Theor Appl Genet (2005) 110: 789–801
DOI 10.1007/s00122-004-1813-7

ORIGINAL PAPER

K. S. Aitken · P. A. Jackson · C. L. McIntyre

A combination of AFLP and SSR markers provides extensive map coverage and identification of homo(eo)logous linkage groups in a sugarcane cultivar

Received: 3 January 2004 / Accepted: 27 August 2004 / Published online: 8 February 2005

Abstract Sugarcane varieties are complex polyploids carrying in excess of 100 chromosomes and are derived from interspecific hybridisation between the domesticated *Saccharum officinarum* and the wild relative *S. spontaneum*. A map was constructed in Q165[♂1], an Australian cultivar, from a segregating F_1 population, using 40 amplified fragment length polymorphism (AFLP) primer combinations, five randomly amplified DNA fingerprints (RAF) primers and 72 simple sequence repeat (SSR) primers. Using these PCR-based marker systems, we generated 1,365 polymorphic markers, of which 967 (71%) were single-dose (SD) markers. Of these SD 967 markers, 910 were distributed on 116 linkage groups (LGs) with a total map length of 9,058.3 cM. Genome organisation was significantly greater than observed in previously reported maps for *Saccharum* spp. With the addition of 123 double-dose markers, 36 (3:1) segregating markers and a further five SD markers, 1,074 markers were mapped onto 136 LGs. Repulsion phase linkage detected preferential pairing for 40 LGs, which formed 11 LG pairs and three multi-chromosome pairing groups. Using SSRs, double-dose markers and repulsion phase linkage, we succeeded in forming 127 of the 136 LGs into eight homo(eo)logy groups (HG). Two HGs were each represented by two sets of LGs. These sets of LGs potentially correspond to *S. officinarum* chromosomes, with each set aligning to either end of one or two larger LGs. The larger chromosomes in the two HGs potentially correspond to *S. spontaneum* chromosomes. This suggestion is consistent with the different basic chromosome number of the two species that are hybridised to form sugarcane cultivars, *S. spontaneum* ($x = 8$) and *S. officinarum* ($x = 10$), and illustrates the structural relationship between the genomes of these two species. The discrepancy of coverage between HGs highlights the difficulty in mapping large parts of the genome.

Communicated by P. Langridge

K. S. Aitken (✉) · C. L. McIntyre
Queensland Bioscience Precinct,
CSIRO Plant Industry, 306 Carmody Rd,
St Lucia, QLD, 4067, Australia
E-mail: Karen.Aitken@csiro.au
Tel.: +61-7-32142360
Fax: +61-7-32142288

P. A. Jackson
Davies Laboratory, CSIRO Plant Industry,
University Drive, Townsville, QLD, 4814, Australia

[1]Denotes variety covered by Australian plant breeding rights.

Introduction

Modern sugarcane cultivars are derived from interspecific hybridisation between the domesticated species *Saccharum officinarum* L. ($2n = 80$) and the wild species *S. spontaneum* L. ($2n = 40$–128), both of which are polyploid. The cultivated hybrids have undergone a series of backcrosses to *S. officinarum* as the recurrent female parent in a process known as 'nobilisation' to recover agronomically adapted genotypes with a high sugar yield. The high sugar yield trait is recovered quickly because *S. officinarum* clones transmit their somatic chromosome number during the first interspecific cross and the first backcross (Bhat and Gill 1985). The nobilisation process represented a major breakthrough in sugarcane improvement, resulting in sugarcanes with increased resistance to diseases and additional benefits in yield increases and improved ratooning (Roach 1972). Several other species have been involved to a lesser extent in the development of modern sugarcane, including *S. robustum* Brandes and Jeswiet ex Grassl, which is thought to be the progenitor of *S. officinarum*. *S. barberi* Jesw. and *S. sinense* Roxb., which are thought to be natural hybrids of *S. spontanteum* and *S. officinarum* (Irvine 1999; D'Hont et al. 2002), have also contributed.

Most sugarcane cultivars have between 100 and 130 chromosomes, of which 15–25% are contributed by *S. spontaneum* (D'Hont et al. 1996). Using fluorescence in

situ hybridisation, D'Hont et al. (1998) showed that the basic chromosome number of *S. officinarum* is $x = 10$, while that of *S. spontaneum* is $x = 8$.

As sugarcane is a polyploid and highly heterozygous, restriction fragment length polymorphisms (RFLPs) and simple sequence repeats (SSRs) reveal several markers or alleles for a given locus. As these markers can be single dose (SD) (single copy present at a locus), double dose (two copies) or multi-dose, it is often not possible to identify all of the individual alleles at a locus. It is still unclear whether sugarcane is disomic (i.e. an allopolyploid with a preferential pairing of chromosomes within a homology group) or polysomic [i.e. an autopolyploid with random pairing of chromosomes within each homology group (HG)], or a combination of both (Grivet et al. 1996; Hoarau et al. 2001). To date, a number of linkage maps have been constructed in sugarcane using SD markers (Wu et al. 1992). SD markers in coupling are used to construct a map, and repulsion is used to determine chromosome assortment. The first map was developed on the wild relative and smaller genome of a *S. spontaneum*, SES208 ($2n = 64$), using random amplified polymorphic DNA (RAPD) markers (Al-Janabi et al. 1993) and RFLPs (da Silva et al. 1993, 1995). A RAPD map has been constructed for *S. officinarum* (Mudge et al. 1996), and four RFLP maps have been constructed for crosses between *S. officinarum* and *S. spontaneum* (Ming et al. 1998, 2002). Using comparative mapping to sorghum, Ming et al. (1998) were able to estimate that only 39.5–46% of the *Saccharum* genome was covered by each of these maps, although collectively they covered approximately 70% of the genome. An RFLP map has also been constructed on a selfed population of the cultivar R570 (Grivet et al. 1996). This map was composed of 408 markers, generated by 120 RFLP probes, which were placed onto 96 linkages groups (LGs). These LGs could be assembled into ten basic HGs. A more extensive map was constructed in this same cultivar with a different population. This map is composed of 887 AFLP markers distributed on 120 LGs (Hoarau et al. 2001). Using AFLP markers in common between the two maps and linkage in repulsion, 34 of these LG could be assigned to one of the ten HGs already defined in the previous RFLP map. Although the AFLP map was 5,849 cM long, it was estimated to cover approximately only a third of the genome. Rossi et al. (2003) added a further 236 resistance gene analogue (RGA) and SSR markers to a subset of the population used to construct the AFLP genetic map by Hoarau et al. (2001). These formed 128 cosegregation groups and were placed into seven HGs with 53 (41%) cosegregation groups still unassigned. SSRs in sugarcane are highly polymorphic, thereby facilitating the determination of HGs, which previously had depended on the use of the more time-consuming RFLP marker system.

The objective of the investigation reported here was to take advantage of the recent release of a large number of SSRs for *Saccharum* species to construct another high-coverage genetic map of a sugarcane cultivar. This genetic map, which has been constructed with both AFLPs and SSRs, has allowed a greater allocation of LGs to HGs than any other existing map. The use of SSRs has also enabled the first comparison of two sugarcane cultivar genetic maps that each incorporate over 1,000 markers.

Materials and methods

Plant material

The *Saccharum officinarum* clone IJ76-514 ($2n = 80$) was used as the female in a cross with Q165$^{\phi}$ ($2n \sim 115$), an Australian commercial variety. The cross was carried out in 1999, and 227 progeny were planted in a replicated field trial in Kalamia, Burdekin (19°32′S, 147°24′E) in 2000.

PCR assay for SSR, AFLP and randomly amplified DNA fingerprints

For DNA extraction, young leaves were sampled from 227 progeny in the field, freeze-dried, then ground to a powder and stored at −20°C. Genomic DNA was extracted following the CTAB method described by Hoisington (1992). SSR primers were obtained from the Sugarcane Microsatellite Consortium collection (Cordeiro et al. 2000). These were screened against the parents of this population to determine polymorphism levels, and the SSRs selected were scored across the population using the following protocol.

PCR reactions for the amplification of SSRs were carried out in a total volume of 20 µl containing 25 ng template DNA, 0.2 μM of forward and reverse primers, 0.2 m *M* of each dNTP, 3.0 m *M* $MgCl_2$ and 0.4 U Tth Plus *Taq* polymerase (Biotech International, Western Australia) in the buffer supplied. During PCR the SSR products were labeled with α-[^{33}P]dCTP (3,000 ci/mmol). Reactions were run on a Gene Amp PCR System 2700 thermocycler (PE Applied Biosystems, Foster City, Calif.) programmed with the cycling conditions: one cycle of 3 min at 94°C, 35 cycles of 1 min at 94°C, 2 min at the appropriate annealing temperature (ranging from 50°C to 56°C depending on the SSR) and 1 min at 72°C, with a final extension step of 5 min at 72°C. The amplified products were mixed with an equal volume of loading dye, denatured at 95°C for 5 min, and 3.6-µl samples run on a denaturing 5% polyacrylamide (20:1) gel at 90 W for 2 h. The gels were subsequently dried using a gel dryer for 25 min at 80°C and exposed to Kodak X-Omat X-ray film for 3–4 days.

AFLP analysis was performed as described by Vos et al. (1995) with some modifications. Genomic DNA (300–400 ng) was double digested with *Eco*RI and *Mse*I enzymes and ligated to the specific adapters. Primers with one selective nucleotide were used for pre-amplification.

The pre-amplification products were diluted 1:20 with water. PCR products were separated as described above, but exposure of the gel was on a high-resolution Kodak Biomax MR X-ray film for 3–4 days. Randomly amplified DNA fingerprints (RAFs) were carried out according to Bentley and Bassam (1996) with the modification that one of the primers was radiolabelled. The 227 progeny were characterised with five RAF primers, 40 AFLP primer combinations and 72 SSR primers.

Marker scoring and analysis

All segregating bands that were distinct and unambiguous were scored (1 for present and 0 for absent). Each marker system identified both monomorphic and polymorphic markers. Each marker was tested against the expected ratios using a χ^2 test for SD markers or duplex markers. SD markers are present only once in the genome, either in a 1:1 ratio (markers present once in one parental genome) or in a 3:1 ratio (marker present in both parents segregating as SD markers). Duplex markers are present twice in one parental genome, either in a 11:3 ratio (for $x = 8$) or in a 7:2 ratio (for $x = 10$) (da Silva et al. 1993). Polymorphic RAF markers were named using the primer number of the manufacturer followed by the number of polymorphic bands identified by the same primer in descending molecular-weight order. AFLP markers were identified by the primer combination, consisting of the three selective nucleotides in the *Eco*RI primer, followed by the three selective nucleotides in the *Mse*I primer, followed by number of polymorphic bands in descending molecular-weight order. SSR markers were labeled m followed by the name and identity number from the Sugarcane Microsatellite Consortium collection and a letter denoting the allele in descending molecular weight.

Map construction

Linkage analysis was carried out with SD markers in coupling phase using MAPMAKER—ver. 2.0 for the Macintosh or ver. 3.0 for a PC (Lander et al. 1987). The markers were grouped at a LOD of 5 and a θ of 0.25 using two-point analysis. In an effort to condense some of the fragmented LGs into single LGs, the LOD score was reduced to 4. With a population of 227 a recombination fraction of 0.43 for linkage in coupling phase can be detected at $P \leq 0.01$ (Wu et al. 1992). Linkage was only retained at this LOD score when markers from the same SSRs had already been identified as being linked on another LG in the HG at an LOD of 5. The markers were ordered within the groups using standard multipoint analyses procedures. The order was considered ambiguous when the most likely order was less than tenfold more likely than the second one, and this was highlighted on the map (Fig. 1). The Haldane function was used to derive map distances from the recombination fractions. The SD map was used as a framework map to assign both the double-dose markers as described in da Silva et al. (1995) and the markers present in both parents as SD markers and segregating in a 3:1 ratio. The double-dose markers would be present twice in the genome and could therefore also be used to determine HGs. All double-dose markers and those segregating in a 3:1 ratio were tested for linkage against each other and SD markers. LGs formed within each marker class were added to the map but double-dose markers were only assigned to the map in their most likely position. Where the markers segregating 3:1 or double-dose markers formed new LGs with SD markers, they were added to the map but not included in the calculation for map distance.

To investigate pairing between LGs, we analysed markers in repulsion. The segregation data matrix was doubled after inverting the scores for each marker. A two-point analysis was then carried out on the new matrix with a LOD = 5 and $\theta = 0.25$. Linkage was identified between the first set of data and the reverse phase data to identify pairs of markers in repulsion. The distribution of these markers was analysed, and repulsion was only accepted where large parts of the LG were involved, as artefactual repulsion is more likely to involve single pairs of markers (Wu et al. 1992; Qu and Hancock 2001).

Construction of HGs

A single LG is based only on markers in coupling and therefore will correspond to a single chromosome among all of the chromosomes within a HG. LGs were initially assembled into HGs if they contained two or more of the same SSRs. In addition, LGs were putatively added if they contained an SSR locus in common with the HG. LGs were also included in a HG if they were linked in repulsion with a LG already assigned to a HG.

Results

Comparison of marker systems

A total of 163 SSRs were screened against the parents of this population to determine polymorphism levels. Of these, 117 (72%) were polymorphic between the parents of this cross with at least one band difference. Across the population 72 SSRs were scored (Table 1), producing 350 polymorphic bands of which 238 (75.6%) were SD. The number of SD markers identified with each SSR ranged from 1 to 16 with an average of 3.3 (Table 1). The SSR markers mapped onto between 1 and 9 different LGs (Table 1).

The majority of the markers were generated using 40 AFLP primer combinations. These detected both monomorphic and polymorphic bands, including

Table 1 Simple sequence repeats (SSRs) screened across progeny of Q165[♂], polymorphism levels revealed, numbers of single- and double-dose markers, number of markers mapped and location on homo(eo)logy groups (HGs) and linkage groups (LGs)

SSR	Number of polymorphic bands					HG	LG
	Total	Single dose	Single dose (%)	3:1	Double dose		
mSSCIR1	16	6	37.5	0	5	2	12a, 6, 30, 31c,
mSSCIR8	6	5	83.3	0	1	1	17, 42, 105, 19, 20
mSSCIR9	4	3	75.0	0	1	3	86, 24, 65
mSSCIR10	4	1	75.0	2	1	3	23, 29
mSSCIR12	8	5	75.0	1	1	2	12a, 8, 58, 30, 107
mSSCIR14	4	3	75.0	0	0	4	56, 78, 29,
mSSCIR15	3	2	100.0	1	0	6	82, 71
mSSCIR16	3	2	66.7	0	0	7	60, 26
mSSCIR17	3	3	100.0	0	0	2	35, 12a, 8
mSSCIR18	3	0	33.3	1	1	–	–
mSSCIR19	3	2	66.7	0	0	4	78, 12b
mSSCIR21	6	3	50.0	0	1	1	95, 43, 50
mSMC21SA	7	5	85.7	1	1	4	3, 1, 36, 31, 44
mSSCIR23	4	2	75.0	1	0	5	33
mSSCIR24	3	1	100.0	2	0	3	67
mSSCIR25	8	6	87.5	1	0	1, 2, 3, 5	19, 53, 4, 69, 94, 55
mSSCIR26	2	1	50.0	0	1	2	12a
mSSCIR27	4	2	50.0	0	1	4	47, 75
mSSCIR28	8	7	100.0	1	0	2	16, 8, 58, 30, 22, 31c,
mSSCIR29	4	3	75.0	0	1	1	43, 50
mSSCIR30	2	1	50.0	0	1	4	80
mSSCIR32	4	4	100.0	0	0	3	93, 28, 63
mSSCIR33	5	2	80.0	2	0	2	35, 8, 50
mSSCIR34	5	4	80.0	0	1	5	106, 79, 57, 55
mSSCIR35	4	2	50.0	0	2	3	4, 40, 74
mSSCIR36	6	5	100.0	1	0	1	17, 42, 9, 19, 114, 27
mSC36BUQ	7	7	100.0	0	0	1	42, 81, 34, 47
mSSCIR37	3	2	66.7	0	1	3	72, 110
mSSCIR38	4	1	75.0	2	1	5	57, 44
mSSCIR39	4	4	100.0	0	0	5	106, 40, 76, 57
mSSCIR41	2	2	100.0	0	0	3	4, 74
mSSCIR42	5	3	80.0	1	1	3	16, 15, 41
mSSCIR43	9	6	77.8	1	1	1	17, 42, 105, 27, 73, 50, 20
mSSCIR44	2	2	100.0	0	0	2	35, 12a
mSSCIR45	2	0	0.0	0	1	3	4, 12
mSSCIR46	3	3	100.0	0	0	2	35, 12a, 53
mSSCIR50	5	1	40.0	1	2	4	12b
mSSCIR51	4	3	75.0	0	0	3	93, 77, 63
mSSCIR52	3	1	33.3	0	1	3	83, 21
mSSCIR54	7	4	57.1	0	1	2, 3	15, 12b, 87
mSSCIR55	4	4	100.0	0	0	6	62, 46
mSSCIR60	8	8	100.0	0	0	3	69, 11, 51, 65, 72, 45
mSMC63SA	5	4	100.0	1	0	3	4, 12
mSSCIR64	4	4	100.0	0	0	6	62, 46 122
mSSCIR66	3	2	100.0	1	0	1	42, 9
mSSCIR74	3	1	66.7	1	0	2	14, 39
mSSCIR75	3	3	100.0	0	0	3	4, 39, 7
mSMC75SA	8	7	87.5	0	0	3	4, 69, 67, 90, 101, 41
mSSCIR77	6	5	83.3	0	0	1	17, 105, 43, 50, 20
mSMC119SA	1	0	100.0	1	0	3	21
mSMC179SA	6	6	100.0	0	0	1	17, 42, 9, 19, 50, 114
mSC286CS	6	6	100.0	0	0	3	4, 52
mSMC292MS	10	10	100.0	0	0	1–7	17, 50, 13, 94, 7, 44, 79, 71, 32
mSC334BS	2	0	50.0	1	0	–	–
mSC336BS	6	4	66.7	0	2	2	87, 13, 30, 31c
mSC597CS	4	3	100.0	1	0	3	4, 94, 74
mSC851MS	5	5	100.0	0	0	2	12a, 8, 6, 31c, 107
mSMC1047HA	4	4	100.0	0	0	4	44, 36, 31, 13
mSMC1608CL	6	4	100.0	2	0	3	65, 24, 108
mSMC1120HA	7	4	85.7	2	1	4	3, 48, 56, 111, 51
mSMC1149HA	5	4	80.0	0	1	4, 5	31, 55, 33 115
mSMC1218LA	2	0	0.0	0	2	2	39
mSMC1232LA	4	1	25.0	0	3	4	64
mSMC1237FL	5	3	60.0	0	2	4	3, 79, 25,

Table 1 (Contd.)

SSR	Number of polymorphic bands					HG	LG
	Total	Single dose	Single dose (%)	3:1	Double dose		
mSMC1481CL	4	3	100.0	1	0	3	69, 21, 65, 47
mSMC1493CL	5	5	100.0	0	0	4	3, 25, 80, 64, 73,
mSMC1527CL	5	0	60.0	3	2	8	28, 71, 55
mSMC1572CL	4	2	50.0	0	2	2	30, 16
mSMC1814LA	8	5	75.0	1	2	2	6, 38, 85, 30
mSMC1825LA	9	6	88.9	2	1	2	35, 12a, 8, 6, 53, 57,
mSMC2039FL	2	2	100.0	0	0	4	107, 64
mSMC2055FL	5	4	80.0	0	1	3	4, 94, 28, 63,
Total	350	238	–	36	47	8	101
Mean	4.9	3.3	75.6	0.5	0.7	1.1	3.2
Range	1–16	0–10	–	0–3	0–5	1–7	0–9

multi-dose markers. In total these primers generated 972 polymorphic bands, including 704 (72%) SD markers (Table 2). The number of polymorphic bands per primer pair varied from 13 to 39 with an average of 24.3 (Table 2). SD markers generated by individual AFLP primer pairs mapped to between 7 and 22 different LGs (Table 2). Five RAF primers generated 43 polymorphic bands, of which 26 (60%) were SD. This system gave lower numbers of markers than AFLPs and the markers detected were less consistent. For this reason, efforts were concentrated on AFLP and SSR markers. Both AFLP and SSR primers detected the same proportion of SD markers, 72% and 75.6% respectively. They also detected similar numbers of double-dose markers, 16.4% and 13.4%, respectively (Tables 1, 2). AFLPs were used to generate the majority of markers as a single gel could produce up to 39 segregating markers. In contrast, the maximum number of markers detected with an SSR was 16. Although the products of up to three SSR primer pairs could be run on one gel, this could only be achieved if the products were of a sufficient size difference, which was complicated by the large number of alleles generated per primer pair in sugarcane. The advantage of SSR over AFLP is the ability to determine HGs.

Map construction

Using 40 AFLP primer combinations, 72 SSR primers and five RAF primers, we scored 1,365 polymorphic markers on this population. Of these, 967 (71%) segregated as SD markers (1:1). As a large number of markers were analysed with an expectation of over 100 LGs, a stringent LOD score of 5 was used to avoid false linkage. Nine hundred and ten SD markers formed 116 LGs (Fig. 1) with 57 unlinked markers. The SD LG varied in length from 0.8 cM to 297.3 cM, with a cumulative length of 9,058.3 cM. The average distance between loci was 9.9 cM, but distribution along the chromosomes was irregular. There were both densely mapped regions and a number of gaps in the map that were larger than 40 cM. However, with a population size of 227, linkage at this distance is still significant at $P \leq 0.01$ (Wu et al. 1992). The large population size also ensured that ambiguous marker orders were strictly limited to those regions on the map where linkage was tight.

Assignment of HGs using SSR markers

Of the 116 SD LGs, 96 formed into HGs using the SSR markers (Fig. 1). In general, the order of SSRs was consistent between LGs found within a HG. Order change did occur with closely linked markers, but a larger population would be needed for precise positioning within these small distances. In three cases marker order was reversed between LGs. In HG 2, markers mSC851MS and mSSCIR12 on LG 12a are reversed on LG 57. In HG 3, the order of mSSCIR60 and mSMC1481CL on LG 69 is reversed compared to that found on LG 65, and on LG 4 the marker order of mSMC2055FL and mSSCIR32 is reversed compared to that found on LG 63. These cases all involve markers at the end of a LG, which are always more difficult to order due to less co-informative data, as reported in other species (Gardiner et al. 1993). Ten of the SSRs detected duplication within a LG (mSSIR1, mSSIR10, mSSIR23, mSMC1120HA, mSMC1527CL, mSMC1825LA, mSSCIR55, mSSCIR60, mSMC63SA and mSMC75SA). These were detected on HGs 2, 3, 4, 6 and 8, with four of the ten located on HG 3. In all but four cases the SSR markers mapped only to one HG. Markers generated by SSR mSMC292MS mapped to all but one HG, while markers generated by mSSCIR25 mapped to four HGs and mSSCIR54 and mSMC1149HA to two HGs each (Table 1, Fig. 1).

Addition of double-dose and 3:1 markers

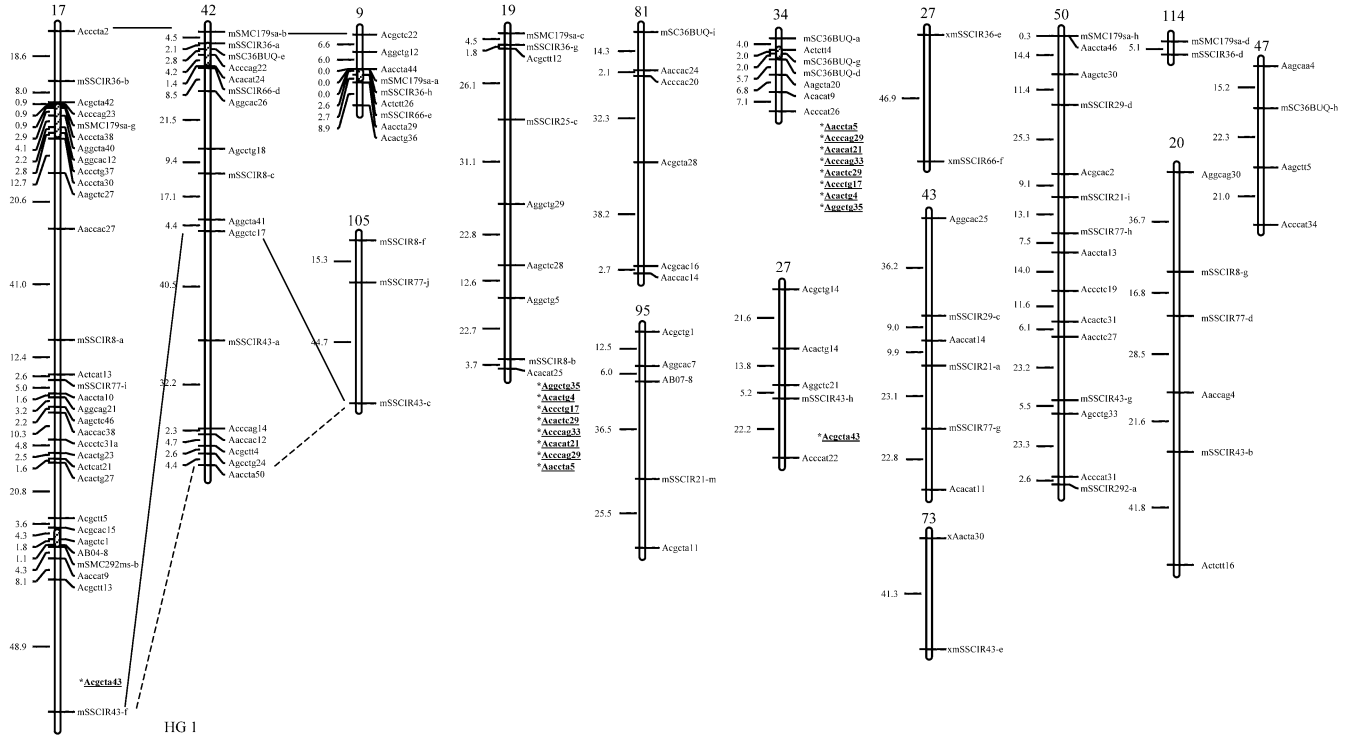
The addition of 77 double-dose markers (123 markers, as 31 mapped to one LG only) and 36 3:1 segregating

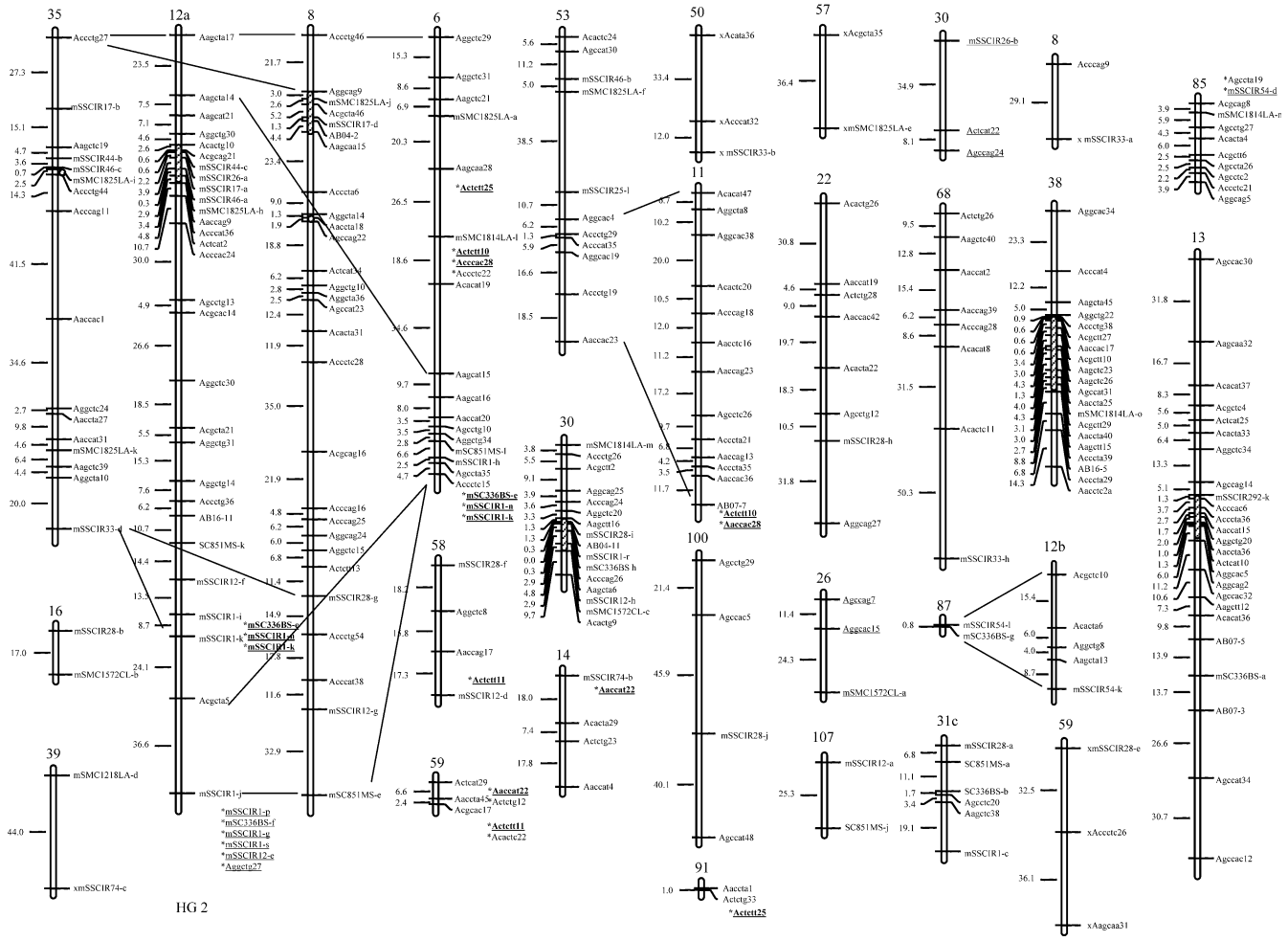
Table 2 Amplified fragment length polymorphism (AFLP) primer pairs screened across progeny of Q165$^{\phi}$, polymorphism revealed and numbers of single- and double-dose markers

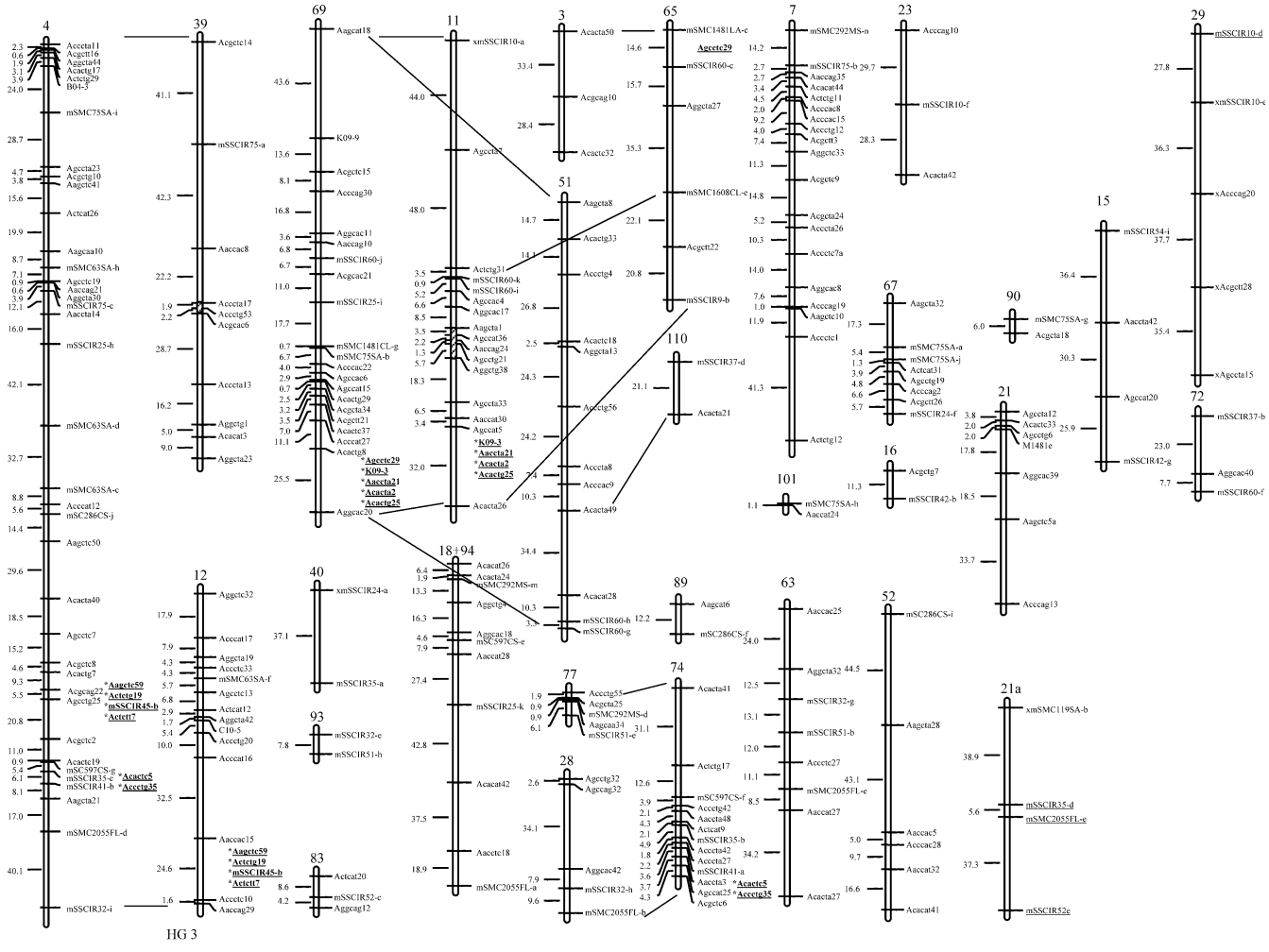
Primer pair *Eco*RI/*Mse*I	Number of polymorphic bands					Number of linkage groups
	Total	Single dose	Single dose (%)	Double dose	3:1	
Aaccac	29	21	72.4	4	2	19
Aaccag	25	18	72.0	4	2	14
Aaccat	27	18	66.7	7	5	20
Aaccta	33	24	72.7	7	5	21
Aacctc	15	9	60.0	2	4	8
Aagcaa	16	12	75.0	1	5	12
Aagcat	17	11	64.7	3	3	7
Aagcta	26	17	65.4	5	10	14
Aagctc	39	27	69.2	7	4	20
Aagctt	16	10	62.5	4	5	10
Acacat	33	28	84.8	2	2	22
Acacta	30	24	80.0	3	5	21
Acactc	28	16	57.1	6	2	13
Acactg	30	19	63.3	6	1	16
Acccac	18	12	66.7	4	1	11
Acccag	30	25	83.3	3	2	20
Acccat	26	20	76.9	4	4	17
Accta	32	27	84.4	3	1	19
Accctc	32	23	71.9	6	5	17
Accctg	33	25	75.8	3	4	21
Acgcac	17	13	76.5	3	1	11
Acgcag	20	15	75.0	2	0	11
Acgcta	28	24	85.7	3	7	20
Acgctc	13	12	92.3	1	1	10
Acgctg	15	14	93.3	1	2	10
Acgctt	22	17	77.3	2	2	13
Actcat	27	19	70.4	5	0	16
Actctg	16	10	62.5	4	7	9
Actctt	17	10	58.8	7	1	10
Agccac	18	12	66.7	4	4	11
Agccag	22	11	50.0	7	3	10
Agccat	24	19	79.2	2	3	17
Agccta	27	18	66.7	6	5	12
Agcctc	17	11	64.7	5	1	10
Agcctg	22	16	72.7	4	2	17
Aggcac	34	20	58.8	6	2	17
Aggcag	16	12	75.0	3	1	10
Aggcta	35	25	71.4	6	1	19
Aggctc	24	21	87.5	2	5	17
Aggctg	23	19	82.6	2	2	18
Total	972	704	–	159	122	–
Mean	24.3	17.6	72.4	4.0	3.05	14.8
Range	13–39	9–28	–	1–7	1–10	7–22

markers enabled another 20 LGs to be identified (Fig. 1). These consisted of seven LGs that contained both double-dose and 3:1 segregating markers, three that contained only double-dose markers and ten that consisted of only 3:1 markers or 3:1 and SD markers. Double-dose markers were also used to determine HGs as they mapped to two LGs within the same HG. Of the 77 double-dose markers assigned to the map, 46 were located on two different LGs. These were used to confirm the location of 37 LGs and added a further eight LGs to the HGs (Fig. 1). The addition of the double-dose and 3:1 segregating markers revealed another HG (HG 8), which contained no SD markers. Finally, the detection of repulsion between LGs was also used to assign LGs to a HG. Using this method LG 3 and LG 113 were assigned to HG 3. Two small HGs were combined into HG 6 as LG 71 and LG 62 are in repulsion, and LG 32 was added to HG 7. In total, there

Fig. 1 A genetic linkage map of sugarcane cv. Q165$^{\phi}$ generated from 227 individuals. The *numbers* on the *left* are the genetic distances in Haldane centiMorgans (cM) between markers. Marker names are on the *right*: double-dose markers start with a *star* and are *underlined* and placed in their most likely position; 3:1 segregating markers have an *x* at the *beginning* of the name. Double-dose markers that were used to confirm or build HGs are in *bold*. Uncertain orders (alternative orders not ruled out at LOD = 1) are represented by *hatches*. The LGs are placed in their HGs using both SSRs and double-dose markers. Linkage in repulsion between linkage groups is represented by *lines*. All markers within the *two solid associated lines* were in repulsion with all other markers on the alternative linkage groups at a LOD≥5. The *dashed lines* represent repulsion at a LOD≥3.


17
42
9
19
81
34
27
50
114
47
105
95
43
20
73
HG 1


35
12a
8
6
53
50
57
30
8
85
22
68
38
11
13
16
58
30
14
100
26
87
12b
107
31c
59
39
91
HG 2



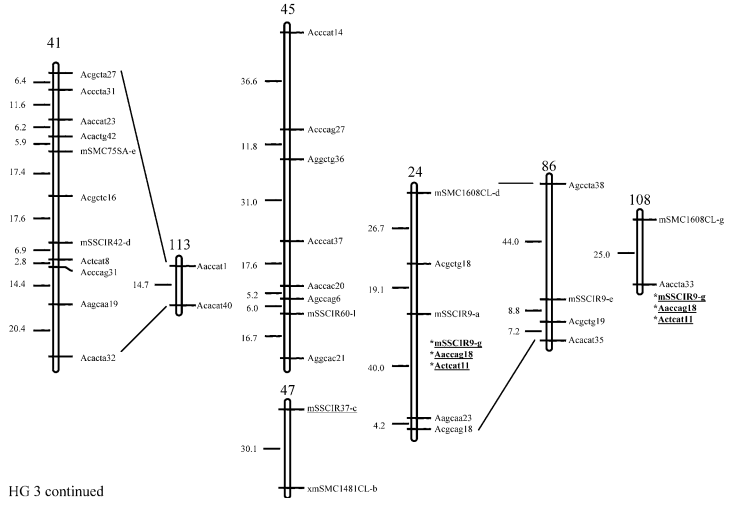


Fig. 1 (Contd.)

were 136 LGs formed with 1,074 markers, of which 915 were SD markers, 123 were double-dose markers and 36 were 3:1 segregating markers. These LGs formed eight HGs with nine LGs unassigned (Fig. 1).

Analysis of HGs incorporating all markers

Of the 136 LGs, 127 formed eight HGs, each containing between 3 and 37 LGs. Markers were not evenly

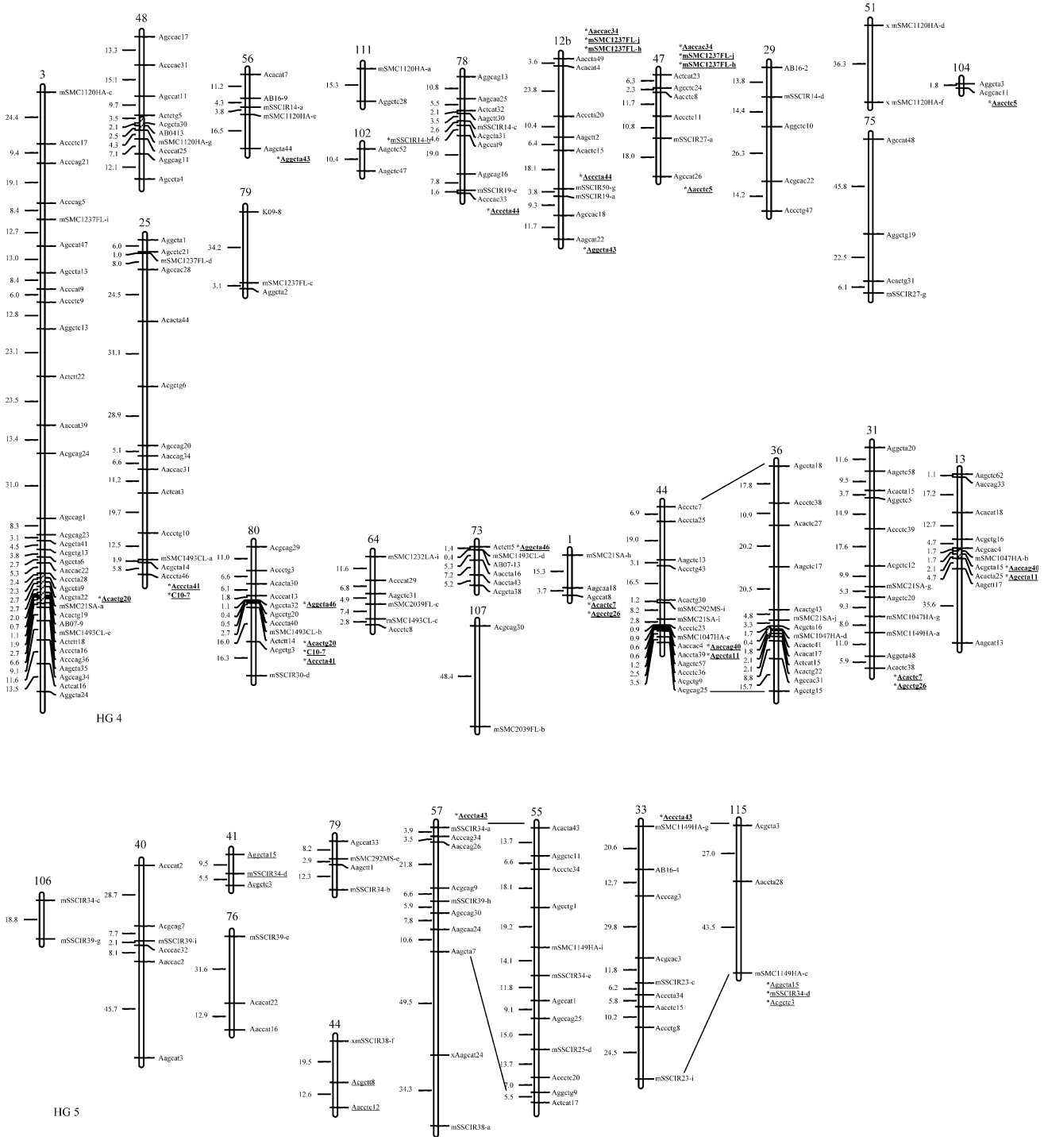


Fig. 1 (Contd.)

distributed over the different HGs (Fig. 1), with the largest HG containing 301 markers and the smallest containing 11. For HG 4, two larger LGs had homology to two separate groups of smaller LGs. The SSRs demonstrated a consistent marker order within each subgroup. (The top group contained mSMC1120HA, mSSCIR14, mSSCIR50, mSSCIR19, mSMC1237FL, and the lower group contained mSMC21SA, mSMC1493CL, mSMC1047HA) (Fig. 1, HG 4). The double-dose markers located in this group were always located to two LGs within a subgroup and not across the subgroups. Decreasing the LOD score to 2 did not result in any of these groups coalescing into a single LG. A similar situation exists for HG 3 although it is not as clearly defined as HG 4. Each of the subgroups has its own conserved group of SSRs across the LG. The

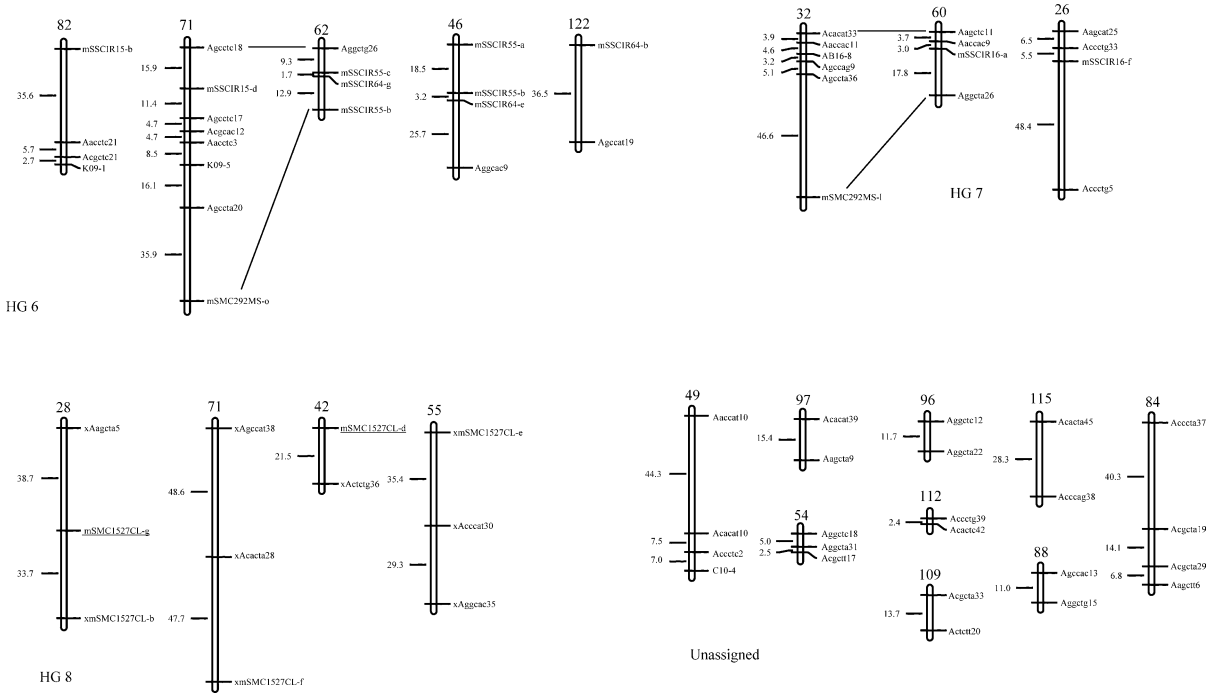


Fig. 1 (Contd.)

double-dose markers also reflect the subgroups and are located on two LGs in one or the other subgroup but not across them (Fig. 1). Interestingly, the SSR mSMC292MS, which are located on all HGs with the exception of HG 8, and mSSCIR25, located on four other HGs, are on both subgroups in different positions.

Double-dose markers were not evenly distributed on the linkage map, with HG 5 having two markers and HG 3 having 33 markers (Fig. 1). This could be a reflection of map coverage with some groups having much higher genome coverage than others. Some groups had large regions that contained just double-dose markers. HG 1 has a group of 8 double-dose markers that are linked to LG 19 and LG 34 that span 91.2 cM (data not shown).

Chromosome pairing

Linkage was examined between all markers in repulsion and all other markers in the map. A total of 40 LGs were linked in repulsion with other LGs at a LOD≥5 (Fig. 1). The high LOD score was used to minimise spurious linkage. Repulsion was detected in every HG, with the exception of HG 8, which is a consequence of it consisting of only 3:1 and double-dose markers. In all cases, the repulsion was detected over the majority of the LGs involved (Fig. 1). In HG 1, repulsion was detected in four LGs, which formed a multi-chromosome pairing group. LG 42 was in repulsion over the whole of its length with LG 17. Repulsion was also detected between LG 42 and LG 9 and LG 105, which are probably another pair as LG 9 and LG 105 are probably parts of a single LG (Fig. 1). In HG 2, four LG formed two pairs, LG 12b with LG 87 and LG 11 with LG 53. Another four LGs, LG 8, LG 6, LG 12a and LG 35, all paired with each other, forming another multi-chromosome pairing group. The strongest repulsion was between LG 12a and LG 8 and the weakest between LG 35 and LG 6 (Fig. 1). In HG 3, repulsion was detected between 16 LGs, ten of which formed four pairs, LG 113 paired with LG 41; LG 74 paired with LG 77 and LG 28, which are probably part of the same LG; LG 4 paired with LG 12 and LG 39, which again could be part of the same LG; LG 24 paired with LG 86 (Fig. 1). Repulsion was detected between another six LGs, which formed a multi-chromosome pairing group; these included between LG 3 and LG 51, which could possibly be part of one LG and repulsion to LG 65 and LG 110, which again could be part of the same LG. LG 51 also paired with LG 69, which itself paired with LG 11. Repulsion was also detected between LG 11 and LG 65. Interestingly, repulsion was detected between LGs within the subgroup of HG 3 and not between the subgroups (Fig. 1). In HG 4, lower levels of repulsion were detected. LG 44 and LG 36 were in repulsion and, interestingly, strong repulsion was detected between LG 13 in HG 2 and LG 13 and LG 12b in HG 4. Again repulsion was detected between LGs in the subgroups and not between subgroups (Fig. 1). In HG 5, repulsion was detected between four LGs; these formed two pairs—LG 57 and LG 55, and LG 33 and LG 115 (Fig. 1). In HG 6, repulsion was detected between LG 71 and LG 62, and in HG 7, between LG 60 and LG 32 (Fig. 1).

Discussion

Linkage map

The map reported here consists of 915 SD markers, leaving only 5.4% SD markers unlinked. The map length using these markers is 9058.3 cM. With the inclusion of double-dose and 3:1 segregating markers, 1,074 markers form 136 LGs, of which 127 (93%) can be placed into eight HGs. Even with this number of markers, the map is obviously not saturated, as there are still 52 unlinked SD markers, 27 double-dose markers that map to only one LG and an obvious difference in coverage of the HGs (Fig. 1). One way to estimate the total genome length would be to take the mean length of the longest chromosome maps, which are assumed to be saturated, and multiply these by the number of chromosomes (Hoarau et al. 2001). However, due to the digenic composition of the sugarcane cultivar genome, the longest LGs in each HG are likely to be those of *S. spontaneum*, which contribute only a small proportion of the genome. Previous studies have shown that *S. spontaneum* LGs are generally better covered by markers due to their more polymorphic nature (D'Hont et al. 1994; Grivet et al. 1996). If these LGs are excluded and the calculation based on the longest remaining LGs, the average is 150 cM (150×120 = 18,000 cM). This is only a rough estimation and is based on a polyploid plant with at least partial polysomy. Using this estimate, the present map probably covers around one-half of the genome. We have no information on the origin of markers and therefore cannot determine which chromosomes are *S. officinarum* in origin and which are *S. spontaneum*. However, the structure of the linkage map allows some comparison to other studies. HGs 1 to 4 all contain one or two LGs that are longer than the rest and have more markers, which probably correspond to the *S. spontaneum* fraction of the genome. In this study, the *S. officinarum* parent had much lower levels of polymorphism than the cultivar parent (data not shown). The majority of the LGs in this map are the smaller, probably *S. officinarum* chromosomes, which is in agreement with other mapping studies (Grivet et al. 1996; Hoarau et al. 2001) and is related to the lower diversity seen in this species (Lu et al. 1994) and the higher redundancy of the *S. officinarum* part of the genome in modern cultivars.

Homo(eo)logy groups

There is a marked difference in coverage of LGs across the HGs. HGs 1 to 4 have at least ten LGs per HG, while HG 5 has eight LGs. HGs 6 to 8 are poorly covered with markers and have only three to five LGs (Fig. 1). This lower number of LGs could also be due to aneuploidy, (i.e. there may be an unequal number of chromosomes in each HG). The lack of markers in these groups is probably due to low levels of polymorphism in this region of the *S. officinarum* part of the genome, as has been seen in the other linkage maps produced for *Saccharum* species (Grivet et al. 1996; Ming et al. 1998; Hoarau et al. 2001). Ming et al. (1998) report that in regions corresponding to sorghum LGs A, B, C, D, F and G, *Saccharum* showed polymorphism levels of 71–89%, whereas regions corresponding to LGs E, H, I and J had polymorphism levels of 42–47%. The reason for this difference could be that fewer SD markers are available for mapping in certain parts of the *S. officinarum* genome. This could be a result of the 2*n* gamete transmission during nobilisation and high inbreeding due to only small numbers of *S. officinarum* lines being used in the ancestry of sugarcane. An indication of this is that HG 8 is only composed of markers that are present in both parents but segregating as SD or double-dose markers. da Silva et al. (1995) also identified a HG composed of only double-dose markers in their *S. spontanteum* map, indicating that the duplication of regions of chromosomes occurs in both the *S. spontaneum* and *S. officinarum* genomes.

Modern cultivars like Q165$^{\phi}$ are hybrids between *S. officinarum* ($x = 10$) and *S. spontaneum* ($x = 8$) (D'Hont et al. 1998) and contain 100–130 chromosomes, of which 15–25% are contributed by *S. spontaneum* (D'Hont et al. 1996). Previous linkage maps of sugarcane cultivars have identified ten (Grivet et al. 1996), nine, 13 (Ming et al. 2002), eight (Hoarau et al. 2001) and seven (Rossi et al. 2003) HGs. The Q165$^{\phi}$ map has a similar number of markers as the map of cv. R570 (Rossi et al. 2003). In the Q165$^{\phi}$ map, eight HGs were formed. However, HGs 3 and 4 both appear to contain *S. spontaneum* chromosomes that have homology to two separate sets of smaller chromosomes that may derive from *S. officinarum*, thereby giving ten homology groups in *S. officinarum*. This observation is consistent with the bi-specific origin of sugarcane and with the different basic chromosome numbers of the ancestral species (Grivet and Arruda 2001).

Addition of double-dose and 3:1 segregating markers

Double-dose markers were used to confirm HGs, and using these and the 3:1 segregating markers, we were able to add another 18 LGs to the map. This method was used by da Silva et al. (1995) to determine HGs with similar results. Linkage analysis within the double-dose markers identified 11 LGs. These11 LGs mapped to four of the HGs and demonstrate that large sections of the chromosomes are duplicated in sugarcane (Fig. 1). For this reason any maps constructed for sugarcane cultivars would benefit from incorporating multi-dose markers to give high genome coverage. Six of the HGs contained double-dose markers, but this inclusion is probably related to the higher coverage of these HGs rather than the higher levels of duplication in some HGs than in others.

Chromosome pairing

In this study 40 LGs were linked in repulsion with each other. In all cases the LOD scores were high. These formed into 11 pairs and three multi-chromosome pairing groups (Fig. 1). The pairs were detected in all HGs except for HG 8, but the multi-chromosome pairing groups were only in HGs 1, 2, and 3. This could be due to the higher coverage of markers in these groups. Interestingly, HG 3 is one of the HGs that have single *S. spontaneum* chromosomes with homology to two sets of *S. officinarum* chromosomes. In comparison to the results of Hoarau et al. (2001), this study detected more repulsion phase linkage and also detected multi-chromosome pairing groups; this could be due to higher genome coverage in the present map. The present study also detected repulsion between LGs in two HGs, HGs 2 and 4. This could be an example of a translocation, which is common in polyploids. For example, the maize genome is a tetraploid, but no completely homologous chromosomes have been identified, implying that considerable reorganisation has taken place since polyploidy first occurred (Moore et al. 1995). Although the detection of all repulsion phase linkage is precluded by the lack of whole genome coverage, this study indicates that sugarcane has at least a partial preferential pairing of chromosomes within a HG and that recombination between HGs probably occurs. Mudge et al. (1996) also detected repulsion phase linkage in *S. officinarum*, but no repulsion phase linkage was detected in the *S. spontaneum* map of SES208 (da Silva et al. 1995).

Comparison of two sugarcane cultivar genetic maps

The use of SSRs also allows the existing R570 linkage map (Rossi et al. 2003), the largest sugarcane map published to date, and the only other map constructed in a sugarcane variety, to be compared with the map generated in this study. The map generated in this study has 49 fewer markers than the R570 map but is 1,213 cM longer. This could be because a number of the RGAs on the R570 map map to the same place as each other (Rossi et al. 2003). Thirty-four common SSRs were used in the construction of the two maps. Of the seven co-segregation groups (CGs) identified in the R570 map, six can be tentatively aligned with five of the eight HGs identified in the Q165$^{\Phi}$ map on the basis of shared common SSRs. However, there are a number of examples of SSRs mapping to different CGs or HGs in the two maps. For example, CG I on the R570 map has five common SSRs, three of which map to HG 4 and two to HG 3 in the Q165$^{\Phi}$ map. CG II has six common SSRs, two of which map to HG 3, two to HG 5, one to HG 4 and one to HG 1 in the Q165$^{\Phi}$ map. CG III has four common SSRs, two map to HG 6 and one each to HG 3 and HG 2 in the Q165$^{\Phi}$ map. CG IV has one SSR in common with the Q165$^{\Phi}$ map, and it maps to HG 2. CG VI has five common SSRs, all of which map to HG 3 in the Q165$^{\Phi}$ map. CG VII has three common SSRs, all of which map to HG 1, and CG VIII has nine common SSRs, seven of which map to HG 2, one to HG 5 and one to HG 3 in the Q165$^{\Phi}$ map. As both maps identified SSRs that mapped to more than one HG, it could be that in some cases the loci that map to different HGs are the result of different polymorphic markers being detected with the same SSR in the two populations. Another explanation could be that translocation occurs in the development of sugarcane cultivars. Cuadrado et al. (2004) carried out in situ hybridisation on three sugarcane cultivars, and subsequently were able to identify numerous examples of chromosome remodelling, including different numbers of recombinant chromosomes in the three cultivars. The rearrangement in the map positions of SSRs in the two maps of R570 and Q165$^{\Phi}$ could also be a result of remodelling between HGs. The repulsion detected in this study between HGs 2 and 4 also provides evidence of possible recombination between HGs. Although a larger number of common markers between the two maps are needed to verify the results found in this study, our results suggest that genetic maps produced for different cultivars will reveal different chromosome arrangements, which is likely to impact on the use of markers for sugarcane breeding.

To obtain sufficient genome coverage for QTL analysis in sugarcane, breeders require large numbers of markers. The extensive genome coverage of this map will facilitate the detection of QTL for traits of interest in sugarcane.

Acknowledgements This project was supported in part by a research grant from the Australian Sugar Research and Development Corporation.

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

Bentley S, Bassam BJ (1996) A robust DNA amplification fingerprinting system applied to analysis of genetic variation within *Fusarium oxysporum* f.sp. *cubense*. J Phytopathol 144:207–213

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

Cordeiro GM, Taylor GO, Henry RJ (2000) Characterisation of microsatellite markers from sugarcane (*Saccharum* sp.), a highly polymorphic species. Plant Sci 155:161–168

Cuadrado A, Acevedo R, Moreno Díaz de la Espina S, Jouve N, de la Torre C (2004) Genome remodelling in three modern *S. officinarum* × *S. spontaneum* sugarcane cultivars. J Exp Bot 55:847–854

D'Hont A, Lu YH, Gonzālez De Lēon D, Grivet L, Feldmann P, Lanaud C, Glaszmann JC (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, Glaszmann J-C (1998) Determination of basic chromosome numbers in the genus *Saccharum* by physical mapping of ribosomal RNA genes. Genome 41:221–225

D'Hont A, Paulet F, Glaszmann J-C (2002) Oligoclonal interspecific origin of 'North Indian' and 'Chinese' sugarcanes. Chromosome Res 10:253–262

Gardiner JR, Coe EH, Melia-Hancock S, Hoisington DA, Chao S (1993) Development of a core RFLP map in maize using immortalized F_2 population. Genetics 134:917–930

Grivet L, Arruda P (2001) Sugarcane genomics: depicting the complex genome of an important tropical crop. Curr Opin Plant Biol 5:122–127

Grivet L, D'Hont A, Roques D, Feldmann P, Lanaud C, Glaszmann JC (1996) RFLP mapping in cultivated sugarcane (*Saccharum* spp.): genome organisation in a highly polyploid and aneuploid interspecific hybrid. Genetics 142:987–1000

Hoarau JY, Offmann B, D'Hont A, Risterucci AM, Roques D, Glaszmann JC, Grivet L (2001) Genetic dissection of a modern sugarcane cultivar (*Saccharum* spp.). 1. Genome mapping with AFLP markers. Theor Appl Genet 103:84–97

Hoisington D (1992) Laboratory protocols. CIMMYT Applied Molecular Genetics Laboratory, Mexico City

Irvine JE (1999) *Saccharum* species as horticultural classes. Theor Appl Genet 98:186–194

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

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

Ming R, Liu SC, Lin YR, 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 organisation of closely related diploid and polyploid genomes. Genetics 150:1663–1682

Ming R, Liu SC, Bowers JE, Moore PH, Irvine JE, Paterson AH (2002) Construction of a *Saccharum* consensus genetic map from two interspecific crosses. Crop Sci 42:570–583

Moore G, Devos KM, Wang Z, Gale MD (1995) Cereal genome evolution, grasses, line up and form a circle. Curr Biol 5:737–739

Mudge J, Andersen WR, Kehrer RL, Fairbanks DJ (1996) A RAPD genetic map of *Saccharum officinarum*. Crop Sci 36:1362–1366

Qu L, Hancock JF (2001) Detecting and mapping repulsion-phase linkage in polyploids with polysomic inheritance. Theor Appl Genet 103:136–143

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

Rossi M, Araujo PG, Paulet F, Garsmeur O, Dias VM, Chen H, Van Sluys M-A, D'Hont A (2003) Genomic distribution and characterization of EST-derived resistance gene analogs (RGAs) in sugarcane. Mol Genet Genomics 269:406–419

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

da Silva JAG, 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

Vos P, Hogers R, Bleeker M, Reijams 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, Sorrells 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