Theor Appl Genet (1999) 99:1053–1060

N. Jannoo · L. Grivet · A. Dookun · A. D'Hont
J.C. Glaszmann

Linkage disequilibrium among modern sugarcane cultivars

Received: 6 March 1999 / Accepted: 25 March 1999

Abstract Modern sugarcane cultivars are derived from a few interspecific hybrids created early in this century. Linkage disequilibrium was investigated in a population of 59 cultivars representing the most important commercial clones bred in Mauritius as well as a few old cultivars involved in their genealogy. Thirty-eight probes scattered over the sugarcane genome map were used to reveal RFLPs. Forty-two cases of bilocus associations were observed involving a total of 33 loci. Most of them are separated by less than 10 cM. All the corresponding allele couples were found in at least 1 of the originally created cultivars, suggesting that they depict ancient associations. This global disequilibrium is interpreted as the result of the foundation bottleneck related to the first interspecific crosses; the preferential allele associations thus created have been maintained through subsequent crosses when the loci were closely linked. This phenomenon is likely also to apply to genes of agricultural interest. A practical consequence is that markers can be used to track known QTLs in modern breeding materials without the necessity to repeatedly study segregating progenies. This structure gives high value to the correlation between molecular markers and agricultural traits among cultivars.

Key words Sugarcane cultivars · Linkage disequilibrium · Foundation bottleneck · Interspecific hybrids · RFLP

Communicated by H.F. Linskens

N. Jannoo · L. Grivet · A. D'Hont (✉) · J.C. Glaszmann
CIRAD, Centre de Coopération Internationale
en Recherche Agronomique pour le Développement,
Avenue Agropolis, BP 5035, F-34032 Montpellier cedex 1, France
e-mail: dhont@cirad.fr
Fax: 334 67615605

A. Dookun
MSIRI, Mauritius Sugar Industry Research Institute,
Réduit, Mauritius

Introduction

Sugarcane is the main sugar-producing crop. It belongs to the genus *Saccharum* of the Poacea family. The genus *Saccharum* is characterised by a high ploidy level and aneuploidy. Two species, *S. officinarum* (2n=80) (Bremer 1930; Li and Price 1967; Price and Daniels 1968; Jagathesan et al. 1970) and *S. spontaneum* (2n=40 to 128) (Sreenivasan et al. 1987), have played a major role in the production of commercial sugarcane cultivars. *S. officinarum*, also known as the 'noble' cane, is the sugar-producing species; *S. spontaneum* is a wild species, poor in sugar but resistant to various biotic and abiotic stresses.

Propagation of sugarcane varieties has been performed exclusively by vegetative means. Until the end of the 19th century, cultivated canes consisted mainly of *S. officinarum* clones, collected from native Papuan and Indonesian gardens. Clones of two other horticultural groups, *S. barberi* and *S. sinense*, probably natural hybrids between *S. officinarum* and *S. spontaneum* (Brandes 1956), were also cultivated. Stimulated by the discovery of sexual fertility in Barbados at the beginning of this century and the ravages caused by the 'sereh' disease at the same period, breeders initiated interspecific hybridisation programs around 1920 in Java and in India. These largely involved *S. officinarum* and *S. spontaneum* clones. The hybrid progenies were repeatedly backcrossed with *S. officinarum* to minimise the negative effect of the wild parent. This procedure is referred to as 'nobilisation' in sugarcane breeding. During nobilisation an asymmetric chromosome transmission occurred: the F_1 as well as the BC_1 hybrids received the somatic chromosome number from the female parent, *S. officinarum* whereas they received a gametic complement from the male parent. The result was a rapid decrease in the number of *S. spontaneum* chromosomes in the hybrids (Bremer 1961). The transmission only became normal starting from the BC_2 generation. Only a few *S. officinarum* and *S. spontaneum* clones, either directly or through *S. barberi*, were used to create the first interspecific hybrids (Arcenaux 1967). The material derived from this process proved successful in increasing productivity and resolving some of the disease problems. Until recently, sugarcane breeding has relied essentially on intercrossing some of these first nobilised clones, produced in Java and India, and their derived progenies. Consequently, current cultivars are characterised by a high chromosome number (100–130) with around 80% of them derived from *S. officinarum* and 20% from *S. spontaneum* (Price 1957; Roach 1969; D'Hont et al. 1996) and are highly polyploid

like their ancestors. A genetic map of the cultivar 'R570' indicates that recombination is as frequent in sugarcane as it is in common related diploid species (Grivet et al. 1996). There is occasional preferential pairing among chromosomes derived from *S. spontaneum* (Jannoo et al. in preparation), although some interspecific intrachromosomal recombination exists (Grivet et al. 1996; D'Hont et al. 1996). The *S. officinarum*-derived fraction of the genome displays chromosome assortments that are incompatible with a complete disomy but with significant and uneven deviations from polysomy depending on the genome region (Jannoo et al. in preparation). Given that breeders usually make crosses between proven genotypes and that 12–15 years are necessary for complete evaluation before variety release, modern cultivars result from less than 10 meioses (between 5 and 7) from the first interspecific crosses.

Modern cultivars may therefore have kept traces of the initial genotypic structure of the first hybrids. A previous study has detected the existence of linkage disequilibrium among cultivars (Lu et al. 1994b). However, new information available from the map of cv 'R570' (Grivet et al. 1996) indicates that these disequilibria involved the mostly tightly linked loci. Here, we report a restriction fragment length polymorphism (RFLP) analysis of modern cultivars using probes regularly scattered over the sugarcane genome to determine the extent to which linkage disequilibrium persists in cultivars.

Materials and Methods

Plant material

The material surveyed consisted of 59 cultivars (Table 1) derived mainly from the Mauritian breeding programme. A few clones bred in different breeding stations were also included: 7 clones were from Reunion island, 2 from Barbados, 1 each from Australia, Hawaii and Taiwan and 8 cultivars corresponding to the first interspecific hybrids created mainly in Java and India at the beginning of this century. These cultivars have all been commercially exploited in Mauritius at one time or another. They were obtained from the Mauritius Sugar Industry Research Institute (MSIRI) in Mauritius.

DNA isolation and restriction

The procedures for extraction of genomic DNA and Southern hybridisation have been presented elsewhere (Hoisington 1992; D'Hont et al 1994). DNA was restricted with two of the following three restriction enzymes, *Dra*I, *Hin*dIII and *Sst*I. Restricted DNA samples (10 µg/lane) were separated by electrophoresis in 0.8% TAE agarose gels at 1.7 V/cm for 24 h and blotted overnight on a nylon membrane (Hybond N+).

Probes

Thirty-eight sugarcane probes were selected, taking advantage of the map of sugarcane cultivar 'R570' (Grivet et al. 1996). Probes were chosen to be regularly distributed over each of the nine linkage groups of this map and to yield hybridisation patterns with strong well-resolved bands. They were derived from sugarcane genomic DNA libraries (SSCIR and SG probes) and from a cDNA library (CDS probes). SSCIR probes were developed at CIRAD (Grivet et al. 1996). SG and CDS probes were kindly supplied by Drs. J. da Silva and W. Burnquist (Copersucar, Brazil), Dr. P. Moore (USDA-HSPA, USA) and Dr. M. Sorrells (Cornell University, USA) (da Silva et al. 1993). All but 3 of these probes were

Table 1 List of cultivars together with their breeding station

Identification	Clones	Breeding Station
Mauritian cultivars		
1	M134/32	Mauritius
2	M213/40	Mauritius
3	M241/40	Mauritius
4	M147/44	Mauritius
5	M31/45	Mauritius
6	M202/46	Mauritius
7	M93/48	Mauritius
8	M305/51	Mauritius
9	M356/53	Mauritius
10	M13/56	Mauritius
11	M377/56	Mauritius
12	M351/57	Mauritius
13	M555/60	Mauritius
14	M907/61	Mauritius
15	M574/62	Mauritius
16	M2173/63	Mauritius
17	M376/64	Mauritius
18	M657/66	Mauritius
19	M1156/66	Mauritius
20	M3035/66	Mauritius
21	M695/69	Mauritius
22	M292/70	Mauritius
23	M1205/70	Mauritius
24	M1557/70	Mauritius
25	M298/71	Mauritius
26	M1030/71	Mauritius
27	M1236/71	Mauritius
28	M50/75	Mauritius
29	M134/75	Mauritius
30	M536/75	Mauritius
31	M791/75	Mauritius
32	M1077/75	Mauritius
33	M1176/77	Mauritius
34	M52/78	Mauritius
35	M624/78	Mauritius
36	M725/78	Mauritius
37	M1658/78	Mauritius
38	M2077/78	Mauritius
39	M2229/80	Mauritius
Imported cultivars		
40	B34104	Barbados
41	B3337	Barbados
42	H32-8560	Hawaii
43	Q96	Australia
44	R472777	Reunion island
45	R570	Reunion island
46	R575	Reunion island
47	R576	Reunion island
48	R577	Reunion island
49	R578	Reunion island
50	R579	Reunion island
51	S17	Taiwan
First Indian cultivars		
52	Co213	India
53	Co281	India
54	Co290	India
55	Co312	India
56	NCo310	India
57	NCo376	India
First Indonesian cultivars		
58	POJ2878	Java
59	PT4352	Java

present in one copy on the sugarcane genome; the 3 exceptions were duplicated on two different linkage groups. The probes were labelled with [^{32}P]-dCTP (ICN Pharmaceuticals) using the Amersham Megaprime labelling kit.

Statistical analysis

For each probe/enzyme combination, the polymorphic markers found over all cultivars were scored as 1 for presence and 0 for absence. These bands are expected to mark as unambigously as possible the allelic molecular sequences borne by the original founders of the varietal population. These sequences – termed here as 'alleles' – are present in discrete numbers given the bottleneck described in the Introduction. Various types of analyses were performed in to answer a sequence of specific questions:

Can consanguinity among subsets of cultivars confuse the analysis?

A factor analysis of correspondences (FAC) (Benzecri 1973) was performed on the clones using all of the polymorphic markers identified in order to provide a synthetic representation of the diversity between sugarcane cultivars. In this analysis, each marker was considered as a variable, and each cultivar was characterised by all the markers it displayed. The FAC identifies several independent axes, each one successively accounting for the largest part of the variation. These axes are linear combinations of the variables, and each cultivar can be located along the axes. Very rare or frequent markers (frequency below 5% or over 95%, respectively) can be placed as inactive variables so as not to unbalance the analysis. The loading of the variables on the axes measures their discriminative power. The distribution of the cultivars along the axes was examined in order to detect contingent varietal groups within the germplasm under survey. Markers that were the main contributors to each of the first three axes were identified. Their linkage relationships along the map were analysed in order to evaluate the extent to which disequilibrium takes part in the structure among the cultivars.

How to detect the significant associations among markers?

In order to test the linkage disequilibrium between the various loci analysed, a Fisher exact test (Mehta and Patel 1983; SAS Institute 1990, FREQ procedure) was performed systematically using all of the polymorphic markers identified. This analysis consists of testing, at a chosen significant threshold, the hypothesis of independent distribution between 2 markers. For each pair of markers a 2×2 contingency table (presence versus absence) was constructed. The probability corresponds to the probability of obtaining the observed contingency table as well as tables which are less probable in the hypothesis of independence, thereby fixing the marginal distributions (Sokel and Rohlf 1981). The significance threshold over which the independence hypothesis was refuted was 5%. However, as more than one comparison was tested, the threshold was modified by dividing 5% by the number of comparisons performed in order to discard random associations. This threshold was unmodified when the Bonferroni test (Rice 1988) was applied due to the important number of comparisons computed.

Is physical linkage the essential source of associations?

In the first phase, the distribution of probes yielding significant association on the map was analysed to test whether the associated markers derived preferentially from loci residing on the same linkage group. In this phase, the significance threshold was determined using the total number of comparisons performed.

In the second phase, the objective was to consider each linkage group separately and to test whether strong bilocus associations appeared more often with closely linked loci. The significance threshold was thus estimated individually for each group on the basis of the number of comparisons performed.

How resolving is the molecular technique used for the analysis?

A particular set of data was considered separately: the associations between markers derived from the same probe. In this case, it is expected that strong associations will correspond to markers that reveal the same allele; these associations will be complete unless at least 1 of the 2 markers reveals more than one allele. Incomplete associations will pinpoint cases where RFLPs, with the restriction enzymes used, are unable to resolve two or more distinct alleles. This result will give an indication on the resolving power of the method, *i.e.* the frequency of the cases where the method actually enables accurate testing of the linkage disequilibrium.

Is linkage disequilibrium evenly distributed among the two specific genome components?

The *S. officinarum* or *S. spontaneum* origin of some of the alleles in disequilibrium was roughly identified. In a previous study, the *S. officinarum* part of cultivars' genome was investigated (Jannoo et al. 1999), and data are available for 12 out of the 38 probes used here; they were hybridised to a sample of 53 *S. officinarum* clones and 109 cultivars. The origin of the markers was identified based on the criteria defined in this previous study (Jannoo et al. 1999): *i.e.* a marker present in at least 2 *S. officinarum* clones was considered to be a *S. officinarum* marker, otherwise it was considered as a *S. spontaneum* marker.

Results

General structure

A complete set of data was obtained for 72 RFLP probe/enzyme combinations revealing 1057 polymorphic markers over the 59 cultivars. The distribution of the frequency of the markers is presented in Fig. 1. Although 25% of the markers appeared in fewer than 10% of the cultivars, a rather homogeneous distribution of the markers was observed among the various frequency classes. A factor analysis of correspondences (FAC) was carried out on 920 markers and 59 cultivars. Axes 1 and 2 of the FAC explained 6.5% and 4.7% of the variability, respectively. While there was no particular structure to the distribution of the cultivars, which were homogeneously distributed along the first two axes (Fig. 2), the general trend corresponded to a separation between cultivars bred in Mauritius and those imported from other breeding stations including old Indian and Indonesian hybrids and cultivars bred in Reunion. The separation was not complete since some non-Mauritian cultivars were also placed together with Mauritian ones, and vice versa. It is also noteworthy that those which were most extreme on the third axis (explaining 4.2% of the variability, data not shown) were highly related, with all having at least one common parent.

The markers contributing the most to the first three axes were identified using their loading on each axis. On axis one, 12 markers were identified as corresponding to 8 probes scattered over 6 of the nine linkage

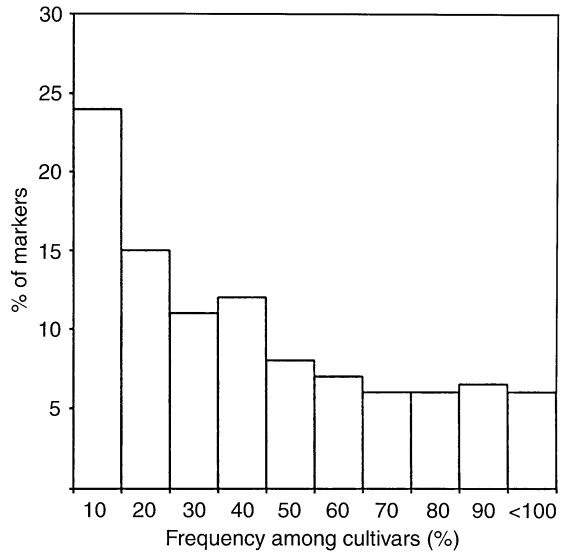

Fig. 1 Distribution of the 1057 polymorphic markers based on their frequency among the 59 cultivars

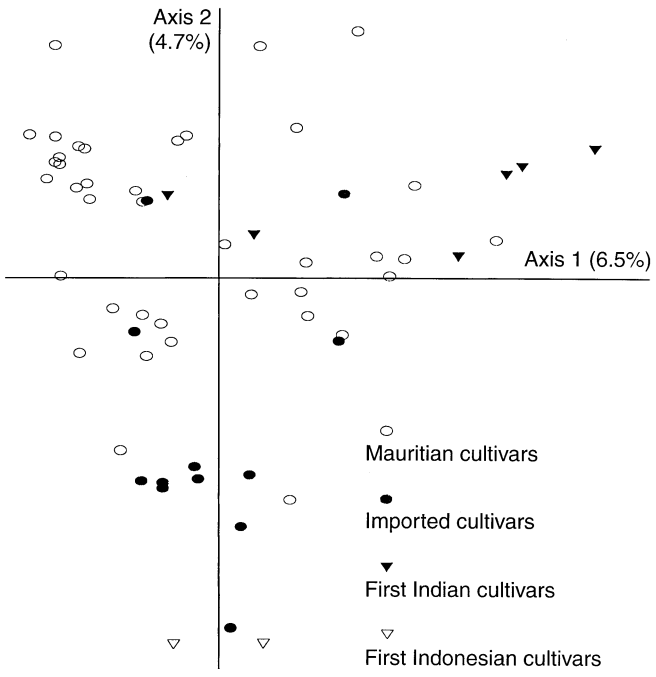

Fig. 2 First plane of a factor analysis of correspondences (FAC) among 59 cultivars using 920 active markers

groups. On the second axis, 15 markers revealed by 9 probes and scattered over four linkage groups were the main contributors. The third axis was essentially explained by 23 markers corresponding to 16 probes and, conversely of axes one and two, some of the markers contributed negatively to axis 3, and some positively: they involved 8 probes spread over seven linkage groups and 12 probes scattered on six linkage groups, respectively. For each of the three axes, the distribution of marker pairs derived from different probes, according to genetic distance, is illustrated in Table 2. The first axis was essentially defined by independent markers. On the second and third axes, although many independent probes were involved, more pairs of physically linked markers were observed, suggesting that linkage disequilibrium related to physical linkage starts to play a significant role in the determination of multilocus genetic diversity.

Linkage disequilibrium

Bilocus associations.

A total number of 540,688 2×2 Fisher exact tests was performed corresponding to all possible 2×2 comparisons of polymorphic markers revealed by the different probes. The individual significance threshold was $P=8.96\times10^{-8}$ for each comparison to permit an overall threshold of $P=0.05$. A total of 59 cases of significant association was recorded, depicting disequilibrium in the population analysed. Of these, 51 involved couples of loci which were physically linked on the map of cv 'R570', these loci were scattered over all the sugarcane genome. In the other 8 cases, the markers in disequilibrium corresponded to loci mapped on different linkage groups, there by excluding physical linkage as an explanation. Most of these markers were not among the main contributors of the first three axes of the FAC; only 3 of them were among those explaining the first axis. Therefore, germplasm relatedness has a limited responsability for their existence. These few unlinked loci could be associated in this sample by chance.

The distribution of the number of linkage disequilibria observed based on the distance between loci is presented in Fig. 3. Almost 75% corresponded to loci less than 10 cM from one another. In 2 other instances, the markers represented loci separated by more than 90 cM, making it doubtful that the associations are due to linkage. Altogether, more than 80% of the associated markers involved loci separated by less than 30 cM from one

Table 2 Number of pairs of markers which contribute the most to the first three axes of the FAC based on the distance of their corresponding loci

Axes	Distance (cM) of linked markers					Total	Number of unlinked markers
	0–20	20–40	40–60	60–80	>80		
1	2	–	1	–	–	3	56
2	6	11	–	3	9	29	69
3	7	4	–	–	2	13	107

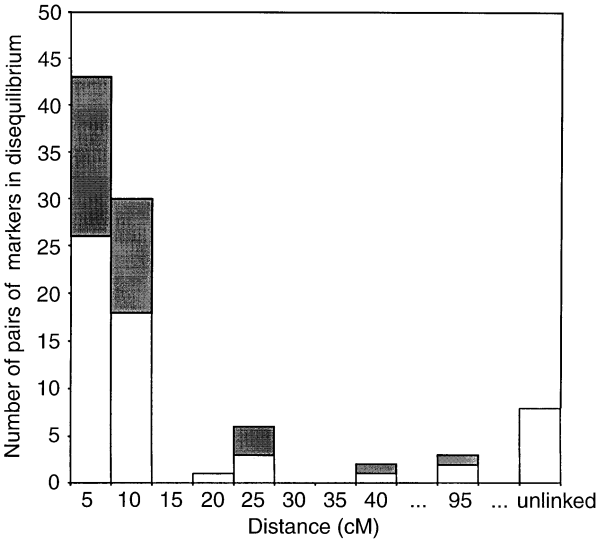

Fig. 3 Distribution of number of associated pairs of markers according to the distance. *Plain bars* correspond to associations detected in the global analysis. Additional associations when only markers of a same linkage group were analysed at a time are shown as *shaded bars*

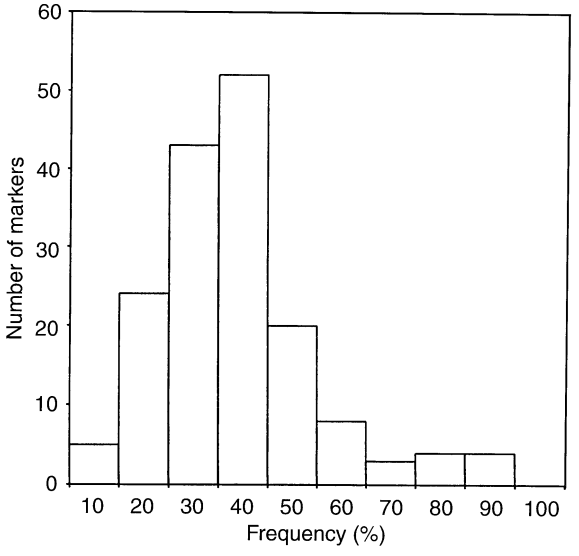

Fig. 4 Distribution of the frequency of the markers involved in linkage disequilibrium

another, thus showing that linkage is the main phenomenon that causes the existence of disequilibria.

The data were then analysed separately for each linkage group, leading to specific significance thresholds depending on the number of comparisons. Eighty-five linkage disequilibria involving 2 distinct probes were recorded, 34 of which had not been observed in the first of analysis. Most of them involved tightly linked loci (Fig 3).

All marker pairs detected as being in disequilibrium were in a coupling phase, the association often being complete. The loosest situation found in the global analysis involved 11 cases of dissociations (1 marker was present while the other was absent) compared to 25 cases of associations (2 markers being either both present or both absent). A similar case with 12 instead of 11 dissociations displayed a probability higher than the threshold in the global analysis but lower when the analysis was carried out separately on each linkage group. The distribution of the frequency of the markers involved in disequilibrium is presented in Fig. 4. Most of them were present in fewer than half of the cultivars; the majority being present between 30% and 40%of the time.

After analysis of tri- and tetra-marker associations, the 85 original pairs of cosegregating markers were reduced to 42 different groups involving 33 loci scattered over all the entire sugarcane genome (Fig 5). Six of these involved alleles at more than 2 loci, 5 groups implicated 3 loci and 1 group 4. While 80% of the disequilibrium concerned loci which were less than 10 cM apart, a few disequilibria were also observed between loci separated by up to 90 cM.

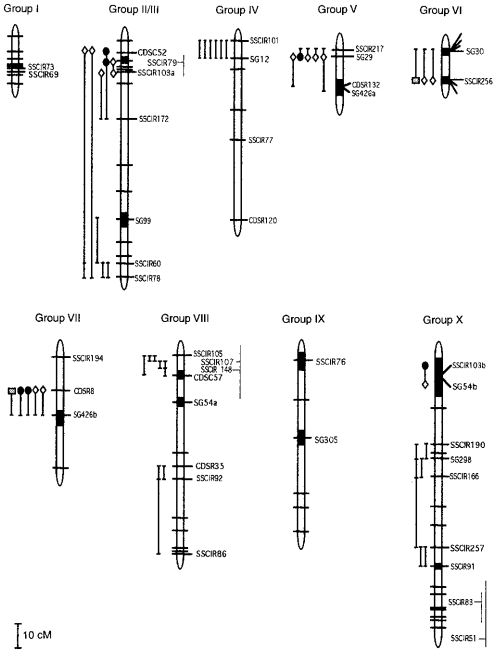

Fig. 5 Linkage disequilibrium detected among 59 cultivars and represented along the composite map of 'R570' (Grivet et al. 1996). Linkage groups II and III are linked (D'Hont et al. unpublished data). The sugarcane map is therefore represented by nine chromosomes. The 38 probes used, corresponding to 41 loci, are shown on the map. Where probes are duplicated, each locus is identified by a *letter*, *a* or *b*. Uncertain positions of isolated loci are indicated by a *T bar* to the *right* of the linkage group. Uncertain orders of clustered loci are denoted in *black* on the chromosome. Loci between which linkage disequilibria were observed are indicated by a *bar* to the *left* of the linkage group. *Different bars* between the loci indicate that several alleles at these loci are in disequilibrium. The origin of the alleles are indicated when known: ● for *S. officinarum*, ◇ for *S. spontaneum*, □ when 2 markers corresponding to the same allele display different origin, *S. spontaneum* in one case and *S. officinarum* in the other

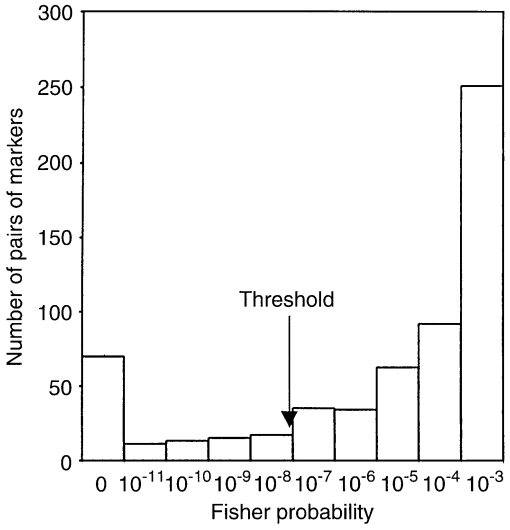

Fig. 6 Distribution of pairs of markers yielded by the same probe according to their Fisher probability. The threshold for analysis considering all markers together is indicated

Monolocus associations

In relation to the high ploidy level of sugarcane, an average of 15 markers were revealed per probe/enzyme combination. In total, 17,408 comparisons involving markers derived from the same probe were considered. The distribution of marker pairs according to the Fisher probability is presented in Fig. 6. A continuous distribution was observed for probabilities of 0–0.01. Seventy cases of complete association were recorded (type a in Table 3), both markers most likely corresponding to only one allele. In addition, other cases of strong but incomplete associations were noted. For example, 44 pairs were revealed under a Fisher probability of 10^{-7}, to correspond approximately to the threshold retained when all the markers were analysed together. Among them, there were 25 cases in which all the individuals that displayed 1 marker also displayed the other, but where the absence of 1 marker was not always associated with the absence of the second (type b in Table 3); the second marker, therefore, probably revealed the same allele as the first but also other alleles. There were 19 cases where the dissociation could be of either situation (type c in Table 3); both markers probably revealed several alleles each, some being common, others being specific.

The results illustrate the limited frequency of optimal conditions for the detection of linkage disequilibrium – *i.e.* when 1 marker corresponds to one allele – in relation to less favourable conditions. A case like that of markers C060H09 and C060D10 (Table 3), which are at the fringe of significance according to the threshold retained for the first step of the bilocus analysis, illustrates how a partial allele resolution by 1 of the markers can mask an actual full association.

Origin of the associations

All of the associations observed within the whole sample corresponded to markers present in at least one of the old cultivars used as progenitors in the breeding programme in Mauritius. The specific origin of 24 alleles involved in some disequilibrium could be examined. Fifteen alleles were found to be likely derived from *S. spontaneum* and only 7 were from *S. officinarum*. In 2 cases there was ambiguity; 2 markers identified by two different restriction enzymes and related to the same allele displayed contradictory specificity on the basis of the criteria used (Fig 5).

Discussion

The aim of this study was to test whether the breeding history of sugarcane has installed a general linkage disequilibrium still perceptible today. This is essential to know for future applications of common genetic analyses of traits of agricultural interest.

The material under survey was chosen to represent a homogeneous set of commercially cultivated clones. A previous study has shown the existence of general differences in the RFLP patterns of cultivars of different geographical origin (Jannoo et al. 1999). This could be related to various specificities such as particular breeding habits or adaptation to local environments. These differences could result in disequilibrium that is not due to genetic linkage. Therefore, the material was chosen mainly from one breeding programme that, conducted in Mauritius, complemented by foreign cultivars with either similar adaptations (like those from Reunion) or involved in the genealogy of most Mauritian cultivars. Most of the genotypes surveyed are therefore related to each other to various degrees; this is confirmed by the homogeneous distribution of the markers in the different frequency classes and by the weakness of the structure observed among the clones under survey.

The conditions were not optimal for detecting disequilibrium. Due to both the high polyploidy of sugarcane cultivars and the limited resolution by the two restriction enzymes used, it was common to find that 1 marker corresponded to several alleles, which confounds the detection of contingent disequilibrium. Despite these hazy conditions, many instances of disequilibrium were uncovered. This phenomenon involved markers usually present in 30–40% of the cultivars and yielded by loci altogether scattered over the entire sugarcane genome. Increasing the number of restriction enzymes can improve the resolution of the different alleles and the detection of linkage disequilibrium in target regions; for example, we were able to reveal up to ten distinct alleles in a study on chromosome assortment in the self-progeny of a single cultivar (Jannoo et al. in preparation). Associations between physically linked loci are thus a reality in this sample, and there exists considerable scope for refining their detection in the future.

Table 3 Examples of pattern associations among allelic markers

Situation[a]	Marker	Cultivars																													
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30
(a)	C101S11	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	1	1	0	0	0	0	1
	C101D21	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	1	1	0	0	0	0	1
(b)	C060H09	1	1	0	1	0	1	1	1	1	0	0	1	1	0	1	1	1	1	0	1	0	0	1	1	0	0	0	0	0	1
	C060D10	1	1	0	1	0	1	1	1	1	0	0	0	1	0	0	1	0	1	0	1	0	0	1	1	0	0	0	0	0	1
(c)	C107S08	0	1	1	1	0	0	1	0	1	1	0	1	1	0	1	0	1	0	0	0	0	1	1	0	0	0	0	1	0	0
	C107D03	0	1	1	1	0	0	1	0	1	1	0	1	1	0	1	0	1	0	0	0	0	1	1	1	0	0	0	0	0	0

Table 3 Continued

Situation[a]	Marker	Cultivars																													Fisher
		31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	Probability
(a)	C101S11	0	0	0	0	0	1	1	0	0	0	1	0	0	0	1	0	0	0	0	0	1	0	0	0	1	1	1	0	0	0
	C101D21	0	0	0	0	0	1	1	0	0	0	1	0	0	0	1	0	0	0	0	0	1	0	0	0	1	1	1	0	0	
(b)	C060H09	1	1	0	1	0	1	1	1	1	0	0	1	1	0	1	0	1	1	0	1	1	1	1	1	1	1	1	1	1	$7{,}83 \cdot 10^{-8}$
	C060D10	0	0	0	1	0	0	0	1	1	0	0	1	1	0	1	0	1	1	0	1	1	1	1	0	0	0	0	1	1	
(c)	C107S08	0	0	1	1	1	1	1	1	0	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1	1	1	0	0	0	$2{,}30 \cdot 10^{-10}$
	C107D03	0	0	1	1	1	1	1	0	0	0	0	0	1	0	0	0	1	1	1	1	1	1	1	1	0	0	0	0	0	

[a] (a) The two markers revealed by the same probe are completely associated. Each marker most likely correspond to one and only one allele. (b) The two markers are dissociated in several cultivars. Marker C060H09 probably reveals one allele more than C060D10. (c) Both markers probably reveal several alleles each, some being common, others being specific

All the marker combinations in disequilibrium were present in at least 1 of the first interspecific hybrids included in this study, thus pointing to the likely donors of the chromosome segments that bear the markers. The loci involved are generally separated by less than 10 cM on the reference genetic map. The few associations between distantly linked or unlinked loci may correspond partly to escapes through the statistical screening for significance and partly to associations derived from the recurrent use of the first Indian interspecific hybrids in the breeding programme in Mauritius. The whole figure gives an indication that 10 cM is the approximate genetic size of the chromosome segments that can be marked throughout the genealogy of modern cultivars.

It is noteworthy that two thirds of the alleles involved in disequilibrium (and whose origin was determined) were derived from *S. spontaneum*, although only about 20% of the genome in cultivars is inherited from this species. This is probably related to the higher resolution of molecular markers on the part of the genome that is both less represented and more polymorphic (Lu et al. 1994a), thus less prone to marker ambiguity. A parallel observation was made while mapping the genome of a cultivar (Grivet et al. 1996). It is also noteworthy that the significant multiallelic associations that relate the most losely linked loci comprise *S. spontaneum*-specific alleles. This may be connected to other observations on the mode of chromosome assortment and the frequency of intrachromosomal interspecific recombination in cultivars. There is increasing evidence that chromosomes contributed by the same species tend to pair together (Grivet et al. 1996; Jannoo et al. in preparation) and that, probably as a result, intrachromosomal interspecific recombination has been relatively rare (D'Hont et al. 1996). It is possible that linkage disequilibrium will be stronger within a genome compartment that is more restricted, possibly being occasionally limited to a single chromosome per homology group, such as the *S. spontaneum* compartment. This hypothesis is worth testing in future studies, because *S. spontaneum* chromosomes are thought to bear many genes of particular agricultural interest.

The scope of application of our findings resides essentially in the possibility of tracking alleles at quantitative trait loci (QTLs) in different cultivars once the associated markers have been identified. Conversely, it will be interesting to specifically intensify the study of molecular diversity in the vicinity of loci where some QTLs of interest have been shown to reside. A survey of cultivars with several probes mapped in such a region in combination with a high number of restriction enzymes may have a high resolution power. Many alleles may be revealed for each locus and may give access to most or all the multilocus haplotypes present in the few ancestral clones at the origin of modern cultivars. There may be opportunities to relate this molecular diversity to the variation scored for traits of value for plant breeders and sugarcane growers.

References

Arceneaux G (1967) Cultivated sugarcanes of the world and their botanical derivation. Proc Int Soc Sugarcane Technol 12: 844–854

Benzecri JP (1973) L'analyse des correspondances. (L'analyse de données, tome 2). Dunod, Paris

Brandes EW (1956) Origin, dispersal and use in breeding of the Melanesian garden sugarcanes and their derivatives, saccharum officinarum l. Proc Int Soc Sugarcane Technol 9:709–750

Bremer G (1930) The cytology of *Saccharum*. Proc Int Soc Sugar Cane Technol 3:408–415

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

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

D'Hont A, Lu YH, Gonzàles de Leòn D, Grivet L, Feldmann P, Lanaud C, Glaszmann JC (1994) A molecular approach to unraveling the genetics ofsugarcane, a complex polyploid of the Andropogoneae tribe. Genome 37:222–230

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

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

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

Jagathesan D, Sreenivasan TV, Ratnambal MJ (1970) Chromosome numbers in noble sugarcanes. Indian J Hered 2: 89–96

Li HW, Price S (1967) Chromosome numbers of noble sugarcane clones. Proc Int Soc Sugar Cane Technol 12:884–886

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

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

Metha CR, Patel NR (1983) A network algorithm for performing Fisher's exact test in r x c contingency tables. J Am Statist Assoc 78:427–434

Price S (1957) Cytological studies in *Saccharum* and allied genera. III. Chromosome numbers in interspecific hybrids. Bot Gaz 118:146–159

Rice W (1988) Analyzing tables of statistical tests. Evolution 43:223–225

Roach BT (1969) Cytological studies in *Saccharum*. Chromosome transmission in interspecific and intergeneric crosses. Proc Int Soc Sugar Cane Technol 13:901–920

SAS Institute (1990) SAS procedures guide, version 6, 3rd edn. SAS Institute Inc,Cary, N.C.

Sokal RR, Rohlf FJ (1981) Biometry. WH Freeman, San Francisco

Sreenivasan TV, Ahloowalia BS, Heinz DJ (1987) Cytogenetics. In: Heinz DJ (ed) Sugarcane improvement through breeding. Elsevier, Amsterdam Oxford New York Tokyo, pp 211–253