Theor Appl Genet (2001) 103:84–97

J.-Y. Hoarau · B. Offmann · A. D'Hont
A.-M. Risterucci · D. Roques · J.-Ch. Glaszmann
L. Grivet

Genetic dissection of a modern sugarcane cultivar (*Saccharum* spp.). I. Genome mapping with AFLP markers

Received: 31 August 2000 / Accepted: 16 October 2000

Abstract Sugarcane cultivars are polyploid, aneuploid clones derived from interspecific hybridization between *Saccharum officinarum* and *S. spontaneum*. Their genome has recently started to be unravelled as a result of the development of molecular markers. We constructed an AFLP genetic map based on a selfing population of a specific cultivar, R570. Using 37 AFLP primer pairs, we detected 1,185 polymorphic markers of which 939 were simplex (segregated 3:1); these were used to construct the map. Of those 939, 887 were distributed on 120 cosegregation groups (CGs) based on linkages in coupling, while 52 remained unlinked. The cumulative length of all the groups was 5,849 cM, which is probably around one-third of the total genome length. Comparison with reference *S. officinarum* clones enabled us to assign 11 and 79 CGs to *S. spontaneum* and *S. officinarum*, respectively, whereas 11 CGs were probably derived from recombination between chromosomes of the two ancestral species. The patchy size of the groups, which ranges from 1 to 232 cM, illustrates the difficulty to access large portions of chromosomes, particularly those inherited from *S. officinarum*. Repulsion phase linkages suggested a high preferential pairing for 13 CG pairs. Out of the 120 CGs, 34 could be assigned to one of the 10 homo(eo)logy groups already defined in a previous RFLP map owing to the use of a small common marker set. The genome coverage was significantly increased in the map reported here. Implications for quantitative trait loci (QTL) research and marker-assisted breeding perspectives are discussed.

Communicated by H.F. Linskens

J.-Y. Hoarau (✉)
CERF, B.P. 315, 97494 Ste-Clotilde, Reunion
e-mail: cerf@guetali.fr
Fax: +31-0262-290507

B. Offmann · A. D'Hont · A.-M. Risterucci · D. Roques
J.-C. Glaszmann · L. Grivet
CIRAD, TA 40/03, Avenue Agropolis,
34398 Montpellier Cedex 5, France

Keywords AFLP · Chromosome pairing · Genetic map · Polyploidy · Sugarcane

Introduction

Cultivated sugarcane has a complex polyploid aneuploid genome derived from interspecific hybridization. Stimulated by disease outbreaks, modern sugarcane breeding emerged early in this century and quickly made use of interspecific crosses. The first interspecific hybrids involved the domesticated species, *S. officinarum* (SO) (2n=80), and the wild species, *S. spontaneum* (SS) (2n=40–128), and a series of backcrosses to SO as the recurrent female parent were performed. This process of introgression, called 'nobilization' (a return to the 'noble' SO type), has provided the major breakthrough in sugarcane improvement, solving some of the disease problems but also providing additional benefits in increasing yields, improving ratooning ability and enlarging adaptability to abiotic stress (Roach 1972). The restoration of high sugar-producing types was speeded up by the fact that the female SO parents transmitted their somatic chromosome number to the F_1 and BC_1 generations (Bhat and Gill 1985; Bremer 1961). *S. barberi* and *S. sinense*, two taxonomical groups probably derived from natural interspecific hybridization between SO and SS, have also contributed to the modern cultivar gene pool. In addition, there has been a marginal contribution of *S. robustum*, the putative wild ancestor of SO, in a few breeding stations. Current sugarcane breeding programs rely on extensive intercrossing of these derived elite cultivars. Contemporary cultivars have 100–130 chromosomes, among which 15–25% are contributed by SS (D'Hont et al. 1996). Recent cytogenetic studies have clarified our knowledge of the basic chromosome numbers in *Saccharum* genus, giving x=10 for SO and x=8 for SS (D'Hont et al. 1998).

Up to now, molecular genetic mapping of *Saccharum* species has been limited, being complicated by the coexistence of simplex and multiplex alleles and irregular chromosome numbers in the various homo(eo)logy classes due

to aneuploidy. Based on simplex markers (Wu et al. 1992), linkage maps have been developed for SS with random amplified polymorphic DNAs (RAPDs) (Al-Janabi et al. 1993) and restriction fragment length polymorphisms (RFLPs) (Da Silva et al. 1993; Da Silva et al. 1995; Ming et al. 1998), for SO with RAPDs (Mudge et al. 1996) and RFLPs (Ming et al. 1998) and for *S. robustum* with RFLPs (Ming et al. 1998). The construction of a map for a current cultivar, R570, has also been undertaken (Grivet et al. 1996) on the basis of a selfed progeny. R570 is the leading commercial variety in Reunion, Mauritius and Guadeloupe, and its rapid extension is to be found in some African countries and in Vietnam. The map encompassed 408 markers, generated by 120 RFLP probes, and these were placed onto 96 cosegregation groups (CGs). These CGs, which correspond to individual chromosomes or pieces of chromosomes, were tentatively assembled into ten basic homo(eo)logy or linkage groups (LGs). This latter effort provided the first opportunity to map genes of agricultural interest. A first gene could be tagged for resistance to the rust disease (Daugrois et al. 1996) and is being finely mapped (Asnaghi et al. 2000).

We are currently using R570 to investigate the genetic control of the main sugar yield components. This requires the use of a large-sized population characterized with many markers. We report here the construction of a genetic map with amplified fragment length polymorphism (AFLP) markers. This investigation provides insight into the genome organization of this typical representative of present-day cultivars, particularly with regard to the respective contributions of ancestral species, SO and SS, and the mode of chromosome assortment.

Material and methods

Plant material

The progeny analyzed in this study was derived from the self-fertilization of the modern cultivar, R570. R570 is derived from a cross between H32–8560 [Hawaiian Sugar Planters Association (HSPA), Hawaii] and R445 [Centre d'Essai de Recherche et de Formation (CERF), Réunion]. The total number of chromosomes in R570 was estimated to be around 112 on the basis of karyotypic observations (D'Hont et al. 1996). The progeny population consisted of a set of 295 individuals which were used for marker characterization and field evaluation. Plant material was provided by CERF. Due to constraints related to the availability of planting materials, this population did not include the individuals previously used by Grivet et al. (1996). This selfed population can be seen as a pseudo F_2 progeny.

In order to study the species origin of AFLP bands, we also characterized 12 SO clones supposed to be in the ancestry of current cultivars (Arceneaux 1965) with AFLPs: Banjarmasin Hitan, Black Cheribon, EK28 and Lothers – all presumably involved in the genealogy of R570 – and Badila, Crystallina, Fiji24, Kaludai Boothan, Korpi, Mauritius Guinghan, NG5752 and Vellai. These clones cover well the genetic diversity of SO that is present in modern cultivars (Jannoo et al. 1999).

AFLP protocol

Genomic DNA isolation from fresh leaves was according to Hoisington (1992). An aliquot for each genotype was standardized at 50 ng/µl prior AFLP manipulations. AFLP analysis (Vos et al. 1995) was performed with the Gibco Brl kit genome I, as recommended by the manufacturer, with some modifications designed to optimize the readability of the gels. Pre-amplification products were diluted 1:25 in water (Merck) instead of 1:50. Reaction products were analyzed on 5% denaturing polyacrylamide (20:1) gels with 7 *M* urea in 1×TBE buffer. In order to increase AFLP signals, we loaded a maximum of 6 µl of each sample onto the gel instead of the 2 µl recommended using a 62-sharkstooth comb. Electrophoresis was performed at 70 W, for approximately 2 h, on a 31×38.5 cm gel apparatus. The gels were subsequently placed for 20 min at 80°C in a gel dryer. Exposure of the gel to a Kodak Biomax MR X-ray film provided high resolution for those fragments present as a single copy in the genome (see below).

The 295 progeny individuals were characterized using the 37 AFLP primer combinations shown Table 1. The SO clones were characterized on the same gels for 35 primer combinations.

Marker scoring and analysis

Each polymorphic AFLP marker was identified by the primer combination consisting of six letters plus a band number indicated as a suffix. The first three letters represent *Eco*RI selective nucleotides and the last three, the *Mse*I selective nucleotides. Clear and unambiguous bands were scored in a presence versus absence fashion. We distinguished simplex markers among all possible cases. In the absence of segregation distortion, the expected segregation ratio for simplex markers is 3:1, irrespective of the type of chromosome pairing. For duplex markers, the expected segregation ratio is 15:1 in the case of disomic inheritance and slightly larger for polysomic inheritance. For triplex markers or multiplexes of higher levels, the segregation ratio is always higher, and it increases rapidly as the level of multiplexing increases. In order to distinguish the simplex markers from all the others, we retained all markers with a segregation ratio lower than 6.7:1 ($\sqrt{3x15:1}$). This ratio gives equal χ^2 statistical value for both 3:1 and 15:1 hypotheses (Mather 1957), the latter being the smallest theoretical ratio for all non-simplex markers. Simplex markers showing segregation distortion toward lower values on the basis of a χ^2 test at P=0.05 were considered as skewed and were marked on the map.

Map construction

The linkage relationships of simplex markers in coupling phase were determined using MAPMAKER 3.0 (Lander et al. 1987). Two-point analyses were performed at a LOD score threshold of 5 and a recombination fraction threshold of 0.35. The groups of linked markers identified at this stage are referred to as co-segregation groups (CGs). CGs were then ordered by multipoint analyses using standard procedures. When the most likely hypothesis was less than tenfold more likely than the second one, the order was considered to be ambiguous and this was denoted on the map. Map distances were derived from the recombination fraction using the Haldane function.

Determination of specific marker and segment origin

It is very useful to be able to trace whether a particular marker was transmitted from SO or from SS in order to delineate interspecific intra-chromosomal recombination patterns. It is, however, not straightforward because of the recognized uncertainty of the identity of some of the parents involved in the genealogy, in particular the SS clones. Therefore a procedure had to be worked out.

As a reference for the SO gene pool involved in modern cultivars, the set of noble clones described earlier was considered to be highly representative, given the structure of the variation known in this species (Jannoo et al. 1999); a marker absent from this set has a high probability of being derived from SS. Two such markers close to one another certainly indicate a segment derived from SS. Conversely, the presence of a marker in one or several of these

clones indicates that it can be derived from SO. The main informative elements for evaluating the attached probability are:

- three clones from SS have been extensively used, either directly or via a *S. barberi*/*S. sinense* intermediate, and another five more limitedly (Roach 1972);
- the probability that an RFLP marker present in SO is also present in a given SS clone is 0.20 (Lu et al. 1994a); RFLP markers have the same molecular basis of polymorphism as AFLP markers;
- in global terms, SS is expected to have contributed roughly one-fifth of the genome of R570. Computing this probability would require oversimplistic assumptions. However, it is reasonable to say that the first situation (absence of a marker) is more conclusive that the second one (presence of a marker).

In order to assign a tentative origin to a given region in a CG that bears markers of the two origins, we tested several procedures and finally chose to consider the following marker series (s and o for putative SS-specific or SO-specific, respectively):

- ...ss... or ...sos..., indicative of a SS-derived segment;
- ...ooo... or ...oosoo..., indicative of a SO-derived segment;
- the others being considered as ambiguous.

A quality index was assigned to this assessment of the whole CG segmental specificity, evaluated with the percentage of markers which follow the final succession of specific segments.

Investigation of chromosome assortment

Potential preferential pairing between chromosomes was investigated using a multi-step procedure. The first one was to systematically test all pairwise linkage repulsions between markers. To this end, the segregation data matrix was doubled after inverting the scores for each of the markers. A two-point analysis was performed on the new matrix with the threshold of LOD=3 and θ=0.35. We looked for linkages between the first and the newly created marker sets. This yielded couples of markers in repulsion. As a benchmark, a typical segregation of two alleles in a complete disomic situation with no missing data (74+/–:147+/+:74–/+) yields a LOD score of 10.9. The LOD score increases slightly if there is a departure from a 1:2:1 segregation, and it slightly decreases in the case of missing data. It drops to 9 when a -/- genotype appears (e.g. 74+/–:147+/+:73–/+:1–/–). The second step consisted of analyzing the distribution of such couples in repulsion. For those CGs that are sufficiently long, a preferential pairing is expected to result in an even distribution of repulsions along the CG pair. Conversely, artefactual repulsions have little chance of spanning long segments or entire CGs. The third step consisted of an additional test in the cases of large LOD scores. If two chromosomes systematically pair with one another, then each gamete will transmit one of any two allelic marker pairs borne by these chromosomes, and there will be no case of double absence of the two alleles (–/–) among the progeny. This kind of evidence can help differentiate between complete disomic behavior and incomplete preferential pairing.

This whole analysis is complicated by the fact that AFLP markers cannot be grouped into allele classes, as can RFLP markers. The cases of detected repulsions are all the more indicative of preferential pairing.

Results

Map construction

A total of 4,611 discrete bands were generated using the 37 primer combinations (Table 1). Of these, 26% (1,180) were polymorphic. The number of bands per primer pair ranged from 85 to 166 (mean of 125), and the number of segregating bands ranged from 18 to 43 (mean of 32). The level of polymorphism for individual primer combinations varied from 15% to 40%. According to the criterion we had selected, 939 of the 1,180 segregating markers were found to be simplex (Fig. 1). These markers were used to build the map based on linkages in coupling phase. Among them, 79 were skewed toward lower presence frequencies at P=0.05.

In order to avoid false linkages, cosegregation groups were built up at a stringent LOD score threshold of 5. This high value was chosen with respect to the high number of markers and the high number of expected CGs. As a result, 887 simplex markers were assigned to 120 CGs (Fig. 2), and 52 markers remained unlinked. The length of these CGs spans a very wide range, from 1 to 232 cM. Each single primer combination generated markers which were scattered over between 13 and 29 different CGs (Table 1). The cumulative length of all CGs was 5,849 cM, with an average distance between two markers of 6.5 cM. However, the distribution of markers along the groups was irregular. Six gaps were slightly larger than 35 cM, and these corresponded to a two-point distance slightly inferior than this threshold. Given the large number of progeny surveyed, ambiguous marker orders were strictly limited to some regions where linkage was tight.

Of the 79 skewed markers, 72 were scattered on 27 cosegregation groups; the remaining 7 were unlinked. Forty-one markers clustered on five chromosomal segments, namely CGs 4, 9, 66, 71 and 78. Note that the number of skewed chromosomal regions (27+7=34) is not very different from what would be expected by the sole random effect according to the Bonferoni procedure (23).

Comparison with the RFLP map

The correspondence between LGs of the AFLP and the RFLP (Grivet 1996) maps was established on the basis of 45 AFLP markers used as anchors between the two mapping populations (unpublished results). Of 120 CGs identified here, 27 were assigned to a LG of the RFLP map (Table 2). The Roman numbers (I–X) refer to linkage groups (LGs), and each CG is identified by the Roman number of the LG to which it belongs plus a specific Arabic number. Anchor markers are in boxes in Fig. 2. Correspondence between groups was assessed on the basis of one to six bridge markers. The genetic map presented in Fig. 2 is organized according to LGs when possible. Groups II and III as well as groups V and VI were each assembled into a single LG to account for the progress in RFLP mapping (unpublished results).

Mapping coverage has been greatly improved compared to the RFLP map. We used a population almost fourfold larger (295 individuals compared to 77), and the number of markers had been doubled (887 AFLP markers as opposed to 408 RFLP markers) compared to the

Table 1 AFLP primer pairs used and polymorphism revealed in this survey of selfed progeny of cv. R570

Primer pair *Eco*RI/*Mse*I	Number of visible bands	Number of polymorphic bands	% of polymorphism	Number of simplex markers	Number of cosegregation groups covered
Eaac/Mcac	135	36	26.7	28	21
Eaag/Mcaa	154	29	18.8	24	18
Eaag/Mcac	136	27	19.8	18	16
Eaag/Mcag	153	36	23.5	27	20
Eaag/Mcat	142	34	23.9	25	20
Eaag/Mcta	166	40	24.1	34	29
Eaag/Mctc	104	32	30.8	25	20
Eaag/Mctt	154	40	26.0	35	26
Eaca/Mcaa[a]	148	23	15.5	16	15
Eaca/Mcta	120	22	18.3	20	17
Eaca/Mctc	98	33	33.7	29	27
Eaca/Mctg	105	43	40.9	33	25
Eaca/Mctt[a]	132	29	21.9	23	20
Eacc/Mcaa	127	35	27.5	25	22
Eacc/Mcac	85	27	31.8	19	17
Eacc/Mcat	122	41	33.6	32	27
Eacc/Mctc	121	25	20.7	20	17
Eacc/Mctg	116	41	35.3	31	22
Eacc/Mctt	129	27	20.9	25	21
Eacg/Mcag	148	28	18.9	22	19
Eacg/Mcta	106	29	27.3	22	19
Eacg/Mctt	154	44	28.6	36	29
Eact/Mcac	106	35	33.0	34	26
Eact/Mcag	134	23	17.2	18	16
Eact/Mcat	125	35	28.0	29	26
Eact/Mctg	106	39	36.8	27	18
Eact/Mctt	98	29	29.6	25	19
Eagc/Mcaa	115	28	24.3	24	20
Eagc/Mcag	118	18	15.2	17	13
Eagc/Mcat	148	32	21.6	25	21
Eagc/Mcta	146	32	21.9	24	21
Eagc/Mctc	94	25	26.6	22	18
Eagc/Mctt	105	39	37.1	32	23
Eagg/Mcat	128	39	30.5	31	21
Eagg/Mcta	116	36	31.0	27	23
Eagg/Mctc	103	28	27.2	19	15
Eagg/Mctg	114	21	18.4	16	14
Total	4,611	1,180	–	939	120
Mean	124.6	32	25.6	25.5	20.6
Range	85–166	18–43	–	16–36	13–29

[a] AFLP combination for which the R570 banding pattern was not compared with the banding pattern of the 12 *S. officinarum* clones (see text)

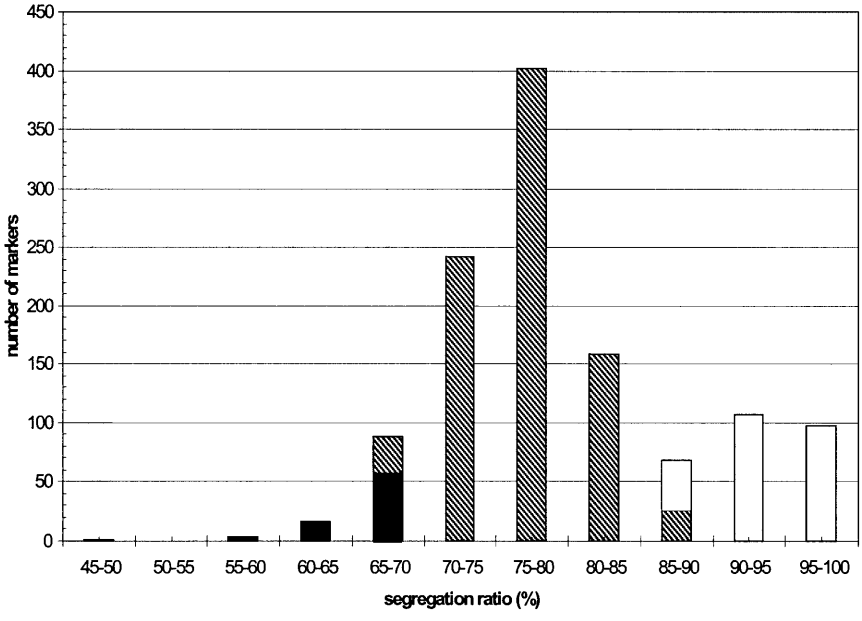

Fig. 1 Distribution of the segregation ratios (percentage of progeny showing the marker) of the 1,180 segregating markers as revealed by 37 AFLP primer combinations in the self progeny of R570. The 939 markers retained as simplex for map construction are shown by the *hatched area*. The 79 markers skewed toward lower values are in *black* (see text for threshold determination)

Table 2 Tentative species origin of CG segments based on the putative origin of AFLP markers and comparison with previous results obtained with RFLP. The AFLP-based origin of segments was attributed as described in the text. The origin of chromosome segments based on the RFLP map was derived from Grivet et al. (1996)

AFLP map					RFLP map		
CG	Length (cM)	Succession of putative specific markers[a]	Segmental specificity	R[b]	CG[c]	Length (cM)	Segmental specificity
29	232	oSSooSoSSOOOOSSSSoSSSSo	SOS	17/23	X3(3)	14	S
15	191	ooSSSoSoSoSSoSOOsOOO	SO	13/20	VIII2(1)	78	S
74	180	SSSSSSOOOOOOSSS	SOS	15/15			
28	174	oSSSSOOOOOOOOSSSS	SOS	16/17	X2(0)	131	OSOS
14	173	SSoSSSSSoSSSSSSoo	S	13/17	VIII1(3)	89	S
4	155	SSSSOOOSoSSoSS	SOS	12/14	III1(4)	45	S
9	147	oSSOOOOOOOOOSoSoS	SOS	14/17	IV2(6)	79	S
26	145	OOOOSSSSoSSoSSS	OS	13/15	X1(1)	98	S
56	145	SSSoSoSSSSS	S	9/11			
24	136	OOOSSoSSSSSoo	OS	10/13	IX2(1)	73	S
8	136	SoSSSSSSoS	S	8/10	IV1(2)	86	S
82	131	OOOOOOOOOsOOOsOOO	O	15/17			
55	126	OOOOOOO	O	7/7			
23	123	OOOOOOOOO	O	9/9			
59	120	OsOOOO	O	5/6			
22	118	oSSSSoS	S	5/7	IX4(1)	9	S
52	105	OOOOO	O	5/5			
66	102	s					
53	100	SSSSSSSoSSo	S	9/11			
94	96	OOOOO	O	5/5			
90	91	OOOSSSoSo	OS	7/9			
57	88	OOsOOSSSoo	OS	7/10			
93	86	OOOOOOO	O	7/7			
2	79	OOsOOOOOSS	OS	9/10	I9(1)	9	S
78	79	OOOOOOO	O	7/7			
51	77	OOOOOOOOOOOOOO	O	14/14			
30	73	OOOOOOOOO	O	9/9	X12(1)	10	O
32	73	OOOOOO	O	6/6	X9(1)	29	O
63	71	OOOOOOO	O	7/7			
10	71	OOOOO	O	5/5	V2(1)	13	
61	70	OOOOOOOOOOO	O	11/11			
19	68	sOOO	O	3/4	VIII10(2)	7	O
88	64	OOOO	O	4/4			
72	63	OOsOO	O	4/5			
71	56	OOOOO	O	5/5			
87	55	OOOO	O	4/4			
107	53	OOOO	O	4/4			
79	53	OOOO	O	4/4			
76	50	OOOO	O	4/4			
75	48	OOOOO	O	5/5			
105	46	OOOO	O	4/4			
111	43	SoS	S	2/3			
73	40	OOOO	O	4/4			
20	40	OOO	O	3/3	VIII5(1)	15	O
1	39	OOOsOO	O	5/6	I4(2)	5	O
33	39	SSSS	S	4/4	X1(2)	102	S
64	38	OOO	O	3/3			
92	38	OOO	O	3/3			
80	36	OOO	O	3/3			
42	36	OOO	O	3/3			
81	35	OO	O	2/2			
119	35	OOO	O	3/3			
106	35	OOO	O	3/3			
108	34	OOOOOO	O	6/6			
101	34	oso					
38	34	OO	O	2/2			
91	34	OOO	O	3/3			
27	33	OOOO	O	4/4	X11(1)	7	O
65	33	OO	O	2/2			
96	33	OO	O	2/2			
3	31	OOOOO	O	5/5	I5(1)	9	O
120	31	OOOO	O	4/4			
7	30	OO	O	2/2	III2(0)	25	S

Table 2 (continued)

AFLP map					RFLP map		
CG	Length (cM)	Succession of putative specific markers[a]	Segmental specificity	R[b]	CG[c]	Length (cM)	Segmental specificity
68	30	OOOOO	O	5/5			
89	30	OOO	O	3/3			
60	30	OOOOO	O	5/5			
77	29	OO	O	2/2			
35	29	OOOOO	O	5/5			
62	29	OOOOOOO	O	7/7			
69	28	so					
109	26	OOO	O	3/3			
37	26	oSS	S	2/3			
16	25	oSSoSSS	S	5/7	VIII11(2)	22	
99	24	OO	O	2/2			
40	23	so					
86	23	o					
58	23	o					
47	22	o					
67	21	SoSSS	S	4/5			
45	20	OOO	O	3/3			
34	18	o					
48	18						
116	17	o					
118	17	OOO	O	3/3			
6	17	OOO	O	3/3	III6(1)	3	
25	17	OOOs	O	3/4	X10(2)	6	S
17	16	o			VIII8(1)	30	
70	16	OOO	O	3/3			
102	16	OO	O	2/2			
49	16	OO	O	2/2			
31	14	OOO	O	3/3			
5	14	OOOO	O	4/4	II6(1)	16	
84	13	OOO	O	3/3			
39	13	OO	O	2/2			
95	13	OOOO	O	4/4			
43	13	o					
110	12	o					
113	12	OO	O	2/2			
11	11	OO	O	2/2	VII3(1)	21	O
54	10	OOOO	O	4/4			
98	10						
50	10	OOOO	O	4/4			
97	9	OOO	O	3/3			
115	8	o					
85	8	OOO	O	3/3			
21	7	o					
114	7	OO	O	2/2			
13	6	OO	O	2/2	VII1(1)	50	O
44	5	OO	O	2/2			
104	5	so					
103	4	OO	O	2/2			
100	4	OO	O	2/2			
36	4	OO	O	2/2			
112	3	OO	O	2/2			
18	3	OO	O	2/2			
12	2	SS	S	2/2	VII2(1)	35	O
46	2	o					
117	2						
41	2	OO	O	2/2			
83	1	OO	O	2/2			
Total	5,850						

[a] In capitals those that determine the final tentative segmental specificity: s, *S. spontaneum*, o, *S. officinarum*

[b] R, Ratio of the number of informative markers for determining segmental specificity over the total number of markers for which origin was determined in the CG

[c] Number of anchor markers between RFLP and AFLP map is given in parenthesis

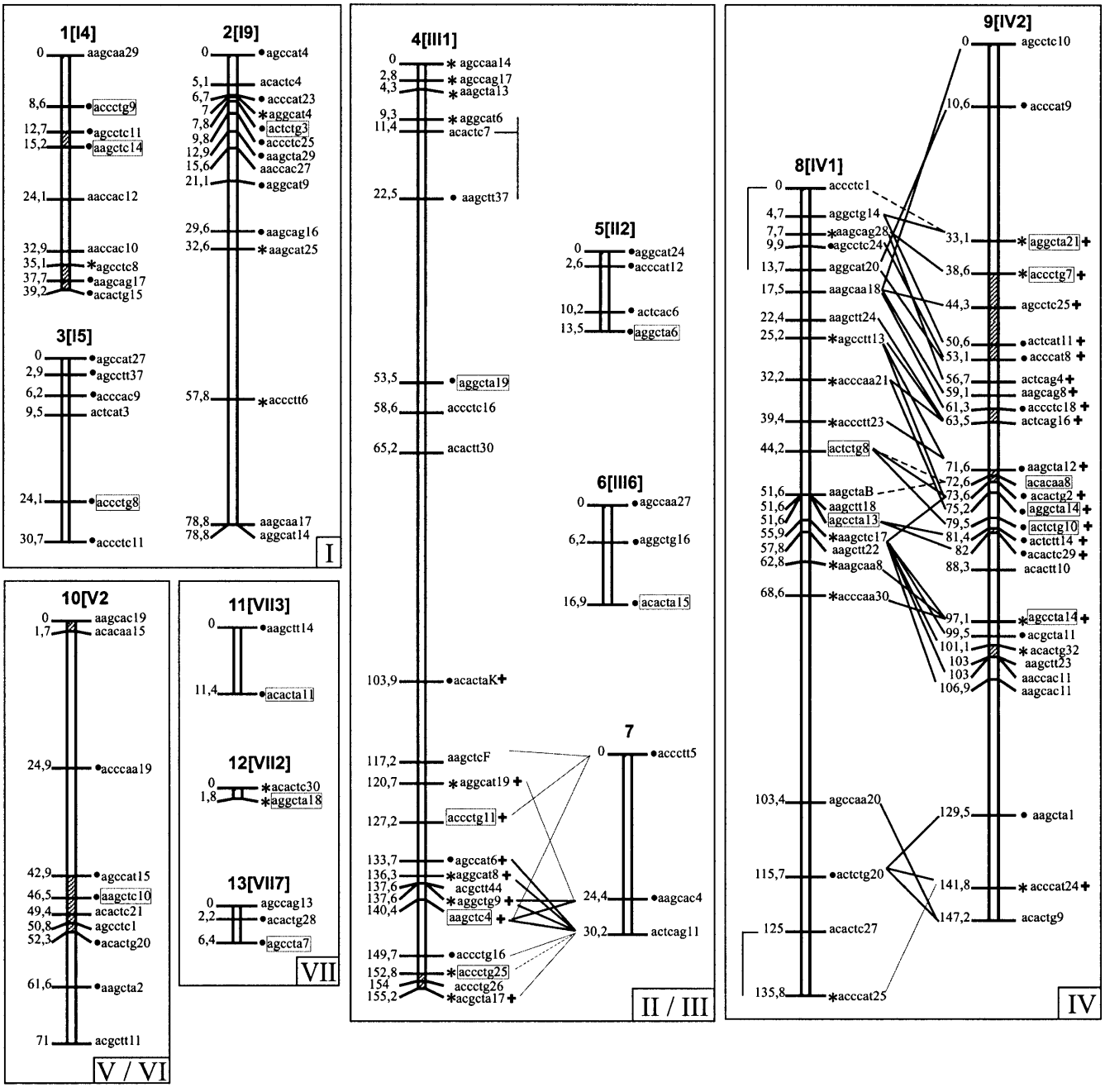

Fig. 2 AFLP linkage map of sugarcane cv. R570 generated from a self-progeny of 295 individuals. Segregation data permitted 887 markers generated by 37 primer pairs to be assembled into 120 CGs. The threshold for the LOD score and maximum recombination fraction were 5 and 35 cM, respectively. The numbers on the *left* of a CG are the cumulated genetic distances in Haldane centiMorgans. Marker names are on the *right*. When possible, map presentation is organized according to LGs (identified by a Roman numeral, from *I* to *X*) defined in a previous RFLP map performed on another population derived from the same cross (Grivet et al. 1996). CGs of a same LG are enclosed in a *large rectangle*. Markers in *boxes* represent AFLP bridge markers with the previous RFLP map. CGs are identified by a number at the *top*; second label in *square brackets* indicates the corresponding name on the RFLP map (see Table 2). Uncertain orders (alternate orders not ruled out at LOD=1) are represented by *hatches*. When the position of a single marker was not clearly determined according to the same threshold, a «*T*» *bar* was attached to the relevant marker on the *left* or on the *right* of the group. Linkage in repulsion between two markers is represented by a *dashed*, a *continuous* or a *bold line*, indicating moderate (3≤LOD<4), strong (4≤LOD<5) or very strong (LOD≥5) linkage, respectively. When the species origin of a marker was known, it is indicated with: * for a putative *S. spontaneum* origin, or ● for a putative *S. officinarum* origin. A *plus* (+) indicates that the segregation of a marker is skewed toward lower values at P=0.05

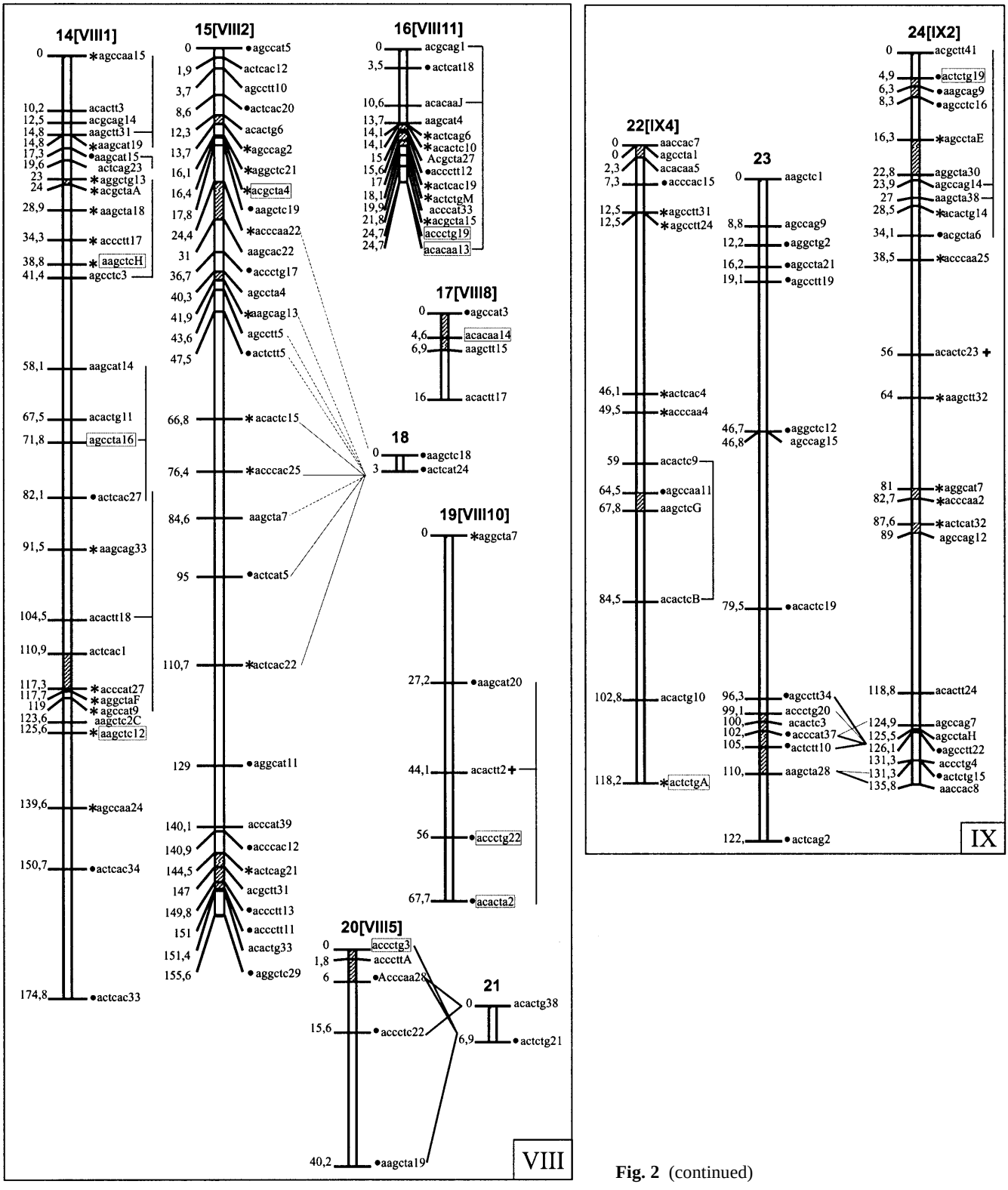

Fig. 2 (continued)

RFLP map. This enabled us to obtain 120 CGs as opposed to 96. The total length of the map was increased by 3,842 cM (5,850 cM as compared to 2,008 cM). This increase is accounted for by 24 additional groups and a large extension of the majority of the previously existing groups. However, 41 (33%) of the 120 groups still have a length inferior to 20 cM, and 76 groups (63%) are shorter than 40 cM (Table 2).

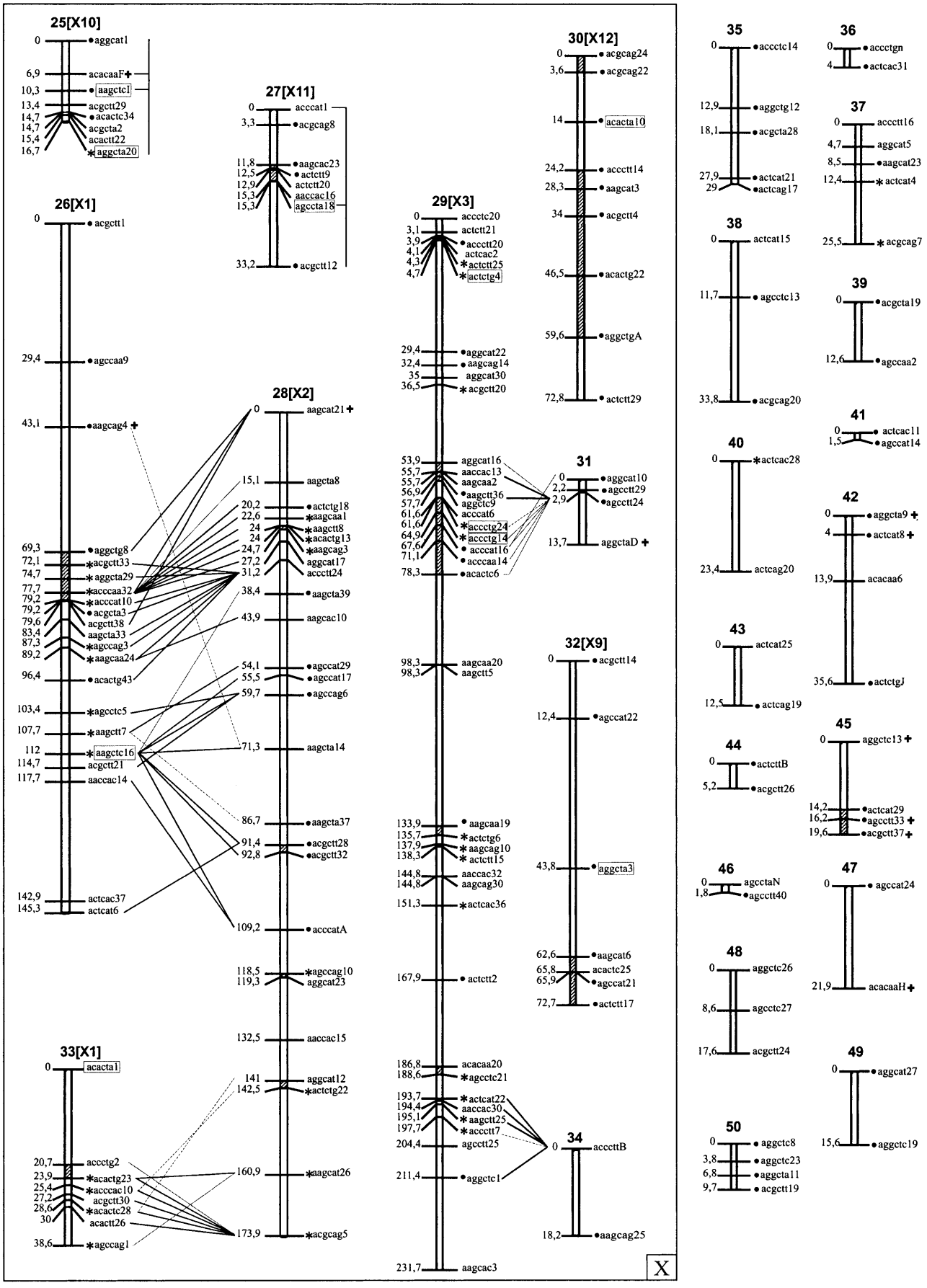

Fig. 2 (continued)

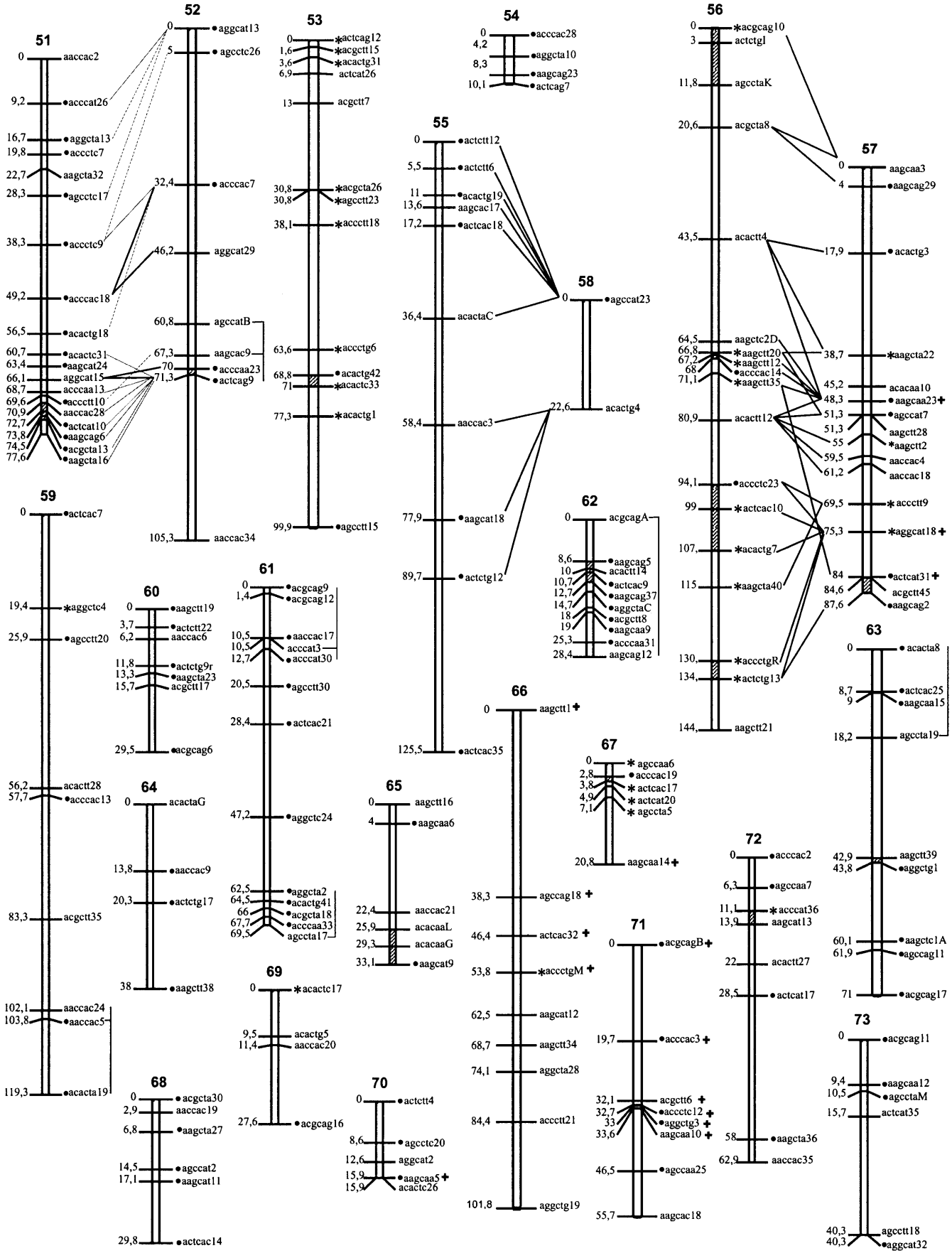

Fig. 2 (continued)

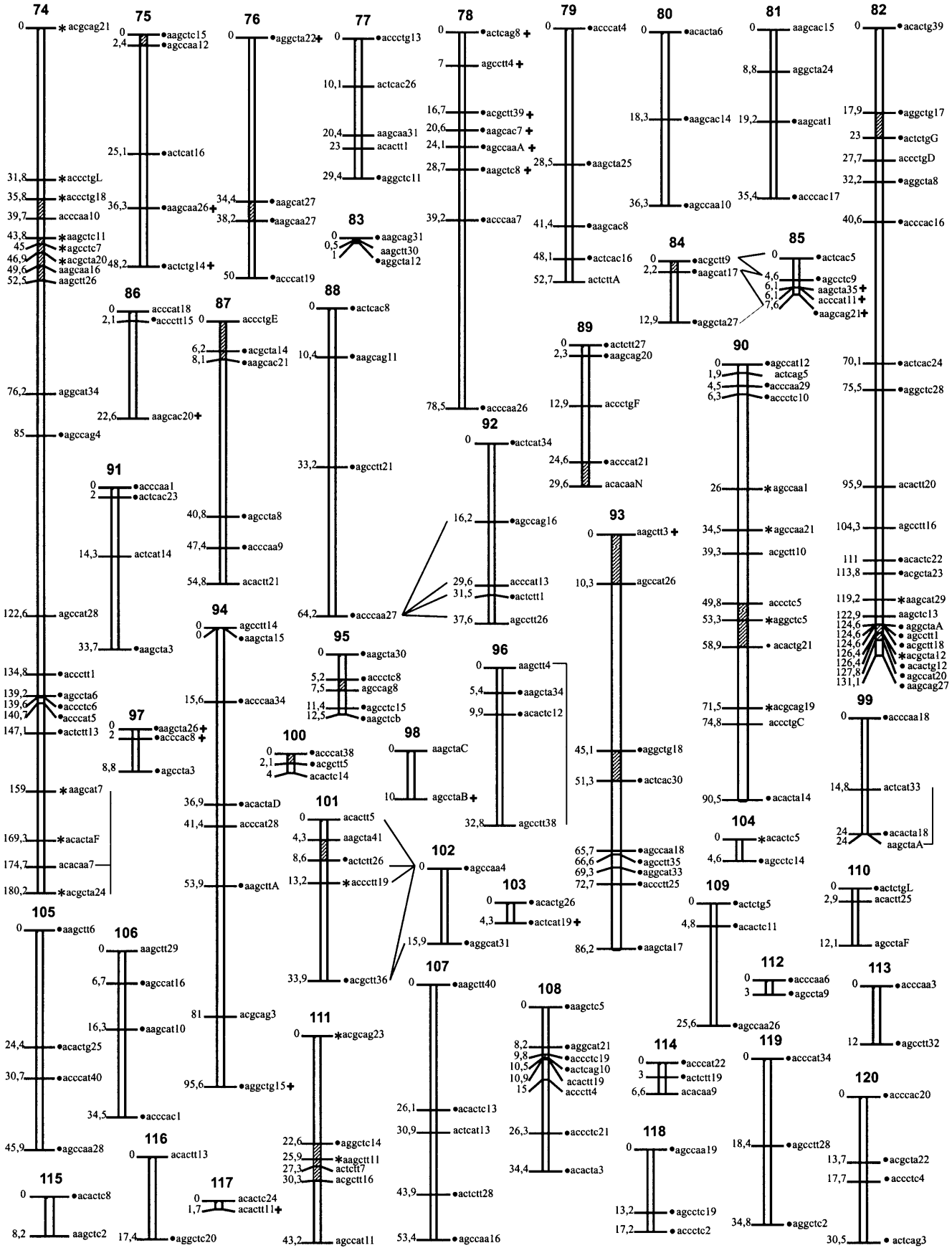

Fig. 2 (continued)

Species origin of markers

The sample of 12 SO clones putatively involved in the ancestry of modern cultivars was compared with R570 AFLP patterns for 35 of the 37 primer pairs used in this study. This allowed us to investigate the species origin of 591 of the 887 markers of the map, for which corresponding bands were clearly scorable without any interference with additional bands over the 12 clones surveyed. Of the 591 markers studied, 155 were assigned to SS and 436 to SO. The information available for each of the 120 CGs is summarized in Table 2. The specificity of the CG was considered for those groups with at least two markers with a tentative specific origin. As shown in Table 2, this gave a sole SS origin for 11 CGs, a sole SO origin for 79 CGs and a recombined origin with chromosomes segments of both ancestral species for 11 CGs. No origin could be attributed to 19 CGs.

Among the 11 CGs with a sole SS origin, the concordance of origin with the RFLP map was good except for a single CG of the second group for which a SO origin had been assessed based on a single RFLP marker. These CGs can be divided in two groups: (1) 6 CGs (including CG33 on the basis of its correspondence to X1) with a size larger than 100 cM and a species origin assessed on the basis of a minimum of five markers; (2) 5 CGs smaller than 43 cM with a species origin assessed on the basis of two to five markers. The first set may account for chromosomes inherited from SS with only limited or no recombination with SO. The second set may account for small pieces of SS chromosomes introgressed into SO chromosomes that are not targeted by markers in the present map. The CGs with a sole SO origin ranged in size between 1 and 131 cM, with 71% of them being smaller than 40 cM. The correspondence with CGs of the RFLP map was possible for 11 of them, and the congruence of origin was good except for two, namely CGs 7[III2] and 25[X10], for which one marker had been attributed a SS origin with RFLPs. The general small size of SO CGs is in agreement with what was observed on the RFLP map and is related to the lower diversity in this species (Lu et al. 1994a, b) and/or the higher redundancy in the SO fraction kept in modern cultivars, which result in fewer simplex markers, as already discussed (D'Hont et al. 1994; Grivet et al. 1996). The 11 CGs with a recombined origin ranged in size between 79 and 232 cM, and they were all included among the longest 24 CGs identified on the map. Correspondence with the RFLP map was possible for 8 CGs out of the 11. Seven of them had been assessed a sole SS origin with RFLPs, whereas one of them, CG 28 [X2], had been attributed a recombined origin. This discrepancy is probably due to the much higher number of markers screened for origin in the present study, which enabled a finer depicting of segment origin along CGs.

Chromosome pairing behavior

We examined the linkage in repulsion of each marker with all other markers of the map. A total of 28 CGs bore markers that were linked in repulsion with other markers at a LOD≥3. Among these, 22 CGs were linked two-by-two, giving 11 pairs (Fig. 2). The other six CGs corresponded to two cases where two small CGs were linked in repulsion with a large one: CG 28, in repulsion with CGs 26 and 33, and CG 29, in repulsion with CGs 31 and 34 (LG X). In both cases the two small CGs are in repulsion with different portions of the large CG, indicating that they may belong to a same chromosome not evenly covered by markers. Indeed, this hypothesis of a single chromosome is confirmed in the case of CGs 26 and 33, owing to the cosegregation on the same RFLP CG of anchor markers borne by each segment.

For each marker of each CG involved in a preferential pairing pair, we retained only the strongest repulsion linkage (Fig. 2). LOD scores reached higher than 5, even 10 or more in a few cases: CG8 and CG9, CG20 and CG21, CG56 and CG57, and CG101 and CG102. When the length of the CGs involved in repulsion was high, significant linkages were regularly distributed, which is concordant with a preferential pairing of corresponding chromosomes at meiosis. The absence of –/– appeared to be the rule, with the most extreme case (CG8 and CG9) displaying 35 couples of markers with no –/– segregant. This makes all of these pairs good candidates for complete disomic behavior. The same is probably true for CG26 and CG28, despite the slightly lower LOD scores. The observed LOD values are higher than those found in our previous RFLP map (Grivet et al. 1996) due to the larger size of the progeny population used here, but the data are compatible; the repulsion between IV1 and IV2 and between X1 and X2 found with RFLPs is clearly confirmed here.

The CGs linked in repulsion with 1 of the 27 CGs already assigned to a LG were also assigned to the corresponding LG. These CGs are 7, 18, 21, 23, 28, 31 and 34, which are assigned to LGs III, VIII, VIII, IX, X, X and X, respectively. According to repulsions observed in the RFLP map, groups 7 (LG III) and 28 (LG X) are probably the former III2 and X2, respectively. Finally, the number of CGs assigned to a particular LG was increased up to 34 as a result of repulsion linkages, while 86 groups were left unassigned (Fig. 2, Table 2). Some of the pairs of CGs identified here may already have been identified in the RFLP map, but correspondence may not have been established due to a loose coverage by the anchor markers.

Among the 13 pairs of chromosomes that displayed preferential pairing, 7 involved repulsion between CGs, most probably of sole SO origin, and 6 involved at least one recombined CG. Among the latter, the most spectacular instance is that of CG8 and CG9, which display a very strong and continuous repulsion between a chromosome probably derived from SS and a chromosome derived from interspecific recombination, thus exemplifying the opportunities for fine interspecific recombinations.

Discussion

The total length of the present map has been increased by 3,842 cM compared to the previous RFLP map performed on the same cross, and the proportion of unlinked simplex markers has been substantially reduced, from 19.2% (97/505) to 5.5% (52/939). Nevertheless, the map is obviously not saturated, given the small size of many CGs and the 52 markers that remained unlinked. A rough but simple way to estimate the complete length of the map would be to multiply the number of homo(eo)logous chromosomes by the mean length of the longest CGs, which are probably close to saturation. The difference in the basic chromosomal numbers in SO and SS, which are x=10 and x=8, respectively, has, however, to be considered. This difference is probably due to simple Robertsonian events that may have carried out in two cases to one SS chromosome being collinear to two SO chromosomes. Candidate LGs that may account for this are LGs VIII and the X (Grivet et al. 1996; Glaszmann et al. 1997). It is noteworthy that four of the five largest CGs (14, 15, 28 and 29) belong to those two LGs and that they have a potential SS structural organization, as indicated by their SS or mixed SO/SS origin. We excluded those four CGs from the computation plus CG 74 for which the LG is unknown. The following largest CGs were distributed among three different LGs (LGs II/III, IV, IX) and had a length of around 150 cM, which is the usual size for chromosome maps in diploids. Therefore an estimate of the global length of the map can be made by multiplying this length by the total chromosome number: $112 \times 150 \approx 17{,}000$ cM. This value is very high but should not be compared with those obtained in diploids or amphidiploids where pairing is strictly disomic and where by definition the length of two homologous chromosomes is counted once. Based on this computation, the present map would cover around one-third (5,849/17,000) of the genome length of R570. There is an obvious deficiency in the coverage of the SO fraction of the genome. Mapping in any sugarcane cultivar should lead to a similar discrepancy. In this perspective, producing a saturated map for a cultivar remains an enormous task despite the rapid improvement of available marker technologies.

This study enabled a reconsideration of the species origin of chromosome segments based on an increased resolution due to the high number of markers investigated. Our determination of marker origin was based on a survey of 12 SO clones and is only tentative. The procedure for studying the linear organization of this specificity was designed to be rather conservative by considering several adjacent markers, that is several independent pieces of information. On this basis, 21 CGs had at least one unambiguous SS inherited fragment. Among these, 6 large CGs have a SS origin, 11 large CGs have a double SS and SO origin and the 5 small CGs with a solely SS origin may depict small SS segments borne by one to five predominantly SO chromosomes, as discussed above. It is unlikely that an entire chromosome inherited from SS had escaped mapping given the higher level of polymorphism in this species, and six can be considered to be the maximum number of full-length SS chromosomes. These data should be compared with results obtained using GISH (D'Hont et al. 1996), which identified for R570 some 12 SS chromosomes, with 11 chromosomes issuing from a single or a double recombination between the two ancestral species. The difference suggests that GISH efficiently identifies SS genome portions but may have a resolution power that is insufficient to identify all recombined chromosomes; this information will be useful for future GISH analysis of other modern cultivars.

Our investigation of repulsion linkage identified 13 cases of CG pairs that preferentially pair at meiosis. In most cases the LOD scores were high, giving little opportunities for artifacts. Moreover, when a comparison was possible, the pairs were consistent with those detected in the RFLP map, which was constructed from a distinct progeny sample derived from the same cross (Grivet et al. 1996). The number of pairs detected here was, however, higher due to a twofold increase in the number of markers, which permitted a better map coverage, and to a fourfold increase in the number of progeny, which permitted a higher power for the detection of repulsion. The numerous chromosome segments not yet covered by markers prevent an exhaustive overview of pairing along the whole map and between all the homo(eo)logs. These data do confirm the atypical pairing behavior that was perceptible in the RFLP analysis. The existence of some preferential pairing definitively rejects the hypothesis of complete polysomy; the data even suggest the possibility of complete local disomy.

The use of the map for a search for QTLs raises the matter of coverage efficiency. The map covers about one-third of the genome. For a QTL analysis, the distal markers of unsaturated CGs (about a hundred) can be considered to be efficiently tagging an additional 10 cM. Similarly, the unlinked markers would cover 20 cM. This extends by about 3,000 cM the coverage of the map for tagging useful genes and brings the efficient coverage to about one-half of the genome of R570. However, additional efforts will be necessary for the determination of homo(eo)logy relationships and for a more systematic allele resolution in regions of interest. The use of RFLP or microsatellite loci well-scattered along the basic map should provide more opportunity for a criss-cross coverage of the genome thanks to a fine allele resolution. The identification of candidate regions for QTLs will help target the resolution efforts. For that purpose one will benefit from the genetics of other grasses (Glaszmann et al. 1997), especially sorghum whose genome is very closely related and highly collinear with the one of sugarcane (Dufour et al. 1997; Ming et al. 1998).

References

Al-Janabi SM, Honeycutt RJ, McClelland M, Sobral BWS (1993) A genetic linkage map *of Saccharum spontaneum* L. 'SES 208'. Genetics 134:1249–1260

Arceneaux G (1965) Cultivated sugarcanes of the world and their botanical derivation. Proc Int Soc Sugar Cane Technol 12: 844–854

Asnaghi C, Paulet F, Kaye C, Grivet L, Deu M, Glaszmann J-C, D'Hont A (2000) Application of synténie across Poaceae to determine the map location of a rust resistance gene of sugarcane. Theor Appl Genet 101:962–969

Bhat SR, Gill BS (1985) The implication of 2n egg gametes in nobilisation and breeding of sugarcane. Euphytica 34:377–384

Bremer G (1961) Problems in breeding and cytology of sugar cane. Euphytica 10:59–78

Da Silva JAG, Burnquist WL, Tanksley SD (1993) RFLP linkage map and genome analysis of *Saccharum spontaneum*. Genome 36:782–791

Da Silva J, Honeycutt RJ, Burnquist W, Al-Janabi SM, Sorrells ME, Tanksley SD, Sobral BWS (1995) *Saccharum spontaneum* L. 'SES 208'genetic linkage map combining RFLP- and PCR-based markers. Mol Breed 1:165–179

Daugrois JH, Grivet L, Roques D, Hoarau JY, Lombard H, Glaszmann JC, D'Hont A (1996) A putative major gene for rust resistance linked with an RFLP marker in sugarcane cultivar R570. Theor Appl Genet 92:1059–1064

D'Hont A, Lu YH, Gonzalez de Leon D, Grivet L, Feldmann P, Lanaud C, Glaszmann J-C (1994) A molecular approach to unravelling the genetics of sugarcane, a complex polyploid of the Andropogoneae tribe. Genome 37:222–230

D'Hont A, Grivet L, Feldmann P, Rao S, Berding N, Glaszmann J-C (1996) Characterisation of the double genome structure of modern sugarcane cultivars (*Saccharum* spp.) by molecular cytogenetics. Mol Gen Genet 250:405–413

D'Hont A, Ison D, Alix K, Roux C and Glaszmann J-C (1998) Determination of basic chromosome numbers in the genus *Saccharum* by physical mapping of ribosomal RNA genes. Genome 41:221–225

Dufour P, Deu M, Grivet L, D'Hont A, Paulet F, Bouet A, Lanaud C, Glaszmann J-C, Hamon P (1997) Construction of a composite sorghum map and comparative genetic mapping with sugarcane, a related complex polyploid. Theor Appl Genet 94:409–418

Glaszmann J-C, Dufour P, Grivet L, D'Hont A, Deu M, Paulet F, Hamon P (1997) Comparative genome analysis between several tropical grasses. Euphytica 96:13–21

Grivet L, D'Hont A, Roques D, Feldmann P, Lanaud C, Glaszmann J-C (1996) RFLP mapping in a highly polyploid and aneuploid interspecific hybrid. Genetics 142:987–1000

Hoisington D (1992) Laboratory protocols. CIMMYT applied molecular genetics laboratory, CIMMYT, Mexico, D.F.

Jannoo N, Grivet L, Seguin M, Paulet F, Domaingue R, Rao PS, Dookun A, D'Hont A, Glaszmann JC (1999) Molecular investigation of the genetic base of sugarcane cultivars. Theor Appl Genet 99:171–184

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

Lu YH, D'Hont A, Walker DIT, Rao PS, Feldmann P, Glaszmann J-C (1994a) Relationships among ancestral species of sugarcane revealed with RFLP using single copy maize nuclear probes. Euphytica 78:7–18

Lu YH, D'Hont A, Paulet F, Grivet L, Arnaud M, Glaszmann J-C (1994b) Molecular diversity and genome structure in modern sugarcane varieties. Euphytica 78:217–226

Mather K (1957) *The measurement of linkage in heredity*. Methuen, London.

Ming R, Liu S-C, Lin Y-R, Da Silva J, Wilson W, Braga D, Van Deynze A, Wenslaff TF, Wu KK, Moore PH, Burnquist W, Sorrells ME, Irvine JE, Paterson AH (1998) Detailed alignment of *Saccharum* and *Sorghum* chromosomes: comparative organization of closely related diploid and polyploid genomes. Genetics 150:1663–1682

Mudge J, Andersen WR, Kehrer RL, Fairbanks DJ (1996) A RAPD genetic map of *Saccharum officinarum*. Crop Sci 36: 1362–1363

Roach B (1972) Nobilisation of sugarcane. Proc Int Soc Sugar Cane Technol 14:206–216

Vos P, Hogers R, Bleeker M, Reijans M, van de Lee T, Hornes M, Frijters A, Pot J, Peleman J, Kuiper M, Zabeau M (1995) AFLP: a new technique for DNA fingerprinting. Nucleic Acids Res 23:4407–4414

Wu KK, Burnquist W, Sorrels ME, Tew TL, Moore PH, Tanksley SD (1992) The detection and estimation of linkage in polyploids using single-dose restriction fragments. Theor Appl Genet 83:294–300