB8

Enzyme	*Eco*RV	*Eco*RI	*Hin*dIII	*Dra*I	*Xba*I	Total
Number of probes	92	89	22	8	2	213

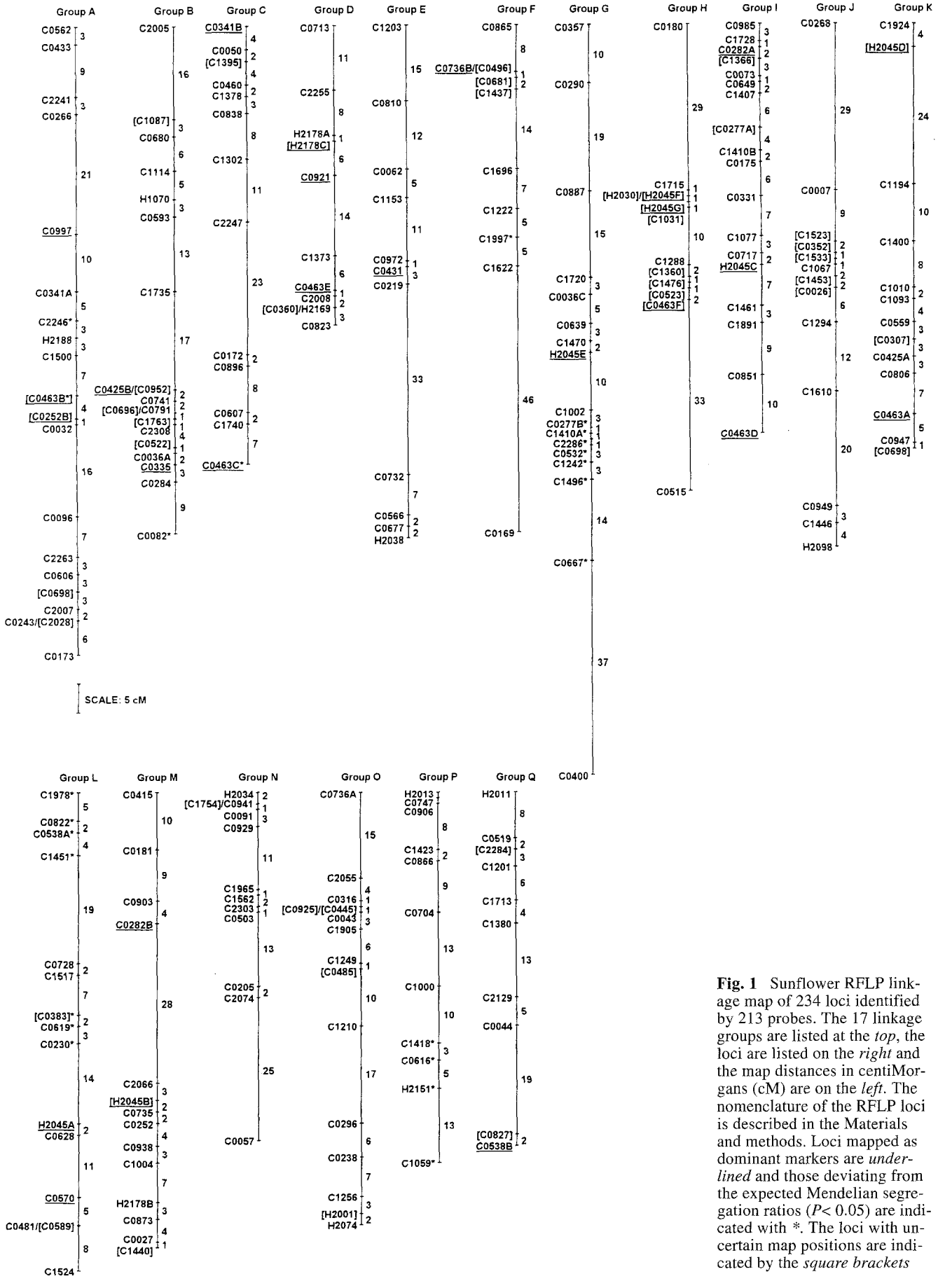

Fig. 1 Sunflower RFLP linkage map of 234 loci identified by 213 probes. The 17 linkage groups are listed at the *top*, the loci are listed on the *right* and the map distances in centiMorgans (cM) are on the *left*. The nomenclature of the RFLP loci is described in the Materials and methods. Loci mapped as dominant markers are *underlined* and those deviating from the expected Mendelian segregation ratios ($P< 0.05$) are indicated with *. The loci with uncertain map positions are indicated by the *square brackets*

Table 2 Loci showing distorted segregation ($P< 0.05$) from the expected ratios of 1:2:1 or 3:1 (*A* HA89 homozygous class · *H* heterozygous class · *B* ZENB8 homozygous class)

Locus	Group	A	H	B	χ^2
C0277B	G	81	147	52	6.71
C1410A		80	155	50	8.51
C2286		82	152	51	8.01
C0532		80	155	52	7.31
C1242		75	161	49	9.55
C1496		82	150	49	9.04
C0667		74	157	48	9.24
C1978	L	72	159	48	9.58
C0822		72	153	48	8.21
C0538A		73	157	51	7.32
C1451		65	161	52	8.18
C0383		88	141	50	10.38
C0619		88	145	51	9.77
C0230		86	139	55	6.88
C1418	P	70	159	46	10.91
C0616		71	161	51	8.20
H2151		67	164	54	7.67
C1059		61	172	53	12.21
C2246	A	67	153	53	6.12
C0463B		–	226	51	6.41
C0082	B	65	161	56	6.25
C0463C	C	97	181	–	14.51
C1997	F	63	152	50	7.02

Table 3 Sunflower RFLP probes detecting multiple loci

Probe	Number of loci	Groups
C0036	2	B, G
C0252	2	M, A
C0277	2	G, I
C0282	2	I, M
C0425	2	B, K
C0463	6	K, A, C, I, D, H
C0538	2	L, Q
C0736	2	O, F
C0341	2	A, C
C1410	2	G, I
H2045	7	L, M, I, K, G, H
H2178	3	D, M

Description of the map

Analysis of the segregation data using MAPMAKER revealed that 234 loci were arranged in 17 linkage groups (Fig. 1), covering 1380 cM of the sunflower genome. One locus (C0592) segregated independently. Linkage groups F and Q have the lowest number of mapped loci. The average map interval was 5.9 cM, but there were 11 regions on 10 linkage groups that were greater than 20 cM. In general, the probes that detected duplicated regions hybridized to 2 unlinked loci (Table 3). However, 2 probes H2045 and C0463 revealed loci on 6 different linkage groups.

The best order of markers on each linkage group (i.e. the one which gave the highest LOD score) is given in Fig. 1; however these orders are not unequivocal. Loci whose positions are uncertain are shown in square brackets on the map because they were either placed at a LOD score of 2.0 or were excluded on the basis of the three-point data.

Discussion

We report here a genetic linkage map for *H. annuus* that comprises 234 marker loci detected by 213 RFLP probes. The presence of duplicated loci within the sunflower genome (Table 3) supports cytogenetic evidence that suggests *H. annuus* has evolved through allopolyploidy (Heiser and Smith 1955; Jackson and Murray 1983). However, more duplicated loci will have to be mapped in order to detect the conservation of linkage blocks between chromosomes as seen in species such as maize (Helentjaris et al. 1988) and *Brassica* (Slocum et al. 1990). The 234 loci were organized into 17 linkage groups (Fig. 1), which probably correspond to the 17 haploid chromosomes of *H. annuus*. The letter used here in identifying each linkage group is completely arbitrary as no classical genetic map exists for sunflower. The coverage of 1380 cM probably represents 60–80% of the sunflower genome, based on comparisons with a random amplified polymorphic DNA (RAPD) map constructed for *H. anomalus* (an interspecific hybrid between *H. annuus* and *H. petiolaris*) that covers 2338 cM (Rieseberg et al. 1993) and our other linkage maps (unpublished data).

The distribution of markers between the 17 groups is fairly uniform with, in general, the largest groups containing the most markers. The differences between the overall lengths of the linkage groups (e.g. group G is 2.5 times bigger than group D) may be related to chromosome size differences, as is the case in tomato (Bernatzky and Tanksley 1986), or to incomplete sampling of the genome. Although the majority of loci are well dispersed (average interval of 5.9 cM), there are 11 regions where the distance between pairs of adjacent markers exceeds 20 cM. Similar "gaps" have been reported in most plant RFLP maps e.g. *Brassica rapa* (Song et al. 1991) and barley (Graner et al. 1991), and they probably represent regions of high recombination in the genome or reflect an under-representation of clones from these areas in the libraries used as probe sources.

The majority of mapped loci were scored as codominant markers and followed the expected 1:2:1 segregation ratio in the F_2 population. However, there were a number of loci that showed distorted segregation (Table 2), and these were found to be localized to four chromosomal regions on linkage groups G, L and P (Fig. 1). Similar findings have also been reported in the RFLP linkage maps of *Brassica napus* (Landry et al. 1991), lettuce (Landry et al. 1987), rice (McCouch et al. 1988) and *Phaseolus vulgaris* (Nodari et al. 1993), among others, and this phenomenon is thought to be due to the presence of deleterious loci in these regions. However, in this study distortion was always due to a reduction in the ZENB8 homozygous class (Table 2). This selection may have been gametic in origin or may

have occurred when the F_2 population was thinned from 2 plants to 1 plant per hill due to the removal of less vigorous plants.

The successful detection of QTL in different segregating populations of sunflower will depend on the continued development of the RFLP linkage map presented here. The mapping of additional cDNA clones, which appear to be a richer source of RFLP markers than genomic clones in sunflower (Berry et al. 1994), should improve the genome coverage of the map and allow the saturation of genic regions with marker loci.

Acknowledgements We would like to thank Simon Griffiths, Carolyn Truman, Angela Woods and Heather Dunne for their technical assistance; Richard Allen and Luke Taylor for their statistical and computing support; and Florin Stoenescu for growing the F_2 population in his sunflower nursery.

References

Bernatzky R, Tanksley SD (1986) Toward a saturated linkage map in tomato based on isozymes and random cDNA sequences. Genetics 112:887–898

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

Botstein D, White RL, Skolnick M, Davis RW (1980) Construction of a genetic linkage map in human using restriction fragment length polymorphisms. Am J Hum Genet 32:314–331

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

Diers BW, Keim P, Fehr WR, Shoemaker RC (1992) RFLP analysis of soybean seed protein and oil content. Theor Appl Genet 83:608–612

Edwards MD, Stuber CW, Wendel JF (1987) Molecular marker-facilitated investigations of quantitative trait loci in maize. I. Numbers, genomic distribution, and types of gene action. Genetics 116:113–125

Feinberg AP, Vogelstein B (1983) A technique for radiolabelling endonuclease fragments to high specific activity. Anal Biochem 132:6–13

Gebhardt RE, Ritter E, Debener T, Schachtschabel U, Walkemeier B, Uhrig H, Salamini F (1989) RFLP analysis and linkage mapping in *Solanum tuberosum*. Theor Appl Genet 78:65–75

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

Graner A, Jahoor A, Schondelmaier J, Seidler H, Pillen K, Fischbeck G, Wenzel G (1991) Construction of an RFLP map of barley. Theor Appl Genet 83:250–256

Heiser CB, Smith DM (1955) New chromosome numbers in *Helianthus* and related genera. Proc Ind Acad Sci 64:250–253

Helentjaris T (1987) A genetic map for maize based on RFLPs. Trends Genet 3:217–221

Helentjaris T, Weber D, Wright S (1988) Identification of the genomic location of duplicate nucleotide sequences in maize by analysis of restriction fragment length polymorphisms. Genetics 118:353–363

Heun M, Kennedy AE, Anderson JA, Lapitan NLV, Sorrells ME, Tanksley SD (1991) Construction of a restriction fragment length polymorphism map for barley (*Hordeum vulgare*). Genome 34:437–447

Jackson RC, Murray BG (1983) Colchicine induced quadrivalent formation in *Helianthus*: Evidence for ancient polyploidy. Theor Appl Genet 64:219–222

Kahler AL, Lay CL (1985) Genetics of electrophoretic variants in the annual sunflower. J Hered 76:335–340

Kosambi DD (1944) The estimation of map distances from recombination values. Ann Eugen 12:172–175

Krauter R, Steinmetz A, Friedt W (1991) Efficient interspecific hybridisation in the genus *Helianthus* via "embryo rescue" and characterisation of the hybrids. Theor Appl Genet 82:521–525

Lander ES, Green P, Abrahamson J, Barlow A, Daly MJ, Lincoln SE, Newburg L (1987) MAPMAKER: An interactive computer package for constructing primary genetic linkage maps of experimental and natural populations. Genomics 1:174–181

Landry BS, Kesseli RV, Farrara B, Michelmore R (1987) A genetic linkage map of lettuce (*Lactuca sativa* L.) with restriction fragment length polymorphism, isozyme, disease resistance and morphological markers. Genetics 116:331–337

Landry BS, Hubert N, Etoh T, Harada JJ, Lincoln SE (1991) A genetic map for *Brassica napus* based on restriction fragment length polymorphisms detected with expressed DNA sequences. Genomics 34:543–552

Leclerq P (1966) Une sterilite male utilisable pour la production d'hybrids simples de tournesol. Ann Amelior Plant 16:135-144

Liu YG, Tsunewaki K (1991) Restriction fragment length polymorphism (RFLP) analysis in wheat. II. Linkage maps of the RFLP sites in common wheat. Jpn J Genet 66:617–633

McCouch SR, Kochert G, Yu ZH, Wang ZY, Khush GS, Coffman WR, Tanksley SD (1988) Molecular mapping of rice chromosomes. Theor Appl Genet 76:815–829

Miller JF (1992) Update on inheritance of sunflower characteristics. In: Proc 13TH Int Sunflower Conf, vol. 2, Pisa, Italy, pp 905–945

Nodari RO, Tsai SM, Gilbertson RL, Gepts P (1993) Towards an integrated map of common bean. 2. Development of an RFLP-based linkage map. Theor Appl Genet 85:513–520

Paterson AH, Lander ES, Hewitt JD, Peterson S, Lincoln SE, Tanksley SD (1988) Resolution of quantitative traits into Mendelian factors by using a complete linkage map of restriction fragment length polymorphisms. Nature 335:721-726

Rieseberg LH, Choi H, Chan R, Spore C (1993) Genomic map of a diploid hybrid species. Heredity 70:285–293

Rogers SO, Bendich AJ (1985) Extracting DNA from milligram amounts of fresh, herbarium and mummified plant tissues. Plant Mol Biol 5:69–76

Shoemaker RC, Guffy RD, Lorenzen LL, Specht JE (1992) Molecular genetic mapping of soybean: Map utilization. Crop Sci 32:1091–1098

Slocum MK, Figdore SS, Kennard WC, Suzuki JY, Osborn TC (1990) Linkage arrangement of restriction fragment length polymorphism loci in *Brassica oleracea*. Theor Appl Genet 80:57–64

Song KM, Suzuki JY, Slocum MK, Williams PH, Osborn TC (1991) A linkage map of *Brassica rapa* (syn. *campestris*) based on restriction fragment length polymorphism loci. Theor Appl Genet 82:296–304

Vallejos CE, Sakiyama NS, Chase CD (1992) A molecular marker-based map of *Phaseolus vulgaris* L. Genetics 131:733–740