

# Inheritance of inflorescence architecture in sorghum

P. J. Brown · P. E. Klein · E. Bortiri · C. B. Acharya ·  
W. L. Rooney · S. Kresovich

Received: 6 February 2006 / Accepted: 21 June 2006 / Published online: 18 July 2006  
© Springer-Verlag 2006

**Abstract** The grass inflorescence is the primary food source for humanity, and has been repeatedly shaped by human selection during the domestication of different cereal crops. Of all major cultivated cereals, sorghum [*Sorghum bicolor* (L.) Moench] shows the most striking variation in inflorescence architecture traits such as branch number and branch length, but the genetic basis of this variation is little understood. To

study the inheritance of inflorescence architecture in sorghum, 119 recombinant inbred lines from an elite by exotic cross were grown in three environments and measured for 15 traits, including primary, secondary, and tertiary inflorescence branching. Eight characterized genes that are known to control inflorescence architecture in maize (*Zea mays* L.) and other grasses were mapped in sorghum. Two of these candidate genes, *Dw3* and the sorghum ortholog of *ramosa2*, colocalized precisely with QTL of large effect for relevant traits. These results demonstrate the feasibility of using genomic and mutant resources from maize and rice (*Oryza sativa* L.) to investigate the inheritance of complex traits in related cereals.

Communicated by R. Bernardo

**Electronic Supplementary Material** Supplementary material is available to authorised users in the online version of this article at <http://dx.doi.org/10.1007/s00122-006-0352-9>.

P. J. Brown (✉) · C. B. Acharya · S. Kresovich  
Institute for Genomic Diversity, Cornell University,  
158 Biotechnology Building, Ithaca, NY 14853, USA  
e-mail: pjb34@cornell.edu

P. J. Brown · S. Kresovich  
Department of Plant Biology, Cornell University, Ithaca,  
NY 14853, USA

P. E. Klein  
Institute for Plant Genomics and Biotechnology,  
Texas A&M University, College Station, TX 77843, USA

P. E. Klein  
Department of Horticulture, Texas A&M University,  
College Station, TX 77843, USA

E. Bortiri  
Plant Gene Expression Center, Department of Plant and  
Microbial Biology, University of California-Berkeley,  
Albany, CA 94710, USA

W. L. Rooney  
Department of Soil and Crop Sciences, Texas A&M  
University, College Station, TX 77843, USA

**Keywords** Inflorescence architecture · Sorghum ·  
Quantitative trait locus · Candidate gene

## Introduction

With 10,000 species, the grass family (*Poaceae*) is one of the most species-rich groups of Angiosperms. *Poaceae* includes major cereal crops such as rice, wheat (*Triticum aestivum* L.), and maize, which, together with other cereals, provide over 70% of the food consumed by humans. Grasses bear grains in spikelets, a key reproductive innovation consisting of a short shoot with two bracts called glumes subtending a variable number of reduced flowers. Spikelets are borne on the main inflorescence rachis or on branches formed off the main rachis, in a pattern that is highly variable among grass species. Thus, the genetic control of the number and growth of branches, or inflorescence architecture, is very important in this family because it directly affects grain yield.

Maize and rice are two of the leading model systems for grass research, and inflorescence mutants from these species have been used to characterize a number of genes involved in the control of grass inflorescence architecture (Bommert et al. 2005). Some of these genes also appear to affect quantitative variation in inflorescence traits (Upadyayula et al. 2005) and their exploitation for cereal improvement is just beginning (Ashikari et al. 2005). Robertson (1989) was an early proponent of the idea that the genes underlying quantitative and mutant/qualitative variation in a species might be the same. By extension, mutants from one species might also be used to identify candidate genes in related species. Today, this logic underlies an approach which has met with considerable success in identifying genes involved in agronomic traits in the grasses, including maize flowering time (Thornsberry et al. 2001), wheat vernalization (Trevaskis et al. 2003), sorghum photoperiod sensitivity (Childs et al. 1997), sorghum lignin composition (Bout and Vermerris 2003), and inflorescence architecture in both maize and rice (Gallavotti et al. 2004; Komatsu et al. 2003a). There are many reasons why a given candidate gene might not affect a quantitative trait. The expression of functional variation at a candidate locus could be masked by alleles at epistatic loci (Lauter and Doebley 2002), or the functional variation itself might have been lost due to selection or drift (Clark et al. 2004). However, the co-localization of candidate genes and quantitative trait loci (QTL) for relevant traits provides a means of prioritizing specific genomic regions for fine-mapping and/or association mapping (Salvi and Tuberosa 2005).

Sorghum is a tropical  $C_4$  grass with a relatively compact genome of 736 MB which will soon be sequenced (<http://www.jgi.doe.gov>), and is the world's fifth most important cereal crop, with extensive subsistence use in some of the world's poorest countries in sub-Saharan Africa. Although sorghum itself has few characterized genes or mutant lines, it is only 11.9 million years diverged from maize (Swigonova et al. 2004), which boasts an exceptional collection of inflorescence mutants. Sorghum inflorescences, called panicles, are more extensively branched than those of maize or rice, and there is tremendous variation within sorghum both in the degree of inflorescence branching and in the degree of rachis and branch elongation. Variation in panicle morphology is a primary factor in the subdivision of domesticated sorghum into five major races: bicolor, caudatum, kafir, durra, and guinea (Harlan and De Wet 1972).

The objective of this study was to investigate the inheritance of sorghum inflorescence architecture by

determining trait heritabilities, correlations between traits, and identifying QTL of large effect. We employed an integrative approach similar to that used previously with foxtail millet (Doust et al. 2005) to assess the possible contribution of known grass inflorescence genes to genetic variation in sorghum inflorescence morphology. Eight candidate genes from maize and other grasses were mapped, along with inflorescence QTL, in a population of sorghum recombinant inbred lines (RILs) that shows extensive variation in inflorescence architecture. The co-localization of candidates and QTL in this study provides a starting point for isolating genes with quantitative effects on inflorescence architecture in the grasses.

## Materials and methods

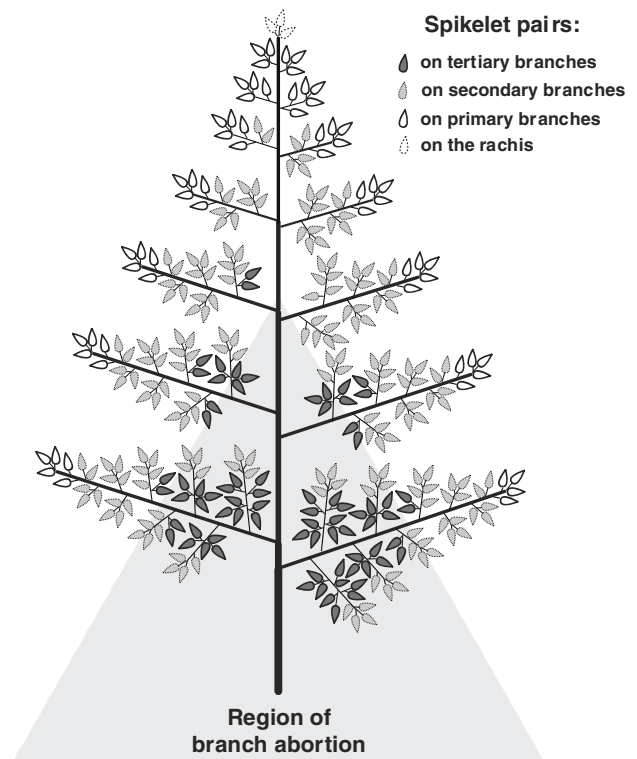
### Plant materials and growing conditions

A series of RILs were developed from the cross (BTx623 × IS3620C) during the creation of a high-density molecular map of sorghum (Menz et al. 2002). In this study, we grew and phenotyped 119 of these lines at the  $F_9$ – $F_{10}$  stage. The female RIL parent, BTx623, is a subtropically-adapted parental line with a stout rachis, numerous short primary branches, extensive secondary and tertiary branching, and a copious number of large seeds. The male RIL parent, IS3620C, is an exotic accession with a thin rachis and long primary branches that have reduced secondary and tertiary branching, and a reduced number of small seeds. The BTx623 panicle has a compact overall appearance, whereas the IS3620C panicle is often described as “open” or “loose”. Although most elite sorghum varieties have relatively compact panicles and are high yielding, the effect of panicle architecture on yield components has not been adequately addressed. IS3620C is a “converted” line, meaning that the original line from which it was derived was crossed to a temperately-adapted line and then backcrossed to the original exotic while selecting for short stature and photoperiod-insensitivity (Stephens et al. 1967). RILs and parent lines were planted in three locations (Weslaco, College Station, and Halfway, TX, USA) in 2004. A randomized complete block design was used, with two replicates per location and 2 rows of each parent per replicate, for a total of 246 rows per location. The Weslaco location (26°N, 98 °W) was planted in February, harvested in June, received 36.1 cm of rainfall and experienced an average daily temperature of 23.6°C. The College Station location (30.5 °N, 96 °W) was planted in April, harvested in August, and received

72.6 cm of rainfall with an average daily temperature of 25.6°C. The Halfway location (34 °N, 101.5 °W) was planted in May, harvested in September, and received 39.1 cm of rainfall with an average daily temperature of 22.9°C (<http://www.ncdc.noaa.gov>).

### Phenotyping of panicle traits

Mature panicles were harvested and dried before measurement of panicle traits. Rachis diameter was measured with calipers at the bottom-most branch of the panicle. Primary branches were removed individually and counted, and rachis length was measured as the distance from the bottom-most branch to the end of the primary axis. Next, two branches were randomly selected from the “long-branch zone” in the bottom third of each panicle. On each of two long branches, we measured branch length, secondary branch number, and tertiary branch number. Tertiary branch number was measured as the number of secondary branches containing tertiary branches. The trait percent tertiary branching was calculated as (number of secondary branches with tertiary branching/total number of secondary branches)  $\times$  100. A “branch” is defined as an axis bearing two or more spikelet pairs; an axis bearing a single spikelet pair is considered a pedicel (Ikeda et al. 2004). Thus, for the bottom-most branch in Fig. 1, secondary branch number is 10, tertiary branch number is 6, and percent of tertiary branching is 60. After undergoing branch abortion in the shaded area, the same panicle branch would yield a secondary branch number of 5, a tertiary branch number of 1, and a percent of tertiary branching of 20. All panicle traits described were measured in three to five panicles per row, and row means were used for quantitative analysis. For traits measured in two branches per panicle, means were calculated for each panicle prior to calculating a row mean. Traits evaluated on a per row basis included plant height, peduncle length (the distance from the flag leaf to the lowest inflorescence branch), flowering time, panicles per row, seed set, 100-seed weight, seeds per panicle, and total seed weight. To assess 100-seed weight, seeds per panicle, and total seed weight, 5–8 panicles were threshed from each row. Seeds per panicle was calculated as (seed weight per panicle/100-seed weight)  $\times$  100. Total seed weight was calculated as (panicles per row)  $\times$  (seed weight per panicle). All traits were measured in three environments, with the exception of flowering time, which was not measured in Halfway, and the seed traits (100-seed weight, seeds per panicle, seed set, and total seed weight) which were not measured in College Station because of damage from sorghum midge (*Contarinia sorghicola*).



**Fig. 1** Inflorescence branching in sorghum. Spikelet pairs may be borne on primary, secondary, or tertiary branches, or directly on the rachis, as indicated. The shaded triangle indicates the region of branch abortion observed in IS3620C

### Scanning electron microscopy

Inflorescence meristems were dissected from 4 to 8 week old, greenhouse-grown plants, washed overnight in FAA (45% EtOH, 5% acetic acid, 10% formalin), and transferred to 95% EtOH. Fixed meristems were critical-point dried, mounted, sputter-coated, and photographed using a Hitachi S4700 field emission SEM at the Plant Gene Expression Center in Albany, CA, USA

### Quantitative analysis

Trait variances were partitioned using the random effects ANOVA model  $y = \mu + E + G + G \times E + \text{error}$ , where E represents environment, G represents genotype, and  $G \times E$  represents the genotype by environment interaction. The error term includes the variance between row means for the two replicates of each genotype at each location. Heritability estimates for each trait were calculated as  $(\sigma^2_G)/(\sigma^2_G + \sigma^2_{G \times E}/L + \sigma^2_{\text{error}}/LR)$ , where L and R represent the number of locations and the number of replicates per location, respectively. ANCOVA was used to partition the phenotypic covariance between traits into genetic and

environmental components. Genetic correlations were calculated between traits as  $(\text{cov}_{T_1T_2})/\sqrt{(\text{var}_{T_1}\text{var}_{T_2})}$ , where  $T_1$  and  $T_2$  represent the traits being compared, and the numerator and denominator represent the genetic components from the ANCOVA and ANOVA analyzes, respectively. All statistical analyzes were performed using R 2.0.1 (<http://www.r-project.org>).

### QTL analysis

The RILs used in this study have been genotyped at the  $F_6$  stage for  $\sim 3,000$  AFLP, RFLP, and SSR markers (Menz et al. 2002), of which we used 396 framework markers to perform QTL analysis. The complete (BTx623  $\times$  IS3620C) linkage map, including the framework markers used in this study, is available online at <http://www.sorgblast2.tamu.edu>. Composite interval mapping (CIM) was performed using WinQTLCart version 2.5 (Wang et al. 2005). Genotypic coefficients from the ANOVA models were used to perform CIM analysis on each trait across all three environments. Seed set was included as a covariate in the models for both seeds per panicle and total seed weight. Cofactors for all analyzes were selected by stepwise regression using the forward selection, backward elimination method with a cutoff of 0.05 for cofactor selection and elimination. A window size of 10 cM and a walk speed of 0.5 cM were used for all analyzes. Significance was established by performing 1,000 permutations and setting a highly stringent cutoff of  $\alpha = 0.001$ .

### Candidate gene analysis

DNA was extracted from leaf tissue pooled from 8 to 10 plants per RIL using a modification of the CTAB protocol of Doyle and Doyle (1987). Genomic contigs for inflorescence candidate genes from *Zea*, *Oryza*, and *Lolium* were assembled, in most cases from publicly-available sequences, and BLAST analysis was performed against the SAMI database (Sorghum Assembled genoMic Islands, <http://www.magi.plantgenomics.iastate.edu>), which is an assembly of methylation-filtered GSS reads from Orion Genomics (Bedell et al. 2005). Primers were designed manually in regions of the resulting gene contigs that showed high similarity across all species and/or all copies of the gene within a given species. Multiple fragments from each gene contig were treated with exonuclease I and shrimp alkaline phosphatase and sequenced on an ABI3730 to search for polymorphism between the mapping parents. Wherever possible, RIL genotypes at candidate loci were scored using indel polymorphisms assayed on

either an ABI3730 or on agarose gels. If only SNPs were found between the mapping parents, RILs were genotyped using the derived cleaved amplified polymorphic sequences (dCAPS) technique (Michaels and Amasino 1998). Genotyped candidates were placed on the (BTx623  $\times$  IS3620C) map with MapPop (Brown and Vision 2000).

## Results

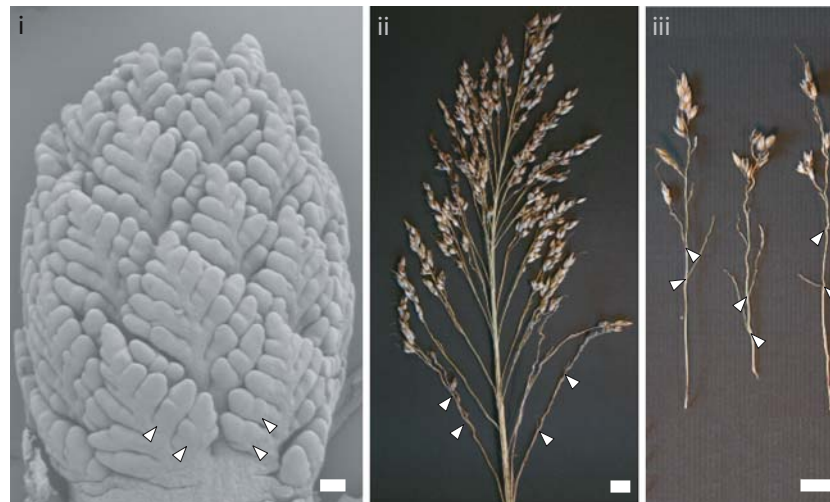
### Anatomy of the sorghum inflorescence

Sorghum inflorescences show primary, secondary, and tertiary branching, with primary branches becoming progressively shorter and less branched towards the top of the panicle (Fig. 1). In IS3620C, the lower primary branches often bear spikelet pairs only at their tips due to the abortion of basal secondary and tertiary branches, a phenomenon known to sorghum breeders as “blasting”. The effect of branch abortion on panicle architecture is clearly evident when comparing scanning electron microscopy photographs of immature IS3620C inflorescences with mature IS3620C panicles (Fig. 2). Whereas the scanning electron microscopy photos reveal extensive secondary and tertiary branching in the lower primary branches, mature panicles have empty nodes or short aborted branches without spikelets on the lower portion of the corresponding primary branches. At flowering time, many of these branches are still visible as white, stunted tissue that falls off as the panicle dries and matures.

### Quantitative genetic analysis of panicle traits

Genotype, environment, and genotype by environment interactions contributed significantly to the variance of all 15 traits. Means, standard deviations, and ranges of the traits measured are provided as supplemental data (Table S1). BTx623 panicles consistently display more primary, secondary, and tertiary branching than IS3620C panicles. Branch length is shorter in BTx623, although rachis length is similar between the two parents. BTx623 has heavier seeds, increased seed number, and higher total seed weight, while IS3620C bears many more panicles per row. BTx623 and IS3620C have similar values for plant height and flowering time across all three environments. However, transgressive segregation is observed for all traits in the RILs and is especially pronounced for height and flowering time.

With regard to traits involved in elongation, heritabilities are much higher for plant height, rachis length,



**Fig. 2** Branch abortion in IS3620C. *White arrowheads* indicate analogous secondary branches in each frame. **i** SEM image of an immature IS3620C panicle, with prolific secondary and tertiary branching on the lower primary branches; *scale bar* 0.1 mm. **ii** Mature IS3620C panicle, with a “loose” overall appearance pro-

duced in part by sparse spikelet production on the lower primary branches; *scale bar* 1 cm. **iii** Primary branches from the lower portion of a mature IS3620C panicle. The lower secondary branches on each primary branch have aborted without producing spikelets; *scale bar* 1 cm

and branch length than for peduncle length (Table 1). With regard to inflorescence branching, heritability is much higher for primary branch number than for secondary and tertiary branching traits. Flowering time and seed-related traits have much higher  $G \times E$  variance components than the other traits.

A number of trait pairs show significant genetic correlations (Table 2), indicating that their inheritance is linked either functionally (pleiotropy) or physically (linkage disequilibrium). Rachis length and branch length are highly correlated with each other, but show

non-significant or negative correlations with plant height and peduncle length. Similarly, the genetic correlation between primary and secondary branch number is estimated at only 0.34, and the correlation between primary and tertiary branch number is not significant. Seeds per panicle is strongly and positively correlated with all three orders of branch number. Much of the variance for total seed weight in this population is due to differences in seed set [ $r_{a(\text{Total seed weight:Seed set})} = 0.62$ ]. Rachis diameter shows significant positive correlations with all three orders of branch number. When rachis diameter was

**Table 1** Quantitative genetic statistics for sorghum inflorescence traits

Trait type	Trait name	Variance components (% of total $\sigma^2_p$ )				Heritability ( $h^2$ )
		$\sigma^2_G$	$\sigma^2_E$	$\sigma^2_{G \times E}$	$\sigma^2_{\text{Error}}$	
Branching	Primary branch number	71	9	9	11	0.94
	Secondary branch number	33	2	13	52	0.72
	Tertiary branch number	44	2	11	43	0.80
	% tertiary branching	45	8	10	37	0.83
Elongation	Plant height	78	6	8	8	0.94
	Peduncle length	46	11	9	34	0.84
	Rachis length	70	6	8	16	0.93
	Branch length	62	12	4	22	0.93
Seed	100-seed weight	56	9	22	13	0.85
	Seeds per panicle	49	8	22	21	0.82
	Seed set	20	30	31	19	0.60
	Total seed weight	14	44	25	17	0.56
Miscellaneous	Rachis diameter	55	11	4	30	0.90
	Flowering time	65	5	19	11	0.89
	Panicles per row	20	41	3	36	0.74

**Table 2** Genetic correlations ( $r_A$ ) between traits

	Primary branch number	Secondary branch number	Tertiary branch number	% tertiary branching	Plant height	Peduncle length	Rachis length	Branch length	100-seed weight	Seeds per panicle	Seed set	Total seed weight	Rachis diameter	Flowering time
Primary branch number														
Secondary branch number	<b>.34</b>													
Tertiary branch number	-.13	<b>.65</b>												
% tertiary branching	<b>-.33</b>	<b>.27</b>	<b>.90</b>											
Plant height	.08	.09	-.07	-.15										
Peduncle length	<b>-.26</b>	<b>-.22</b>	<b>-.27</b>	<b>-.23</b>	<b>.22</b>									
Rachis length	.06	.06	.15	.17	-.16	.01								
Branch length	<b>-.23</b>	.06	<b>.33</b>	<b>.39</b>	<b>-.23</b>	<b>-.19</b>	<b>.71</b>							
100-seed weight	.14	.06	.09	.10	.11	.01	.07	-.01						
Seeds per panicle	<b>.44</b>	<b>.41</b>	<b>.36</b>	<b>.24</b>	<b>.28</b>	-.06	<b>.22</b>	.09	<b>.42</b>					
Seed set	-.04	.00	.08	.10	<b>.32</b>	.10	.03	-.10	<b>.35</b>	<b>.57</b>				
Total seed weight	<b>.32</b>	<b>.20</b>	.09	.01	<b>.46</b>	.11	.02	-.16	<b>.52</b>	<b>.82</b>	<b>.62</b>			
Rachis diameter	<b>.58</b>	<b>.50</b>	<b>.24</b>	.04	<b>-.19</b>	-.15	<b>.23</b>	.06	<b>.30</b>	<b>.55</b>	.05	<b>.26</b>		
Flowering time	<b>.28</b>	<b>.33</b>	<b>.21</b>	.09	-.07	<b>-.28</b>	-.05	.16	<b>-.29</b>	<b>-.35</b>	<b>-.59</b>	<b>-.50</b>	<b>.20</b>	
Panicles per row	-.14	<b>-.35</b>	<b>-.49</b>	<b>-.44</b>	<b>.33</b>	<b>.29</b>	<b>-.24</b>	<b>-.37</b>	<b>-.33</b>	<b>-.29</b>	.01	.16	<b>-.49</b>	<b>-.27</b>

Significant correlations are shown in bold

included as a covariate in the linear models for primary, secondary, and tertiary branch number, it accounted for 25, 15, and 6% of the phenotypic variance, respectively.

#### Detection of QTL

Nineteen QTL, each accounting for 8.5–38.4% of the phenotypic variance for a given trait, were detected at  $P < 0.001$  (Fig. 3; Table S2). Of the nineteen QTL, twelve are for traits that show clear differences between the two parents, such as primary branch number, branch length, and 100-seed weight. Of these 12, nine show effects in the “expected” direction, with the BTx623 allele contributing to increased branching or seed weight, for example.

#### Genomic organization of QTL

Several regions of the genome display clustering of QTL with large effects (Fig. 3). A cluster on the long arm of chromosome 7 includes QTL for plant height, rachis length, and branch length with peaks within a few cM of each other. The closest marker to these QTL peaks is the *Dw3* gene.

A cluster on chromosome 6 includes QTL for plant height, rachis diameter, and primary branch number. The latter two QTL have large effects in the opposite direction of what would be expected based on parental phenotypes: the BTx623 allele is associated with

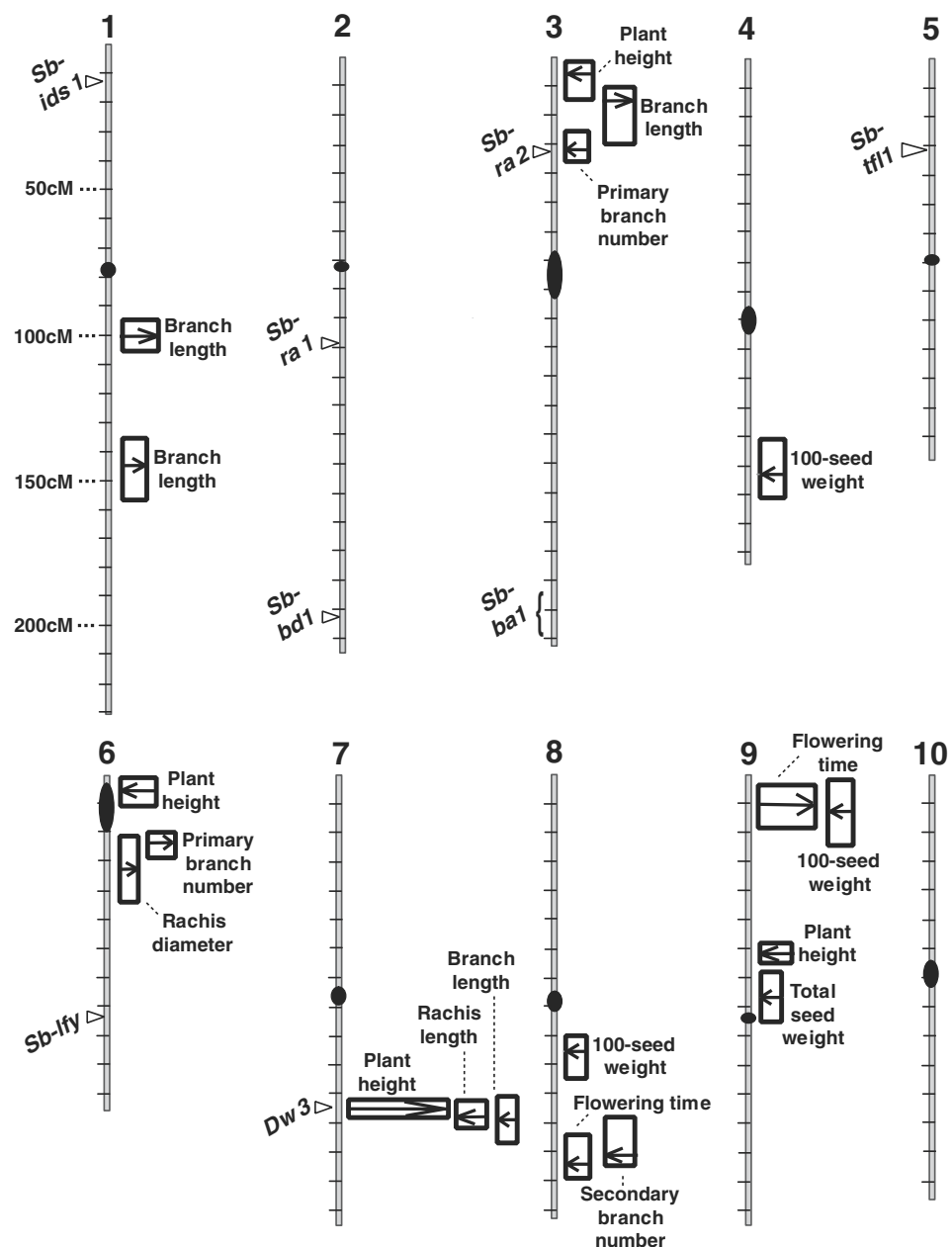
reduced primary branching and reduced rachis diameter, even though BTx623 has more primary branches and its rachis is almost twice as thick as that of IS3620C (Fig. 3).

A cluster on chromosome 3 includes QTL for plant height, branch length, and primary branch number. The closest marker to the primary branch number QTL is the *Sb-ra2* gene.

#### Candidate gene mapping

Map locations for all eight candidate genes (Table 3; Fig. 3) agree well with current models of the sorghum genome’s colinearity with maize and rice (Devos 2005). Candidate gene map locations could not have been predicted unambiguously using colinearity alone, due to the complex history of duplications and deletions between grass genomes (Gaut 2002). The primer pairs designed amplified single bands for every candidate gene with the exception of *Sb-ids1*. The original primers used to amplify *Sb-ids1* amplified a single band that yielded clean sequence data, but subsequent primers designed for mapping detected a second, *Sb-ids1*-like locus on the short arm of chromosome 10 (data not shown). One of the candidate genes, *Sb-ba1*, showed no polymorphism between the mapping parents over a 2.1 kb contig which included the entire coding region and approximately 1.4 kb of upstream sequence. Since this was unusual given the genetic distance between these lines, we sequenced the corresponding fragments

**Fig. 3** Candidate genes and QTL for sorghum inflorescence traits. The ten sorghum chromosomes are shown in order and calibrated at 10 cM intervals. *Black ellipses* represent centromeres and pericentromeric heterochromatin as reported by kim et al. (2005). Mapped candidate genes and QTL are shown to the left and right, respectively, of each chromosome. The *Sb-bal* gene location is shown with a *bracket* to indicate that the precise location of this gene on this map is not known. Each QTL is indicated by a *box* and *arrowhead* that represent the 2-LOD interval and LOD peak, respectively. The direction of the *arrowhead* indicates the direction of the QTL effect (*left-pointing arrow*: BTx623 allele increases trait value; *right-pointing arrow*: IS3620C allele increases trait value), and the width of the QTL box is proportional to the LOD score



in the exotic progenitor of IS3620C (NSL51039), which is tall and photoperiod-sensitive, and in the “conversion parent” (BTx406), the elite line which was used to donate dwarfing and photoperiod-insensitive alleles. The sequencing results showed that IS3620C had been “converted” at the *Sb-bal* locus (data not shown). Therefore, *Sb-bal* was mapped using a different sorghum population (BTx623 × *Sorghum propinquum*, in collaboration with J. Bowers and A. Paterson), and its position on the present map is inferred based on the alignment of these two sorghum genetic maps (Feltus et al. 2006).

#### Co-localization of candidate genes and QTL

Two of the eight mapped candidate genes co-localized with QTL. Given that the 2-LOD intervals of the detected QTL cover 235 cM, or 14% of the 1,677 cM genome (Table S2), the proportion of candidates co-localizing with QTL is not significantly greater than what would be expected due to chance alone using a chi-square test ( $P = 0.37$ ). Even if 1-LOD QTL intervals are considered, the detected QTL collectively cover 127 cM or 7.5% of the genome, and the proportion of candidates co-localizing with QTL is still not

**Table 3** Candidate genes for grass inflorescence architecture

Sorghum candidate		Putative orthologs							
Name	Chr. (cM)	Amino acids	Spp.	Gene	Reference	Amino acids	% ID w/sorghum protein	Homology	Function
<i>Sb-ba1</i>	3 (174–199) <sup>a</sup>	224	<i>Zea</i>	<i>ba1</i>	gallavotti et al. (2004)	211	87	bHLH TF	Primary branch initiation
<i>Sb-ra1</i>	2 (99.1)	113/184 <sup>b</sup>	<i>Oryza</i>	<i>LAX</i>	komatsu et al. (2003b)	215	62		Branch determinacy
<i>Sb-ra2</i>	3 (30.5)	268	<i>Zea</i>	<i>ra1</i>	vollbrecht et al. (2005)	174	71/71	EPF-class C <sub>2</sub> H <sub>2</sub> zinc finger LOB domain	Branch determinacy
			<i>Zea</i>	<i>ra2</i>	bortiri et al. (2006)	261	87		
			<i>Oryza</i>	<i>BAB92516.1</i>	( <a href="http://www.gramene.org">http://www.gramene.org</a> )	251	77		
<i>Sb-bd1</i>	2 (196.0)	326	<i>Zea</i>	<i>bd1</i>	chuck et al. (2002)	313	83	1 AP2-domain ERF	Spikelet determinacy
			<i>Oryza</i>	<i>fzp</i>	komatsu et al. (2003a)	318	73		
<i>Sb-ids1</i>	1 (13.2)	432	<i>Zea</i>	<i>ids1</i>	chuck et al. (1998)	433	94	2 AP2-domain ERF	Spikelet determinacy
			<i>Oryza</i>	<i>BAC21448.1<sup>c</sup></i>	<i>Gramene</i>	436	64		
			<i>Oryza</i>	<i>AC104433</i>	<i>Gramene</i>	434	69		
<i>Sb-lfy</i>	6 (82.7)	? <sup>d</sup>	<i>Zea</i>	<i>zfl-2</i>	bombliet et al. (2003)	391/394	?	<i>FLORI-CAULA/LEAFY</i>	Reproductive transition/info. branching
			<i>Oryza</i>	<i>rfl</i>	kyozuka et al. (1998)	389	?		
			<i>Lolium</i>	<i>LtLFY</i>	coical et al. (2001)	400	?		
<i>Sb-ffl1</i>	5 (27.5)	173	<i>Zea</i>	<i>Zmcen1-3</i>	this paper	177/173 <sup>?</sup>	92/92 <sup>?</sup>	<i>TFL1/CEN (PEBP)</i>	Reproductive transition/info. branching
			<i>Oryza</i>	<i>Oscen2</i>	zhang et al. (2005)	173	91		
			<i>Lolium</i>	<i>LpTFL1</i>	jensen et al. (2001)	173	89		
<i>Dw3</i>	7 (116.5)	1,402	<i>Zea</i>	<i>br2</i>	multani et al. (2003)	1,394	91	P-glycoprotein	Auxin transport
			<i>Oryza</i>	<i>BAD12940.1</i>	<i>Gramene</i>	1,344	82		

<sup>a</sup> *Sb-ba1* was mapped in a population of 48 (BTx623 × *Sorghum propinquum*) F<sub>2</sub>s courtesy of J. Bowers and A. Paterson

<sup>b</sup> The *Sb-ra1* locus is tandemly duplicated (Vollbrecht et al. 2005)

<sup>c</sup> When a single orthologous copy of a candidate gene could not be unambiguously identified all putative orthologs from that grass species are listed

<sup>d</sup> Question marks are included where the full amino acid sequence of a protein is not known



significant ( $P = 0.06$ ). Both co-localizing candidates mapped within 2 cM of the respective QTL peaks, whereas the 2-LOD intervals averaged 15 cM.

## Discussion

### Genetic architecture of sorghum panicle traits

Variation in sorghum inflorescence architecture results from differences in the branching, elongation, and branch abortion of the developing inflorescence. Although the genic basis of these differences is unknown, the high heritabilities of easily-measured inflorescence traits like branch length and primary branch number suggest that inflorescence architecture may be quite amenable to quantitative analysis. Since the sorghum panicle is more highly branched than the inflorescences of rice or maize, and it shows great variability among inbred lines, it is an excellent model with which to study quantitative variation in higher-order inflorescence branching in grasses.

### *Inflorescence branching*

Primary branch number had significantly higher heritability than secondary and tertiary branching traits in this study. There are several probable reasons for this. First, only a subset of the primary branches were used to estimate secondary and tertiary branching. Second, secondary and tertiary branching are affected not only by meristem initiation but also by the abortion of basal branches, which takes place later in development and is presumably controlled by a distinct set of genes. Subsequent studies could potentially disentangle these phenomena by phenotyping secondary branch initiation separately from secondary branch abortion.

Analysis of both genetic correlations and QTL supports the idea that primary, secondary, and tertiary inflorescence branching in sorghum are largely under separate genetic control. There is no overlap between QTL detected for primary and secondary branch number, in keeping with the many reports of specific genes that control specific stages of inflorescence branching (Bommert et al. 2005).

Three QTL for inflorescence branching were detected, two for primary branch number and one for secondary branch number. The primary branch number QTL on chromosome 6 maps within a few cM of a QTL for rachis diameter. When rachis diameter was included as a covariate in the linear model for primary branch number, this QTL was no longer significant at  $P < 0.001$ . Therefore, this primary branch number QTL

may be detected due to its proximity to the rachis diameter QTL. Alternatively, these two QTL may represent a single locus that exerts a concomitant increase in primary branch number and rachis diameter.

This genomic region is also known to harbor *Mal*, a major gene for photoperiod-sensitivity, as well as the major height gene *Dw2* (Lin et al. 1995). Since most exotic grain sorghum has been selected for photoperiod-insensitivity through crossing to an insensitive line, this region of the genome is usually converted to an elite state in exotic germplasm, and is absent from some sorghum maps (e.g. see Sanchez et al. 2002). Some of the QTL in this genomic region may be conditioned by elite alleles. In particular, the plant height QTL is likely *Dw2*, which is dominant in BTx623 but recessive in the conversion parent BTx406. The QTL for rachis diameter and primary branch number could be conditioned by either elite or exotic alleles in this population, but their proximity to *Mal* and *Dw2* suggests that at least some converted lines of sorghum may have been unintentionally converted at these inflorescence QTL.

### *Inflorescence elongation*

Inflorescence elongation appears to be under separate genetic control from plant height, as has been noted previously (Doggett 1988). In this population, elite alleles of the QTL near *Dw3* cause moderate elongation of the rachis and primary branches and a drastic compaction of the lower vegetative internodes, consistent with the reported effects of *Dw3* (Multani et al. 2003). Aside from these putative *Dw3* QTL, no close co-localization of QTL for plant height, rachis length, and branch length was observed. Thus, not only is the elongation of the inflorescence distinct from that of the vegetative internodes, but rachis elongation and rachis branch elongation also appear to be distinct processes. Four QTL were detected for branch length, more than for any other trait except plant height. Branch length is of relevance to grain mold resistance and bird predation, and is a promising target for further quantitative genetic study in sorghum.

### *Inflorescence branch abortion*

The overall appearance of panicle “compactness”, which is used as a racial indicator in sorghum, is influenced not only by inflorescence branching and elongation, but also by the abortion of higher-order branches prior to spikelet maturity. In this study, branch abortion was assessed by measuring secondary and tertiary branching on two basal branches per panicle: since a “branch” was defined as an axis bearing multiple spikelet

pairs, aborted branches were not counted. The expectation for a QTL resulting from differences in branch abortion is that it should affect both secondary and tertiary branching. However, only a single QTL was detected for secondary branch number, and no QTL were detected for tertiary branching. Future studies that measure branch abortion directly may have more success in uncovering the genetic basis of this complex trait.

#### *Relationship between inflorescence and seed traits*

Grass inflorescence architecture is of primary interest to breeders because it contributes directly to yield, yield stability, and grain quality. The strong positive correlation between the number of seeds per panicle and all three orders of inflorescence branching suggests that this yield component can be selected for through increased inflorescence branching. None of the three QTL detected for 100-seed weight were linked to inflorescence architecture QTL, suggesting that selection for both increased seed weight and increased inflorescence branching is potentially feasible. However, these results should be interpreted with caution, and the relationship between seed weight, seed number, and inflorescence architecture should be reassessed in a sorghum population that does not show such wide variation in seed set.

#### *Mapping of grass inflorescence genes in sorghum*

Eight candidate genes for grass inflorescence architecture were mapped in sorghum. Based on both genomic colinearity and sequence similarity, it is likely that the sorghum genes are orthologous to their counterparts from other grasses. While Southern blots are still needed for absolute verification of copy number, results from BLAST analysis of the SAMI database also suggest that most of these genes are present in a single copy in the sorghum genome (data not shown). Finally, given the high degree of similarity between the predicted sorghum proteins and their putative orthologs in maize, rice, and ryegrass, it is likely that the proteins perform similar functions. If functional variation exists at these candidate gene loci in sorghum, then they have the potential to contribute to quantitative variation in inflorescence architecture.

During the mapping of *Sb-ids1*, we detected an additional *Sb-ids1*-like locus on the short arm of chromosome 10 (data not shown). Intriguingly, this genomic region in sorghum has several markers collinear with maize 9S near the map location of the *rgo1* mutant (Kaplin and Freeling 2003). Since *ids1* and *rgo1* show non-allelic non-complementation in maize, it is likely that they encode similar proteins.

*Sb-bal* could not be mapped in our population because of the introgression of an elite allele indistinguishable from that of BTx623. For this reason, we had little ability to assess the possible contribution of *Sb-bal* to phenotypic variation in sorghum inflorescence architecture. The conversion of *Sb-bal* in IS3620C is likely due to linkage with a height gene rather than direct selection on *Sb-bal* itself, since the *Sb-bal* region co-localizes with previously reported height QTL and conversion events (Lin et al. 1995), and functional characterization of *bal* in maize does not suggest that it has any effect on plant height or flowering time (Gallavotti et al. 2004).

#### *Correspondence between sorghum QTL and candidate genes*

Two of eight mapped candidate genes were found to co-localize with QTL, which is not significantly different from the expectation due to chance alone. Both these cases of co-localization merit further examination, however, since these two candidates map very close to the peaks of QTL for traits that closely match their characterized gene function. The QTL reported here also suggest a short list of other candidate genes for subsequent mapping in sorghum. For example, a flowering time QTL detected in this study maps to the short arm of chromosome 9 between *rz390* and *cdo580*, which both map to maize bin 8.05 within a few cM of the maize flowering time QTL *Vgt1* (Salvi et al. 2002).

The precise co-localization of the *Dw3* gene with QTL for plant height, rachis length, and branch length strongly suggests that IS3620C carries a wild-type *Dw3* allele. The BTx623 *Dw3* allele, which has an ~ 800 bp tandem duplication in exon 5, is similar in sequence to the unstable, dwarfing allele characterized by Multani et al. (2003), whereas the IS3620C *Dw3* allele does not contain this duplication (data not shown). Therefore, we hypothesize that the QTL at this locus are conditioned by the ~ 800 bp duplication in exon 5 of the *Dw3* gene.

The *Sb-ra2* gene co-localizes closely with one of only two QTL detected for primary branch number across the genome. A ~ 1.9 Kb fragment of the *Sb-ra2* locus was sequenced from each parent, including the entire *Sb-ra2* protein-coding region, ~ 400 bp of upstream sequence, and ~ 700 bp of downstream sequence. Only a single polymorphism was found, a 4 bp indel 72 bp downstream of the stop codon (data not shown). Thus, the predicted *Sb-RA2* proteins are identical between mapping parents, although it is possible that the parents differ in expression and/or translation of the *Sb-ra2* transcript.

Since the evidence from QTL and candidate gene co-localization is purely coincidental, the possible contributions of *Dw3* and *Sb-ra2* to variation in sorghum inflorescence architecture merit closer examination through fine-mapping and/or association mapping. The latter approach offers several advantages in that it is much quicker and can evaluate a much wider range of alleles, which is particularly important for a trait potentially subjected to diversifying selection. However, sorghum is a predominantly-selfing organism with moderate linkage disequilibrium (Hamblin et al. 2005), so association mapping may not be able to isolate single polymorphisms conditioning the phenotypes of interest. Association mapping is also potentially disadvantaged, relative to fine-mapping, when the target is a non-coding polymorphism not in close physical proximity to the gene whose expression it affects. At the B locus in maize, for example, expression is affected by tandem repeats over 100 KB upstream of the start codon (Stam et al. 2002). Used together, the fine-mapping and association mapping approaches can provide a significant increase in power (Meuwissen et al. 2002).

It is worth noting that the two candidate genes of greatest promise, *Dw3* and *Sb-ra2*, are both genes that were identified as maize mutants. In contrast to *LEAFY* and *TFL1*, which were discovered as *Arabidopsis* and *Antirrhinum* mutants, and to *bal1/LAX* and *bd1/fzp*, which were identified in both maize and rice, *br2/Dw3* and *ra2* (as well as *ra1* and *ids1*) are genes with mutant phenotypes described only in maize. Although firm inference cannot be drawn from such a small sample, this study nonetheless suggests that phylogenetic proximity is a critical consideration when selecting candidate genes from other species.

## Conclusions

In this study, we have mapped a number of candidate genes for grass inflorescence development, several of which co-localize closely with QTL for relevant inflorescence traits in sorghum across multiple environments. These initial results suggest that the approach employed here, which uses candidate genes from related model organisms to investigate quantitative variation in a target species with high phenotypic variance, may be highly effective. Since inflorescence architecture in sorghum is a complex trait with dramatic intraspecific variation, the possible contribution of candidate genes from other grasses, not limited to the small number described here, needs to be evaluated across a broader range of sorghum germplasm. Ultimately, the genes identified in sorghum may help explain how human selection, acting upon the common

genetic and physical architecture of the grass inflorescence, produced a range of viable, alternative outcomes in different domesticated cereals.

**Acknowledgments** We thank Toby Kellogg, Sarah Hake, and Gael Pressoir for critical advice, Robert Klein for providing plant materials, Delilah Wood for assistance with scanning electron microscopy, and Mark Sorrells and David Benschler for providing threshing equipment. The MORPH Research Coordination Network provided financial assistance to PJB.

## References

- Ashikari M, Sakakibara H, Lin S, Yamamoto T, Takashi T, Nishimura A, Angeles W, Qian Q, Kitano H, Matsuoka M (2005) Cytokinin oxidase regulates rice grain production. *Science* 309:741–745
- Bedell JA, Budiman MA, Nunberg A, Citek RW, Robbins D, Jones J, Flick E, Rohlfing T, Fries J, Bradford K, MeMenamy J, Smith M, Holeman H, Roe BA, Wiley G, Korf IF, Rabinowicz PD, Lakey N, McCombie WR, Jeddloh JA, Martienssen RA (2005) Sorghum genome sequencing by methylation filtration. *PLoS Biol* 3:e13
- Bombliks K, Wang R-L, Ambrose BA, Schmidt RJ, Meeley RB, Doebley JF (2003) Duplicate *FLORICAULA/LEAFY* homologs *zfl1* and *zfl2* control inflorescence architecture and flower patterning in maize. *Development* 130:2385–2395
- Bommert P, Satoh-Nagasawa N, Jackson D, Hirano H-Y (2005) Genetics and evolution of inflorescence and flower development in grasses. *Plant Cell Physiol* 46:69–78
- Bortiri E, Chuck G, Vollbrecht E, Rocheford T, Martienssen R, Hake S (2006) *ramosa2* encodes a LATERAL ORGAN BOUNDARY domain protein that determines the fate of stem cells in branch meristems of maize. *Plant Cell* 18:574–585
- Bout S, Vermerris W (2003) A candidate-gene approach to clone the sorghum *Brown Midrib* gene encoding caffeic acid O-methyltransferase. *Mol Genet Genom* 269:205–214
- Brown DG, Vision T (2000) MapPop 1.0: software for selective mapping and bin mapping. <http://www.bio.unc.edu/faculty/vision/lab/mappop/>
- Childs KL, Miller FR, Cordonnier-Pratt MM, Pratt LH, Morgan PW, Mullet JE (1997) The sorghum photoperiod sensitivity gene, *Ma3*, encodes a phytochrome B. *Plant Physiol* 113:611–619
- Chuck G, Meeley RB, Hake S (1998) The control of maize spikelet meristem fate by the *APETALA2*-like gene *indeterminate spikelet1*. *Genes Dev* 12:1145–1154
- Chuck G, Muszynski M, Kellogg E, Hake S, Schmidt RJ (2002) The control of spikelet meristem identity by the *branched silkless1* gene in maize. *Science* 298:1238–1241
- Clark RM, Linton E, Messing J, Doebley JF (2004) Pattern of diversity in the genomic region near the maize domestication gene *tb1*. *Proc Natl Acad Sci* 101:700–707
- Devos KM (2005) Updating the ‘crop circle’. *Curr Opin Plant Biol* 8:155–162
- Doggett H (1988) Sorghum, 2nd edn. Longman Scientific & Technical, Essex
- Doust AN, Devos KM, Gadberry MD, Gale MD, Kellogg EA (2005) The genetic basis for inflorescence variation between foxtail and green millet (*Poaceae*). *Genetics* 169:1659–1672
- Doyle J, Doyle JL (1987) Genomic plant DNA preparation from fresh tissue: CTAB method. *Phytochem Bull* 19:11

- Feltus FA, Hart GE, Schertz KF, Casa AM, Kresovich S, Klein PE, Paterson AH (2006) Alignment of genetic maps and QTLs between inter- and intra-specific sorghum populations. *Theor Appl Genet* (in press)
- Gallavotti A, Zhao Q, Kyojuka J, Meeley RB, Ritter MK, Doebley JF, Pe ME, Schmidt RJ (2004) The role of *barren-stalk1* in the architecture of maize. *Nature* 432:630–635
- Gaut BS (2002) Evolutionary dynamics of grass genomes. *New Phytol* 154:15–28
- Gocal GF, King RW, Blundell CA, Schwartz OM, Andersen CH, Weigel D (2001) Evolution of floral meristem identity genes. Analysis of *Lolium temulentum* genes related to APETALA1 and LEAFY of *Arabidopsis*. *Plant Physiol* 125:1788–1801
- Hamblin MT, Salas Fernandez MG, Casa AM, Mitchell SE, Paterson AH, Kresovich S (2005) Equilibrium processes cannot explain high levels of short- and medium-range linkage disequilibrium in the domesticated grass *Sorghum bicolor*. *Genetics* 171:1247–1256
- Harlan JR, De Wet JMJ (1972) A simplified classification of cultivated sorghum. *Crop Sci* 12:172–176
- Ikeda K, Sunohara H, Nagato Y (2004) Developmental course of inflorescence and spikelet in rice. *Breeding Sci* 54:147–156
- Jensen CS, Salchert K, Nielsen KK (2001) A *TERMINAL FLOWER1*-like gene from perennial ryegrass involved in floral transition and axillary meristem identity. *Plant Physiol* 125:1517–1528
- Kaplinsky NJ, Freeling M (2003) Combinatorial control of meristem identity in maize inflorescences. *Development* 130:1149–1158
- Kim J-S, Islam-Faridi MN, Klein PE, Stelly DM, Price HJ, Klein RR, Mullet JE (2005) Comprehensive molecular cytogenetic analysis of sorghum genome architecture; distribution of euchromatin, heterochromatin, genes and recombination in comparison to rice. *Genetics* 171:1963–1976
- Komatsu M, Chujo A, Nagato Y, Shimamoto K, Kyojuka J (2003a) *FRIZZY PANICLE* is required to prevent the formation of axillary meristems and to establish floral meristem identity in rice spikelets. *Development* 130:3841–3850
- Komatsu K, Maekawa M, Ujiie S, Satake Y, Furutani I, Okamoto H, Shimamoto K, Kyojuka J (2003b) *LAX* and *SPA*: major regulators of shoot branching in rice. *Proc Natl Acad Sci* 100:11765–11770
- Kyojuka J, Konishi S, Nemoto K, Izawa T, Shimamoto K (1998) Down-regulation of *RFL*, the *FLO/LFY* homolog of rice, accompanied with panicle branch initiation. *Proc Natl Acad Sci* 95:1979–1982
- Lauter N, Doebley JF (2002) Genetic variation for phenotypically invariant traits detected in teosinte: implications for the evolution of novel forms. *Genetics* 160:333–341
- Lin Y-R, Schertz KF, Paterson AH (1995) Comparative analysis of QTLs affecting plant height and maturity across the *Poaceae*, in reference to an interspecific sorghum population. *Genetics* 141:391–411
- Menz MA, Klein RR, Mullet JE, Obert JA, Unruh NC, Klein PE (2002) A high-density genetic map of *Sorghum bicolor* (L.) Moench based on 2926 AFLP, RFLP, and SSR markers. *Plant Mol Biol* 48:483–499
- Meuwissen THE, Karlsten A, Lien S, Olsaker I, Goddard ME (2002) Fine mapping of a quantitative trait locus for twinning rate using combined linkage and linkage disequilibrium mapping. *Genetics* 161:373–379
- Michaels SD, Amasino RM (1998) A robust method for detecting single-nucleotide changes as polymorphic markers by PCR. *Plant J* 14:381–385
- Multani DS, Briggs SP, Chamberlin MA, Blakeslee JJ, Murphy AS, Johal GS (2003) Loss of an MDR transporter in compact stalks of maize *br2* and sorghum *dw3* mutants. *Science* 302:81–84
- Robertson DS (1989) Understanding the relationship between qualitative and quantitative genetics. In: Helentjaris T, Burr B (eds) Development and application of molecular markers to problems in plant genetics. Cold Spring Harbor Lab Press, New York, pp 81–87
- Salvi S, Tuberosa R (2005) To clone or not to clone plant QTLs: present and future challenges. *Trends Plant Sci* 10:297–304
- Salvi S, Tuberosa R, Chiapparino E, Maccaferri M, Veillet S, van Beuningen L, Isaac P, Edwards K, Phillips RL (2002) Toward positional cloning of Vgt1, a QTL controlling the transition from the vegetative to the reproductive phase in maize. *Plant Mol Biol* 48:601–613
- Sanchez AC, Subudhi PK, Rosenow DT, Nguyen HT (2002) Mapping QTLs associated with drought resistance in sorghum (*Sorghum bicolor* L. Moench). *Plant Mol Biol* 48:713–726
- Stam M, Bebele C, Ramakrishna W, Dorweiler JE, Bennetzen JL, Chandler VL (2002) The regulatory regions required for B' paramutation and expression are located far upstream of the maize b1 transcribed sequences. *Genetics* 162:917–930
- Stephens JC, Miller FR, Rosenow DT (1967) Conversion of alien sorghums to early combine genotypes. *Crop Sci* 7:396
- Swigonova Z, Lai J, Ma J, Ramakrishna W, Llaca V, Bennetzen JL, Messing J (2004) Close split of sorghum and maize genome progenitors. *Genome Res* 14:1916–1923
- Thornsberry JM, Goodman MM, Doebley JF, Kresovich S, Nielsen D, Buckler ES (2001) *Dwarf8* polymorphisms associate with variation in flowering time. *Nat Genet* 28:286–289
- Trevaskis B, Bagnall DJ, Ellis MH, Peacock WJ, Dennis ES (2003) MADS box genes control vernalization-induced flowering in cereals. *Proc Natl Acad Sci* 100:13099–13104
- Upadyayula N, da Silva HS, Bohn MO, Rocheford TR (2005) Genetic and QTL analysis of maize tassel and ear inflorescence architecture. *Theor Appl Genet* (in press)
- Vollbrecht E, Springer PS, Goh L, Buckler ES, Martienssen RA (2005) Architecture of floral branch systems in maize and related grasses. *Nature* 436:1119–1126
- Wang S, Basten CJ, Zeng Z-B (2005) Windows QTL Cartographer 2.5. Statistical Genetics, NCSU
- Zhang S, Hu W, Wang L, Lin C, Cong B, Sun C, Luo D (2005) *TFL1/CEN*-like genes control intercalary meristem activity and phase transition in rice. *Plant Sci* 168:1393–1408