Construction of an RFLP Linkage Map for Cultivated Sunflower

Brady A. VICK, C. C. JAN, Jerry F. MILLER, Eugene T. BUTLER*, and Alex L. KAHLER*
USDA-ARS, Northern Crop Science Laboratory, Fargo, ND 58105, and *Biogenetics Services, Inc., Brookings, SD 57006

Abstract

An RFLP linkage map of cultivated sunflower, *Helianthus annuus* L., was constructed based on the segregation of 269 loci detected by 228 cDNA probes. Ninety-three F_2 plants of a cross between inbred lines RHA 271 and HA 234 were used as the mapping population. The linkage analyses were conducted using MAPMAKER 3.0. The genetic markers defined 20 linkage groups, covering 1129 centiMorgans of the sunflower genome. One locus was unlinked and four loci failed to be assigned onto their respective sequences. One hundred ninety-nine of the 269 loci had codominance segregation, with the rest showing dominance segregation. Among the mapped 226 probes (264 loci), 32 segregated for multiple loci, and 43 clusters of tightly linked markers with 0 cM distance among loci. This map has an average marker to marker distance of 4.6 cM, with seven markerless regions exceeding 20 cM. More work is needed to provide a better marker coverage of the entire genome, and to reduce the linkage group to 17, the $2n$ chromosome number of cultivated sunflower.

Key Words: RFLP (restriction fragment length polymophism), linkage map, *Helianthus annuus* L., cDNA

Introduction

Cultivated sunflower (*Helianthus annuus* L., $2n=2x=34$) is one of the few major crops indigenous to the U.S. Sunflower developed into a major crop in the early 1900's, primarily through the breeding efforts of Russian scientists, and presently ranks as the second largest oilseed crop in the world. However, the narrow genetic base of cultivated sunflower has been a common concern among sunflower researchers and could limit its future agronomic success. Improvement of sunflower as a commercial crop will largely depend on a continuous introduction of useful genetic diversity from wild species into breeding programs, and on deciphering the wealth of genetic information encoded in the sunflower genomes. Up to now, only a limited number of genes controlling important traits such as disease resistance, fertility restoration, branching, fatty acid composition, plant height, maturity, and a few morphological markers have been studied in detail (Miller, 1992). A linkage map of classical genes in *Helianthus* does not exist. Relating genes to specific chromosomes is not yet possible, and very few genes are available for linkage studies. Development of a linkage map using only classical genetic markers is laborious and time consuming, and may not be practical for sunflower. Molecular RFLP markers offer great opportunity for the sunflower geneticist to circumvent this difficult situation.

Several recent studies demonstrated successful application of molecular markers in sunflower evaluation. Gentzbittel et al. (1992) described the phylogeny of 44 *Helianthus* species with the use of 10 DNA probes. Relationships among inbred lines of cultivated sunflower were studied by Gentzbittel et al. (1994) and by Berry et al. (1994) using low-copy genomic and cDNA probes. Rieseberg et al. (1995), using 197 RAPD markers, concluded that gene introgression in *Helianthus* involved both chromosomal and genic barriers. Berry et al. (1995) reported the first RFLP map of cultivated sunflower using an F_2 population of ZENB8 x HA 89. The objective of the project reported here was to evaluate polymorphisms among 23 USDA inbred lines and to construct a low density RFLP map for cultivated sunflower.

Materials and Methods

Plant Material. A total of 23 USDA inbred lines, including 12 restorer (RHA) lines, 10 maintainer (HA) lines, and one cms line, were tested in an initial study to evaluate the degree of polymorphism in cultivated sunflower. A total of 409 cDNA probes were evaluated with these lines using three restriction enzymes, *Eco*RI, *Eco*RV, and *Hin*dIII. An additional 206 probes were screened on a subset of three RHA and three HA lines that represented the highest polymorphism among lines. These six lines were grown in the field in the summer of 1992 at Fargo, ND. A half-diallel cross combination was completed using single plants as male and female parents. F_1 plants were grown in the winter of 1993, and F_2 plants were sampled for the RFLP mapping project in the summer of 1994. For this report, 93 F_2 plants of cross RHA 271 x HA 234 were selected for the RFLP linkage mapping.

cDNA Library. A sunflower cDNA library was developed using seedling leaves of HA 89 following the methods of Ullrich et al. (1977), Gilsin et al. (1974), and Gubler and Hoffman (1983). Total RNA was extracted by homogenization of fresh tissue in a guanidinium thiocyanate solution and purified by sedimentation through CsCl. The poly(A)$^+$ RNA fraction was obtained by oligo(dT)-cellulose chromatography and employed as the template in a reverse transcription reaction primed with a dT-tailed *Not*I primer/adaptor oligonucleotide. The cDNA was constructed with 5'-*Eco*RI and 3'-*Not*I sequences and was ligated with the pGEM-11Zf vector. The sunflower cDNAs were inserted in an "anti-sense" orientation relative to the *lac* promoter in the vector to avoid bias against sequences that could generate toxic fusion products in the *E. coli* host, DH5-alpha. Following transformation and selection, the presence and lengths of sunflower cDNA inserts were determined by agarose gel electrophoresis of polymerase chain reaction (PCR) products that were produced either with T7 and SP6 promoter primers or the pUC/M13 "forward" and "reverse" primers.

DNA extraction, digestion, and Southern blotting. Leaf material from plants was rapidly frozen in liquid nitrogen, lyophilized, and ground to a fine powder using a sample mill equipped with a 0.4 mm screen. DNA extractions and purifications were performed by procedures described by Rogers and Bendich (1985) and Doyle and Doyle (1990). Restriction fragments of the various DNAs were "de-salted" prior to agarose gel electrophoresis and Southern blotting to minimize lane-to-lane variability of migration rates.

Transfer of DNA to charge-modified nylon membranes was performed by capillary action under alkaline conditions, and the DNA was fixed to the membrane via ultraviolet irradiation.

RFLP probes. From the total 615 cDNA probes screened, 228 cDNA probes were identified to be single or low-copy polymorphic probes for the two parents, RHA 271 and HA 234, and were used for this study. For each probe, only the best probe-enzyme combination was used for screening the F_2 segregating population.

Linkage analysis. Autoradiographs were scored twice by two different persons and any discrepancies between the two scorings were corrected. The same procedure was followed for entering the raw data to a Lotus ASCII file.

The linkage map was constructed using MAPMAKER 3.0 (Lander et al. 1987). Linkage groups were obtained using the "group" command with a recombination value of less than 0.30 and a LOD score of less than 3.5. For each group, three point analysis was first performed, followed by the "order" command, and the resulting marker order was examined using the "ripple" command to obtain the unique marker sequences. Loci not in the unique sequence were placed using the "try" command. Unlinked loci, and loci that were excluded by the "try" command because of their three point analysis, were not mapped. Unexpected mistakes in scoring, data entering, or double cross-over were further checked with the "error detection" command. Multiple loci detected by a single probe had suffixes a, b, c,... added following the regular probe names.

Results and Discussion

The 228 cDNA probes used for the RFLP linkage study represented 37% of the 615 screened probes. These probes segregated for 269 polymorphic loci. We chose 84 probes for use with *Eco*RI, 72 with *Eco*RV, and 72 with *Hin*dIII. Of the 269 total loci, 199 (74%) loci had codominance segregation and the rest had dominance segregation. We have mapped 226 probes (264 loci) onto 20 linkage groups, covering 1129 cM of the sunflower genome, with an average marker-to-marker distance of 4.6 cM (Fig.1). One marker, 9C2, was not linked to any group, and markers 7A2a, 7D6a, 13A3, and 22D4c failed to be placed properly onto their respective groups. The number of markers per group varied from 2 to 27, and the average between-marker distance varied from 2.21 to 15.8 cM. There were 43 clusters of tightly linked markers, 2 to 10, throughout the genome with 0 cM distance between markers within each cluster. The clustering of so many markers reduced the resolving power of this map, as well as its practical utilization for sunflower improvement.

Among the 226 probes, there were 194 single-loci probes, 27 2-loci probes, four 3-loci probes, and one 4-loci probe. The multiple loci probes represented 14% of the total probes and 28% (70) of the markers on the map. Since the majority of the loci of the multiple probes segregated independently, carefully scored multiple-loci from single probes will likely add more usable markers to the map and should be considered for our future work. In general, the multiple loci of each probe linked to different groups, or to the same group

but with substantial distance from each other. The only exceptions were the loci of probes 1A6, 8B3, and 13B3 which were relatively close to each other. In addition, this duplication of loci may be also an indication supporting the polyploid origin of diploid *H. annuus* (Jackson and Murray, 1983).

There were 13 between-marker distances over 20 cM covering 339.1 cM map distance, which is 30% of the total mapped genome. These results agreed well with those of Berry et al. (1995), who reported 11 regions with between-marker distance of greater than 20 cM. Most of the large gaps should disappear when more probes are added to the map. The use of cDNA probes from other parts of the sunflower plant, the use of genomic DNA probes, or the use of different mapping populations may also help to reduce the number of gaps.

It is apparent that our 20 linkage groups do not match the 17 haploid chromosome number of the cultivated sunflower. The map of Berry et al. (1995) covered 1380 cM, which they believed to represent 60-80% of the sunflower genome. With addition of more markers in the future, the extra linkage groups in our map should merge with other groups to give a total of 17. In addition, RFLP markers will be eventually linked to particular chromosomes of the sunflower genome. If markers of two linkage groups can be shown to link to the same chromosome, then those two groups should be part of one linkage group. We are presently constructing a set of sunflower trisomics (Jan et al., 1988) which we will test with representative probes of the 19 linkage groups. Such studies have the potential to not only sort the trisomics into linkage groups for easier cytological studies, but will offer the opportunity to connect two or more linkage groups together if they are all located on the same chromosome.

Due to the lack of cytological tool stocks such as trisomics and the limited number of available classical genetic trait markers, RFLP markers will be an ideal tool for placing genes onto the linkage map and eventually to the chromosomes. Mouzeyar et al. (1995) successfully linked a downy mildew resistance gene to RFLP and RAPD markers. Future work in mapping simply inherited genes such as resistance to different races of downy mildew and rust, fertility restoration, and nuclear male sterility is urgently needed. RFLP probes will also be valuable for studying interspecific gene introgression, and interspecific and intraspecific relationships (Rieseberg et al., 1995,Gentzbittel et al., 1992, Berry et al., 1994).

Acknowledgements

The authors would like to thank James A. Anderson for his many hours of valuable advice and suggestions throughout this work. Assistance of Lynn S. Dahleen and Phillip McClean is also greatly appreciated. Mark Daly's assistance with the MAPMAKER program and Alicia Meyer's assistance with the computer work greatly helped the analysis and the preparation of this manuscript.

References

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

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 marker analysis of *Helianthus annuus* L. 2. Construction of an RFLP linkage map for cultivated sunflower. Theor. Appl. Genet. 91:195-199.

Doyle, J.J., and J.L. Doyle. 1990. Isolation of plant DNA from fresh tissue. Focus 12:13-15.

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

Gentzbittel, L., Y.-X. Zhang, F. Vear, and B. Griveau. 1994. RFLP studies of genetic relationships among inbred lines of the cutivated sunflower, *Helianthus annuus* L.: evidence for distinct restorer and maintainer germplasm pools.

Gilsin, V., R. Crkvenjakov, and B. Cyrus. 1974. Ribonucleic acid isolation by cesium chloride centrifugation. Biochemistry 13:2633-2637.

Gubler, U. and B.J. Hoffman. 1983. A simple and very efficient method for generating cDNA libraries. Genet. 25:263-269.

Jackson, R.C. and B.G. Murray. 1983. Colchicine induced quadrivalent formation in *Helianthus*: Evidence for ancient polyploidy. Theor. Appl. Genet. 64:219-222.

Jan, C.C., J.M. Chandler, and S.A. Wagner. 1988. Induced tetraploidy and trisomics production of *Helianthus annuus* L. Genome 30:647-651.

Lander, E.S., P. Green, J. Abrahamson, H. Barlow, M. Daly, S. Lincoln, and L. Newbury. 1987. MAPMAKER: an interactive computer program for constructing genetic maps of experimental and natural populations. Genomics 1: 174-181.

Miller, J.F. 1992. Update on inheritance of sunflower characteristics. In: Proc. 13th Int. Sunflower Conf. vol.2, Pisa, Italy, pp.905-945.

Mouzeyar, S., P. Roeckel-Drevet, L. Gentzbittel, J. Philippon, D. Tourvieille De Labrouhe, F. Vear, and P. Nicolas. 1995. RFLP and RAPD mapping of the sunflower *Pl*1 locus for resistance to *Plasmopara halstedii* race 1. Theor. Appl. Genet. 91:733-737.

Rieseberg, L.H., C.R. Linder, and G.J. Seiler. 1995. Chromosomal and genic barriers to introgression in *Helianthus*. 141:1163-1171.

Rogers, S.O. and A.J. Bendich. 1985. Extracting DNA from milligram amounts of fresh herbarium and mummified plant tissues. Plant Molecular Biology 5:69-76.

Ullrich, A., J. Shine, J. Chirqwin, R. Pictet, E. Tischer, W.J. Rutler, and H.M. Goodman. 1977. Rat insulin genes: construction of plasmids containing coding sequences. Science 196:1313-1315.

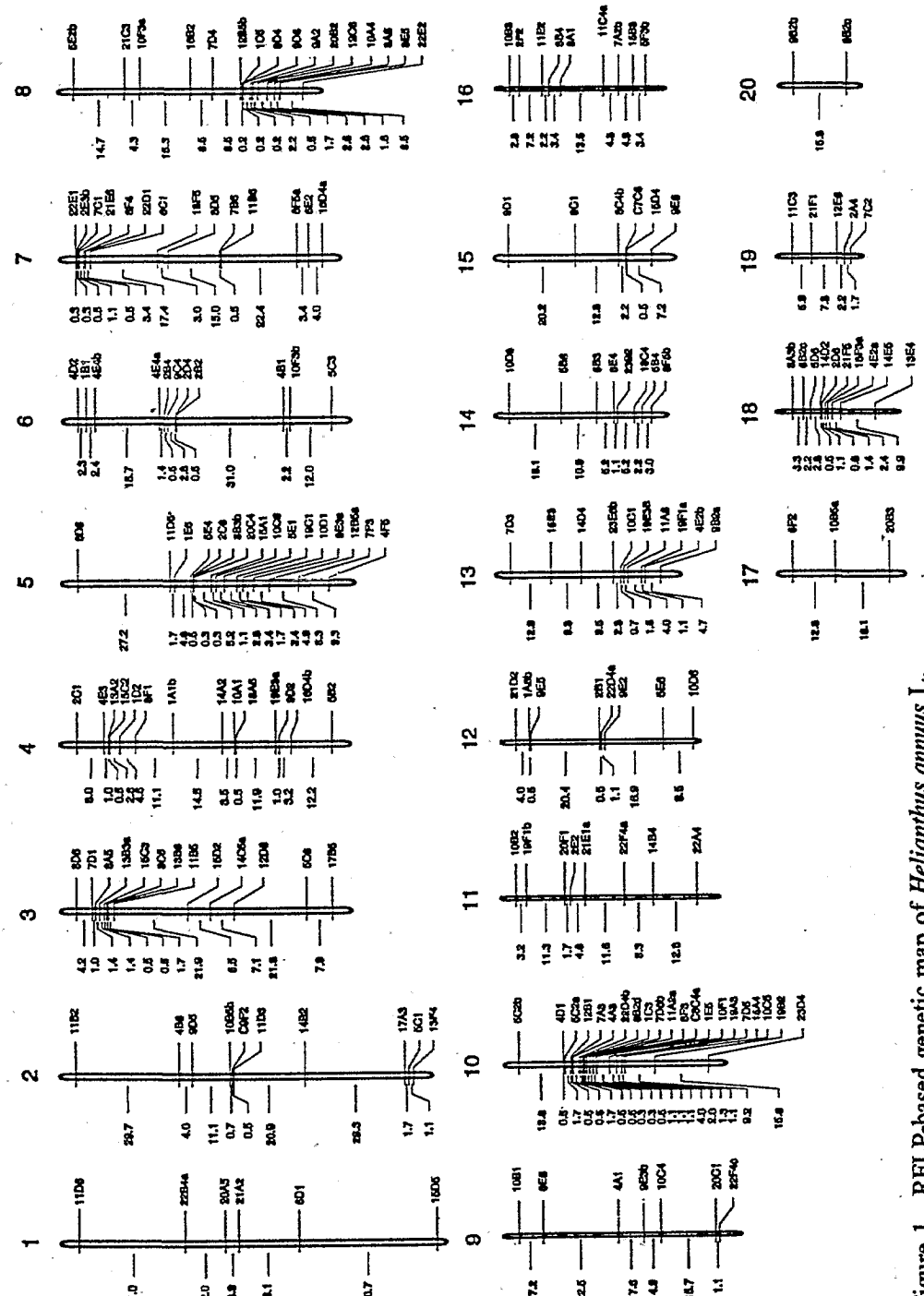

Figure 1. RFLP-based genetic map of *Helianthus annuus* L.