Theor Appl Genet (1999) 99: 218–234

L. Gentzbittel · E. Mestries · S. Mouzeyar
F. Mazeyrat · S. Badaoui · F. Vear
D. Tourvieille de Labrouhe · P. Nicolas

A composite map of expressed sequences and phenotypic traits of the sunflower (*Helianthus annuus* L.) genome

Received: 27 August 1998 / Accepted: 28 December 1998

Abstract A map of the sunflower genome, based on expressed sequences and consisting of 273 loci, was constructed. The map incorporates data from seven F_2 populations, for a total of 1115 individuals. Two hundred and fourty five loci corresponding to 170 anonymous cDNA markers and four loci for morphological markers were mapped. We also mapped 18 loci corresponding to previously described genes or to sequences obtained through homology cloning. The unit maps vary from 774 cM to 1060 cM, with an average value of 14 major linkage groups. The integrated map is arranged in 17 major linkage groups including 238 loci, plus four small segments with 2–5 marker loci; and covers 1573 cM with an overall average marker interval of 7 cM. Thirty five percent of the markers were dominant in nature and 30% showed inter-linkage group duplication without any indication of homoeologous linkage groups. Evidence is provided for the independence of two distinct fertility restoration genes, for the presence of two loosely linked branching loci, and for marker tightly linked to the *Rf1* restoration locus. This map provides an efficient tool in breeding applications such as disease-resistance mapping, QTL analyses and marker-assisted selection.

Key words Sunflower · Linkage mapping · cDNA · RFLP · Phenotypic traits · Composite mapping

Communicated by G. Wenzel

L. Gentzbittel[1]
Gie Cartisol, 20 rue Bachaumont, 75002 Paris, France

E. Mestries · S. Mouzeyar · F. Mazeyrat · S. Badaoui
F. Vear · D. Tourvieille de Labrouhe · P. Nicolas
Groupe de Recherche Auvergne Tournesol-INRA,
Station d'Amélioration des Plantes et de Pathologie végétale,
Domaine de Crouelle, 234 avenue du Brézet,
63000 Clermont-Ferrand, France-UA UBP-INRA OVGV,
Campus des Cézeaux, 24 Avenue des Landais,
63177 Aubiere, France

Present address:
[1] INP-ENSAT, laboratoire de Biotechnologie et Amélioration des Plantes, Avenue de l'Agrobiopole, BP 107 Auzeville-Tolosane,
31326 Castanet Cedex, France
Fax: 33-5-62-19-35-89
E-mail: gentz@ensat.fr

Introduction

The sunflower (*Helianthus annuus* L.) is one of the four major annual world crops grown for edible oil. This oil is rich in unsaturated fatty acids, contains large amounts of vitamin E and is easy to refine. Most cultivars are single-cross hybrids but, unlike crops such as maize, its breeding history is quite recent and the first commercial hybrids appeared in 1969. Breeding programmes have focused basically on yield and oil content, but breeding for disease resistance is also very important. Disease resistance is particularly important in Western Europe, as sunflowers of North American origin were first bred in Russia under much drier conditions. Despite the importance of the sunflower as a food crop, the *Asteraceae* is not a very important crop-plant family and intensive linkage mapping studies have only been made for lettuce (Kesseli et al. 1994; Witsenboer et al. 1997). To-date, three linkage maps of cultivated sunflower have been published (Berry et al. 1995; Gentzbittel et al. 1995; Jan et al. 1998) but only one of these seems to be saturated (Berry et al. 1995).

Molecular marker-based linkage maps are powerful tools for plant-breeding programmes. Genetic maps have been described for a wide range of plants including tomato (Bernatzky and Tanksley 1986) and maize (Helentjaris 1987) as pioneer types, and with avocado (Sharon et al. 1997) and a high-density map of rice (Harushima et al. 1998) as more recent examples. The rapid accumulation of markers and mapping populations is a challenge to the management of information and the merging of separate sets of data in order to

accumulate more valuable information for further research and the better use of genetics. Combined maps provide an easy and convenient way of comparing the component maps and offer important information about the reliability of marker order and the distances between markers. Molecular-based linkage maps have made possible the identification of markers tightly linked to genes of major economic interest, such as race-specific disease resistance loci (Kasuga et al. 1997; Mohan et al. 1997). To-day, the field where molecular markers are used for the study of quantitatively inherited phenotypes is considerable, varying from aroma in rice (Lorieux et al. 1996) or eating quality in maize (Azanza et al. 1996) to yield components (Li et al. 1997) or complex disease resistances (William et al. 1997). At present, one of the major goals is to isolate and characterize genes involved in the expression of important agronomic traits. A candidate-gene strategy is a putative way of relating a phenotype (quantitatively inherited or Mendelian-inherited) to its biochemical basis by demonstrating a co-segregation between this candidate-gene and a phenotype of interest. Beside expressed sequence tags (ESTs), a candidate-gene sequence can be obtained by homology cloning. However this strategy requires preliminary knowledge and has to be pursued using transgenic experiments to validate the candidate-gene function.

With the belief that a detailed molecular linkage map for sunflower would facilitate the mapping of disease resistances and other important agronomic traits, which will, in turn, improve the efficiency of breeding programmes, we report here an expressed-sequence-based linkage map constructed using anonymous cDNA markers as well as known genes. The present study aims at integrating seven individuals maps and we present our lastest release of the sunflower genome mapping project. It consists of a near-saturated linkage map, based on RFLPs and describing linkage analyses of phenotypic traits, with the merging of seven F_2 populations. In order to go further in the creation of a prospective tool, we also attempt to provide evidence for the linkage mapping of known genes, with the idea of creating maps of identified sequences. The mapping of probes involved in defense systems (HSP-70-related, chitinase and PAL) is described. Other stress-related probes (HSP17.6, HSP17.9, ubiquitine and calmodulin) are also mapped. Using some of the maps described in this paper, our research group has mapped downy mildew resistance loci (Mouzeyar et al. 1995; Roeckel-Drevet et al. 1996; Vear et al. 1997), characterized quantitative trait loci (QTLs) for *Sclerotinia sclerotiorum* resistances (Mestries et al. 1998), and identified candidate-genes for downy mildew and *S. sclerotiorum* resistance (Gentzbittel et al. 1998).

Materials and methods

Plant materials

Four F_2 populations were developed by crossing five different parents. The origins of these four progenies and the sample sizes used for the map constructions are presented in Table 1. Three previously described mapping populations were also included in the analyses (Table 1). They were C1 (HA89 × RHA266), C2 (CX × RHA266) and C5 (PAC2 × RHA266) (Gentzbittel et al. 1995). The seven populations are coded as S1, S2, S3, S4, C1, C2 and C5 respectively.

RFLP analysis

Plant DNA isolation and Southern hybridization were adapted from Gentzbittel et al. (1995). Sunflower DNA was extracted from frozen green leaves of field-grown plants by a combination of the proteinase K and CTAB methods. Sample DNA was hydrolysed by *Eco*RI, *Eco*RV, *Hin*dIII and *Bgl*II restriction enzymes. The digested DNAs were separated on 0.8% agarose gels and transferred onto charged Nylon membranes (Nplus, Amersham) using 20 × SSC as a transfer buffer. The DNA was bound to the filters by UV-cross-linking. DNA probes were radio-labelled (20 μCi per probe) by random priming (Amersham) and filters were hybridized overnight at 65°C under standard conditions. Autoradiographies were obtained by exposing membranes to X-ray films with two intensifying screens, for 1–20 days as needed.

Source of molecular markers

RFLP probes

RFLP probes were essentially chosen among those described in a previous experiment (Gentzbittel et al. 1995) both for their

Table 1 Segregating populations used for composite mapping of the sunflower genome

Cross	Genotypes	Characteristics	Number of F_2 plants studied
S1	SD × PAC1	*b1* (apical branching)	139
S2	SD × CP73		175
S3	CP73 × PAC1	*Msc1* (fertility restoration), *b1* (apical branching)	280
S4	GH × PAC2	*Msc1* (fertility restoration), *b1* (apical) and *bbr* (basal branching)	161
C1[a]	HA89 × RHA266	*Rf1* (fertility restoration), *b1* (apical branching)	80
C2[a]	CX × RHA266	*Rf1* (fertility restoration)	130
C5[a]	PAC2 × RHA266	*Rf1* (fertility restoration), *b1* (apical branching)	150

[a] Data from Gentzbittel et al. (1995)

polymorphism between the parental lines and their distribution throughout the previously described map. The probes are anonymous cDNAs selected from different libraries including etiolated seedlings, green leaves, ovaries and florets. Previously unpublished probes (coded numbers above S250) were mapped on at least one of the seven populations.

Genes and characterized sequences

We mapped a set of characterized sequences on the basis of their putative biological functions related to disease or stress resistances. Defense-related genes are phenylalanine ammonia lyase (PAL, Mazeyrat et al. 1998), chitinase (GenBank accession U96640) and a HSP70-related sequence (GenBank accession U96641), all three obtained through homology cloning using degenerate primers. Stress-related genes were ubiquitine (Binet et al. 1989), HSP 17.6 and HSP17.9 (Coca et al. 1994), and calmodulin (Courbou et al. 1997). We also mapped a set of anther-specific expressed genes: SF2 (Domon et al. 1990), SF3 (Baltz et al. 1996), SF15 (Dudareva et al. 1996), SF16, SF17, SF18 (Domon et al. 1990), SF18.61 and SF21 (Krauter-Canham et al. 1997).

Morphological markers

Observations of male sterility/male fertility were made by noting the presence or absence of pollen due to the presence or absence of the *Msc1* (Leclercq 1971) restorer gene in the F_3 plants. The male fertility restoration gene *Rf1* (Kinman 1970; Fick and Zimmer 1974) was previously mapped to linkage group 6 (Gentzbittel et al. 1995). The apical – *b1* gene (Putt 1964) and basal branching were also recorded on F_3 plants. The *b1* gene was previously reported to be located on linkage group 7 (Gentzbittel et al. 1995).

Map construction

Individual maps were constructed for each progeny using the MapMaker/Exp V3.0 software (Lander et al. 1987). A minimum LOD score of 4.0 and a threshold recombination fraction value of 0.35 were chosen. The gene orders were then estimated using a LOD threshold of 2.0. When the most-likely gene order was obtained, putative alternative orders were examined by the use of the 'ripple' command. Distances were expressed in Kosambi centiMorgans. In order to create a composite map, we assumed that markers that mapped to similar linkage-group regions in the different populations identified the same chromosome locus. For composite mapping, it is possible to pool the scores across populations similarly derived (i.e. all F_2 populations) by assigning missing values to individuals in populations where the corresponding locus is non-segregating (Beavis and Grant 1991). As the seven populations mapped were segregating F_2 populations, we thus created a supra-population from the seven F_2 unit populations from which we computed pooled segregation scores. The data set for the composite map simply consisted in joining (for each locus) the unit data sets. Comparisons of marker order beetween the composite map and the unit maps was made by comparison of the corresponding LOD scores in the different populations (based on the set of informative markers). When a locus order in a population was found to be different from that of the composite map, and if the LOD score difference between the two orders was not greater than 3.0, the composite map order was taken as the unit map order. The UNIX or DOS version of MapMaker/EXP 3.0 (Lander et al. 1987) was employed to construct the unit maps as well as the composite map. The Mapmaker/Exp package was used for two reasons. Firstly, the algorithm of MapMaker/Exp appears to give the best solution of the gene ordering and, secondly, this software is distributed free of charge. Genetics maps were drawn with the use of the Drawmap package (van Ooijen 1994).

Results and discussion

Nomenclature of markers and linkage groups

In the present study, much attention was paid to ascertaining whether markers with different names in different populations represented the same locus. Alternatively, markers with the same name might represent different loci. This situation is complicated by the fact that, in sunflower, numerous duplicated locations of RFLP probes have been described (Berry et al. 1995; Gentzbittel et al. 1995; Jan et al. 1998). The nomenclature of the loci followed that of Gentzbittel et al. (1995): for example, S080H3_1 represents the locus indexed 1 for the cDNA probe SUN080 when the DNA was hydrolyzed with *Hin*dIII. In an attempt to test for the existence of a unique genome localisation for a probe, much attention was paid to promote hybridization of the same probe by using different restriction endonucleases on the same individuals. When a locus was identified at the same chormosomal location when digested with several restriction enzymes, it follows a more complete nomenclature: for example S012E1H3 represents the unique locus detected by cDNA probe SUN012 when the DNA was hydrolyzed either by *Eco*RI or *Hin*dIII. When a locus revealed by a multiple-copy probe was identified as being the same when the DNA was digested with different enzymes, its nomenclature is more complex: for example, S124E1_1H3_2 represents the indexed locus 1 with restriction endonuclease *Eco*RI and the indexed locus 2 with restriction endonuclease *Hin*dIII, when probing with the cDNA probe SUN124.

It is sometimes difficult to precisely relate the different loci revealed by a probe within the different mapping populations. This problem is not trivial when working with seven different populations. As a rule of thumb, when two loci identified in different segregating populations and with the same core name were mapped within a 5-cM distance, they were considered to represent only one locus and the name was adjusted accordingly. This strategy for reducing the number of non-informative loci was also described for barley composite mapping (Qi et al. 1996). The linkage group designation was as in Gentzbittel et al. (1995).

Unit maps

Seven individual maps were generated by running Mapmaker/Exp 3.0. The characteristics of these maps are reported in Table 2. When compared with the original maps of the C1, C2 and C5 crosses (Gentzbittel et al. 1995) the present maps were shorter, mainly

Table 2 Characteristics of the seven unit sunflower maps generated: number of detected loci, number of linkage groups, length of the map

Segregating population	Number of loci detected	Number of linkage groups	Map length (in cM)
S1	86	17	774
S2	106	18	1056
S3	111	16	1068
S4	95	18	1008
C1[a]	76	14	176
C2[a]	99	18	582
C5[a]	144	21	763
Core map	124	17	1106
Composite map	273	21	1573

[a] Data modified from Gentzbittel et al. (1995)

because of the use of the Kosambi function instead of the Haldane function, as previously described. A few loci were added to these populations, consisting of characterized sequences. The maps for the C1, C2 and C5 crosses were significantly smaller (mean length: 507 cM) than those computed for the S1–S4 populations (mean length: 976 cM). This may result from a higher degree of polymorphism between the parental inbreds of the second series of crosses, which allowed new regions of the genome to be revealed. Secondly, there were large differences in the number of individuals genotyped, ranging for example from 80 F_2 plants for the C1 population to 280 F_2 individuals for the S3 cross. None of the different unit maps are saturated. The number of linkage groups detected varies from 14 in the C1 population to 21 in the C5 cross. It is noteworthy that the quality of the unit maps largely depends on the number of data points generated. Between populations, the locus order of the linkage groups was almost the same, as no significant differences in the LOD score orders were recorded.

The S1 population confirmed the mapping of the *b1* locus, responsible for the apical branching, to linkage group 7. Basal branching (*bbr*), an as yet unlocalised phenotype, was assigned to the lower extremity of linkage group 7 in the S4 population, linked to the *b1* locus (32 cM). The fertility restoration locus *Msc1* was assigned to linkage group 12, in the S3 and S4 populations.

As previously reported and highlighted (Berry et al. 1995; Gentzbittel et al. 1995; Jan et al. 1998), about 35% of the probes show dominant segregation patterns and about 30% of the probes detected duplicated loci throughout the genome. The proportion of such loci is nearly identical in the three independent studies, apparently not related to the source of probes and the genetic background, suggesting a structural property of the sunflower genome (Sossey-Alaoui et al. 1998).

Composite mapping and morphological traits

Composite mapping is a pre-requisite in exploiting all the available information generated through different mapping programmes and much attention was paid to the different steps of map construction. We built the composite map in three steps.

(1) In order to minimize computation time, and to ensure a valid estimation of recombination fractions, a subset of informative loci was defined. It consisted of mostly co-dominant loci, that were mapped in at least three of the seven populations. For each linkage group, the best order was computed with the 'compare' command, then checked by the 'ripple' command. Based on 124 loci, a 'core map' was thus obtained, whose characteristics are given Table 2.

(2) The remaining loci were added to this backbone by the 'build' command.

(3) A 'ripple' command was then performed on each linkage group, in order to check if the proposed order was not a local minimum of the ordering function.

The composite map of the sunflower genome is proposed in Fig. 1. This map exhibits 17 major linkage groups and four small frements, for a total length of 1573 cM. It comprises 273 loci, including four morphological traits and 18 loci corresponding to previously described genes. Four small segments with 2–5 marker loci are described. The description of new probes and loci results in the joining of different linkage groups previously found as independent: the present linkage group 12 is elongated by the addition of previously described linkage group 17; previous linkage group 24 was fused with present linkage group 16.; lastly, the present linkage group 17 was obtained by the joining of previous linkage groups 18, 20 and 21 (Gentzbittel et al. 1995).

The mapping density is about 7 cM, but reflects very different situations according to the linkage group (Table 3). For example linkage group 9 contains 25 loci for 103 cM, in comparison with linkage group 12 with 13 loci for 118 cM total. In addition, the distribution of loci along the linkage groups does not appear to be homogeneous, with some regions exhibiting a high density of loci. Examples are given by linkage group 7 or linkage group 3. A clustering of markers around the centromeric regions has been described (Chao et al. 1989; Tanksley et al. 1992; Qi et al. 1996), probably in relation to the centromeric suppression of recombination (Tanksley et al. 1992). However, for sunflower the centromeric positions are unknown, so that it is not possible to go further in the analyses. Only a recent description of four acrocentric chromosomes has been

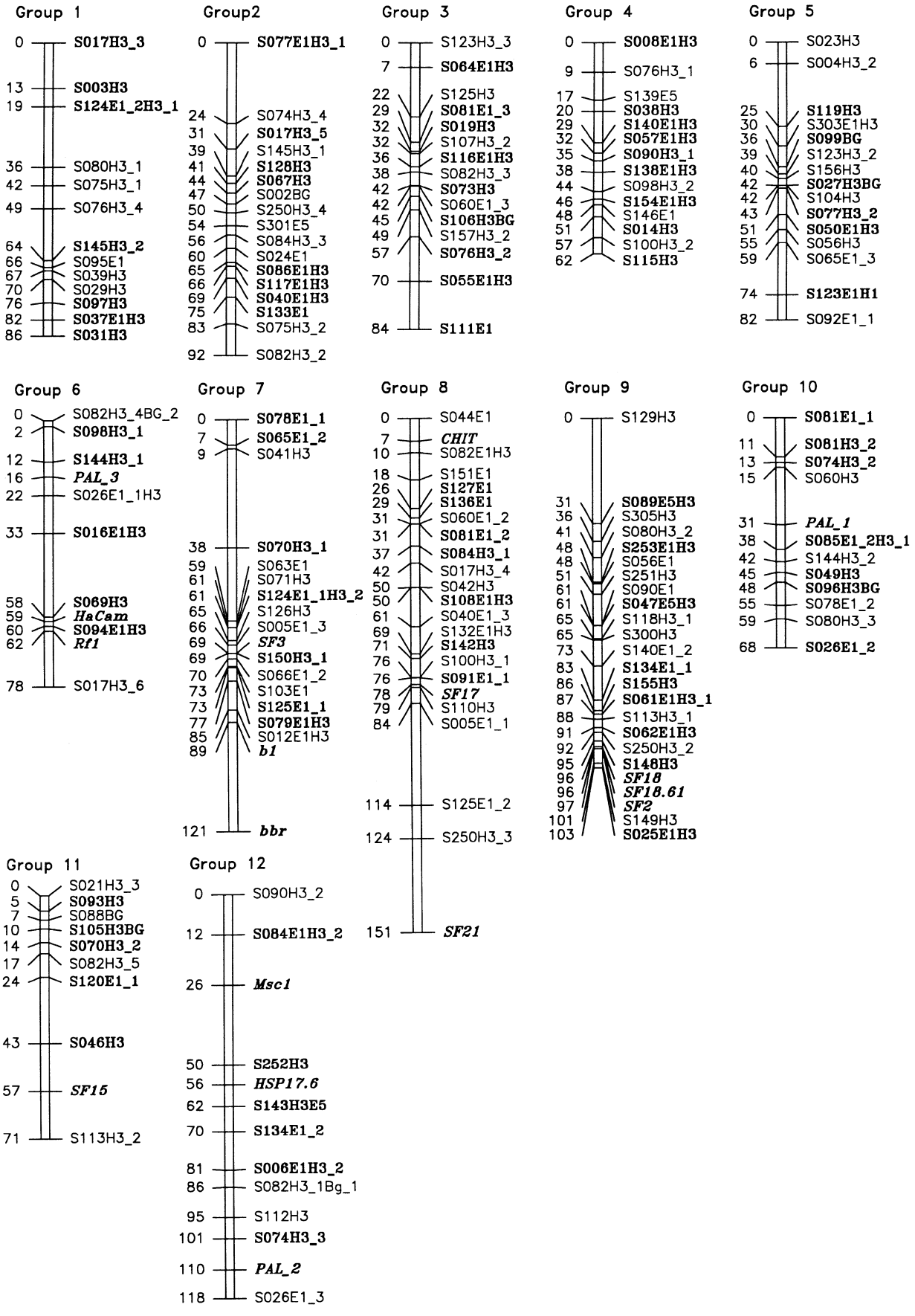

Fig. 1 Continued

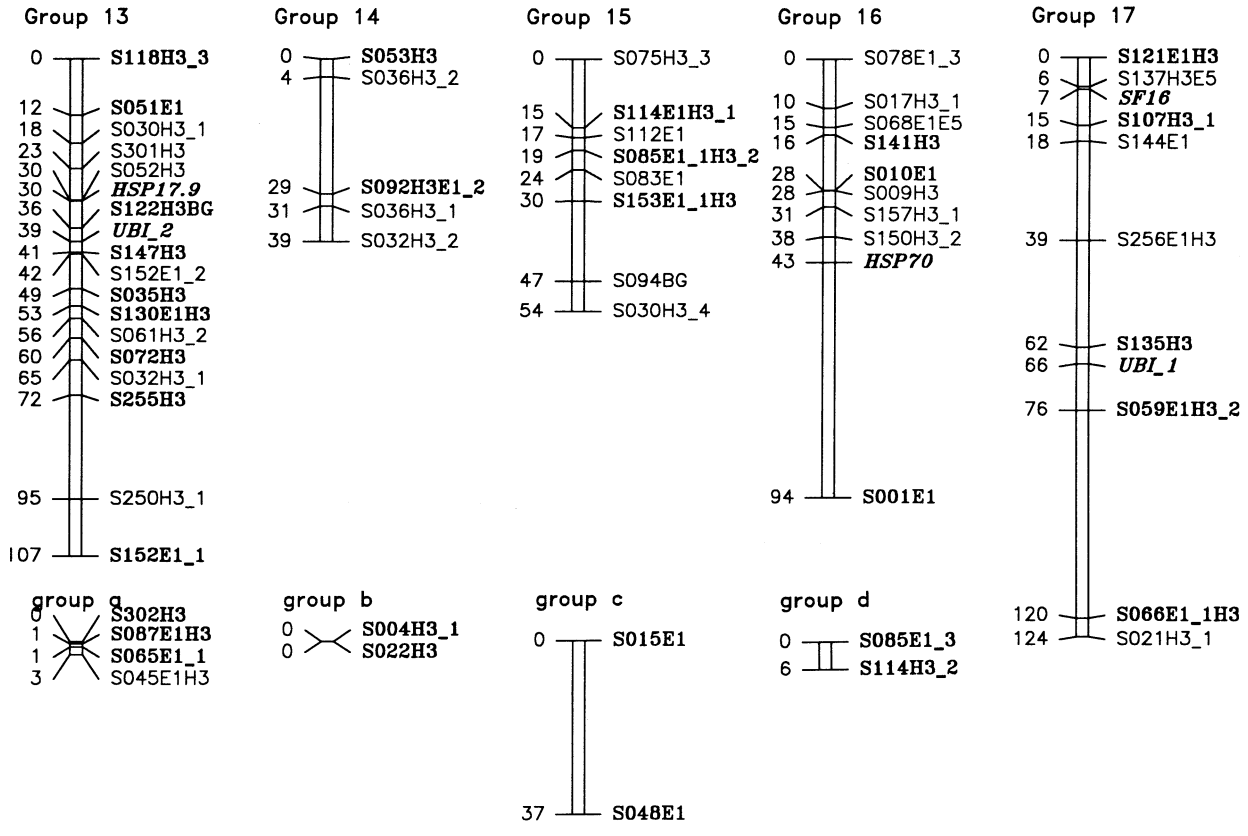

Fig. 1 Composite molecular linkage map of the sunflower genome. Linkage groups are arbitrarily oriented and cumulative distances are expressed in Kosambi cM. Markers *in bold* are backbone loci. Characterized sequences are *italicized*

Table 3 Sunflower RFLP linkage map: linkage map size, number of markers and average marker interval per major linkage group

Linkage group	Size (cM)	Number of markers	Average marker interval (cM)
1	86	14	6.1
2	92	17	5.4
3	84	16	5.2
4	62	14	4.1
5	82	15	5.5
6	78	11	7.1
7	121	18	6.7
8	151	23	6.0
9	103	24	4.1
10	68	12	5.2
11	71	10	7.1
12	118	13	9.1
13	107	18	5.9
14	39	5	7.8
15	54	8	6.7
16	94	9	9.4
17	124	11	11.3
Totals and mean	1534	238	6.7

published for cultivated sunflower (Schrader et al. 1997). The presence of a large number of probes hybridizing different loci on the sunflower genome is now well-documented (Berry et al. 1995; Gentzbittel et al. 1995; Jan et al. 1998) and the question of an ancient polyploid origin for sunflower is still under study (Sossey-Alaoui et al. 1998). The proposed model, based upon the amphipolyploid origin of large diploid genomes, is similar in nature to the results obtained in onion (King et al. 1998). From a practical point of view, the interest for probes revealing a few loci in a single hybridization is counter-balanced by the difficulties generated through the composite mapping process as it would be difficult to clearly identify the different loci segregating in different mapping populations.

We report the presence of two distinct fertility restoration loci of the PET1 cytoplasmic male sterility (Leclercq 1969): *Rf1* on linkage group 6 and *Msc1* on linkage group 12 . We have no evidence as to whether these two loci are duplications of a single gene, but their genetic independence would make it possible to combine the two loci in an inbred without difficulty. Three probes closely linked to the *Rf1* gene are reported, which could promote a chromosome-landing approach. However, as this gene could have been introgressed in sunflower from a related species (Kinman 1970; Fick and Zimmer 1974) one cannot exclude the possibility of supression of recombination, due to the presence of foreign DNA. The tight linkage of *HaCam*, S094 and S069 could thus be an artifact due to local mis-pairing during meiosis. For branching, the situation is quite different: apical branching has long been

reported (Putt 1964) but basal branching (*bbr*) has usually been neglected, more particularly because only few cultivated inbreds exhibit this phenotype. We demonstrate that apical and basal branching are quite different loci, linked on linkage group 7. Fertility restoration (either by *Rf1* or *Msc1*) is almost always associated with apical branching in the R pool of cultivated inbred genotypes (Berry et al. 1994; Gentzbittel et al. 1994). This situation demonstrates the effectivness of breeding in keeping these two phenotypic traits associated, for practical purposes of pollen delivery, although being independent. The presence of a disease resistance cluster on linkage group 1 was previously reported (Vear et al. 1997; Gentzbittel et al. 1998) and is beyond the scope of this paper.

Mapping of characterized genes

Defense-related genes

A PCR product of about 450 bp (GenBank accession U96640) and sharing high levels of homology with previously described chitinases was cloned and mapped in the two crosses S1 and S3. Autoradiographs (data not shown) revealed a very simple pattern, suggesting a unique sequence. The chitinase locus co-segregated with loci assigned to the lower distal region of linkage group 8. Members of the HSP 70 family are often involved in stress reactions. A PCR product of about 900-bp long (GenBank accession U96641) was obtained, exhibiting homologies with previously described HSP70 genes. It was mapped for the C5, S2 and S3 crosses. This probably unique sequence (data not shown) showed co-segregation with loci assigned to linkage group 16. An internal fragment of 700-bp of the PAL cDNA of sunflower (Mazeyrat et al. 1998) was mapped on the S2 and S4 populations. At least three loci were mapped, on linkage groups 6, 10 and 12.

Mapping of stress-related probes

Homologous probes for ubiquitine and HSP17.9 were mapped on the C1, C5, S2, S3, S4 and C5, S1 and S4 crosses respectively, while HSP17.6 was mapped on the C1 population. All three probes revealed complex hybridization patterns, suggesting multigenic families. A locus corresponding to a polymorphic band of the HSP17.9 probe was localised to linkage groups 13, while two distinct loci of the ubiquitine probe were assigned to linkage groups 13 and 17. We were able to localise a locus for HSP17.6 linked to the *Msc1* locus on linkage group 12. The calmodulin probe, *HaCam* (Courbou et al. 1997), appears to be tightly linked to the *Rf1* locus, on linkage group 6.

Homology cloning was found to be an efficient way of obtaining significant parts of well-known and characterized genes, allowing fruitful genetic analysis of the sequences studied. This approach, although based on previous knowledge for the identification of homologous regions, allows a rapid and efficient targeting of sequences with a biological meaning. It was used recently for targeting genes for disease resistance in tomato (Ohmori et al. 1998), bean (Geffroy et al. 1998) and sunflower (Gentzbittel et al. 1998). However, a large proportion of the mapped sequences appear to consist of multiple-copy probes, sometimes exhibiting dominant segregation patterns. This fact limits expression-analyses of the genes, for example through Northern experiments, because of the existence of putative pseudogenes or duplicated expressed genes, making it impossible to distinguish between active and unexpressed genes until locus-specific probes or primers are developed.

Mapping of SF probes

Over the eight mapped probes, a small cluster appears on linkage group 9: SF18, SF18.61 and SF2. Other localisations are on linkage group 7 (SF3), linkage group 8 (SF21 and SF17), linkage group 15 (SF15) and linkage group 17 (SF16). Four different cDNAs homologous to SF15 were described (Dudareva et al. 1996). However, we did not find any evidence for several genomic locations of this probe, suggesting that the different expressed genes are located at the same genetic locus. SF2 and SF18 have been described to be distinct sequences, exclusively hybridizing with RNA from anthers in the late developmental stage, and highly expressed in male-fertile plants (Domon et al. 1990). Some features suggested that they could be functionally related. Interestingly, these two probes are closely linked on linkage group 9. This is in accordance with results depicting the frequent linkage of genes involved in similar biological functions.

Reliability of the composite map

Markers that are unique to a particular population can, of course, only be positioned on the basis of information from that single population. Therefore, the ordering of unique markers on the composite maps is less accurate than the ordering of anchor loci, in particular in the regions where overlapping of the linkage groups is weak. An explicit assumption was made that one probe will hybridize with the same locus in different mapping populations. Another assumption was that gene order is conserved throughout the different genotypes employed. Comparative genome mapping studies in sunflowers are not available, but numerous studies have shown an obvious conservation of gene structure, in particular in grasses (Guimaraes et al. 1997; Wang et al. 1992), the *Solanaceae* (Bonierbale et al. 1988) and *Brassica* (Bohuon et al. 1996). We thus hypothesize a conservation in the order of loci between

the nine different inbred lines used as parents for the mapping populations. Different methods exist for comparing map orders obtained from different populations (Beavis and Grant 1991; Causse et al. 1996). In the present case, the variances of the estimates of recombination fractions are quite different, as the effective numbers of individuals genotyped were very different between populations (Table 1). LOD-score comparison

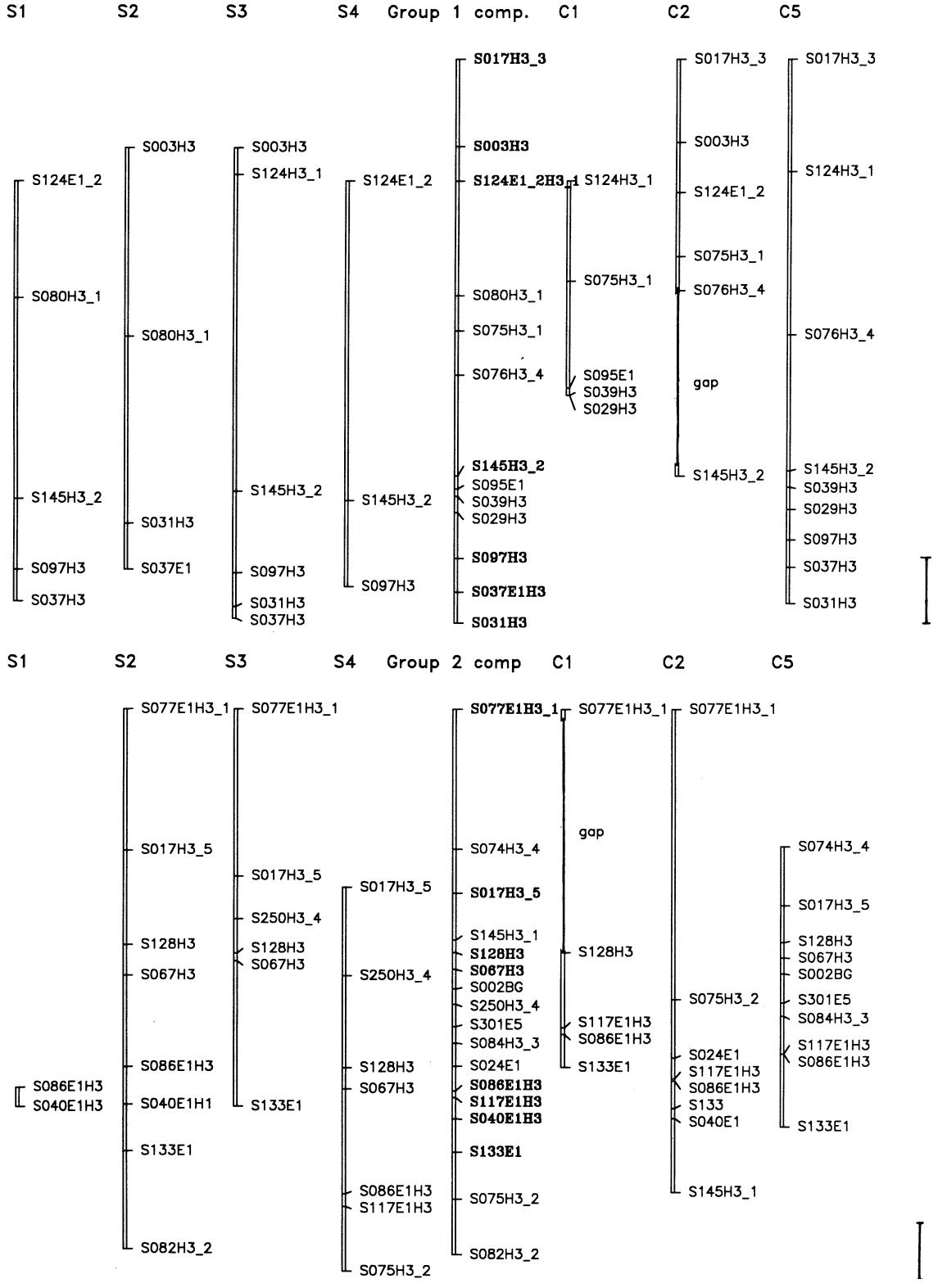

Fig. 2 See page 233 for legend

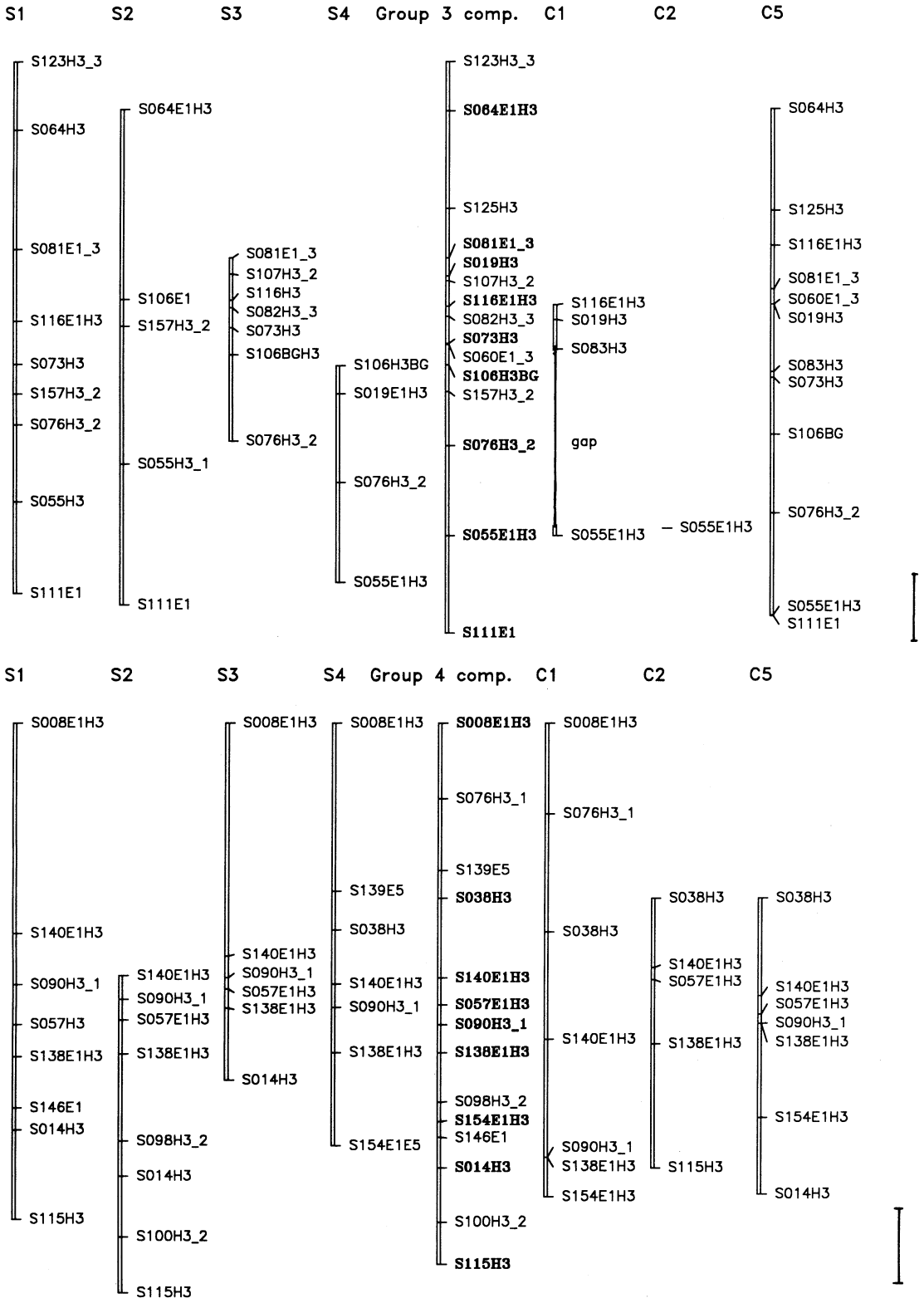

Fig. 2 See page 233 for legend

methods, based on the assumption of homogeneous variances of estimates, could not be used. We thus consider the observed heterogeneity as a consequence of heterogeneity in the size of the populations mapped.

The reliability of integrated maps largely depends on the number of common loci shared by the individual maps. The composite map presented in Fig. 1 was created on the basis of 124 markers common to at least

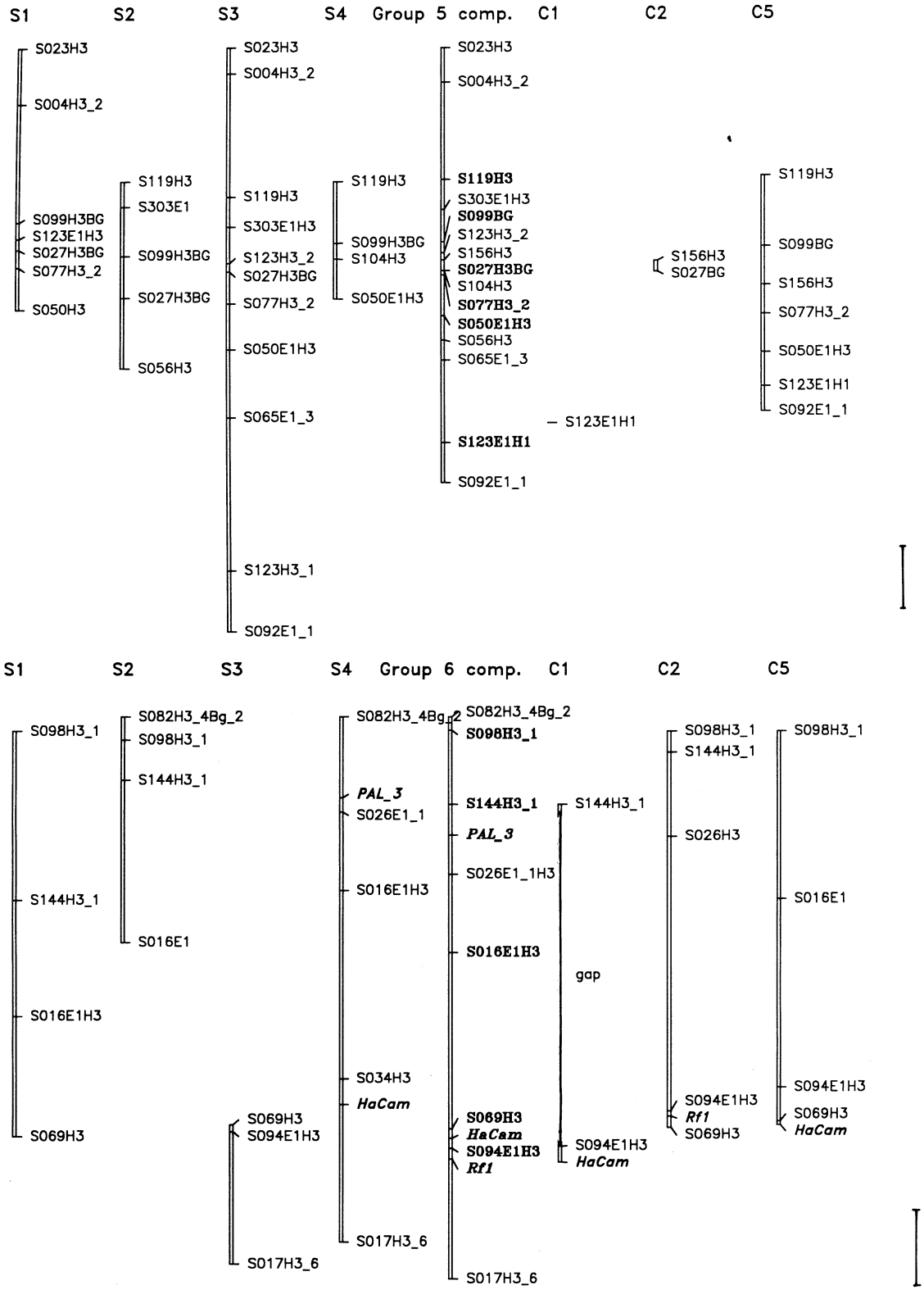

Fig. 2 See page 233 for legend

three of the unit-mapping populations. In Fig. 2 we present the backbone structure of the 17 linkage groups. Anchor loci were chosen on the basis of their co-dominant genetics as well as their 'strategic placement' in regions of weak genome coverage. As previously described (Gentzbittel et al. 1995), genome coverage is very different in the different crosses studied. This fact highlights the need for molecular

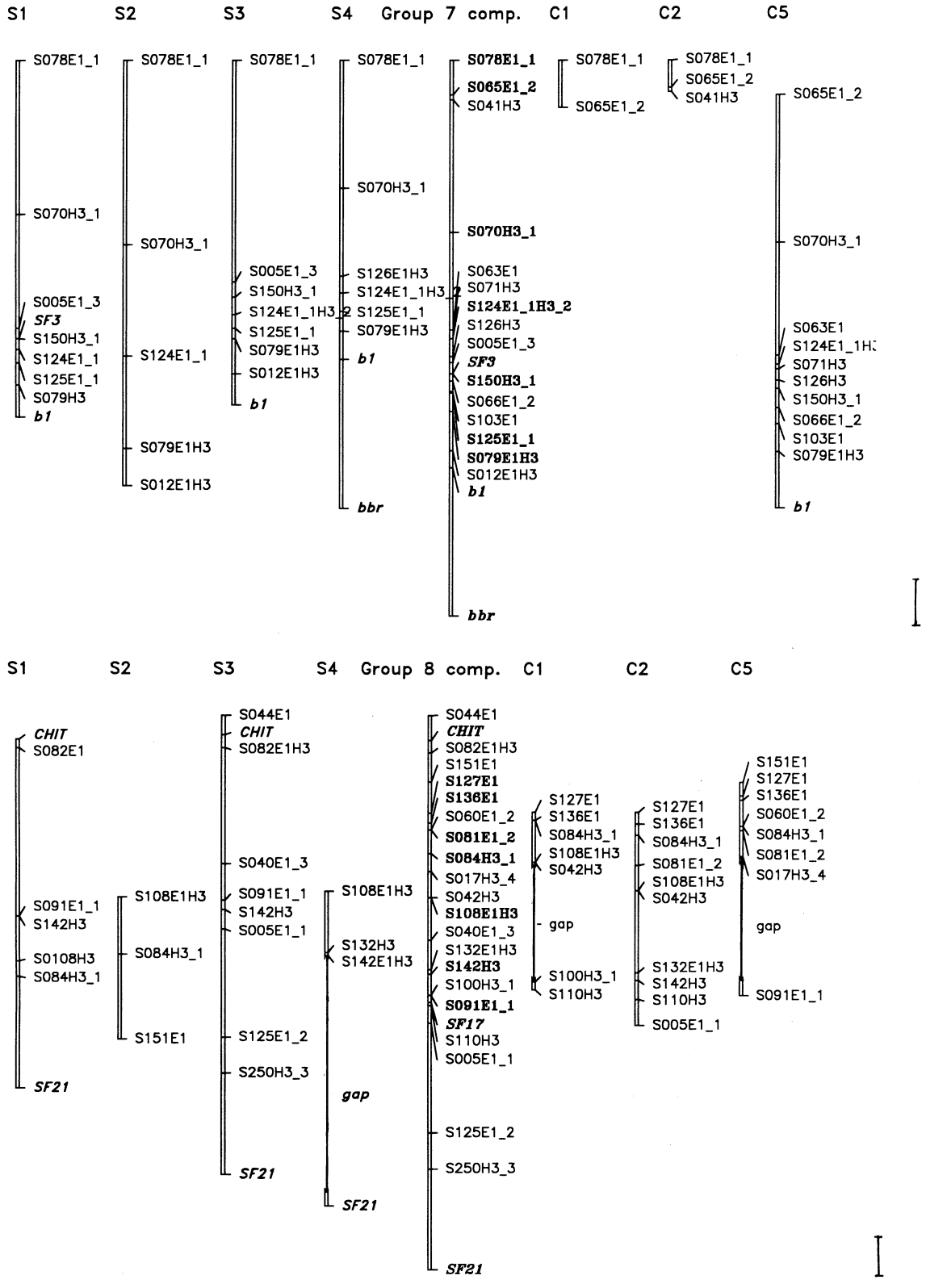

Fig. 2 See page 233 for legend

markers having a high polymorphic information content (PIC). As an example in the present case, hypothesized RFLP mapping of only a few of the seven populations leads to a map with less interest, as large regions of the genome would not be revealed. For sunflower, it appears that molecular markers such as AFLPs (Hongtrakul et al. 1997) or SSRs could have greater interest for mapping as they present better PIC

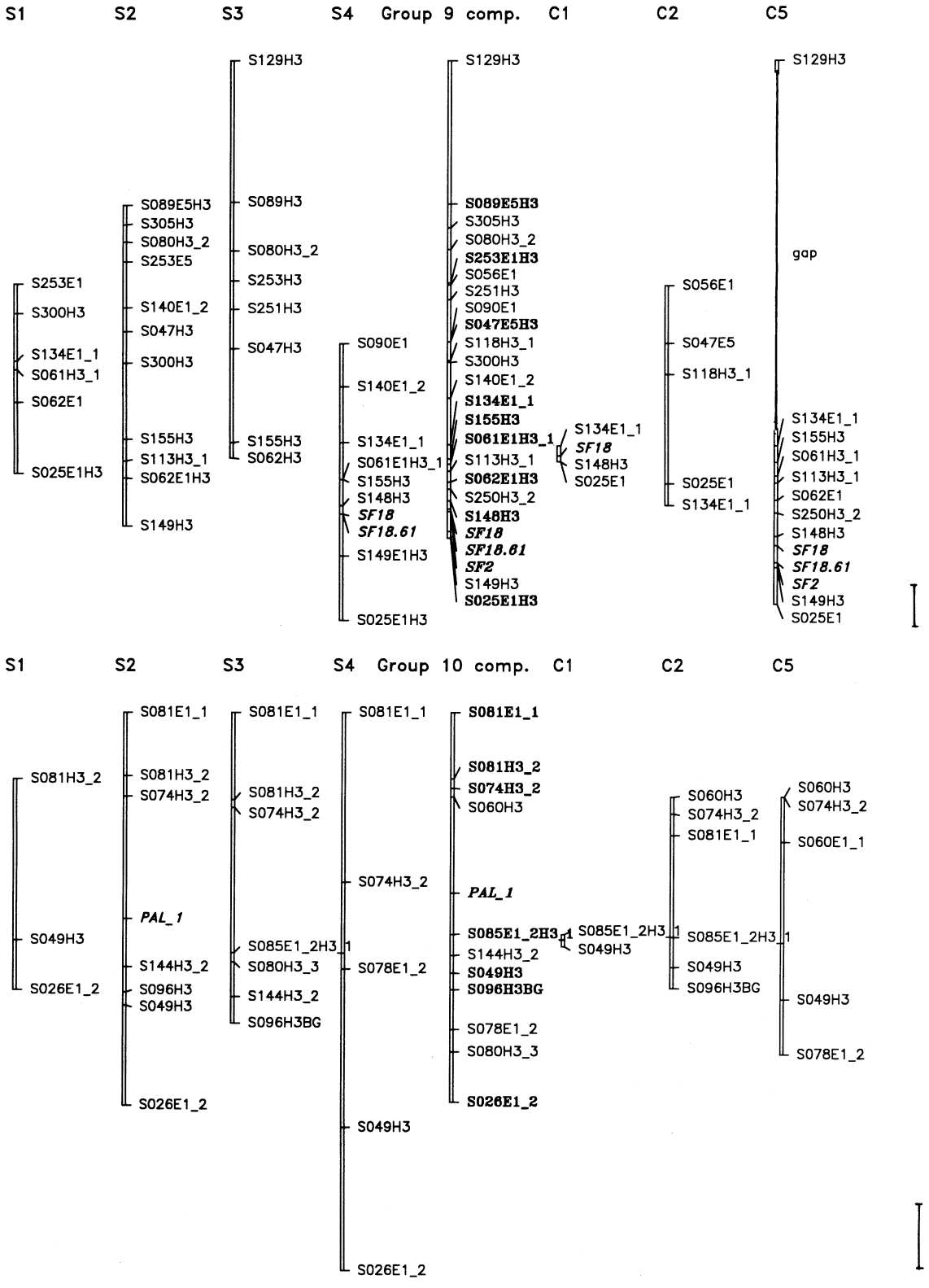

Fig. 2 See page 233 for legend

values. By definition, the integration of maps originating form different populations is only feasible if common markers are available. The efficiency of RFLPs in such a situation should be emphasized. These molecular markers, despite their relatively poor polymorphism and the complicated technnique involved in their quasi-routine utilization, are the sole tool, with SSRs, allowing such composite maps to be made. A recent

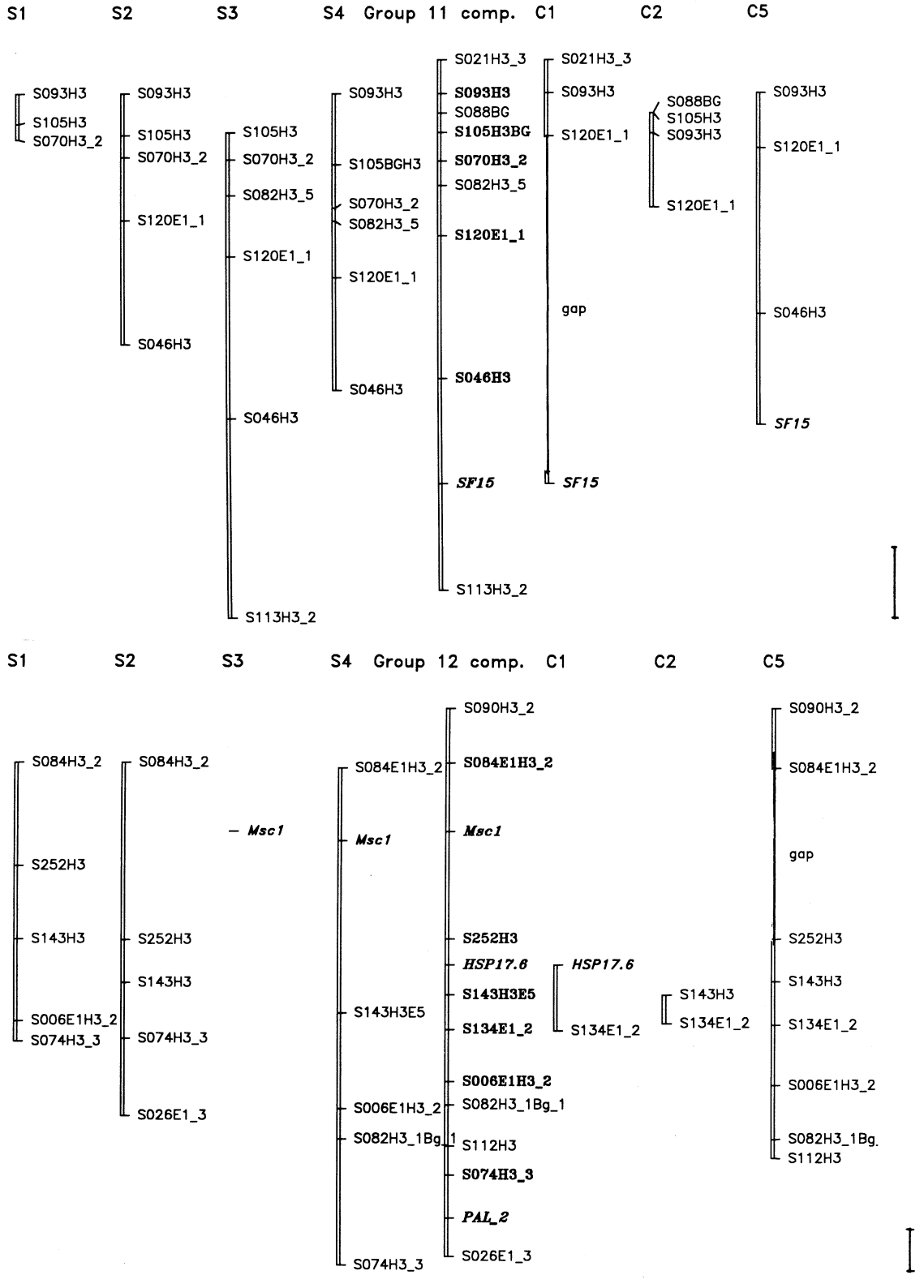

Fig. 2 See page 233 for legend

paper (Van der Voort et al. 1997) describes the integration of distinct linkage maps using AFLPs.

Estimating the sunflower genome to be 18 Morgans long, we computed, with a circular chromosome approximation (Lange and Boehnke 1982), that at least 206 loci are necessary to have a 90% probability that a given gene will be located at no more than 10 cM from a probe. If we consider an average number of 100

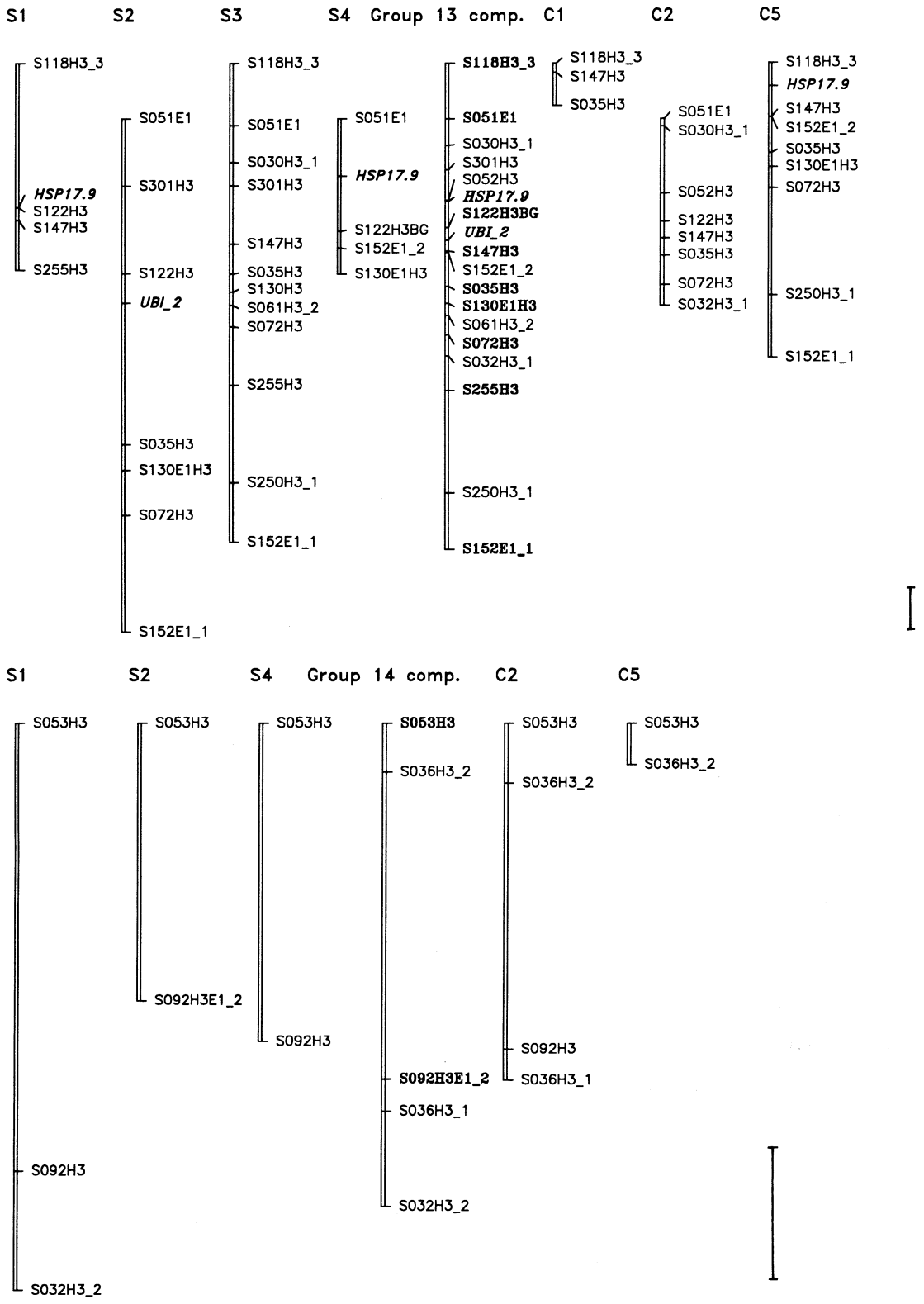

Fig. 2 See page 233 for legend

loci identified in each cross, the mean distance between a gene and a random locus (Martin et al. 1991) is estimated to be 9 cM, with a 95% upper-limit of 27 cM. In fact, we mapped downy mildew resistance genes in such a confidence interval (Mouzeyar et al. 1995; Roeckel-Drevet et al. 1996).

From the composite map, a subset of cDNA markers, as evenly distributed as possible over

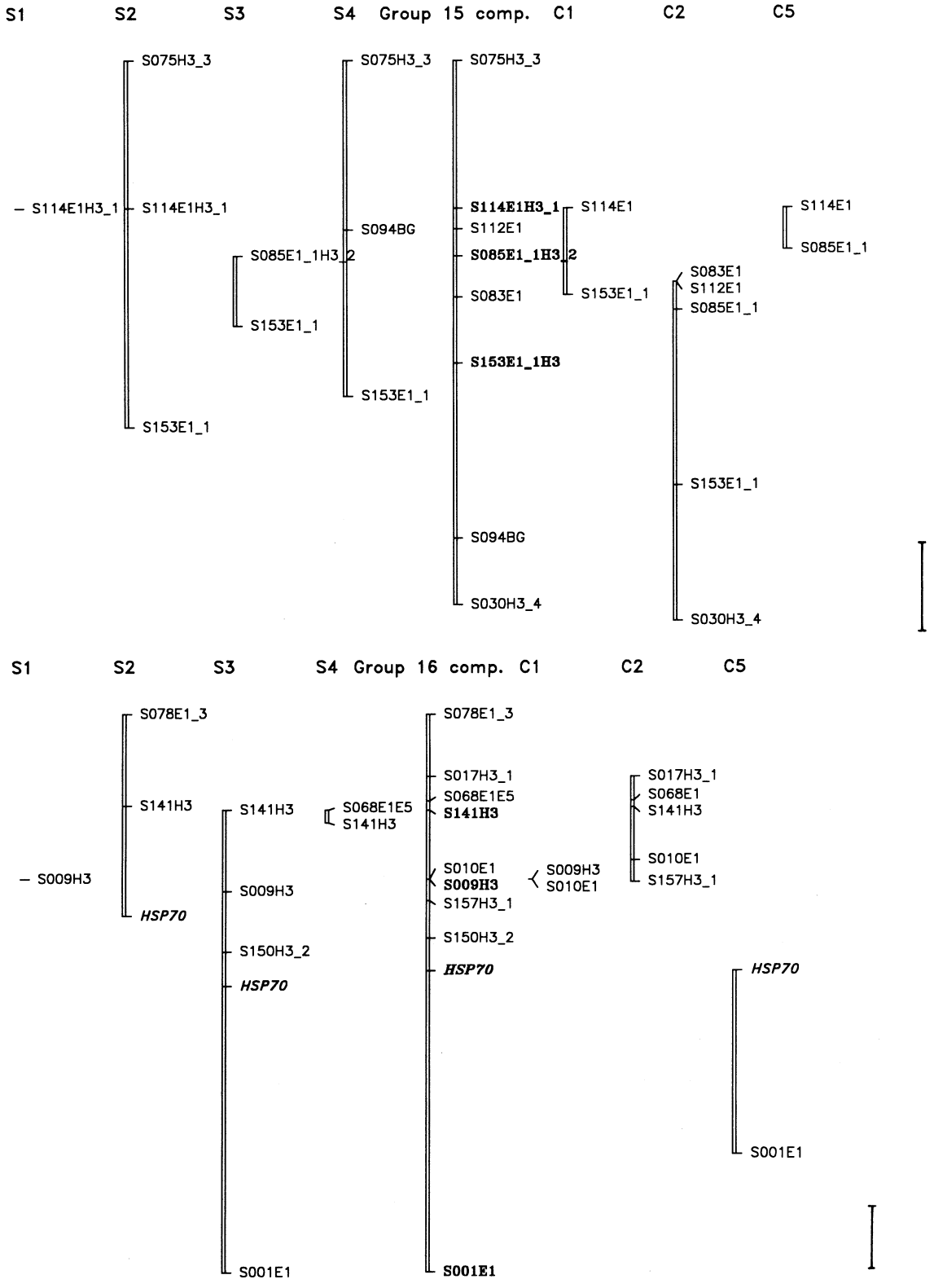

Fig. 2 See page 233 for legend

the genome, has been selected as anchor loci, which will be used for linkage-group landmarks and the adding of AFLP and SSR markers to this core map. The definition of a reliable set of 'bridge markers' (Qi et al. 1996) is of upmost importance for maximizing the probability of detecting all the QTLs affecting a character in a particular mapping population.

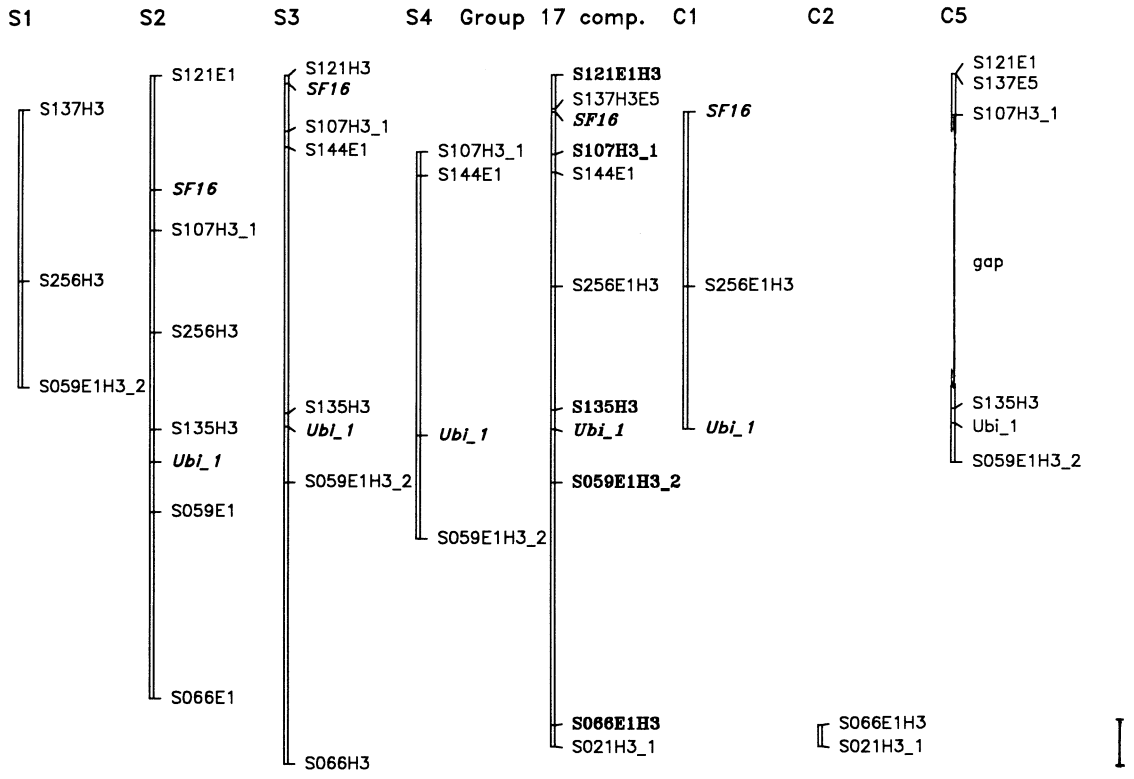

Fig. 2 Backbone structure of the composite map of sunflower. Only anchor loci for individual maps are indicated. The "*gap*" in some individual maps indicates that linkage is not detected in that particular population. Distances are expressed in Kosambi cM

Acknowledgements We thank Annie Perrault and Véronique de Conto for generating the molecular data. Jacqueline Philippon provided excellent field data. Prof. A. Steinmetz and Dr. Krauter are also acknowledged for their gracious gift of the SF probes. This work was financed in part by the GIE CARTISOL (a consortium of 11 seed companies including Advanta, Agrosem, Cargill, Caussade Semences, KWS, Limagrain, Maïsadour, Pioneer, RAGT, Rustica-Prograin-Génétique, Verneuil). It was also supported in part by grants from the CETIOM, the Ministère de l'Agriculture et de la Pêche (contrat de branche 1994–1996), the Ministère de l'Education Nationale, de l'Enseignement Supérieur et de la Recherche (Actions Concertées Coordonnées – Sciences du Vivant) and a State-Region contract "Semences et Plants".

References

Azanza F, Tadmor Y, Klein BP, Rocheford TR, Juvik JA (1996) Quantitative trait loci influencing chemical and sensory characteristics of eating quality in sweet corn. Genome 39:40–50

Baltz R, Evrard J, Bourdon V, Steinmetz A (1996) The pollen-specific LIM protein PLIM-1 from sunflower binds nucleic acids in vitro. Sex Plant Reprod 9:264–268

Beavis WD, Grant D (1991) A linkage map based on information form four F_2 populations of maize (*Zea mays* L.). Theor Appl Genet 82:636–644

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 RJ, Barnes SR, Caligari PDS (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 ST, Leon AJ, Hanfrey CC, Challis P, Burkholz A, Barnes SJ, 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

Binet M-N, Steinmetz A, Tessier L-H (1989) The primary structure of sunflower (*Helianthus annuus*) ubiquitin. Nucleic Acids Res 17:2119

Bohuon E, Keith D, Parkin I, Sharpe A, Lydiate D (1996) Alignment of the conserved C genomes of *Brassica oleracea* and *Brassica napus*. Theor Appl Genet 93:5–6

Bonierbale MW, Plaisted RL, Tanksley SD (1988) RFLP maps based on a common set of clones reveal modes of chromosomal evolution in potato and evolution. Genetics 120:1095–1103

Causse M, Santoni S, Damerval C, Maurice A, Charcosset A, Deatrick J, de Vienne D (1996) A composite map of expressed sequences in maize. Genome 39:418–432

Chao S, Sharp PJ, Worland A, Warham EJ, Koebner RMD, Gale MD (1989) RFLP-based genetic maps of wheat homeologous group-7 chromosomes. Theor Appl Genet 78:495–504

Coca MA, Almoguera C, Jordano J (1994) Expression of sunflower low-molecular-weight heat-shock proteins during embryogenesis and persistence after germination: localisation and possible functional implications. Plant Mol Biol 25:479–492

Courbou I, Badaoui S, Mouzeyar S, Gentzbittel L, Nicolas P (1997) RT-PCR cloning of a sunflower calmodulin (accession U79736) gene. (PGR97-072). Plant Physiol 114:395

Domon C, Evrard JL, Herdenberger F, Pillay DTN, Steinmetz A (1990) Nucleotide sequence of two anther-specific cDNAs from sunflower (*Helianthus annuus* L). Plant Mol Biol 15:643–646

Dudareva N, Reddy JT, Kraüter R, Bronner R, Steinmetz A, Pillay DTN (1996) A flower-specific gene family whose expression is

regulated temporally and spatially during flower development in sunflower. Plant Sci 120:161–173
Fick GN, Zimmer DE (1974) Fertility restoration in confectionery sunflowers. Crop Sci 14:603–604
Geffroy V, Creusot F, Falquet J, Sévignac M, Adam-Blondon A-F, Bannerot H, Gepts P, Dron M (1998) A family of LRR sequences in the vicinity of the *Co-2* locus for anthracnose resistance in *Phaseolus vulgaris* and its potential use in marker-assisted selection. Theor Appl Genet 96:494–502
Gentzbittel L, Zhang X-Y, 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 Y-X, Bervillé A, Nicolas P (1995) Development of a consensus linkage RFLP map of cultivated sunflower (*Helianthus annuus* L.). Theor Appl Genet 90:1079–1086
Gentzbittel L, Mouzeyar S, Badaoui S, Mestries E, Vear F, Tourvieille de Labrouhe D, Nicolas P (1998) Cloning of molecular markers for disease resistance in sunflower, *Helianthus annuus* L. Theor Appl Genet 96:519–525
Guimaraes CT, Sills GR, Sobral BWS (1997) Comparative mapping of Andropogoneae: *Saccharum L.* (sugarcane) and its relation to sorghum and maize. Proc Natl Acad Sci USA 94:14261–14266
Harushima Y, Yano M, Shomura P, Sato M, Shimano T, Kuboki Y, Yamamoto T, Lin SY, Antonio BA, Parco A, Kajiya H, Huang N, Yamamoto K, Nagamura Y, Kurata N, Khush GS, Sasaki T (1998) A high-density rice genetic linkage map with 2275 markers using a single F-2 population. Genetics 148:479–494
Helentjaris T (1987) A genetic linkage map for maize based on RFLPs. Trends Genet 3:217–221
Hongtrakul V, Huestis GM, Knapp SJ (1997) Amplified fragment length polymorphisms as a tool for DNA fingerprinting sunflower germplasm: genetic diversity among oilseed inbred lines. Theor Appl Genet 95:400–407
Jan CC, Vick BA, Miller JF, Kahler AL, Butler ETI (1998) Construction of an RFLP linkage map for cultivated sunflower. Theor Appl Genet 96:15–22
Kasuga T, Salimath SS, Shi JR, Gijzen M, Buzzell RI, Bhattacharyya MK (1997) High-resolution genetic and physical mapping of molecular markers linked to the *Phytophthora* resistance gene *Rps1-k* in soybean. Mol Plant-Microbe Interact 10:1035–1044
Kesseli RV, Paran I, Michelmore RW (1994) Analysis of a detailed genetic linkage map of *Lactuca sativa* (Lettuce) constructed from RFLP and RAPD markers. Genetics 136:1435–1446
King JJ, Bradeen JM, Bark O, Mccallum JA, Havey MJ (1998) A low-density genetic map of onion reveals a role for tandem duplication in the evolution of an extremely large diploid genome. Theor Appl Genet 96:52–62
Kinman ML (1970) New develoments in the USDA and state experiment station sunflower breeding programs. Proc 4th Int Sunflower Conf, Memphis, Tennessee: pp 181–183
Krauter-Canham R, Bronner R, Evrard JL, Hahne G, Friedt W, Steinmetz A (1997) A transmitting tissue- and pollen-expressed protein from sunflower with sequence similarity to the human RTP protein. Plant Sci 129:191–202
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
Lange K, Boehnke M (1982) How many polymorphic genes will it take to span the human genome? Am J Hum Genet 34:842–845
Leclercq P (1969) La stérilité mâle cytoplasmique du tournesol. Ann Amélior Plant 19:99–106
Leclercq P (1971) La stérilité mâle cytoplasmique du tournesol. I. Premières études sur la restauration de la fertilité. Ann Amélior Plant 21:45–54
Li Z, Pinson S, Park W, Paterson A, Stansel J (1997) Epistasis for three grain-yield components in rice (*Oryza sativa* L). Genetics 145:453–465
Lorieux M, Petrov M, Huang N, Guiderdoni E, Ghesquiere A (1996) Aroma in rice: genetic analysis of a quantitative trait. Theor Appl Genet 93:1145–1151
Martin GB, Williams JGK, Tanksley SD (1991) Rapid identification of markers linked to a *Pseudomonas* resistance gene in tomato by using random primers and near-isogenic lines. Proc Natl Acad Sci USA 88:2336–2340
Mazeyrat F, Salles S, Drevet J, Roeckel-Drevet P, Tourvieille D, Ledoigt G (1998) Isolation of a complete PAL cDNA from sunflower (PGR98-108). Plant Physiol 117:719
Mestries E, Gentzbittel L, Tourvieille de Labrouhe D, Nicolas P, Vear F (1998) Analysis of quantitative trait loci associated with resistance to *Sclerotinia sclerotiorum* in sunflowers (*Helianthus annuus* L.) using molecular markers. Mol Breed 4:215–226
Mohan M, Sathyanarayanan PV, Kumar A, Srivastava MN, Nair S (1997) Molecular mapping of a resistance-specific PCR-based marker linked to a gall midge resistance gene (*Gm4t*) in rice. Theor Appl Genet 95:777–782
Mouzeyar S, Roeckel-Drevet P, Gentzbittel L, Philippon J, Tourvieille de Labrouhe D, Vear F, Nicolas P (1995) RFLP and RAPD mapping of the sunflower *Pl1* locus for resistance to *Plasmopara halstedii* race 1. Theor Appl Genet 91:733–737
Ooijen JW van (1994) DrawMap: a computer program for drawing genetic linkage maps. J Hered 85:66
Putt ED (1964) Recessive branching in sunflowers. Crop Sci 4:444–445
Qi X, Stam P, Lindhout P (1996) Comparison and integration of four barley genetic maps. Genome 39:379–394
Roeckel-Drevet P, Gagne G, Mouzeyar S, Gentzbittel L, Philippon J, Nicolas P, Tourvieille de Labrouhe D, Vear F (1996) Colocation of downy mildew (*Plasmopara halstedii*) resistance genes in sunflower (*Helianthus annuus* L.). Euphytica 91:225–228
Schrader O, Ahne R, Fuchs J, Schubert I (1997) Karyotype analysis of *Helianthus annuus* using Giemsa banding and fluorescence in situ hybridization. Chrom Res 5:451–456
Sharon D, Cregan PB, Mhameed S, Kusharska M, Hillel J, Lahav E, Lavi U (1997) An integrated genetic linkage map of avocado. Theor Appl Genet 95:911–921
Sossey-Alaoui K, Serieys H, Tersac M, Lambert P, Schilling E, Griveau Y, Kaan F, Bervillé A (1998) Evidence for several genomes in *Helianthus*. Theor Appl Genet 97:422–430
Tanksley SD, Ganal MW, Prince JP, de Vincente MC, Bonierbale MW, Broun P, Fulton TM, Giovannonni JJ, Grandillo S, Martin GB, Messeguer R, Miller JC, Miller L, Paterson AH, Pineda O, Roder MS, Wing RA, Wu W, Young ND (1992) High-density molecular linkage maps of the tomato and potato genomes. Genetics 132:1141–1160
Van der Voort JNAMR, Vanzandvoort P, Vaneck HJ, Folkertsma RT, Hutten RCB, Draaistra J, Gommers FJ, Jacobsen E, Helder J, Bakker J (1997) Use of allele specificity of comigrating AFLP markers to align genetic maps from different potato genotypes. Mol Gen Genet 255:438–447
Vear F, Gentzbittel L, Philippon J, Mouzeyar S, Mestries E, Roeckel-Drevet P, Tourvieille de Labrouhe D, Nicolas P (1997) The genetics of resistance to five races of downy mildew (*Plasmopara halstedii*) in sunflower (*Helianthus annuus* L.). Theor Appl Genet 95:584–589
Wang ML, Atkinson MD, Chinoy CN, Devos KM, Gale MD (1992) Comparative RFLP-based genetic maps of barley chromosome 5 (1H) and rye chromosome 1R. Theor Appl Genet 84:339–344
William HM, Hoisington D, Singh RP, Gonzalez de Leon D (1997) Detection of quantitative trait loci associated with leaf rust resistance in bread wheat. Genome 40:253–260
Witsenboer H, Vogel J, Michelmore RW (1997) Identification, genetic localization, and allelic diversity of selectively amplified microsatellite polymorphic loci in lettuce and wild relatives (*Lactuca* spp.). Genome 40:923–936