Theor Appl Genet (2011) 123:999–1011
DOI 10.1007/s00122-011-1642-4

ORIGINAL PAPER

QTL for fibre-related traits in grain × sweet sorghum as a tool for the enhancement of sorghum as a biomass crop

Amukelani L. Shiringani · Wolfgang Friedt

Received: 28 October 2010 / Accepted: 22 June 2011 / Published online: 8 July 2011

Abstract Compared to maize and temperate grasses, sorghum has received less attention in terms of improving cell wall components. The objectives of this study were to identify quantitative trait loci (QTL) with main effects, epistatic and pleiotropic effects along with QTL × environment (QE) interactions controlling fibre-related traits in sorghum. Neutral detergent fibre (NDF), acid detergent fibre (ADF), acid detergent lignin (ADL), cellulose, hemicellulose, fresh leaf mass, stripped stalk mass, dry stalk mass, fresh biomass and dry biomass were analysed from a population of 188 grain × sweet sorghum recombinant inbred lines. A genetic map consisting of 157 DNA markers was constructed, and QTL were detected using composite interval mapping (CIM). CIM detected more than 5 additive QTL per trait explaining 7.1–24.7% of the phenotypic variation. Abundant co-localization of these QTL was observed across all chromosomes, and the highest cluster was identified on chromosome 6. Searching for candidate genes using the confidence interval of our QTL clusters reveals that these clusters might comprise a set of genes that are tightly linked. Some QTL showed multiple effects; however, the allele for each trait was favouring the parent with the increasing effect. QE interactions were observed for QTL showing multiple effects. Additive × additive interaction was observed for 7 out of 10 traits, indicating the importance of epistatic analysis. However, the phenotypic variation explained by digenic interactions was lower compared to the individual QTL. Our results indicate that various genetic components contribute to fibre-related traits and should be considered during the enhancement of sorghum for lignocellulosic biomass.

Communicated by A. Paterson.

Electronic supplementary material The online version of this article (doi:10.1007/s00122-011-1642-4) contains supplementary material, which is available to authorized users.

A. L. Shiringani · W. Friedt (✉)
Department of Plant Breeding, Research Centre for BioSystems, Land Use and Nutrition (IFZ), Justus-Liebig University-Giessen, Heinrich-Buff-Ring 26-32, 35392 Giessen, Germany
e-mail: wolfgang.friedt@agrar.uni-giessen.de

Present Address:
A. L. Shiringani
Centre of Novel Agricultural Products, Department Biology Area7, Plant Breeding and Genetics, University of York, PO Box 373, York Yo105YW, UK

Introduction

Grasses display significant genetic variation in digestibility, structure and components of cell wall (Grabber et al. 2004). Plant cell walls contain cellulose fibrils entrenched within a matrix of lignin, hemicelluloses, inorganic solvents, proteins and phenolics. Cellulose is the most abundant polymer on earth (Lerouxel et al. 2006; Taylor 2008) and make up 40% of the dry mass in secondary cell walls compared to 15–30% in primary cell walls (Sticklen 2008). Recently, high-throughput genomic technologies made a significant contribution in understanding the molecular details involved in cellulose biosynthesis (Dhugga 2007; Lerouxel et al. 2006). Genome sequencing revealed that all higher plants have multiple cellulose synthase (CESA) catalytic subunit proteins encoded by the CESA genes.

Unlike cellulose, an unbranched polymer that contains anhydrous glucose, hemicellulose is a branched polymer containing glucose, xylose, mannose, arabinose and galactose. Cosgrove (2005) reported that xyloglucan and arabinoxylan are the most abundant hemicelluloses.

Structural similarities between ß-1,4-glucan chains of cellulose and the backbones of various ß-linked hemicellulosic polysaccharides suggest that cellulose synthase-like (CSL) genes might be involved in the biosynthesis of hemicelluloses (Lerouxel et al. 2006).

Lignin is an amorphous polymer consisting of the three aromatic alcohols (monolignols) *p*-coumaryl, coniferyl and sinapyl alcohols (Buranov and Mazza 2008). The respective aromatic constituents of these alcohols in the polymer are *p*-hydroxyphenyl (H), guaiacyl (G) and syringyl (S) moieties (a part or half a molecule), which are transported from the cytosol to the apoplast. During the lignification process, these monolignols produce a complex three-dimensional amorphous lignin polymer via linkages which lack the regular and ordered units, such as those in cellulose or protein.

Unlike in woody plants, lignin biosynthesis has not been well studied in grasses; the structure and properties of lignin and the precise interrelationships with other cell wall components are not well understood (Buranov and Mazza 2008). Efforts are being made to understand lignin biosynthesis in grasses because of the growing interest in bioethanol production from renewable biomass. Among grasses, maize has received the most attention in genetic and genomic studies related to cell wall lignification and degradability (e.g. Barrierè et al. 2003), and a better understanding of the pathways and genes involved in sorghum will be accumulated because sorghum genome is sequenced (Paterson et al. 2009).

There is a growing interest in enhancing plants for conversion to biofuel, to replace petroleum-based fuels while reducing greenhouse gases. Among the potential plants such as maize, sugarcane and switchgrass, sorghum has attractive characteristics such as high phenotypic and genetic diversity (Assar et al. 2005; Mace et al. 2008) and better tolerance of drought, waterlogging and saline-alkali soil, for development as a model biomass crop (Sarath et al. 2008). Sorghum offers lignocellulosic biomass such as stover, an abundant and inexpensive source of fermentable sugars. Carpita and McCann (2008) reported that biomass quality is influenced by cell wall composition and structure, and can be determined by content and composition of lignin, cellulose and hemicellulose and how they are cross-linked, whereas biomass yield and stability will be determined by agronomic traits such as plant height and diameter, lodging and diseases and pest resistance.

Sugars obtained from cell wall polysaccharides such as cellulose and hemicelluloses can be hydrolysed by hemicellulose- or cellulose-degrading enzymes, a process known as saccharification (Lin and Tanaka 2006). The monosaccharides obtained are fed to microorganisms in fermenters similar to the systems used for starch- or sugarcane-derived sugars. Vermerris et al. (2007) reported that while energy stability of processing stover-to-ethanol has been questioned, technical and economic analysis show that the production of ethanol from lignocellulosics could be beneficial. To make this technology economically competitive, cost-effective and efficient means of transportation and storage of biomass yield and processing strategy must be developed. Besides, it is mandatory to understand the physical and chemical properties, biosynthesis and molecular genomics of plant cell walls and also how to manipulate them in the candidate crop(s).

A successful combination of desired genes and exclusion of deleterious genes have been achieved in many crops through conventional and molecular plant breeding. Improvements in biomass quality in sorghum are dependent on genetic variability within the species, the heritability of the trait(s), selection intensity and the ability of plant breeders to understand the genetic architecture controlling these traits. The objectives of this study were (1) to identify chromosomal regions associated with fibre and their related agronomic traits, (2) to identify epistatic and pleiotropic QTL associated with these traits and (3) to identify stable QTL across environments.

Materials and methods

Plant material

A mapping population consisting of 188 recombinant inbred lines (RIL) derived from a cross between the inbred lines SS79 (sweet sorghum) and M71 (grain sorghum Macia-SA), bred by Dr Willy Wenzel at the Agricultural Research Council of South Africa (ARC), was used for this study. SS79 has lower levels of stem fibre components than M71, such as neutral detergent fibre (NDF), acid detergent fibre (ADF), acid detergent lignin (ADL), cellulose and hemicelluloses but higher levels of stem sugars such as sucrose, glucose and fructose. M71 is characterized by a reduction in fibre yield traits such as plant height, stripped stalk mass, dry stalk mass, fresh leaf mass, fresh biomass and dry biomass compared to SS79. Each RIL was derived from a single F_2 plant following the single seed descent (SSD) method until the $F_{5:6}$ generation (W. Wenzel, pers. communication).

Field layout and data collection

The RIL and the parental lines were sown on 04 May 2007 and 14 May 2008, respectively, at two locations in Germany: Rauischholzhausen/Ebsdorfergrund (2 Mitscherlich pot trials) and Gross Gerau (2 field plot trials). In each environment, experiments were planted in a randomized complete block design (RCBD) with three replications. At

Gross Gerau (GG), a one-row plot of 5 m row length and a three-row plot of 3 m row length were planted in 2007 and 2008, respectively. In both years, the inter-row space was 0.3 m and seeds were sown at 3 cm depth in 0.15 m intervals. The plants were thinned to 10 plants per m^2. In the pot experiment at Rauischholzhausen (RH), two seeds were sown per 20 cm pot at 2 cm depth. Previously, we reported differences in the biophysical data among the four abovementioned environments, such as average cumulative temperature, cumulative rainfall, latitude, longitude, altitude, soil type and soil pH (Shiringani et al. 2010).

Plants were harvested manually by cutting the stem at the base with scissors. Because some lines were photoperiod sensitive, and to harvest the plants at the same physiological stage, plots were harvested 40 days after anthesis. In GG, three plants were sampled at random from the middle of each row for trait measurements and the average from the three plants was used for evaluations. On the other hand, in the pot trials at RH, both plants per container were harvested and results averaged. The entire above-ground plant material was weighed to estimate the fresh biomass. The leaves and panicle of each of the plants were stripped from the stalks. Total fresh leaf weight was recorded per plant followed by weighing the stripped stalks. Panicles, leaves and stripped stalk masses (SSM) were milled and then dried in an oven at 105°C for 2–3 days. After drying, the samples were weighed to estimate the dry stripped stalk mass (DSM) and the dry biomass (DBM).

For fibre content analysis, the ground and dried stripped stalks were further milled to pass a 0.5 mm (GG) or 1 mm (RH) sieve, respectively. From the milled sample, 130 genotypes from GG and RH were selected based on extreme differences among agronomical traits to calibrate and validate near infrared reflectance spectroscopy (FOSS NIRSystem model 6500-C, Silver spring, MD, USA) for ADF, NDF and ADL. Initial quantifications were performed according to van Soest et al. (1991) with minor modifications. NDF estimates the sum of cellulose, hemicellulose and lignin, ADF the sum of cellulose and lignin, while ADL measures lignin only. Therefore, NDF minus ADF yields the hemicellulose content, and ADF minus ADL corresponds to cellulose content.

Subsequent estimates of fibre compounds in the sorghum stalks were obtained via NIRS. Approximately 1.5 g of the milled samples was scanned in the NIRSystem using a 3.7 cm diameter sample holder over a wavelength of 400–2,500 nm with a 2 nm interval. A set of 130 selected genotypes were split into two subsets: (1) a calibration set, in which the experimental values were used to develop the best NIRS calibration equations through WinISI II software (Infrasoft international, LLC, PA, USA) and (2) a validation set, which was used to test the ability of the calibration equations to predict the trait value. The best equation should have low values for the standard deviation (SD), the standard error of calibration (SEC), the standard error of cross-validation (SECV) and the systematic difference between the two sets, respectively, but a higher coefficient of determination in the calibration (R^2). After obtaining the best equations, the samples of 188 genotypes from 4 environments were scanned to estimate fibre component levels. Detailed information on parameters (SD, SEC, SECV, R^2, etc.) required to obtain the best equation for each trait is provided in Supplementary Table S1.

Phenotypic data analysis

Descriptive statistics and analysis of variance were done applying the statistical software SAS® 9.1. The Proc GLM procedure was used to test differences among the RILs across environments. Broad-sense heritability (H^2) on a RIL-mean basis was estimated for all traits, as suggested by Littell et al. (2006) and Fehr (1987). Estimation of the variance components used to calculate H^2 was analysed by restricted maximum likelihood method, where genotypes, harvesting date, environment and genotype × environment interactions were treated as random effects. Coefficients of determination (R^2) were obtained from the linear regression of individual yields in different environments on the mean yields of all the genotypes in each environment (Léon and Becker 1988).

Genotyping

DNA was extracted according to Doyle and Doyle (1990) from the leaf tissue of greenhouse-grown seedlings of parental lines and 188 RILs. Simple sequence repeat (SSR) primer pairs (Bhattramakki et al. 2000; Dean et al. 1999; Kong et al. 2000; Schloss et al. 2002) and EST-SSR markers (Srinivas et al. 2008, 2009) from sorghum were used to screen the parents of the mapping population. The PCR protocol and thermocycler touchdown conditions were adopted from Hasan et al. (2006) and Xu et al. (2005). Most of the SSR primers were labelled with M13 tailing as described by Berg and Olaisen (1994). The fluorescently labelled universal M13 primer 5′-AGGGTTTTCCCAGT CACGACGTT-3′ was added to the PCR reaction, and the forward primer of each SSR was appended with the sequence 5′-TTTCCCAGTCACGACGTT-3′. After the first cycle of PCR, fragments were successively amplified with the labelled primer (Hasan et al. 2008; Rygulla et al. 2007).

Amplified fragment length polymorphism (AFLP) amplification was performed according to Vos et al. (1995). During the amplification steps, the sequences of the primers +0 EcoR1, +0 Mse1, +1 EcoR1 and +1 Mse1 were 5′-GAC TGC GTA CCA ATT C-3′, 5′-GAT GAG

TCC TGA GTA A-3′, 5′-GAC TGC GTA CCA ATT CA-3′ and 5′-GAT GAG TCC TGA GTA AC-3′, respectively, as reported by Uptmoor et al. (2003).

All SSR and AFLP amplification products were separated and visualised using the LI-COR 4200 DNA Analyser (MWG Biotech, Ebersberg, Germany).

Linkage map construction and QTL analysis

Linkage analysis was performed applying Joinmap® 4.0 software using 215 markers and 188 RILs. Individual markers missing more than 25% of genotype data and those that showed segregation distortion were eliminated from the data set. Loci were assigned to linkage groups (LG) with a minimum LOD score of 3.0 and a maximum Kosambi distance of 40 centiMorgans (cM). The linkage map was compared to the sorghum maps published by Bhattramakki et al. (2000) and Mace et al. (2009) and assigned to chromosomes according to the sorghum chromosome nomenclature of Kim et al. (2005).

QTL analysis was carried out using composite interval mapping (CIM) within the software PLABQTL version 1.2 (Utz and Melchinger 1994), which searches for QTL that are evaluated subsequently in multiple regression models. To generate empirical thresholds for LOD values, 1,000 permutations were performed for all traits. A LOD threshold value of 3 was found to detect QTL at a standard genome-wide significance of 5% ($\alpha_g = 0.05$). All markers suggested by default in the first stepwise regression were used as cofactors. These markers were chosen based on per cent of phenotypic variation explained and simple correlation of the phenotypic observation with the selected marker. At LOD 3, all putative QTL associated with the traits are identified; from these putative QTL, those that show a significant likelihood ratio that a QTL is present/absent were identified based on the multiple regression of the final simultaneous fit function of the software, herein reported as 'detected QTL'. After a further regression analysis, some of the detected QTL were declared significant based on the phenotypic variance explained and QTL effect, herein reported as 'significant QTL'. Additive × additive QTL (*aa*) in all traits were calculated according to MODEL AA of the PLABQTL. The interactions were estimated from the set of detected QTL in a stepwise regression based on the f-to-enter value by using the Bonferroni bound at $\alpha = 0.05$. For QE analysis, the detected QTL undergo a simultaneous fit for each environment and the results are presented in an ANOVA. The difference between the fits of the data from individual environments and the means across environments gives the mean squares of QE interactions. The significance of QE interactions was tested using an F-test with Bonferroni adjustment.

To identify putative candidate genes involved in lignin and cellulose biosynthesis, we first identified the physical position of the QTL confidence intervals (http://www.phytozome.net/sorghum.php) by BLASTing the sequence of the microsatellites that are closest to the QTL position and by using the in-silico SSR map published by Li et al. (2009). To 'Free text' search for major biochemical pathway genes involved in lignin and cellulose biosynthesis, key words were applied to identify the physical position of these genes based on the updated sorghum annotation (http://mips.helmholtz-muenchen.de/plant/sorghum).

Results

Phenotyping results

NIRS results (Supplementary S1; Table 1) showed that the distribution of the reference values in the calibration and validation sets covered a great variability, suggesting that the sampled genotypes represented the overall genetic/phenotypic diversity of this population. NIRS equations

Table 1 Mean values, standard deviations and range for fibre components and fibre-related agronomic traits in recombinant inbred lines across environments

Trait	Parental lines		Recombinant inbred lines (RILs)			
	SS79	M71	Mean	Standard deviation	Minimum	Maximum
ADF (g/dry weight)[a]	29.72	36.11	33.67	2.37	28.27	40.65
ADL (g/dry weight)[a]	5.76	7.50	7.23	0.64	5.73	8.86
NDF (g/dry weight)[a]	46.58	57.99	52.01	3.19	44.93	60.81
Cellulose (g/dry weight)[a]	24.03	29.23	26.44	2.03	22.03	33.12
Hemicellulose (g/dry weight)	16.53	19.90	18.34	1.01	16.11	21.93
Fresh leaf mass (g/plant)	122.48	59.85	106.96	22.61	60.42	192.59
Stripped stalk mass (g/plant)	642.54	191.57	375.86	140.02	118.04	754.90
Dry stalk mass (g/plant)	73.65	35.28	53.90	21.48	7.44	123.26
Fresh biomass (g/plant)	1,102.74	472.72	570.76	130.1	277.46	944.09
Dry biomass (g/plant)	177.94	81.02	92.34	22.82	25.59	158.67

[a] Stripped sorghum stalks

Table 2 Mean squares, heritabilities, repeatabilities and coefficient of variation for fibre components and fibre-related agronomic traits in recombinant inbred lines across environments

Trait	Mean squares					h_B^{2c}	R^2
	Genotype	Environment	G × E[a]	Harvest date	CV[b]		
ADF (g/dry weight)[d]	***	***	***	***	6.14	0.84	0.92
NDF (g/dry weight)[d]	***	***	***	NS	10.64	0.70	0.76
ADL (g/dry weight)[d]	***	***	***	***	10.35	0.67	0.84
Cellulose (g/dry weight)[d]	***	***	***	***	6.89	0.82	0.91
Hemicellulose (g/dry weight)	**	***	NS	NS	27.32	0.24	0.24
Fresh leaf mass (g/plant)	***	***	***	NS	34.37	0.51	0.61
Stripped stalk mass (g/plant)	***	***	***	NS	33.84	0.79	0.90
Dry stalk mass (g/plant)	NS	NS	NS	NS	–	0.005	0.004
Fresh biomass (g/plant)	***	***	***	NS	33.49	0.34	0.52
Dry biomass (g/plant)	***	***	***	NS	28.29	0.13	0.36

*** Significant at 0.001 probability level, NS non significant level at 0.05 probability level

[a] *G* genotype, *E* environment

[b] *CV* coefficient of variation

[c] Broad-sense heritability on an entry-mean basis

[d] Stripped sorghum stalks

showed better prediction information for NDF ($r^2 = 0.92$) and ADF ($r^2 = 0.88$) than for ADL ($r^2 = 0.72$). These r^2 values are similar to what has been reported by Murray et al. (2008) for stem NDF, ADF and ADL.

The average performance of parents and RILs across locations is presented in Table 1. NIRS results showed that the male parent M71 has higher stem fibre content such as cellulose, hemicellulose, NDF, ADF and ADL contents than the female parent SS79, which is characterized by a high stem sugar content. Although cell wall components were not quantified in these parents before, this was expected because sweet sorghum genotypes have been selected in Limpopo, the South African region where SS79 originates, to be easy-to-peel by teeth. In terms of biomass components and/or total biomass, SS79 weighed more than M71. The average performance of the RILs fell within the range of the parents, although transgressive segregation was also observed.

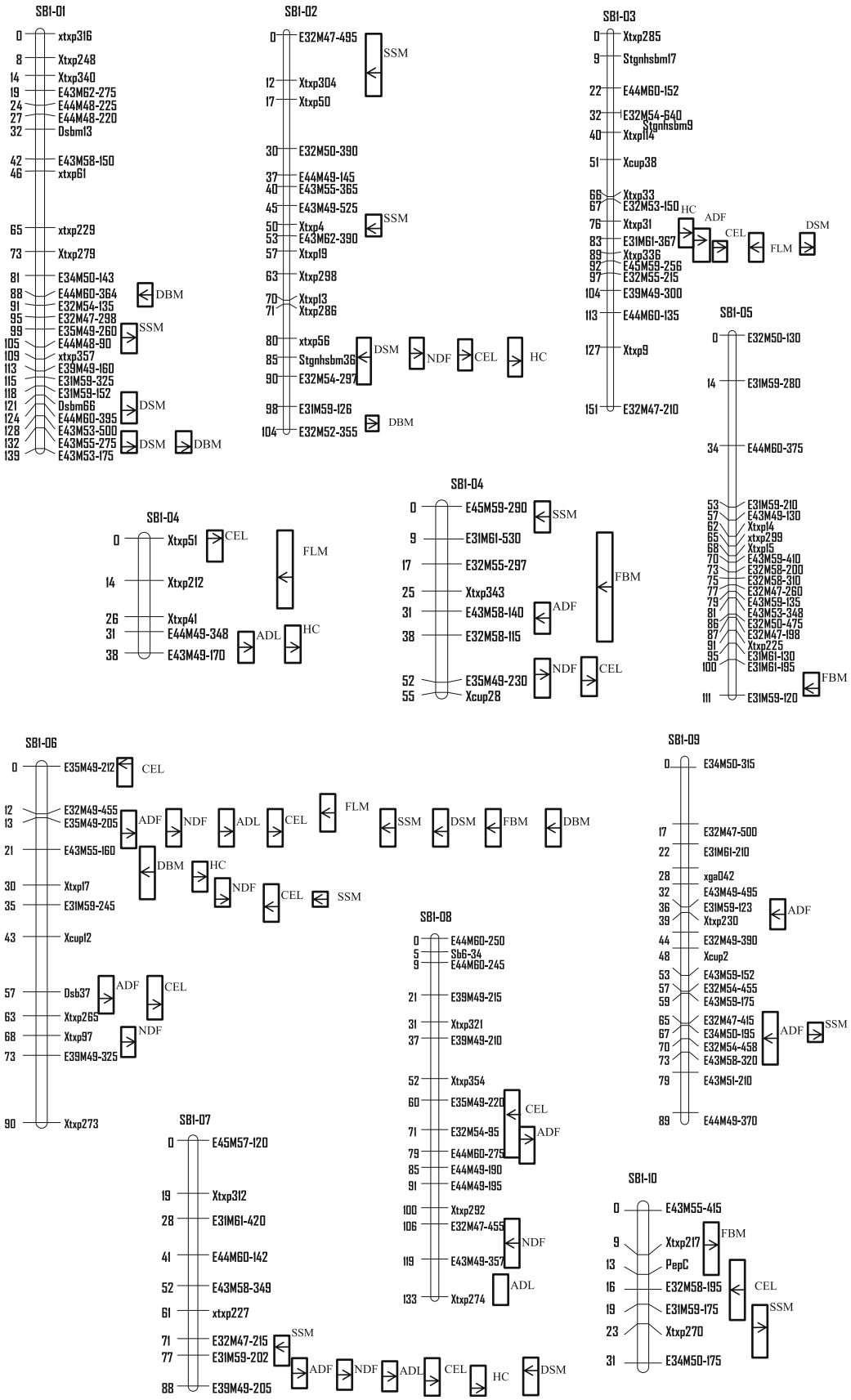
The mean squares, heritabilities and repeatabilities for the trait data are shown in Table 2. Analysis of variance using the general linear model showed that genotypes and environments were highly significant ($P < 0.001$) in all the traits investigated except for dry stalk mass. There was a significant interaction between genotype and environment ($P < 0.001$) in all traits except hemicellulose and dry stalk mass. Although the harvest date was expected to be the most important source of variation in most traits, it was significant only for ADF, ADL and cellulose. The coefficient of variation ranged from 6.14 to 34.37%, with ADF showing the lowest variation and fresh leaf mass the highest (Table 2).

High broad-sense heritability values were estimated for ADF ($H^2 = 0.84$), cellulose ($H^2 = 0.82$), stripped stalk mass ($H^2 = 0.79$) and NDF ($H^2 = 0.70$), while low values were observed for dry stalk mass ($H^2 = 0.005$), dry biomass ($H^2 = 0.13$) and hemicellulose ($H^2 = 0.24$). Traits with higher heritability also showed higher repeatability (Table 2). Low heritability and repeatability can be the result of sampling error coupled with environmental effects, which might lessen success in enhancing these traits.

Linkage map construction and QTL analysis

The genetic map was constructed from M71 × SS79 RILs and consists of 157 AFLP, SSR and EST-SSR markers (Fig. 1). A total of 59 markers were excluded because of either high missing values or divergence from the expected segregation ratio. The map covered 1,029 cM with an average distance of 6.55 cM. The significant QTL detected by CIM are shown in Fig. 1 and full information of the detected additive QTL is presented in Supplementary Table S2. Environments 01, 02, 03 and 04 throughout the chapter represent the field experiments in GG 2007, GG 2008, RH 2007 and RH 2008, respectively. A supplementary file (Supplementary info 1) showing the results of QTL detected at each environment is also supplied. This file can add a valuable input for the sorghum research community in terms of QTL comparison. The QTL described in this manuscript are those obtained by analysing the average data of each trait because we aim at identifying robust QTL that are less specific to environments.

A total of 72 additive QTL associated with fibre quality traits were detected on 10 chromosomes using composite interval mapping. Of those, 17 additive QTL were detected


SB1-01
0 xtxp316
8 Xtxp248
14 Xtxp340
19 E43M62-275
24 E44M48-225
27 E44M48-220
32 Dsbm13
42 E43M58-150
46 xtxp61
65 xtxp229
73 Xtxp279
81 E34M50-143
88 E44M60-364
91 E32M54-135
95 E32M47-298
99 E35M49-260
105 E44M48-90
109 xtxp357
113 E39M49-160
115 E31M59-325
118 E31M59-152
121 Dsbm66
124 E44M60-395
128 E43M53-500
132 E43M55-275
139 E43M53-175
DBM
SSM
DSM
DSM
DBM
SB1-02
0 E32M47-495
12 Xtxp304
17 Xtxp50
30 E32M50-390
37 E44M49-145
40 E43M55-365
45 E43M49-525
50 Xtxp4
53 E43M62-390
57 Xtxp19
63 Xtxp298
70 Xtxp13
71 Xtxp286
80 xtxp56
85 Stgnhsbm36
90 E32M54-297
98 E31M59-126
104 E32M52-355
SSM
SSM
DSM
NDF
CEL
HC
DBM
SB1-03
0 Xtxp285
9 Stgnhsbm17
22 E44M60-152
32 E32M54-640
40 Xtxp114
51 Xcup38
66 Xtxp33
67 E32M53-150
76 Xtxp31
83 E31M61-367
89 Xtxp336
92 E45M59-256
97 E32M55-215
104 E39M49-300
113 E44M60-135
127 Xtxp9
151 E32M47-210
Stgnhsbm9
HC
ADF
CEL
FLM
DSM
SB1-04
0 Xtxp51
14 Xtxp212
26 Xtxp41
31 E44M49-348
38 E43M49-170
CEL
FLM
ADL
HC
SB1-04
0 E45M59-290
9 E31M61-530
17 E32M55-297
25 Xtxp343
31 E43M58-140
38 E32M58-115
52 E35M49-230
55 Xcup28
SSM
FBM
ADF
NDF
CEL
SB1-05
0 E32M50-130
14 E31M59-280
34 E44M60-375
53 E31M59-210
57 E43M49-130
62 Xtxp4
65 xtxp299
68 Xtxp15
70 E43M59-410
73 E32M58-200
75 E32M58-310
77 E32M47-260
79 E43M59-135
81 E43M53-348
86 E32M50-475
87 E32M47-198
91 Xtxp225
95 E31M61-130
100 E31M61-195
111 E31M59-120
FBM
SB1-06
0 E35M49-212
12 E32M49-455
13 E35M49-205
21 E43M55-160
30 Xtxp17
35 E31M59-245
43 Xcup12
57 Dsb37
63 Xtxp265
68 Xtxp97
73 E39M49-325
90 Xtxp273
CEL
ADF
NDF
ADL
CEL
FLM
SSM
DSM
FBM
DBM
DBM
HC
NDF
CEL
SSM
ADF
CEL
NDF
SB1-07
0 E45M57-120
19 Xtxp312
28 E31M61-420
41 E44M60-142
52 E43M58-349
61 xtxp227
71 E32M47-215
77 E31M59-202
88 E39M49-205
SSM
ADF
NDF
ADL
CEL
HC
DSM
SB1-08
0 E44M60-250
5 Sb6-34
9 E44M60-245
21 E39M49-215
31 Xtxp321
37 E39M49-210
52 Xtxp354
60 E35M49-220
71 E32M54-95
79 E44M60-275
85 E44M49-190
91 E44M49-195
100 Xtxp292
106 E32M47-455
119 E43M49-357
133 Xtxp274
CEL
ADF
NDF
ADL
SB1-09
0 E34M50-315
17 E32M47-500
22 E31M61-210
28 xga042
32 E43M49-495
36 E31M59-123
39 Xtxp230
44 E32M49-390
48 Xcup2
53 E43M59-152
57 E32M54-455
59 E43M59-175
65 E32M47-415
67 E34M50-195
70 E32M54-458
73 E43M58-320
79 E43M51-210
89 E44M49-370
ADF
ADF
SSM
SB1-10
0 E43M55-415
9 Xtxp217
13 PepC
16 E32M58-195
19 E31M59-175
23 Xtxp270
31 E34M50-175
FBM
CEL
SSM

◀ **Fig. 1** Genetic linkage map of the M71 × SS79 RIL population consisting of 102 AFLP, 49 SSR and 6 EST-SSR markers. Markers prefixed with *E* are AFLPs, *X* and *Sb* are SSR markers, while *Stg* and *Dsb* are EST-SSRs. QTL were detected by composite interval mapping (CIM). All QTL shown exhibited significant effects in different environments and also using the combined environments (4) for each trait. *ADF* acid detergent fibre, *NDF* neutral detergent fibre, *ADL* acid detergent lignin, *CEL* cellulose, *HC* hemicelluloses, *FBM* fresh biomass, *DBM* dry biomass, *DSM* dry biomass, *FLM* Fresh leaf mass, *SSM* stripped stalk mass. The *top* and *bottom* borders of the QTL *boxes* represent the approximate support intervals of the detected QTL associated with the respective trait on the right side of the *box*. The peak of each QTL is shown by the position of the *arrow* inside the *box*, and the direction indicates the allele contribution by parent SS79 (pointing left) or by parent M71 (pointing right)

on all chromosomes associated with ADF. Eight of these QTL, with LOD scores ranging from 3.14 to 7.99, significantly influenced the phenotype across environments. A QTL with a LOD score of 7.99 on SBI-03 position 86 cM contributed most to the phenotypic variation (17.8%), followed by a QTL with LOD score 6.91 on SBI-06 position 58 cM, which contributed 15.6%. Higher additive effects among the detected QTL were found on SBI-06, left flanked by E35M49-205 (1.14 g/dry weight), and SBI-07, left flanked by E31M59-202 (1.21 g/dry weight). The positive sign of the additive effects indicates that the positive alleles were contributed by the M71 parent.

A total of 14 additive QTL, distributed on all chromosomes, were detected associated with NDF. Seven significant QTL influenced NDF in different environments. The phenotypic variation explained by the detected QTL ranged from 8.3 (LOD score 3.52 on SB1-06, left flanked by Xtxp265) to 18.8% (LOD score 8.18 on SBI-07, left flanked by E31M59-202). Three QTL detected on SBI-02, SBI-06 and SBI-07 showed high additive effects of 1.21, 1.79 and 1.23 g/dry weight, respectively. These effects showed the contribution of the M71 alleles in increasing neutral detergent fibre. For ADL, five out of 15 detected additive QTL were significant in influencing the phenotype across the environments; significant QTL were detected on SBI-04, SBI-06, SBI-07 and SBI-08. The phenotypic variation explained by all detected QTL ranged from 7.1 (LOD score 3.1 on SBI-09 at 72 cM) to 18.9% (LOD score 4.28 on SBI-08 at 130 cM).

A total of 16 additive QTL controlling cellulose content were detected distributed on all chromosomes. The largest cluster was observed on SBI-06, where four QTL significantly influenced the phenotype either in environments 01, 02 or 04. In total, 11 out of 16 QTL significantly influenced the phenotype. The phenotypic variation explained by these QTL ranged from 7.4 to 17.6%. A total of 67% of the detected QTL showed a positive sign of the additive effects value, emphasising the contribution of alleles from the grain sorghum parent M71 in increasing cellulose quantity.

At a LOD of 3.0, eight QTL were detected associated with hemicelluloses content. These QTL were distributed across all chromosomes except chromosome 1 and 9. Five of the 8 detected QTL significantly influenced the phenotype in different environments. The detected QTL explained between 7.4 and 23.6% of the phenotypic variation, with the highest contribution from a QTL at 28 cM on SBI-06 with LOD score 10.99 and an additive effect of 0.448 g/dry weight. Most of the QTL (7 of 9) showed a positive additive effect value, emphasizing the contribution of the M71 parent in increasing hemicelluloses content.

A total of 56 additive QTL controlling fibre-related traits were found across all chromosomes. Of those, five QTL associated with fresh leaf mass were detected on SBI-02, SBI-03, SBI-04 and SBI-06. Three QTL significantly influenced the phenotype in overall environments, particularly in environments 01, 02 and 03. The phenotypic variation explained by these QTL was 47.8%. The sign of additive effect values of all the detected QTL was negative, suggesting that SS79 contributed most of the alleles for increasing fresh leaf mass.

A total of 15 additive QTL, distributed on all chromosomes except SBI-08, were detected associated with stripped stalk mass. Ten of these QTL showed significant effects on the trait across environments. Clusters with three QTL each were observed on SBI-01 and SBI-06, whereby all QTL in the cluster on SBI-06 were significant. The phenotypic variation explained by these QTL ranged from 7.9 to 17.4%. The additive effects of the detected QTL showed that both parents contributed alleles for decreasing or increasing stripped stalk weight.

Dry stalk mass was shown to be significantly influenced by six out of 10 additive QTL detected on 5 chromosomes in all environments. The phenotypic variation explained by all of those QTL ranged from 7.0 to 16.7%. The highest phenotypic variation was explained by a QTL at 14 cM on SBI-06 with a LOD score of 7.46, followed by a QTL at 82 cM on SBI-07 (LOD 5.72). Two out of 10 QTL showed positive additive effects, indicating that most of the alleles were contributed by SS79, the parent with high stalk mass.

For fresh biomass, 10 additive QTL were detected where 7 significantly influenced the trait across all environments. These significant QTL were distributed on six chromosomes, whereby two were found on chromosome SBI-01. The phenotypic variation explained by these QTL ranged from 7.8 to 24.7%. Additive effects of the detected QTL ranged from −0.04 to −52.56 g/plant favouring SS79, the high biomass parent, and 2.84 to 57.50 g/plants favouring M71, the low biomass parent.

A total of 16 QTL associated with dry biomass were detected on 8 chromosomes. Clusters with 4 QTL each were found on chromosomes SBI-01 and with 3 QTL were found on SBI-02; five of the 16 QTL were found significant

Table 3 Digenic epistasis associated with fibre-related traits

Trait	Chrom	Left marker	Chrom	Left marker	Add. effect	Mean	R^2%
ADF (g/dry weight)[a]	None						
NDF (g/dry weight)[a]	SBI-01	*E43M53-500*	SBI-03	*E31M61-367*	−0.77	51.96	6.7
ADL (g/dry weight)[a]	SBI-02	*E32M47-495*	SBI-07	*E31M59-202*	−0.12	7.26	3.3
Cellulose (g/dry weight)[a]	SBI-01	*E35M49-260*	SBI-05	*E44M60-375*	0.50	26.59	7.4
	SBI-01	*Xtxp279*	SBI-06	*Dsb37*	−0.42		5.0
Hemicelluloses (g/dry weight)[a]	SBI-02	*Stgnhsbm36*	SBI-05	*E31M59-210*	−0.27	18.46	7.6
	SBI-4b	*E44M49-348*	SBI-05	*E31M59-210*	−0.23		5.4
	SBI-4b	*E44M49-348*	SBI-06	*E43M55-160*	−0.33		11.3
	SBI-07	*E31M59-202*	SBI-08	*E39M49-215*	0.24		6.0
Fresh leaf mass (g/plant)	None						
Stripped stalk mass (g/plant)	SBI-01	*E35M49-260*	SBI-09	*E43M59-152*	31.28	362.29	5.1
	SBI-01	*Xtxp279*	SBI-06	*Xtxp17*	41.15		5.5
Dry stalk mass (g/plant)	SBI-01	*E43M53-175*	SBI-08	*Xtxp292*	−4.36	51.12	3.4
	SBI-02	*Xtxp56*	SBI-07	*E31M59-202*	−4.52		3.8
	SBI-4b	*E44M49-348*	SBI-06	*E43M55-160*	5.38		6.8
Fresh biomass (g/plant)	SBI-01	*E43M53-175*	SBI-09	*E31M59-123*	29.14	559.48	4.3
	SBI-01	*E35M49-260*	SBI-02	*Xtxp50*	−36.24		4.3
	SBI-01	*E35M49-260*	SBI-09	*E32M47-415*	29.72		4.4
	SBI-01	*Xtxp248*	SBI-10	*E43M55-415*	−26.96		3.7
Dry biomass (g/plant)	None						

ADF Acidic detergent fibre, *NDF* Neutral detergent fibre, *ADL* Acidic detergent lignin, *Chrom* Chromosome, R^2% phenotypic variation explained in percentage, *Add. effect* additive effect

[a] Stripped sorghum stalks

in combined environments or environment 01. The phenotypic variation explained by these QTL ranged from 7.5 to 12.5% and the additive effects showed that the two parents each contributed alleles with both positive and negative effects on dry biomass.

QE interaction and epistasis

QE interaction is an important component of quantitative traits. QTL detected in one environment but not in another might indicate QE interactions (Veldboom and Lee 1996). A total of 25 QTL showing significant QE interactions were detected across all traits except for hemicellulose, NDF and dry biomass (Supplementary Table S2). Dry biomass was quantified in only 2 environments, which may explain why no QE interaction was detected. A total of 7 QTL showing significant QE interactions were detected for stripped stalk mass, 4 each for dry stalk mass and fresh biomass, 3 were for fresh leaf weight, 3 for cellulose and 2 for ADF and ADL. QTL with significant QE interaction also showed pleiotropic effects. For example, the QTL at 14 cM on SBI-06 exhibited significant QE interaction for almost all traits.

Epistasis is important for studying complex traits, and genetic models for QTL mapping that assume lack of epistasis can lead to a biased estimation of QTL parameters (Cao et al. 2001). A total of 17 significant digenic, additive × additive interactions were detected for the studied cell wall components and related agronomic traits (Table 3). Most of the additive × additive interactions detected involved SBI-01, showing the importance of this chromosome in digenic interactions. A total of 4 digenic epistatic QTL were detected associated with hemicellulose and fresh biomass, respectively, 3 with dry stalk mass, 2 each with stripped stalk mass and cellulose and one each with ADL and NDF. No significant additive × additive interactions were detected for ADF, dry biomass or fresh leaf mass. According to Cao et al. (2001), the epistatic pairwise QTL detected for a particular trait depend on the developmental stage of the plants. The authors concluded that interaction between QTL and background or modifying loci might be the prevalent form of epistasis affecting the behaviour of quantitative traits.

The parts of total phenotypic variation explained by interactions were as follows: for ADL, 3.3% explained by one digenic interaction; for NDF, 6.7% explained by one digenic interaction; for stripped stalk mass, 10.6% explained by two digenic interactions; for cellulose, 12.4% explained by two digenic interactions; for dry stalk mass, 14% explained by three digenic interactions; for

hemicellulose, 30.3% explained by 4 digenic interactions; and for fresh biomass, 16.7% were explained by 4 digenic pairs. Thus, the total phenotypic variation explained by the interactions was lower than that caused by the corresponding main effects. The importance of additive × additive interaction effects among total genetic effects is therefore expected to be trait-dependent.

Co-localization and pleiotropic effects

A number of QTL clusters with effects on multiple traits were detected on different chromosomes (Supplementary Table S2, and Fig. 1). The largest clusters were observed on SBI-03 (83 cM), SBI-06 (14 cM) and SBI-07 (77 cM). The QTL on SBI-06 influenced all of the traits except hemicellulose and fresh leaf mass, explaining 8.4 to 17.4% of the phenotypic variation with lowest effect on cellulose and highest on stripped stalk mass. The QTL on SBI-07 influenced all traits except dry biomass, stripped stalk mass and fresh leaf mass and explained 9.0 to 18.9% of the phenotypic variation for the various traits.

In the two clusters observed on SBI-06 and SBI-07, the additive effect sign for cell wall components (cellulose, hemicelluloses, ADL, ADF, NDF) was opposite to that of agronomic yield (fresh and dry biomass, stripped stalk mass and dry stalk mass). This indicates that parents contributed in opposition to increasing fibre content and agronomic mass. These chromosomal regions also contributed to ADL content, which is not necessarily desired by the breeder because lignin prohibits saccharification of cellulose and hemicelluloses. However, distinguishing between pleiotropy and linkage is difficult and would require validation by conducting a large number of crosses that may or may not produce recombinants.

Discussion

The increasing demand for food and rising concerns about climate change and energy security pushes *Sorghum bicolor* to the top of global agendas. The main attraction is its ability to provide renewable energy products, industrial commodities, as well as food and animal feed. Carpita and McCann (2008) reported that sorghum, like maize, has the potential to provide abundant and sustainable resources of lignocellulosic biomass for the production of ethanol as biofuel. Furthermore, it can also serve as a genetic model for the improvement of C_4 bioenergy grasses. Both maize and sorghum are annual grasses and thus give farmers flexibility for crop production. Maize has a historical depth of genetic knowledge, and the genomes of both sorghum and maize have recently been sequenced (Paterson et al. 2009; Schnable et al. 2009).

In this study, we quantified and detected QTL that are linked with cell wall content and biomass yield traits. Until now, there is limited literature in this field. To our knowledge, only Murray et al. (2008) conducted a similar study using 176 inbred lines derived from a cross between grain and sweet sorghum. As in our work, Murray et al. (2008) reported QTL associated with cellulose content and hemicelluloses on SBI-03, SBI-07 and SBI-08, QTL associated with ADL on SBI-03, SBI-06, SBI-07 and SBI-08, and QTL associated with NDF on SBI-03 and SBI-07. Both studies identified two QTL on SBI-08 for cellulose, ADL and NDF; however, in the present study, different additive effects of the QTL were observed in some cases.

In comparison to other crop plants, maize has received particular attention regarding fibre content research. Before biofuel gained attractiveness, maize cell wall components were studied because of digestibility in ruminants (Moore and Hatfield 1994; Cardinal et al. 2003). Comparative genomics studies indicated that sorghum is closely related to maize, having diverged from each other approx. 11.9 Mya (Swigonova et al. 2004; Moore et al. 1995). Krakowsky et al. (2005) studied a population of 191 maize RILs for ADL, ADF and NDF content in stem tissues. They reported 16 QTL associated with NDF, 18 QTL with ADF and 10 QTL with ADL. Using the comparative map published by Mace et al. (2009), we could identify where some of these loci map to the sorghum genome. In maize, the loci *umc58* and *umc34* were shown to be responsible for increases in NDF, NDF and ADF and explained 7% (umc58/NDF), 23% (umc34/NDF) and 20% (umc34/ADF) of the phenotypic variation, respectively (Krakowsky et al. 2005). These two markers map on sorghum chromosome SBI-06 (Mace et al. 2009). We have identified significant QTL on chromosome SBI-06 explaining 12.8 and 10.1% of the phenotypic variation for increases in ADF and NDF, respectively. The loci *umc4* and *phi12*, mapping on maize chromosome 2, are associated with an increase in NDF, ADF and ADL, explaining 13, 10 and 16% of the phenotypic variation in maize, respectively. The two markers, *umc5* and *umc4*, flanking these QTL were mapped on sorghum chromosome SBI-02; thus, we can hypothesise that this region in maize is synthenic to that on SBI-02 of sorghum. We have identified QTL on SBI-02 at position 72–88 cM that are associated with increased ADF, NDF and ADL, explaining 11.8, 16.5 and 7.2% of the phenotypic variation, respectively. Although genome reorganisations such as gene amplification, movement rearrangements or retrotransposition (Feuillet and Keller 2002; Song et al. 2002) may frequently happen, such particular regions could have been conserved during evolution.

Murray et al. (2008) also found co-localization of QTL associated with fibre-related traits. However, unlike in our study, their QTL were distributed across only nine

chromosomes and no QTL were detected on chromosome SBI-02. In our study, using an additive × additive model resulted in detection of more QTL with high additive effects. Additive effects are particularly important for breeding because they can be directly exploited in selection for the respective trait(s). In maize, Cardinal et al. (2003) reported clustering and overlapping of QTL for cell wall components such as lignin, NDF and ADF in the stalk, and Barrierè et al. (2008) reported co-localization of QTL associated with lignin content, lignin monomer composition and cell digestibility. In our study, co-localization and overlapping of QTL were also observed. This suggests that genes controlling cell wall components may either be linked in the sorghum genome or act in a pleiotropic manner.

Flowering date and plant height are important characteristics in crop production and breeding (Lin et al. 1995). These traits correlate with various physiological and quality traits in different species because different sets of alleles or loci are being expressed under different growing conditions. In previous work (Shiringani et al. 2010), we reported a weak but positive correlation between plant height and flowering date in this population ($r = 0.183$, $P < 0.01$), both of which were positively correlated with sugar components (Brix, glucose content, sucrose content and total sugar content). In the present study, we found that plant height and flowering date are negatively related with fibre components (data not shown). Thus, early flowering and short plants are important selection characteristics for fibre components particularly in this population where the female parent SS79 has lower fibre content, higher sugar content and is taller than the male parent M71.

QTL for plant height and flowering date co-localized with those of sugar and fibre components across the genome (Lin et al. 1995; Pereira and Lee 1995; Murray et al. 2008). For the same sorghum population, we previously reported five QTL associated with flowering date on SBI-03, SBI-04, SBI-06, SBI-06, SBI-07 and SBI-08 (Shiringani et al. 2010). The QTL on SBI-03 were associated with reduced time to flowering, while the QTL on SBI-07 were associated with retarded flowering as well as with a decrease of glucose and sucrose content. Thus, some of the genotypes in the population may carry a combination of alleles from both parents for these two QTL. Such genomic loci that showed multiple effects between plant height, sugar and fibre components were detected on SBI-02, SBI-06 and SBI-07.

When considering the markers flanking the QTL, some loci were specific to the traits in this population (Supplementary Table S2; Shiringani et al. 2010) and thus should be investigated further for specific trait selection and improvement of sorghum. For example, marker *Xtxp285* on SBI-03 was associated only with early flowering date, *E44M48-225* on SBI-01 was associated solely with high ADF content, *Xtxp212* on SBI-04b was only associated with an increase in plant height, *Xtxp265* on SBI-06 was only associated with high NDF content, *E35M49-212* on SBI-06 was associated with decreased cellulose, respectively, and *Xtxp336* on SBI-03 was associated with high ADL content only. On the other hand, although these are considered promising markers, they should be used with caution for selection purposes because of overlaps in the QTL confidence intervals with other traits.

The role of epistasis in crop improvement is captivating because it is evident that gene interaction may have a strong effect on complex traits. Although epistasis is difficult to estimate, understanding the amount and type of epistasis present can enhance the reliability of predictions and the design of plant breeding programmes (Lamkey and Lee 1993). Using suitable software such as PLABQTL, it is possible to detect QTL with additive and epistatic effects along with QE interactions. In our study, epistasis was detected in seven out of the 10 traits studied. However, the inability to detect epistasis cannot be taken as evidence for the total absence of epistasis, because of possible swapping or even cancelling of epistatic effects of different loci, genotype × environmental interactions, sample size and estimates using the analysis of variance approach (Lamkey and Lee 1993). The genetic basis of heterosis has been attributed to overdominance at a single locus, or dominance complementation, along with multilocus epistasis (Garcia et al. 2008; Li et al. 2001). The results of additive × additive interactions in our study showed the contribution of both parents, while the additive effect per trait was not consistent, but rather differed among pairs. However, it is desired to fix additive × additive effects that increase lignocellulosic biomass for marker-assisted selection.

The reference genome of Sorghum BTx623 has been sequenced and annotated (Paterson et al. 2009). This sequence information is a great asset for genetic analysis and molecular breeding approaches such as identification of candidate genes or regulatory elements in QTL hot spot regions of the genome. There have been several successful reports on using reference genome sequence to identify functional candidate genes associated with carbohydrate traits after QTL were identified using biparental QTL mapping, association mapping and QTL expression (Prioul et al. 1999; Wilson et al. 2004; Thomas et al. 2010). We have found a possible candidate gene on SBI-02 at the interval 72–88 cM, encoding the catalytic subunits of the cellulose synthase-like (CSL) protein (OsCS1F3), which is particularly involved in hemicellulose biosynthesis. The CSL proteins have amino acid sequences similar to those of cellulose synthase (CESA) enzymes that are involved in cellulose biosynthesis. QTL for cellulose and hemicelluloses overlapped on SBI-02, SBI-03 and SBI-06, and these

regions seem to harbour corresponding genes encoding the CESA/CSL enzymes. During cellulose biosynthesis, at least three CESA/CSL proteins are involved in the same cell and developmental stage. The gene(s) encoding the enzyme UDP-glucosyltransferase are also located within the clusters on SBI-02, SBI-03 and SBI-06. This UDP-glucosytransferase in cellulose biosynthesis is reported to be specific to the Golgi apparatus (http://www.gramene.org/Sorghum_bicolor), where the hexametric complexes (rosettes) that consist of six functional units of glucans are assembled.

Various candidate genes associated with lignin biosynthesis are possibly located in the regions of these sorghum QTL. We identified a gene associated with cinnamate-4-hydroxylase (C4H), a key enzyme in lignin biosynthesis that catalyses the conversion of cinnamate into 4-hydroxycinnamate, on SBI-03 (position 90–96 cM) and SBI-04b (position 30–38 cM). Down-regulation of this gene resulted in a decrease of lignin content in alfafa (Reddy et al. 2005). A gene associated with 4-coumarate-CoA ligase2 (4CL), an enzyme that catalyses the conversion of hydroxycinnamic acids to hydroxycinnamyl-CoA thioesters, was found on SBI-04b (position 30–38 cM). Hu et al. (1999) reported that down-regulation of the gene encoding 4CL in aspen decreased lignin but increased cellulose content. The gene controlling the enzyme similar to caffeoyl-CoA O-methyltransferase (CCOMT) that is involved in the methylation of monolignol caffeoyl-CoA OMTs was found on SBI-07. Blaschke et al. (2004) reported that down-regulating the CCOMT gene increased the stem biomass in tobacco without a change of lignin content, while Guo et al. (2001) reported a decrease of lignin in alfalfa. These results suggest that in the clusters on SBI-02, SBI-03 and SBI-06, unlike a common gene that regulates these traits, a set of tightly linked genes controlling fibre content may be grouped in some or all of these chromosomes.

When searching for the regions that involve digenic epistasis, we found cellulose synthase-like A4 on SBI-01 and cellulose synthase-like H1 on SBI-06. The interval of the locus E32M47-495 on SBI-02 detected during digenic epistasis coincides with the location of genes controlling ferulate-5-hydroxylase and 4-coumarate-CoA ligase. Saballos et al. (2009) reported that five out of 14 genes enconding cinnamyl alcohol dehydrogenase (CAD) in sorghum are found in this region. The CAD enzyme catalyses the reduction of cinnamyl aldehyde precursors, the last step in monolignol biosynthesis. The down-regulation of the gene encoding CAD resulted in a decrease of lignin content but did not affect plant development in tobacco (Chabannes et al. 2001). As we detected this region based on the additive × additive model, it validates that valuable gene loci likely to be missed by QTL analysis might be detected by digenic epistasis.

The main goal of plant breeders regarding bioenergy is to develop cultivars with high biomass yield as well as cellulose and hemicellulose content, which will be stable across years and locations. However, the plants interact with their environment from germination to maturity, and since this interaction differs from one genotype to another the genotype × environment interaction complicates selection and breeding. The identification of QTL that are not environment-specific, or with only minor QE effects, should be particularly useful for marker-assisted selection. In our study, it was observed that some QTL are stable across environments and others are more or less influenced by environment. QE interaction was more frequently observed for agronomic traits than cell wall components and for QTL that showed high pleiotropic effects.

The ultimate goal in this study was to identify QTL associated with fibre-related traits as a first step towards enhancing sorghum as a bioenergy source. Pronounced phenotypic variation was observed among the genotypes, and significant QTL were detected among traits. Many QTL associated with these traits co-localized and clustered, especially on chromosomes SBI-02, SBI-03, SBI-06 and SBI-07 with the allele derived from the particular parent showing the respective high phenotype. The combination of favourable alleles for high fibre quality with high biomass in single RILs represents interesting new sorghum breeding material for the production of biomass and bioenergy. Epistastic, additive × additive interactions were found to play a major role in the genetic basis of fibre-related traits. The identified QTL regions provide an opportunity for the application of marker-assisted selection to enhance lignocellulosic biomass in sorghum. Now that we found that the cluster regions identified contain prospective candidate genes, it will be consequent to study these regions in more detail.

Acknowledgments The authors wish to thank Dr Willy Wenzel, ARC Potchefstroom, South Africa, for his invaluable advice regarding plant material and field experiments. We would like to gratefully acknowledge Dr Rod Snowdon for critically reading the manuscript, Benjamin Wittkop for his valuable input during fibre content analysis and Wubishet Bekele for his contribution in searching for candidate genes. We are also thankful for skilled technical assistance of Nelly Weiss, Swetlana Renner and Adriana Ochoa Fandiño in fibre content analysis. This work has been primarily supported by the German Academic Exchange Service (DAAD), the Agricultural Research Council (ARC) and the National Research Foundation (NRF) of South Africa.

References

Assar HA, Uptmoor R, Abdelmula AA, Salih M, Ordon F, Friedt W (2005) Genetic variation in sorghum germplasm from Sudan, ICRISAT, and USA assessed by simple sequence repeats (SSRs). Crop Sci 45:1636–1644

Barrierè Y, Gulliet C, Goffner D, Pichon M (2003) Genetic variation and breeding strategies for improved cell wall digestibility in annual forage crops. A review. Anim Res 52:193–228

Barrierè Y, Thomas J, Denoue D (2008) QTL mapping for lignin content, lignin monomeric composition, p-hydroxycinnamate content, and cell wall digestibility in the maize recombinant inbred line progeny F838 × F286. Plant Sci 175:585–595

Berg ES, Olaisen B (1994) Hybrid PCR sequencing–sequencing of PCR products using a universal primer. Biotechniques 17:896–901

Bhattramakki D, Dong J, Chhabra A, Hart G (2000) An integrated SSR and RFLP linkage map of *Sorghum bicolor* (L.) Moench. Genome 43:988–1002

Blaschke L, Legrand M, Mai C, Polle A (2004) Lignification and structural biomass production in tobacco with suppressed caffeic/5-hydroxy ferulic acid-*O*-methyl transferase activity under ambient and elevated CO_2 concentrations. Physiol Plant 121:75–83

Buranov AU, Mazza G (2008) Lignin in straw of herbaceous crops. Ind Crops Prod 28:237–259

Cao GQ, Zhu J, He CX, Gao YM, Wu P (2001) QTL analysis for epistatic effects and QTL × environment interaction effects on final height of rice (*Oryza sativa* L). Acta Genet Sin 28:135–143

Cardinal AJ, Lee M, Moore KJ (2003) Genetic mapping and analysis of quantitative trait loci affecting fibre and lignin content in maize. Theor Appl Genet 106:866–874

Carpita NC, McCann MC (2008) Maize and sorghum: genetic resources for bioenergy grasses. Trends Plant Sci 13:415–420

Chabannes M, Barakate A, Lapierre C, Marita JM, Ralph J, Pean M, Danoun S, Halpin C, Grima-Pettenati J, Boudet AM (2001) Strong decrease in lignin content without significant alteration of plant development is induced by simultaneous down-regulation of cinnamoyl CoA reductase (CCR) and cinnamyl alcohol dehydrogenase (CAD) in tobacco plants. Plant J 28:257–270

Cosgrove DJ (2005) Growth of the plant cell wall. Nat Rev Mol Cell Biol 6:850–861

Dean RE, Dahlberg JA, Hopkins MS, Mutchell SE, Kresovich S (1999) Genetic redundancy and diversity among 'Orange' accessions in the U.S. National sorghum collection assessed with simple sequence repeats (SSR) markers. Crop Sci 39:1215–1221

Dhugga KS (2007) Maize biomass yield and composition of biofuel. Crop Sci 47:2211–2227

Doyle JJ, Doyle JL (1990) Isolation of plant DNA from fresh tissue. Focus 12:13–15

Fehr WR (1987) Principles of cultivar development. MacMillan, New York

Feuillet C, Keller B (2002) Comparative genomics in the grass family: molecular characterization of grass genome. Ann Bot 89:3–10

Garcia AAF, Wang S, Melchinger AE, Zen Z-B (2008) QTL mapping and the genetic basis of heterosis in maize and rice. Genetics 180:1707–1724. doi:10.1534/genetics.107.082867

Grabber JH, Ralph J, Lapieree C, Barrierè Y (2004) Genetic and molecular basis of grass cell-wall degradability I Lignin-cell wall matrix interactions. CR Biol 327:455–465

Guo D, Chen F, Inoue K, Blount JW, Dixon RA (2001) Downregulation of caffeic acid 3-*O*-methyltransferase and caffeoyl CoA 3-*O*-methyltransferase in transgenic alfalfa: impacts on lignin structure and implications for the biosynthesis of G and S lignin. Plant Cell 13:73–88

Hasan M, Seyis F, Badani AG, Pons-Kühnemann J, Friedt W, Lühs W, Snowdon RJ (2006) Analysis of genetic diversity in the *Brassica napus* L. gene pool using SSR markers. Genet Resour Crop Evol 53:793–802

Hasan M, Friedt W, Pons-Kühnemann J, Freitag NM, Link K, Snowdon RJ (2008) Association of gene-linked SSR markers to seed glucosinolate content in oilseed rape (*Brassica napus* ssp. napus). Theor Appl Genet 116:1035–1049

Hu W-J, Harding SA, Lung J, Popko JL, Ralph J, Stokke DD, Tsai C-J, Chiang VL (1999) Repression of lignin biosynthesis promotes cellulose accumulation and growth in transgenic trees. Nat Biotechnol 17:808–812

Kim J-S, Klein PE, Klein RR, Price HJ, Mullet JE, Stelly DM (2005) Chromosome identification and nomenclature of Sorghum bicolor. Genetics 169:1169–1173

Kong L, Dong J, Hart GE (2000) Characteristics, linkage-map positions, and allelic differentiation of Sorghum bicolor (L.) Moench DNA simple sequence repeats (SSRs). Theor Appl Genet 101:438–448

Krakowsky MD, Lee M, Coors JG (2005) Quantitative trait loci for cell wall components in recombinant inbred lines of maize (*Zea mays* L.) I: stalk tissue. Theor Appl Genet 111:337–346

Lamkey KR, Lee M (1993) Quantitative genetics, molecular markers, and plant improvement. In BC Imrie, JB Hacker (ed) Focused plant improvement: towards responsible and sustainable agriculture. Proceedings of 10th Australian plant breeding conference, Gold Coast, 18–23 April 1993, Australian convention Travel Service: Canberra, pp 104–115

Léon J, Becker HC (1988) Repeatability of some statistical measures of phenotypic stability—correlation between single year results and multi years results. Plant Breed 100:137–142

Lerouxel O, Cavalier DM, Liepman AH, Keegstra K (2006) Biosynthesis of plant cell wall polysaccharides—a complex process. Curr Opin Plant Biol 6:621–630

Li Z-K, Luo LJ, Mei HW, Wang DL, Shu QY, Tabien R, Zhong DB, Ying CS, Stansel JW, Khush GS, Paterson AH (2001) Overdominant epistatic loci are the primary genetic basis of inbreeding depression and heterosis in Rice. I. Biomass and grain yield. Genetics 158:1737–1753

Li M, Yuyama N, Luo L, Hirata M, Cai H (2009) In silico mapping of 1758 new SSR markers developed from public genomic sequences for sorghum. Mol Breed 24:41–47

Lin Y, Tanaka S (2006) Ethanol fermentation from biomass resources: current state and prospects. Appl Microbiol Biotechnol 69:627–642

Lin Y-R, Schertz KF, Paterson A (1995) Comparative analysis of QTLs affecting plant height and maturity across *Poaceae*, in reference to an interspecific sorghum population. Genetics 141:391–411

Littell RC, Milliken GA, Stroup WW, Wolfinger RD, Schabenberger O (2006) SAS® for mixed models, 2nd edn. SAS Institute, Cary

Mace ES, Xia L, Jordan DR, Halloran K, Parh DK, Huttner E, Wenzl P, Kilian A (2008) DArT markers: diversity analyses and mapping in Sorghum bicolor. BMC Genomics 9:26. doi:10.1186/1471-2164-9-26

Mace ES, Rami J-F, Bouchet S, Klein PE, Klein RR, Kilian A, Wenzl P, Xia L, Halloran K, Jordan DR (2009) A consensus genetic map of sorghum that integrates multiple component maps and high-throughput diversity array technology (DarT) markers. BMC Plant Biol 9:13. doi:101186/1471-2229-9-13

Moore KJ, Hatfield RD (1994) Carbohydrates and forage quality. In: Fahey GC et al (eds) Forage quality, evaluation, and utilization. ASA, CSSA, and SSSA, Madison

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

Murray SC, Rooney WL, Mitchell SE, Sharma A, Klein PE, Mullet JE, Kresovich S (2008) Genetic improvement of sorghum as biofuel feedstock II QTL for stem and leaf structural carbohydrates. Crop Sci 48:2181–2193

Paterson AH, Bowers JE, Bruggmann R, Dubchak I, Grimwood J et al (2009) The sorghum bicolor genome and the diversification of grasses. Nature 457:551–556

Pereira MG, Lee M (1995) Identification of genomic regions affecting plant height in sorghum and maize. Theor Appl Genet 90:380–388

Prioul J-L, Pelleschi S, Sene M, Thévenot C, Causse M, de Vienne D, Leonardi A (1999) From QTLs for enzyme activity to candidate genes in maize. J Exp Bot 50:1281–1288

Reddy MSS, Chen F, Shadle G, Jackson L, Aljoe H, Dixon RA (2005) Targeted down-regulation of cytochrome p450 enzymes for forage quality improvement in alfalfa (*Medicago sativa* L.). Proc Natl Acad Sci USA 102:16573–16578

Rygulla W, Snowdon RJ, Eynck C, Koopmann B, van Tiedemann A, Lühs W, Friedt W (2007) Broadening the genetic basis of *Verticillium longisporum* resistance in *Brassica napus* by interspecific hybridisation. Phytopathology 97:1391–1396

Saballos A, Ejeta G, Sanchez E, Kang C, Vermerris W (2009) A genomewide analysis of the cinnamyl alcohol dehydrogenase family in sorghum (*Sorghum bicolor* (L.) Moench) identifies SbCAD2 as the brown midrib6 gene. Genetics 181:783–795

Sarath G, Mitchell RB, Sattler SE, Funnell D, Pedersen JF, Graybosch RA, Vogel KP (2008) Opportunities and roadblocks in utilization forages and small grains for liquid fuels. J Ind Microbiol Biotechnol 35:343–354

Schloss SJ, Mitchell SE, White GM, Kukatla R, Bowers JE, Paterson AH, Kresovich S (2002) Characterization of RFLP probe sequences for gene discovery and SSR development in *Sorghum bicolor* (L.) Moench. Theor Appl Genet 105:912–920

Schnable PS, Ware D, Fulton R, Stein JC, Wei F et al (2009) The B73 maize genome: complexity, diversity, and dynamics. Science 326:1112–1115

Shiringani AL, Frisch M, Friedt W (2010) Genetic mapping of QTLs for sugar-related traits in a RIL population of *Sorghum bicolor* L Moench. Theor Appl Genet 121:333–336

Song R, Llaca V, Messing J (2002) Mosaic organization of orthologous sequence in grass genomes. Genome Res 12:1549–1555

Srinivas G, Satish K, Mohali SM, Reddy RN, Madhusudhana R, Balakrishna D, Venkatesh BB, Howarth CJ, Seetharama N (2008) Development of genic-microsatellite markers for sorghum staygreen QTL using comparative genomic approach with rice. Theor Appl Genet 117:703–717

Srinivas G, Satish K, Madhusudhana R, Seetharama N (2009) Exploration and mapping of microsatellite markers from subtracted drought stress ESTs in *Sorghum bicolor* (L.) Moench. Theor Appl Genet 118:703–717

Sticklen MB (2008) Plant genetic engineering for biofuel production: towards affordable cellulosic ethanol. Nat Rev Genet 9:433–443

Swigonova Z, Lai J, Ma J, Ramakrishna W, Llaca V, Benetzen JL, Messing J (2004) Close split of sorghum and maize genome progenitors. Genome Res 14:1916–1923

Taylor NG (2008) Cellulose biosynthesis and deposition in higher plants. New Phytol 178:239–252

Thomas J, Guillaumie S, Verdu C, Denoue D, Pichon M, Barrière Y (2010) Cell wall phenylpropanoid-related gene expression in early maize recombinant inbred lines differing in parental alleles at a major lignin QTL position. Mol Breed 25:105–124

Uptmoor R, Wenzel WG, Friedt W, Donalson G, Ayisi K, Ordon F (2003) Comparative analysis on the genetic relatedness of *Sorghum bicolor* accessions from southern Africa by RAPDs, AFLPs and SSRs. Theor Appl Genet 106:1316–1325

Utz HF, Melchinger AE (1994) Comparison of different approaches to interval mapping of quantitative trait loci. In van Ooijen JW, Jansen J (eds) Biometrics in plant breeding: applications of molecular markers. Wageningen, pp 195–204

Van Soest PJ, Robertson JB, Lewis BA (1991) Methods for dietary fibre, neutral detergent fibre, and nonstarch polysaccharides in relation to animal nutrition. J Dairy Sci 74:583–3597

Veldboom LR, Lee M (1996) Genetic mapping of quantitative trait loci in maize in stress and nonstress environments: II plant height and flowering. Crop Sci 36:1320–1327

Vermerris W, Saballos A, Ejeta G, Mosier NS, Ladisch MR, Carpita NC (2007) Molecular breeding to enhance ethanol production from corn and sorghum stover. Crop Sci 47:S143–S153

Vos P, Hogers R, Bleeker M, Reijans M, van de Lee T, Hornes M, Frijters A, Pot J, Peleman J, Kuiper M (1995) AFLP: a new technique for DNA fingerprinting. Nucleic Acids Res 21:4407–4414

Wilson LM, Whitt SR, Ibáñez AM, Rocherford TR, GoodMan MM, Buckler IV ES (2004) Dissection of maize kernel composition and starch production by candidate gene association. Plant Cell 16:2719–2733

Xu XY, Bai GH, Carver BF, Shaner GE (2005) Mapping of QTLs prolonging the latent period of *Puccinia triticina* infection in wheat. Theor Appl Genet 110:244–251