Theor Appl Genet (2006) 112: 1306–1317
DOI 10.1007/s00122-006-0233-2

ORIGINAL PAPER

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

Quantitative trait loci identified for sugar related traits in a sugarcane (*Saccharum* spp.) cultivar × *Saccharum officinarum* population

Received: 26 February 2005 / Accepted: 29 January 2006 / Published online: 1 March 2006

Abstract The identification of markers linked to quantitative trait loci (QTLs) for increased sugar accumulation could improve the effectiveness of current breeding strategies in sugarcane. Progeny from a cross between a high sucrose producing cultivar, Q165$^{\phi}$ (denotes Australian plant breeding rights), and a *Saccharum officinarum* clone, IJ76-514 were grown in two field experiments in different years, and evaluated in the early and mid-season phases of crop maturity, to identify robust QTLs in Q165$^{\phi}$ affecting sucrose content in cane. Using an extensive genetic map constructed for Q165$^{\phi}$ with over 1,000 AFLP and SSR markers, a total of 37 QTLs were identified for brix and pol of which, 16 were detected in both experiments. Of these 37 QTL, 30 were clustered into 12 genomic regions in six of the eight homo(eo)logous groups. Each QTL explained from 3 to 9% of the phenotypic variation observed. Both positive and negative effects were identified and the location of the QTLs on linkage groups belonging to the same homo(eo)logy group suggested that a number of the QTLs were allelic forms of the same genes. Of the 37 QTLs identified, the majority were significant in both early and mature cane, but 8 were identified as early specific QTLs and 9 as mature cane QTLs. In total, 97 interactions were significant ($P < 10^{-5}$) and these were localised to 32 genomic regions of which 6 were detected with both years' data. Models including all the QTLs explained from 37 to 66% of the total phenotypic variation, depending on the trait. The results will be subsequently applied in marker assisted breeding.

Q165$^{\phi}$ denotes variety covered by Australian plant breeding rights.

Communicated by E. Guiderdoni

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

Introduction

Sugarcane cultivars are polyploids, aneuploids and outbreeding. An individual cultivar is highly heterozygous and clonally propagated. Modern cultivars are derivatives from the high sucrose containing *Saccharum officinarum* ($2n = 80$) and the hardier but low sugar content wild species *Saccharum spontaneum* ($2n = 40$–128). These cultivars and parents in commercial breeding programs trace back to a relatively small number of key ancestors used in the first interspecific hybridisations in the early 1900s (Roach 1984). The initial interspecific hybrids were subjected to a series of backcrosses to *S. officinarum* or to other hybrids to recover the high sugar producing type. This process of introgression is known as nobilization and provided a major breakthrough in sugarcane breeding, improving yields, ratooning ability and increasing adaptability to stress, compared with prior sugarcane breeding based on the use of only *S. officinarum* (Roach 1989). The result of this breeding history is that modern sugarcane cultivars generally contain between 100 and 120 chromosomes, of which 15 to 25% are contributed by *S. spontaneum* (D'Hont et al. 1996).

A major objective of sugarcane variety improvement programs is obtaining higher sugar yields. Sugar yield is the product of cane yield and sugar content in the harvested material. Improvement of sugar content in sugarcane is highly attractive from a commercial perspective since it increases revenue from increased sugar production without increased harvesting, cane transport and milling costs that are associated with higher cane yields. A high priority is therefore placed on improving sugar content in sugarcane breeding programs. However despite this, and the high heritability of sugar content, there has been limited progress in improving sugar content in sugarcane breeding programs over the last 40 years (Jackson 2005).

A number of sugarcane molecular marker maps have been developed (Ming et al. 1998, 2002a; Hoarau et al. 2001; Aitken et al. 2005) but even the most complete of these is estimated to still only cover about 60% of the genome. These maps have been used for preliminary quantitative trait loci (QTL) analysis for a number of traits related to sucrose content (Ming et al. 2001, 2002b; Hoarau et al. 2002) and found many QTLs of small effect contributing to sugar content. The QTL analysis undertaken by Hoarau et al. (2002) used a population derived from selfing of the French sugarcane variety, R570. By contrast, Ming et al. (2001, 2002b) reportedly used populations from crosses between *S. officinarum* and *S. spontaneum*, although the chromosome counts (>80) in the "*S. officinarum*" parents would suggest they could be interspecific hybrids.

Effects of genes controlling physiological processes, such as sucrose accumulation, are expected to be highly context dependent, interacting with the background genotype and physiology (Peccoud et al. 2004). Therefore QTL studies are arguably best done directly within the genetic population(s) in which subsequent marker assisted breeding is planned (Podlich et al. 2004). Past reports on QTL analysis of sucrose content in sugarcane have used clones derived from a *S. spontaneum* parent (Ming et al. 2001, 2002b) or arising from self pollination of a parent (Hoarau et al. 2002). While such studies provide useful information they do differ to those normally used to generate cultivars in commercial breeding programs. In this study we used a population derived from an elite commercial parent and a *S. officinarum* clone. The performance of the progeny (sugar content and cane yield) were similar to those generated in commercial breeding programs. In this report, we focus on characterising the distribution of QTL controlling sucrose accumulation in the elite parent genome.

The level of sucrose accumulation at the start of the harvesting season (normally at some time in autumn or early winter in most industries) is relatively low in most sugar industries. There would be large commercial benefits in cultivars with higher sucrose when harvesting currently begins or before. This could improve industry profitability early in the harvesting season, and could facilitate longer harvesting seasons, and improved use of industry harvesting and milling capital, in future. Cox et al. (1990) found that in their experiment on a range of commercial and experimental clones, there was greater genetic variation for sucrose content in early months compared to late months. On the other hand, several studies have shown high genetic correlations among unselected sugarcane clones between sugar content very early in the season with mid seasonal measurements (Jackson and Morgan 2003), suggesting common genes controlling sucrose content across contrasting times.

We report the identification of QTLs for sugar related traits measured over two independent field trials in early and mid season. The purpose was to determine if repeatable QTL controlling sucrose accumulation could be detected between environments, and also if different QTLs control the accumulation of sugar early compared to later in the growing season. This work formed part of a broader research objective to understand if and how marker assisted breeding approaches can be applied within a sugarcane-breeding program.

Materials and methods

Mapping population

The *S. officinarum* clone IJ76-514 ($2n = 80$) was used as the female in a cross, made in 1999, with Q165$^{\phi}$ ($2n \sim 115$), an Australian cultivar and elite parent. Two hundred and thirty progeny clones were evaluated in two field trials. Both trials were of a randomised complete block design with four replicates of single row 4 m plots at the Kalamia estate, Ayr, (147.4°E–19.5°S) Queensland. The trials were planted in September 2000 and August 2001.

Phenotyping

Brix, pol and commercial cane sugar (CCS) were measured on the extracted juice from four random stalks taken from each plot and crushed through a small mill using standard procedures (BSES 1984). Brix is an estimate of total dissolved solids in juice, while pol is commonly used in sugar industries as an estimate of sucrose content. In the Australian sugar industry CCS is an estimate of commercially extractable sucrose and is determined from a standard function of brix, pol and fibre (BSES 1984). Given the high costs of measuring fibre and relatively small effect of genetic variation in fibre affecting CCS relative to pol and brix, a constant fibre level of 13% was assumed following procedures routinely used in Australian sugarcane breeding programs. All measurements on juice from four stalks were made during April and August in 2001 and April and July in 2002 on plant cane from the separate field trials.

Genotyping and data analysis

Generation of the marker data and construction of the genetic map of Q165$^{\phi}$ were reported by Aitken et al. (2005). AFLPs and SSRs were used to generate 2,238 polymorphic markers that were scored for this population. Of these, 1,365 were present in Q165$^{\phi}$ and were used to generate the map reported by Aitken et al. (2005). 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 numbers of polymorphic bands in descending molecular-weight order. SSR markers were labelled m then the name and identity number from the Sugarcane Microsatellite Consortium collection and a letter denoting the allele by descending molecular weight. A total of 951 single dose (present

only once in the genome and segregating 1:1) and 123 double dose markers (present twice in one parental genome) formed 136 linkage groups (LG) of which 127 LG could be assigned into 8 homo(eo)logy groups (HG). This number of HGs is equivalent to the basic chromosome number of *S. spontaneum* (D'Hont et al. 1996). In two HGs two sets of small LGs aligned to single larger LGs forming in total the ten basic chromosomes of *S. officinarum* (Aitken et al. 2005).

Phenotypic data (CCS, brix, pol) from each experiment were analysed using the SAS statistical package (SAS Institute Inc., NC, USA). For analyses of variance, the following model was assumed:

$$y_{ij} = \mu + b_j + g_i + e_{ij},$$

where μ, b_j, g_i and e_{ij} are the grand mean, block effect, genotype effect and error effect, respectively. Genotypes were considered to be random effects, generating variance component σ^2_g.

Analyses of covariance were done for selected pairs of traits. This was done in the same manner as the analyses of variance except that sums of cross products and mean cross products were determined, with appropriate covariance components and mean cross products substituted for variance components and mean squares.

Broad sense heritabilities (h^2) for each trait were determined from (Falconer and Mackay 1996):

$$h^2 = \frac{\sigma^2_g}{\sigma^2_p},$$

where σ^2_g = genetic variance and σ^2_p = phenotypic variance. Phenotypic variance was determined from:

$$\sigma^2_p = \sigma^2_g + \frac{\sigma^2_e}{r},$$

where σ^2_e = error variance and r = number of replicates.

Genetic correlations between characters X and Y, i.e. $r_{g(x,y)}$ were determined from:

$$r_{g(x,y)} = \frac{\mathrm{cov}_{g(x,y)}}{\sigma_{g(x)} \cdot \sigma_{g(y)}},$$

where

$\mathrm{cov}_{g\ (x,y)}$	genetic component of covariance between characters X and Y,
$\sigma_{g(x)}$ and $\sigma_{g(y)}$	genetic standard deviation for characters X and Y, respectively.

Detection of putative QTL was carried out for each marker using a one-way analysis of variance to identify significant associations between the present or absence of a marker and the trait. Both single dose and multi dose (present as two or more copies per genome) markers were used for this analysis. The threshold used for detection of a QTL was calculated using 1,000 permutations (Churchill and Doerge 1994). The QTLs detected are identified as suggestive and significant and correspond to the genome-wide probabilities proposed by Lander and Kruglyak (1995). Where suggestive linkage corresponds to linkage that would occur one time at random in a genome scan and significant linkage where linkage would be expected to occur 0.05 times in a genome scan (that is, with probability 5%). This allowed the detection of putative QTL which in the continuation of this study will be verified in further generations. As sugarcane is a polyloid up to 12 alleles could be segregating at a locus which means that only the most significantly different alleles are likely to be detected although other alleles at that locus may contribute to the trait. To test for association at alleles of a QTL a second round of analysis was carried out with a modified threshold corresponding to $P = 0.01$. The QTL effect was estimated as the average difference in phenotype of the individuals carrying the marker compared to those without. To identify epistatic interactions, every single dose marker was tested for digenic linear × linear interactions. Due to the relatively small sample size, estimations of the four class means are poor so a stringent threshold of $P < 0.00001$ was used to identify an interaction. The phenotypic variance explained by all the QTL was determined by multiple regression.

Results

Analysis of phenotypic data

For all traits, (brix, pol and CCS), the clone effect was highly significant ($P < 0.0001$) in both field trials. Broad-sense heritabilities were generally higher in the 2002 field trial probably due to a more even establishment (Table 1). However, in both years the heritabilities were high indicating good control of within-trial environmental variation and experimental error. Estimated genetic correlations were high (> 0.7) between brix, pol and CCS for all sampling occasions (Table 2). There were also high genetic correlations estimated between pol (Table 2) and other traits (not shown) for each time of measurement in both experiments. QTL identified for CCS added no new information and therefore the results for CCS are not reported further in this study. Frequency histograms of trait performance across progeny showed approximately normal distributions across both years (Fig. 1). There was a shift to higher values from the samplings early in the year in 2002 compared with 2001 reflecting better environmental conditions for sucrose accumulation in 2002.

Quantitative trait loci identified for Q165$^{(i)}$ using single dose markers

In the initial analysis, measurements from early and mid-season sampling were averaged across years. This was to increase the chance of detecting robust QTL that

Table 1 Mean, error variance, and broad sense heritability of each trait in each of the 2000 and 2001 planted experiments

Trait	Time of sampling	Experiment	Mean	Error variance	Broad-sense heritability
Brix (%)	Early	2000	11.6	0.87	0.76
		2001	14.1	0.47	0.84
	Mid-season	2000	18.7	1.40	0.59
		2001	18.5	0.39	0.88
Pol reading	Early	2000	30.0	20.3	0.93
		2001	39.5	14.2	0.82
	Mid-season	2000	67.7	40.6	0.59
		2001	65.1	12.7	0.90
CCS (%)	Early	2000	1.67	0.85	0.74
		2001	4.04	0.71	0.81
	Mid-season	2000	12.6	1.21	0.63
		2001	9.61	0.59	0.90

Table 2 Estimates of genetic correlations from analyses of covariance for measurements of (1) brix versus CCS, (2) pol versus CCS and (3) pol versus brix; each for the four experiment × sampling time combinations

Experiment	Time of sampling	Brix versus CCS	Pol versus CCS	Brix versus pol	Pol early versus pol mid
2000	Early	0.71	1.00	0.97	0.84
	Mid	0.95	0.96	0.99	
2001	Early	0.90	0.99	0.93	1.00
	Mid	0.94	0.99	0.97	

Also shown are estimates of genetic correlations for pol measured between early and mid sampling times for the 2000 and 2001 experiments

were less experiment specific. No QTL were identified as significant but at the suggestive level, for combined early data, five QTLs were identified that affected pol, and six that affected brix (Table 3). These represented eight different QTLs, which were located in HG 1, 2 and 3 and explained from 3 to 5% of the phenotypic variation. Three of these QTLs were detected using both traits; two were unique to pol and three to brix (Table 3). These QTLs were represented by from one to two markers at this significance level. When the threshold level was decreased to $P < 0.01$, 11 and 14 putative QTLs were identified for pol and brix respectively (Table 3). These were represented by 14 different QTLs, eleven of which affected both traits and three unique to brix. Again they explained from 3 to 5% of the variation and were represented by from one to five markers. These were located on 14 LG, which belonged to HG 1–6 (Table 3). LG 4 from HG 3 contained two QTLs, which were located at different ends of the LG with opposite effects on the traits. The rest of the LGs contained only one QTL although the chance of more than one QTL located on the same linkage group cannot be ruled out. The reduction in stringency from suggestive to $P \leq 0.01$ added a further six putative QTLs but also increased the likelihood of the different traits detecting the same QTLs.

Analysis of the mid-season measurements identified no QTL as significant but at the suggestive level seven and eight QTLs for pol and brix were identified. These were represented by from one to five markers located on eight different LGs in HGs 1–5 (Table 3). Seven of these QTLs affected both traits, with one unique to brix, and explained from 4 to 7% of the phenotypic variation. Reducing the stringency to $P \leq 0.01$ identified a further eight putative QTLs using the combined mid season data. In total, for the mid season data, 16 QTLs were located on 16 LGs in HGs 1–5. Thirteen of these QTLs were identified by both traits. For both the early and mid season data, 23 QTLs were identified which mapped to 22 LGs, which were located in HGs 1–6. Seven of these QTLs were detected with both early and mid season data, seven were unique to early season measurements and nine unique to mid season measurements (Table 3).

To examine the repeatability of the results between the two trials, the data was reanalysed for separate years. Only one QTL was identified at the significant level on HG1 which explained a maximum of 9% of the phenotypic variation and was only identified in 1 year. A total of 104 markers were identified at the suggestive level for the eight traits. These were located to 24 QTLs with from 1 to 5 markers at each location, each of which explained from 4 to 9% of the variation (Table 3). Only three QTLs located on HG 2 LG 6, HG 2 LG100 and HG 3 LG18 were significant in both years at this stringency. When the stringency was reduced to $P \leq 0.01$ a further four QTL were identified (Table 3). At this stringency 14 (50%) of the QTLs were detected in both years' data with at least one of the traits. If the level was dropped to $P \leq 0.05$ then 23 (82%) were detected in both years' data with at least one trait. Both positive and negative effects were identified for QTL and in all cases these effects were consistent across traits and years. Using the individual year data at the suggestive level the number of QTLs detected varied from 3 to 9 for each trait and each explained from 3 to 9% of the variation (Table 3). The majority of the QTLs had been previously detected with the combined data; only five QTLs were unique to this data set. Reducing the stringency to $P \leq 0.01$ resulted in from 8 to 15 QTL detected by the traits; only 3 of these additional QTLs were not detected at the suggestive level with at least one trait. In one case, two QTLs of opposite effect were located on the same

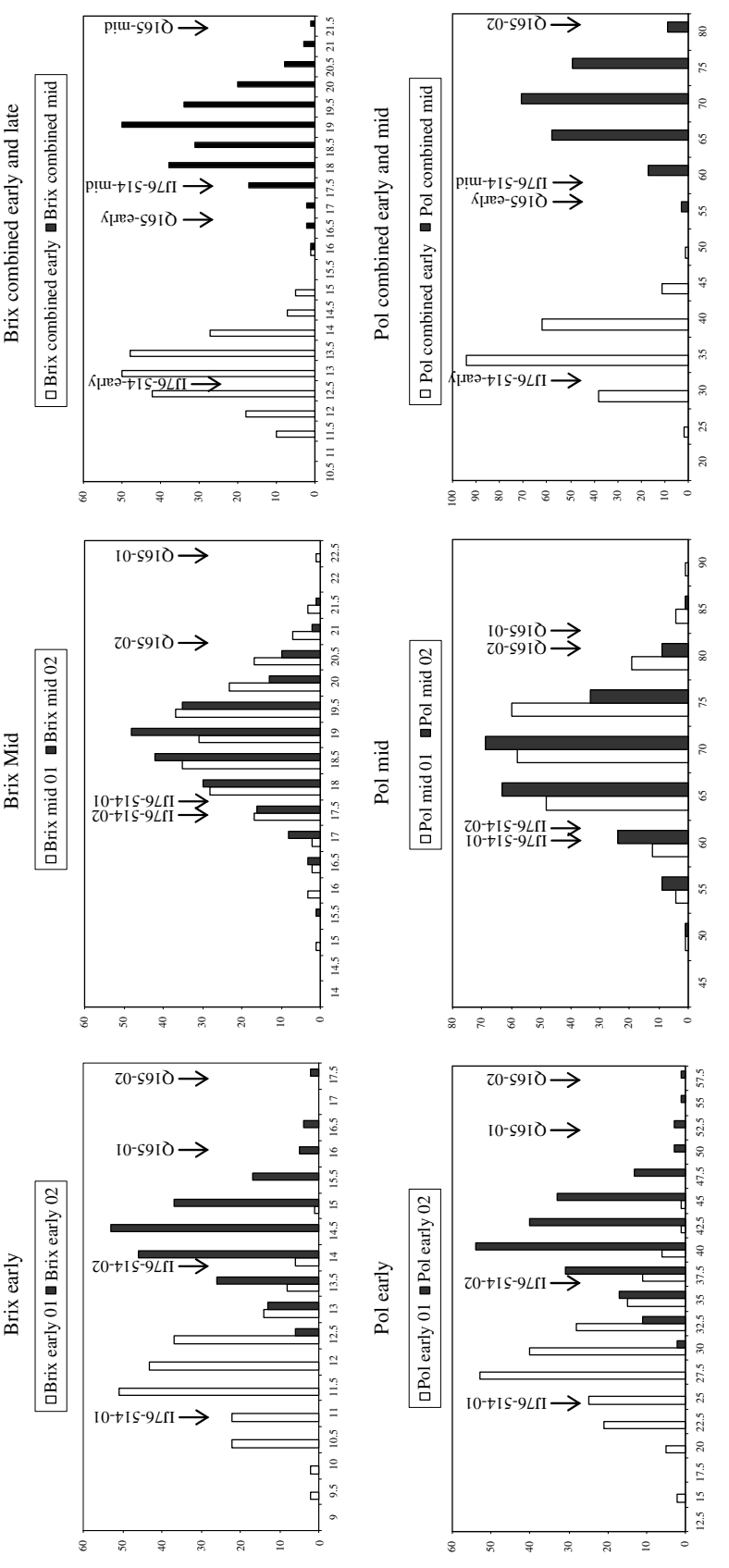

Fig. 1 Frequency distribution of the eight traits for the IJ76-514 × population showing means of the two parents in both years. Early = early season data, mid = mid season data

Table 3 The significant and suggestive associations between single and multi dose markers for brix and pol readings for the combined early, mid season and individual years data

Marker[a]	HG/LG[b]	QTL[c]	Early season										Mid season									
			Pol						Brix				Pol				Brix					
			2001	r[d]	2002	r	Combined[e]	r	2001	r	2002	r	Combined r	2001	r	2002	r	Combined r	2001	r	2002	r	Combined r
Single dose markers																						
Aaccac24	1, 81	Q-1	−2.12***[g]	5	−1.43*[h]	4	−1.68***	4	−0.38***	4	−0.24*	3	−0.28**	3								
mSSCIR36b[f]	1, 17	Q-1	1.94***	4					0.38***	4												
mSMC179sah	1, 50	Q-1	1.44*	2					0.30**	3				1.63*	2	1.97**	3	1.79**	3		0.33***	3	0.31** 3
mSSCIR8c	1, 42	Q-2												−3.50****	8			−2.28***	5	−0.66****	9		−0.39*** 5
mSSCIR 8b	1, 19	Q-2	−1.34*	2	−1.78***	4	−1.46**	3	−0.25*	2	−0.33**	4	−0.27** 3	−2.21**	3	−2.70***	6	−2.36***	5	−0.42**	3	−0.40***	5 −0.42*** 5
Aggctc29	2, 6	Q-3	1.76**	3	1.92***	4	1.84***	5	0.31**	3	0.35***	4	0.34*** 5	2.37***	4	1.62*	2	1.83**	3	0.42**	4		0.31** 3
mSMC1825LAj	2, 8	Q-3	1.39*	2	1.50*	3	1.35**	3	0.30*	3	0.32**	3	0.28** 3	2.40***	4					0.44***	4		
mSMC1825LAk	2, 35	Q-3	−1.77**	3					−0.36***	4				−2.12**	3	−1.68*	2	−1.88**	3	−0.35*	2		−0.30* 3
Acacta6	2, 12b	Q-4	−1.67**	3			−1.35*	3	−0.36***	4			−0.30** 4	−2.04**	3	−1.81*	3			−0.40**	3	−0.35**	4
mSC336BSg	2, 87	Q-4	1.75*	3	1.50*	3	1.55*	3	0.40***	5	0.35**	4	0.35*** 5	2.14*	3	2.03*	3	1.91*	3	0.41*	3	0.39**	4 0.36** 4
Agcctg29	2, 100	Q-4												−2.37***	4	−2.18***	4	−2.12***	4	−0.42**	3	−0.35**	4 −0.36*** 4
mSMC292MSk	2, 13	Q-4			1.75***	4					0.33**	4		2.01*	3	2.04**	3	1.83**	3	0.38*	3	0.32**	3 0.33** 3
Aagcta45	2, 38	Q-5									0.37***	4									0.34**	3	
Acccat35	2, 53	Q-5			1.80***	4	1.55**	4			0.41***	5	0.34*** 5	1.65*	3	2.73***	6	2.59***	6	0.33*	2	0.48***	7 0.43*** 7
Acgctc2	3, 4	Q-6	2.00***	4	1.72**	4	1.64***	4	0.30*	3	0.30**	3	0.28** 3	2.06**	3	1.88*	3			0.35*	2	0.26*	2
Aggctg4	3, 18	Q-6	1.48*	2	1.57**	3	1.55**	4	0.34**	3	0.35***	4	0.33*** 4	2.33***	4	1.94**	3	2.18***	4	0.35*	2	0.34**	3 0.39**** 5
Actctg29	3, 4	Q-7	−1.55*	3	−1.91***	4	−1.64***	4	−0.29*	3	−0.39***	5	−0.33*** 4										
Acacat40	3, 113	Q-7	−1.28*	2	−1.65**	3	−1.59***	4	−0.25*	2	−0.31**	3	−0.31*** 4										
Aagctc5a	3, 21	Q-8												−2.38***	4					−0.43**	4		
Acccac9	3, 51	Q-8	2.12*	3					0.44**	3				2.34***	4			1.84**	3	0.43**	4	0.28*	2 0.34** 3
Agccag1	4, 3	Q-9	1.91***	4	1.60**	3	1.47**	3	0.32**	3	0.31**	3	0.27** 3	2.99***	6	1.93**	3	1.91**	3	0.56***	6	0.29*	2 0.33** 3
Aagctc13	4, 44	Q-9												−2.65***	5	−1.94**	3	−2.40***	5	−0.48***	5		−0.37*** 4
Acacat7	4, 56	Q10	1.46*	2			1.31*	3								1.93**	3	1.88**	3				0.29* 3
Aagctc52	4, 102	Q10			1.49*	3					0.32**	3		2.07**	3					0.36*	3		
mSSCIR34e	5, 55	Q11			−1.29*	2	−1.39**	3			−0.28*	3	−0.30** 3	−2.12**	3	−2.03**	3	−1.94**	3	−0.38*	3	−0.35**	3 −0.35*** 4
mSSCIR34a	5, 57	Q11														1.98**	3	1.99***	4	0.34*	2	0.37***	4 0.39*** 5
mSSCIR64c	6, u[i]	Q12	2.1***	5					0.40***	5													
Agcctc17	6, 71	Q12	−1.44*	2	−1.65**	3	−1.29*	2	−0.31**	3	−0.37***	5	−0.30** 4										
Multi dose markers																							
MmSSCIR14b[j]	4, u	Q10							0.26**	3	0.27**	3	0.26*** 4										
MAggcta43	4, u	Q10														2.13***	5				0.29**	4	
MAcgcac24	u[g]				1.43*	3					0.29**	3		1.57*	2	2.15***	4	1.92***	4				0.33*** 4
MAcccat15	u													2.13***	5					0.38***	4		
MAgccac20	u		2.03***	4			1.64**	4	0.37***	4	0.26*	2	0.32** 4	1.76*	2	2.13**	3	1.98**	3		0.34**	3	0.32** 3
MAaccag19	u								−0.25**	4	−0.24*	3	−0.23** 3	−1.48*	2	−1.78**	3			−0.30*	2	−0.35***	5 −0.27** 3

Table 3 (Contd.)

Marker[a]	HG/LG[b]	QTL[c]	Early season									Mid season								
			Pol					Brix				Pol					Brix			
			2001	r^2[d]	2002	r	Combined[e]	r	2001	r	2002	r	Combined	r	2001	r	2002	r	Combined	r
MActcat30	u			−1.37**	4					−0.28***	4									
MAacctc10	u		1.73**	3	2.31***	6	1.85**	5	0.31*	3	0.45***	6	0.36***	5	2.17**	3	2.05**	3	2.23***	4
MActcat6	u														−2.22***	6	−1.37*	2	−1.87***	5

When an association was significant for a given trait the association for the other traits is also reported up to $P \leq 0.05$. For each association the percentage of variance (r) and the additive effect attributed to that marker is also reported

[a]The most significant marker of the group associated with the trait

[b]The homology group and linkage group from Aitken et al. (2005)

[c]Assignment of a QTL to a homo(eo)logous locus

[d]The phenotypic variation explained by that marker (r^2)

[e]QTL identified with combined data from both years

[f]All SSR markers start with an m all other markers are AFLPs

[g]Genome wide significant levels, suggestive linkage—statistical evidence that would be expected to occur one time at random in a genome scan (***); significant linkage—statistical evidence expected to occur 0.05 times in a genome scan (that is with probability 5%) (****)

[h]Level of significance for marker-trait association of alleles of a QTL: $0.01 \leq P \leq 0.05$*, $0.005 \leq P \leq 0.01$**

[i]Unknown location

[j]M in front of a marker name indicates a multi dose marker

LG, HG 3 LG 4, which were detected with both the combined and individual years' data. Of the 28 QTL that were detected using single dose markers with all the traits, 17 were detected with both the early and mid season data; six with only early data and five with only the mid season data.

Multi dose markers

The 239-multi dose markers were tested for association with the traits by ANOVA as for the single dose markers. For all the traits a total of nine markers were significant at the suggestive level of which eight were detected by both traits (Tables 3, 4). Four of these markers were consistent across years at $P \leq 0.01$ with at least one trait, four were detected by both early and mid season data, three by mid season data only and two by early season data only. Seven of these markers fitted the expected ratios for double dose markers while the other two were higher dosage. Two of these markers could be placed on the map, and both mapped to HG4. Aggcta43 is linked to LG56 with an existing QTL and another LG not containing a QTL (HG4 LG12b). The other double dose marker mSSCIR14b was linked to LG102 only (Aitken et al. 2005).

Digenic interactions

For the combined data, nine digenic interactions were detected at $P \leq 0.00001$ and eight affected more than one trait. From 1 to 4 independent interactions were detected for a given trait (Table 4). A significant QTL was detected at or near 10 of the 18 markers involved in the interactions. The interactions detected were always between at least one marker located near or at a QTL; in one case the interaction was between two QTLs. The r^2 of interactions individually ranged from between 11 and 13% and when combined from 24 to 36% (data not shown and Table 4). The HG could be assigned for all markers involved in the interactions apart from one. Interactions were detected both within and between HGs. HG2 interacted with HG3 and HG4 and an unassigned LG as well as within HG2. HG3 interacted with LGs within HG3 and the unassigned LG and HG1 interacted with HG3. In both cases of interactions within HG they could be between alleles of the same loci but the rest of the interactions are evidence of epistasis between alleles belonging to different loci.

When the individual year's data was analysed, a total of 97 interactions were significant at $P \leq 0.00001$. Out of the 194 markers involved in these interactions, 4 were unlinked markers, 49 were markers not linked to QTL and 141 were linked to QTL. None of the interactions were detected in both years at a threshold of $P \leq 0.00001$ level but all the interactions affected more than one trait. The majority of interactions were detected in 2001 with only 12 interactions at this significance level detected in

Table 4 Number of single dose QTL, multi dose QTL (M QTL) at different significant levels and interactions detected for brix and pol and the amount of variation explained by the models

	Pol						Brix					
	E01[g]	M01	E02	M02	CE	CM	E01	M01	E02	M02	CE	CM
Single dose markers[a]	9 (3)	23 (8)	10 (1)	11 (3)	8 (2)	19 (9)	11 (1)	16 (5)	27 (4)	15 (5)	10 (1)	20 (7)
Single dose markers common to both years[a]	0	1					0	0				
No. of QTL[a, b]	5	9	5	3	5	7	6	4	6	3	7	8
No. of QTL[c]	8	15	10	11	11	15	12	10	14	11	14	14
QTL common to both years[c]	3	6					5	3				
M QTL[a]	1 (0)	2 (2)	1 (1)	2 (1)	0 (0)	3 (1)	1 (0)	2 (0)	2 (1)	2 (0)	2 (1)	3 (1)
M QTL common too both years[c]	1	1					1	3				
Int QTL pairs[d]	3	13	1	1	2	4	2	13	4	1	3	4
r^2 of multiple regression[e]												
QTL[a]	17% (199)	31% (192)	17% (202)	13% (207)	20% (203)	27% (177)	20% (152)	21% (193)	25% (183)	12% (183)	28% (143)	29% (177)
QTL[c]	27% (186)	39% (184)	24% (198)	27% (202)	33% (177)	40% (169)	42% (77)	34% (192)	42% (137)	33% (144)	41% (133)	38% (136)
QTL and M QTL[f]	31% (185)	43% (184)	28% (198)	38% (201)	38% (176)	48% (168)	43% (77)	38% (192)	51% (137)	43% (143)	50% (133)	49% (135)
Int-QTL	27% (202)	56% (199)	13% (203)	11% (206)	24% (146)	36% (179)	18% (206)	55% (199)	37% (129)	13% (206)	26% (201)	33% (178)
QTL, MQTL and int-QTL	43% (180)	66% (182)	37% (194)	43% (199)	51% (125)	55% (166)	51% (77)	62% (189)	65% (85)	47% (141)	58% (128)	54% (132)

[a]Detected at the suggestive level
[b]One marker counted per cluster
[c]Detected at $P \leq 0.01$ and previously detected with another trait at the suggestive level
[d]Interactions detected at $P \leq 0.00001$
[e]In parenthesis the total degrees of freedom of the regression model representing the number of no missing molecular data across all factors
[f]QTL identified with single dose and multi dose markers
[g]E = early season data and M = mid season data in 2001 and 2002, CE = combined early season data, CM = combined mid season data

2002. The individual interactions explained from 11 to 15% of the variation (data not shown and Table 4). The markers involved in the interactions were located on five LGs in HG1, eleven in HG2 and HG3, six in HG4, two in HG5 and one in HG6. In total, 43% of the interactions were represented by more than one marker, and regions located on HG1 LG50, HG2 LG12b, HG3 LG51 and HG3 LG74 had four markers involved in the interaction while HG6 LG71 had five markers and HG4 LG3 six markers. The majority of interactions were between two markers only but six regions interacted with three or more other LGs; all of these apart from one are located at a QTL. One region on HG1 LG42 interacted with ten other regions, HG2 LG53 with five other regions, HG1 LG81 with four other regions, and HG3 LG4, LG51 and LG74 with three other regions. Out of the 32 location-by-location interactions, which represent all the interactions, only five were between two LGs in the same HG. Although the same markers were not repeatable between the 2 years, six out of the 32 regions were involved in interactions in both years.

Location of QTLs within homo(eo)logous groups

Of the 136 linkage groups generated in the Q165$^{\phi}$ map (Aitken et al. 2005), 127 were placed into the 8 HGs that represent the basic chromosome number of *S. spontaneum*. All of the single dose QTLs could be assigned to a map location (Table 3). There were five QTLs located on HG1. Three of these were in similar positions on LG17, LG81 and LG50 and could correspond to different alleles of the same homo(eo)logous loci (Table 3). The other two QTLs from HG1 were also in similar positions on LG42 and LG19 (Table 3). The most significant marker was the SSR mSSCIR8 suggesting that these are alleles of the same homo(eo)logous loci. Nine QTLs were located to HG2. The QTLs on LGs 35, 8 and 6 all have QTL either at or adjacent to the SSR mSMC1825LA and are probably alleles of the same homo(eo)logous loci. A Further three QTLs on LGs 53, 38 and 13 are again in similar positions and a third group of QTL on LGs 100, 87 and 12b, with LG 87 and 12b in repulsion correspond to two more homo(eo)logous loci (Table 3). HG 3 contains 6 QTLs, two of which are on LG4. Interestingly this HG group has a longer *S. spontaneum*-like chromosome, which appears to be in alignment with two sets of smaller *S. officinarum* chromosomes (Aitken et al. 2005). One QTL on LG4 is in a similar position to the QTL on LG18 and is probably an allele of the same homo(eo)logous loci. Interestingly there were indications of other QTLs in this region on other LGs but they did not reach the significance thresholds used. The rest of the QTL are located on the other set of *S. officinarum* chromosomes. The other QTL on LG4 aligns with LG113, while the QTLs on LG 51 and LG 21 also align. Thus potentially, there are three loci on HG 3 that influence sucrose accumulation in Q165$^{\phi}$ (Table 3). Four QTLs were identified for HG 4. This was another HG that contains *S. spontaneum* chromosomes, which align with two sets of *S. officinarum* chromosomes. The QTL on LG 3 aligns to the QTL on LG 44, again there were indications of other putative QTLs on other LGs in this region although they did not meet the threshold and were not defined as QTL. The QTL on LG 56 is in the other set of *S. officinarum* chromosomes and aligns to the double dose marker mSSCIR14b, which was located to LG 102. The other multi dose marker that could be located on the map (Aggcta43) aligns to LG56 and is already represented by this QTL and LG 12b, which does not have a QTL, so there appears to be at least two homo(eo)logous loci on HG 4. Two QTLs were identified for HG 5 and are probably alleles of the same homo(eo)logous loci as the most significant marker in both case was mSSCIR34. Two QTLs were detected on HG 6, one on LG 71 and the unlinked SSR mSSCIR64. These are probably alleles of the same homo(eo)logous loci as LG 71 is in repulsion to LG 62, which contains other alleles of mSSCIR64. The fewer QTL detected on HG5 and 6 could be due to lower genome coverage as these groups are sparsely covered with markers. In total, there appears to be a minimum of 12 loci that effect sucrose accumulation in Q165$^{\phi}$.

Multiple regression analysis

Using multiple regression on all traits, the single dose QTLs explained from 13% (three QTL) to 31% (nine QTL) of the variation (Table 4). If the stringency was dropped to include all significant alleles at a QTL ($P \leq 0.01$) then 8–15 QTL explained 24–42% of the variation. The addition of the multi dose markers resulted in a slight increase of the variation explained but less than 10% up to a maximum of 51% (Table 4). The amount of variation explained by the interactions varied from 11% ($df = 206$) to 56% ($df = 199$) and was proportional to the number of interactions included in the model for that trait. In three out of the 12 traits, the amount of variation explained by the interactions exceeded the amount explained by the single dose QTL (Table 4). The proportion of total genetic variance explained by all the effects in a single model ranged from 37 to 66%. In most cases, these estimates were based on over half the population but in two cases they were below 100. The markers accounted for similar amounts of phenotypic variation in 2001 and 2002 and also in the early verus the mid season data.

Allele contributions

In total for both single and multi dose markers, 37 QTLs were identified for the two sucrose related traits studied. Out of these, 22 QTLs had a positive effect on the traits and 14 had a negative effect. The brix combined mid season data was examined to determine the relationship between the phenotype and the number of QTLs with

positive and negative allele effects. The individual plants were grouped on the basis of brix content and the average number of positive and negative QTLs determined within each group (Fig. 2). The number of negative QTLs decreased as the brix increased and the converse was true of positive QTLs (Fig. 2). For the group with on average the highest brix value, the plants had 16 positive QTLs and 4 negative QTLs.

Discussion

Brix and pol are known by sugarcane breeders to be moderately to highly heritable traits and this is further illustrated in this study where heritability ranged from 0.59 to 0.93 (Table 1). Despite this, and intensive selection for high sucrose, sugarcane breeders have made limited improvement in increased sucrose for parents or sugarcane cultivars for the last 40 years (Rattey et al. 2004; Jackson 2005). Q165$^{\phi}$ is considered by sugarcane breeders to represent a typical elite parent in Australian breeding programs, sharing common ancestors with other elite parents widely used. The absence of segregating large QTL effects for brix or pol in an elite parent is not surprising, since alleles with large positive or negative additive effects, if present at all in ancestral clones, should have been readily accumulated to a higher dosage level (if positive effect), or eliminated from prior breeding within several breeding generations (if negative) for a highly heritable trait like brix and pol. However, if a number of segregating QTLs with small or moderate additive effects still remain in elite parents, as suggested by the results of this study, then breeding and selection should still result in incremental genetic gains, until the more favourable alleles at each locus are fixed. Therefore, assuming the results observed for Q165$^{\phi}$ are similar to that of other elite parents in our breeding programs, it is unclear why clear genetic improvement in brix or pol of cultivars has not been readily achieved over recent decades. However, some potential reasons may include: (1) epistasis effects where QTL with observed positive effects across a whole population have much smaller effects within high sucrose genotypes due to other overriding genetic-physiological limits at higher levels of sucrose accumulation, (2) that insufficient selection pressure has been applied for brix and pol in breeding programs in comparison with other traits (Jackson 2005), or (3) that some favourable QTL for brix and pol have unfavourable pleiotropic effects on other important traits under selection (e.g. cane yield). Further research is required to assess if these explanations, or others, lie behind limited breeding progress for CCS.

In this study we chose to identify QTL using a threshold derived from permutation testing and then reduce these thresholds to identify further alleles of the QTL which would be present in this polyploid (Lander and Kruglyak (1995). It is interesting to note that after lowering the stringency to $P \leq 0.01$ the majority of QTL detected at the lower stringency had already been detected with another trait at the higher threshold level. We detected a total of 29 QTLs with the combined data, the majority of which were detected with more than one trait. The number of QTLs detected increased to 37 when the individual year's data were analysed. Each QTL only explained a small amount of variation from 3 to 9%, which is in the same range as observed in another sugarcane cultivar, R570 (Hoarau et al. 2002). The low r^2 is likely to be a product of the high ploidy coupled with high heterozygosity, which results in many alleles at homologous loci. In contrast to a diploid plant where the comparison is between two alleles per locus, in sugarcane there could be up to 12 or more segregating factors per locus. Another study on crosses between *S. spontaneum* × hybrid clones (derived from *S. officinarum* × *S. spontaneum*) detected some larger effects probably due to the use of the wild relative with very low sucrose content (*S. spontaneum*) but most were within the same range (Ming et al. 2001, 2002b).

The fairly high map coverage of Q165$^{\phi}$ allowed the QTL to be aligned into HG. All of the single dose markers and the double dose markers that could be assigned a map position were aligned into 12 QTL positions across 6 HG (Table 3) In all cases an allele of the loci assigned as Q1–Q12 (Table 3) was detected at the significant or suggestive level indicating that a gene or genes at this location has an effect on sucrose accumulation. Ming et al. (2002b) mapped 50 QTLs affecting sugar yield to 12 genomic regions in sugarcane. Using the SSRs mapped in both Rossi et al. (2003) and Aitken

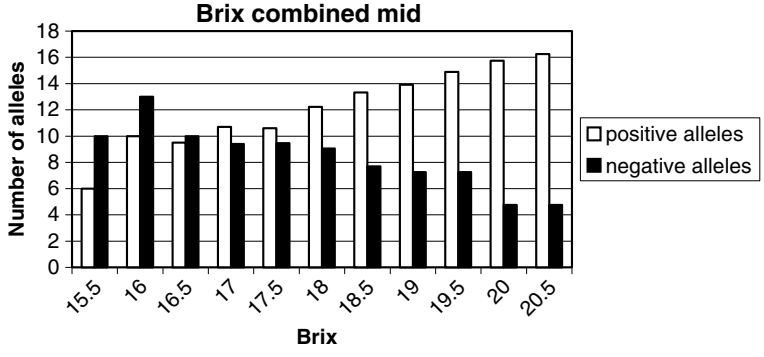

Fig. 2 Allele frequency distribution based on brix combined mid season values for Q165$^{\phi}$

et al. (2005), some of the QTL positions in Q165$^{\phi}$ and R570 could be aligned. Out of the nine QTL identified for brix at $P \leq 0.005$ in Hoarau et al. (2002) four could be assigned to a homology group. Of these, one QTL mapped to CGII close to mSSCIR36, which maps to our HG1 where a QTL for brix was identified at the suggestive level (Q1). The others map to HGs where there are not enough markers in common to compare the locations.

A total of 37 QTLs were identified when the data was analysed by year. At $P \leq 0.01$, 18 (49%) of the QTL were identified in both years with at least one of the traits. If the significance level was dropped to $P \leq 0.05$ then 29 (78%) of the QTL were detected in both years with at least one of the traits. Whatever the significance level, the effect was always in the same direction for all of the QTL identified in both years. This consistency of QTL detection between years is similar to that seen in an equivalent study (Hoarau et al. 2002). In both studies, interactions were detected and contributed to a significant amount of the variation explained in the final model. In both cases, most of the interactions were not consistent between years although the direction of the effect was always consistent. Six of the 32 regions involved in interactions in Q165$^{\phi}$ were detected in both years. It is interesting to note that of the 194 markers involved in the interactions, 73% were linked to QTL. This was not the case in the previous sugarcane study where only 6 of the 82 markers involved in the interactions were associated with a QTL (Hoarau et al. 2002). Studies on other crops have identified both QTL involved with interactions and interactions not involving QTL (Kulwal et al. 2004; Li et al. 1998). Liao et al. (2001) demonstrated that in rice epistasis is more sensitive to genetic background and environments than main effect QTLs. The interactions in this study have to be taken cautiously due to the unbalanced size of the genotypic classes compared to the single marker QTL detection. The addition of these interactions into the multiple regression models increased the variance explained by a substantial amount. Further work is needed with larger populations to determine the exact significance of these interactions.

Early and mid season QTL

Seven QTLs (19%) were consistently detected with only the early season data and 8 (22%) with only the mid season data. The majority of the QTLs (59%) had an effect at both early and mid season sampling, which explains the relatively high correlations between these sampling dates (Table 2). Out of the seven QTLs identified for early season brix and pol, four had a negative effect on sucrose and three a positive effect. Selection for the positive QTLs and away from the negative QTLs should result in high early brix and pol, which results in higher mature cane sucrose (Jackson and Morgan 2003). An increase in early season brix and pol would be of major economic benefit to the sugar industry (Cox et al. 1994). Out of all the QTLs detected, 22 had positive effects on the traits. For brix combined mid season data, the progeny with the highest brix contained on average 16 positive effect QTLs but four negative effect QTLs (Fig. 2). In the population studied, no progeny existed which contained all 22 positive alleles. However, marker assisted breeding strategies should offer breeders better ways of selecting parental clones with higher breeding values, and such strategies will be tested in future experiments, in progress.

References

Aitken KS, Jackson PA, McIntyre CL (2005) A combination of AFLP and SSR markers provides extensive map coverage and identification of homo(eo)logous linkage groups in a sugarcane cultivar. Theor Appl Genet 110:789–801

Bureau of Sugar Experiment Stations (1984) The standard laboratory manual for Australian sugar Mills. Volume 1. Principles and practices'. BSES, Brisbane

Churchill GA, Doerge RW (1994) Empirical threshold values for quantitative trait mapping. Genetics 138:963–971

Cox MC, Hogarth DM, Mullins RT (1990) Clonal evaluation of early sugar content. Proc Aust Soc Sugar Cane Technol 12:90–98

Cox MC, Hogarth DM, Hansen PB (1994) Breeding and selection for high early season sugar content in a sugarcane (*Saccharum* spp. hybrids) improvement program. Aust J Agric Res 45:1569–1575

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

Falconer DS, Mackay TFC (1996) Introduction to quantitative genetics. Longman Group, Harlow

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

Hoarau J-Y, Grivet L, Offmann B, Raboin L-M, Diorflat J-P, Payet J, Hellmann M, D'Hont A, Glaszmann J-C (2002) Genetic dissection of a modern sugarcane cultivar (*Saccharum* spp.). II. Detection of QTLs for yield components. Theor Appl Genet 105:1027–1037

Jackson PA (2005) Breeding for improved sugar content in sugarcane. Field Crops Res 92:277–290

Jackson PA, Morgan TE (2003) Early stage selection for commercial cane sugar (CCS) in sugarcane clones: effects of time of sampling and irrigation. Aust J Agric Res 54:389–396

Kulwal PL, Singh R, Balyan HS, Gupta PK (2004) Genetic basis of pre-harvest sprouting tolerance using single-locus and two locus QTL analyses in bread wheat. Genomics 4:94–101

Lander E, Kruglyak L (1995) Genetic dissection of complex traits: guidelines for interpreting and reporting linkage results. Nat Genet 11:241–247

Li Z, Pinson SRM, Stansel JW, Paterson AH (1998) Genetic dissection of the source–sink relationship affecting fecundity and yield in rice (*Oryza sativa* L.). Mol Breed 4:419–426

Liao CY, Wu P, Hu B, Yi KK (2001) Effects of genetic background and environment on QTLs and epistasis for rice (*Oryza sativa* L.) panicle number. Theor Appl Genet 103:104–111

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 S-C, Moore PH, Irvine JE, Paterson AH (2001) QTL analysis in a complex autopolyploid: genetic control of sugar content in sugarcane. Genome Res 11:2075–2084

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

Ming R, Wang Y-W, Draye X, Moore PH, Irvine JE, Paterson AH (2002b) Molecular dissection of complex traits in autopolyploids: mapping QTLs affecting sugar yield and related traits in sugarcane. Theor Appl Genet 105:332–345

Peccoud J, Vander Velden K, Podlich DW, Winkler CR, Arthur WL, Cooper M (2004) The selection values of alleles in a molecular network model are context dependent. Genetics 166:1715–1725

Podlich DW, Winkler CR, Cooper M (2004) Mapping as you go: an effective approach for marker-assisted selection of complex traits. Crop Sci 44:1560–1571

Rattey AR, Jackson PA, Wei X-M, Cox M, Stringer J (2004) Opportunities to increase rates of parent improvement in Australian sugarcane breeding programs. In: Proceedings of the Australian society sugar cane technical conference, vol 26 (CD-ROM)

Roach BT (1984) Conservation and use of the genetic resources of sugarcane. Sugarcane 2:7–11

Roach BT (1989) Origin and improvement of the genetic base of sugarcane. Proc Aust Soc Sugar Cane Technol 10:34–47

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