Theor Appl Genet (2002) 105:1124–1136
DOI 10.1007/s00122-002-0989-y

S. Tang · J.-K. Yu · M.B. Slabaugh · D.K. Shintani
S.J. Knapp

Simple sequence repeat map of the sunflower genome

Received: 21 December 2001 / Accepted: 28 March 2002 / Published online: 6 August 2002

Abstract Several independent molecular genetic linkage maps of varying density and completeness have been constructed for cultivated sunflower (*Helianthus annuus* L.). Because of the dearth of sequence and probe-specific DNA markers in the public domain, the various genetic maps of sunflower have not been integrated and a single reference map has not emerged. Moreover, comparisons between maps have been confounded by multiple linkage group nomenclatures and the lack of common DNA markers. The goal of the present research was to construct a dense molecular genetic linkage map for sunflower using simple sequence repeat (SSR) markers. First, 879 SSR markers were developed by identifying 1,093 unique SSR sequences in the DNA sequences of 2,033 clones isolated from genomic DNA libraries enriched for $(AC)_n$ or $(AG)_n$ and screening 1,000 SSR primer pairs; 579 of the newly developed SSR markers (65.9% of the total) were polymorphic among four elite inbred lines (RHA280, RHA801, PHA and PHB). The genetic map was constructed using 94 RHA280 × RHA801 F_7 recombinant inbred lines (RILs) and 408 polymorphic SSR markers (462 SSR marker loci segregated in the mapping population). Of the latter, 459 coalesced into 17 linkage groups presumably corresponding to the 17 chromosomes in the haploid sunflower genome ($x = 17$). The map was 1,368.3-cM long and had a mean density of 3.1 cM per locus. The SSR markers described herein supply a critical mass of DNA markers for constructing genetic maps of sunflower and create the basis for unifying and cross-referencing the multitude of genetic maps developed for wild and cultivated sunflowers.

Keywords Microsatellite · Simple sequence repeat · *Helianthus* · Sunflower

Communicated by A.L. Kahler

Supplementary material Sunflower SSR primer and polymorphism database.

Electronic supplementary material to this paper can be obtained by using the Springer Liuth server located at http://www.dt.doi.org/10.1007/s00122-002-0989-y

S. Tang · J.-K. Yu · M.B. Slabaugh · S.J. Knapp (✉)
Department of Crop and Soil Science, Oregon State University, Corvallis, OR 97331 USA
e-mail: steven.j.knapp@orst.edu
Tel.: +1-541-737-5842, Fax: +1-541-737-1334

D.K. Shintani
Department of Biochemistry, University of Nevada, Reno, NV 89557 USA

Introduction

Cultivated sunflower (*Helianthus annuus* L.), a diploid ($x = 17$) annual, is a globally important source of edible oil, a universally known and widely grown ornamental, and one of 48 species of *Helianthus* (subtribe Heliantheae, subfamily Asteroideae, family Compositae) indigenous to and distributed throughout North America (Heiser and Smith 1955; Heiser 1976; Seiler and Rieseberg 1997). Globally, sunflower ranks fourth or fifth among oilseed crops in land area under production and seed production (Food and Agricultural Organization of the United Nations Statistical Databases 2001; http://www.fao.org). Between 1996 and 2000, about 21.1 million hectares of oilseed sunflower were produced per annuum. Because of the economic and ecological importance of wild and cultivated sunflowers, *H. annuus* has been one of the primary models for genetics and genomics research in the Compositae (Knapp et al. 2001), a family comprised of approximately 2,000 genera and approximately 20,000 species (Stebbins 1974; Cronquist 1977; Jansen et al. 1991).

The key genetic maps of sunflower are anchored by three sets of proprietary restriction fragment length polymorphism (RFLP) markers (Berry et al. 1994; Gentzbittel et al. 1994; Jan et al. 1998). Conservatively, 18 genetic maps of varying density and completeness have been constructed in wild and cultivated sunflowers using about 1,100 RFLP markers (Berry et al. 1995,

1996, 1997; Gentzbittel et al. 1995, 1999; Jan et al. 1998) and several hundred random amplified polymorphic DNA (RAPD) and amplified fragment length polymorphism (AFLP) markers (Rieseberg et al. 1993; Peerbolte and Peleman 1996; Gedil et al. 2001). Despite the wealth of DNA markers and mapping resources in sunflower, a dense genetic map comprised of openly shared, single- or low-copy DNA markers has not been developed, and a single reference map has not emerged. Moreover, the autonomy of the maps has produced a multitude of complete and partial linkage-group assemblages and four linkage-group nomenclatures (68 linkage-group codes), and confounded comparisons between maps (Knapp et al. 2001). The development of a molecular genetic linkage map comprised of openly shared, single- or low-copy DNA markers is crucial for unifying and cross-referencing maps, establishing a universal linkage group nomenclature, and rapidly and efficiently identifying DNA markers for constructing new maps.

The goal of the present research was to construct a dense molecular genetic-linkage map for sunflower based on simple sequence repeat (SSR) markers. The abundance, molecular diversity, and genotyping characteristics of SSRs in plant and animal genomes have been well chronicled (Hamada et al. 1982; Tautz and Renz 1984; Weber and May 1989; Akkaya et al. 1992; Lagercrantz et al. 1993; Morgante and Olivieri 1993; Wang et al. 1994; Röder et al. 1995; Liu et al. 1996; Powell et al. 1996a, b; Taramino and Tingey 1996; Milbourne et al. 1998; Cho et al. 2000). Most importantly, SSR markers form the backbone of second-generation molecular genetic-linkage maps in rice (*Oryza sativa* L.) (Temnykh et al. 2000), barley (*Hordeum vulgare* L.) (Ramsay et al. 2000), soybean (*Glycine max* L.) (Cregan et al. 1999b) and other crop plants. The supply of SSR markers has heretofore been insufficient in sunflower for constructing complete or saturated genetic maps. Using unenriched and enriched genomic DNA libraries, 210 SSR markers have been developed thus far (Brunel 1994; Whitton et al. 1997; Hongtrakul et al. 1998; Gedil 1999; Yu et al. 2002). Of the total, less than half are known to be polymorphic among elite inbred lines (Gedil 1999; Yu et al. 2002). Because the genetic map of sunflower is long (about 1,650 cM or longer) and DNA marker polymorphism rates typically range from 10 to 36% per cross (Berry et al. 1994; Gentzbittel et al. 1994; Yu et al. 2002), several hundred additional SSR markers are needed to saturate the genetic map of sunflower. The development of 879 new SSR markers and a molecular genetic-linkage map for sunflower are described herein. The map was constructed using 459 SSR marker loci.

Materials and methods

SSR discovery, DNA sequence analysis, and primer design

Genomic DNA libraries enriched for AC- or AG-repeats were constructed by Genetic Identification Services (Chatsworth, Calif., USA) using DNA isolated from leaves harvested from 4-week old seedlings of the sunflower inbred line HA370 (Miller and Gulya 1990). DNA was isolated using a modified CTAB method (Webb and Knapp 1990) and enriched using the procedures described by Karagyozov et al. (1993) and Edwards et al. (1996).

Fifty to 100 μl of the genomic DNA library stock solutions were spread onto Bluo-Gal/IPTG/ampicillin (75 μg/ml)-LB agar plates. White clones were isolated and inoculated on ampicillin-LB agar medium in 96-well format plates. The plates were incubated at 37 °C overnight. Clone scrapes were used as templates for colony PCR. The sunflower inserts were amplified using universal M13 forward and reverse primers. Clone insert lengths and numbers were checked on 1.5% agarose gels loaded with 5 μl of PCR solution. Clones with single inserts ranging in length from 350 to 800 bp were selected for sequencing. The amplicons were purified using QIAquick PCR purification kits (Qiagen, Valencia, Calif., USA) processed on a Qiagen BioRobot 3000 (Qiagen, Valencia, Calif., USA). Roughly 25 to 50 ng of each purified amplicon was used as a template for DNA sequencing. We single-pass sequenced 2,033 genomic DNA clone inserts from one end using 1.5 to 3.0 pmol of the universal M13 forward primer. Half of the sequences were produced on an ABI377 Sequencer (Applied Biosystem, Perkin Elmer, Foster City, Calif., USA) in the Oregon State University Center for Gene Research and Biotechnology Central Services Laboratory, and the other half were produced on an ABI3700 Sequencer (Applied Biosystem, Perkin Elmer, Foster City, Calif., USA) in the University of Nevada, Reno, Department of Biochemistry, Genomics Laboratory.

Plasmid sequences were trimmed and the edited DNA sequences were screened for the presence of SSRs. The PILEUP function of the Genetics Computing Group software (Madison, Wis., USA) was used to align sequences harboring common motifs to identify redundant and unique sequences. SSR genotyping primers were designed using PRIMER 3.0 (http://www-genome.wi.mit.edu/cgi-bin/primer/ primer3_www.cgi). Target amplicon lengths, based on the length of repeat and flanking sequences in the reference allele sequence, were systematically varied to create a uniform series of SSR marker-length bins from 130 to 530 bp to facilitate multiplexing.

SSR marker genotyping

SSR primers were screened for utility, functionality, and length polymorphisms on agarose and polyacrylamide gels using two public inbred lines (RHA280 and RHA801) (Fick et al. 1974; Roath et al. 1981) and two proprietary inbred lines (PHA and PHB) developed by Pioneer Hi-Bred International Inc. (Johnston, Iowa). RHA280 is a confectionery fertility restorer line, and RHA801, PHA and PHB are oilseed fertility restorer lines.

SSR genotyping primers were synthesized by MWG Biotech (High Point, N.C., USA) and Perkin Elmer (Foster City, Calif., USA). Forward primers were modified by adding fluorophores (6FAM, HEX, TET or NED) to the 5′ ends. Fluorophores were uniformly distributed among length bins to facilitate multiplexing. SSR genotyping assays were performed using ABI Prism 377 DNA sequencing (polyacrylamide) gels and GeneScan 2.1 and Genotyper 2.0 software (Applied Biosystems, Foster City, Calif., USA). PCRs were performed by using 15 μl of reaction mixture containing 1× PCR buffer, 2.5 mM Mg^{++}, 0.2 μM each of dNTPs, 0.1% Tween-20, 5 to 7.5 pmol of each primer, 0.5–0.75 units of *Taq* polymerase (Qiagen, Valencia, Calif., USA) and 10 to 20 ng of genomic DNA. 'Touchdown' PCR (Don et al. 1991) was used to reduce spurious amplification. The initial denaturation step was performed at 95 °C for 3 min, followed by 1 cycle of 94 °C for 30 s, 64 °C for 30 s and 72 °C for 30 s. The annealing temperature was decreased 1 °C per cycle in subsequent cycles until reaching 58 °C (the base annealing temperature varied from 54 to 60 °C depending on the primer pair). Products were subsequently amplified for 32 cycles at 94 °C for 30 s, 58 °C for 30 s, and 72 °C for 30 s with a final extension for 20 min.

SSR markers were screened for polymorphisms using three-color multiplexes; amplicons labelled with 6FAM, HEX or TET,

or 6FAM, HEX or NED, were separately produced, diluted 15- to 25-fold, and pooled. Samples were prepared for analysis by combining 0.5 µl of diluted amplicon, 0.5 µl of a GeneScan 500 internal-lane standard labelled with TAMRA or ROX, and 50% formamide. Samples were heated to 92 °C for 5 min, chilled on ice for 5 to 10 min, and loaded into 48- or 96-well format gels. GeneScan Filter Set C and the TAMRA 500 internal-lane standard were used for analyses of amplicons labeled with 6FAM, HEX or TET. GeneScan Filter Set D and the ROX 500 internal-lane standard were used for analyses of amplicons labeled with 6FAM, HEX or NED. Genotyper 2.0 was used for allele scoring. Heterozygosities were estimated for each SSR marker as described by Ott (1991).

Genetic mapping

The genetic map was constructed using 94 RHA280 × RHA801 F_7 recombinant inbred lines (RILs). RHA280 was manually emasculated to produce the F_1. The RILs were developed by single-seed descent in summer and winter nurseries, and greenhouses, in Corvallis, Oregon, and Balcarce, Argentina, between 1995 and 1998. SSR genotyping assays were performed, as described earlier, by multiplexing six to 14 SSR markers per lane.

Statistical analyses were performed and the map was constructed using Mapmaker 3.0 (Lander et al. 1987) and G-MENDEL 3.0 (Holloway and Knapp 1993). Chi-square-tests for segregation distortion were performed for each locus using log-likelihood ratio statistics (G). Loci were assembled into groups using likelihood odds (LOD) ratios with a LOD threshold of 3.0 and a maximum recombination frequency threshold of 0.4. Multiple locus-order estimates were produced for each linkage group using Mapmaker (Lander et al. 1987) and G-MENDEL (Holloway and Knapp 1993). The likelihoods of different locus orders were compared and the locus-order estimate with the highest likelihood was selected for each linkage group. Map distances (cM) were calculated from recombination frequency estimates using the Kosambi (1944) mapping function. Recombination frequency estimates in the RIL mapping population were corrected for multiple meioses under selfing using $r = R/(2 - 2R)$, where $R = n_r/(n_r + n_n)$ is the uncorrected recombination frequency and n_r and n_n are the number of recombinant and non-recombinant genotypes, respectively (Haldane and Waddington 1931).

Table 1 Throughput in the process of developing and screening simple sequence repeat markers for sunflower from $(AC)_n$- and $(AG)_n$-enriched genomic DNA libraries

Development stage	AC-Library	AG-Library	Total
Clones sequenced	967	1,066	2,033
Clone sequences harboring SSRs	829	980	1,809
Unique SSR sequences	457	636	1,093
Primer pairs designed and tested	417	583	1,000
SSR markers developed	368	511	879
Polymorphic SSR markers	238	341	579

Results

SSR marker development

The inserts from 2,394 genomic DNA clones enriched for $(AC)_n$ or $(AG)_n$ were amplified by colony PCR. Of the total, 344 had multiple inserts or were too short (less than 350 bp) or long (greater than 800 bp) for marker de-

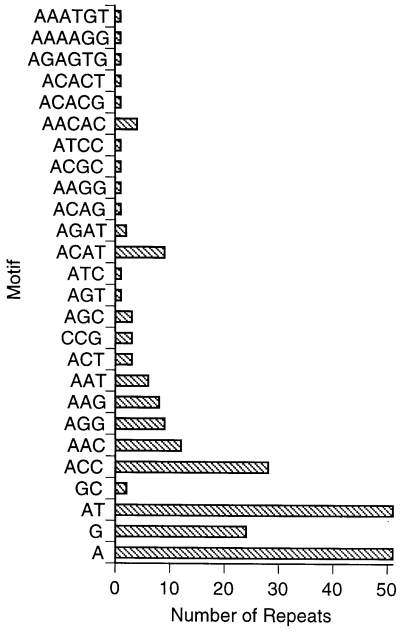


Fig. 1 Spectrum and number of simple sequence repeats other than $(AC)_n$ and $(AG)_n$ identified in the DNA sequences of genomic DNA clones enriched for $(AC)_n$ and $(AG)_n$ in sunflower

Table 2 The number of simple sequence repeat (SSR) primers tested, the number and percentage of functional and polymorphic SSR markers, mean heterozygosities (H) of SSR markers among four inbred lines (RHA280, RHA801, PHA, and PHB), and mean repeat numbers $(\bar{n})$ in reference allele sequences for different repeat types and motifs for 879 SSR markers developed from $(AC)_n$- and $(AG)_n$-enriched genomic DNA libraries in sunflower

Class	Primers tested	SSR markers developed		Polymorphic SSR markers		H	$(\bar{n})$
		Number	%	Number	%		
Compound	117	103	88.0	78	75.7	0.51	15.9
Imperfect	265	222	83.8	129	58.1	0.49	14.5
Perfect	618	554	89.6	372	67.1	0.53	12.4
$(AC)_n$	371	327	88.1	203	62.1	0.51	11.7
$(AG)_n$	539	472	87.6	327	69.3	0.52	14.6
Mononucleotide	8	6	75.0	5	83.3	0.53	10.7
Dinucleotide	918	807	87.9	534	66.2	0.52	13.4
Trinucleotide	50	47	94.0	23	48.9	0.44	8.0
Tetranucleotide or longer	24	19	79.2	17	89.5	0.57	24.9
Total	1,000	879	87.9	579	65.9	0.52	13.3

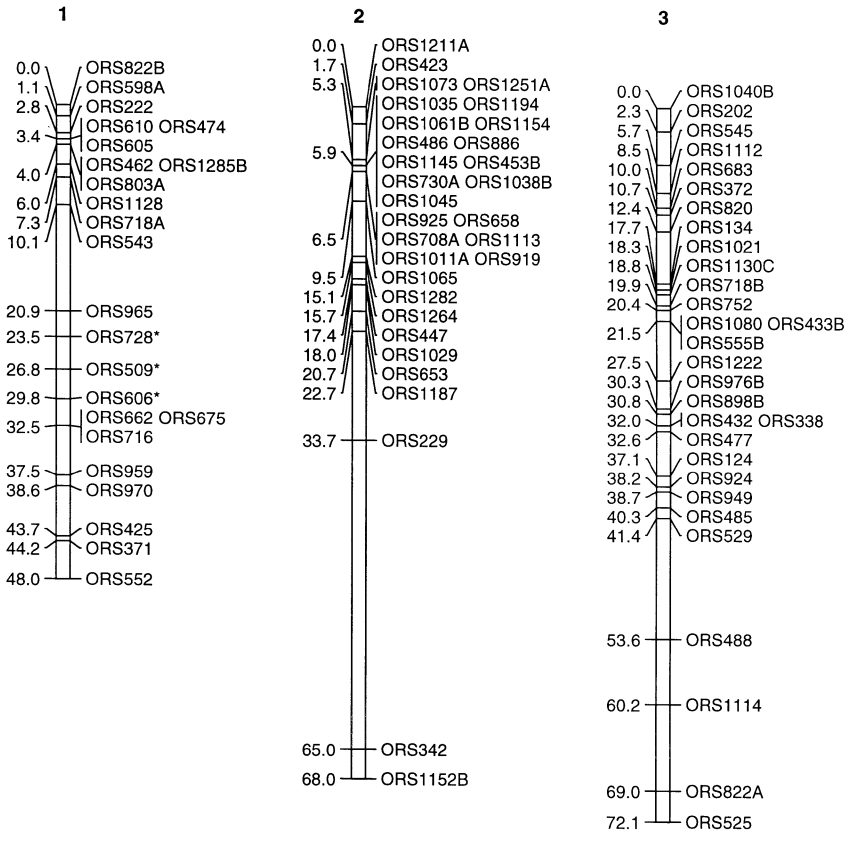


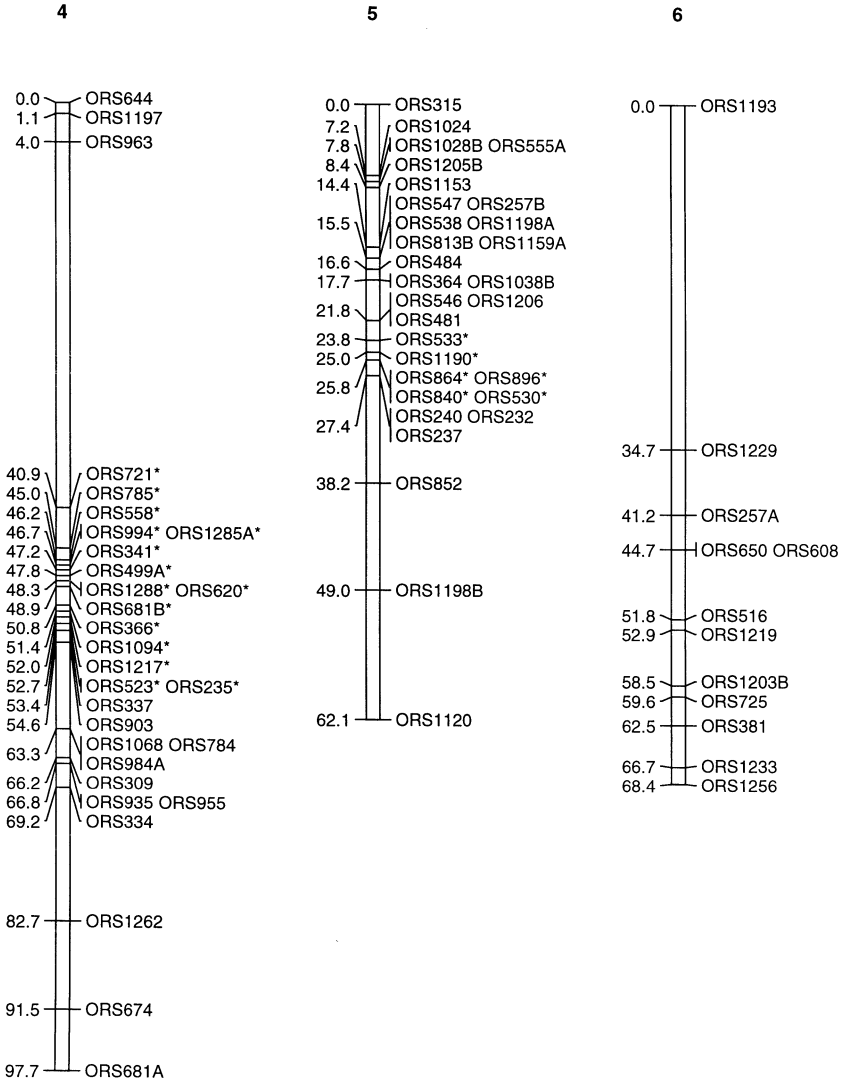
Fig. 2 Molecular genetic linkage map of sunflower based on 459 simple sequence repeat marker loci and 94 RHA280 × RHA801 recombinant inbred lines. Loci with distorted segregation ratios are highlighted with *asterisks*

velopment or one-pass sequencing, respectively. We purified and sequenced 2,050 clone inserts, 977 from the AC- and 1,073 from the AG-enriched library (Table 1). Complete DNA sequences were produced for 2,033 clone inserts, 967 from the AC- and 1,066 from the AG-enriched library. Of the total, 829 clones from the AC library (85.7%) and 980 clones from the AG library (91.9%) harbored $(AC)_n$, $(AG)_n$ or other SSRs. We identified 1,093 unique SSR sequences (60.4%) (Table 1); 61.8% were perfect, 11.7% were compound, and 26.5% were imperfect repeats (Weber 1990) (Table 2). Predictably, the most prevalent repeats were $(AC)_n$ and $(AG)_n$, the motifs used to enrich the genomic DNA libraries (Table 2). We identified 226 non-target repeats (Fig. 1). The most prevelant non-target dinucleotide and trinucleotide repeats were $(AT)_n$ and $(ACC)_n$, respectively (Fig. 1). Four of the 51 $(AT)_n$ repeats were perfect. The others were associated with $(AC)_n$ and $(AG)_n$ in compound repeats, motif combinations commonly found in other plant genomes, e.g., rice (*O. sativa* L.) (Temnykh et al. 2000), maize (*Zea mays* L.) (Taramino and Tingey 1996) and potato (*Solanum tuberosum* L.) (Milbourne et al. 1998). $(AT)_n$ is one of the most prevalent repeats in plant genomes (Lagercrantz et al. 1993; Morgante and Olivieri 1993; Wang et al. 1994); however, because poly-AT probes self-anneal, enrichment methods have not been developed for $(AT)_n$, and $(AT)_n$ repeats have not been specifically targeted in SSR marker development.

The mean number of repeat units $(\bar{n})$ was 13.4 for dinculeotide repeats, 8.0 for trinucleotide repeats, and 24.9 for tetranucleotide, pentanucleotide and hexanucleotide repeats (Table 2). The number of repeat units (n) ranged from 5 to 95, excluding four short dinucleotide repeats ($2 \leq n \leq 4$) found in compound repeats, e.g., $(GT)_6(AT)_4(GT)_{12}(AT)_4$ $(GT)_{20}$ (ORS338). The longest perfect dinucleotide, trinucleotide and tetranucleotide repeats were $(CT)_{89}$ (ORS1250), $(ACC)_{20}$ (ORS1062), and $(GTAT)_{95}$ (ORS908), respectively. The longest repeat (485 bp) was a compound pentanucleotide repeat, $(GTAGT)_{80}(GTCGT)_{17}$(ORS921).

Primer pairs were designed for 1,000 of 1,093 unique SSRs (49.2% of the original clone sequences and 91.5% of the unique SSR sequences); primers could not be designed for 93 unique SSRs because of short flanking sequences. The 1,000 primer pairs yielded 879 SSR markers; 121 primer pairs produced complex or messy amplicons or failed to produce amplicons of the predicted lengths (Table 2). Of the SSR markers, 120 (13.7%) produced null alleles in one or more of the four lines, 351 (39.9%) amplified two or more loci, and 579 (65.9%) were polymorphic among the four inbred lines (Table 1). Mean heterozygosities ranged from 0.44 for trinucleotide to 0.57 for tetranucleotide and longer repeats (Table 2). Mean heterozygosities for dinucleotide, $(AG)_n$, $(AC)_n$, perfect, imperfect and compound repeats ranged from 0.49 to 0.53. While these estimates were based on only four inbred lines (the parents of two RIL mapping populations), they were nearly identical to the mean heterozygosities reported by Yu et al. (2002) for 130 SSR markers among 16 inbred lines (2,080 SSR marker-in-

Fig. 2 (continued)
Legend see page 1127



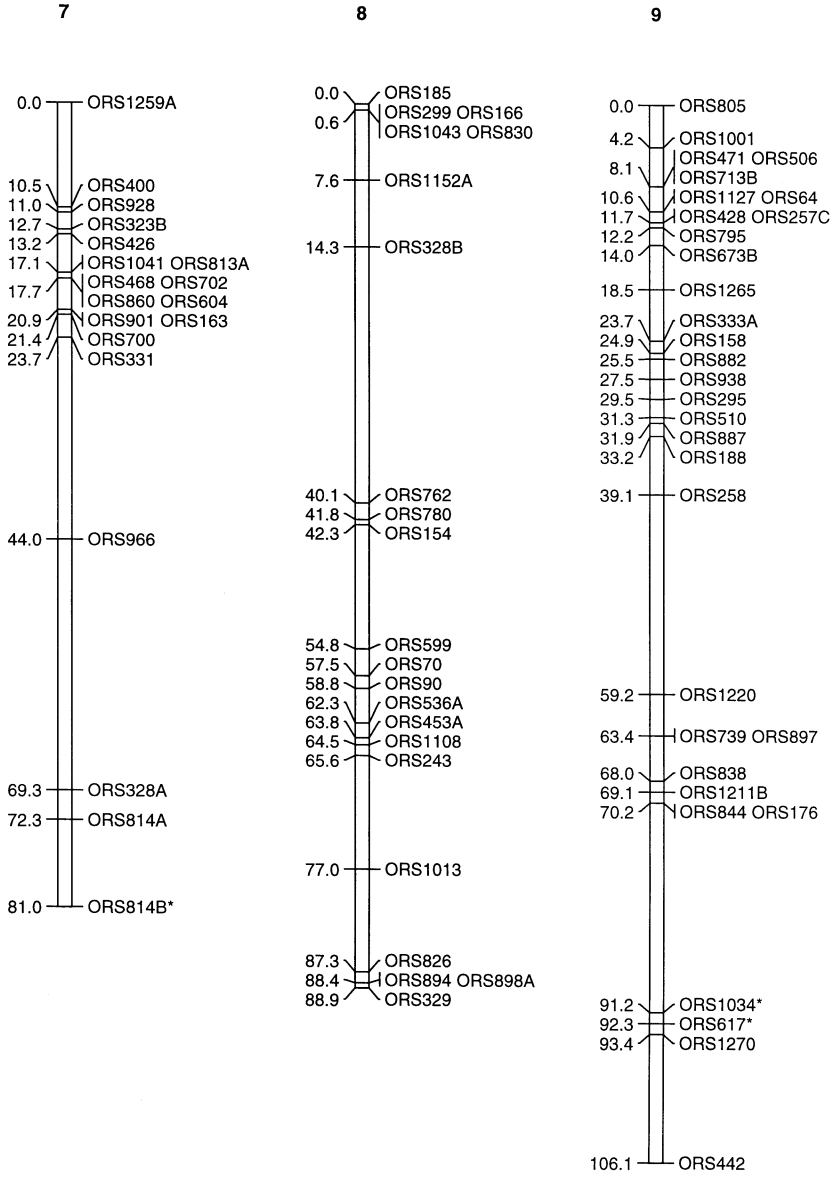
bred line combinations). Because of the multilocus nature of many of the SSRs, 100% of the null alleles could not be unequivocally identified from inbred line fingerprints, e.g., non-allelic bands produced by two inbreds can appear to be allelic (codominant) on fingerprinting gels and segregate as dominant SSR marker loci in mapping populations (see below). Of 3,516 SSR marker-inbred line combinations (4 × 879), 228 produced no bands; thus, assuming complete homozygosity of the inbreds, the frequency of null alleles based on the 'no band' count was 228/3,516 = 0.065.

The genetic linkage map

Of the newly developed SSR markers, 362 (41.2%) were polymorphic in RHA280 × RHA801. When combined with 46 polymorphic SSR markers from the sets developed by Gedil (1999) and Yu et al. (2002), a total of 408 out of 1,089 SSR markers were polymorphic in RHA280 × RHA801. The polymorphic SSR markers produced 462 segregating SSR marker loci. The genetic linkage map was constructed using 459 SSR marker loci (Fig. 2); three loci were unlinked (ORS467A, ORS512B and ORS601). The 459 loci coalesced into 17 linkage groups ranging in length from 48.0 to 106.1 cM (Fig. 2 and Table 3). The map was 1,368.3-cM long and had a mean density of 3.1 cM/locus. Linkage group 6 (LG6) was the least dense (5.7 cM/locus), while LG17 was the most dense (1.6 cM/locus). The number of SSR marker loci per linkage group ranged from 12 on LG6 to 47 on LG10. The longest gap (36.9 cM) flanked ORS963 and ORS721 on LG4 (Fig. 2). The next widest gaps, and the only gaps longer than 30 cM, were between ORS229 and ORS342 on LG2 (31.3 cM), ORS1193 and ORS1229 on LG6 (34.7 cM), and ORS976A and ORS215 on LG13 (30.7 cM). The other pronounced gaps or sparse regions spanned ORS331, ORS966 and ORS328A on LG7, ORS328B and ORS762 on LG8 and ORS782, ORS694 and ORS1086 on LG14.

The 17 linkage groups on the RHA280 × RHA801 map (Fig. 2) correspond to 17 linkage groups on the

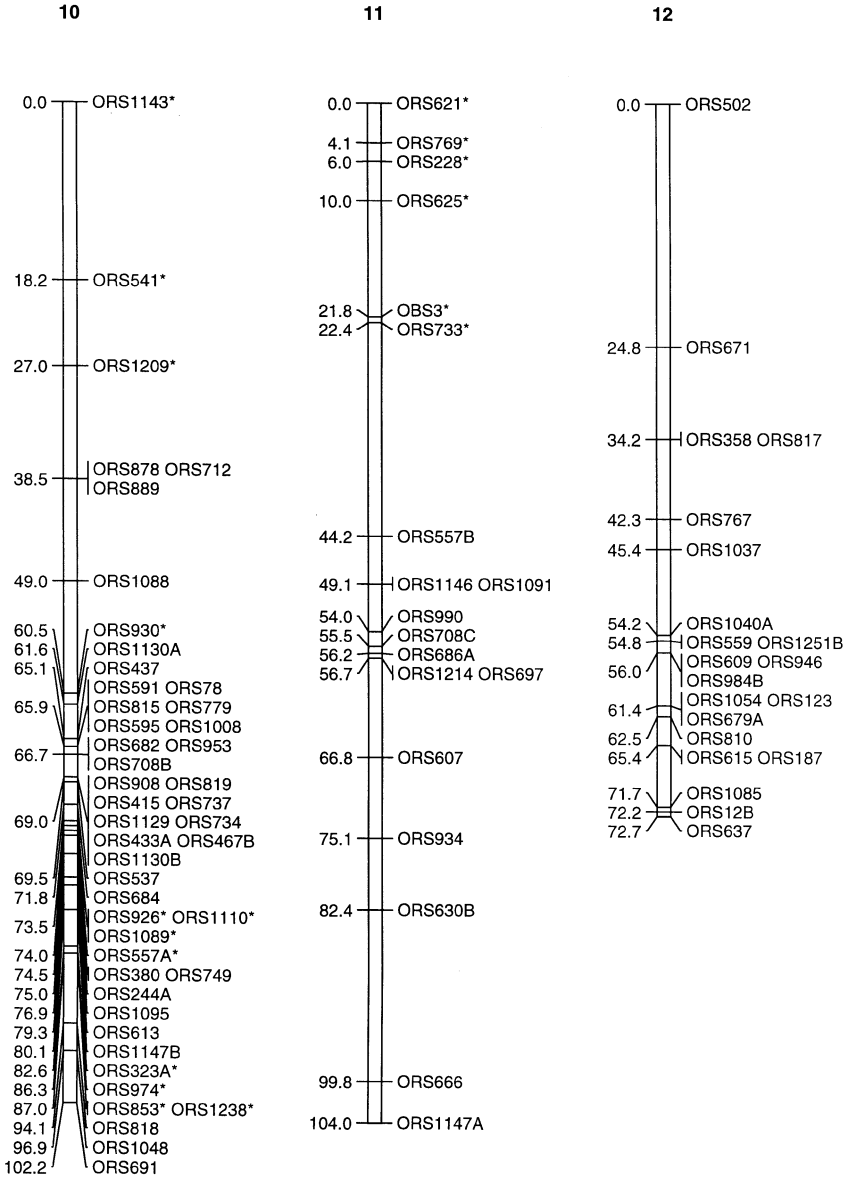
Fig. 2 (continued)
Legend see page 1127



RFLP-SSR map described by Yu (2001) and RFLP maps described by Berry et al. (1996, 1997) and Gedil et al. (2001). The RFLP-SSR map of Yu (2001) was used to cross-reference linkage groups on the present map and the reference RFLP map of Berry et al. (1996, 1997). Presumably, each linkage group corresponds to one of the 17 chromosomes in the haploid sunflower genome ($x = 17$). The probability that the same linkage group or groups are missing on every one of the aforementioned maps or that two or more of the linkage groups belong to one chromosome seems infinitesimally small, particularly since the RHA280 × RHA801 SSR and composite RFLP (Berry et al. 1997) maps alone are based on a total of 1,094 RFLP and SSR marker loci.

Of the SSR markers genotyped in RHA280 × RHA801, 360 amplified a single polymorphic locus and 48 SSR markers amplified two or three polymorphic loci. Of the 462 SSR marker loci genotyped, 263 were codominantly scored and 199 SSR marker loci were dominantly scored. Many of the dominant SSR marker loci were produced by multilocus SSR markers that produced bands in both parents. Once mapped, the bands were identified as being non-allelic and dominant. By using the number of segregating null alleles (199) and an estimate of the minimum number of alleles produced by the complete set of SSR markers (2 × 1,089), the frequency of null alleles in the mapping population was estimated to be 199/2,178 = 0.091. This slightly overestimates the frequency of null alleles because the actual number of alleles produced by the SSR markers was greater than 2,178. SSR markers typically cannot be strictly classified as dominant or codominant because many are multiallelic and, when one of the alleles is a null, are dominant in some crosses and codominant in others. Many of the SSR markers mapped as dominant markers in RHA280 × RHA801 can be codominantly mapped in other crosses.

Fig. 2 (continued)
Legend see page 1127



The mean frequency of heterozygotes (0.0183) was not significantly different from the predicted frequency for F_7 RILs (0.0156). The mean segregation ratio for SSR marker loci across the genome (43,146 datapoints) was 0.99 *AA*: 1.01 *aa*, where *AA* is the RHA280 homozygote and *aa* is the RHA801 homozygote. The mean segregation ratios for SSR marker loci on individual linkage groups ranged from 0.86 *AA*: 1.14 *aa* on LG2, to 1.18 *AA*: 0.82 *aa* on LG11 (Fig. 3). The segregation ratios for 58 SSR marker loci (12.6% of the total) were significantly distorted; 14 produced an excess of RHA801 homozygotes and 44 produced an excess of RHA280 homozygotes. The distorted loci are identified on the genetic linkage map by asterisks appended to the locus names (Fig. 2). LG2, 3, 6, 8, 12, 16 and 17 had no SSR marker loci with significantly distorted segregation ratios (Figs. 2 and 3). Of the 58 distorted SSR marker loci, 49 were clustered on LG1, 4, 5, 10, 11 and 14 (Figs. 3 and 4). The other distorted loci (9 out of 58) were unlinked (ORS467A and ORS512B) or dispersed singly or in tightly linked pairs among other linkage groups. LG7, 9 and 13 had one and LG9 and 15 had two, significantly distorted SSR marker loci each (Fig. 2). The nine dispersed or unlinked loci constituted less than 2% of the mapped markers (less than the probability of Type I errors); thus, the distorted segregation ratios seem to have been predominantly caused by selection operating against one genome or the other in six genomic regions.

RHA280 genotype frequency (f_{AA}) means (boxes in Fig. 3) were less than 0.5 for SSR marker loci on ten linkage groups, greater than 0.5 for SSR marker loci on five linkage groups, and equal or nearly equal to 0.5 for SSR marker loci on two linkage groups (Fig. 3). RHA280 genotype frequencies for the two most-distorted SSR marker loci on each linkage group, in both directions (upper and lower whisker points), are shown in

Fig. 2 (continued)
Legend see page 1127

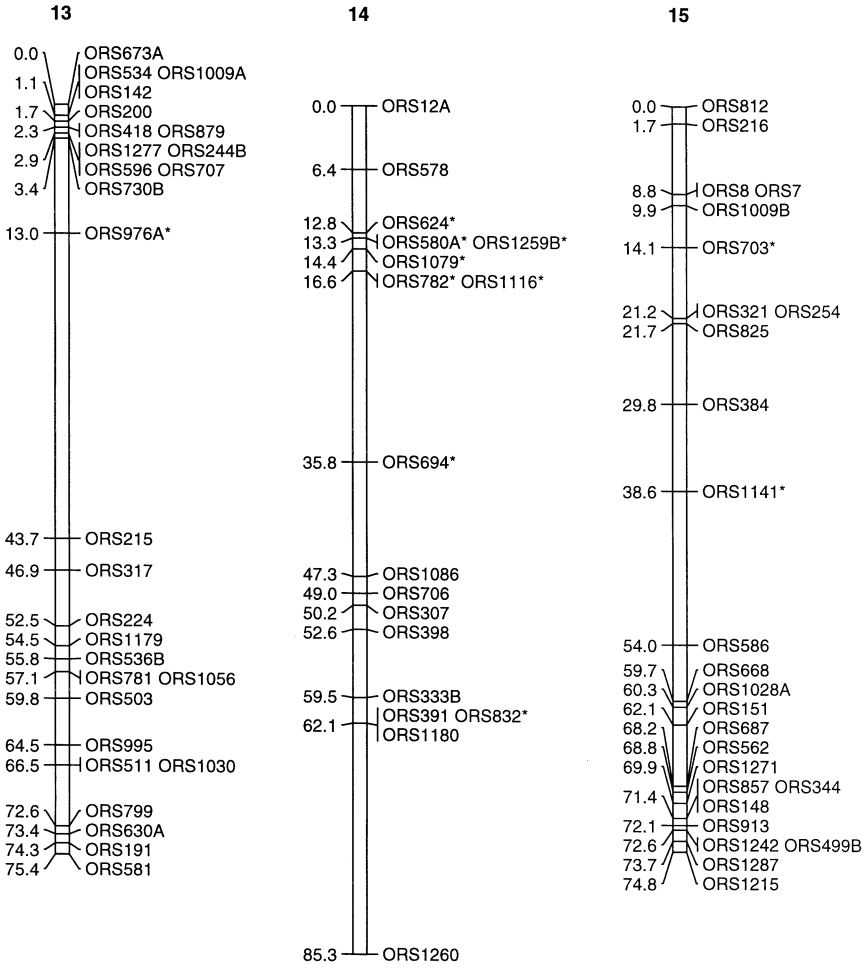


Fig. 3. The segregation ratio extremes were narrowest on LG3, 6, 8, 9, 12 and 17, chromosomes with no or minimal segregation distortion, and greatest on LG4 and 10, chromosomes with significant segregation distortion in both directions. The segregation ratios for linked groups of distorted loci on LG1, 4, 5, 10, 11 and 14 were distorted in one direction and systematically decreased as distances from the most-strongly selected loci increased (Fig. 4). Such patterns are commonplace in plants and the hallmark of differential transmission of gametes, e.g., see Faris et al. (1998) and Konduri et al. (2000). The long gaps on both ends of LG4, the upper end of LG10, and the lower end of LG14 were bordered by genomic regions where the segregation ratios of the SSR marker loci were distorted in opposite directions (Fig. 4). Specifically, on LG4, selection operated against RHA280 homozygotes in upper and lower telomeric regions, and against RHA801 homozygotes in the putative centromeric region. On LG10, selection produced an excess of RHA280 genotypes in the putative centromeric region (bordered on the upper end by ORS930) and a deficiency of RHA280 genotypes in the upper telomeric region encompassing ORS1209. On LG14, selection operated against RHA280 homozygotes in the lower telomeric region (ORS1260) and against RHA801 homozygotes in the upstream region bordering the 23.2-cM gap between ORS1260 and ORS391. The patterns of segregation distortion on LG4, 10 and 11 suggest the presence of strongly selected or "segregation distorter" loci (Lyttle 1991; Faris et al. 1998) flanking regions of high recombination (Fig. 4). Theoretically, the segregation distortion caused by such loci does not adversely affect the estimation of recombination frequencies or locus orders in backcross, doubled-haploid, or RIL progeny. Bailey (1961) showed that the maximum-likelihood estimators of recombination frequency (r_{ML}) were identical for likelihoods based on the differential transmission of gametes (segregation distortion) and the equal transmission of gametes (no segregation distortion) in backcross progeny, and, consequently, doubled-haploids and RILs.

Discussion

The SSR markers described herein supply a critical mass of sequence-based DNA markers for constructing and cross-referencing genetic maps in sunflower, and increase the total number of SSR markers in the public domain to 1,089. Clone and primer sequences, allele lengths, heterozygosities, and other supplemental data for the SSR

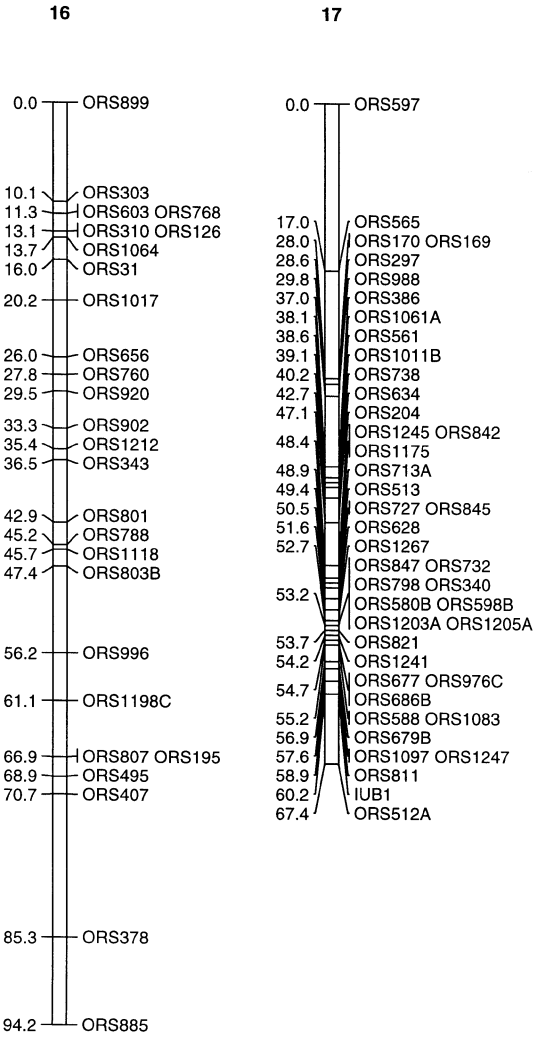


Fig. 2 (continued)
Legend see page 1127

Table 3 Linkage-group length and simple sequence repeat (SSR) marker number and density for each linkage group on the RHA280 × RHA801 genetic linkage map of sunflower

Linkage group	Number of loci	Length (cM)	Density (cM/locus)
1	24	48.0	2.0
2	31	68.0	2.2
3	30	72.1	2.4
4	30	97.7	3.3
5	30	62.1	2.1
6	12	68.4	5.7
7	19	81.0	4.3
8	22	88.9	4.0
9	32	106.1	3.3
10	47	102.2	2.2
11	19	104.0	5.5
12	21	72.7	3.5
13	28	75.4	2.7
14	18	85.3	4.7
15	26	74.8	2.9
16	27	94.2	3.5
17	43	67.4	1.6

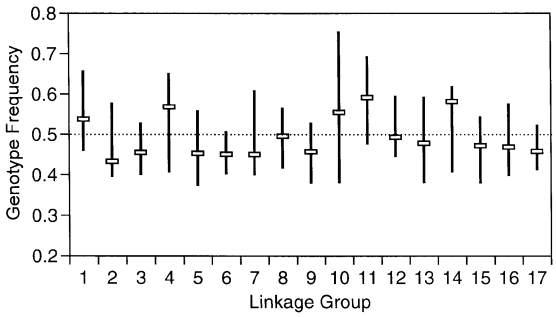


Fig. 3 RHA280 genotype (AA) frequency means (*boxes*) and ranges (*whiskers*) for SSR marker loci on the RHA280 × RHA801 molecular genetic linkage map of sunflower. The upper end of the range is the AA genotype frequency for the SSR marker locus with the greatest excess of RHA280 homozygotes, whereas the lower end of the range is the AA genotype frequency for the SSR marker locus with the greatest excess of RHA801 homozygotes

markers have been deposited in two public databases: http://www.compositdb.ucdavis.edu (the Compositae Genome Database) and http://www.css.orst.edu/knapp-lab. Using 100% of the polymorphic SSR markers between RHA280 and RHA801 (408 SSR markers and 459 SSR marker loci), we constructed a dense and nearly complete molecular genetic linkage map for sunflower (Fig. 2).

Several factors played important roles in the development of the SSR map. First, the polymorphism rate for the RHA280 × RHA801 cross was near the upper end of the range for crosses between elite inbred lines. RHA280 and RHA801 were selected for producing the RILs on the basis of pedigree (Cheres and Knapp 1998) and market class (confectionery and oilseed) differences. Based on subsequent analyses of SSR allelic diversity, RHA280 × RHA801 was found to be one of the most-polymorphic elite × elite crosses in sunflower (Yu et al. 2002; unpublished data). Presumably, the polymorphism rates are comparable for several other confectionery × oilseed crosses. Oilseed × oilseed crosses tend to be less polymorphic than the RHA280 × RHA801 cross; however, some are only slightly less polymorphic. The polymorphism rate for PHA × PHB, a cross between oilseed fertility restorer lines, was 33.2% (362 out of the 1,089 of the SSR markers were polymorphic). Similarly, 285 of the SSR markers (26.1%) were polymorphic in HA370 × HA372 (Yu 2001), a cross between oilseed sterility maintainer lines. Second, the sheer number of SSR markers in the public collection created the critical mass needed to coalesce 17 linkage groups, fill gaps, and produce a dense map (one SSR marker locus per 3.1 cM). Third, the process of genotyping and mapping dominant SSR marker loci (40% of the segregating loci) was facilitated and simplified by using RILs, and potential genotyping errors caused by mistyping of heterozygotes were reduced. The RILs create a permanent, perpetual mapping resource for sunflower. Fourth, genotyping throughput was maximized (by present day standards) by using a multicolor assay system and designing the SSR primers

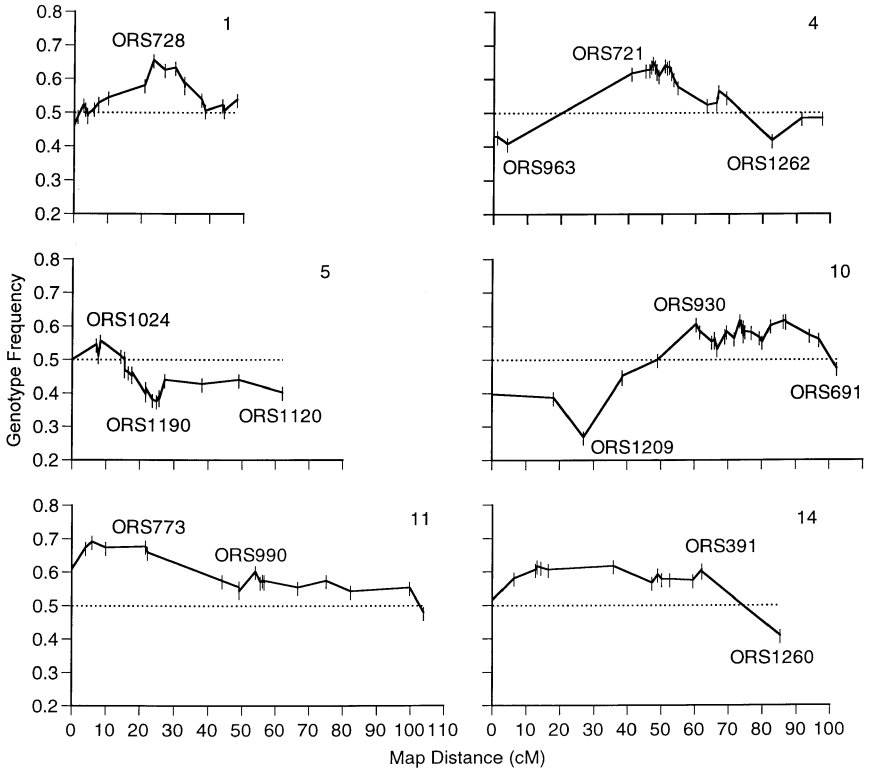


Fig. 4 Frequencies of homozygous RHA280 genotypes for simple sequence repeat marker loci on six linkage groups (LG1, 4, 5, 10, 11 and 14) on the RHA280 × RHA801 molecular genetic linkage map of sunflower

to facilitiate 'pooled amplicon multiplexing' by length. Genotyping assays of mapping population progeny were performed using pools of 6 to 14 SSR markers per lane.

RFLP and SSR markers are rarely dominant in crop plants (Akkaya et al. 1992; Kleinhofs et al. 1993; Morgante and Olivieri 1993; Röder et al. 1995; Liu et al. 1996; Powell et al. 1996a, b; Taramino and Tingey 1996; Smith et al. 1997; Milbourne et al. 1998; Cregan et al. 1999a, b; Ramsay et al. 2000). Sunflower, however, is an oddity in that a significant percentage of RFLP and SSR markers produce null alleles; 9.0% of the SSR markers described by Yu et al. (2002) and 11 to 30% of the RFLP markers described by Berry et al. (1994), Gentzbittel et al. (1994) and Jan et al. (1998) produced null alleles among elite inbred lines. The 11% estimate was for a selected sample of RFLP markers (Berry et al. 1994), whereas the 30% estimate was for a random sample of RFLP markers (Gentzbittel et al. 1994; Jan et al. 1998); therefore, roughly one-third of the randomly selected single- and low-copy RFLP markers were 'dominant' in some sunflower crosses. Dominant SSRs, while rare, are not isolated to sunflower, e.g., 9.0% of the rice (*O. sativa* L.) SSR markers described by Cho et al. (2000) produced null alleles. Null alleles are presumably caused by DNA polymorphisms in primer sites. DNA sequences flanking SSRs in non-coding regions are extraordinarily polymorphic in maize (*Z. mays* L.) (Mogg et al. 2002). We speculate that sequences flanking SSRs in non-coding regions are similarly hypervariable in sunflower and that the null alleles (primer annealing failures) in the present study were primarily caused by primer-site DNA polymorphisms.

Based on length comparisons to dense RFLP maps and a direct comparison between the RHA280 × RHA801 SSR map (Fig. 2), and the HA370 × HA372 RFLP-SSR and PHA × PHB SSR maps described by Yu (2001), the density and distribution of loci supplied by the public SSR markers should be sufficient to assign virtually any new locus to the genetic linkage map of sunflower. The RHA280 × RHA801 map is the densest single-public RFLP or SSR map developed thus far for sunflower and is similar in length (1,368.3 cM) to the densest RFLP maps (Berry et al. 1995, 1996, 1997; Gentzbittel et al. 1995, 1999; Jan et al. 1998). Of the previously developed RFLP maps, the longest (1,573 cM) was the composite map produced by merging seven F_2 maps comprised of 238 RFLP loci (Gentzbittel et al. 1999). The densest and next longest (1,472 cM) was the composite map produced by merging nine F_2 maps comprised of 632 RFLP marker loci (Berry et al. 1996). Of the 16 individual F_2 maps used to construct the two composite RFLP maps, the 234-locus ZENB8 × HA89 F_2 RFLP map of Berry et al. (1995) is the longest (1,380 cM).

Using the reference RFLP map of Berry et al. (1997) as the basis for comparison, two telomeric regions on the RHA280 × RHA801 map seem to lack SSR marker loci, an approximately 20-cM region distal to ORS552 on LG1 and an approximately 36-cM region distal to ORS899 on LG16 (Fig. 2 and Yu 2001). RFLP markers from the public reference map (Berry et al. 1997; Gedil et al. 2001) flank the missing telomeric segments (ZVG4-1 on LG1 and ZVG71-16 on LG16). Conversely, RFLP markers are lacking in an approximately 32-cM

telomeric region on LG3 distal to ORS432 and ORS898 on the RHA280 × RHA801 SSR map (Fig. 2), and ZVG9-3 on the RFLP and RFLP-SSR maps (Berry et al. 1997; Yu 2001). The other gaps in the SSR map (described earlier) spanned internal map segments. Based on the predicted length of the genetic linkage map of sunflower (1,650 cM) (Gentzbittel et al. 1995; Berry et al. 1996), the three unlinked SSR markers and the aforementioned gaps, we speculate that the RHA280 × RHA801 SSR map spans 82% or more of the sunflower genome (Fig. 2). Collectively, RFLPs and SSRs undoubtedly span more than 90% of the sunflower genome. With a dense framework of SSR marker loci in place, the focus needs to shift towards the targeted development of DNA markers to close gaps, and increase marker densities in regions flanking biologically and economically important-trait loci (Cregan et al. 1999b; Cardle et al. 2000).

The non-random physical distribution of SSRs in plant genomes (Ramsay et al. 1999; Elsik and Williams 2001) reduced recombination in centromeric regions (Areshechenkova and Ganal 1999; Künzel et al. 2000; Ramsay et al. 2000), and the genomic origin of DNA sequences used for SSR marker development affect the distribution of SSRs on genetic linkage maps. SSRs isolated from genomic DNA libraries (enriched and unenriched) tend to be concentrated in retrotransposon and dispersed repetitive-element DNA (Ramsay et al. 1999), a phenomena manifested by dense clustering of SSR marker loci in some species (Paglia et al. 1998; Ramsay et al. 2000) but not others (Cregan et al. 1999b; Temnykh et al. 2000).

SSR clustering was minimal in sunflower and did not hinder the development of a ‘complete’ genetic linkage map. LG4, 10 and 17 had dense clusters, LG2, 3, 5, 13 and 15 had moderately dense clusters, and the other nine linkage groups had no appreciable clusters of SSR marker loci (Fig. 2). Temnykh et al. (2000) found virtually no clustering of SSR marker loci in rice. Conversely, SSR marker loci were densely clustered on every linkage group on the genetic map of barley (Ramsay et al. 2000). The clustering phenomena was attributed to the frequent association of SSRs with “retrotransposons and other dispersed repetitive elements”, a high density of such elements in centromeric regions, and the phenomena of reduced recombination in centromeric regions (Cuadrado and Schwarzacher 1998; Areshechenkova and Ganal 1999; Ramsay et al. 1999, 2000; Künzel et al. 2000). Similar non-random associations have been described in conifers and have hindered the coalescence of linkage groups into numbers corresponding to the haploid chromosome number (Paglia et al. 1998; Elsik and Williams 2001). The clustering phenomena were substantially less dramatic in sunflower. Moreover, centromeric clustering of DNA markers is not restricted to SSRs in sunflower. Dense, presumably centromeric, clusters of RFLP marker loci (based on 582 cDNA and 30 genomic DNA probes) were present in several linkage groups on the genetic map described by Berry et al. (1996). The parallel patterns of clustering of SSRs and RFLPs in the sunflower genome (the SSRs sampled 459 non-coding regions and the RFLPs sampled about 600 coding regions) starkly contrasted with the patterns reported in barley (Ramsay et al. 2000). Perhaps the forces underlying genomic rearrangements in sunflower have been more dynamic.

The scarcity of colinearly arranged tracts of duplicated genes in sunflower, an ancient polyploid (Heiser and Smith 1955; Rieseberg 1991; Gentzbittel et al. 1994, 1995; Zhang et al. 1995; Berry et al. 1996; Seiler and Rieseberg 1997; Sossey-Alaoui et al. 1998; Bert et al. 2001), suggests that the genome has undergone substantial reorganization since hybridization. Using 133 duplicated RFLP marker loci, Berry et al. (1996) found “no apparent conservation of linkage blocks between homoeologous chromosomes” in sunflower. They proposed two possible mechanisms to account for the absence of conserved linkage blocks, the loss of the progenitor genomes through diploidization (Zhu et al. 1994) and the formation of pseudogenes (Gentzbittel et al. 1995) through the activity of *copia*-like retrotransposons. Predictably, we found no syntenic linkage blocks of duplicated SSR marker loci either (48 SSR primers amplified two or more segregating loci in RHA280 × RHA801); however, ORS536A and B mapped to the duplicated regions on LG8 and 13 described by Bert et al. (2001). Finally, we identified about 600 single-copy SSR markers (40% of the SSR primers amplified multiple loci). The single-copy SSRs supply a critical mass of sequence-based DNA markers for unequivocally identifying and mapping homologous loci across populations, establishing a universal linkage-group nomenclature, and unifying and integrating the multitude of independent genetic maps developed for wild and cultivated sunflowers.

Acknowledgements We are grateful to Alberto Leon of Advanta Seeds (Balcarce, Argentina) for assisting with the development of the RHA280 × RHA801 RILs. This research was funded by grants to S.J. Knapp from the Plant Genome Program of the United States Department of Agriculture National Research Initiative Competitive Grants Program (#98-35300-6166), Advanta, Dow Agrosciences, Pioneer Hi- Bred International and Syngenta. Oregon Agricultural Experiment Station Technical Paper No. 11,828.

References

Akkaya MS, Bhagwat AA, Cregan PB (1992) Length polymorphisms of simple sequence repeat DNA in soybean. Genetics 132:1131–1139

Areshechenkova T, Ganal MW (1999) Long tomato microsatellites are predominantly associated with centromeric regions. Genome 42:536–544

Bailey NTJ (1961) Introduction to the mathematical theory of genetic linkage. Clarendon Press, Oxford

Berry ST, Allen RJ, 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

Berry ST, Leon AJ, Hanfrey CC, Challis P, Burkholz, Barnes SR, Rufener GK, Lee M, Caligari PDS (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 ST, Leon AJ, Challis P, Livini C, Jones R, Hanfrey CC, Griffiths S, Roberts A (1996) Construction of a high density, composite RFLP linkage map for cultivated sunflower (*Helianthus annuus* L). In: Proc 14th Int Sunflower Conf Vol 2, Beijing, China, pp 1155–1160
Berry ST, Leon AJ, Peerbolte R, Challis P, Livini C, Jones R, Feingold S (1997) Presentation of the Advanta sunflower RFLP linkage map for public research. In: Proc 19th Sunflower Res Workshop, Fargo, North Dakota, USA, pp 113–118
Bert PF, Tourvieille de Labrouhe D, Philippon J, Mouzeyar S, Jouan I, Nicolas P, Vear F (2001) Identification of a second linkage group carrying genes controlling resistance to downy mildew (*Plasmopara halstedii*) in sunflower (*Helianthus annuus* L.). Theor Appl Genet 103:992–997
Brunel D (1994) A microsatellite marker in *Helianthus annuus* L. Plant Mol Biol 24:397–400
Cardle L, Ramsay L, Milbourne D, Macaulay M, Marshall D, Waugh R (2000) Computational and experimental characterization of physically clustered simple sequence repeats in plants. Genetics 156:847–854
Cheres MT, Knapp SJ (1998) Ancestral origins and genetic diversity of cultivated sunflower: coancestry analysis of public germplasm sources. Crop Sci 38:1476–1482
Cho YG, Ishii T, Temnykh S, Chen X, Lipovich L, McCouch SR, Park WD, Ayres N, Cartinhour S (2000) Diversity of microsatellites derived from genomic libraries and GenBank sequences in rice (*Oryza sativa* L.). Theor Appl Genet 100:713–722
Cregan PB, Mudge J, Fickus EW, Marek LF, Danesh D, Denny R, Shoemaker RC, Matthews BF, Jarvik T, Young ND (1999a) Targeted isolation of simple sequence repeat markers through the use of bacterial artificial chromosomes. Theor Appl Genet 98:919–928
Cregan PB, Jarvik T, Bush AL, Shoemaker RC, Lark KG, Kahler AL, Kaya N, VanToai TT, Lohnes DG, Chung J, Specht JE (1999b) An integrated genetic linkage map of the soybean genome. Crop Sci 39:1464–1490
Cronquist A (1977) The Compositae revisited. Brittonia 29:137–153
Cuadrado A, Schwarzacher T (1998) The chromosomal organization of simple sequence repeats in wheat and rye chromosomes. Chromosoma 107:587–594
Don RH, Cox PT, Wainwright BJ, Baker K, Mattick JS (1991) 'Touchdown' PCR to circumvent spurious priming during gene amplification. Nucleic Acids Res 19:4008
Edwards KJ, Barker JH, Daly A, Jones C, Karp A (1996) Microsatellite libraries enriched for several microsatellite sequences in plants. BioTechniques 20:758–760
Elsik CG, Williams CG (2001) Families of clustered microsatellites in a conifer genome. Mol Genet Genomics 265:535–542
Faris JD, Laddomada B, Gill BS (1998) Molecular mapping of segregation distortion loci in *Aegilops tauschii*. Genetics 149:319–327
Fick GN, Zimmer DE, Kinman ML (1974) Registration of six sunflower parental lines. Crop Sci 14:912
Gedil MA (1999) Marker development, genome mapping, and cloning of candidate disease resistance genes in sunflower, *Helianthus annuus* L. PhD thesis, Oregon State University, Corvallis
Gedil MA, Wye C, Berry ST, Seger B, Peleman J, Jones R, Leon A, Slabaugh MB, Knapp SJ (2001) An integrated RFLP-AFLP linkage map for cultivated sunflower. Genome 44:213–221
Gentzbittel L, Zhang YX, Vear F, Griveau B, Nicolas P (1994) RFLP studies of genetic relationships among inbred lines of the cultivated sunflower, *Helianthus annuus* L.: evidence for distinct restorer and maintainer germplasm pools. Theor Appl Genet 89:419–425
Gentzbittel L, Vear F, Zhang YX, Berville A, Nicolas P (1995) Development of a consensus RFLP linkage map of cultivated sunflower (*Helianthus annuus* L.). Theor Appl Genet 90: 1079–1086
Gentzbittel L, Mestries E, Mouzeyar S, Mazeyrat F, Badaoui S, Vear F, Tourvieille de Labrouhe D, Nicolas P (1999) A composite map of expressed sequences and phenotypic traits of the sunflower (*Helianthus annuus* L.) genome. Theor Appl Genet 99:218–234
Haldane JBS, Waddington CH (1931) Inbreeding and linkage. Genetics 16:357–374
Hamada H, Petrino MG, Kakunaga T (1982) A novel repeated element with Z-DNA-forming potential is widely found in evolutionarily diverse eukaryotic genomes. Proc Natl Acad Sci USA 79:6465–6469
Heiser CB (1976) The sunflower. University of Oklahoma Press, Norman
Heiser CB, Smith DM (1955) New chromosome numbers in *Helianthus* and related genera. Proc Indiana Acad Sci 64:250–253
Holloway JL, Knapp SJ (1993) G-MENDEL 3.0 user's guide. Oregon State University, Corvallis
Hongtrakul V, Slabaugh MB, Knapp SJ (1998) DFLP, SSCP, and SSR markers for Δ9-stearoyl-acyl-carrier protein desaturase strongly expressed in developing seeds of sunflower: intron lengths are polymorphic among elite inbred lines. Mol Breed 4:195–203
Jan CC, Vick BA, Miller JF, Kahler AL, Butler ET (1998) Construction of an RFLP linkage map for cultivated sunflower. Theor Appl Genet 96:15–22
Jansen RK, Michaels HJ, Palmer JD (1991) Phylogeny and character evolution in the Asteraceae based on chloroplast DNA restriction-site mapping. Systematic Bot 16:98–115
Karagyozov L, Kalcheva ID, Chapman VM (1993) Construction of random small-insert genomic libraries highly enriched for simple sequence repeats. Nucleic Acids Res 21:3911–3912
Kleinhofs A, Kilian A, Saghai Maroof MA, Biyashev RM, Hayes PM, Chen FQ, Lapitan N, Fenwick A, Blake TK, Kanazin V, Ananiev E, Dahleen L, Kudrna D, Bollinger J, Knapp SJ, Liu B-H, Sorrells M, Heun M, Fanckowiak JD, Hoffman D, Skadsen R, Steffenson BJ (1993) A molecular, isozyme, and morphological map of the barley (*Hordeum vulgare*) genome. Theor Appl Genet 86:705–712
Knapp SJ, Berry S, Rieseberg LH (2001) Genetic mapping in sunflowers. In: Phillips RL, Vasil IK (eds) DNA markers in plants. Kluwer, Dordrecht, the Netherlands, pp 379–403
Konduri V, Godwin ID, Liu CJ (2000) Genetic mapping of the *Lablab purpureus* genome suggests the presence of a 'cuckoo' gene(s) in this species. Theor Appl Genet 100:866–871
Kosambi DD (1944) The estimation of map distance from recombination values. Ann Eugen 12:172–175
Künzel G, Korzun L, Meister A (2000) Cytologically intergrated physical restriction fragment length polymorphism maps for the barley genome based on translocation breakpoints. Genetics 154:397–412
Lagercrantz U, Ellegren H, Andersson L (1993) The abundance of various polymorphic microsatellite motifs differs between plants and vertebrates. Nucleic Acids Res 21:1111–1115
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
Liu ZW, Biyashev RM, Saghai Maroof MA (1996) Development of simple sequence repeat DNA markers and their integration into a barley linkage map. Theor Appl Genet 93:869–876
Lyttle TW (1991) Segregation distorters. Annu Rev Genet 25: 511–557
Milbourne D, Meyer RC, Collins AJ, Ramsay LD, Gebhardt C, Waugh R (1998) Isolation, characterization and mapping of simple sequence repeat loci in potato. Mol Gen Genet 259: 233–245
Miller JF, Gulya TJ (1990) Registration of ten oilseed sunflower germplasm lines. Crop Sci 30:430–431
Mogg R, Batley J, Hanley S, Edwards D, Sullivan HO, Edwards KJ (2002) Characterization of the flanking regions of *Zea mays* microsatellites reveals a large number of useful sequence polymorphisms. Theor Appl Genet (in press)

Morgante M, Olivieri AM (1993) PCR-amplified microsatellites as markers in plant genetics. Plant J 3:175–182

Ott J (1991) Analysis of human genetic linkage. John Hopkins University Press, Baltimore, Maryland

Paglia P, Oliver AM, Morgante M (1998) Towards second-generation STS (sequence-tagged site) linkage maps in conifers: a genetic map of Norway spruce. Mol Gen Genet 258:466–478

Peerbolte RP, Peleman J (1996) The CARTISOL sunflower RFLP map (146 loci) extended with 291 AFLP markers. In: Proc 18 Sunflower Res Forum, Fargo, North Dakota, pp 174–178

Powell WG, Machray C, Provan J (1996a) Polymorphism revealed by simple sequence repeats. Trends Plant Sci 1:215–222

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

Ramsay L, Macaulay M, Cardle L, Morangte M, Ivanissevich DS, Maestri E, Powell W, Waugh R (1999) Intimate association of microsatellite repeats with retrotransposons and other dispersed repetitive elements in barley. Plant J 17:415–425

Ramsay L, Macualay M, Ivanissevich SD, MacLean K, Cardle L, Fuller J, Edwards KJ, Tuvesson S, Morgante M, Massari A, Maestri E, Marmiroli N, Sjakste T, Ganal M, Powell W, Waugh R (2000) A simple sequence repeat-based linkage map of barley. Genetics 156:1997–2005

Rieseberg LH (1991) Homoploid reticulate evolution in *Helianthus* (Asteraceae): evidence from ribosomal genes. Am J Bot 78:1218–1237

Rieseberg LH, Choi H, Chan R, Spore C (1993) Genomic map of a diploid hybrid species. Heredity 70:285–283

Roath WW, Miller JF, Gulya TJ (1981) Registration of RHA 801 sunflower germplasm. Crop Sci 21:479

Röder MS, Plaschke J, König SU, Börner A, Sorrels ME, Tanksley ST, Ganal MW (1995) Abundance, variability and chromosomal location of microsatellites in wheat. Mol Gen Genet 246:327–333

Seiler GJ, Rieseberg LH (1997) Systematics, origin, and germplasm resources of the wild and domesticated sunflower. In: Schneiter AA (ed) Sunflower technology and production. Am Soc Agron, Madison, Wisconsin, pp 21–66

Smith JSC, Chin ECL, Shu H, Smith OS, Wall SJ, Senior ML, Mitchell SE, Kresovich S, Ziegle J (1997) An evaluation of the utility of SSR loci as molecular markers in maize (*Zea mays* L.): comparisons with data from RFLPs and pedigree. Theor Appl Genet 95:163–173

Sossey-Alaoui K, Serieys H, Tersac M, Lambert P, Schilling E, Griveau Y, Kaan F, Berville A (1998) Evidence for several genomes in *Helianthus*. Theor Appl Genet 97:422–430

Stebbins GL (1974) Flowering plants: evolution above the species level. Belknap Press Harvard, Cambridge, Massachusetts

Taramino G, Tingey S (1996) Simple sequence repeats for germplasm analysis and mapping in maize. Genome 39:227–287

Tautz D, Renz M (1984) Simple sequences are ubiquitous repetitive components of eukaryotic genomes. Nucleic Acids Res 12:4127–4137

Temnykh S, Park WD, Ayres N, Cartinhour S, Hauck N, Lipovich L, Cho YG, Ishii T, McCouch SR (2000) Mapping and genome organization of microsatellite sequences in rice (*Oryza sativa* L.). Theor Appl Genet 100:697–712

Wang Z, Weber JL, Zhong G, Tanksley SD (1994) Survey of plant short tandem DNA repeats. Theor Appl Genet 88:1–6

Webb DM, Knapp SJ (1990) DNA extraction from a previously recalcitrant plant genus. Mol Biol Rep 8:180–185

Weber JL (1990) Informativeness of human $(dC\text{-}dA)_n\cdot(dG\text{-}dT)_n$ polymorphisms. Genomics 7:524–530

Weber JL, May PE (1989) Abundant class of human DNA polymorphisms which can be typed using the polymerase chain reaction. Am J Hum Genet 44:388–396

Whitton J, Rieseberg LH, Ungerer MC (1997) Microsatellite loci are not conserved across the Asteraceae. Mol Biol Evol 14: 204–209

Yu JK (2001) Simple sequence repeat marker development and mapping in cultivated sunflower. PhD thesis, Oregon State University, Corvallis

Yu JK, Mangor J, Thompson L, Edwards KJ, Slabaugh MB, Knapp SJ (2002) Allelic diversity of simple sequence repeat markers among elite inbred lines in cultivated sunflower. Genome (in press)

Zhang YX, Gentzbittel L, Vear F, Nicolas P (1995) Assessment of inter- and intra-inbred line variability in sunflower (*Helianthus annuus*) by RFLPs. Genome 38:1040–1048

Zhu T, Schupp JM, Oliphant A, Keim P (1994) Hypomethylated sequences: characterization of the duplicate soybean genome. Mol Gen Genet 144:638–645