CrossMark

REVIEW

QTL mapping including codominant SNP markers with ploidy level information in a sugarcane progeny

E. A. Costa · C. O. Anoni · M. C. Mancini · F. R. C. Santos ·
T. G. Marconi · R. Gazaffi · M. M. Pastina · D. Perecin ·
M. Mollinari · M. A. Xavier · L. R. Pinto · A. P. Souza ·
A. A. F. Garcia

Abstract Quantitative trait locus (QTL) mapping contributes to sugarcane (*Saccharum* spp.) breeding programs by providing information about the genetic effects, positioning and number of QTLs. Combined with marker-assisted selection, it can help breeders reduce the time required to develop new sugarcane varieties. We performed a QTL mapping study for important agronomic traits in sugarcane using the composite interval mapping method for outcrossed species. A new approach allowing the 1:2:1 segregation ratio and different ploidy levels for SNP markers was used to construct an integrated genetic linkage map that also includes AFLP and SSR markers. Were used 688 molecular markers with 1:1, 3:1 and 1:2:1 segregation ratios. A total of 187 individuals from a biparental cross (IACSP95-3018 and IACSP93-3046) were assayed across multiple harvests from two locations. The evaluated yield components included stalk diameter (SD), stalk weight (SW), stalk height (SH), fiber percentage (Fiber), sucrose content (Pol) and soluble solid content (Brix). The genetic linkage map covered 4512.6 cM and had 118 linkage groups corresponding to 16 putative homology groups. A total of 25 QTL were detected for SD (six QTL), SW (five QTL), SH (four QTL), Fiber (five QTL), Pol (two QTL) and Brix (three QTL). The percentage of phenotypic variation explained by each QTL ranged from 0.069 to 3.87 %, with a low individual effect because of the high ploidy level. The mapping model provided estimates of the segregation ratio of each

E. A. Costa, C. O. Anoni, and M. C. Mancini have contributed equally to this work.

E. A. Costa · M. C. Mancini · A. P. Souza
Centro de Biologia Molecular e Engenharia Genética (CBMEG), Departamento de Genética e Evolução, Universidade Estadual de Campinas (UNICAMP), Cidade Universitária ZeferinoVaz, CP 6010, Campinas, SP 13083-875, Brazil

C. O. Anoni · T. G. Marconi · M. Mollinari ·
A. A. F. Garcia (✉)
Departamento de Genética, Escola Superior de Agricultura Luiz de Queiroz (ESALQ), Universidade de São Paulo (USP), CP 83, Piracicaba, SP 13400-970, Brazil
e-mail: augusto.garcia@usp.br

F. R. C. Santos · M. A. Xavier · L. R. Pinto
Centro Avançado da Pesquisa Tecnológica do Agronegócio de Cana, IAC/Apta, CP 206, Ribeirão Prêto, SP 14001-970, Brazil

R. Gazaffi
Centro de Ciências Agrárias, Universidade Federal de São Carlos, Rodovia Anhanguera, Km 174, Araras, SP, Brazil

M. M. Pastina
Núcleo de Biologia Aplicada, Embrapa Milho e Sorgo, Rodovia MG 424, Km 65, Sete Lagoas, MG 35701-970, Brazil

mapped QTL (1:2:1, 3:1 or 1:1). Our results provide information about the genetic organization of the sugarcane genome and constitute the first step toward a better dissection of complex traits.

Introduction

Sugarcane is a complex autopolyploid and outbred species with a high level of heterozygosity. Modern sugarcane varieties are derived from the interspecific hybridization of *Saccharum officinarum* (2n = 80) and *Saccharum spontaneum* (2n = 40–128), resulting in highly polyploid and aneuploid plants. The introduction of this hybridization scheme represents a large breakthrough in modern sugarcane breeding, solving disease problems, providing increased yields and, adaptability to grow under several environmental conditions and improving ratooning ability (Paterson et al. 2013).

Genetic linkage map construction and QTL mapping provide information about the genetic architecture of traits, linkage and pleiotropy (Zeng et al. 1999).

D. Perecin
Universidade Estadual Paulista (UNESP), Departamento de Ciências Exatas, Jaboticabal, SP, Brazil

A. P. Souza
Departamento de Biologia Vegetal, Instituto de Biologia, Universidade Estadual de Campinas, Campinas, SP CEP 13083-892, Brazil

Present Address:
E. A. Costa
Federal de São Paulo (UNIFESP), unidade Parque Tecnológico, Avenida Cesare Mansueto Giulio Lattes, 1201, Eugênio de Mello, São José dos Campos, SP CEP 12247-014, Brazil

Present Address:
T. G. Marconi
Texas A&M University System, Agrilife Research and Extension Center, 2415 E. Hwy 83, Weslaco, TX 78596, USA

Present Address:
M. Mollinari
North Carolina State University, Bioinformatics Research Center, 332 Ricks Hall,, Raleigh, NC 27607, USA

However, the construction of genetic maps in sugarcane is more complicated and laborious than in diploid species because (i) the high level of polyploidy and aneuploidy results in complex chromosomal segregation patterns during meiosis (Heinz and Tew 1987); (ii) mapping progeny are derived from bi-parental crosses between highly heterozygous outbred parents, in which there are different numbers of alleles per locus, resulting in a mixture of marker segregation patterns in the progeny (Wu et al. 2002; Lin et al. 2003; Pastina et al. 2010); and (iii) the linkage phase between markers and between markers and QTL is unknown (Pastina et al. 2012; Gazaffi et al. 2014).

Wu et al. (1992) proposed the development of genetic linkage maps based solely on a segregation analysis of single-dose markers (SDM). These markers represent alleles that are present in one copy in one of the parents and segregating in a 1:1 ratio in the progeny, or in one copy in both parents and segregating in a 3:1 ratio. The first genetic maps were published by da Silva et al. (1993) and Al-Janabi et al. (1993) using a population derived from a cross between the double-haploid 'ADP85-0068' and *S. spontaneum* 'SES208', trying maximize SDM detection. In general, the majority of published sugarcane linkage maps have been constructed based on SDM and on the double *pseudo-testcross* strategy (Grattapaglia and Sederoff 1994), which results in two individual maps, one for each parent (Daugrois et al. 1996; Ming et al. 2001, 2002a; Hoarau et al. 2002; Reffay et al. 2005; Aitken et al. 2006, 2008; Al-Janabi et al. 2007; as reviewed by Pastina et al. 2010; Shing et al. 2013). Garcia et al. (2006) presented an integrated sugarcane linkage map that was constructed based on the methodology proposed by Wu et al. (2002) and combined information provided by markers segregating in 1:1 and 3:1 ratios, resulting in an integrated linkage map for both parents. Integrated linkage maps are advantageous because more saturated maps are obtained, allowing for better estimates of QTL locations and the ability to estimate linkage and linkage phases more accurately. This approach was also used for the sugarcane linkage map presented by Oliveira et al. (2007), Palhares et al. (2012) and Pastina et al. (2012).

To date, approximately 22 linkage maps (reviewed by Alwala and Kimbeng 2010 and Pastina et al. 2010) for sugarcane exist. The majority of these maps were constructed using molecular markers that are read as

dominant markers in polyploids, such as restriction fragment length polymorphism (RFLP), random amplified polymorphism (RAPD), amplified fragment length polymorphism (AFLP), and simple sequence repeat (SSR) markers. SSRs behave like dominant markers in complex polyploids because they do not allow the identification of the different alleles or the marker dosage (Garcia et al. 2013). Modern technologies, such as the Sequenom iPLEX MassARRAY® (Sequenom Inc., San Diego, CA, USA), allow for the evaluation of single-nucleotide polymorphisms (SNPs) throughout the genome considering the relative abundance of each allele, i.e., the allelic dosage (Serang et al. 2012; Garcia et al. 2013; Bargary et al. 2014; Mollinari and Serang 2015). Knowledge of the dosage and ploidy level of an SNP can significantly increase the information imparted by each locus and provides several advantages for genetic analysis. None of the published sugarcane genetic linkage maps are believed to be saturated, and the only map made with SNP markers uses only single dosages and does not consider ploidy level (Aitken et al. 2014).

Most of the economically important agronomic traits have a quantitative nature with polygenic inheritance (Falconer and Mackay 1996; Lynch and Walsh 1998) that is highly influenced by environmental conditions. The first QTL mapping study in sugarcane was performed by Sills et al. (1995). Other works studies making important contributions to QTL mapping include Ming et al. (2001, 2002a, b), Hoarau et al. (2002), McIntyre et al. (2005), da Silva and Bressiani (2005), Reffay et al. (2005), Aitken et al. (2006, 2008), Piperidis et al. (2008), Pastina et al. (2012), and Shing et al. (2013). Specifically, Pastina et al. (2012) proposed a mixed model for QTL mapping in sugarcane in order to include complex G × E interactions. Unlike previously reported studies, these authors considered the presence of genetic correlations between locations and harvests by testing a specific variance and covariance matrix for the genetic effects. In this sense, it was possible to consider genotype-by-environment, QTL-by-location, QTL-by-harvest and QTL-by-harvest-by-location interactions and thus to distinguish stable-effect QTL from those with interaction effects, contributing to a better understanding of the genetic basis of important traits of sugarcane.

Recently, Gazaffi et al. (2014) developed an approach for composite interval mapping (CIM) in full-sib crosses. In this method, linkage phase and segregation patterns are determined in addition to QTL location, allowing for the simultaneous analysis of QTLs with different segregation patterns. Additionally, Gazaffi et al. (2014) included markers as cofactors into the mapping procedure, enabling more precise estimates of putative QTL locations in the genome. This provides a useful framework for QTL mapping in sugarcane based on integrated maps.

In this study, we constructed an integrated genetic linkage map from a bi-parental cross between two Brazilian commercial sugarcane varieties with SSRs, AFLPs and SNPs. The SNP calling approach provides information about ploidy level, which could be used to infer the size of identified homology groups. The SNPs are also displayed as codominant markers (1:2:1 segregation ratio), which have never before been included in a sugarcane genetic map. These markers are more informative than the ones segregating in 1:1 or 3:1 patterns. We then performed QTL mapping using CIM after evaluating the field trials using mixed models. This methodology permits the reporting of the segregation pattern, linkage phase and additive and dominance effects for each putative QTL.

Materials and methods

Mapping population and field experiments

Full-sib progeny were obtained from a bi-parental cross between the elite clone IACSP95-3018 (female parent) and the variety IACSP93-3046 (male parent), which were developed by the Sugarcane Breeding Program at the Instituto Agronômico de Campinas (IAC). IACSP95-3018 is a promising clone that is also used as parent in the IAC Sugarcane Breeding Program, whereas IACSP93-3046 has a high level of sucrose, resistance to rust, good tillering and an erect stool habit and is recommended for mechanical harvest. The full-sib progeny were planted in 2007 at Sales de Oliveira (São Paulo, Brazil) and in 2011 at Ribeirão Preto (São Paulo, Brazil) in a randomized complete block design with four (in 2007) and three (2011) replicates and plots of 2-m-rows spaced 1.5 m apart. Both parents and two varieties (SP81-3250 and RB835486) were included in each replicate as checks. A preliminary analysis was made in all of the full-sib

progeny using ten SSRs to avoid using individuals who not belong in the bi-parental cross.

Molecular marker data and map construction

A total of 187 individuals from the full-sib progeny were screened with SSRs, AFLPs and SNPs. Of the 140 SSRs that were used, 105 were EST-SSRs that had been developed from sequences in the Brazilian Sugarcane EST Project (SUCEST) database and presented in Pinto et al. (2004, 2006), Oliveira et al. (2007, 2009), and Marconi et al. (2011). The other 35 were genomic SSRs that had been developed either by CIRAD (*Centre de Cooperation Internationale Recherché Agronomique pour le Développement*, Montpellier, France) and described in Rossi et al. (2003) or by Cordeiro et al. (2000). A total of 25 AFLP markers (*Eco*RI/*Mse*I) were screened according to Vos et al. (1995). The 531 SNP markers that were genotyped had also been developed from the SUCEST database, described in Garcia et al. (2013). The SSR and AFLP amplification products were separated by electrophoresis on 6 % denaturing polyacrylamide gels and visualized by silver staining (Creste et al. 2001).

SSR and AFLP markers were scored based on their presence (1) or absence (0) in the parents and segregating progeny. The 1:1 and 3:1 marker segregation was verified through a Chi square test using Bonferroni correction at a significance level of 0.05 to control for type I errors for multiple tests (Province 1999). The SNP genotyping data were generated using a Sequenom iPLEX MassARRAY® (Sequenom Inc., San Diego, CA, USA) as described in Garcia et al. (2013), and each SNP locus was assessed for ploidy level and allelic dosage (number of copies of each form) using SuperMASSA software (Serang et al. 2012). This software simultaneously considers all of the available information and the genetic constraints that the derived results must fulfill, i.e., the possible genotypes to be observed given the ploidy level and the parental genotypes, the ratio between allele intensities, and the expected complete polysomic segregations (Garcia et al. 2013). Loci were classified as single dose when they had an SNP with only a single copy of one of the alleles in one parent, being nulliplex in the other (thus segregating in a 1:1 ratio in the progeny) or when both parents had a single copy of the same allele (segregating in a 1:2:1 ratio).

The linkage map was constructed using the OneMap package (Margarido et al. 2007). To avoid false-positive linkages, a LOD Score of 5.8 and a recombination frequency of 0.50 were used to determine the linkage groups (LGs). To determine the marker order and linkage phases in each linkage group with five markers or less, all possible orders were compared, and the most likely order was selected. For LGs with more than five markers, all possible orders for the five most informative markers were compared and the most likely order identified; the remaining markers were then sequentially added to the position of highest likelihood with respect to that order. The distances in the genetic map were expressed in centimorgans (cM) based on the Kosambi mapping function (Kosambi 1944).

Phenotypic data analysis

The field experiment was evaluated in 2008 (plant cane) and 2009 (ratoon cane) at Sales de Oliveira (São Paulo, Brazil) and in 2012 (plant cane), 2013 (ratoon cane) and 2014 (ratoon cane) at Ribeirão Preto (São Paulo, Brazil).

Important agronomic and economic traits in sugarcane production were measured: soluble solid content (Brix), fiber percentage (Fiber), sucrose content (Pol), stalk weight (SW), stalk diameter (SD) and stalk height (SH). Each measurement was obtained from a sample of 10 stalks harvested from each individual plot according to the methods described in Consecana (2006) for both plant cane and ratoon crops. Brix denotes the total dissolved solids in cane juice, whereas Fiber refers to the water-insoluble matter that is present in the stalk.

To evaluate the phenotypic data obtained in different locations and harvests, an appropriate mixed model was adjusted by comparing different structures for the variance–covariance (VCOV) matrices of genetic (G) and non-genetic effects (R). The statistical model used for each trait separately was (underlined terms indicate random effects):

$$\underline{Y}_{ijrkm} = \mu + L_k + H_m + LH_{km} + \underline{B}_{j(km)} + \underline{G}_{i(km)} + \underline{e}_{ijrkm}$$

where $\underline{Y}_{ijrkm}$ is the phenotype of the ith genotype ($i = 1, \ldots, n$) in the jth block ($j = 1, 2, 3$) of the rth replicate ($r = 1, \ldots, 4$), the kth location ($k = 1, 2$)

and the mth harvest ($m = 1, 2$); μ is the trait mean; L_k is the location effect; H_m is the harvest effect; LH_{km} is the location by harvest interaction; $\underline{B}_{j(km)}$ is the effect of the jth block in the kth location and mth harvest; $\underline{G}_{i(km)}$ is the effect of the ith individual in the kth location and mth harvest; and $\underline{e}_{ijrkm}$ is the non-genetic effect. The individuals can be separated into two groups by genotype: $n = n_g + n_c$. The number of genotypes is represented by n_g ($i = 1, \ldots, n_g$), and the number of checks is represented by n_c ($i = n_g + 1, \ldots, n_g + n_c$). Similar to that of Pastina et al. (2012), the model for $\underline{G}_{i(km)}$ is:

$$\underline{G}_{i(km)} \begin{cases} g_{i(km)} & i = 1, \ldots, n_g \\ c_{i(km)} & i = n_g + 1, \ldots, n_g + n_c \end{cases}$$

With $g_{i(km)}$ representing the random genetic effect of the ith genotype at the kth location and mth harvest and $c_{i(km)}$ representing the fixed effect of the ith check at the kth location and mth harvest.

The VCOV matrix G was obtained via the Kronecker direct product of $G_M \otimes I_{ng}$, as also described in Pastina et al. (2012), where $G_M = G^L \otimes G^H$, $\otimes$ is the Kronecker direct product of genetic effects, and I_{ng} is an identity VCOV matrix of genotypes. To compare the factorial combination between locations and harvests, we also considered $G_M = G^{L-H}$. Likewise, the R_M matrix was obtained via the Kronecker direct product of R^L (residual effects between locations), R^H (residual effects between harvests) and R^B (residual effects between blocks) matrices; thus, $R_M = R^L \otimes R^H \otimes R^B$. It was assumed that $\mathbf{e} MVN(0, R)$ and $gMVN(0, G)$, where $e = (e_{11111}, \ldots, e_{IJRKM})$ and $\left(|1(11), \ldots, g_{I(KM)}\right)$, respectively.

Different models for the VCOV structures were compared via the Akaike information criterion (AIC) (Akaike 1974) and Bayesian information criterion (BIC) (Schwarz 1978). The G_M matrix was G^L; G^L; G^L, and different VCOV structures were tested for G^H; G^H matrix; G_M and the model R_M. For the factorial combination $G_M = G^{L-H}$, comparisons of different VCOV structures were performed in a single step, and the R_M matrix was used to model the G^L, G^H and G^{L-H} matrices, including identity, diagonal, uniform, uniform heterogeneous, analytic factor of order 1, auto-regressive of order 1 and unstructured, analogous to the modeling performed by Pastina et al. (2012).

Likewise, the same VCOV structures were compared for R^L, and R^H, whereas for R^B, factor analytic of order 1 and auto-regressive of order 1 models were not included. Once the best model was selected, joint adjusted means were obtained for each trait via best linear unbiased predictor (BLUP), combining the information of different harvests and locations, and the variance components were estimated by residual maximum likelihood (REML). All of the phenotypic analyses were performed in the GenStat (v. 16.1) software (VSN International, 2011).

The broad-sense heritability coefficient was calculated for each trait. Genetic correlations were estimated for each pair of traits based on the BLUP means and using the Pearson correlation coefficient as implemented in the R software (R Core Team 2013).

QTL mapping

To test the association between phenotype and genotype, each trait was analyzed separately using the joint adjusted means obtained via BLUP and the composite interval mapping (CIM) model (Zeng 1993), and the extended approach described by Gazaffi et al. (2014) was applied. QTL mapping analysis was performed as follows: (i) conditional multipoint QTL genotype probabilities for all positions in a discrete grid of evaluation points with a step size of 1 cM along the genome were estimated via hidden Markov models (HMMs) implemented in the OneMap package (Margarido et al. 2007); (ii) a genome scan was performed (in a grid of 1 cM) to detect QTLs; and (iii) subsequent to QTL identification, the significance of additive (α_p, α_q) and dominance (δ_{pq}) effects was verified, and segregation patterns were estimated (Gazaffi et al. 2014). The inclusion of cofactors was based on a multiple linear regression analysis using stepwise selection with the AIC. We considered a maximum of 27 cofactors and a window size of 20 cM.

The threshold for the detection of a QTL was calculated using 1000 permutations and a significance level of 0.05 (Churchill and Doerge 1994) considering the distribution of the second highest peak for each linkage group (Chen and Storey 2006). For positions with evidence of putative QTL, significant marginal effects for α_p, α_q and δ_{pq} were verified using LOD scores equivalent to a significance level of 0.05 (Gazaffi et al. 2014). The proportion of the phenotypic variance (R^2) explained by each detected QTL was

obtained for all effects simultaneously. All of the analyses were performed in the R environment (R Development Core Team 2013) using an R test package under development.

Results

Linkage map

Using 140 SSR and 25 AFLP primer combinations, we scored 1102 polymorphic markers in the mapping population. Of these markers, 634 (57.5 %) segregated as single-dose markers (SDMs) in 1:1 (377) and 3:1 (257) ratios. Moreover, 531 SNP markers underwent dosage and ploidy level estimation using the SuperMASSA software (Serang et al. 2012). Of these, 54 (10.17 %) were SDMs with segregation ratios of 1:1 (30) and 1:2:1 (24) (Supplementary Material—Table S1). This low number of SDMs is agrees with the results reported by Garcia et al. (2013) and indicates that SDMs are not the most abundant markers in the sugarcane genome when modern genotyping technologies and appropriate analytical methodologies are used. In total, 688 polymorphic SDMs were used to construct the genetic linkage map.

The estimated linkage map has 421 markers (61.2 %) on 118 LGs, with 267 (38.8 %) markers remaining unlinked. The LGs vary in length from 1.0 to 142.9 cM, with an average length of 38.2 cM per LG and an accumulated length of 4512.6 cM. The average distance between markers is 10.7 cM, with average of 3.6 markers per LG, distributed irregularly along the chromosomes. Fourteen gaps are present, ranging from 30 to 39.7 cM. These gaps show that parts of the genome are only partially covered by the markers. The final linkage map is smaller than that constructed for R570 (Hoarau et al. 2001), Q165 (Aitken et al. 2005, 2014) and SP80-180 × SP80-4966 (Oliveira et al. 2007), which have 5849, 9058.3, 9774.4, and 6261.1 cM, respectively.

The 118 LGs were used to establish putative homology groups (HGs) based on at least one pair of common SSR-derived markers that were shared. In total, 87 LGs were organized into 16 putative HGs, each containing between 2 and 17 LGs (Supplementary Material—Figure S1). In accordance with

previous studies (Hoarau et al. 2001; Aitken et al. 2005, 2014; Raboin et al. 2006; Garcia et al. 2006; Oliveira et al. 2007), the markers were not distributed equally within the different HGs. The largest HG contains 54 markers distributed along 11 LGs, and the smallest HG comprises six markers distributed between two LGs. The remaining 31 LGs contain no pairs of SSR-derived markers in common to allow for assignment into HGs. In some cases, the marker order was maintained among LGs. For example, in HGI, the markers scb060, SMC1047HA and scb262 were found to occur in the same order on LG4 and LG8. In HGIII, the order of the cv038, Cir001 and Cir012 markers was the same on LG30 and LG40. In HGIV, the order of the Cir067, Cir036 and scb312 markers on LG42 was preserved on LG44. Gardiner et al. (1993), Aitken et al. (2005, 2014) and Oliveira et al. (2007) also reported the preservation of marker order in some LGs belonging to the same HG.

This is the first genetic map in which SNP markers segregating in a 1:2:1 fashion, and therefore being codominant, was included together with information about ploidy level. Of the 54 SDM SNPs, nine are linked on the map with segregation ratios of 1:2:1 (five) and 1:1 (four) and ploidy levels ranging from 6 to 12 (Table 1). These ploidy levels agree with chromosome number estimates for modern sugarcane cultivars, which fall between 6 and 14 (D'Hont et al. 1996; D'Hont 2005; Garcia et al. 2013).

Table 1 Description of the single-dose marker SNPs linked on the genetic map in the full-sib sugarcane progeny (IACSP95-3018 vs. IACSP93-3046), presenting the segregation ratio, ploidy level, linkage group (LG) and homology group (HG) for each SNP

ID_SugSNP	Segregation	Ploidy[a]	LG	HG
SugSNP_0032	1:2:1	8	47	IV
SugSNP_0062	1:1	6	108	Not identified
SugSNP_0208	1:1	6	105	Not identified
SugSNP_0209	1:2:1	12	50	V
SugSNP_0270	1:2:1	6	105	Not identified
SugSNP_0503	1:1	8	26	II
SugSNP_0729	1:1	6	25	II
SugSNP_0764	1:2:1	12	68	IX
SugSNP_0828	1:2:1	8	113	Not identified

[a] Ploidy level was estimated using SuperMASSA software (Serang et al. 2012)

Phenotypic analysis

Different VCOV models for the G and R matrices were compared in the phenotypic analysis, and the best one, i.e., that with the lowest AIC and BIC values, was selected (Table 2). For SD, SW, SH and Fiber, the best model for the G_M matrix was Uns–AR1, which is based on the Kronecker direct product of the G^L and G^H matrices for the locations and the harvests, respectively. For Brix and Pol, the best model led to a uniform structure, considering each harvest-location combination as an environment. The models selected for SD, SW, SH and Fiber consider a particular genetic variance for each location and a specific covariance between different locations, whereas the correlation between harvest decay and time, i.e., consecutive harvests, are more correlated. Likewise, the model selected for Brix and Pol allows for homogeneous genetic variances across environments and a common genetic covariance between pairs of environments. Different models for the R_M (non-genetic effects) matrices were also compared for all of the traits that were evaluated in this study. For SD, SW, SH and Fiber, $R_M = R^L \otimes R^H \otimes R^B$, whereas for Brix and Pol, $R_M = R^{L-H} \otimes R^B$. In this way, both adjusted BLUP means and variance components were obtained for each trait, allowing for the estimation of genetic parameters (Tables 3, 4).

Broad-sense heritability coefficients were high: 80.02 for SD, 71.98 for SW, 69.58 for SL, 54.48 for Fiber, 53.99 for Brix and 53.20 for Pol. Heritability estimates were consistent with the coefficient of variation (CV), showing that the field experiments were conducted under good conditions. The highest CV value was associated with SW (21.033), as expected because this trait is strongly influenced by environmental conditions. The CV values were low for the other traits, ranging from 5.57 for Brix to 12.14 for SH.

A total of nine significant pairwise correlation coefficients were found. The highest significant phenotypic correlation was found for Brix and Pol (0.91). Intermediate significant genotypic correlations were reported for SD-SW (0.66), SH-SW (0.55), SD-Fiber (−0.39), SW-Fiber (−0.25), Fiber-Brix (0.25), and SW-Pol (0.22). The lowest significant correlation was observed between the traits SD and Pol (0.18) (Table 4).

QTL mapping

In total, 421 single-dose markers (1:1, 3:1 and 1:2:1) were considered in the QTL mapping procedure. Based on the information from these markers, a total of 25 QTLs were detected for SD, SW, SH, Fiber, Pol and Brix using the CIM approach and the integrated genetic map constructed in this study (Table 5, Fig. 1). For all of the evaluated traits in plant cane and ratoon cane, 1000 permutation tests were performed, resulting in LOD score thresholds of 4.40 for SD, 3.98 for SW, 4.31 for SH, 4.89 for Fiber, 5.75 for Pol, and 6.34 for Brix. QTLs were identified in 22 linkage groups and 10 distinct homology groups. As expected for sugarcane, the proportion of the phenotypic variance (R^2) explained by each QTL was low, ranging from 0.069 to 3.87 %. Mapped QTL exhibited 1:2:1, 3:1 and 1:1 segregation ratios, which are also the

Table 2 Most appropriated selected structures for the variance–covariance matrices of genetic (G_M matrix) and non-genetic (R_M matrix) effects used in the mixed model analysis considering each trait separately

Trait	G_M matrix	G_M structures	R_M matrix	R_M structures
SD	$G_M = G^L \otimes G^H$	Uns ⊗ AR1	$R_M = R^L \otimes R^H \otimes R^B$	Uns ⊗ Diag ⊗ Id
SW	$G_M = G^L \otimes G^H$	Uns ⊗ AR1	$R_M = R^L \otimes R^H \otimes R^B$	Diag ⊗ Uns ⊗ Id
SH	$G_M = G^L \otimes G^H$	Uns ⊗ AR1	$R_M = R^L \otimes R^H \otimes R^B$	Id ⊗ Unif_Het ⊗ Unif
Fiber	$G_M = G^L \otimes G^H$	Uns ⊗ Uns	$R_M = R^L \otimes R^H \otimes R^B$	Diag ⊗ Uns ⊗ Diag
Brix	$G_M = G^{L-H}$	Unif	$R_M = R^{L-H} \otimes R^B$	Unif_Het ⊗ Uns
Pol	$G_M = G^{L-H}$	Unif	$R_M = R^{L-H} \otimes R^B$	Unif_Het ⊗ Diag

SD Stalk diameter, *SW* stalk weight, *SH* stalk height, *Fiber* fiber percentage, *Brix* soluble solid content, *Pol* sucrose content, *Id* identity, *Diag* diagonal, *Unif* uniform, *Unif_Het* uniform heterogeneous, *AR1* auto-regressive of order 1, *Uns* unstructured

Table 3 Estimation of genetic parameters: broad-sense heritability coefficient (h^2), genotypic variance (σ_G^2), phenotypic variance (σ_P^2), residual coefficient of variation (CV) and mean (μ), based on BLUP values for each trait separately in the full-sib sugarcane progeny (IACSP95-3018 vs. IACSP93-3046)

Parameter	SD	SW	SH	Fiber	Brix	Pol
h^2	80.02	71.98	69.58	54.48	53.99	53.20
σ_G^2	0.021	0.958	99.4	0.199	0.169	0.171
σ_P^2	0.092	6.38	697.7	1.13	1.55	1.66
CV	9.518	21.033	12.140	7.377	5.570	7.862
μ	2.68	10.678	191.13	12.75	20.61	15.45

SD Stalk diameter, *SW* stalk weight, *SH* stalk height, *Fiber* fiber percentage, *Brix* soluble solid content, *Pol* sucrose content

Table 4 Correlation matrix for all of the traits that were evaluated in the full-sib sugarcane progeny (IACSP95-3018 vs. IACSP93-3046) based on BLUP values

Trait	SW	SH	Fiber	Brix	Pol
SD	**0.66**	−0.10	**−0.39**	0.00	**0.18**
SW		**0.55**	−0.25	0.10	**0.22**
SH			0.09	0.12	0.11
Fiber				**0.25**	−0.02
Brix					**0.91**

Values displayed below the diagonal were omitted because they are identical to the values shown above the diagonal

Bold values are significant at 5 % ($\alpha = 0.05$)

SD Stalk diameter, *SW* stalk weight, *SH* stalk height, *Fiber* fiber percentage, *Brix* soluble solid content, *Pol* sucrose content

marker segregation patterns available in our genetic map.

Of these 25 QTLs, five were detected for SW, two for Pol, three for Brix, six for SD, four for SH and five for Fiber (Fig. 1; Table 5). The highest LOD score (9.48) was associated with the QTL SW.1, whereas the lowest score (4.49) was associated with the QTL SW.5. Most of the QTLs had a 1:1 segregation pattern, and the proportion of phenotypic variation (R^2) explained for all QTLs were higher (7.11 %) for fiber content than for the remaining traits. QTL B.2 shows the lowest R^2 value (0.069), and QTL F.2 shows the highest R^2 value (3.87). QTL F.2 accounts for almost half of the phenotypic variance explained by the mapped QTL for fiber content and exhibits a 1:2:1 segregation pattern. This QTL was detected by an SNP marker (SugSNP_0729) and is classified as hexaploid (Table 1).

Discussion

QTL mapping is a useful tool for evaluating genotypes of importance for breeding programs because the inheritance of the majority of quantitative traits in sugarcane is complex. However, successful QTL mapping depends on the construction of reliable genetic linkage maps. In general, most of the 118 LGs present in our genetic map had reduced size (average length of 38.2 cM) and had few markers per LG (average of 3.57). This information if combined with the 241 markers that remained unlinked indicates that the genetic linkage map is still not saturated. This lack of saturation can probably be attributed to the low level of polymorphism found in some regions of the *S. officinarum* genome, from which a large part of the genome of modern sugarcane varieties originated as a consequence of the nobilization process (Alwala and Kimbeng 2010). Grivet et al. (1996), Ming et al. (1998), Hoarau et al. (2001), Aitken et al. (2005), Garcia et al. (2006), Oliveira et al. (2007), Palhares et al. (2012) and Pastina et al. (2012) also observed small LGs with few linked markers in their genetic maps. Although the genetic map not be completely saturated, the number of LGs agrees with the chromosome number expected in modern sugarcane varieties derived from *S. officinarum* (2n = 80) and *S. spontaneum* (2n = 40–128), which have a genome composed of approximately 70–80 % *S. officinarum*, 10–20 % *S. spontaneum* and 5–17 % recombinant chromosomes (D'Hont et al. 1996; Grivet et al. 1996; D'Hont and Glaszmann 2001).

When only single-dose polymorphisms are considered, gaps are expected. In our linkage map, gaps ranging from 30 to 39.7 cM were observed; these

Table 5 Quantitative trait loci estimated with the composite interval mapping (CIM) model considering each trait separately in the full-sib sugarcane progeny (IACSP95-3018 vs. IACSP93-3046)

QTL	LG	Position (cM)	LOD[a]	α_p	α_q	δ_{pq}	Segregation	R^2 (%)
SW.1	36	15	9.48	**−0.36**	**−0.25**	–	3:1	0.88
SW.2	108	11	4.94	−0.10	**0.38**	–	1:2:1	0.37
SW.3	75	59.10	4.89	**0.22**	−0.07	0.11	1:1	3.43
SW.4	77	16	4.54	**−0.12**	**0.38**	–	3:1	0.16
SW.5	09	6.47	4.29	**0.20**	–	–	1:1	0.09
								4.93
P.1	111	15.94	8.14	**−0.11**	–	–	1:1	0.34
P.2	46	10	7.63	**−0.072**	**0.087**	**0.051**	3:1	2.30
								2.64
B.1	46	33	8.25	**−0.10**	**0.061**	**0.0508**	3:1	1.24
B.2	61	18.40	7.29	**−0.18**	0.033	–	1:2:1	0.069
B.3	49	47	6.78	0.033	**−0.090**	**−0.074**	1:2:1	1.85
								3.16
SD.1	12	43	6.45	**0.026**	−0.007	**−0.42**	1:2:1	0.34
SD.2	43	3.45	5.64	–	**−0.040**	–	1:1	0.18
SD.3	1	83.01	5.27	**0.055**	**0.022**	–	3:1	0.16
SD.4	67	4	5.04	0.050	**−0.040**	–	1:1	0.51
SD.5	75	50	4.92	**0.036**	0.013	**0.031**	1:2:1	2.14
SD.6	90	79.03	4.56	−0.018	**−0.036**	–	1:1	0.74
								4.07
SH.1	84	23	5.83	**3.17**	–	–	3:1	1.38
SH.2	45	12.61	5.52	**−3.36**	1.35	–	1:1	2.04
SH.3	118	3.87	5.35	**1.87**	**2.28**	–	1:1	0.48
SH.4	92	37.87	4.95	**1.35**	**2.20**	**0.018**	3:1	0.17
								4.07
F.1	61	9	6.84	−0.012	**−0.135**	–	1:1	0.27
F.2	25	11.34	6.46	**−0.20**	−0.03	–	1:2:1	3.87
F.3	2	49	6.46	**0.11**	–	–	1:1	0.26
F.4	15	10	6.12	**−0.128**	–	–	1:1	2.35
F.5	116	0	5.86	**−0.042**	**−0.189**	–	3:1	0.36
								7.11

Bold values denote significant marginal effects, Segregation Mendelian ratio

SW Stalk weight, *B* soluble solid content, *P* sucrose content, *SD* stalk diameter, *SH* stalk height, *F* fiber percentage, *LG* linkage group, *Position (cM)* mapped position in centiMorgans, α_p additive effect in parent IACSP95-3018, α_q additive effect in parent IACSP93-3046, δ_{pq} dominance effect, R^2 Percentage of phenotypic variance explained by each QTL

[a] LOD score for declaring significance obtained for each trait separately (4.03 for SD, 4.15 for SW, 4.77 for F, 5.8 for P, and 6.79 for B) based on permutation

values are smaller than the 40 cM gaps reported by Hoarau et al. (2001). Another possible explanation for the large number of unlinked markers is the use of progeny derived from a cross between two commercial varieties. This type of cross has complex meiotic behavior, including aneuploidy and non-pairing of chromosomes, which could result in a large proportion of unlinked markers on the genetic map (Garcia et al. 2006). Using alleles derived from the same SSR or EST-SSR locus that was mapped on multiple LGs, we identified 16 HGs. This is significantly more than the expected basic number of chromosomes for the genus *Saccharum*, which falls between x = 8 and x = 10

(D'Hont et al. 1998; Irvine 1999; Grivet and Arruda 2001). Five HGs contained only two LGs. The small size of the linkage groups may reflect chromosome fragmentation that hinders correct HG grouping, further suggesting that the map is not saturated and reinforcing the need to use multiplex markers.

It is extremely difficult to determine the ploidy level of polymorphisms using molecular markers, such as AFLPs and SSRs, due to the dominant nature and dominance behavior in complex polyploids, respectively, of these markers. Garcia et al. (2013) suggested that only 30.5 % of all the markers in the sugarcane genome are SDMs when considering SNPs
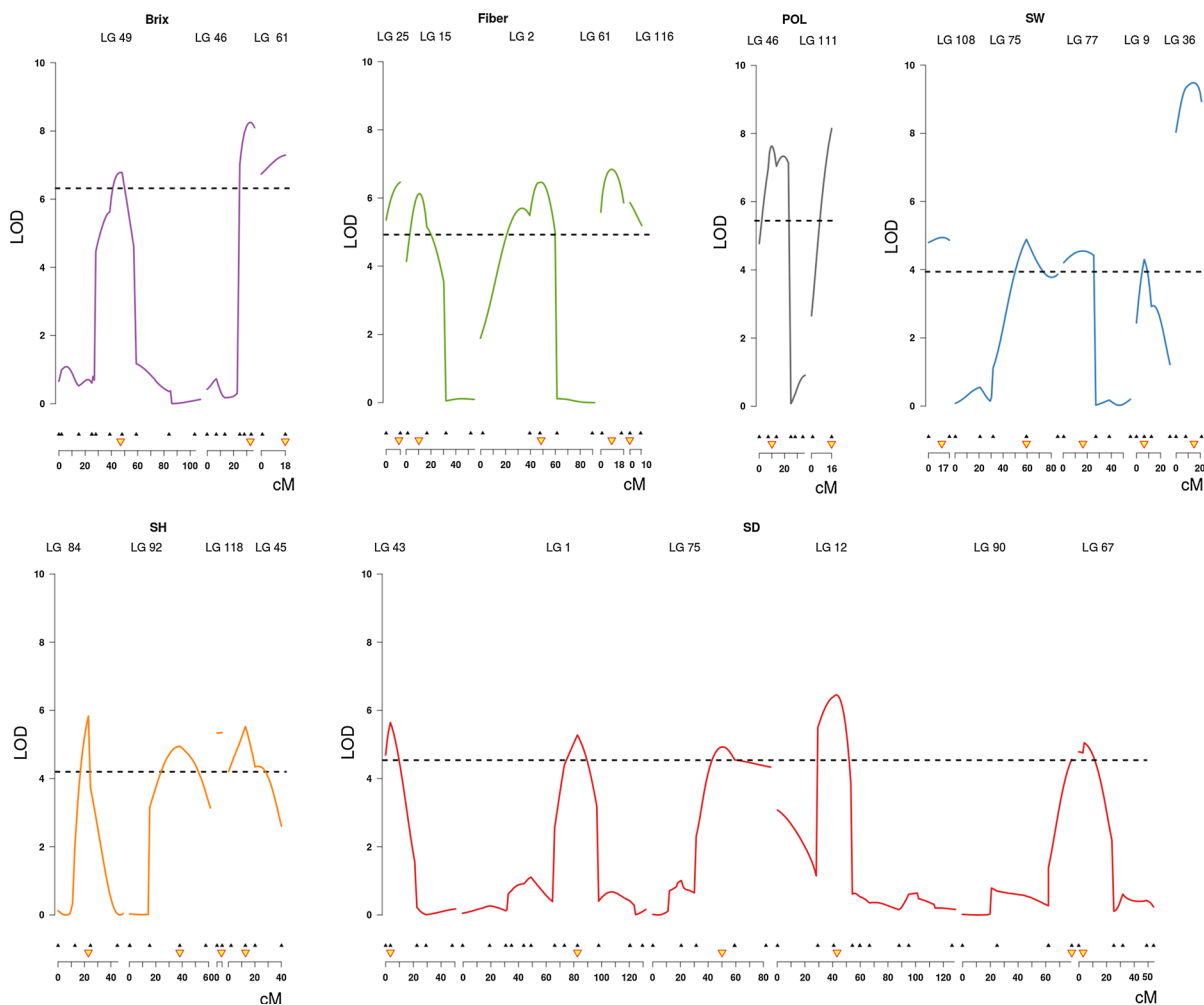
Fig. 1 QTLs identified (*inverted triangles*) and associated with stalk diameter (SD), stalk weight (SW), stalk height (SH), fiber, Pol and Brix in the full-sib sugarcane progeny (IACSP95-3018 vs. IACSP93-3046)

of all ploidy levels classified by SuperMASSA software. This result is in contrast to the results of others studies, such as Aitken et al. (2014), which reported that 83 % of all SNPs markers are in SDMs when not considering information about the ploidy level. Of the 531 SNPs genotyped in our mapping population, 54 (10.16 %) were SDMs, and only nine of these were linked on the genetic map. However, even if the analysis is restricted to SDM SNPs, knowing the ploidy level of the SNPs is advantageous for mapping the sugarcane genome because the ploidy level can indicate the number of chromosomes present in the homology group to which the SNP was mapped. This ploidy information can now be ascertained, improving the genetic mapping of sugarcane.

As shown in Table 1, SugSNP_0032 was estimated to have a ploidy level of 8 and mapped to LG47 in HGIV (Fig. 2). This HG comprises seven LGs; the ploidy level of 8 for SugSNP_0032 suggests that the HG is not yet saturated and will likely contain eight LGs when the map is saturated. SugSNP_0503 and SugSNP_0729 are found in HGII on LG25 (11.3 cM) and LG26 (8.3 cM), respectively, and have ploidy levels of 8 (SugSNP_0503) and 6 (SugSNP_0729). However, this HG has 17 LGs (Fig. 2). The LGs to which the SNP markers were mapped are very small, suggesting that this HG will be reorganized as the map becomes more saturated.

The variation in ploidy level (Table 1) of the SNPs mapped to LGs inside the HGs agrees with previously
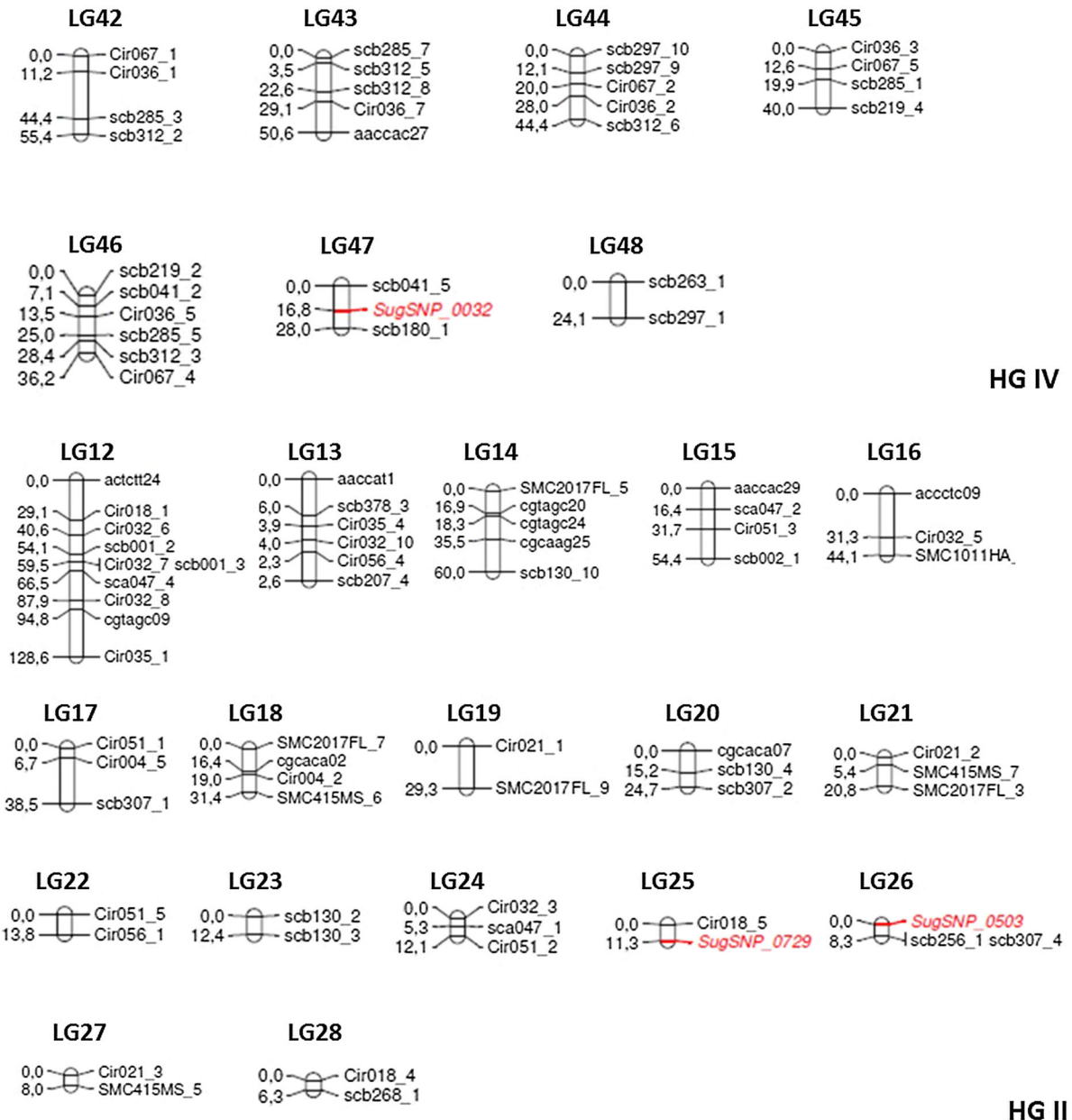
Fig. 2 Two homology groups (HGs) with SNP markers (*in red*) mapped in the full-sib sugarcane progeny (IACSP95-3018 vs. IACSP93-3046). In HG IV, SugSNP_0032 was found on LG47. In HG II, SugSNP_0729 and SugSNP_0503 were mapped to LG25 and LG26, respectively

published data indicating that the sugarcane homology groups have different numbers of chromosomes (Grivet and Arruda 2001). An estimation of the ploidy level for each SNP is essential for further analysis because saturated linkage maps cannot be obtained for sugarcane if molecular markers with higher doses are not considered. The use of these multiplex markers widely distributed throughout the genome will probably contribute to increased linkage map coverage.

Because the association between genotype and phenotype in sugarcane is based on data from studies at different locations across multiple harvests, it is

important to take into account appropriate assumptions regarding variance and covariance matrices for genotype effects and residual effects (Smith et al. 2007). The fitted VCOV structure for the genetic effects matrix for SD, SW, SH and Fiber was the Uns-AR1 model, showing that each interaction, namely, genotype-location-harvest (see Table 3 in results), is inherent for each location, whereas the correlation between harvests decays over time. By applying mixed models to data from trials performed in different environments, it is possible to detect heterogeneity in genetic variance and correlations between environments (Malosetti et al. 2013), allowing for a more realistic understanding of the genotypic effects. Analyses that account for joint-adjusted means, obtained via BLUPs, for different environments should improve the detection of significant marker-phenotype associations.

CIM offers several advantages over single-marker and interval mapping approaches (Zeng 1993, 1994; Jansen and Stam 1994). However, few studies have reported the use of CIM in QTL mapping for sugarcane yield-related traits (Aitken et al. 2008; Shing et al. 2013). Gazaffi et al. (2014) presented a model for outcrossing species using integrated maps. QTL mapping results are more informative when the CIM method is performed with an integrated genetic linkage map and the genotypes at QTL are obtained via multipoint conditional probabilities. This method was employed here, and it proved to be an excellent approach to QTL mapping in outbred species using full-sib progeny obtained from two non-inbred parents. Our QTL mapping results provided estimates of additive effects, dominance effects and segregation patterns for QTL, as well as linkage phases. Using this approach, it is possible to obtain the segregation pattern of QTLs, including those located in less informative regions. In addition, our results are very useful for marker-assisted selection because it is possible to identify, in each parent, alleles that contribute to increased or decreased values of a specific phenotypic trait.

Because our study involved field experiments with the same mapping population in two locations, a validation process ensured that the genotypes in both places were consistent. Although a low number of QTLs were detected, the confirmation of the relation between genotype and phenotype in both field locations gives us more confidence for the mapped QTL. A total of 25 QTLs were identified, including six for SD, five for SW, four for SH, five for Fiber, two for Pol and three for Brix. All of the QTLs for a given trait were located in different linkage groups. The values of the estimated heritability coefficients for Brix (53.99) and Pol (53.20) may have caused a decrease in the statistical power of the QTL mapping. Comparisons between QTL mapping results from different sugarcane studies are difficult due to differences in parental variety, experimental design and the environment, among other factors. However, several studies have reported the detection of a low number of QTLs associated with cane yield and sugar yield traits because all of the genetic maps in sugarcane use only single-dose markers and are not yet saturated (Sills et al. 1995; Ming et al. 2002a, b; Jordan et al. 2004; da Silva and Bressiani 2005; Aitken et al. 2006; Piperidis et al. 2008). By including additional markers as cofactors in the mapping model, the variation associated with QTLs located outside the mapping interval and the detection of false-positive QTLs were both reduced. Note that all QTLs were detected when the average of the environments (harvest location × harvest) was considered. Because we used BLUPs in the QTL mapping procedure, we cannot draw conclusions about QTL stability across different combinations of environments (i.e., location and harvest). Such claims can be made with mapping models that take the QTL-location-harvest interaction into account, as suggested by Pastina et al. (2012).

All of the QTLs showed significant additive and/or dominance effects (with a predominance of additive effects), indicating that the associated alleles contribute to the variation observed in the investigated traits. In particular, these effects were detected for IACSP95-3018 because it is a promising clone used as a parent in the IAC Sugarcane Breeding Program. The significant dominance effects detected for the QTLs SD.1, SD.5, P.2, B.1 and B.3 are important for understanding the genetic complexity of sugarcane. The percentage of phenotypic variation explained by each QTL is low, ranging from 0.069 (B.2) to 3.87 (F.2). Given the high level of polyploidy in sugarcane, individual QTLs are expected to have limited effects on phenotype. Hoarau et al. (2002) found R^2 values ranging from 3 to 7 %; Aitken et al. (2006) reported variation between 2 and 8 %; and Aitken et al. (2008) observed R^2 values between 2 and 10 %. Ming et al. (2002a) reported phenotypic variation with greater

influence on evaluated traits, with R^2 values ranging from 3.8 to 16.2 %. The inclusion of more molecular markers and the use of mapping models that consider multiple traits and multiple environments, as well as QTL dosage, would likely increase the proportion of phenotypic variation explained by each QTL.

The QTLs that had a 1:1 segregation pattern each exhibited only a significant effect (α_p, α_q or δ_{pq}), except for SH.3. Six QTLs displayed a 1:2:1 segregation pattern, of which three had two significant effects and three had only one significant effect. QTLs that had a 3:1 segregation pattern showed all significant effects (α_p, α_q and δ_{pq}) according to Gazaffi et al. (2014) and segregated in a 3:1 ratio. This information helps us understand the behavior of the QTL alleles in the progeny.

QTLs for correlated traits that map to neighboring or overlapping regions of the same linkage group are important for future investigations involving linkage and pleiotropy. SD and SW have a significant positive phenotypic correlation, and QTLs for both (SD.5 and SW.3) were detected in close proximity on LG75 (HGXI). These QTLs may actually be a single pleiotropic QTL. Brix and Pol also have a significant positive phenotypic correlation, and QTLs for both (P.2 and B.1) mapped to the same linkage group (LG46 and HGIV). The LOD profiles for these QTL were similar, possibly indicating a second pleiotropic QTL. QTLs for Fiber and Brix (F.2 and B.2) mapped to LG61 (HGVII). These traits had a significant positive phenotypic correlation.

Several factors hinder the dissection of polyploid genomes by QTL mapping, including the small number of markers; the use of SDMs only; the large proportion of unlinked markers, resulting in poorly saturated maps; and the absence of inbred lines, which reduces mapping accuracy. We used mixed models applied to phenotypic data, which allowed us to model VCOV structures for genetic effects to predict genotype values and to avoid an unbalanced data scenario. Moreover, QTL genotypes were estimated via multi-point conditional probabilities, providing advantages, such as an increase in the statistical power for QTL detection. We also mapped QTLs into an integrated genetic linkage map via a CIM approach that included additive and dominance effects, enabling us to estimate segregation patterns and linkage phases for all 25 QTL. The limits on the power of QTL detection could be overcome by the advent of single-nucleotide polymorphism (SNP) technology that uses a statistical

innovation to interpret data from polyploids (Debibakas et al. 2014), moving from dominant markers to bi-allelic SNPs with ploidy and dosage information (Garcia et al. 2013). The SuperMASSA software (Serang et al. 2012) was able to infer ploidy level and allelic dosage for all SNPs. However, only the SNPs that were classified as SDMs were used for linkage map construction (Supplementary Material—Table S1). The linkage map did not provide a full coverage of the genome, and the QTL number was likely underestimated.

To the best of our knowledge, this study is the first to include SNPs segregating in a codominant pattern (1:2:1) for sugarcane, allowing a better integration of the map and better statistical tests for QTL. Presumably, the utilization of multiple dose markers will lead to more precise QTL localization and better estimates of QTL effects, segregation patterns and interactions, in addition to a more saturated genetic linkage map. Therefore, the development of new approaches that include ploidy and dosage information is necessary to better understand complex polyploid genomes. This knowledge could then be used as a first step to improve the methods for understanding the genetic and genomic mechanisms associated with agronomic traits.

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

Aitken KS, Jackson PA, McIntyre CL (2006) Quantitative trait loci identified for sugar related traits in a sugarcane (*Saccharum* spp.) cultivar× *Saccharum officinarum* population. Theor Appl Genet 112:1306–1317

Aitken KS, Hermann S, Karno K, Bonnett GD, McIntyre LC, Jackson PA (2008) Genetic control of yield related stalk traits in sugarcane. Theor Appl Genet 117:1191–1203

Aitken KS, Mc Neil MD, Hermann S, Bundock PC, Kilian A, Heller-Uszynska K, Henry RJ, Li J (2014) A comprehensive genetic map of sugarcane that provides enhanced map coverage and integrates high-throughput Diversity Array Technology (DArT) markers. BMC Genom 15:152

Akaike H (1974) A new look at the statistical model identification. IEEE Trans Autom Control 19:716–723

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

Al-Janabi SM, Parmessur Y, Kross H, Dhayan S, Saumtally S, Ramdoyal K, Autrey LJC, Dookun-Saumtally A (2007) Identification of a major quantitative trait locus (QTL) for yellow spot (*Mycovellosiella koepkei*) disease resistance in sugarcane. Mol Breed 19:1–14

Alwala S, Kimbeng CA (2010) Molecular genetic linkage mapping in *Saccharum*: strategies, resource and achievements. In: Henry R, Kole C (eds) Genetics, genomics and breeding of sugarcane. Science Publishers, New Hampshire, pp 69–96

Bargary N, Hinde J, Garcia AAF (2014) Finite mixture model clustering of SNP data. In: Mackenzie G, Peng D (eds) Statistical modelling in biostatistics and bioinformatics, 1st edn. Springer, Switzerland, pp 139–157

Consecana – Conselho nacional dos produtores de cana-de-açúcar, açúcar e álcool do estado de São Paulo (2006) Manual de instruções – CONSECANA-SP. Editora CONSECANA, Piracicaba, p 112

Chen L, Storey JD (2006) Relaxed significance criteria for linkage analysis. Genetics 173:2371–2381

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

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

Creste S, TulmannNeto A, Figueira A (2001) Detection of single sequence repeats polymorphisms in denaturing polyacrylamide sequencing gel by silver staining. Plant Mol Biol Rep 19:299–306

D'Hont A (2005) Unravelling the genome structure of polyploids using FISH and GISH: examples in sugarcane and banana. Cytogenet Genome Res 109:27–33

D'Hont A, Glaszmann JC (2001) Sugarcane genome analysis with molecular markers, a first decade of research. Proc Int Soc Sugarcane Technol 24:556–559

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

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

da Silva JA, Bressiani JA (2005) Sucrose synthase molecular marker associated with sugar content in elite sugarcane progeny. Genet Mol Biol 28:294–298

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

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

Debibakas S, Rocher S, Garsmeur O, Toubi L, Roques D, D'Hont A, Hoarau J-Y, Daugrois JH (2014) Prospecting sugarcane resistance to sugarcane yellow leaf virus by genome-wide association. Theor Appl Genet 127:1719–1732

Falconer DS, Mackay TF (1996) Introduction to quantitative genetics, 4th edn. Editora Longman Group, Londres, p 464

Garcia AAF, Kido EA, Pinto LR, Meza AN, Silva GD, Souza HMD, Pinto LR, Pastina MM, Leite CS, da Silva JAGD, Ulian EC, Figueira A, Souza AP (2006) Development of an integrated genetic map of a sugarcane (*Saccharum* spp.) commercial cross, based on a maximum-likelihood approach for estimation of linkage and linkage phases. Theor Appl Genet 112:298–314

Garcia AAF, Mollinari M, Marconi TG, Serang OR, Silva RR, Vieira MLC, Vicentini R, Costa EA, Mancini MC, Garcia MOS, Pastina MM, Gazaffi R, Martins ERF, Dahmer N, Sforça DA, Silva CBC, Bundock P, Henry R, Souza GM, van Sluys MA, Landell MGA, Carneiro MS, Vincentz MAG, Pinto LR, Vencovsky R, Souza AP (2013) SNP genotyping allows an in-depth characterization of the genome of sugarcane and other complex autopolyploids. Sci Rep 3:3399

Gardiner JM, Coe EH, Melia-Hancock S, Hoisington DA, Chao S (1993) Development of a core RFLP map in maize using an immortalized F(2) population. Genetics 134:917–930

Gazaffi R, Margarido RAG, Pastina MM, Mollinari M, Garcia AAF (2014) A model for quantitative trait loci mapping, linkage phase, and segregation pattern estimation for a full-sib progeny. Tree Genet Genomes 10:791–801

GenStat for Windows 16th Edition.VSN International, Hemel Hempstead, UK. www.GenStat.co.uk

Grattapaglia D, Sederoff R (1994) Genetic linkage maps of *Eucalyptus grandis* and *Eucalyptus urophylla* using a pseudo-testcross: mapping strategy and RAPD markers. Genetics 137:1121–1137

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 organization in a highly polyploid and aneuploid interspecific hybrid. Genetics 142:987–1000

Heinz DJ, Tew TL (1987) Hybridization procedures. In: Heinz DJ (ed) Sugarcane improvement through breeding. Elsevier, Amsterdam, pp 313–342

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.). I. Genome

mapping with AFLP markers. Theor Appl Genet 103:84–97

Hoarau JY, Grivet L, Offmann B, Raboin LM, Diorflar JP, Payet J, Hellmann M, D'Hont A, Glaszmann JC (2002) Genetic dissection of a modern sugarcane cultivar (*Saccharum* spp.). II. Detection of QTL for yield components. Theor Appl Genet 105:1027–1037

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

Jansen RC, Stam P (1994) High resolution of quantitative traits into multiple loci via interval mapping. Genetics 136:1447–1455

Jordan DR, Casu RE, Besse P, Carroll BC, Berding N, McIntyre CL (2004) Markers associated with stalk number and suckering in sugarcane collocates with tillering and rhizomatous ness QTL in sorghum. Genome 47:988–993

Kosambi DD (1944) The estimation of map distances from recombination values. Annu Eugene 12:172–175

Lin M, Lou X, Chang M, Wu R (2003) A general statistical framework for mapping quantitative trait loci in no model systems: issue for characterizing linkage phases. Genetics 165:901–913

Lynch M, Walsh B (1998) Genetics and analysis of quantitative traits. Sinauer Associates, Sunderland

Malosetti M, Ribaut J-M, van Eeuwijk FA (2013) The statistical analysis of multi-environment data: modeling genotype-by-environment interaction and its genetic basis. Front Physiol 4:44. doi:10.3389/fphys.2013.00044

Marconi TG, Costa EA, Miranda HRCAN, Mancini MC, Cardoso-Silva CB, Oliveira KM, Pinto LR, Molinari M, Garcia AAF, Souza AP (2011) Functional markers for gene mapping and genetic diversity studies in sugarcane. BMC Res Notes 4:264

Margarido GRA, Souza AP, Garcia AAF (2007) OneMap: software for genetic mapping in outcrossing species. Hereditas 144:78–79

McIntyre CL, Whan VA, Croft B, Magarey R, Smith GR (2005) Identification and validation of molecular markers associated with Pachymetra root rot and brown rust resistance in sugarcane using map- and association-based approaches. Mol Breed 16:151–161

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 organization of closely related diploid and polyploid genomes. Genetics 150:1663–1682

Ming R, Liu SC, 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, Wang W, Draye X, Moore H, Irvine E, Paterson H (2002a) Molecular dissection of complex traits in autopolyploids: mapping QTL affecting sugar yield and related traits in sugarcane. Theor Appl Genet 105:332–345

Ming R, Del Monte TA, Hernandez E, Moore PH, Irvine JE, Paterson AH (2002b) Comparative analysis of QTL affecting plant height and flowering among closely-related diploid and polyploidy genomes. Genome 45:794–803

Mollinari M, Serang O (2015) Quantitative SNP genotyping of polyploids with MassARRAY and other platforms. In: Batley J (ed) Genotyping: methods and protocols, methods in molecular biology, vol 1245. Springer, New York, pp 215–241

Oliveira KM, Pinto LR, Marconi TG, Margarido GRA, Pastina MM, Teixeira LHM, Figueira AV, Ulian EC, Garcia AAF, Souza AP (2007) Functional integrated genetic linkage map based on EST markers for a sugarcane (*Saccharum* spp.) commercial cross. Mol Breed 20:189–208

Oliveira KM, Pinto LR, Marconi TG, Mollinari M, Ulian EC, Chabregas SM, Falco MC, Burnquist W, Garcia AAF, Souza AP (2009) Characterization of new polymorphic functional markers for sugarcane. Genome 52:191–209

Palhares AC, Rodrigues-Morais TB, Van Sluys MA, Domingues DS, Maccheroni W Jr, Jordao H Jr, Souza AP, Marconi TG, Mollinari M, Gazaffi R, Garcia AA, Vieira ML (2012) A novel linkage map of sugarcane with evidence for clustering of retrotransposon-based markers. BMC Genet 13:51

Pastina MM, Pinto LR, Oliveira KM, Souza AP, Garcia AAF (2010) Molecular mapping of complex traits. In: Henry R, Kole C (eds) Genetics, genomics and breeding of sugarcane. Science Publishers, Enfield, pp 117–148

Pastina MM, Malosetti M, Gazaffi R, Mollinari M, Margarido GRA, Oliveira KM, Pinto LR, Souza AP, van Eeuwijk FA, Garcia AAF (2012) A mixed model QTL analysis for sugarcane multiple-harvest-location trial data. Theor Appl Genet 124:835–849

Paterson AH, Moore PH, Tew TL (2013) The gene pool of *Saccharum* species and their improvement. In: Paterson AH (ed) Genomics of the Saccharinae, vol 11. Springer, Berlin, pp 43–72

Pinto LR, Oliveira KM, Ulian EC, Garcia AAF, Souza AP (2004) Survey in the sugarcane expressed sequence tag database (SUCEST) for simple sequence repeats. Genome 47:795–804

Pinto LR, Oliveira KM, Marconi TG, Garcia AAF, Ulian EC, De Souza AP (2006) Characterization of novel sugarcane expressed sequence tag microsatellites and their comparison with genomic SSRs. Plant Breed 125:378–384

Piperidis N, Jackson PA, D'Hont A, Besse P, Hoarau JY, Courtois B, Aitken KS, McIntyre CL (2008) Comparative genetics in sugarcane enables structured map enhancement and validation of marker-trait associations. Mol Breed 21:233–247

Province MA (1999) Sequential methods of analysis for genome scan. In: Rao DC, Province MA (eds) Dissection of complex traits. Academic Press, San Diego, p 583

R Core Team (2013) R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria

Raboin LM, Oliveira KM, Raboin LM, Lecunff L, Telismart H, Roques D, Butterfield M, Hoarau JY, D'Hont A (2006) Genetic mapping in sugarcane, a high polyploid, using biparental progeny: identification of a gene controlling stalk colour and a new rust resistance gene. Theor Appl Genet 112:1382–1391

Reffay N, Jackson PA, Aitken KS, Hoarau JY, D'Hont A, Besse P, McIntyre CL (2005) Characterization of genome regions incorporated from an important wild relative into Australian sugarcane. Mol Breed 15:367–381

Rossi M, Araujo PG, Paulet F, Garsmeur O, Dias VM, Chen H, Van Sluys MA, D'Hont AD (2003) Genomic distribution

and characterization of EST82 derived resistance gene analogs (RGAs) in sugarcane. Mol Genet Genom 269:406–419

Schwarz G (1978) Estimating the dimension of a model. Ann Stat 6:461–464

Serang O, Mollinari M, Garcia AAF (2012) Efficient exact maximum a posteriori computation for bayesian SNP genotyping in polyploids. PLoS ONE 7:e30906

Shing S, Shing RP, Sridhar B, Huerta-Espino J, Eugenio LVE (2013) QTL mapping of slow-rusting, adult plant resistance to race Ug99 of stem rust fungus in PBW343/Muu RIL population. Theor Appl Genet. doi:10.1007/s00122-013-2058-0

Sills GR, Bridges W, Al-Janabi SM, Sobral BWS (1995) Genetic analysis of agronomic traits in a cross between sugarcane (*Saccharum officinarum* L.) and its presumed progenitor (*S. robustum* Brandes & Jesw. ex Grassl). Mol Breed 1:355–363

Smith AB, Stringer JK, Wei X, Cullis BR (2007) Varietal selection for perennial crops where data relate to multiple harvests from a series of field trials. Euphytica 157:253–266

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

Wu K, Burnquist W, Sorrels M, Tew T, Moore P, Tanksley S (1992) The detection and estimation of linkage in polyploids using single-dose restriction fragments. Theor Appl Genet 83:294–300

Wu R, Ma CX, Painter I, Zeng ZB (2002) Simultaneous maximum likelihood estimation of linkage and linkage phases in outcrossing species. Theor Popul Biol 61:349–363

Zeng ZB (1993) Theoretical basis of separation of multiple linked gene effects on mapping quantitative trait loci. Proc Natl Acad Sci 90:10972–10976

Zeng ZB (1994) Precision mapping of quantitative trait loci. Genetics 136:1457–1468

Zeng ZB, Kao CH, Basten CJ (1999) Estimating the genetic architecture of quantitative traits. Genet Res 74:279–289