

Supermodels: sorghum and maize provide mutual insight into the genetics of flowering time

E. S. Mace · C. H. Hunt · D. R. Jordan

Received: 7 June 2012 / Accepted: 8 February 2013 / Published online: 5 March 2013
© Her Majesty the Queen in Right of Australia 2013

Abstract Nested association mapping (NAM) offers power to dissect complex, quantitative traits. This study made use of a recently developed sorghum backcross (BC)-NAM population to dissect the genetic architecture of flowering time in sorghum; to compare the QTL identified with other genomic regions identified in previous sorghum and maize flowering time studies and to highlight the implications of our findings for plant breeding. A subset of the sorghum BC-NAM population consisting of over 1,300 individuals from 24 families was evaluated for flowering time across multiple environments. Two QTL analysis methodologies were used to identify 40 QTLs with predominately small, additive effects on flowering time; 24 of these co-located with previously identified QTL for flowering time in sorghum and 16 were novel in sorghum. Significant synteny was also detected with the QTL for flowering time detected in a comparable NAM resource recently developed for maize (*Zea mays*) by Buckler et al. (Science 325:714–718, 2009). The use of the sorghum BC-NAM population allowed us to catalogue allelic variants

at a maximal number of QTL and understand their contribution to the flowering time phenotype and distribution across diverse germplasm. The successful demonstration of the power of the sorghum BC-NAM population is exemplified not only by correspondence of QTL previously identified in sorghum, but also by correspondence of QTL in different taxa, specifically maize in this case. The unification across taxa of the candidate genes influencing complex traits, such as flowering time can further facilitate the detailed dissection of the genetic control and causal genes.

Introduction

Flowering time in the grasses is a complex trait that controls adaptation of plants to their environment by tailoring vegetative and reproductive growth phases to local climatic effects. In determinant plants, such as cereal crops, flowering needs to occur when conditions for pollination and seed development are optimal and consequently most plants restrict flowering to a specific time of the year. This control is commonly achieved by using two environmental cues; day length (photoperiod) and temperature, in addition to endogenous cues, primarily related to development. The response of the plants to these different cues is controlled by four interacting pathways: photoperiod, temperature, autonomous and plant hormonal pathways (e.g. Mouradov et al. 2002; Amasino and Michaels 2010). The photoperiod pathway gives broad adaptation to a particular climatic region based on the day length cues either causing flowering to occur faster as day length declines (short day plants) or as day length increases in long day plants. The temperature pathway controls the fine scale progression to flowering, reacting to local changes in temperature that may occur in one season and not the next. In contrast, the

Communicated by R. Snowdon.

Electronic supplementary material The online version of this article (doi:10.1007/s00122-013-2059-z) contains supplementary material, which is available to authorized users.

E. S. Mace (✉) · C. H. Hunt
Department of Agriculture, Forestry and Fisheries,
Hermitage Research Station, 604 Yangan Road,
Warwick, QLD 4370, Australia
e-mail: Emma.Mace@daff.qld.gov.au

D. R. Jordan
Queensland Alliance for Agriculture and Food Innovation,
The University of Queensland, St. Lucia, QLD 4072, Australia
e-mail: david.jordan@uq.edu.au

autonomous pathway promotes flowering indirectly by facilitating responsiveness to cues that actively promote flowering (Simpson 2004), and the hormonal pathway controls the timing of flowering in relation to reproductive development. Farmers have exploited variation in flowering time to expand the areas where particular crops are grown, well beyond their natural range, and to fine-tune adaptation of particular varieties to specific environmental and management conditions.

Sorghum is a short day C_4 grass native to Africa, grown as a crop for its grain, forage and biomass. Under domestication, its range as a cereal crop extends from the equator up to approximately 50° of latitude and is grown at elevations up to 2,500 m. To achieve such wide adaptation, flowering time in cultivated grain sorghum varieties ranges from as little as 45 days to more than 120 days. To create varieties that can be grown effectively in temperate environments, breeders in developed countries have systematically eliminated photoperiodism from elite grain sorghum germplasm (Chanterreau et al. 2001).

Because of its quantitative nature, economic importance and its ease of scoring, the genetic basis of flowering time has been extensively studied in sorghum. Classical genetic studies have identified six major maturity loci Ma_1 , Ma_2 , Ma_3 and Ma_4 (Quinby 1967) and Ma_5 and Ma_6 (Rooney and Aydin 1999), the latter two interact epistatically to increase photoperiod sensitivity and lengthen the duration of the vegetative phase. Sorghum is grown as a hybrid crop in the developed world and as a result the degree of dominance and heterosis for flowering time is of practical importance. If photoperiod effects are excluded studies indicate that the genetic control of flowering time in sorghum is largely additive, although a degree of dominance is also observed leading to hybrid vigour for early flowering with estimates of mid-parent heterosis varying from 1 to 6 % for (Kirby and Atkins 1968; Liang and Walter 1968; Wenzel 1988).

The underlying genes and molecular mechanisms controlling flowering have been more extensively studied in the model dicot *Arabidopsis* (e.g. Greenup et al. 2009; Imaizumi 2010), with over 80 genes in the four parallel gene pathways described and characterised (e.g. Flowers et al. 2009; Higgins et al. 2010; Putterill et al. 2004). Comparative genomics studies have identified conserved homologues of the genes in these pathways in other cereals, including rice, *Brachypodium* and barley and in particular, aspects of the photoperiod and autonomous pathways are well conserved across species (e.g. Higgins et al. 2010; Lagercrantz 2009; Colasanti and Coneva 2009; Cockram et al. 2007; Jung and Muller 2009).

Although the number of genes present in the flowering time pathway of the *Arabidopsis* model provides a likely upper limit to the number of flowering genes present in

sorghum, it is likely that the number of genes that have allelic series and can therefore be detected as QTL is lower. Theoretically, only a relatively small number of genetic loci with allelic series would be required to generate the range of phenotypic variation in flowering observed in cultivated sorghum. The actual number of loci and their allelic diversity will be a product of the diversity within the wild population and selection pressures that occurred during domestication. Potentially, these pressures may have reduced variation at some loci, alternatively the maintenance of variation at many loci may have been favoured because of the flexibility it provide farmers to fine tune their varieties while allowing selection for QTL for other traits linked to flowering time loci.

Because the advent of molecular marker technologies over two decades ago, 14 QTL and association mapping studies for flowering time in sorghum have been undertaken (Brown et al. 2006; Chanterreau et al. 2001; Crasta et al. 1999; Feltus et al. 2006; Hart et al. 2001; Kebede et al. 2001; Kim 2003; Lin et al. 1995; Mannai et al. 2011; Murray et al. 2008; Parh 2005; Ritter et al. 2008; Shirin-gani et al. 2010; Srinivas et al. 2009). To date, the majority of sorghum QTL studies have been based on the crosses between parental lines selected for their diverse phenotypes (often interspecific crosses). To date, these 14 studies have identified 78 QTL for flowering time, representing approximately 27 unique genomic regions, with an average of 5.8 QTL detected per study. In addition, a recent positional cloning study (Murphy et al. 2011) identified the causal gene at the Ma_1 locus in sorghum. However, most of what we know to-date about the genetic architecture of flowering time in sorghum is based on the traditional QTL linkage mapping in bi-parental populations, and while such bi-parental crosses provide valuable insights, the current view of the genetic architecture of flowering time in sorghum, particularly as it pertains to cultivated sorghum, is very limited. A more global investigation of the genetic control of this complex trait requires resources that enable multiple alleles to be detected simultaneously. A large sorghum backcross nested association mapping (BC-NAM) population has recently been developed (Jordan et al. 2011). The BC-NAM population captures a substantial proportion of the global genetic diversity of sorghum inbred lines as measured by DArT markers (Jordan et al. 2011) and provides the power to resolve complex, quantitative traits. The sorghum BC-NAM panel consists of a set of 4,000 BC-RILs from crosses between the elite line R931945-2-2 and 56 other diverse inbreds. A subset of this resource has been genotyped with DArT markers. The high allele diversity and large sample size provide power for detection and resolution of QTL. A comparable NAM resource has recently been developed in maize (Buckler et al. 2009; McMullen et al. 2009) consisting of 25

recombinant-inbred line (RIL) populations; based on 25 founder lines each crossed in a half-sib design to the common reference parent B73. The power of the maize NAM resource has been successfully demonstrated through the genetic dissection of a number of complex traits, including flowering time (Buckler et al. 2009), northern (Poland et al. 2011) and southern (Kump et al. 2011) leaf blight resistance, leaf architecture (Tian et al. 2011) and kernel composition (Cook et al. 2012).

The objective of the current study was to use the sorghum BC-NAM population in order to dissect the genetic architecture of flowering time in sorghum; to compare the QTL identified with other genomic regions identified in previous sorghum and maize flowering time studies and to highlight the implications of our findings for sorghum domestication and breeding.

Materials and methods

Germplasm

The study used a backcross nested association mapping population (BC-NAM) previously developed by the Department of Agriculture Fisheries and Forestry (DAFF), formerly the Queensland Department of Primary Industries (QDPI) (Jordan et al. 2011). The development of the population is described in detail by Jordan et al. (2011), and the 24 BC-NAM families used in the current study are detailed in Table 1. The 24 diverse lines used for population development for the current study included 16 lines from breeding programs around the world, 6 landraces converted to temperate adaptation by the sorghum conversion program (Stephens et al. 1967), one landrace and one weedy species. The sorghum BC-NAM used for this study is a reference design with each subpopulation being produced by crossing a single elite parent (R931945-2-2) with a diverse line and backcrossing the resulting F_1 to a genetic male sterile (ms_3) version of the elite parent to produce a large BC_1F_1 population. These seeds were grown in a single long row of well-spaced plants. Fifty to one hundred plants were selected from each row on the basis appropriate agronomic type. The selection focused predominantly on choosing plants with height and maturity within the acceptable range required for performance in Australian cropping systems. Each selected BC_1F_1 plant was self-pollinated and the resulting BC_1F_2 seed from each plant was grown in a 4-m row containing approximately 40 plants. One fertile plant was selected from each row and self-pollinated. Again selection was imposed for appropriate agronomic type which reduced genetic variance but did not generate non-normality in the phenotype within individual population and thus is still valid for analysis.

This process of head to row selection was repeated once more to produce BC_1F_4 seed. Selection for male fertility at each generation insured that most rows were homozygous for the fertile allele of the ms_3 . The BC_1F_4 seed from the between 30 and 90 lines from each population was grown in a single row and top cross hybrid seed was produced by taking bulk pollen from each BC_1F_4 row and crossing it to a cytoplasmic genetic male sterile line (B923296). The female parent was chosen to represent female parental lines typically used in Australia with good general combining ability for grain yield, sorghum midge resistance and stay green.

Phenotypic data and testing

Data in this paper were produced from trials grown in the Australian summer cropping seasons of 2004/05, 2005/06, 2006/07 and 2007/08. Hybrids were evaluated each year in trials planted at ten locations across the 4 years with 20 trials overall in Queensland and Northern NSW covering the region between 29 and 32 degrees of latitude (Table 2; ESM Table S1). Trials were planted between November and February each season. Days to flowering was defined as the time taken from planting for 50 % of the plants in a plot to commence flowering. There were 14, 4, 14 and 12 BC-NAM families in each of the four seasons, respectively, making an unbalanced set of data for an overall comparison of BC-NAM families.

Statistical analysis of breeding trials

The trials used partially replicated designs (Cullis et al. 2006) with the number of genotypes ranging from 674 to 1,197 across the locations, with 30 % of the genotypes replicated two or more times and single plots of the remaining 70 %. Each individual trial had a different design and each design attempted to diminish the possibility of spatial error effects occurring within each trial. Concurrence of genotypes and populations between years allows groups of trials from multiple years to be analysed as a single multi-environment trial (MET). A MET containing the 20 trials was analysed by fitting a linear mixed model using the package ASReml (Butler et al. 2009) and the R statistical software (R Development Core Team 2012). The model consists of a fixed effect for flowering at each trial, random effects for trial by genotype and the spatial error for each individual trial using procedures, such as those discussed in Smith et al. (2001). The genotype by environment (GxE) interaction was examined by fitting a second-order factor analytic structure to the trial by genotype interaction. The MET included additional genotypes that were in the trials, but not included

Table 1 Description of the characteristics of the diverse parental lines used to develop the 24 BC-NAM families included in current study

| NRP | Country of origin | Type | No. individuals per family | No. trials |
|------------------------|-------------------|--------------------|----------------------------|------------|
| <i>S. arundinaceum</i> | Australia | Weedy species | 96 | 8 |
| SC23 | Ethiopia | Converted landrace | 50 | 5 |
| Karper669 | USA | Breeding line | 61 | 14 |
| KS115 | USA | Breeding line | 57 | 4 |
| Kuyuma | Zambia | Breeding line | 50 | 4 |
| TX2903 | USA | Breeding line | 58 | 7 |
| SDS1948-3 | Zimbabwe | Breeding line | 49 | 4 |
| LR2931-2 | China | Breeding line | 66 | 13 |
| SC56-14E | Sudan | Converted landrace | 65 | 11 |
| ICSV400 | Mali | Breeding line | 56 | 4 |
| SC62C | Sudan | Converted landrace | 54 | 13 |
| ICSV745 | India | Breeding line | 58 | 10 |
| SC35C | Ethiopia | Converted landrace | 59 | 11 |
| SC108C | Ethiopia | Converted landrace | 47 | 4 |
| LR2490-3 | China | Breeding line | 62 | 10 |
| Sureno | Honduras | Breeding line | 52 | 4 |
| Malisor84-7 | Mali | Breeding line | 64 | 11 |
| Rio | USA | Breeding line | 53 | 11 |
| SC326-6 | Ethiopia | Converted landrace | 49 | 11 |
| QL12 | Australia | Breeding line | 56 | 11 |
| Ai4 | China | Breeding line | 85 | 11 |
| Dorado | El Salvador | Breeding line | 46 | 11 |
| Macia | South Africa | Breeding line | 48 | 11 |
| IS3614-3 | Nigeria | Landrace | 48 | 6 |

for this study. Although some trials included only a very small number of BC-NAM families, for example trials from 2006, all the data from these trials was included to establish a more extensive analysis of the relationship between each trial. The analysis resulted in a genetic variance for each trial along with a set of loadings that represent factor analytic loadings. The flowering times for the trials showed high levels of correlation across sites thus ranking the genotypes similarly. As a result, a single estimate of flowering time was produced for each genotype using the MET analysis.

Generalised repeatability estimates on a hybrid mean basis were calculated for flowering time using the method proposed by Cullis et al. (2006) (Table 2). This method was used because the standard definitions of heritability assume balanced data, where trials are analysed using models with independent random effects for blocks and plots and are not applicable to the partially replicated designs used in this study. For example, the standard definitions of heritability do not allow for the use of spatial models that were used for these analyses. Such models, by their nature, imply complex variance covariance structures pertaining to experimental units.

Markers

Five plants of each of the 1,389 individuals from the 24 BC-NAM families (BC_1F_4 generation) were bulked for DNA extraction and screened with 1,005 DArT markers, following the methodology described in Mace et al. (2008) and Jordan et al. (2011). A subset of 932 high quality markers, with minimal missing data across all lines, were selected for subsequent analyses, spanning 95 % of the consensus genetic linkage map distance (Mace et al. 2009a), averaging 0.6 markers/1 cM (Table 3). The number of polymorphic markers per BC-NAM family ranged from 192 to 489 (ESM Table S2).

QTL analyses

Owing to the small size of the individual populations and the selection imposed for maturity and height genes during their development, the standard mapping approaches typically applied to bi-parental crosses were not appropriate. Therefore, we used two different approaches, one novel and one already established, and compared the results of the two methodologies.

Table 2 Description of the phenotypic field trial data, including the total number of genotypes per trial, the number of BC-NAM families per trial, the number of BC-NAM individuals per trial, the mean days to flowering (DTF) per trial, the genetic variance, heritability and loading per trial site

| Closest town | Sowing date | Total # genotypes | # BC-NAM families | # BC-NAM individuals | Mean DTF | Genetic Variance | Repeatability | Loading |
|------------------|-------------------|-------------------|-------------------|----------------------|----------|------------------|---------------|---------|
| Biloela | 2 February 2005 | 1,136 | 14 | 756 | 52.02 | 0.338 | 75.7 | 0.723 |
| Biloela | 26 January 2006 | 1,059 | 4 | 186 | 51.75 | 1.146 | 84.9 | 0.88 |
| Biloela | 19 January 2007 | 996 | 14 | 663 | 46.99 | 0.824 | 77.1 | 0.767 |
| Biloela | 22 January 2008 | 1,010 | 10 | 759 | 54.74 | 0.576 | 78 | 0.775 |
| Cecil Plains | 13 December 2005 | 953 | 3 | 180 | 59.19 | 3.6 | 84.5 | 0.876 |
| Clermont | 13 February 2007 | 768 | 14 | 504 | 59.02 | 0.813 | 77.5 | 0.782 |
| Dalby | 21 September 2007 | 903 | 10 | 681 | 77.42 | 3.475 | 89.1 | 0.966 |
| Dalby | 25 November 2004 | 1,164 | 14 | 763 | 59.57 | 4.219 | 86.4 | 0.903 |
| Dalby | 18 October 2007 | 947 | 10 | 715 | 62.44 | 2.051 | 86.4 | 0.92 |
| Dalby | 16 November 2004 | 674 | 7 | 407 | 57.04 | 2.947 | 85.4 | 0.912 |
| Dysart | 21 January 2005 | 1,069 | 14 | 715 | 56.31 | 1.783 | 79 | 0.774 |
| Dysart | 25 February 2007 | 938 | 14 | 617 | 56.95 | 1.32 | 77.7 | 0.77 |
| Warwick | 5 November 2004 | 1,197 | 14 | 780 | 67.93 | 1.605 | 78 | 0.706 |
| Warwick | 29 November 2005 | 1,075 | 3 | 186 | 56.36 | 2.177 | 83.1 | 0.835 |
| Warwick | 21 November 2006 | 1,009 | 14 | 680 | 64.24 | 1.748 | 79.4 | 0.777 |
| Warwick | 12 November 2007 | 1,012 | 10 | 759 | 68.86 | 1.588 | 85.8 | 0.905 |
| Jimbour | 15 November 2004 | 682 | 7 | 412 | 62.4 | 3.633 | 84.1 | 0.884 |
| Kilcummin | 4 February 2008 | 899 | 10 | 680 | 66.58 | 0.289 | 78 | 0.792 |
| Liverpool Plains | 14 November 2007 | 980 | 10 | 742 | 67.12 | 2.441 | 86.2 | 0.912 |
| Srpingsure | 8 January 2008 | 753 | 12 | 569 | 62.45 | 0.939 | 83 | 0.885 |

Multi-populational genomic regional QTL analysis (mpQTL)

The novel mpQTL analysis method calculated an integrated probability statistic to highlight regions of the genome that have a key contribution to the variation of flowering time in the BC-NAM families studied, through a two-step process. The first step involved single marker analysis across the individual BC-NAM families. As all the BC-NAM families were not present in all trial years (ESM

Table S1) and markers were also unbalanced across BC-NAM families, with polymorphism varying across the BC-NAM non-recurrent parental (NRP) lines (ESM Table S2). Therefore, an individual subset of polymorphic markers was selected for each BC-NAM family. A linear mixed model was then conducted for each individual marker within each BC-NAM family. For a more extensive description of this model, see ESM file S1. Each marker was used as a fixed effect and random effects included trial by genotype, spatial effects within each trial and error

Table 3 Details of DArT markers used genome-wide in BC-NAM data analysis

| LG | # markers | Length (cM) |
|--------|-----------|-------------|
| SBI-01 | 99 | 181.7 |
| SBI-02 | 135 | 223.32 |
| SBI-03 | 91 | 153.7 |
| SBI-04 | 98 | 166.7 |
| SBI-05 | 111 | 118.5 |
| SBI-06 | 68 | 164.1 |
| SBI-07 | 72 | 132.8 |
| SBI-08 | 128 | 131.9 |
| SBI-09 | 62 | 135.6 |
| SBI-10 | 68 | 112.2 |
| Totals | 932 | 1,520.52 |

effects within each trial. For each marker, the model estimates a statistic that can be translated into a P value that tests the hypothesis that the two marker alleles have equal flowering effect within the specific BC-NAM family. At the completion of these analyses, a table of P values for marker by BC-NAM family was generated.

To combine the results of the single marker analysis across BC-NAM families, the most significant marker was selected from a sliding window of length 5 cM and a step of 1 cM. This stepwise process was performed on each BC-NAM family separately to generate a series of probability values spaced 1 cM apart along each chromosome. Fisher's combined probability test (Fisher 1932) was then applied to combine the results of the single marker regression across all the BC-NAM families to generate a single probability value representing the approximate presence of significance at each 1 cM point along each LG. In order to overcome the imbalance across the BC-NAM families at each marker location, a false discovery rate adjustment was made to the Fisher P values to allow a consistent 0.01 % significance value ($-\log$ Fisher P value of >3) to declare QTL significance.

Association mapping (AM)

To assist with the confirmation of our results, a secondary, more standard analysis was performed. A mixed model marker–trait association analysis was performed using the statistical package GenStat (VSN International 2011). Population structure was accounted for as a fixed effect using the subpopulation grouping option (equating to each BC-NAM family). In the majority of cases, there was little or no pedigree information for the exotic parents and hence our parentage structure only has a single generation which is equivalent to the sub-population grouping. The model for marker–trait association included an intercept term, the

effects associated with each BC-NAM family, random effects of genotype and the fixed effects of the tested markers. The Wald statistical test was then used for each marker, individually, to test the null hypothesis that the marker's effect was zero. A 1 % false discovery rate ($-\log_{10} P = 2$) was used to declare QTL significance.

Projecting maize QTL onto the sorghum consensus map

A total of 75 significant markers were identified by Buckler et al. (2009) associated with flowering time in maize (36 significant marker/trait associations for days to anthesis (DA; male flowering) and 39 significant marker/trait associations for days to silking (DS; female flowering)), representing a total of 46 unique genomic locations. The significant markers identified were sequence mapped with an in silico mapping strategy analogous to e-PCR (Schuler 1998) using BLAST similarity (Altschul et al. 1990) between the maize marker and the sorghum whole genome sequence (WGS) sequence and conducted via Phytozome (<http://www.phytozome.net>), requiring hits with $E \leq 1e - 10$ based on BLASTn. The predicted cM locations of the identified physical locations of the maize markers on the sorghum genome were calculated using a framework map based on the 504 sequenced markers on the consensus map, averaging 1 marker/3.1 cM or 1 marker/1.3 Mbp, as detailed in Mace and Jordan (2011).

Results

Phenotypic assessment of flowering time in BC-NAM families

A summary of the results from the MET analysis is presented in Table 2. There was a difference of 30 days between the earliest flowering trial (Bil2007; 46.9 days) and the latest flowering trial (Dal2008; 77.4 days). The individual trial repeatabilities were all greater than 75 % with half of the trials having a repeatability greater than 80 % indicating that in general the analysis explained the majority of the genetic variation within each trial. The total variance explained from the fit of the second-order factor analytic model was 81 %, this indicated that we can explain 81 % of the total variation in these data with only two factor components; the first of these factors explained 76 % and have an average trial loading of 0.84 and an average between trial correlation of 0.72. Trial loadings were all higher than 0.7; these values indicate that in general all sites ranked the genotypes in a similar way, indicating minimal genotype by environment interaction, and an across site average for each genotype would be

representative of the performance of these genotypes and could be used for further study.

Figure 1 shows a boxplot representation of the mean and variance of the predicted average across site flowering times for each BC-NAM family. The vertical line at 60.73 represents the average flowering time for the recurrent parent genotype R931945-2-2. All families except for SC326-6 have an average flowering time less than the recurrent parent.

QTL results

QTLs were mapped across the 1,389 individuals comprising the 24 BC-NAM families using the AM and mpQTL approaches. 25 QTL were identified by the AM and 34 QTL by the mpQTL method, with 19 out of a total of 40 unique QTL being in common (Table 4; Fig. 2). In a given region of the genome, the integrated probability statistic, calculated through the mpQTL analysis method, increased with the number of BC-NAM families for which the region displayed significant QTL effects in addition to the proximity of the QTL positions estimated in the different families and the level of significance in different families. When the stringency of the thresholds were relaxed slightly ($-\log_{10} P > 1.5$ for the AM and $-\log$ Fisher $P > 2$ for mpQTL analysis) a further 12 QTL were found to be in common; totalling 31 out of the 40 unique QTL in common across the two methodologies.

All 40 QTL are presented in this study. The majority of flowering time QTL was shared among multiple families,

with 35 out of the total of 40 QTL showing significant effects among 3 to 12 families, with each QTL on average being detected in 5.6 families (Table 4). The 40 QTL identified were distributed across all chromosomes; the highest number of QTL per chromosome was 6 (for SBI-03 and SBI-04), while the lowest was 2 (SBI-10). At least one QTL was located in the heterochromatic region of each chromosome, with the exception of SBI-02. The average length of the support intervals for QTL positions was 4.1 cM, as determined through the extent of the QTL region greater than the defined statistical thresholds in the mpQTL method (ESM Table S3).

Relative to the recurrent parent, the absolute values of significant allelic effects ranged from -1.46 to 1.78 days, averaging -0.70 days across significant alleles contributing to earliness and averaging 0.78 days across significant alleles contributing to lateness (Table 4; Fig. 3). At 75 % of the QTLs, the NRPs contributed both positive (later) and negative (earlier) allele effects relative to R931945-2-2, suggesting that allelic series of at least three alleles each exist at these QTLs (ESM Figure S1). Figure 4 details an example of a putative allelic series for one of the QTL on SBI-10, QDTF_NAM_10_61, which has allele effects ranging from -1.45 days (contributed by SC56-14E) up to 0.76 days (contributed by parent SC23). At 6 QTL, the diverse, NRP lines only contributed negative (earlier) allele effects and at 4 QTL, only positive (later) allele effects were detected (Table 4). The total range of QTL allele effect sizes per QTL, across all BC-NAM families segregating at each loci, ranged from 0.61 days up to 3.08 days (Table 4), with an average range of allele effects per QTL of 1.72 days.

A common consequence of using small populations such as the BC-NAM families for QTL analysis is the tendency for QTL effects to be overestimated (Beavis 1994). To determine the impact of population size on allele effects estimated, we compared the effect sizes determined in the *S. arundinaceum* BC-NAM subset family (96 individuals) with those determined in a larger (214 individuals) population derived from the same cross. The allele effect sizes between the two populations were similar, having an average difference of just 0.17 days, and with no apparent bias towards larger estimates of effect sizes in the small population (ESM Figure S4). We believe this is in part due to the high heritability of the trait simplifying effects of the backcross structure on the composition of individual genotypes within each NAM family.

In terms of allele occurrence, the QTL identified solely by the AM methodology were rarer than those identified by the mpQTL method. Using the AM methodology solely, there were two populations, on average, with a significant marker ($p < 0.05$), underlying the QTL regions, compared to 5.3 for the mpQTL methodology solely, and compared

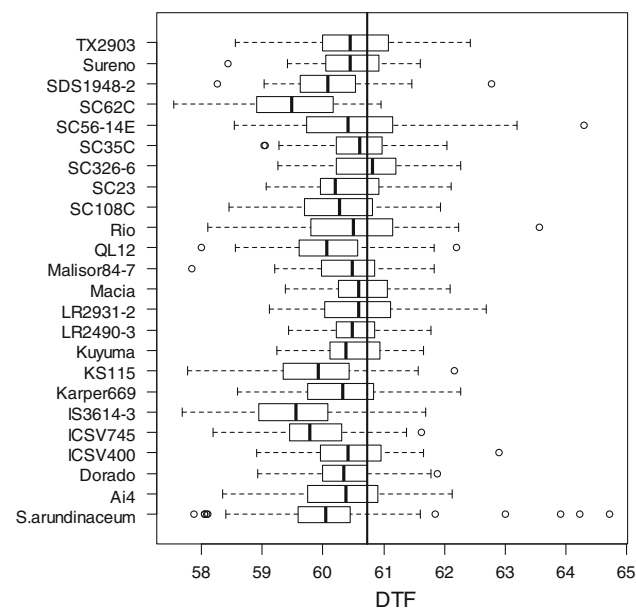


Fig. 1 Box-plots of BLUPs of days to flower (DTF) per BC-NAM family based on data from 20 locations, the vertical line representing the mean of the recurrent parent, R931945-2-2

Table 4 Summary of QTL identified in BC-NAM families, detailing location, probability values (significant *P* values in bold), total range of significant effect size in days to flowering (DTF)

| QTL ID | LG | QTL mid-point (cM) | -log Fisher <i>P</i> | -log ₁₀ <i>P</i> | # sig pops | Early effects (in DTF) | | | Late effects (in DTF) | | | Total range |
|----------------|----|--------------------|----------------------|-----------------------------|------------|------------------------|-------|---------|-----------------------|------|---------|-------------|
| | | | | | | Min | Max | Average | Min | Max | Average | |
| QDTF_NAM_1_10 | 1 | 10 | 4.13 | 0.85 | 5 | -0.78 | -0.48 | -0.61 | 0.89 | 1.12 | 1.01 | 1.90 |
| QDTF_NAM_1_53 | 1 | 53 | 7.19 | 2.95 | 9 | -0.89 | -0.51 | -0.71 | 0.68 | 0.73 | 0.71 | 1.62 |
| QDTF_NAM_1_66 | 1 | 66 | 7.96 | 4.36 | 7 | -1.20 | -0.45 | -0.75 | 0.67 | 0.67 | 0.67 | 1.87 |
| QDTF_NAM_1_112 | 1 | 112 | 4.73 | 1.82 | 6 | -0.80 | -0.43 | -0.61 | 0.55 | 1.78 | 1.04 | 2.58 |
| QDTF_NAM_1_182 | 1 | 182 | 2.27 | 2.20 | 2 | -0.35 | -0.03 | -0.18 | 0.04 | 0.97 | 0.40 | 1.32 |
| QDTF_NAM_2_23 | 2 | 23 | 4.15 | 1.37 | 4 | 0.00 | 0.00 | 0.00 | 0.45 | 1.71 | 0.93 | 1.71 |
| QDTF_NAM_2_123 | 2 | 123 | 5.51 | 0.39 | 6 | -1.15 | -0.60 | -0.84 | 0.44 | 0.74 | 0.60 | 1.89 |
| QDTF_NAM_2_147 | 2 | 147 | 6.86 | 3.70 | 8 | -1.36 | -0.53 | -0.74 | 0.00 | 0.00 | 0.00 | 1.36 |
| QDTF_NAM_3_21 | 3 | 21 | 3.15 | 1.84 | 5 | -0.98 | -0.40 | -0.65 | 0.37 | 0.46 | 0.41 | 1.44 |
| QDTF_NAM_3_34 | 3 | 34 | 5.56 | 2.60 | 5 | -1.00 | -0.46 | -0.61 | 0.60 | 0.60 | 0.60 | 1.60 |
| QDTF_NAM_3_47 | 3 | 47 | 4.13 | 3.18 | 4 | -0.99 | -0.49 | -0.72 | 0.00 | 0.00 | 0.00 | 0.99 |
| QDTF_NAM_3_57 | 3 | 57 | 6.82 | 2.05 | 6 | -0.80 | -0.56 | -0.68 | 0.58 | 1.36 | 0.87 | 2.16 |
| QDTF_NAM_3_68 | 3 | 68 | 12.00 | 1.11 | 8 | -1.36 | -0.40 | -0.68 | 0.45 | 1.04 | 0.74 | 2.40 |
| QDTF_NAM_3_133 | 3 | 133 | 3.69 | 2.67 | 5 | -1.18 | -0.89 | -0.99 | 0.38 | 0.45 | 0.41 | 1.63 |
| QDTF_NAM_4_15 | 4 | 15 | 5.20 | 3.53 | 7 | -0.98 | -0.46 | -0.67 | 0.60 | 0.60 | 0.60 | 1.59 |
| QDTF_NAM_4_43 | 4 | 43 | 2.30 | 3.17 | 2 | -0.95 | -0.08 | -0.37 | 0.41 | 0.61 | 0.51 | 1.56 |
| QDTF_NAM_4_58 | 4 | 58 | 3.68 | 3.75 | 3 | -0.88 | -0.54 | -0.75 | 0.00 | 0.00 | 0.00 | 0.88 |
| QDTF_NAM_4_74 | 4 | 74 | 6.85 | 5.95 | 9 | -1.26 | -0.29 | -0.63 | 0.00 | 0.00 | 0.00 | 1.26 |
| QDTF_NAM_4_87 | 4 | 87 | 6.39 | 4.83 | 7 | -1.33 | -0.34 | -0.68 | 0.00 | 0.00 | 0.00 | 1.33 |
| QDTF_NAM_4_98 | 4 | 98 | 10.15 | 4.51 | 9 | -1.30 | -0.41 | -0.78 | 0.32 | 1.27 | 0.68 | 2.57 |
| QDTF_NAM_5_4 | 5 | 5 | 2.02 | 4.55 | 3 | -0.53 | -0.04 | -0.23 | 0.03 | 1.24 | 0.36 | 1.76 |
| QDTF_NAM_5_10 | 5 | 10 | 4.56 | 3.02 | 4 | -0.96 | -0.70 | -0.79 | 1.09 | 1.09 | 1.09 | 2.05 |
| QDTF_NAM_5_27 | 5 | 27 | 7.86 | 2.67 | 7 | -0.98 | -0.57 | -0.86 | 1.01 | 1.29 | 1.15 | 2.28 |
| QDTF_NAM_5_67 | 5 | 67 | 7.80 | 2.97 | 12 | -1.16 | -0.43 | -0.65 | 0.51 | 1.36 | 0.93 | 2.52 |
| QDTF_NAM_6_10 | 6 | 10 | 12.00 | 4.44 | 8 | 0.00 | 0.00 | 0.00 | 0.38 | 1.72 | 0.76 | 1.72 |
| QDTF_NAM_6_23 | 6 | 23 | 1.82 | 2.27 | 1 | 0.00 | 0.00 | 0.00 | 0.34 | 1.04 | 0.69 | 1.04 |
| QDTF_NAM_6_32 | 6 | 32 | 3.17 | 1.93 | 3 | -0.61 | -0.40 | -0.53 | 0.00 | 0.00 | 0.00 | 0.61 |
| QDTF_NAM_6_43 | 6 | 43 | 3.78 | 1.71 | 3 | -0.67 | -0.65 | -0.66 | 1.14 | 1.14 | 1.14 | 1.81 |
| QDTF_NAM_6_83 | 6 | 83 | 2.34 | 2.65 | 2 | -0.53 | -0.01 | -0.21 | 0.08 | 1.05 | 0.42 | 1.58 |
| QDTF_NAM_7_5 | 7 | 5 | 3.23 | 1.53 | 4 | -0.86 | -0.47 | -0.62 | 0.46 | 0.46 | 0.46 | 1.32 |
| QDTF_NAM_7_59 | 7 | 59 | 3.34 | 1.34 | 5 | -0.72 | -0.44 | -0.54 | 0.72 | 0.72 | 0.72 | 1.44 |
| QDTF_NAM_7_73 | 7 | 73 | 6.49 | 1.74 | 6 | -0.49 | -0.46 | -0.47 | 0.59 | 0.99 | 0.71 | 1.48 |
| QDTF_NAM_8_30 | 8 | 30 | 6.24 | 1.07 | 7 | -0.73 | -0.38 | -0.49 | 0.71 | 0.87 | 0.78 | 1.60 |
| QDTF_NAM_8_66 | 8 | 66 | 9.38 | 1.99 | 12 | -1.44 | -0.54 | -0.76 | 0.54 | 1.65 | 0.82 | 3.08 |
| QDTF_NAM_8_115 | 8 | 115 | 6.05 | 2.85 | 6 | -0.99 | -0.42 | -0.69 | 0.53 | 1.38 | 0.96 | 2.37 |
| QDTF_NAM_9_25 | 9 | 25 | 2.15 | 3.09 | 2 | -1.24 | -0.01 | -0.26 | 0.09 | 0.46 | 0.20 | 1.70 |
| QDTF_NAM_9_75 | 9 | 75 | 5.27 | 0.51 | 5 | -0.74 | -0.54 | -0.64 | 0.45 | 0.55 | 0.50 | 1.29 |
| QDTF_NAM_9_136 | 9 | 136 | 4.51 | 8.56 | 6 | 0.00 | 0.00 | 0.00 | 0.50 | 1.12 | 0.72 | 1.12 |
| QDTF_NAM_10_44 | 10 | 44 | 3.09 | 1.18 | 5 | -1.03 | -0.47 | -0.72 | 0.95 | 0.95 | 0.95 | 1.98 |
| QDTF_NAM_10_61 | 10 | 61 | 9.57 | 4.00 | 9 | -1.46 | -0.45 | -0.82 | 0.55 | 0.77 | 0.69 | 2.22 |

to 6.1 for the QTL detected in both methodologies. The most extreme example of this was for QDTF_NAM_6_23 which was a very rare QTL allele, detected as significant ($p = 0.0165$) by the AM methodology in only one population (Rio). The AM methodology also detected alleles

with smaller effect size on average, compared to the mpQTL methodology (average total range of allelic effects of QTL detected solely by the AM methodology was 1.4 days compared to 1.7 days in QTL detected solely by the mpQTL methodology). The mpQTL methodology, due

Fig. 2 QTL analysis results for flowering time in the BC-NAM population across the genome. The vertical lines indicate the breaks between chromosomes. **a** QTL analysis results based on AM; **b** QTL analysis results based on mpQTL analysis

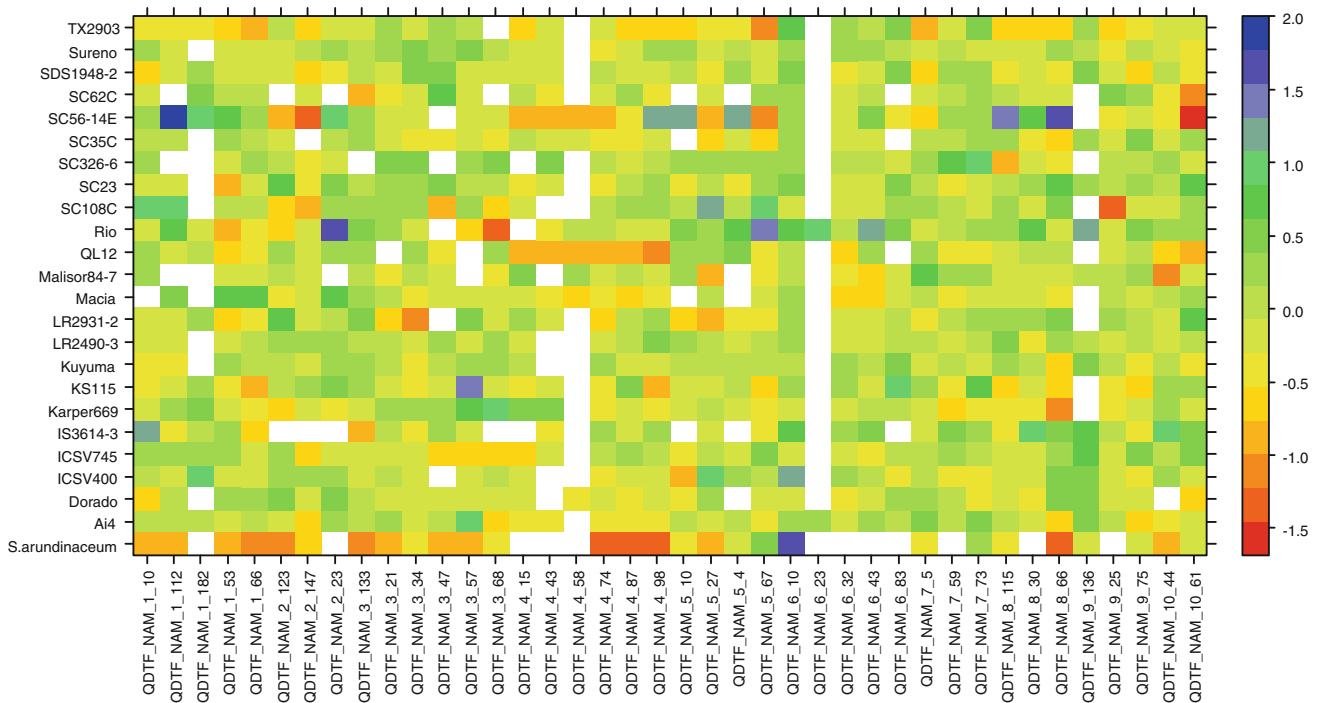
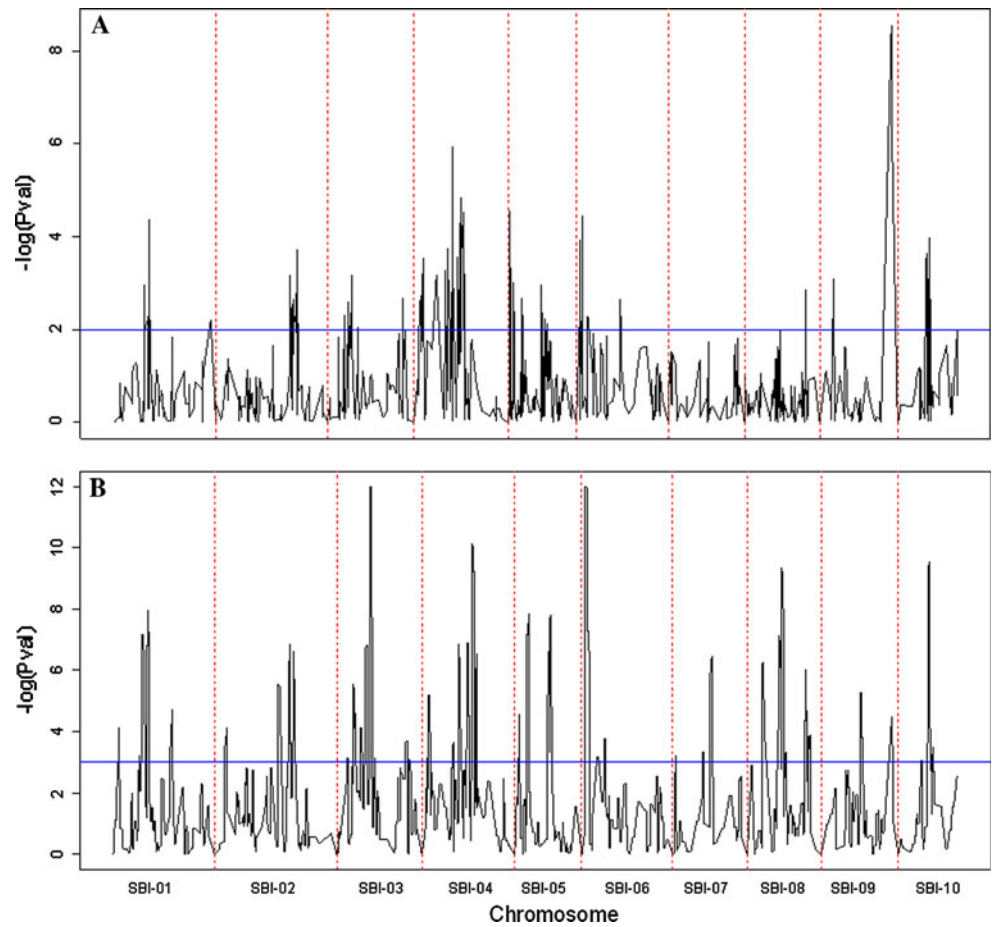


Fig. 3 Heat-map for flowering time QTL effects by chromosomal position and donor parent

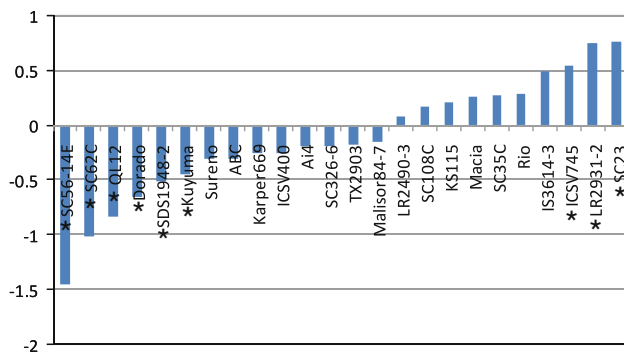


Fig. 4 Allelic series for QTL QDTF_NAM_10_61 across 24 BC-NAM families. Allele effect estimates are relative to R931945-2 allele days to flowering (*asterisk* indicates significant effect)

to the 5 cM sliding window analysis step which attempted to compensate for the disparity in common markers across population, could identify more populations with a significant marker in a given region than the AM approach, and as a consequence the confidence interval (CI) of the QTL identified solely by the mpQTL methodology were on average larger (3.6 cM) compared to the CI of QTL detected solely by the AM methodology (0.2 cM).

The number of QTL detected with significant allelic effects within each BC-NAM family ranged from 2 (SC62C) to 18 (SC56-14E and Tx2903) with an average of 9.45 QTL detected per BC-NAM family (Fig. 5). The total additive effect of all the significant effect QTL alleles per BC-NAM family, calculated by summing significant additive positive and negative QTL effects, ranged from 19.85 days (SC56-14E) to just 1.92 days (SC62C), with an average of 6.9 days summed across positive and negative QTL allele effects per BC-NAM family (Fig. 5). The BC-NAM family with the largest summed negative (early) effect was *S. arundinaceum*, with 16 significant early effect QTL with a total additive effect of -16.5 days. The *S. arundinaceum* population had only one significant positive (late) effect QTL with an estimated allele effect of 1.72 days. In contrast, the BC-NAM family with the largest summed positive (late) effect was SC56-14E, with 9 significant late QTL with a total additive effect of $+10.77$ days. This population also had nine significant early QTL with a total additive effect of -9.08 days. One BC-NAM family, Sureno, only had significant late effect QTL alleles (three significant QTL with a total additive effect of 1.29 days) and one BC-NAM family, SC62C, only had significant early effect QTL alleles (two significant QTL with a total additive effect of -1.92 days). Overall, the BC-NAM families contributed approximately twice as many significant negative (early) effect QTL (146 in total; average of 6.08 per BC-NAM family) compared to significant positive (late) effect QTL (81 in total; average of 3.37 per BC-NAM family).

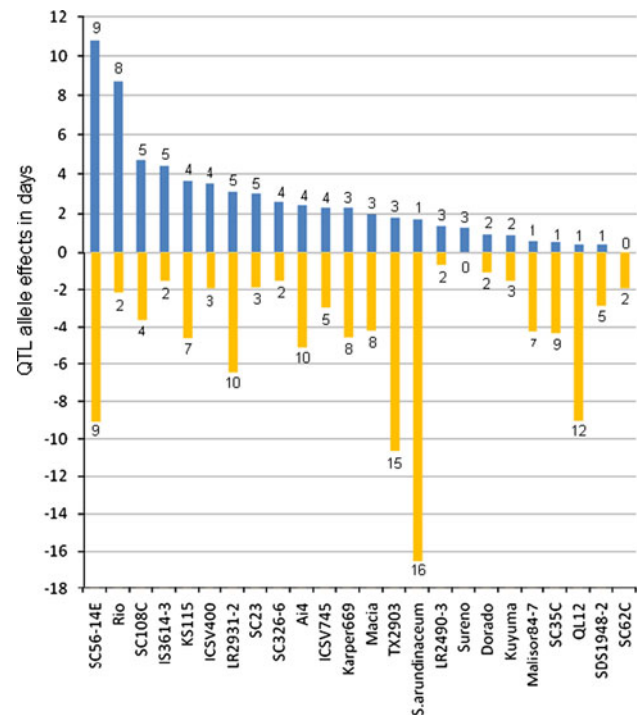


Fig. 5 Sums of the estimated positive (late) and negative (early) QTL allele effects for each BC-NAM family. Number of QTLs next to each bar

We compared flowering time QTL detected in the BC-NAM population to those reported in 13 previous sorghum studies (Table 5). The QTL identified in the 14th study (Murray et al. 2008) were unable to be projected onto the consensus map as the genetic linkage map in this study was constructed predominately with AFLP markers which are non-sequenced based and hence not readily comparable across studies. Nine unique sorghum bi-parental populations and one diversity research set were used across the 13 studies; from these a total of 70 QTL were projected onto the sorghum consensus map (Mace and Jordan 2011), representing 27 unique genomic locations (62 of the QTL representing 20 meta QTL, and 8 unique QTL). On average, 5.8 QTL were detected per study, significantly less than the number detected per BC-NAM family in the current study (average of 9.45 QTL/BC-NAM family). Of the 40 QTL detected in the BC-NAM, 22 co-located with previously identified sorghum QTL, based on the genetic distance of the QTL mid-points being less than 10 cM away from each other (ESM Fig. S2). A further 4 of the 40 detected QTL were within 8 cM or less of the CI of a previously identified sorghum QTL, 16 QTL detected in the BC-NAM populations were novel, and not previously identified in sorghum. In total, 22 of the 27 unique genomic regions previously identified to be associated with flowering time co-located with the BC-NAM QTL. Only five genomic regions containing eight previously identified

Table 5 Details of previously identified flowering time QTL from 13 other studies projected onto the consensus map

| Publication | # projected QTL | Population pedigree | Cross type | Generation | Population size |
|--------------------------|-----------------|------------------------------|-----------------------|----------------|-----------------|
| Brown et al. (2006) | 2 | BTx623/IS3620C | Cultivated/cultivated | RI | 137 |
| Chantreau et al. (2001) | 10 | IS2807/IS7680 | Cultivated/cultivated | RI | 85 |
| Crasta et al. (1999) | 2 | B35/Tx430 | Cultivated/cultivated | RI | 96 |
| Feltus et al. (2006) | 6 | BTx623/IS3620C | Cultivated/cultivated | RI | 137 |
| Feltus et al. (2006) | 2 | BTx623/ <i>S. propinquum</i> | Cultivated/wild | F ₂ | 370 |
| Hart et al. (2001) | 2 | BTx623/IS3620C | Cultivated/cultivated | RI | 137 |
| Kebede et al. (2001) | 2 | SC56/Tx7000 | Cultivated/cultivated | RI | 125 |
| Kim (2003) | 1 | BTx623/IS3620C | Cultivated/cultivated | RI | 137 |
| Lin et al. (1995) | 3 | BTx623/ <i>S. propinquum</i> | Cultivated/wild | F ₂ | 370 |
| Mannai et al. (2011) | 8 | Diversity research set | Cultivated/landraces | Fixed lines | 107 |
| Parh (2005) | 5 | R939145-2-2/IS8525 | Cultivated/cultivated | RI | 146 |
| Ritter et al. (2008) | 4 | R9188/R9403463-2-1 | Cultivated/cultivated | RI | 184 |
| Shiringani et al. (2010) | 5 | M71/SS79 | Cultivated/cultivated | RI | 188 |
| Srinivas et al. (2009) | 18 | 296B/IS18551 | Cultivated/cultivated | RI | 168 |

flowering time QTL have been reported in previous studies but not detected in the BC-NAM population.

We further evaluated the correspondence of the sorghum BC-NAM QTL with the maize NAM QTL recently described (Buckler et al. 2009). Using a multiple-family joint stepwise regression method, Buckler et al. (2009) identified 36 QTL for days to anthesis (DA; male flowering) and 39 QTL for days to silking (female flowering); the comparable traits for days to flower in sorghum, representing a total of 46 unique genomic locations across the 75 significant marker–trait associations. Of the 40 sorghum BC-NAM QTL for flowering time detected in this study, 37 (92.5 %) were less than 10 cM away from the projected location of at least one significant maize marker (based on the mid-point of the sorghum QTL); 30 (75 %) were less than 5 cM away from at least one significant maize marker and 15 (37.5 %) were less than 1 cM away from at least one significant maize marker (Table 6; ESM Fig. S3; ESM Table S3). Each of the 37 sorghum QTL corresponding closely to maize QTL co-located with, on average, 3.05 maize QTL, ranging from one (for seven sorghum QTL) to six (for one sorghum QTL). Of note, one sorghum QTL (QDTF_NAM_6_8) was less than 1 cM away from four syntenous maize QTL, while a further 13 sorghum QTL were less than 5 cM away from at least three syntenous maize QTL, corresponding to the known genomic synteny between the species. Figure 6 demonstrates the general finding of this study that the sorghum QTL correspond to multiple maize QTL detected in Buckler et al. (2009). Figure 6a details the correspondence of QTL detected on SBI-01 with maize flowering time QTL; four of the five sorghum BC-NAM DTF QTL co-locate within 10 cM of at least six maize flowering time QTL. Both Figs. 6a and b demonstrates the apparent conservation of QTL order

between the species however there were a number of examples of striking differences in the relative distances between the QTL. In some cases increased recombination per unit of physical distance in sorghum suggest that some maize QTL may be the result of multiple linked flowering time genes, as demonstrated by Fig. 6b.

We also predicted the genetic linkage map (cM locations) of the sorghum BC-NAM QTL on the maize NAM map (Buckler et al. 2009), based on the projection of the mid-points of the sorghum QTL (Fig. 7). These were calculated based on 838 sequenced markers on the consensus maize NAM map (available via <http://www.panzea.org>). The 40 sorghum QTL corresponded in total to 83 unique loci on the maize genome, with each QTL having at least two locations. When considering an average CI of 4 cM, 54 % (45 of the 83) of the sorghum QTL projected onto the maize genome co-located with one or more maize NAM QTL for flowering time. A χ^2 goodness of fit test indicated that the degree of correspondence between the sorghum and maize QTL was greater than would be expected by chance and highly significant ($p = 0.00142$).

Discussion

This study reports on the use of the sorghum BC-NAM population to identify QTL for flowering time, using two different QTL analysis methodologies. In total, 1,389 BC-NAM individuals, from 24 families, were evaluated for flowering time using data from 20 locations over 4 years. We identified 40 QTL for flowering time from the BC-NAM population and demonstrated a high degree of correspondence between the sorghum NAM QTL for flowering time and the maize NAM QTL for flowering time, in addition to

Table 6 Correspondence to previously identified QTL for flowering time in Sorghum (SbQTL) and *Zea mays* (Zm Chr1 -10) detailing the number of QTL previously identified where the distance between the mid-point of the QTL was less than 10 cM away from the sorghum flowering time QTL detected in current study

| Sb QTL ID | Number of co-locating QTL from previous studies | | | | | | | | | | |
|----------------|---|---------|---------|---------|---------|---------|---------|---------|---------|---------|----------|
| | Sb QTL | Zm Chr1 | Zm Chr2 | Zm Chr3 | Zm Chr4 | Zm Chr5 | Zm Chr6 | Zm Chr7 | Zm Chr8 | Zm Chr9 | Zm Chr10 |
| QDTF_NAM_1_10 | 1 | 2 | | | | 2 | | | | | |
| QDTF_NAM_1_53 | 2 | 2 | | | | 1 | | | | | |
| QDTF_NAM_1_66 | 3 | 2 | | | | 1 | | | | | |
| QDTF_NAM_1_112 | 1 | 1 | | | | | | | | 4 | |
| QDTF_NAM_1_182 | | | | | | | | | | | |
| QDTF_NAM_2_23 | | | | | | | | 1 | | | |
| QDTF_NAM_2_123 | 2 | | | | | | | 1 | | | |
| QDTF_NAM_2_147 | 2 | | 2 | | | | | 2 | | | |
| QDTF_NAM_3_21 | | | | 1 | | | | | | | |
| QDTF_NAM_3_34 | | | | 3 | | | | | 1 | | |
| QDTF_NAM_3_47 | | | | 3 | | | | | | | |
| QDTF_NAM_3_57 | 1 | | | 4 | | | | | 2 | | |
| QDTF_NAM_3_68 | 1 | | | 1 | | | | | 2 | | |
| QDTF_NAM_3_133 | 1 | | | 1 | | | | | | | |
| QDTF_NAM_4_15 | | | | | | 2 | | | | | |
| QDTF_NAM_4_43 | 1 | | | | 1 | 3 | | | | | |
| QDTF_NAM_4_58 | | | | | 1 | 2 | | | | | |
| QDTF_NAM_4_74 | 1 | | | | 2 | 2 | | | | | |
| QDTF_NAM_4_87 | 1 | | | | 1 | 1 | | | | | |
| QDTF_NAM_4_98 | | | | | 1 | | | | | | |
| QDTF_NAM_5_5 | | | 2 | | 1 | | | | | | |
| QDTF_NAM_5_10 | | | 2 | | 1 | | | | | | |
| QDTF_NAM_5_27 | | | 2 | | 2 | | | | | | |
| QDTF_NAM_5_67 | | | 2 | | 2 | | | | | | |
| QDTF_NAM_6_10 | | | 2 | | | | | | | | 2 |
| QDTF_NAM_6_23 | | | 2 | | | | | | | | 2 |
| QDTF_NAM_6_32 | | | 2 | | | | | | | | 2 |
| QDTF_NAM_6_43 | 3 | | 2 | | | | | | | | 2 |
| QDTF_NAM_6_83 | 2 | | 3 | | | | | | | | 2 |
| QDTF_NAM_7_5 | | | | | 1 | | 1 | | | | |
| QDTF_NAM_7_59 | 3 | | | | 1 | | | | | | 2 |
| QDTF_NAM_7_73 | 1 | 2 | | | 1 | | | | | | |
| QDTF_NAM_8_30 | | | | | | | | | | | |
| QDTF_NAM_8_66 | 3 | | | 1 | | | | | | | 2 |
| QDTF_NAM_8_115 | 1 | | | | | | | | | | |
| QDTF_NAM_9_25 | 1 | | | | | | 2 | | 2 | | |
| QDTF_NAM_9_75 | | | | | | | | | 1 | | |
| QDTF_NAM_9_136 | 2 | | | | | | 1 | | 1 | | |
| QDTF_NAM_10_44 | 7 | | | | | | 1 | | | | |
| QDTF_NAM_10_61 | 2 | | | | | | 1 | | | 1 | |

previously identified flowering time QTL in sorghum. Putative allelic series of multiple, relatively small effect alleles were detected at 75 % of the QTL detected in this

study, with the NRPs contributing both positive (later) and negative (earlier) allele effects relative to R931945-2-2. The significant individual QTL allele effect sizes ranged from

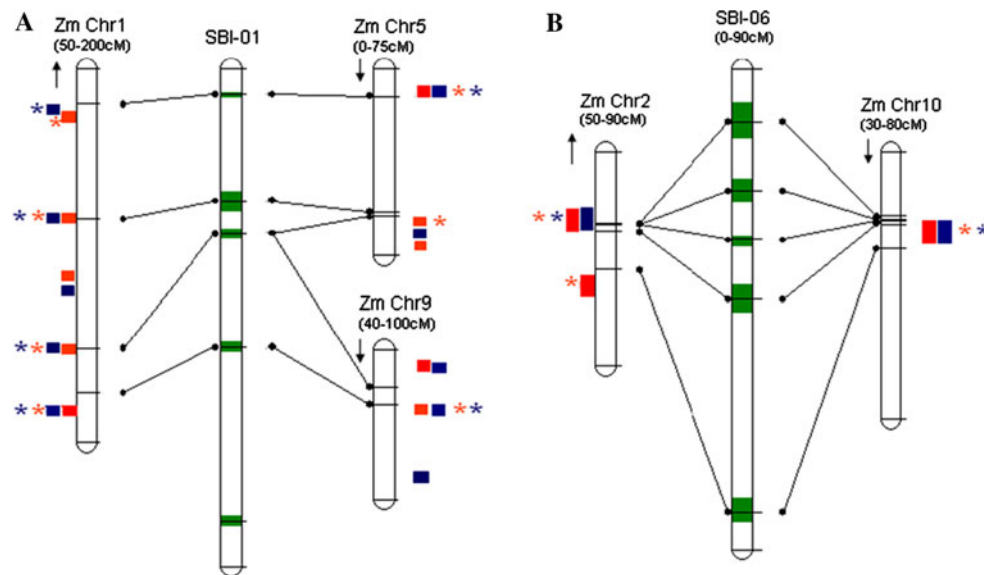


Fig. 6 Sorghum/maize flowering time QTL synteny. **a** Between sorghum SBI-01 and maize chromosomes 1, 5 and 9. **b** Between sorghum SBI-06 and maize chromosomes 2 and 10. Sorghum QTL indicated as *green* segments on sorghum chromosome^a. Projected locations of the mid-point of the sorghum QTL onto the syntenic maize chromosomes indicated by *lines* between chromosomes. Locations of the maize DA (*blue*) and DS (*red*) significant markers

indicated. Co-ordinates of segments of maize chromosomes detailed (based on maize NAM map) and orientation of maize chromosomes, relative to sorghum chromosome, indicated by *arrows*. *Asterisks* indicate maize QTL with mid-points less than 10 cM from the location of the sorghum QTL mid-point^a. For graphical display purposes, when the QTL CI was based on a single marker only, the average CI distance (4.1 cM) was used

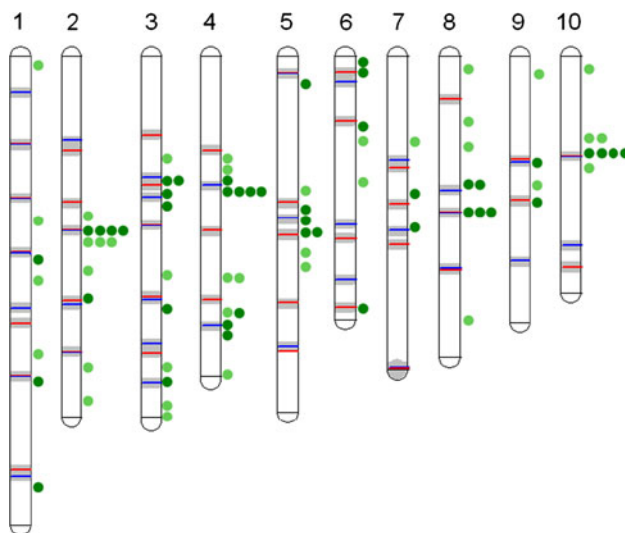


Fig. 7 Genetic linkage NAM map of the maize genome with maize significant markers for flowering time highlighted in red (DS: female flowering) and blue (DA: male flowering) and 4 cM CI highlighted in grey around each significant marker. The predicted locations (based on the projection of the mid-point of the QTL with 4 cM CI) of the sorghum NAM QTL for flowering time are highlighted to the right of each chromosome in *dark green*, where the CIs overlap with a maize QTL, or *light green*, when the CIs don't overlap

0.32 days up to 1.78 days, and additionally both rare and common QTL alleles were detected in our study, ranging from a very rare QTL allele detected in only one population to very common QTL alleles detected in 12 populations. The

integration of the two QTL analysis methodologies used in the current study allowed for the identification of QTL with both rare and common alleles and with small and large effect and also accounted for independent polymorphism patterns across the BC-NAM families. The identification of syntenous QTL in maize provided additional precision, and in many cases permit the identification of underlying candidate genes. The results of this study have successfully contributed to an enhanced understanding of the genetic architecture of flowering time in sorghum, through the identification of QTL in common with previous studies, in both sorghum and maize, and novel QTL, providing unique insights into the genetic control of this trait and its implications for domestication and plant breeding.

Genetic architecture of flowering time in sorghum is controlled by a relatively large number of loci with multiple alleles with small effects

In any QTL experiment, the design of the population or populations studied and the environments where phenotypic data are collected, influence the QTL that are detected, as well as the accuracy of estimates of their locations and effects. The impact of these factors needs to be understood in order for data from experiments to be interpreted appropriately. In this study we used multiple backcross derived populations with a common parent R931945-2-2. The design allows the effects of QTL from

each NRP to be calculated relative to R931945-2-2. Selection, to ensure genotypes fell within a phenotypic range that enabled adaptation to Australian sorghum growing environments, during population development resulted in the removal of genes with very large effects, such as major photoperiod response genes that are known to be present in a number of the NRP, as well as removing some allelic combinations that showed strong epistatic effects such as the known interaction between Ma_5 and Ma_6 (Chanterreau et al. 2001). However, even in the absence of large effect QTL, the analyses undertaken were sufficiently powerful to detect multiple small effect QTL. The backcrossing design also reduces the potential contribution of epistatic interactions between alleles from the NRPs to the genetic variance of the individual BC-NAM families. In addition, our use of F_1 hybrids to evaluate the lines potentially influenced the QTLs we detected due to the potential for dominance of genes in the tester to mask QTL from the NRPs.

The genetic architecture of flowering time in the BC-NAM population was characterised by relatively small allelic effects. Although a few QTL with relatively large QTL effects were detected (ten QTL with a total putative allelic series effect range of >2 days), most of the QTL detected had total estimated allele effect ranges of between 1 and 2 days. The largest individual allele effect size was +1.78 days (QDTF_NAM_6_43). Overall, the effect sizes determined in the BC-NAM population were comparable to the allele effect sizes reported for the previously identified sorghum flowering time QTL (Table 5); the majority (almost 85 %) of the QTL also had QTL allele effect sizes between -2 and +2 days. The largest flowering time QTL allele effect reported for sorghum to date is for Ma_1 , the major photoperiod-sensitivity locus in sorghum, which was reported to have an additive effect of 40.3 days in an interspecific mapping population (*S. bicolor* × *S. proproinquum*) (Lin et al. 1995). However, this QTL is part of the photoperiod pathway and alleles with such large effects would have been removed from our populations by the selection imposed during population development; as evidenced by our finding of that the QTL we identified that co-located with Ma_1 having a much smaller effect size (1.81 days) than previously reported in the literature. This example also highlights the issue of the multiple environmental cues influencing flowering time; specifically temperature and photoperiod. In our study, despite the BC-NAM population being grown in environments with a range of different photoperiods, the genetic correlations between environments were high, indicating low levels of re-ranking of the genotypes which in turn indicates that the majority of the QTL were associated with the temperature component of flowering rather than photoperiod response. Of the 13 QTL studies on flowering time in sorghum to date, only two

have specifically separated the effects of photoperiod and temperature in order to identify QTL for photoperiod sensitivity (Chanterreau et al. 2001; Mannai et al. 2011). Three of the QTL identified in the BC-NAM population correspond to previously described photoperiod sensitivity QTL (2 on SBI-01; QDTF_NAM_1_53 and QDTF_NAM_1_112 and one on SBI-10; QDTF_NAM_10_44). A known flowering time candidate gene in the photoperiod response pathway underlies the QTL on SBI-ten (FT; Sb10g003940), however as multiple QTL co-locate in this genomic region from multiple studies with different environments and photoperiods, it is possible that either this gene influences both the temperature and photoperiod pathways or that it is linked to other flowering time genes in the temperature pathway. The results indicate that the relatively small QTL effects detected was predominantly the result of the removal of photoperiod sensitivity QTL and possibly other large effect QTL by selection imposed during the development of the population. The remaining QTL would be enriched for QTL mainly associated with the other components of the pathway.

This study has revealed that the combined effect of relatively small allelic substitutions across many QTLs can lead to the substantial differences in flowering time observed across the BC-NAM population and by inference cultivated sorghum more generally. We observed large numbers of polymorphic flowering loci. Four BC-NAM families detected 15 or more QTL in the current study; 18 QTL were detected in both the SC56-14E and Tx2908 BC-NAM families, 17 QTL in the *S. arundinaceum* BC-NAM family and 15 in the LR2931-2 BC-NAM family. The genomes of the NRPs were found to be a mosaic of loci for both late and early flowering time effects, leading to total ranges of combined QTL allele effects per population between 19.8 days (SC56-14E) and 1.9 days (SC62C). It should be noted, however, because the populations were backcross derived and selected for maturity, individuals with extreme phenotypes were excluded from the lines that were evaluated. As a result the phenotypic range of flowering time observed in the populations (Fig. 1) was expected to be much less than the range that might be expected if the parental alleles were segregating randomly at a frequency of 0.5.

In general, the numbers of positive and negative alleles detected in a particular NRP were roughly equal suggesting that balancing selection for a particular flowering time was occurring in the breeding populations where the line was developed. The most extreme exception to this general observation was observed in the *S. arundinaceum* population where 19 alleles for early flowering were observed and only one allele for late flowering. This population was the only one that involved a cross between a wild and domesticated sorghum. The accession is an introduced weed in Australia with genes for photoperiod response.

This accession was collected at the southern end of the weed's natural range where the presence of photoperiod genes would potentially prevent the plant flowering at a time when environmental conditions allowed seed production. The weed population was introduced after settlement and has only been in this location for a relatively short period of time (<150 years). The high number of early alleles probably reflects selection for early maturity genes to mitigate the mal-adaptation imposed by the presence of large effect photoperiod sensitivity genes. In the BC-NAM families based on cultivated varieties the detection of large numbers of loci with multiple functional alleles indicate the adaptive benefit to maintaining variation at multiple loci even in cultivated sorghum.

The comparison of our BC-NAM QTL results with previous QTL studies for flowering time in sorghum, revealed that 60 % of the QTL identified in the sorghum BC-NAM population co-located with previously identified flowering time QTL in sorghum. The remaining 40 % (16 in total) was novel QTL for flowering time, likely identified due to the broader range of genetic diversity captured by the BC-NAM, novel segregation between recurrent parent and/or tester, and its greater statistical power compared to that achieved in standard, bi-parental mapping studies. However, it is also possibly due to unique untested environments or statistical anomalies in our approach. In contrast, five genomic regions have been reported in previous studies but were not detected in our BC-NAM population. The QTL in these regions could potentially be related to photoperiod and hence less likely to be detected in the BC-NAM panel; indeed in two-thirds of these QTLs absent from the BC-NAM panel, where flowering time candidate genes underlying the QTL were identified, candidate genes from the photoperiod pathway were present (Table S4); or alternatively these QTL could have been eliminated when selecting for a reduced maturity range in the BC-NAM population. Two of the five QTL not detected in the BC-NAM population were identified in multiple previous studies; however, the remaining three were only detected in single studies, and could represent low frequency alleles contributing novel QTL only detected in specific genetic backgrounds or with specific environmental sensitivity. A range of other factors can also contribute to lack of congruency between studies including mapping resolution influenced by the magnitude of linkage disequilibrium in the mapping population or diversity set and QTL analysis methods used.

Our study validated the majority (22 of the 27) of previously identified unique genomic regions associated with flowering time in sorghum, in addition to identifying 16 novel QTL not previously discovered. In addition to the QTL unique to previous sorghum QTL studies, the total number of unique genomic regions identified to date

associated with flowering time in sorghum is currently just over 43. This is significantly less than the approximately 80 genes currently identified in the flowering time pathway of *Arabidopsis*. However, considering that the majority of flowering time QTL have been identified in temperate sorghum where major effect genes for photoperiod sensitivity have been removed it seems likely that these QTL are predominately associated with temperature rather than photoperiod, whereas 38 genes are known in the photoperiod pathway of *Arabidopsis*, the total number of 43 flowering time QTL identified for sorghum is comparable to the 42 genes currently identified in the temperature, autonomous, and plant hormonal pathways in *Arabidopsis*. This could indicate that a very high proportion of the QTL involved with the genetic control of flowering time associated with temperature have been identified and hence that this study significantly contributes to our global understanding of the genetic architecture of flowering time in sorghum.

High degree of correspondence between sorghum and maize flowering time QTL, as determined through NAM populations

This study revealed a remarkably high degree of correspondence between flowering time QTL identified in both the sorghum and maize NAM populations; two-thirds (50 out of 76) of the significant markers identified for days to anthesis (DA) and days to silking (DS) in the maize NAM panel (Buckler et al. 2009), when projected onto the sorghum consensus map, were within 10 cM of the mid-point of at least one sorghum QTL detected in this study. This is in agreement with a previous study (Lin et al. 1995) that also noted the conservation of genomic regions associated with short-day flowering between many members of the Poaceae family, including sorghum, maize and rice. As noted previously, all the sorghum QTL corresponded to two locations in maize (three sorghum QTL corresponded to three locations), as expected based on the previous studies that have identified that colinear arrangements of chromosomal segments of sorghum (and rice) match maize at a ratio of 1:2, typical of a whole genome duplication event in maize (Messing 2009). Seventy-five percent of the sorghum QTL detected in this study corresponded to a maize flowering time QTL at both syntenic genomic locations. Duplicated regions of genomes frequently accumulate mutations faster than a functional single-copy gene over generations and often lose function (Taylor and Raes 2004); however, the correspondence of the majority of the sorghum QTL to maize QTL in both expected syntenic regions indicate that, in the case of the flowering time genes underlying these QTL, both duplicated genes are likely to be functional, potentially indicating that the

maintenance of functional alleles at duplicated loci is being favoured by natural and/or artificial selection, in contrast to the general observation of biased fractionation in regions of the maize genome derived from whole genome duplication (Woodhouse et al. 2010).

The results of this study highlight the value of using sorghum as a tool to facilitate the dissection of the genetic control of quantitative traits in maize and vice versa. In particular, in situations when the maize syntenous QTL is in the heterochromatin, but the corresponding sorghum syntenous QTL is in the euchromatin (e.g. multiple sorghum QTL corresponding to the heterochromatic region on maize chromosome 2, Fig. 7), the information on the potential candidate genes from the shorter physical fragment underlying the sorghum QTL can be used to inform potential candidates in maize too. Similarly, we can use maize as a model for sorghum, e.g. the sorghum QTL in the heterochromatic region of SBI-05 (QDTF_NAM_5_67) corresponds to two maize QTL, both for DS on two duplicated regions on maize chromosome 4 (DS_4_40.4 and DS_4_105). The mpQTL probability plots based on log Fisher P show a sharp decrease in the P value in the middle of this QTL region, however, it does not fall below the set threshold (Fig. 8). The additional evidence of the two maize QTL provides support for further dissecting the larger sorghum QTL on SBI-05 into two separate QTL; this approach also identified another sorghum QTL on SBI-02 (QDTF_NAM_2_147) that could be further dissected into two separate QTL based on the evidence of multiple maize QTL underlying it. The high degree of correspondence between the maize and sorghum NAM flowering time studies, combined with the high proportion of phenotypic variance explained in the NAM study (89 %), provide further evidence that the majority of the genetic regions controlling flowering time in the sorghum BC-NAM population have been identified.

The genetic architecture of flowering time, therefore, has been demonstrated to have a number of similarities between maize and sorghum based on the number and location of QTL identified for flowering time in both the maize and sorghum NAM populations; however, differences in the magnitude of estimated QTL alleles were observed. In maize, the majority of QTL allele effects were modest with ~80 % of QTL alleles effecting DS by less than 0.5 days. In contrast in sorghum only 22 % of QTL alleles had an effect size of less than 0.5 days, with 18 % having an effect size of greater than 1 day (compared to 2 % in maize). Given that there appears to be a selective advantage to maize in maintaining alleles at both duplicated loci it could be hypothesized that if similar pressures are acting on sorghum then we might see larger effect sizes and a greater number of alleles in each putative allelic series. Additionally, the larger range in most putative

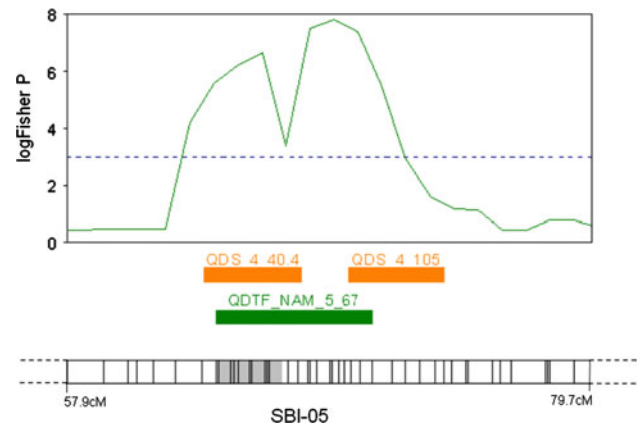


Fig. 8 Segment of SBI-05 showing location of sorghum NAM QTL, QDTF_NAM_5_67 based on log Fisher P plot and two co-locating maize NAM QTL. Grey region on chromosome bar indicates heterochromatic region of SBI-05

allelic series in sorghum in comparison to maize could have been influenced by the selection we carried out for a particular maturity range in our BC-NAM population. Evidence for the putative allelic series, with both early and late effects relative to R931945-2-2, was found in 75 % of the QTL identified in our study, compared to between 69 and 72 % in the maize NAM study. In contrast, in other studies using the maize NAM population to dissect the genetic architecture of other complex traits, the proportion of QTL identified with allelic series was much smaller, e.g. only 37 % of QTL with both positive and negative effects for southern leaf blight resistance (Kump et al. 2011) and only 31–43 % with allelic series for kernel composition traits (Cook et al. 2012). The similarity between the proportion of flowering time QTL identified in the sorghum and maize NAM populations with putative allelic series, despite differences in the magnitude of allele effects, likely reflects the adaptive importance of maintaining variability for this trait, to enable the plant to adjust flowering time in response to changing environmental conditions.

The synteny of QTL within, and particularly across, taxa provides opportunities to fast-track candidate gene identification, through the identification of common candidate genes in syntenous QTL regions. By integrating information on genes involved in flowering time across cereal crops from the reviews of Higgins et al. (2010) and Cockram et al. (2007) with previously identified flowering time genes in sorghum (Mace et al. 2009b), we determined 77 genes involved in flowering time and mapped these sequences *in silico* on the sorghum genome. This yielded a total of 221 loci, out of which 58 corresponded to 18 sorghum BC-NAM QTL for flowering time (ESM Table S4). To demonstrate the utility of this approach, we looked at the closest QTL syntenous between maize and sorghum when projected on to the sorghum genome;

QDTF_NAM_9_25 whose mid-point was just 0.03 cM away from the projected location of the significant marker for the maize QTL DS_6_6.9. One candidate gene for flowering time was located within 3 cM of the syntenous sorghum (QDTF_NAM_9_25) and maize (DS_6_6.9) QTL; APETELA2, a floral identity gene (SbAP2 or Sb09g002080) which in *Arabidopsis* is thought to be a floral repressor, causing late flowering when overexpressed (Putterill et al. 2004). And such an approach, however, assumes that both QTL positions and annotation and gene prediction in the genome databases are accurate; in particular larger QTL intervals are more likely to encompass multiple candidate genes (e.g. of the five sorghum BC-NAM QTL with CI > 10 cM, four had five or more underlying candidate genes for flowering time), so any gene within the interval may be considered a candidate until the interval can be narrowed by either further genetic analysis or cross taxa comparison. In addition, it is possible that QTL represent multiple genes in tight linkage. Nevertheless, our preliminary analysis of candidate genes underlying syntenous QTL in sorghum and maize indicate that unification across taxa provides additional power and resolution to identify candidate genes underlying QTL intervals influencing a complex trait such as flowering time.

Implications for plant breeding

Flowering time is a fundamental trait for plant breeders developing varieties with adaptation to particular target environments and is often the subject of early generation selection in sorghum. Changes in climate brought about by global warming will likely require plant breeders to change the maturity genes present in their elite breeding populations (e.g. Craufurd and Wheeler 2009). To efficiently achieve such change will require detailed knowledge of the genetic architecture of flowering time and its association with other traits. It is critically important that breeders know if particular flowering time alleles are linked to genes for other traits of interest. In such circumstances, early generation selection may aid or restrict progress for those traits. An analysis of the genome-wide distribution of the flowering time QTL known to date in sorghum also indicates that they are not evenly distributed on a centimorgan basis (ESM Fig. S2), with flowering time QTL occurring near the centromere in the heterochromatic regions, which are regions known to suppress recombination (e.g. Akhunov et al. 2003), of every chromosome. Heterochromatic enrichment (on a centimorgan basis) for QTL has been previously observed (Mace and Jordan 2011) and creates conditions favourable for heterosis to occur via pseudo over dominance. Consequently, an understanding of the

impact of flowering time and favourable and unfavourable genetic linkage will be critical for the development of efficient selection strategies, particularly those involving molecular markers.

One further application of detailed knowledge of the flowering time network in sorghum is the capacity to use crop models to predict the consequences of particular combinations of flowering time QTL in particular environments including likely future climate change scenarios. Quantitative crop growth simulation models such as APSIM-sorghum (Hammer et al. 2010), can link changes in the allelic values of genes directly to their phenotypic consequences at the crop level (van Oosterom et al. 2006). Simulation models represent a very important resource for the modern breeder (Sun et al. 2011), facilitating the link between genotype to phenotype and capturing and connecting research at molecular, organism and crop system scales. Until very recently, the level of understanding of the dynamics and functionality of gene networks that determine the link between the genotypic expression of a particular gene (or QTL) and its consequence at the phenotypic level has been lacking and mostly limited to a few model species (van Oosterom et al. 2006; Hammer et al. 2004). Our improved global understanding of the genetic architecture of sorghum, through a meta-analysis of existing QTL studies in addition to the data generated from the current study using the BC-NAM population provides the required input data for the sorghum crop simulation model to dynamically simulate the phenotypic effect of changes in any given genotype, and hence accurately predict the phenotypic consequences of new combinations of genes.

Conclusions

The successful resolution of flowering time genetic architecture in sorghum demonstrates the power of the sorghum BC-NAM resource to catalogue allelic variants at a maximal number of QTL and understand their functional allelic variation and distribution across diverse germplasm. The use of individual populations precludes the generation of such a comprehensive catalogue of both QTL and alleles as it is virtually impossible to create a single bi-parental population that segregates for allelic variants at all genetic loci influencing a complex phenotype (Lin et al. 1995).

The sum of the allele effects presented in this study indicate that the QTL identified cause significant variation in flowering time across the BC-NAM families, with a maximum summed additive effect of QTL alleles of almost 20 days in an individual family. This study provides important information for sorghum breeders and researchers to further manipulate flowering time, which is

likely to be of increasing importance in response to climate change.

The utility of the BC-NAM approach presented here is exemplified by correspondence of QTL previously identified in sorghum and by correspondence of QTL in different taxa. In particular, the significant correspondence of the flowering time QTL identified in the maize and sorghum NAM resources highlights the synergies of studying traits across maize and sorghum, which provide exciting new opportunities and real efficiencies for gene discovery and the dissection of the genetic architecture of many quantitative traits, including but not limited to flowering time.

Acknowledgments The authors would like to thank Erik van Oosterom, Ian Godwin and Chris Lambrides for their critical review of the manuscript and to acknowledge the Queensland Government and the Grains Research and Development Corporation (GRDC) for providing funding for this research.

References

- Akhunov ED, Goodyear AW, Geng S, Qi LL, Echalié B, Gill BS, Miftahudin Gustafson JP, Lazo G, Chao SM, Anderson OD, Linkiewicz AM, Dubcovsky J, La Rota M, Sorrells ME, Zhang DS, Nguyen HT, Kalavacharla V, Hossain K, Kianian SF, Peng JH, Lapitan NLV, Gonzalez-Hernandez JL, Anderson JA, Choi DW, Close TJ, Dilbirli M, Gill KS, Walker-Simmons MK, Steber C, McGuire PE, Qualset CO, Dvorak J (2003) The organization and rate of evolution of wheat genomes are correlated with recombination rates along chromosome arms. *Genome Res* 13:753–763
- Altschul S, Gish W, Miller WP, Myers E, Lipman D (1990) Basic local alignment search tool. *J Mol Biol* 215:403–410
- Amasino RM, Michaels SD (2010) The timing of flowering. *Plant Physiol* 154:516–520
- Beavis WD (1994) The power and deceit of QTL experiments: lessons from comparative QTL studies. In: Proceedings of 49th annual corn sorghum and research conference. American Seed Trade Association, Washington, DC, pp 250–266
- Brown PJ, Klein PE, Bortiri E, Acharya CB, Rooney WL, Kresovich S (2006) Inheritance of inflorescence architecture in sorghum. *Theor Appl Genet* 113:931–942
- Buckler ES et al (2009) The genetic architecture of maize flowering time. *Science* 325:714–718
- Butler DG, Cullis BR, Gilmour AR, Gogel BJ (2009) ASReml-R reference manual release 3. Technical report, QLD Department of Primary Industries and Fisheries, Brisbane, QLD
- Chantreau J, Trouche G, Rami JF, Deu M, Barro C, Grivet L (2001) RFLP mapping of QTLs for photoperiod response in tropical sorghum. *Euphytica* 120:183–194
- Cockram J, Jones H, Leigh FJ, O'Sullivan D, Powell W, Laurie DA, Greenland AJ (2007) Control of flowering time in temperate cereals: genes, domestication and sustainable productivity. *J Exp Bot* 58:1231–1244
- Colasanti J, Coneva V (2009) Mechanisms of floral induction in grasses: something borrowed, something new. *Plant Physiol* 149:56–62
- Cook JP, McMullen MD, Holland JB, Tian F, Bradbury P, Ross-Ibarra J, Buckler ES, Flint-Garcia SA (2012) Genetic architecture of maize kernel composition in the nested association mapping and inbred association panels. *Plant Physiol* 158:824–834
- Crasta OR, Xu WW, Nguyen HT, Rosenow DT, Mullet J (1999) Mapping of post flowering drought resistance traits in grain sorghum: association between QTLs influencing premature senescence and maturity. *Mol Gen Genet* 262:579–588
- Craufurd PQ, Wheeler TR (2009) Climate change and the flowering time of annual crops. *J Exp Bot* 60:2529–2539
- Cullis BR, Smith AR, Coombes NE (2006) On the design of early generation variety trials with correlated data. *J Agric Biol Environ Stat* 11:381–393
- Feltus FA, Hart GE, Schertz KF, Casa AM, Kresovich S, Abraham S, Klein PE, Brown PJ, Paterson AH (2006) Alignment of genetic maps and QTLs between inter- and intraspecific sorghum populations. *Theor Appl Genet* 112:1295–1305
- Fisher RA (1932) Statistical methods for research workers, 4th edn. Oliver and Boyd, Edinburgh
- Flowers JM, Hanzawa Y, Hall MC, Moore RC, Purugganan MD (2009) Population genomics of the *Arabidopsis thaliana* flowering time gene network. *Mol Biol Evol* 26:2475–2486
- Greenup A, Peacock WJ, Dennis ES, Trevaskis B (2009) The molecular biology of seasonal flowering-responses in *Arabidopsis* and the cereals. *Ann Bot* 103:1165–1172
- Hammer GL, Sinclair TR, Chapman SC, van Oosterom EJ (2004) On systems thinking, systems biology and in the in silico plant. *Plant Physiol* 134:909–911
- Hammer GL, van Oosterom E, McLean G, Chapman SC, Broad I, Harland P, Muchow RC (2010) Adapting APSIM to model the physiology and genetics of complex adaptive traits in field crops. *J Exp Bot* 61:2185–2202
- Hart GE, Schertz KF, Peng Y, Syed NH (2001) Genetic mapping of *Sorghum bicolor* (L.) Moench QTLs that control variation in tillering and other morphological characters. *Theor Appl Genet* 103:1232–1242
- Higgins JA, Bailey PC, Laurie DA (2010) Comparative genomics of flowering time pathways using *Brachypodium distachyon* as a model for the temperate grasses. *PLoS ONE* 5:e10065
- Imaizumi T (2010) *Arabidopsis* circadian clock and photoperiodism: time to think about location. *Curr Opin Plant Biol* 13:83–89
- Jordan DR, Mace ES, Cruickshank AW, Hunt CH, Hanzell RG (2011) Exploring and exploiting genetic variation from unadapted sorghum germplasm in a breeding program. *Crop Sci* 51:1444–1457
- Jung C, Muller AE (2009) Flowering time control and applications in plant breeding. *Trends Plant Sci* 14:563–573
- Kebede H, Subadhi PK, Rosenow DT, Nguyen HT (2001) Quantitative trait loci influencing drought tolerance in grain sorghum (*Sorghum bicolor* L. Moench). *Theor Appl Genet* 103:266–276
- Kim J (2003) Genomic analysis of sorghum by fluorescence in situ hybridization. PhD thesis. Texas A&M University, Texas
- Kirby JS, Atkins RE (1968) Heterotic response for vegetative and mature plant characters in grain sorghum, *Sorghum bicolor* (L.) Moench. *Crop Sci* 8:335–339
- Kump KL, Bradbury PJ, Wissner RJ, Buckler ES, Belcher AR, Oropeza-Rosas MA, Zwonitzer JC, Kresovich S, McMullen MD, Ware D, Balint-Kurti PJ, Holland JB (2011) Genome-wide association study of quantitative resistance to southern leaf blight in the maize nested association mapping population. *Nat Genet* 43:163–168
- Lagercrantz U (2009) At the end of the day: a common molecular mechanism for photoperiod response in plants? *J Exp Bot* 60:2501–2515
- Liang GHL, Walter TL (1968) Heritability estimates and gene effects for agronomic traits in grain sorghum, *Sorghum vulgare* Pers. *Crop Sci* 8:77–81
- 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
- Mace ES, Jordan DR (2011) Integrating sorghum whole genome sequence information with a compendium of sorghum QTL studies reveals non-random distribution of QTL and of gene rich regions with significant implications for crop improvement. *Theor Appl Genet* 123:169–191
- 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
- Mace ES, Rami JF, Bouchet S, Klein PE, Klein RR, Kilian A, Wenzl P, Xia L, Halloran K, Jordan DR (2009a) A consensus genetic map of sorghum that integrates multiple component maps and high-throughput diversity array technology (DARt) markers. *BMC Plant Biol* 9:13
- Mace ES, Hunt C, Rodgers D, Jordan DR (2009b) Multi-population QTL analysis in sorghum. *SABRAO J Breed Genet* (41), special supplement. A paper presented at the 14th Australasian plant breeding conference, 10–14 Aug 2009, Cairns, Australia
- Mannai YE, Shehzad T, Okuno K (2011) variation in flowering time in sorghum core collection and mapping of QTLs controlling flowering time by association analysis. *Genet Resour Crop Evol* 58:983–989
- McMullen MD et al (2009) Genetic properties of the maize nested association mapping population. *Science* 325:737–740
- Messing J (2009) The polyploidy origin of maize. *Handbook of maize*, pp 221–238. doi:10.1007/978-0-387-77863-1_11
- Mouradov A, Cremer F, Coupland G (2002) Control of flowering time: interacting pathways as a basis for diversity. *Plant Cell* 14(supplement):S111–S130
- Murphy RL, Klein RR, Morishige DT, Brady JA, Rooney WL, Miller FR, Dugas DV, Klein PE, Mullet JE (2011) Coincident light and clock regulation of pseudoresponse regulator protein 37 (PRR37) controls photoperiodic flowering in sorghum. *Proc Natl Acad Sci USA* 108:16469–16474
- Murray SC, Rooney WL, Mitchell SE, Sharma A, Klein PE, Mullet JE, Kresovich S (2008) Genetic improvement of sorghum as a biofuel feedstock: II. QTL for stem and leaf structural carbohydrates. *Crop Sci* 48:2180–2193
- Parh D (2005) DNA-based markers for ergot resistance in sorghum. PhD thesis. University of Queensland, Brisbane
- Poland JA, Bradbury PJ, Buckler ES, Nelson RJ (2011) Genome-wide nested association mapping of quantitative resistance to northern leaf blight in maize. *Proc Natl Acad Sci USA* 108:6893–6898
- Putterill J, Laurie R, Macknight R (2004) It's time to flower: the genetic control of flowering time. *BioEssays* 26:363–373
- Quinby JR (1967) The maturity genes of sorghum. In: Ag Norman (ed) *Advances in Agronomy*, vol 19. Academic Press, New York, pp 267–305
- R Development Core Team (2012) R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna. ISBN 3-900051-07-0
- Ritter KB, Jordan DR, Chapman SC, Godwin ID, Mace ES, McIntyre CL (2008) Identification of QTL for sugar-related traits in a sweet x grain sorghum (*Sorghum bicolor* L. Moench) recombinant inbred population. *Mol Breed* 22:367–384
- Rooney WL, Aydin S (1999) Genetic control of a photoperiod-sensitive response in *Sorghum bicolor* (L.) Moench. *Crop Sci* 39:397–400
- Schuler GD (1998) Electronic PCR: bridging the gap between genome mapping and genome sequencing. *Trends Biotechnol* 16:456–459
- 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* 121:323–336
- Simpson GG (2004) The autonomous pathway: epigenetic and post-transcriptional gene regulation in the control of *Arabidopsis* flowering time. *Curr Opin Pl Biol* 7:1–5
- Smith A, Cullis B, Thompson R (2001) Analyzing variety by environment data using multiplicative mixed models and adjustments for spatial field trend. *Biometrics* 57:1138–1147
- Srinivas G, Satish K, Madhusudhana R, Nagaraja Reddy R, Murali Mohan S, Seetharama N (2009) Identification of quantitative trait loci for agronomically important traits and their association with genic-microsatellite markers in sorghum. *Theor Appl Genet* 118:1439–1454
- Stephens JC, Miller FR, Rosenow DT (1967) Conversion of alien sorghums to early combine genotypes. *Crop Sci* 7:396
- Sun X, Peng T, Mumm RH (2011) The role and basics of computer simulation in support of critical decisions in plant breeding. *Mol Breeding* 28:421–436
- Taylor JS, Raes J (2004) Duplication and divergence: the evolution of new genes and old ideas. *Ann Rev Genet* 38:615–643
- Tian F, Bradbury PJ, Brown PJ, Hung H, Sun Q, Flint-Garcia S, Rocheford TR, McMulle MD, Holland JB, Buckler ES (2011) Genome-wide association study of leaf architecture in the maize nested association mapping population. *Nat Genet* 43:159–162
- Van Oosterom EJ, Hammer GL, Chapman SC, Doherty A, Mace ES, Jordan DR (2006) Predicting flowering time in sorghum using a simple gene network: functional physiology or fictional functionality? In: *Proceedings of the 5th Australian sorghum conference*. 30 Jan–2 Feb 2006, Gold Coast, Australia
- VSN International (2011) *GenStat for Windows*. 14th edn. VSN International, Hemel Hempstead, UK. <http://GenStat.co.uk>
- Wenzel WG (1988) Relationships between yield, base temperature and drought resistance in grain sorghum. *S Afr J Pl Soil* 5:150–154
- Woodhouse MR, Schnable JC, Pedersen BS, Lyons E, Lisch D, Subramanian S, Freeling M (2010) Following tetraploidy in maize, a short deletion mechanism removed genes preferentially from one of the two homologs. *PLoS Biol* 8:e1000409