

QTL analysis of early-season cold tolerance in sorghum

Joseph Knoll · Nilupa Gunaratna · Gebisa Ejeta

Received: 17 July 2007 / Accepted: 21 November 2007 / Published online: 21 December 2007
© Springer-Verlag 2007

Abstract Cool temperatures during the early-growing season are a major limitation to growing sorghum [*Sorghum bicolor* (L.) Moench] in temperate areas. Several landraces from China have been found to exhibit higher emergence and greater seedling vigor under cool conditions than most breeding lines currently available, but tend to lack desirable agronomic characteristics. The introgression of desirable genes from Chinese landraces into elite lines could be expedited by marker-assisted selection. Using a population of 153 RI lines, developed from a cross between Chinese landrace ‘Shan Qui Red,’ (SQR, cold-tolerant) and SRN39 (cold-sensitive), QTL associated with early-season performance under both cold and optimal conditions were identified by single marker analysis, simple interval mapping (SIM), and composite interval mapping (CIM). Germination was observed under controlled conditions, and other traits were measured in field plantings. Two QTL for germination were identified: one on linkage group SBI-03a, derived from SRN39, was significant under cold and optimal temperatures. The other, on group SBI-07b, showed greater significance under cold temperatures and was contributed by SQR. A region of group SBI-01a, derived from SQR, showed strong associations with seedling emergence and seedling vigor scores under early and late field plantings. A QTL for both early and late emergence was

identified by CIM on SBI-02 which favored the SRN39 allele. SIM identified a QTL for early vigor on SBI-04 favoring the SQR genotype. Further studies are needed to validate the effects of these QTL, but they represent the first step in development of a marker-assisted breeding effort to improve early-season performance in sorghum.

Introduction

Sorghum [*Sorghum bicolor* (L.) Moench] originated in tropical and subtropical regions of Africa (Doggett 1988) and is therefore well adapted to warm growing conditions. It is an important grain crop in many parts of the world, including the United States. Because of its tropical origins, cool temperatures, especially during the early growing season, are the major limitation to sorghum cultivation in the northern United States and other temperate areas (Allegre de la Soujeole and Miller 1984). The ability of sorghum to withstand many other abiotic stresses, such as drought, would otherwise make it an ideal crop for many temperate locations in the northern United States. The development of sorghum cultivars with improved early-season cold tolerance would allow expansion of sorghum to these more northerly latitudes and would also allow for earlier planting in areas where it is currently being grown (Yu and Tuinstra 2001). Improved emergence and early-season vigor would enable better stand establishment, and protect against loss of seedlings during unexpected cold periods in late spring. Because plant population and spacing have been shown to affect grain yield (Willey and Heath 1969), improvement in stand establishment and early-season survival should ultimately lead to higher and more reliable yields of grain sorghum in temperate locations.

Communicated by H. H. Geiger.

J. Knoll · G. Ejeta (✉)
Department of Agronomy, Purdue University,
915 W. State Street, West Lafayette, IN 47907, USA
e-mail: gejeta@purdue.edu

N. Gunaratna
Department of Statistics, Purdue University,
150 N. University Street, West Lafayette, IN 47907, USA

Although most of the available sorghum germplasm is of tropical origin, some sorghums have evolved in isolation in temperate regions of China. These sorghums, referred to as ‘kaoliangs,’ tend to exhibit higher seedling emergence and greater seedling vigor under cool conditions than most other sorghum varieties (Cisse and Ejeta 2003; Singh 1985). Sorghum was most possibly introduced to China through trade with India, perhaps as early as 4,000 years ago (Doggett 1988). In China, natural selection under the temperate environment resulted in evolving the kaoliangs into a distinctly unique group of landrace sorghum cultivars. These landraces form an excellent genetic resource for the improvement of cold tolerance in sorghum (Yu and Tuinstra 2001; Franks et al. 2006). However, many of these cold tolerant Chinese varieties lack favorable agronomic characteristics, such as short stature and disease resistance, required for sorghum to be grown in the United States. To harness the potential of these Chinese landraces for cold tolerance improvement would likely involve intensive backcrossing and selection to eliminate the undesirable traits of the donor parents and to recover the desired cold tolerance characteristics. Marker-assisted breeding could be used to direct the introgression of seedling cold tolerance genes from kaoliangs into elite high-yielding lines while minimizing the introduction of other undesirable traits from the donor parents.

The objective of this research is to use kaoliang germplasm to improve early-season performance of grain sorghum for cultivation in temperate areas through marker-assisted selection and breeding. To identify genetic markers associated with quantitative trait loci (QTL) for seedling vigor traits under cold stress conditions, a recombinant inbred (RI) mapping population was developed from a cross between the Chinese kaoliang ‘Shan Qui Red’ (SQR) and an African caudatum SRN39. SQR exhibits high emergence and vigorous seedling growth under cold temperatures, while SRN39 performs poorly under early-season cold stress (Cisse 1995). In a previous study, Gunaratna (2002) observed that seedling vigor traits under cold stress appear to be under similar genetic control as those observed under more optimal temperatures. In this study, we performed separate QTL analyses for germination, field emergence, and seedling vigor under both cold and more optimal conditions to test this hypothesis, and to identify markers associated with these traits.

Materials and methods

Genetic material

A recombinant inbred (RI) population of sorghum was created by crossing ‘Shan Qui Red’ (SQR), a Chinese

kaoliang, with SRN39, a caudatum of African origin. The F₁ generation of the cross was selfed to produce an F₂ population from which individual F₂ plants were then advanced through subsequent generations using the single-seed descent method of plant breeding. Advanced generations were maintained through selfing and bulking the seed from each line after the F₆ generation (Cisse and Ejeta 2003). Seed of the parental lines and the RI progeny were maintained in a common nursery.

Measurement of phenotypic traits

Germination percentage was measured in controlled conditions for 146 RI lines, both at optimal (30°C for 8 h/20°C for 16 h) and cold (13°C constant) temperatures. The optimal temperature regime is recommended by the Association of Official Seed Analysts (AOSA 1999). Germination under optimal temperature was measured in two replicates in a completely randomized design, and germination under cold temperature was measured in four replicates in a randomized complete block design (blocked by time). In each replicate, 100 seeds of each parent and RI line were treated by soaking overnight in 5% (w/v) suspension of captan {*N*-[(trichloromethyl)thio]-4-cyclohexane-1,2-dicarboximide} fungicide. They were then rinsed with distilled water and placed on moistened Whatman No. 1 filter paper in 100 × 15 mm polystyrene Petri dishes. These were maintained at the proper temperatures in dark Conviron E15 growth chambers (Controlled Environments, Ltd., Winnipeg, Manitoba) at maximum relative humidity. The filter papers were moistened throughout the experiment as needed. Emergence of the radicle from the seed was counted as a successful germination. Final germination was counted 7 days after sowing for the optimal test and at 14 days for the cold test (Gunaratna 2002).

Percent emergence was measured in the field under early (cold) and late (warmer) planting in three successive years (1998–2000) at the Purdue Agronomy Center for Research and Education (ACRE) in West Lafayette, IN. A larger set of RI lines (177 entries) and both parents were evaluated in the field for cold tolerance and other agronomic characteristics. Each planting consisted of two replicates, planted in a randomized complete block design (blocked by location in the field). One hundred seeds were planted in each experimental plot. Final percent emergence was determined as the number of emerged seedlings 40 days after planting. Seedlings were scored for vigor in the field. Vigor scores were assessed visually on a scale of 1–5, with 1 representing very high vigor and 5 representing very poor vigor, as described by Maiti et al. (1981). Vigor scores are based on size and physical appearance of the seedlings, and are not based on seedling emergence percentages.

In all, six sets of phenotypic data (germination at cold temperature, germination at optimal temperature, early field emergence, late field emergence, early vigor score, and late vigor score) were analyzed in this study.

Data analysis

Statistical analyses of the phenotypic data were performed using SAS v. 8.0 (SAS Institute 2001). Prior to analysis, the germination data were transformed to improve the homogeneity of error variances using the following formula:

$$Y' = \arcsin(\sqrt{Y})$$

where Y is the percent germination. For germination at optimal temperature (two replications, completely randomized design) the following ANOVA model was applied using PROC GLM in SAS:

$$Y_{ij} = \mu + \alpha_i + \varepsilon_{ij}$$

where Y_{ij} is the j th observation of the i th genotype, μ is the overall experiment mean, α_i is the effect of the i th genotype, and ε_{ij} represents the normally distributed random error. Other traits were analyzed using randomized complete block designs. For germination at cold temperature (four replications) and emergence in the field at normal planting (one planting and two replications), the following model was used

$$Y_{ij} = \mu + \alpha_i + \beta_j + \varepsilon_{ij}$$

where Y_{ij} is the observed value for the i th genotype in the j th block, μ is the overall experiment mean, α_i is the effect of the i th genotype, β_j is the effect of the j th block, and ε_{ij} represents the normally distributed random error. For all other field data, multiple plantings were analyzed and the following model was used:

$$Y_{ijk} = \mu + \alpha_i + \beta_j + \gamma_k + (\alpha\gamma)_{ik} + \varepsilon_{ijk}$$

where Y_{ijk} is the observed value for the k th genotype in the j th block in the i th planting, μ is the overall experiment mean, α_i is the effect of the i th planting, $\beta_{j(i)}$ is the effect of the j th block nested within the i th planting, γ_k is the effect of the k th genotype, $(\alpha\gamma)_{ik}$ is the interaction between the i th planting and k th genotype, and ε_{ijk} represents the normally distributed random error (Gunaratna 2002).

Heritability estimates were calculated for all traits. For germination in the growth chamber and for emergence under normal planting, the following formula was used:

$$H = \frac{MS_G - MS_E}{MS_G}$$

where MS_G is the mean square for genotype and MS_E is the mean square for error. For other field traits, the heritability was calculated using a slightly different formula:

$$H = \frac{MS_G - MS_{P \times G}}{MS_G}$$

where $MS_{P \times G}$ is the mean square for the genotype by planting interaction (Gunaratna 2002).

Genotyping of recombinant inbred lines

Different types of polymorphic DNA markers including simple sequence repeats (SSR), random amplified polymorphic DNA (RAPD), and restriction fragment length polymorphisms (RFLP), were used to genotype parental and RI lines. For the SSR markers, each 25- μ l PCR mixture contained 15 ng template DNA, 0.2 μ M each forward and reverse primer, 0.05 mM each of dATP, dCTP, dGTP, and dTTP, 1.5 mM $MgCl_2$, 50 mM KCl, 10 mM Tris pH 8.9, 0.01% Triton X-100, 0.02% cresol red dye, 12% sucrose, and 2.5 U *Taq* DNA polymerase. Reactions were carried out in 96-well plates in a PTC-100 Programmable Thermal Controller (MJ Research, Watertown, MA, USA) or in a Primus 96 thermocycler (MWG-Biotech, High Point, NC, USA). Reaction temperature parameters varied slightly for SSR markers depending on the source of the primers. For all SSR markers developed by Kong et al. (2000) and Bhatramakki et al. (2000), the conditions were as follows: initial denaturation at 94°C for 60 s, followed by 35 cycles of 94°C (denaturation) for 10 s, 55°C (annealing) for 35 s, and 72°C (extension) for 45 s. This was followed by a final extension at 72°C for 3 min. For all other SSRs the same PCR conditions were used, except the annealing temperature was 58°C. RAPD reactions were carried out as described by Cisse (1995) and Weerasuriya (1995). All PCR products were visualized by agarose gel electrophoresis and staining with ethidium bromide. For the RFLP markers, electrophoresis and blotting were carried out as described by Hulbert et al. (1990). Probes were radioactively labeled by random primer extension as described by Feinberg and Vogelstein (1983). Hybridization and washing conditions were as described by Yang et al. (1996), and the markers were visualized by exposure to the X-ray film.

All markers with the prefix 'Xt x p' are SSR markers from primers developed at Texas A&M University by Kong et al. (2000) or Bhatramakki et al. (2000). Marker *SvHPR* and all markers with the prefix 'SbAG' are SSR markers developed by Taramino et al. (1997); other markers with the prefix 'Sb' are SSR markers developed by Brown et al. (1996). All markers with the prefix 'OP' are RAPD markers from Operon Technologies (Alameda, CA, USA), and all markers with the prefix 'UBC' are RAPD markers from the Nucleic Acids and Protein Services Unit at the University of British Columbia (Vancouver, BC, Canada). All other markers are RFLPs, which used probe sequences

from maize. Clones were obtained from various sources (Melake-Berhan et al. 1993). Three biallelic phenotypic markers were also included: testa (present/absent), pericarp color (red/white), and seedling plant color (red/green).

Linkage map construction

A total of 143 markers were included in the dataset, but not all RI lines were scored for all markers. Individual RI lines missing a large proportion (>25%) of marker data were thus eliminated from the dataset, giving a working mapping population size of 153 RI lines. The final marker set included 56 SSR, 67 RAPD, 17 RFLP, and three phenotypic markers. All markers were tested for segregation distortion using a chi-square test with one degree of freedom at $\alpha = 0.05$, with a Bonferroni correction for multiple testing of the 143 markers. A total of 11 markers were significantly distorted (2 SSR, 6 RAPD, and 3 RFLP) and were excluded from further analyses.

The remaining 132 markers were used to construct a genetic linkage map using *Mapmaker/EXP* v. 3.0b (Lincoln et al. 1992, 1993). Tentative linkage groups were formed using the ‘group’ command with a minimum LOD score of 3.0 (default) and a maximum genetic distance of 20 Kosambi centiMorgans (cM) (Kosambi 1944). The groups were then built up using the ‘near’ and ‘try’ commands. Local map orders with the shortest distance were determined using the ‘ripple’ command. A relatively large maximum distance was allowed in order to align the map with those previously published by Bhatramakki et al. (2000) and Menz et al. (2002) using common SSR and RFLP markers. Linkage groups with gaps larger than 50 cM were broken into separate groups. The linkage groups were then numbered following the chromosome designations outlined by Kim et al (2005). Separate linkage groups known to reside on the same chromosome were given the appropriate number followed by ‘a’ or ‘b’. One small linkage group, containing three markers, could not be assigned to a chromosome, and was designated as group SBI-U.

Single marker analysis

Single marker analysis was performed using *Windows QTL Cartographer* v. 2.5 (Wang et al. 2006). Simple linear regressions were performed on all traits to test the association of each marker genotype to the trait phenotypic least-squares mean value. Transformed values were used for the germination data. The other data were not transformed prior to this analysis. No corrections were made to control overall Type-I error (chance of identifying false positives),

as this analysis was conducted as a prelude to interval mapping. The latter procedure was used to confirm QTL identified by single marker analysis. Because a linkage map is not required, single marker analysis also allowed for the analysis of the unlinked markers.

Interval mapping

The unlinked markers were eliminated from the dataset prior to interval mapping. *Windows QTL Cartographer* v. 2.5 (Wang et al. 2006) was used to perform simple interval mapping (SIM, model 3) and composite interval mapping (CIM, model 6) (Zeng 1993, 1994). Background markers for inclusion in the CIM model were selected by forward/backward stepwise regression for each trait. The five most significant background markers were then used for analysis (default). The ‘walking speed’ was set at 2 cM for both models, and the ‘window size’ at 10 cM for CIM. The experiment-wise significance thresholds (α) were determined for each trait using 1,000 permutations in the appropriate model, as described by Churchill and Doerge (1994). The additive effects and percentage of variation explained (R^2) for all significant QTL were determined at their peak LR values.

Results

Prior to transformation, the distribution of germination percentages from the growth chamber experiments was highly skewed to the left. Among the RI lines, the mean germination percentages were 79.0% for the cold treatment and 81.3% for the optimal treatment (Table 1). Although the difference in these percentages appears small, a paired two-sample *t* test of the transformed data showed that these treatments were in fact statistically different at $\alpha = 0.05$. They were also highly correlated ($R = 0.883$). In contrast, data for the field-based traits were distributed more symmetrically. The mean seedling emergence for early plantings was 38.2%, and increased to 55.8% for later plantings. A strong correlation was observed between early and late emergence percentages ($R = 0.739$). Seedling vigor also improved slightly, but significantly, with later planting. The mean vigor score was 3.31 for early plantings and 2.97 for later plantings. Early and late vigor scores were also highly correlated ($R = 0.777$). Transgressive segregation was observed among the RI population; however, the parents remained near the phenotypic extremes in each of the trait distributions. SQR had significantly higher germination percentage than SRN39 under both cold and optimal temperatures. Field emergence percentage of SQR was also greater than that of SRN39, and the Chinese

Table 1 Summary of phenotypic data for all traits measured in SRN39, SQR, and the RI population. SRN39 and SQR are significantly different for all traits at $\alpha \leq 0.05$

Trait	SRN39	SQR	RI lines				Heritability
	Mean	Mean	Mean	St. dev.	Min.	Max.	
Cold germination (%)	57.8	95.2	79.0	14.5	27.8	98.9	0.929
Optimal germination (%)	63.5	96.5	81.3	14.4	34.0	98.5	0.936
Early emergence (%)	9.4	53.9	38.2	20.0	0.0	92.0	0.773
Late emergence (%)	26.7	89.4	55.8	19.8	0.0	100.0	0.749
Early vigor (score 1–5)	4.56	1.81	3.31	0.89	1.00	5.00	0.773
Late vigor (score 1–5)	4.42	1.50	2.97	0.85	1.00	5.00	0.728

parent also showed significantly better seedling vigor (lower score) after both early and late planting dates (Table 1). Heritability estimates were very high for all traits (Table 1). The highest heritability estimates were obtained for percent germination in the growth chamber at cold ($H = 0.929$) and optimal temperature ($H = 0.936$), although these could be overestimated due to the fact that any genotype by environment variance is included in the genotype mean square (MS_G) value.

Genetic linkage map

Linkage groups were constructed based on the locations of common SSR and RFLP markers in the maps published previously by Bhatramakki et al. (2000) and Menz et al. (2002). The linkage groups were then numbered according to the chromosome designations of Kim et al. (2005). Three corresponding linkage groups (SBI-01, SBI-03, and SBI-07) were divided into two groups each because of long gaps (>50 cM). One additional linkage group of three markers was also found, and was designated SBI-U (for ‘unlinked’), as its position relative to any published maps could not be determined. The longest distance allowed between two markers within any linkage group was 45.5 cM (on linkage group SBI-01a). There are two other gaps larger than 40 cM (on linkage groups SBI-02, and SBI-07a). Six markers (four RAPDs, one RFLP, and the phenotypic marker for testa) could not be assigned to any linkage group, for a total of 126 mapped markers. The total genetic distance of this map is 2,128 cM. The genetic linkage map of sorghum generated in this analysis is diagrammed in Fig. 1.

Single marker analysis

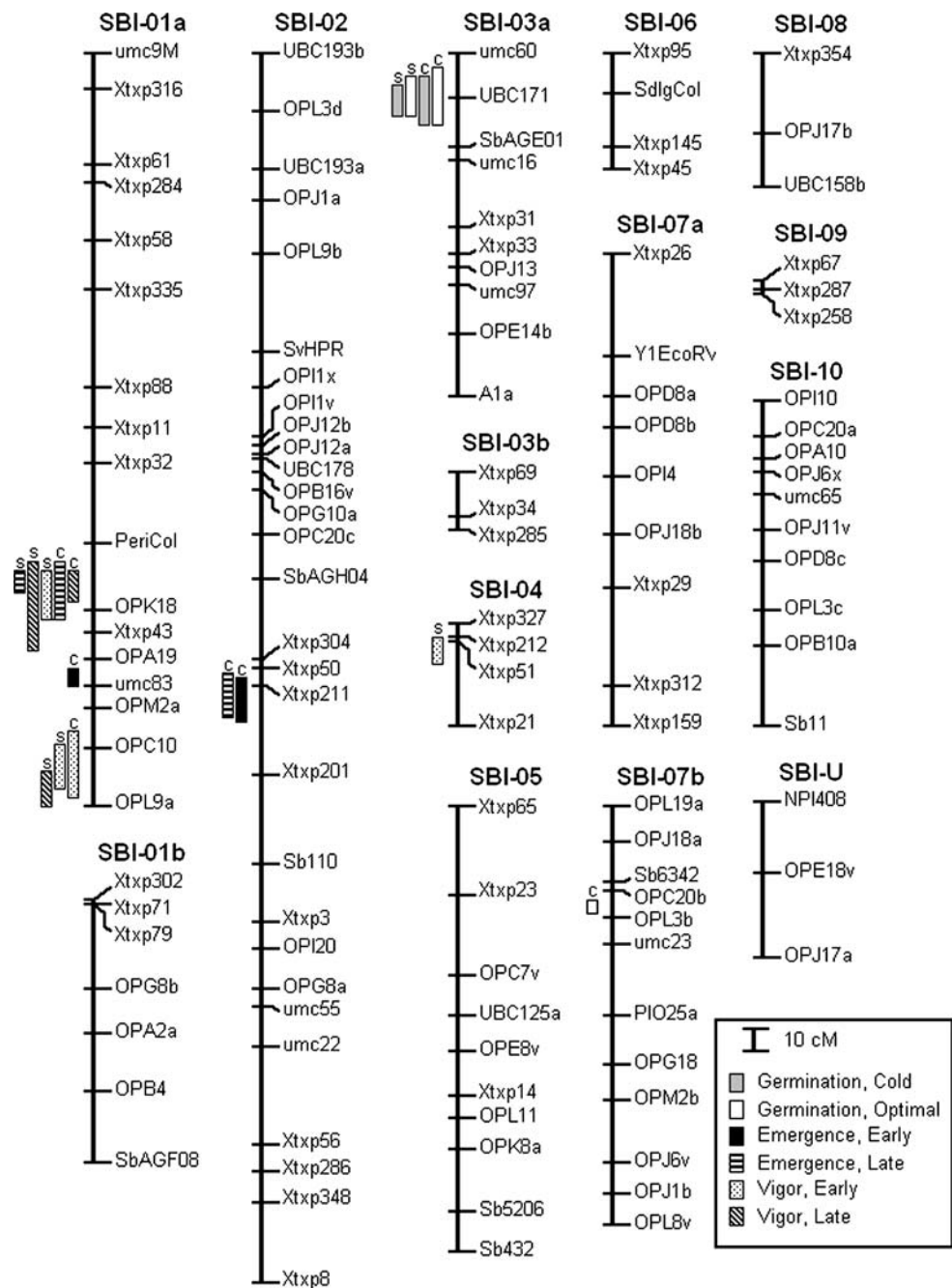
Single marker analysis revealed possible associations between several markers and phenotypes. All significant

markers identified by single marker analysis are summarized in Table 2. The most highly significant marker for germination under cold or optimal temperature was *UBC171* on SBI-03a. Surprisingly, increased seed germination at this locus was associated with the SRN39 allele, as were most of the other markers for germination at either temperature. Linkage group SBI-07b carries several markers showing highly significant association with the SQR allele at both temperature treatments.

Markers significantly associated by single marker analysis with percent emergence in early planting were found on linkage groups SBI-01a, SBI-01b, SBI-02, SBI-03a, SBI-04, SBI-07a, and SBI-09. Some of these markers were also significant for emergence under late planting. SSR marker *Xtxp211* had the highest additive effect for emergence under late planting (4.6% points) and also had a strong effect on emergence under early planting (3.7% points). In both cases, higher emergence at this locus is associated with the SRN39 genotype. Most other markers associated with greater percent emergence are of the SQR genotype.

Most of the significant markers associated with high seedling vigor (low vigor score), at either planting date, favor the SQR genotype. Many of the markers that are significant for vigor after early planting are also significant for vigor after late planting. However, there are notable exceptions. Those with the greatest additive effects are located on linkage groups SBI-01a, SBI-04, and SBI-09 for early planting, and on groups SBI-01a, SBI-02, and SBI-05 for later planting. SSR markers *Xtxp50* and *Xtxp211* on linkage group SBI-02, and three SSR markers on group SBI-09 (*Xtxp67*, *Xtxp287*, and *Xtxp258*), showed significant additive effects for seedling vigor under late planting, but not under early planting, while three markers on group SBI-04 (*Xtxp327*, *Xtxp212*, and *Xtxp51*) were found to be highly significant under early planting, but not late. The markers on groups SBI-02 and SBI-09 were associated with the SRN39 genotype, while those on group SBI-04 were associated with the SQR genotype.

Fig. 1 Genetic linkage map of sorghum, and QTL showing significant ($\alpha \leq 0.10$) associations with early-season traits by interval mapping. Linkage group numbers are indicated above their respective linkage groups, following the nomenclature of Kim et al. (2005). Marker names are indicated on the right of each linkage group, and genetic distance in centiMorgans (cM) is indicated by the *scale bar*. QTL positions are indicated by the *bars* to the left of the linkage groups, and their associated traits are indicated in the *legend*. The letters 's' or 'c' above the QTL indicate detection of the QTL by simple interval mapping (SIM) or composite interval mapping (CIM), respectively



Interval mapping

Significant QTL were identified for all traits using interval mapping. The interval mapping results support many of the putative QTL identified by single marker analysis. Several, but not all, QTL that were significant in single marker analysis were also found to be significant by one or both interval mapping models. Interval mapping did not identify any QTL that were not detected by single marker analysis. The results of SIM and CIM are similar, but there are some notable differences. The results of both interval mapping procedures are summarized in Table 3, and the positions

of the QTL on the genetic map of sorghum are indicated in Fig. 1.

Both SIM and CIM identified a highly significant ($\alpha < 0.05$) QTL for germination at both cold and optimal temperatures on linkage group SBI-03a near RAPD marker *UBC171*. As in single marker analysis, increased germination percentage was associated with the SRN39 allele. As estimated by CIM, this QTL explained 11.5% of the variation at cold temperature and 14.3% in the optimal temperature treatment. Additive effects were similar to those estimated by single marker analysis (6.2 arcsine percentage points for cold and 6.9 for the optimal

Table 2 Additive effect of one SRN39 allele at significant markers for early-season traits, as estimated by single marker analysis. Non-significant effects are not listed

Marker	Linkage group	Germination (arcsin %) ^a		Emergence (%) ^a		Seedling vigor (score 1–5) ^b	
		Cold	Optimal	Early	Late	Early	Late
Xtxp58	SBI-01a	–	–	–0.039**	–	0.152**	0.114*
PeriCol		–	–	–	–0.035*	0.099*	0.119*
OPK18		–	–	–0.031*	–0.042**	0.158***	0.188****
Xtxp43		–	–	–0.032*	–0.036*	0.132**	0.176***
OPA19		–	–	–0.035**	–0.034*	0.138**	0.150**
umc83		–	–	–0.038**	–	0.134**	0.100*
OPC10		–	–	–0.026*	–	0.155***	–
OPL9a		–	–	–	–0.037*	0.123**	0.199****
Xtxp302	SBI-01b	–	–	–	–	–0.095*	–
OPA2a		–	–0.032*	–	–	–	–
SbAGF08		–	–	0.035**	–	–0.119*	–0.124*
Xtxp50	SBI-02	–	–	–	0.037**	–	–0.097*
Xtxp211		–	–	0.037**	0.046**	–	–0.127**
OPG8a		–0.031*	–0.034*	–	–	–	–
umc22		–	–	–0.028*	–	–	–
Xtxp286		0.036*	0.034*	–	–	–	–
Xtxp348		0.037*	0.043**	–	–	–	–
UBC171	SBI-03a	0.065****	0.069****	0.033*	0.031*	–	–
umc16		0.036*	–	–	–	–	–
Xtxp33		–	–	–0.025*	–	–	–
umc97		–	–	–0.035**	–	–	–
A1a		–	–	–	0.036*	–0.098*	–0.108*
Xtxp69	SBI-03b	–	–	–	–	–	0.118*
Xtxp34		0.031*	–	–	–	–	–
Xtxp285		0.040*	–	–	–	–	–
Xtxp327	SBI-04	–	–	–0.029*	–	0.126**	–
Xtxp212		–	–	–0.036**	–0.029*	0.160***	–
Xtxp51		–	–	–0.033**	–0.030*	0.158***	–
OPC7v	SBI-05	–	–	–	–0.034*	–	0.140**
Xtxp14		–	–	–	–	0.101*	–
Sb5206		–	–	–	–	0.098*	–
OPJ18b	SBI-07a	0.041**	0.034*	–	–	–	–
Xtxp29		0.046**	0.040*	–0.037**	–	0.108*	0.114*
OPJ18a	SBI-07b	–0.040**	–0.044**	–	–	–	–
Sb6342		–0.042**	–0.040**	–	–	–	–
OPC20b		–0.046**	–0.046**	–	–0.030*	–	0.116*
OPL3b		–0.043**	–0.033*	–	–	–	0.097*
OPJ17b	SBI-08	0.037*	–	–	–	–	–
Xtxp67	SBI-09	–	–	–	0.031*	–0.108*	–
Xtxp287		–	–	0.031*	–	–0.122**	–
Xtxp258		–	–	0.030*	–	–0.118*	–
Ksu3a	Unlinked	–	–	–	0.033*	–	–0.102*
Testa	Unlinked	–	–	–	–	0.112*	0.119*

*, **, ***, **** Significance at $\alpha = 0.05, 0.01, 0.001, \text{ and } 0.0001$, respectively

^a A negative sign indicates that a negative effect (decreased germination or emergence) is associated with the SRN39 allele. Percents are expressed as decimals

^b A negative sign indicates that higher vigor (lower score) is associated with the SRN39 allele

Table 3 Quantitative trait loci (QTL) associated with early-season traits, detected by simple interval mapping (SIM) and composite interval mapping (CIM)

Trait	Linkage group	Position (cM)	Flanking markers		Additive effect ^a	R ² ^b
<i>Simple interval mapping</i>						
Cold germination (arcsin %)	SBI-03a	20.6	UBC171	SbAGE01	0.065	0.127**
Optimal germination (arcsin %)	SBI-03a	20.0	umc60	UBC171	0.072	0.152**
Early emergence (%)	—	—	—	—	—	—
Late emergence (%)	SBI-01a	237.0	PeriCol	OPK18	−0.080	0.206**
Early vigor (score 1–5)	SBI-01a	239.0	PeriCol	OPK18	0.252	0.203**
	SBI-01a	323.8	OPC10	OPL9a	0.248	0.197**
	SBI-04	11.3	Xtxp51	Xtxp21	0.193	0.120*
Late vigor (score 1–5)	SBI-01a	239.0	PeriCol	OPK18	0.309	0.277**
	SBI-01a	331.8	OPC10	OPL9a	0.253	0.185**
<i>Composite interval mapping</i>						
Cold germination (arcsin %)	SBI-03a	20.6	UBC171	SbAGE01	0.062	0.115**
	SBI-07b	44.8	OPC20b	OPL3b	−0.054	0.094*
Optimal germination (arcsin %)	SBI-03a	22.6	UBC171	SbAGE01	0.070	0.143**
Early emergence (%)	SBI-01a	279.1	OPA19	umc83	−0.044	0.080*
	SBI-02	283.5	Xtxp201	Sb110	0.046	0.081**
Late emergence (%)	SBI-01a	237.0	PeriCol	OPK18	−0.076	0.186**
	SBI-02	281.2	Xtxp211	Xtxp201	0.059	0.105**
Early vigor (score 1–5)	SBI-01a	321.8	OPC10	OPL9a	0.269	0.227**
Late vigor (score 1–5)	SBI-01a	239.0	PeriCol	OPK18	0.261	0.175**

*, ** Significance at $\alpha = 0.10$ and 0.05 respectively

^a Additive effect of one SRN39 allele

^b Percentage (expressed as decimal) of variance explained by the QTL

temperature treatment.). Composite interval mapping also revealed a marginally significant ($\alpha \leq 0.10$) QTL for germination at cold temperature on linkage group SBI-07b, between two RAPD markers *OPC20b* and *OPL3b*. For this QTL, increased germination percentage was associated with the SQR genotype, and it explained 9.5% of the observed variation.

Simple interval mapping did not identify any significant QTL for emergence in the field under early planting, but two were identified by CIM. A marginally significant ($\alpha = 0.10$) QTL was located on group SBI-01a between *OPA19* and *umc83*. The other QTL is located on linkage group SBI-02 between markers *Xtxp201* and *Xtxp211*. This second QTL explained 8.1% of the observed variation with an additive effect of 4.6% points. This same region of the genome was also found to be significant for emergence after optimal planting by CIM, but not by SIM. For optimal planting, the percentage of variation explained by this QTL was 10.5% and the additive effect was 5.9% points. At this locus, higher emergence was associated with the SRN39 genotype. A QTL for field emergence after optimal planting was identified by both interval-mapping models on linkage group SBI-01a between the phenotypic marker for pericarp color (*PeriCol*) and RAPD

marker *OPK18*. This QTL had an additive effect of −7.6% points, and explained 18.6% of the observed variation as estimated by CIM. The negative sign of the additive effect indicates that higher emergence at this locus is associated with the SQR allele.

Simple interval mapping identified two QTL on linkage group SBI-01a for association with seedling vigor under both early and late plantings. The first QTL is in the same region as the QTL for seedling emergence between *PeriCol* and *OPK18*. Under early planting this QTL explained 20.3% of the variation in vigor score, and 27.7% under later planting, as estimated by SIM. The second QTL is located between *OPC10* and *OPL9a*. The variation explained by this QTL was 19.7% under early planting and 18.5% under later planting. When analyzed by CIM, the first QTL was only significant for vigor under late planting, with an additive effect of 0.261, and explained 17.5% of the variation in seedling vigor scores. With CIM, the second QTL on group SBI-01a was only significant for vigor under early planting, explaining 22.7% of the variation, with an additive effect of 0.269. SIM also identified a QTL for early seedling vigor on SBI-04, near SSR markers *Xtxp212* and *Xtxp51*, explaining an estimated 12.0% of the observed variation. For all these seedling vigor QTL, the

SQR allele is associated with lower vigor score (more vigorous seedlings).

Discussion

The genetic map of sorghum generated in this study was constructed to follow the orders and groupings in previously published maps (Bhatramakki et al. 2000; Menz et al. 2002; Kim et al. 2005). Most of the map orders are the same as those previously reported, but our map distances are much larger. The total genetic distance of our map was 2,128 cM. In contrast, the map of Bhatramakki et al. (2000) is only 1,406 cM in total genetic distance, and the map of Menz et al. (2002) is 1,713 cM.

A QTL for germination at both cold and optimal temperatures is located on linkage group SBI-03a, near *UBC171*. This suggests that a common genetic mechanism controls germination at either temperature regime at this locus. Surprisingly, higher germination is associated with the SRN39 allele of this QTL, even though this parent has lower germination than SQR. Most possibly, SRN39 possesses genes at other loci that reduce the effects of this QTL. Single marker analysis indicated several candidate QTL for this trait at various locations. A QTL for germination at cold temperature was identified by CIM on group SBI-07a, and it was associated with the SQR genotype. This QTL was only significant for the cold treatment when analyzed by CIM, but single marker analysis indicated significance in this region of the genome under both temperature treatments. It is possible that germination at this locus is slightly affected by temperature.

Using interval-mapping procedures, no QTL were found to be common to both emergence and germination at any temperature regime, despite the apparent relatedness of these traits. Closer examination of our data shows that percent germination in the growth chamber at cold or optimal temperature is only weakly correlated with percent emergence in the field at normal planting time ($R = 0.260$ and $R = 0.275$, respectively), and is not significantly correlated under early field planting. Past evidence varies as to the correlation between seed germination in controlled conditions and emergence in the field. A lack of correlation between controlled environment and field-based seedling characteristics was reported by Alegre de la Soujeole and Miller (1984) in sorghum and by McConnell and Gardner (1979) in maize. More recently, however, Tiryaki and Andrews (2001) reported a strong correlation ($R = 0.66$) between rate of germination at 15°C and rate of emergence in an early-planted field test, although a correlation between final percentages was not reported. Yu et al. (2004) assayed 30 sorghum hybrids and reported a moderate correlation ($R = 0.44$) between germination in a

controlled soil-free environment (15°C) and emergence in the field. This general lack of agreement between controlled environment and field-based studies is possibly due to the high level of variability under field conditions. Nonetheless, germination and emergence, although seemingly related, are possibly under separate genetic control. Although high germination is an important trait, from a practical standpoint it would be more useful to the breeder to select directly for field emergence, or for markers associated with this trait. If markers are to be used, however, QTL must be identified which can maintain their influence despite the highly variable conditions found in the field.

A region of linkage group SBI-01a contains QTL for seedling emergence and seedling vigor, all of which are associated with the SQR genotype. The QTL for emergence and vigor under optimal planting overlap, which may suggest a common genetic mechanism controlling both emergence and subsequent seedling growth. Using SIM, two QTL for seedling vigor were located on linkage group SBI-01a for both early and late planting. However, with CIM, two separate seedling vigor QTL emerged, one for early planting, and one for late planting. Also, CIM identified an additional QTL for early seedling emergence, which lies between these two other QTL. A QTL for seedling emergence, significant under both planting times, was located on SBI-02 by CIM, but not by SIM. These examples highlight the importance of utilizing and comparing different models to locate and analyze QTL. Due to the relatively small sample size of this population, it is not certain whether this region of SBI-01a truly contains one or multiple genes controlling vigor. It is also not clear whether the QTL for early emergence in this same region actually represents the same locus for later emergence. These QTL might all represent the same gene, and the multiple peaks could be an effect of sample size. The sample size may be too small to precisely determine the exact locations of these QTL (Beavis 1998). Nonetheless, this region of the SQR genome contains markers, which should be useful for marker assisted selection to improve early-season performance in sorghum. Our data show that percent seedling emergence in the field was highly correlated with seedling vigor score ($R = -0.822$ under early planting and $R = -0.780$ under later planting). The finding of linked or overlapping QTL for both emergence and vigor supports this result, and it is certainly possible that a common genetic mechanism influences both high emergence and increased seedling vigor. CIM identified two linked QTL on SBI-01a for seedling vigor, one for early and one for late. SIM also identified these two QTL, but under this model, both were significant under early and late plantings. Either model is consistent with our finding that seedling vigor scores under early and late plantings are

highly correlated ($R = 0.777$), and should indicate that the QTL derived from SQR show stable performance across varying temperature environments.

A potential QTL was detected on linkage group SBI-04 near *Xtsp51* that was associated only with seedling vigor in early planting, but not in later planting. This QTL may represent a temperature-specific cold tolerance mechanism, derived from SQR, which influences early-season vigor. This QTL was highly significant under single marker analysis, significant at the 0.10 level in SIM, but not significant in CIM. This linkage group only contains four markers, which could possibly explain this result. Further investigation is needed to verify the effect of this QTL.

The underlying genetic mechanisms responsible for the QTL associated with early-season traits in sorghum are not yet known. Using a sample from the RI population described in this study, Cisse (1995) found that phenolic compounds in the seed were significantly positively correlated with seedling emergence, seedling vigor, seedling height, and germination at 22°C. Cisse (1995) also found that the concentration of flavan-4-ols, tannins, and total phenols in the seed were associated with RAPD marker *OPA19* and several other markers on the same linkage group. In this study, *OPA19* and the phenotypic marker for pericarp color (*PeriCol*) are located in the region of SBI-01a that shows strong associations with seedling emergence and vigor. This suggests that the presence of phenolic compounds in the seed, and particularly in the seed coat, may contribute to emergence and seedling vigor in SQR, perhaps by suppressing soil-borne pathogens, which are often present in cold, wet soils. Interestingly, the presence of a testa, where much of the phenolic compounds in sorghum seeds are found, was only associated with seedling vigor score ($\alpha = 0.05$), but not with emergence, by single marker analysis. This marker could not be assigned to any linkage group, and thus it was not included in the interval mapping analyses. The presence of additional QTL on other linkage groups indicates that phenolic compounds are probably not the only mechanism influencing seedling emergence and vigor in this sorghum population. Additional studies will be needed to elucidate the functions of the genes underlying these QTL, and to verify the role of phenolic compounds in enhancing early-season performance.

Markers linked to the QTL identified in this study will be used as selection tools in breeding early season vigor and cold tolerance from SQR into elite sorghum breeding lines. Validation of the effects of these molecular markers across other populations and environments will be an additional, but very important, step in developing a marker-assisted breeding effort. Markers which can be further validated will be introgressed into elite breeding lines, which can then be evaluated in hybrid combinations.

In a further evaluation of other Chinese kaoliangs in our collection, lines with seedling vigor as good as, or better than, SQR have been identified (unpublished). To use these lines in a marker-assisted breeding program, it would be necessary to determine if these lines carry the same or different QTL for early season traits as SQR. Yang et al. (1996) investigated diversity among a collection of 34 Chinese sorghum lines using RAPD, RFLP, and inter-simple sequence repeat (ISSR) markers. The level of similarity between lines was 30 to 94%, with an average of 70%, when data from all three types of markers (100 markers total) were combined. Thus, it would be expected that most of the lines in our collection should share at least a few common genes with SQR at important QTL regions. Lines which carry known QTL could be introduced into the breeding program. However, great diversity has been observed within our collection of Chinese kaoliangs, and it is likely that investigation of these lines would reveal additional QTL conferring early season cold tolerance.

Similar to landraces from tropical areas, the kaoliang landraces tend to be tall also, often in excess of two or three meters, because of selection by traditional farmers for use as building material and fuel in addition to grain production (Qingshan and Dahlberg 2001). They are also susceptible to an array of sorghum diseases including leaf rust (*Puccinia purpurea* Cooke) and anthracnose (*Colletotrichum graminicola*). Despite these limitations, the kaoliangs have been shown to be an excellent source of genes for early-season cold tolerance. In a set of hybrids generated with diverse pollinators which included two kaoliangs, SQR and IS4225, Yu and Tuinstra (2001) observed significant and favorable general combining ability (GCA) for early season seedling cold tolerance among kaoliangs, highlighting the potential of these lines as a source of cold tolerance genes to generate improved hybrids. Marker-assisted selection will expedite the introgression of these useful genes from kaoliang landraces into elite breeding lines, and should help to minimize the introduction of undesirable traits from the donor parents.

Acknowledgments We are grateful for the assistance provided by several colleagues. Dr. Cécile Grenier collated all of the RFLP, RAPD, and SSR data from previous studies into a common, usable dataset. Terry Lemming assisted in planting of the field tests. Prof. R. W. Doerge provided instruction on using the mapping and QTL analysis software. Funding for this project was provided by International Sorghum and Millets (INTSORMIL)-USAID Grant # DAN254-G-00-002.

References

- Alegre de la Soujeole AA, Miller FR (1984) Cold tolerance of sorghum during early developmental stages. Proc Ann Corn Sorghum Res Conf 39:18–32

- Association of Official Seed Analysts (AOSA) (1999) Rules for testing seeds. Association of Official Seed Analysts. Las Cruces
- Beavis WD (1998) QTL analysis: power, precision, and accuracy. In: Paterson AH (eds) Molecular dissection of complex traits. CRC Press, Boca Raton, pp 145–162
- Bhattaramakki D, Dong J, Chhabra AK, Hart GE (2000) An integrated SSR and RFLP linkage map of *Sorghum bicolor* (L.) Moench. *Genome* 43:988–1002
- Brown SM, Hopkins MS, Mitchell SE, Senior ML, Wang TY, Duncan RR, Gonzalez-Candelas F, Kresovich S (1996) Multiple methods for the identification of polymorphic simple sequence repeats (SSRs) in sorghum [*Sorghum bicolor* (L.) Moench]. *Theor Appl Genet* 93:190–198
- Churchill GA, Doerge RW (1994) Empirical threshold values for quantitative trait mapping. *Genetics* 138:963–971
- Cisse N (1995) Heritability estimates, genetic correlation, and identification of RAPD markers linked to seedling vigor and associated agronomic traits in sorghum. Ph.D. thesis. Purdue University, West Lafayette
- Cisse N, Ejeta G (2003) Genetic variation and relationships among seedling vigor traits in sorghum. *Crop Sci* 43:824–828
- Doggett H (1988) Sorghum, 2nd edn. Wiley, New York
- Feinberg A, Vogelstein B (1983) A technique for radiolabelling DNA restriction endonuclease fragments to high specific activity. *Anal Biochem* 132:6–13
- Franks CD, Burow GB, Burke JJ (2006) A comparison of U.S. and Chinese sorghum germplasm for early season cold tolerance. *Crop Sci* 46:1371–1376
- Gunaratna N (2002) Early season cold tolerance in sorghum. M.S. thesis, Purdue University, West Lafayette, IN
- Hulbert SH, Richter TE, Axtell JD, Bennetzen JL (1990) Genetic mapping and characterization of sorghum and related crops by means of maize DNA probes. *Proc Natl Acad Sci USA* 87:4251–4255
- Kim J-S, Klein PE, Klein RR, Price HJ, Mullet JE, Stelly DM (2005) Chromosome identification and nomenclature of *Sorghum bicolor*. *Genetics* 169:1169–1173
- Kong L, Dong J, Hart G (2000) Characteristics, linkage-map positions, and allelic differentiation of *Sorghum bicolor* (L.) Moench DNA simple-sequence repeats (SSRs). *Theor Appl Genet* 101:438–448
- Kosambi DD (1944) The estimation of map distances from recombination values. *Ann Eugen* 12:172–175
- Lincoln S, Daly M, Lander E (1992) Constructing genetic maps with Mapmaker/EXP 0. Technical report. 3rd edn. Whitehead Institute, Cambridge
- Lincoln S, Daly M, Lander ES (1993) Mapmaker/EXP, version 3.0b. Whitehead Institute, Cambridge
- Maiti RK, Raju PS, Bidinger FR (1981) Evaluation of visual scoring for seedling vigor in sorghum. *Seed Sci Technol* 9:613–622
- McConnell RL, Gardner CO (1979) Selection of cold germination in two corn populations. *Crop Sci* 19:765–768
- Melake-Berhan A, Hullbert SH, Butler LG, Bennetzen JL (1993) Structure and evolution of the genomes of *Sorghum bicolor* and *Zea mays*. *Theor Appl Genet* 86:598–604
- Menz MA, Klein RR, Mullet JE, Obert JA, Unruh NC, Klein PE (2002) A high-density genetic map of *Sorghum bicolor* (L.) Moench based on 2926 AFLP, RFLP and SSR markers. *Plant Mol Biol* 48:483–499
- Qingshan L, Dahlberg JA (2001) Chinese sorghum genetic resources. *Econ Bot* 55:401–425
- SAS Institute (2001) SAS for Windows, Version 8. SAS Inst. Inc., Cary
- Singh SP (1985) Sources of cold tolerance in grain sorghum. *Can J Plant Sci* 65:251–257
- Taramino G, Tarchini R, Ferrario S, Lee M, Pe ME (1997) Characterization and mapping of simple sequence repeats (SSRs) in *Sorghum bicolor*. *Theor Appl Genet* 95:66–72
- Tiryaki I, Andrews DJ (2001) Germination and seedling cold tolerance in sorghum I. Evaluation of rapid screening methods. *Agron J* 93:1386–1391
- Wang S, Basten CJ, Zeng Z-B (2006) Windows QTL Cartographer 2.5. Department of Statistics, North Carolina State University, Raleigh
- Weerasuriya YM (1995) The construction of a molecular linkage map, mapping of quantitative trait loci, characterization of polyphenols, and screening of genotypes for *Striga* resistance in sorghum. Ph.D. thesis, Purdue University, West Lafayette, IN
- Willey RW, Heath SB (1969) The quantitative relationships between plant population and crop yield. *Adv Agron* 21:281–321
- Yang W, de Oliveira AC, Godwin I, Schertz K, Bennetzen JL (1996) Comparison of DNA marker technologies in characterizing plant genome diversity: variability in Chinese sorghums. *Crop Sci* 36:1669–1676
- Yu J, Tuinstra MR (2001) Genetic analysis of seedling growth under cold temperature stress in grain sorghum. *Crop Sci* 41:1438–1443
- Yu J, Tuinstra MR, Claassen MM, Gordon WB, Witt MD (2004) Analysis of cold tolerance in sorghum under controlled environment conditions. *Field Crop Res* 85:21–30
- Zeng Z (1993) Theoretical basis for separation of multiple linked gene effects in mapping quantitative trait loci. *Proc Natl Acad Sci USA* 90:10972–10976
- Zeng Z (1994) Precision mapping of quantitative trait loci. *Genetics* 136:1457–1468