Genotyping by Sequencing of 393 Sorghum bicolor BTx623 × IS3620C Recombinant Inbred Lines Improves Sensitivity and Resolution of QTL Detection

WenQian Kong, Changsoo Kim, Dong Zhang, Hui Guo, Xu Tan, Huizhe Jin, Chengbo Zhou, Lan-shuan Shuang, Valorie Goff, Uzay Sezen, Gary Pierce, Rosana Compton, Cornelia Lemke, Jon Robertson, Lisa Rainville, Susan Auckland, and Andrew H. Paterson

Plant Genome Mapping Laboratory, University of Georgia, Athens, GA 30605
ORCID IDs: 0000-0002-4540-6165 (W.K.); 0000-0003-2159-0487 (A.H.P.)

ABSTRACT We describe a genetic map with a total of 381 bins of 616 genotyping by sequencing (GBS)-based SNP markers in a F6-F8 recombinant inbred line (RIL) population of 393 individuals derived from crossing S. bicolor BTx623 to S. bicolor IS3620C, a guinea line substantially diverged from BTx623. Five segregation distorted regions were found with four showing enrichment for S. bicolor alleles, suggesting possible selection during formation of this RIL population. A quantitative trait locus (QTL) study with this number of individuals, tripled relative to prior studies of this cross, provided resources, validated previous findings, and demonstrated improved power to detect plant height and flowering time related QTL relative to other published studies. An unexpected low correlation between flowering time and plant height permitted us to separate QTL for each trait and provide evidence against pleiotropy. Ten non-random syntenic regions conferring QTL for the same trait suggest that those QTL may represent alleles at genes functioning in the same manner since the 96 million year ago genome duplication that created these syntenic relationships, while syntenic regions conferring QTL for different trait may suggest sub-functionalization after duplication. Collectively, this study provides resources for marker-assisted breeding, as well as a framework for fine mapping and subsequent cloning of major genes for important traits such as plant height and flowering time in sorghum.

KEYWORDS quantitative trait loci genetic mapping plant architecture flowering Paleo-duplication Multiparent Advanced Generation Inter-Cross (MAGIC) multiparental populations MPP

The most drought resistant of the world’s top five cereal crops, sorghum, contributes 26-29% of calories in the human diet in semi-arid areas of Africa (FAO), on some of the world’s most degraded soils and often with limited water inputs. A multi-purpose crop, sorghum has been traditionally used as grain and straw, and is also a promising crop for bioenergy production from starch, sugar, or cellulose on marginal lands with limited water and other resources (Rooney et al. 2007). Botanically, Sorghum bicolor is a model for plants that use C4 photosynthesis, improving carbon assimilation especially at high temperature, complementary to C3 model plants such as Oryza sativa. The sequenced ~730 megabase sorghum genome (Paterson et al. 2009) has not experienced genome duplication in an estimated ~96 million years (Wang et al. 2015), making it particularly useful to study other C4 plants with large polyploid genomes such as maize and many other grasses in the Saccharinae clade including Miscanthus and Saccharum (sugarcane).

Evolution, natural selection and human improvement of sorghum have contributed to great morphological diversity of Sorghum bicolor spp. Cultivated forms of this species can be classified into five botanical
races, bicolor, guinea, caudatum, kafir and durra; with ten intermediate races recognized based on inflorescence architecture and seed morphology (de Wet and Hukcbaay 1967; Harlan and Dewey 1972). *Sorghum bicolor* also has many wild relatives such as *S. propinquum* (2n = 2x = 20), diverged from *S. bicolor* 1-2 million years ago (Feltus et al. 2004); *Sorghum halepense* (2n = 4x = 40), an invasive and weedy species formed by unintentional crossing of *S. bicolor* and *S. propinquum* (Paterson et al. 1995); and many hybrids between these species. These many possible intra- and interspecific crosses have made sorghum a particular interesting model for dissecting the genetic control of complex traits such as plant height and maturity (Lin et al. 1995; Hart et al. 2001; Feltus et al. 2006; Takai et al. 2012), tillering and vegetative branching (Alam et al. 2014; Kong et al. 2014; Kong et al. 2015), perenniality related traits (Paterson et al. 1995; Washburn et al. 2013), sugar composition (Murray et al. 2008a; Murray et al. 2008b; Ritter et al. 2008; Shiringani et al. 2010; Vandenbergink et al. 2013), stay-green leaves and plants (Haussmann et al. 2002; Kassahun et al. 2010), drought resistance (Tuinstra et al. 1998; Kebede et al. 2001; Sanchez et al. 2002), disease and insect resistance (Katsar et al. 2002; Totad et al. 2005).

*Sorghum bicolor* IS3620C is a representative of botanical race ‘guinea’, and is substantially diverged from *S. bicolor* BTx623 among the sorghum races. Prior workers described a genetic map with 323 RFLP and 147 SSR markers (Bhattramakk et al. 2000; Kong et al. 2000), and a quantitative trait locus (QTL) study with 137 F6-8 RILs demonstrated its usefulness for QTL detection by discovering as many as 27 QTL traits of agronomical importance, such as plant height, maturity, number of basal tillers and panicle length (Hart et al. 2001; Feltus et al. 2006).

Next generation sequencing (NGS) has brought new power to revealing allelic differences among individuals by detecting large numbers of single nucleotide polymorphisms (SNP). While it is neither necessary nor practical (yet) to generate whole genome sequences for every individual in a designed mapping population, reduced representation library (RRL) sequencing has been widely utilized and proved to be both a cost and labor efficient tool for genotyping (Andolfatto et al. 2011; Kim et al. 2015). Although current genotyping by sequencing (GBS) platforms and pipelines still face various issues regarding accuracy (especially in polyploids with high levels of heterozygosity), GBS is still a very powerful tool to generate genetic maps with high quality and resolution in nearly homozygous populations with reference genome sequences available (Kim et al. 2015).

In this paper, we describe a genetic map with a total of 381 bins of 616 GBS-based SNP markers in a recombinant inbred line (RIL) population of 393 individuals derived from two divergent *S. bicolor* genotypes, BTx623 and IS3620C. Recent work that utilized digital genotyping to saturate the genetic map of 137 individuals (Morishige et al. 2001; Feltus et al. 2006) has validated QTL for height more precisely but showed little increase in power to detect QTL relative to prior analysis using a lower density of markers (Hart et al. 2001). Our study has tripled the number of individuals described in previous studies (Bhattramakk et al. 2000; Kong et al. 2000; Hart et al. 2001; Morishige et al. 2013), increasing both the number of discernable recombination events and the number of individuals carrying each parental allele, increasing power to detect QTL. A quantitative trait analysis of days to flowering and components of plant height demonstrate the improved power to detect quantitative trait loci (QTL) in this expanded population. Syntenic relationships among sorghum QTL may reflect homoeologous regions retaining the same or similar function after a genome duplication 96 million years ago (Wang et al. 2015). Results of this new QTL mapping build on those of many other studies (Lin et al. 1995; Hart et al. 2001; Kebede et al. 2001; Brown et al. 2006; Srinivas et al. 2009; Shiringani et al. 2010; Morris et al. 2013; Zhang et al. 2015), enriching current resources and providing a better understanding of the genetic control of plant height and days to flowering of *S. bicolor*.

### MATERIALS AND METHODS

#### Genetic stocks

The mapping population is comprised of 399 F7-8 RILs derived by selfing a single F1 plant from *S. bicolor* BTx623 and IS3620C as described (Kong et al. 2000; Hart et al. 2001; Burrow et al. 2011). This RIL population was planted at the University of Georgia Plant Science Farm, Watkinsville, GA, USA on May 10, 2011 and May 18, 2012. Single 3-m row plots of each RIL were machine planted in a completely randomized design. The experiment consists of 44 rows, 10 plots and two representative samples are phenotyped within each row plot.

#### Genotyping

Leaf samples of the RIL population was frozen at -80°C and lyophilized for 48 hr. Genomic DNA was extracted from the lyophilized leaf sample based on Aljanabi et al. (1999).

Our GBS platform is a slightly modified version of Multiplex Shotgun Genotyping (MSG) (Andolfatto et al. 2011) combined with the Tassel GBS analysis pipeline. The basic sequencing platform is an in-house Illumina MiSeq that generates up to 25 million reads of 150 base pairs (bps) fragments per run with single-end sequencing. With one sequencing run of this platform, we obtained 7103 raw SNPs and 691 polymorphic SNPs that were sufficient for genetic mapping for this RIL population (Kim et al. 2015).

SNP ‘calling’ (inference) was based on the reference genome sequence of *Sorghum bicolor* (Paterson et al. 2009). Alignment used the Burrows-Wheeler Aligner (BWA) for single end read (samse). In Tassel GBS, the first 64 bps of each reads were mapped onto a reference genome to decide the position of the reads. SNPs were called based on the alignment of reads to the reference genome. Heterozygosity at a locus is called if two alleles are each inferred to be present at a probability greater than that of sequencing error. Raw SNP data from the Tassel GBS pipeline were further filtered based on several criteria: (a) SNPs were removed if the minor allele frequency is less than 5% or the proportion of missing genotypes greater than 40%; (b) In order to reduce the number of redundant SNPs in studies where strong linkage disequilibrium necessitates only 5–10 cM resolution, we merged SNPs for which pairwise linkage disequilibrium (r2) is greater than 0.9, deriving consensus genotypes in a manner minimizing missing genotypes. SNPs are further merged if the Pearson’s correlation between them is larger than 0.95; (c) For bi-parental populations, the missing genotype of one parental line can be imputed by offspring genotypes if the genotype of the other parent at the locus is known. After these filtering steps, SNP data are used for genetic mapping.

#### Map construction

A genetic map using 616 SNP markers was created using R/qtl (Broman et al. 2003). We further assigned bins for each chromosome to merge markers within 1 cM in genetic distance. Bin genotypes were used for genetic mapping for

---

2564 | W. Kong et al.
Table 1 Summary Statistics for the S. bicolor BTx623 x IS3620C Genetic Map and Bin Map

| Chr | Marker No. | Length (cM) | Avg Spacing (cM) | Max Spacing | Chr | Marker No. | Length | Avg Spacing | Max Spacing |
|-----|------------|-------------|------------------|-------------|-----|------------|--------|-------------|-------------|
| 1   | 140        | 186.6       | 1.3              | 8.5         | 1   | 66         | 189.7  | 2.9         | 8.6         |
| 2   | 83         | 173.2       | 2.1              | 8.9         | 2   | 50         | 180.2  | 3.7         | 9.7         |
| 3   | 80         | 158.6       | 2.0              | 11.0        | 3   | 43         | 160.7  | 3.8         | 11.8        |
| 4   | 57         | 158.9       | 2.8              | 21.1        | 4   | 42         | 164.9  | 4.0         | 21.3        |
| 5   | 23         | 119.2       | 5.4              | 27.6        | 5   | 21         | 123.7  | 6.2         | 27.6        |
| 6   | 53         | 106.9       | 2.1              | 14.6        | 6   | 39         | 111.8  | 2.9         | 14.9        |
| 7   | 44         | 125         | 2.9              | 17.5        | 7   | 30         | 127.8  | 4.4         | 17.7        |
| 8   | 35         | 100.8       | 3.0              | 17.5        | 8   | 21         | 101.6  | 5.1         | 17.8        |
| 9   | 36         | 120.5       | 3.4              | 21.4        | 9   | 23         | 122.4  | 5.6         | 21.6        |
| 10  | 65         | 120.6       | 1.9              | 5.3         | 10  | 46         | 126.1  | 2.8         | 5.9         |
| Overall | 616 | 1370.3   | 2.3              | 27.6        | Overall | 381   | 1408.8 | 3.8       | 27.6        |

QTL mapping

QTL were detected for five traits of interest: plant height (PH), the overall length of a plant; base to flag length (BTF), the length from the base of the plant to the flag leaf; flag to rachis length (FTR), the length from the flag leaf to rachis (a positive sign was assigned if the position of the rachis is taller than the flag leaf; a negative value was assigned if the rachis was ‘buried’ in the flag leaf); number of nodes (ND); and days to flowering (FL), the average days to flowering for the first five plants for each genotype.

We combined the phenotypic data using Best Linear Unbiased Prediction (BLUP) by treating individuals, years, replications nested within years and the interactions between individuals and years as random, since heritability for the traits of interest were relatively high. In 2011, we observed and recorded a soil type change within the experimental field. Since heritability for the traits of interest were relatively high, we expect to see half of the alleles come from each parent, i.e., the expected segregation ratio is 1:1 for each marker locus. A deviation from this ratio may be a result of gametic or zygotic selection. The BTx623 x IS3620C genetic map reveals several clusters of markers experiencing segregation distortion on chromosomes 1, 4, 5, 8, and 9 (Figure 2), peaking at 48.0, 153.6, 64.4, 47.5, and 122.4 cM, corresponding to bin markers. File S4 and S5 contain phenotypes from 2011 and 2012 respectively. Supplemental material available at Figshare: https://doi.org/10.25387/g3.6304538.

RESULTS

Genetic map

A total of 399 RILs were genotyped with 690 SNP markers. Six individuals for which genotyping data suggested three times more than the average number of recombination events were deemed erroneous and removed from the analysis. Marker ordering first follows the published sorghum genome sequence (Paterson et al. 2009). A de novo marker ordering method is also used to compare the order of the genetic map with the reference-based method, but no obvious differences were observed for these two methods in terms of the LOD scores. We excluded 74 unlinked SNPs and obtained an initial genetic map with a total of 616 markers on the ten sorghum chromosomes. As detailed in the methods, we combined SNPs that are within 1cM in genetic distance to construct a genetic map with 381 bins (bin map) with varying genetic distance of 1404.8 cM, with average spacing of 3.8 cM between loci and the largest gap being 27.6 cM on chromosome 5 (Table 1). The percentages of missing genotypes are 24% and 18.4% for the initial and bin maps, respectively. About 54.2% of the alleles of the RIL population come from S. bicolor BTx623, and 45.8% from S. bicolor IS3620C.

The order of the genetic map agrees closely with the physical positions of loci on the genome sequence (Figure 1), suggesting that the GBS method used yields a high quality genetic map in this nearly-homozygous diploid population. Markers are generally more concentrated in the distal regions of each chromosome than central regions (consistent with the general distribution of low-copy DNA sequences in sorghum: Paterson et al. 2009), although distributions of markers on each chromosome vary. For example, we have observed a much larger pericentromeric region on chromosome 7 than on chromosome 1.

Segregation distortion

Segregation distortion occurs when the segregation ratio of offspring at a locus deviates from the Mendelian expectation. In a RIL population, we expect to see half of the alleles come from each parent, i.e., the expected segregation ratio is 1:1 for each marker locus. A deviation from this ratio may be a result of gametic or zygotic selection. The BTx623 x IS3620C genetic map reveals several clusters of markers experiencing segregation distortion on chromosomes 1, 4, 5, 8, and 9 (Figure 2), peaking at 48.0, 153.6, 64.4, 47.5, and 122.4 cM, corresponding to
14.8, 65.1, 49.9, 27.8-38.5 and 59.5 in physical distance, respectively. All regions but the one on chromosome 8 show enrichment of \( S. \text{bicolor} \) alleles. The most extreme segregation distorted region is on chromosome 1, spanning 0-105 cM (Figure 2) with a ratio of 358:35 (\( P = 1.10E-59 \)) at the peak marker, S1_14765342. Peng et al. (1999) also observed this long-spanning segregation distorted region on chromosome 1, despite using a population with a much smaller sample size and marker numbers.

Interestingly, the same general region of extreme segregation distortion on chromosome 1 is also found in two \( S. \text{bicolor} \times S. \text{propinqua} \) derived populations (Bowers et al. 2003; Kong 2013), and remarkably, all three populations mapped this distortion peak at \( \sim 14 \) Mb in physical distance. This correspondence between populations suggests that alleles from \( S. \text{bicolor} \) might be selected for at this genomic position, however, the exact mechanism and the genes involved are unclear.

**QTL Mapping**

The present BTx623 x IS3620C RIL map provides higher power for detecting QTL than a prior map of a subset of these progenies. As examples, we have investigated QTL for five phenotypic traits, plant height (\( \text{PH} \)), base to flag length (\( \text{BTF} \), flag to rachis length (\( \text{FTR} \)), number of nodes (\( \text{ND} \)) and days to flowering (\( \text{FL} \)). Means, standard deviation and other summary statistics are shown in Table S1. Broad-sense heritability estimates for all five traits are relatively high (Table S1). It is interesting that the average \( \text{PH} \) of the progenies is 98.62 cm, greater than the average of either parents, 94.40 for BTx623 and 84.49 for IS3620C. Likewise, \( \text{BTF} \) is greater than the average of either parent. Both height components show substantial genetic variation, indicating that each parent contributes different alleles for \( \text{PH} \) to their progenies (Table S1). Moreover, the relatively high genetic variation in \( \text{PH} \) in this population fosters discovery of QTL, despite that the difference between the two parents is relatively small.

\( \text{PH} \) and \( \text{BTF} \) are highly correlated in this population (Table S2), with a correlation coefficient of 0.8983 (\( P < 0.001 \)). We detect a total of seven and five QTL for \( \text{PH} \) and \( \text{BTF} \), accounting for 40.13% and 41.58% of the total phenotypic variances, respectively (Table S3). Three QTL on chromosomes 3, 6 and 7 overlap for these two traits. The QTL on chromosomes 6 and 7 account for the majority of the phenotypic variance explained for these traits, ~28% and ~32% for \( \text{PH} \) and \( \text{BTF} \), respectively. These two large effect QTL might be related to previously defined \( \text{PH} \) genes, presumably \( dw2 \) on chromosome 6 and \( dw3 \) (Sb07g023730) on chromosome 7 (Quinby and Karper 1945; Multani et al. 2003).

Not only do we detect more QTL than were found in the previous study of a subset of this population (Hart et al. (2001)), seven QTL for \( \text{PH} \) and nine QTL for \( \text{FL} \) in our study vs. five and three from the previous study, but some of the QTL intervals also significantly reduced (Table S3). For example, the 1-LOD interval of QTL on chromosome 7 is narrowed from ~26 cM previously to only ~3 cM in this study (from 57.7 Mb to 59.5 Mb in physical distance), and harbors the gene Sb07g023730 (\( \text{DW3} \)) at ~58.6 Mb. This example indicates that nearly tripling the numbers of individuals and increasing marker density greatly increased the power of QTL detection.

An interesting phenotype that we observed is distance from the flag leaf to the rachis. We distinguished whether or not the rachis is immersed in the flag leaf (see Materials and Methods). The correlation coefficient (Table S2) between \( \text{FTR} \) and \( \text{PH} \) (\( r = 0.0778 \)), though significant at \( P < 0.01 \), is not nearly high as the correlation between the \( \text{BTF} \) and \( \text{PH} \) (\( r = 0.8983 \)), suggesting that the genetic control of these two traits might be different. Indeed, QTL mapping suggests that the genetic control of \( \text{FTR} \) is quite different from that of \( \text{PH} \) and \( \text{BTF} \), with only one QTL (qFTR7.1) on chromosome 7 overlapping with QTL for \( \text{PH} \) and \( \text{FTR} \) (qPH7.1 and qBTF7.1). While the one-LOD QTL interval for qPH10.1 overlaps with qFTR10.1 to some extent, there is no solid evidence to conclude that they are controlled by the same genetic factors, given that the likelihood peaks of the QTL for these two traits are \( \sim 10 \) cM apart. We detected a total of five QTL for \( \text{FTR} \), explaining 28.21% of the total phenotypic variance (Table S3). The additive effect of \( \text{FTR} \) needs to be carefully interpreted, especially when compared to the additive effect of \( \text{PH} \). Since the average value of \( \text{FTR} \) is negative, a negative number for the additive effect for this trait indicates increased \( \text{FTR} \) for a particular allele. For example, the additive effect for qFTR7.1 is -1.19, which
indicates that FTR of plants carrying the IS3620 alleles are actually longer than those of plants carrying the BTx623 alleles. Both alleles from IS3620C for qPH7.1 and qFTR7.1 have the same effect of increasing length, although the sign of their additive effects is different.

We have detected a total of six QTL for number of nodes (ND) in this population (Figure 3), collectively explaining 32.07% of the phenotypic variance. The largest effect QTL is qND8.1, with a LOD score of 12.96 and explaining 11.15% of the variance. QTL for ND rarely overlap with other PH related traits—only qND10.1 shows some correspondence with other height related traits, marginally overlapping with qFTR10.1.

In this study, PH and FL were not significantly correlated (Table S2), an unexpected finding compared to many other sorghum studies (Lin et al. 1995; Murray et al. 2008b; Ritter et al. 2008). A total of nine QTL are detected for FL, substantially more than the 4-6 conventionally thought to influence this trait in a wide range of sorghum genotypes (Quinby and Karper 1945), although only collectively explaining 46.33% of the total phenotypic variance. QTL intervals for FL rarely overlap with those for PH, BTF and FTR. However, five out of six QTL for ND overlap with QTL for FL, and the sign of the allelic effect suggests that plants with more nodes usually flower late. This result indicates that some genes might have pleiotropic effects on these traits or genes for that these two traits are linked and have been selected simultaneously.

**QTL correspondence with other studies**

Traits related to sorghum PH and FL have been extensively studied in many QTL experiments (Lin et al. 1995; Hart et al. 2001; Kebede et al. 2001; Brown et al. 2006; Ritter et al. 2008; Srinivas et al. 2009; Shiringani et al. 2010) and two genome wide association studies (Morris et al. 2013; Kong et al. 2015). The Comparative Saccharinae Genome Resource QTL database (Zhang et al. 2013) aids comparisons of QTL intervals across different studies in sorghum and facilitates validation of QTL for traits of interest. All our QTL detected for PH have been found in other studies. Three PH QTL, qPH6.1, qPH7.1 and qPH9.1, also found in two GWAS studies (Morris et al. 2013; Zhang et al. 2015), are likely to correspond to three dwarfing genes in sorghum, dw2, dw3 and dw1 (Quinby and Karper 1945). The accuracy of our QTL study can be demonstrated by the gene known to cause the dw3 phenotype, Sb07g023730, on chromosome 7. Our QTL study narrowed the 1-lod interval for this QTL to 2Mb (57.7-59.5Mb), with a peak at 58.4Mb, close to the 58.6 Mb location of the causal gene (Multani et al. 2003).

![Figure 2 Segregation distortion pattern in the S. bicolor BTx623 x IS3620C RIL population. y-axis is negative logarithm of p value with a base of 10, x-axis is the genetic distance. Distorted regions are plotted in orange.](image-url)
A total of 6 QTL controlling FL in the BTx623 × IS3620C RILs, qFL1.2, qFL3.1, qFL6.1, qFL8.2, qFL9.1 and qFL10.1, showed correspondence with QTL found in other studies, (Lin et al. 1995; Hart et al. 2001; Brown et al. 2006; Shiringani et al. 2010; Yang et al. 2014) and two QTL, qFL1.1 and qFL6.1, are novel. The QTL with the largest effects on FL, qFL8.2 with a LOD score of 16.1 and explaining 11.16% of phenotypic variance; and qFL9.1 with a LOD score of 11.4 and explaining 7.69% of phenotypic variance, have been consistently found in many independent QTL and GWAS studies (Lin et al. 1995; Brown et al. 2008; Morris et al. 2013; Zhang et al. 2015). Identification of the genes underlying these QTL regions might be especially important. The peak of qFL9.1 is located at /C24 59.3Mb in our study, close to significant peaks at /C24 58.7Mb and /C24 59.0Mb for FL found in a GWAS study (Zhang et al. 2015).

Syntenic study
A total of 30 out of 202 genomic regions contain QTL found in this study were located in colinear locations within sorghum (Paterson et al. 2009) resulting from genome duplication events (Table 2 and Figure 4), as confirmed using the Plant Genome Duplication Database (Lee et al. 2013). Among these, a total of five regions on chromosomes 1 (1), 3 (3), 9 (1) are duplicated within the same chromosome. Among the 25 duplicated genomic regions located on different chromosomes, ten syntenic regions contain the same trait (2 for ND, 3 for PH, 4 for FL and 1 for BTF), which is significantly more than expected to occur by chance (P = 0.0002), and 15 regions contain different traits (Figure 4).

DISCUSSION
The present study providing a relatively dense GBS-based map for 393 individuals of the singularly-important S. bicolor BTx623 × IS3620C RIL population identifies new QTL and increases precision of mapping previously-known QTL, providing both an important resource and new information about the genetic control of important sorghum traits. Benefiting both from an increased sample size and GBS, our study has demonstrated increased power and accuracy of detecting QTL, relative to previous studies of a total of 137 individuals (Hart et al. 2001; Feltus et al. 2006). We also discovered a total of five regions with segregation distortion, possibly due to gametic or zygotic selection during the formation of this RIL population.

There are several advantages of the bin mapping strategy used in this paper. First, it reduces the percentage of missing genotypes by combining the genotypes of adjacent markers. For this experiment, the percentage of missing genotypes is reduced by 5.6% with the bin map. Moreover, QTL intervals are usually 5-10 cM for a typical bi-parental QTL experiment; high marker density does not significantly increase the power of detecting QTL (Lander and Botstein 1989) while combining markers may increase computational processing speed.

Within this single population, we now find more QTL for PH and FL than have been classically thought to segregate in all forms of Sorghum bicolor (Quinby and Karper 1945), reiterating a conclusion from meta-analysis of multiple populations (Zhang et al. 2015) that these traits could not be accounted for by the 4-6 loci suggested by classical studies. Most of the QTL we mapped here correspond to QTL found in other studies, improving confidence in our result. An example is re-identification of the dw3 locus on chromosome 7, a P-glycoprotein auxin transporter (Multani et al. 2003), which proved the power and accuracy of our QTL study by narrowing the QTL interval from ~26 cM to 3 cM known to harbor the causal gene. More generally, QTL mapping complements other data types toward identification of causal genes. For example, a
we also find a QTL controlling FL, qFL9.1, that is ~30cM from qPH9.1. This result again suggests that two separate QTL control these PH and FL effects, a conclusion that is also supported by another study (Thurber et al. 2013).

We discovered a total of 30 syntenic regions within the sorghum genome sequence (Paterson et al. 2009; Lee et al. 2013) containing QTL, with 10 regions containing QTL responsible for the same trait (Table 2 and Figure 4). This non-random correspondence between regions of the genome conferring the same traits indicates that the ten syntenic regions correspond to the same QTL (homeologous) genes that may function in the same ways despite being duplicated 96 million years ago (Wang et al. 2015), while the syntenic regions with different traits may suggest potential sub-functionalization of genes after duplication.

Components of PH and FL have been and will continue to be important for sorghum breeding programs. The past century has witnessed breeding for modern varieties with a particular plant type, for example a semi-dwarf type, to realize striking increases in production such as those which led to the Green Revolution (Evenson and Gollin 2003). The concept of ideotype breeding (Donald 1968), is still an ongoing priority for many breeding programs to increase food and feed production, adapt to climate change and minimize inputs. Genetic components discovered for plant height related traits and flowering time in this study, together with closely-linked diagnostic DNA markers that permit their selection at seedling stages or in non-target environments, may benefit breeding for plant types idealized for the different purposes that sorghum is used. Specifically, little correspondence between PH and FL, together with narrowed QTL intervals, facilitates accurate selection for each trait.
and their correspondence with those from many other studies also provides a framework for fine mapping or subsequent cloning of major genes for PH and FL in sorghum.

ACKNOWLEDGMENTS
The late K. Schertz, USDA-ARS and Texas A&M University, was instrumental in producing the BTx623·IS3620C RILs. We thank the members of the Paterson lab for valuable help and suggestions. We appreciate financial support from the AFRI Plant Growth and Development Program (2009-03477), DOE-USDA Plant Feedstock Genomics Program (2012-03304), and USDA Biotechnology Risk Assessment Program (2012-01658).

LITERATURE CITED
Alam, M. M., E. S. Mace, E. J. van Oosterom, A. Cruickshank, C. H. Hunt et al., 2014 QTL analysis in multiple sorghum populations facilitates the dissection of the genetic and physiological control of tillering. Theor. Appl. Genet. 127: 2253–2266. https://doi.org/10.1007/s00122-014-2377-9
Aljanabi, S. M., L. Forget, and A. Dookun, 1999 An improved and rapid protocol for the isolation of polysaccharide- and polyphenol-free sugarcane DNA. Plant Mol. Biol. Report. 17: 1–8. https://doi.org/10.1023/A:1007692929505
Andolfatto, P., D. Davison, D. Erezyilmaz, T. T. Hu, J. Mast et al., 2011 Multiplexed shotgun genotyping for rapid and efficient genetic mapping. Genome Res. 21: 610–617. https://doi.org/10.1101/gr.115402.110
Bhattramakki, D., J. M. Dong, A. K. Chhabra, and G. E. Hart, 2000 An integrated SSR and RFLP linkage map of Sorghum bicolor (L.) Moench. Genome 43: 988–1002. https://doi.org/10.1139/g00-074
Bowers, J. E., C. Abbey, S. Anderson, C. Chang, X. Draye et al., 2003 A high-density genetic recombination map of sequence-tagged sites for Sorghum, as a framework for comparative structural and evolutionary genomics of tropical grains and grasses. Genetics 165: 367–386.
Broman, K. W., H. Wu, S. Sen, and G. A. Churchill, 2003 R/qtl: QTL mapping in experimental crosses. Bioinformatics 19: 889–890. https://doi.org/10.1093/bioinformatics/btg112
Brown, P. J., P. E. Klein, E. Bortiri, C. B. Acharya, W. L. Rooney et al., 2006 Inheritance of inflorescence architecture in sorghum. Theor. Appl. Genet. 113: 931–942. https://doi.org/10.1007/s00122-006-0352-9
Brown, P. J., W. L. Rooney, C. Franks, and S. Kresovich, 2008 Efficient mapping of plant height quantitative trait loci in a sorghum association population with introgressed dwarfing genes. Genetics 180: 629–637. https://doi.org/10.1534/genetics.108.092239
Burow, G. B., R. R. Klein, C. D. Franks, P. E. Klein, K. F. Schertz et al., 2011 Registration of the BTx623/IS3620C Recombinant Inbred Mapping Population of Sorghum. J. Plant Regist. 5: 141–145. https://doi.org/10.3198/jpr2010.04.0219crmp
Cuevas, H. E., C. Zhou, H. Tang, P. P. Khadke, S. Das et al., 2016 The Evolution of Photoperiod-Insensitive Flowering in Sorghum, A Genomic Model for Panicoid Grasses. Mol. Biol. Evol. 33: 2417–2428. https://doi.org/10.1093/molbev/msw120
de Wet, J. M. J., and J. P. Huckabay, 1967 Origin of Sorghum Bicolor. 2. Distribution and Domestication. Evolution 21: 787–802. https://doi.org/10.1111/j.1558-5646.1967.tb03434.x
Donald, C. M., 1968 The breeding of crop ideotypes. Euphytica 17: 385–403. https://doi.org/10.1007/BF00056241
Evenson, R. E., and D. Gollin, 2003 Assessing the impact of the green revolution, 1960 to 2000. Science 300: 758–762. https://doi.org/10.1126/science.1078710
Feltus, F. A., G. E. Hart, K. F. Schertz, A. M. Casa, P. Brown et al., 2006 Alignment of genetic maps and QTLs between inter- and intraspecific sorghum populations. Theor. Appl. Genet. 112: 1295–1305.
Feltus, F. A., J. Wan, S. R. Schulze, J. C. Estill, N. Jiang et al., 2004 An SNP resource for rice genetics and breeding based on subspecies indica and japonica genome alignments. Genome Res. 14: 1812–1819. https://doi.org/10.1101/gr.4279404
Harlan, J. R., and J. M. J. Dewet, 1972 Simplified Classification of Cultivated Sorghum. Crop Sci. 12: 172. https://doi.org/10.2135/cropsic1972.00111833001200002005x
Hart, G. E., K. F. Schertz, Y. Peng, and N. H. Syed, 2001 Genetic mapping of Sorghum bicolor (L.) Moench QTLs that control variation in tillering...
rhizomatousness in chromosome SBI-01 and analysis of overwintering in a Sorghum bicolor x S. propinquum population. Mol. Breed. 31: 153–162. https://doi.org/10.1007/s11032-012-9778-8

Yamaguchi, M., H. Fujimoto, K. Hirano, S. Araki-Nakamura, K. Ohmae-Shinohara et al., 2016 Sorghum Dw1, an agronomically important gene for lodging resistance, encodes a novel protein involved in cell proliferation. Sci. Rep. 6: 28366. https://doi.org/10.1038/srep28366

Yang, S. S., B. D. Weers, D. T. Morishige, and J. E. Mullet, 2014 CONSTANS is a photoperiod regulated activator of flowering in sorghum. BMC Plant Biol. 14: 148. https://doi.org/10.1186/s12870-015-0477-6

Zhang, D., H. Guo, C. Kim, T. H. Lee, J. P. Li et al., 2013 CGRql, a Comparative Quantitative Trait Locus Database for Saccharinae Grasses. Plant Physiol. 161: 594–599. https://doi.org/10.1104/pp.112.206870

Zhang, D., W. Kong, J. Robertson, V. H. Goff, E. Epps et al., 2015 Genetic analysis of inflorescence and plant height components in sorghum (Panicoideae) and comparative genetics with rice (Oryzoidae). BMC Plant Biol. 15: 107. https://doi.org/10.1186/s12870-015-0477-6

Communicating editor: J. Ma