Molecular Genetics and Genomics
https://doi.org/10.1007/s00438-020-01646-0

ORIGINAL ARTICLE

Genome-wide association study of soybean seed germination under drought stress

Zhangxiong Liu[1] · Huihui Li[1] · Zuowang Gou[2] · Yanjun Zhang[2] · Xingrong Wang[2] · Honglei Ren[3] · Zixiang Wen[4] · Beom-Kyu Kang[5] · Yinghui Li[1] · Lili Yu[1] · Huawei Gao[1] · Dechun Wang[4] · Xusheng Qi[2] · Lijuan Qiu[1]

Received: 22 April 2019 / Accepted: 9 January 2020

Abstract

Drought stress, which is increasing with climate change, is a serious threat to agricultural sustainability worldwide. Seed germination is an essential growth phase that ensures the successful establishment and productivity of soybean, which can lose substantial productivity in soils with water deficits. However, only limited genetic information is available about how germinating soybean seeds may exert drought tolerance. In this study, we examined the germinating seed drought-tolerance phenotypes and genotypes of a panel of 259 released Chinese soybean cultivars panel. Based on 4616 Single-Nucleotide Polymorphisms (SNPs), we conducted a mixed-linear model GWAS that identified a total of 15 SNPs associated with at least one drought-tolerance index. Notably, three of these SNPs were commonly associated with two drought-tolerance indices. Two of these SNPs are positioned upstream of genes, and 11 of them are located in or near regions where QTLs have been previously mapped by linkage analysis, five of which are drought-related. The SNPs detected in this study can both drive hypothesis-driven research to deepen our understanding of genetic basis of soybean drought tolerance at the germination stage and provide useful genetic resources that can facilitate the selection of drought stress traits via genomic-assisted selection.

Keywords Soybean · Drought tolerance · Germination stage · Genome-wide association analysis (GWAS) · Single-nucleotide polymorphisms (SNP)

Introduction

Drought is by many measures the most significant environmental stress adversely affecting world agricultural production (Boyer 1982; Cattivelli et al. 2008; Tuberosa and Salvi 2006), and predictions suggest that water deficit will continue to be a hugely influential abiotic factor affecting global crop yields (Sharma and Lavanya 2002). Moreover, in light of current trends in climate change, it is also appreciated that drought could become more frequent and severe (Mittler and Blumwald 2010). To address the challenges of global

Communicated by Stefan Hohmann.

Zhangxiong Liu, Huihui Li, and Zuowang Gou have contributed equally to this work.

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s00438-020-01646-0) contains supplementary material, which is available to authorized users.

✉ Lijuan Qiu
qiulijuan@caas.cn

Xusheng Qi
qixusheng6608@sina.com

[1] National Key Facility for Gene Resources and Genetic Improvement/Key Laboratory of Crop Germplasm Utilization, Ministry of Agriculture/Institute of Crop Sciences, Chinese Academy of Agricultural Sciences, Beijing 100081, China

[2] Institute of Crop Sciences, Gansu Academy of Agricultural Sciences, Lanzhou 730070, China

[3] Maize Research Institute, Heilongjiang Academy of Agricultural Sciences, Harbin 150086, China

[4] Department of Plant, Soil and Microbial Sciences, Michigan State University, East Lansing 48824, USA

[5] Upland Crop Breeding Research Division, Department of Southern Area Crop Science, National Institute of Crop Science, Miryang 52402, Korea

Published online: 01 February 2020

food security, it is thus critical to increase our understanding of plant responses to abiotic stress; insights in this area can support efforts to breed crops that can maintain higher improved yield under stress conditions (Condon et al. 2004; Morison et al. 2008).

Soybean is the most widely grown seed legume in the world; and it is an inexpensive source of both protein and vegetable oil. Drought stress is one of the most common adverse environmental conditions affecting soybean production (Dogan et al. 2007; Hufstetler et al. 2007; Mohammadi et al. 2012), and it is now understood that drought affects soybean seed yields (to at least some degree) during all growth stages (Boyer 1982). Typical yield losses associated with water deficit are about 40% (Specht et al. 1999); however, depending upon the intensity and stage at which such stress occurs, yield losses can be 80% or higher (Dias et al. 2012; Oya et al. 2004). Due to the high costs of irrigation and/or lack of water availability, irrigation is not a feasible option for most of the soybean growing regions. Therefore, the development of soybean genotypes with drought tolerance, particularly those with the ability to germinate and become established in water-deficit conditions, is viewed as one of the most economical ways to increase the stability and sustainability of soybean production. Thus, drought tolerance has been identified as a major breeding target for crop improvement (Rönnberg-Wästljung et al. 2005; Pennisi 2008; Xiong et al. 2006).

A large number of studies have highlighted that tremendous genetic variability underlies observed differences in soybean drought-tolerance traits (Frederick et al. 2001; Mederski and Jeffers 1973; Specht et al. 1986). Traditionally, drought-tolerant genotypes have been selected based on observations of phenotypes for plants grown under controlled conditions. However, conventional breeding for drought-tolerant crop varieties is time-consuming and labor intensive, due in large part to the complex quantitative nature of drought-tolerance traits and the attendant difficulties in making selection for drought tolerance (Charlson et al. 2009; Ribaut et al. 1997). Thus, if quantitative genetics or other analytical methods can identify genomic regions associated with drought tolerance, breeders would have more useful guidance to facilitate the rapid development of improved cultivars with increased drought tolerance, for example using marker-assisted selection.

QTL mapping on drought-related traits in soybean has been conducted, focusing for example on yields under drought stress conditions (Du et al. 2009a, b), fibrous roots (Abdel-Haleem et al. 2011), water use efficiency (Mian et al. 1996, 1998; Specht et al. 2001), canopy wilting (Abdel-Haleem et al. 2012; Charlson et al. 2009; Hwang et al. 2015, 2016), leaf pubescence density, leaf wilting coefficient, and rate of excised leaf drying (Du et al. 2009c). However, all of these studies have been based on linkage mapping using biparental populations. Few studies were on QTLs underlying drought-tolerance traits at the germination stage. Due to the unknown molecular mechanism of drought tolerance at germination stage in soybean, it is hard for breeders to exploit genetic information on genomic-assisted breeding on drought tolerance.

The present study used association analysis based on genotypic and phenotypic information for a large germplasm diversity panel of 259 released Chinese soybean cultivars and sought to identify genetic mechanisms underlying drought tolerance at the seed germination stage. We used a popular method for phenotyping the drought tolerance of germinating seeds for each cultivar in our association panel, and used a total of 4616 SNPs for our mixed-linear model GWAS. In total, we identified 15 associations between given SNPs and drought-tolerance indices, and three of these SNPs were significantly associated with more than one of the drought-tolerance phenotypic indices. Five of the identified associations have been previously associated with drought-tolerance traits (canopy architecture and seed hardness), supporting both the utility of our GWAS analysis and suggesting that physiological mechanisms controlling drought tolerance in the germination stage may be similar to those known to help plants tolerate drought stress at later stages of the plant lifecycle. Our study thus lays the foundation for the molecular breeding of soybeans with increased drought tolerance and generates multiple functional hypotheses relating to the basic biology of how germinating seeds manage water deficits.

Materials and methods

Plant materials

An association mapping panel composed of 259 cultivars released in the Northeast of China and the Yellow River Valleys was constructed (Liu et al. 2017). Of the 259 Chinese cultivars (which were from 11 provinces), 217 were derived from the Northern spring (Nsp) ecotype, while 42 were derived from the Huang-huai-hai summer (Hsu) ecotype. Notably, these cultivars are known to exhibit both clear geographic differentiation and genetic diversity (Liu et al. 2017). Detailed information for these 259 accessions is presented in Table S1.

Phenotyping

Fundamentally, our strategy for assessing the drought tolerance of the seeds of the various cultivars was based the PEG6000 concentration for soybean seeds drought treatment reported by Shu et al. (2015) (the PEG6000 concentration which we used was 22%) and the data analysis was based on Thabet

et al. (2018). Very briefly, 30 healthy seeds for each cultivar were selected; seeds were put into sterilized Petri dishes ($15 \times 15 \times 2.5$ cm^3) containing two vermiculite layers: the mass of the vermiculite in each of the two layers was 35 g. Next, 120 ml of PEG6000 or control (aqueous) solutions (22% or 0%) were added to the dishes wherein the seeds were sandwiched between the vermiculite layers. Subsequently, the plates were incubated in the dark in a germination chamber at 25 ± 1 °C, and the number of germinate seeds were counted after 7 days of incubation. Note that seeds were considered to have germinated when the radicle was as long as the seed and when the radical was half as long as the seed. For each cultivar, the five germinated seeds with longest radicle were selected; the lengths of these radicles were measured, and their fresh weights were measured and recorded. This seed germination experiment, which also tracked overall germination rates, was repeated three times independently. Therefore, the evaluated germination traits which we examined included the fresh radicle weight (RW), the fresh radicle length (RL), and the germination rate [GR% = (number of germinated seeds/total seed number used in the test) × 100]. We used the data for these traits to generate indices that were defined as the ratio of the germination-related traits (RW, RL, and GR) under drought conditions to the same traits under the control drought-free condition: DT-RW, DT-RL, and DT-GR.

Phenotypic data analysis

For the three germination-related traits (i.e., RW, RL, and GR), phenotypic effect was partitioned into overall mean, genotypic effect (Geno), treatment effect (Treatment), genotype-by-treatment effect (Geno × Treat), replication within treatment effect (Block/Treat), and random error effect in the ANOVA model. Since the three drought-tolerance indices (i.e., DT-RW, DT-RL, and DT-GR) were derived from the two respective treatments, phenotypic effect of each DT-RW, DT-RL, and DT-GR was partitioned into overall mean, genotypic effect (Geno), replication effect, and random error effect in the ANOVA model. Once the linear model of ANOVA is defined, total degree of freedom and total sum square can be partitioned into the components defined in the linear model, from which mean square (MS) of each source of variation can be calculated, and the significance test can be conducted. To calculate the heritability in broad sense (H^2) for DT-RW, DT-RL, and DT-GR, the genetic variance $\left(\sigma^2_{\text{Geno}}\right)$ and error variance $\left(\sigma^2_{\varepsilon}\right)$ can be estimated from the theoretical expectation of MS, and the heritability in broad sense can be estimated from the following equation:

$$H^2 = \frac{\sigma^2_{\text{Geno}}}{\sigma^2_{\text{Geno}} + \sigma^2_{\varepsilon}}$$

The method described above was implemented in tool "ANOVA" in the QTL IciMapping software (Meng et al. 2015).

Genotyping and genotypic data analyses

In general, the genotyping and genotypic data analyses followed the same pipeline as in Liu et al. (2017). DNA samples were extracted from soybean seedlings leaves following a previously described method (Kisha et al. 1997; Liu et al. 2017). The population was genotyped using an Illumina SoySNP6k iSelect BeadChip; this contains 5361 SNPs (Akond et al. 2013). Chromosomal distributions and quality control of these SNPs followed a previous study (Wen et al. 2014). Powermarker 3.25 was used to assess minor allele frequencies (MAFs), polymorphic information content (PIC), heterozygosity, and gene diversity (Liu and Muse 2005). A subset of 4616 SNPs (with missing rates lower than 0.25 and MAF higher than 0.05) was selected from amongst the 5361 SNPs for use in the following analyses.

Multivariate analyses, including principal component analysis (PCA), model-based population structure analysis, and cluster analysis with a neighbor-joining algorithm, were used to cluster the soybean lines of the population into subgroups; the PCA and cluster analysis were implemented in TASSEL 5.0. The Bayesian model-based program STRUCTURE 2.3 (Pritchard et al. 2000) was also used to infer the population structure; this assigned the 259 genotypes into subpopulations based on the aforementioned 4616 polymorphic SNP markers. Ten independent analyses were performed, with each including 100,000 burn-ins and 100,000 MCMC replications. We considered ten different hypothetical numbers of subpopulations (k) (1–10). The number of subpopulations was declared when Δk reached its highest value (an ad hoc statistic; Evanno et al. 2005).

Using Arlequin V3.11, analysis of molecular variance (AMOVA) (Excoffier et al. 1992) and F-statistics (F_{ST}) analysis were performed (Excoffier et al. 2005) to investigate the population differentiations among the subpopulations that were identified using STRUCTURE. For assessing LD, r^2 was calculated using TASSEL 5.0 (Bradbury et al. 2007); the length of an LD block was determined when the decay distance of LD reached $r^2 = 0.1$.

Genome-wide association analysis

For marker–trait associations, a mixed-linear regression model controlling both population structure and kinship matrices (denoted as $Q + K$ model) was adopted in TASSEL 5.0. The kinship coefficient between each accession pair was estimated in TASSEL 5.0 using the Loiselle algorithm (Loiselle et al. 1995). Markers were defined as being significantly associated with the trait by losing the Bonferroni

correction as 2.6 to balance the rates of false positives and false negatives. If many SNPs were in one LD block with the chromosome segment controlling the trait of interest, they all could be called as QTL. In practice, only the SNP having the highest $-\log_{10}(P)$ in the chromosome region was declared as a QTL. In this study, LD decay distance ($r^2 = 0.1$) was used to determine LD blocks; and this was the support interval used to declare significant SNPs associated with a given target trait.

Results

Phenotypic variation and correlation analysis

Three germination-related traits, namely RL, RW, and GR, were measured in seeds for the 259 soybean accessions under 0 (C) or 22% PEG6000 (D). The means, ranges, standard deviations, and coefficients of variation for these traits are shown in Table S2. The means of RL, RW, and GR values for the seeds treated with 22% PEG6000 were 1.5200 cm, 0.0260 g, and 35.6550%, while the means for the controls were of 22.2885 cm, 0.7612 g, and 97.5167%, respectively. Obviously, all of these traits were substantially higher under normal conditions than under drought-stressed conditions. ANOVA indicated significant ($P < 0.001$) variation among genotypes, treatments, and genotype-by-treatment interactions for all three measured traits among the cultivars of the population (Table 1), and genotype-by-treatment interaction had a much smaller variance compared to genotype, indicating that genotypic variation was the major player in the observed phenotypic variation for the three traits. The mean values for the calculated indices (DT-RL, DT-RW, and DT-GR) were 0.0683, 0.0348, and 0.3624 (Table 2), respectively, with significant ($P < 0.001$) variation for all three (Table 3). Given the substantial variation for the distributions of DT-RL, DT-RW, and DT-GR among the cultivars of the population (Fig. 1), the three drought-tolerance indices which we calculated are apparently quantitative traits that are likely controlled by multiple genes.

Table 2 Descriptive statistics of three drought-tolerance indices for the 259 soybean cultivars

Traits	Minimum	Maximum	Mean	SD	CV
DT-RL[a]	0	0.2286	0.0683	0.0474	69.34
DT-RW[b]	0	0.0935	0.0348	0.0213	61.27
DT-GR[c]	0	1.0000	0.3624	0.2681	73.97

SD standard deviation, *CV* coefficient of variation

[a]DT-RL was defined as the ratio of radical length under drought stress conditions and radical length under no-drought stress conditions

[b]DT-RW was defined as the ratio of fresh radical weight under drought stress conditions and fresh radical weight under no-drought stress conditions

[c]DT-GR was defined as the ratio of germination rate under drought stress conditions and germination rate under no-drought stress conditions

Pearson's correlations between the three drought-tolerance indices were analyzed based on the means for the 259 accessions (Table 4). DT-GR was found to be positively correlated with the DT-RL (r of 0.85; $P < 0.001$) and DT-RW (r of 0.79; $P < 0.001$). DT-RL and DT-RW were also positively correlated (r of 0.91; $P < 0.001$). Furthermore, the broad sense heritabilities for DT-RL, DT-RW, and DT-GR were 85.69%, 84.59%, and 92.33%, respectively.

Genetic diversity of the 259 accessions based on 4616 SNPs

Of the 4,616 SNPs of the selected subset from the chip, the mean values for MAF, gene diversity, heterozygosity,

Table 1 Analysis of variance (ANOVA) of three germination-related traits under 0% PEG and 22% PEG conditions for seeds of the 259 soybean cultivars

Trait	Source	DF[d]	Sum of square	Mean square	*F* value	Pr > *F*
RL	Geno	258	4230.71	16.40	17.26	< 0.0001
	Treatment	1	167,456.56	167,456.56	176,248.47	< 0.0001
	Block/treat	4	16.69	4.17	4.39	0.0016
	Geno × treat	258	3280.25	12.71	13.38	< 0.0001
RW	Geno	258	6.16	0.02	13.90	< 0.0001
	Treatment	1	209.49	209.49	121,920.35	< 0.0001
	Block/treat	4	0.50	0.12	72.25	< 0.0001
	Geno × treat	258	5.81	0.02	13.11	< 0.0001
GR	Geno	258	321,549.25	1246.31	39.65	< 0.0001
	Treatment	1	1,483,266.88	1,483,266.88	47,189.53	< 0.0001
	Block/treat	4	638.88	159.72	5.08	0.0005
	Geno × treat	258	261,366.23	1013.05	32.23	< 0.0001

RL radical length, RW fresh radical weight, GR germination rate, *DF* degree of freedom

Table 3 Analysis of variance (ANOVA) of three drought-tolerance indices in the 259 soybean cultivars

Traits	Source	DF	Sum of square	Mean square	*F* value	Pr > *F*
DT-RL[a]	Geno	258	1.7360	0.0067	18.94	< 0.0001
	Replication	2	0.0029	0.0015	4.08	0.0174
DT-RW[b]	Geno	258	0.3494	0.0014	17.38	< 0.0001
	Replication	2	0.0011	0.0005	6.74	0.0013
DT-GR[c]	Geno	258	55.5808	0.2154	36.99	< 0.0001
	Replication	2	0.0702	0.0351	6.02	0.0026

DF degree of freedom

[a]DT-RL was defined as the ratio of radical length under drought stress conditions and radical length under no-drought stress conditions

[b]DT-RW was defined as the ratio of fresh radical weight under drought stress conditions and fresh radical weight under no-drought stress conditions

[c]DT-GR was defined as the ratio of germination rate under drought stress conditions and germination rate under no-drought stress conditions

and PIC were 0.2773, 0.3649, 0.0751, and 0.2904, respectively (Table S3, Fig. S1). The mean LD of the whole genome was $r^2 = 0.2440$. When the LD decay distance was around 2,000 kb, r^2 decreased to half of its maximum value; and when the decay distance was about 10,500 kb, r^2 was below 0.1 (Fig. S2); these results suggest relatively slow LD decay in this population. The mean marker density was 204.27 kb for the whole population and, therefore, was assumed to have adequate power to detect major QTL-related germination drought traits in the association panel.

Consideration of Δk values obtained in STRUCTURE analysis showed that $k = 3$ was the optimal population subdivision, indicating that the accessions formed three major clusters (Fig. 2a, b). To validate and gain further insight into the genetic diversity of the soybean germplasm panel, we constructed a neighbor-joining tree based on the frequency of shared alleles among the accessions. Consistent with the results from STRUCTURE analysis and the geographical origins, the 259 soybean accessions were classified into three major groups (Fig. 2c). One group was composed largely of accessions from the Huang-huai-hai valley region (41 out of 48); one was composed largely of accessions from Heilongjiang province (107 out of 122); and the third group was mainly composed of accessions from Jilin province (76 out of 89) (Table S4).

PCA has been proposed as an alternative to population structure analysis for studying population stratification from genotypic data (Patterson et al. 2006). A PCA of the entire set of 259 accessions with the 4616 SNPs (Fig. 2d) also showed a clear separation of the same three groups that were identified in the STRUCTURE and neighbor-joining tree analyses. The significant pairwise F_{ST} values (Table S5) found between the three subpopulations (0.21, $P < 0.001$) further confirmed the existence of three subpopulations for the 259 accessions. Therefore, a Q matrix with $k = 3$ was used in the subsequent genome-wide association analysis.

QTLs for the three drought-tolerance indices

Fifteen QTLs, distributed in 13 soybean chromosomes were found to be significantly associated with the three drought-tolerance indices, and each of these explained (on average) 5.81% of the phenotypic variation (Table 5, Figs. 3 and 4). There were two QTLs on each of chromosomes 12, and 20. Among the 15 detected QTLs some were associated with more than one trait: 11 were associated with DT-GR; 3 QTLs were associated with DT-RL; and 4 QTLs were associated with DT-RW (Table 5). The QTL located around 16,289,406-bp on chromosome 11 (ss247449682) was commonly associated with DT-GR and DT-RW. The QTL located around 39,325,657-bp on chromosome 17 (ss249472124) was commonly associated with DT-GR and DT-RL. The QTL located around 46,546,046-bp on chromosome 20 (ss250606162) was commonly associated with DT-RW and DT-RL. These three QTLs highlight the aforementioned strong positive correlations between these drought-tolerance indices (Table 4). We consider the potential functional information related to these QTLs below, in the discussion. Notably, 11 of the QTLs we identified are located within 500 kb of loci reported in previous quantitative genetics analyses in soybean, with 5 of them having been previously associated with drought-related traits (Table 5).

Discussion

Seed germination is obviously an essential stage in the growth cycle of most seed plants. The amount of water needed for the germination stage is much smaller than

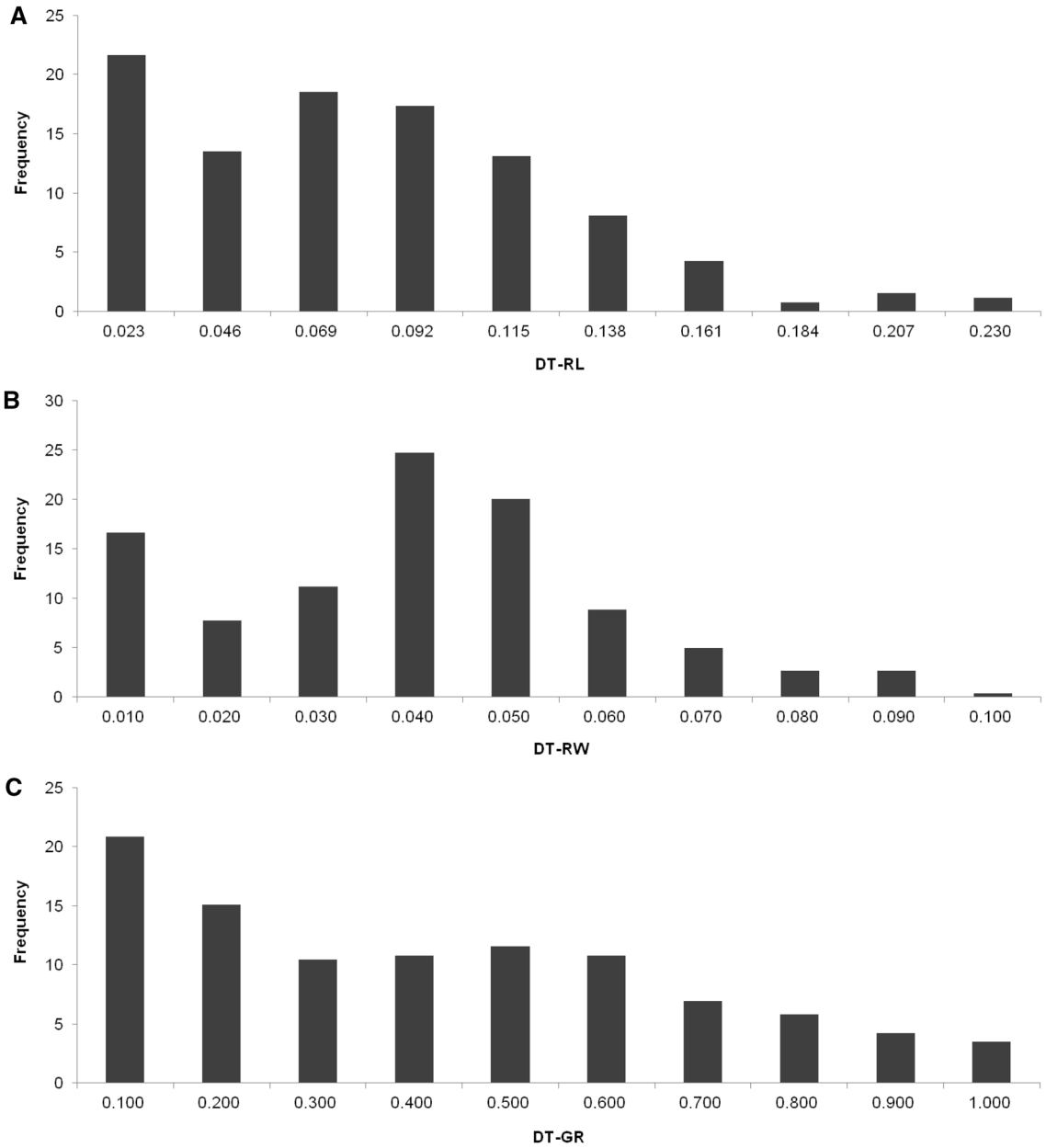

Fig. 1 Frequency distributions of three drought-tolerance indices DT-RL (**a**), DT-RW (**b**), and DT-GR (**c**) in the 259 soybean cultivars

amount needed over the whole soybean life cycle, but it is now understood that water deficits in the germination stage of soybean can cause severe effects on overall growth. Drought occurs frequently in the three main soybean production areas of China, particularly during the sowing time for cultivars based on the Hsu and Nsp ecotypes, thus causing difficulty for sowing and germination. Currently, little is known regarding the genetics of drought tolerance in soybean at the germination stage. In the present study, the distribution of values for three drought-related phenotypic indices revealed substantial diversity in an association panel of 259 released cultivars (Fig. 1), and a genome-wide association analysis based on the 4,616 SNPs identified fifteen QTLs associated with drought tolerance (Table 5).

As noted in the results, our association mapping results for the drought-tolerance indices were compared with QTLs previously reported within a 500 kb vicinity using tools at Soybase (https://www.soybase.org). Five of the QTLs detected in present study, including ss246991464 on Chr. 9, ss247449682 on Chr. 11, ss248468442 on Chr. 14,

Table 4 Phenotypic correlations between drought stress indices based on the means of the traits in the 259 soybean cultivars

Trait	DT-RW[b]	DT-GR[c]
DT-RL[a]	0.9138***	0.8542***
DT-RW[b]		0.7917***

***Significant level under 0.0001 for Pearson correlation test

[a]DT-RL was defined as the ratio of radical length under drought stress conditions and radical length under no-drought stress conditions

[b]DT-RW was defined as the ratio of fresh radical weight under drought stress conditions and fresh radical weight under no-drought stress conditions

[c]DT-GR was defined as the ratio of germination rate under drought stress conditions and germination rate under no-drought stress conditions

ss250269376 on Chr. 19, and ss250606162 on Chr. 20, were located in regions (Table 5) that have been previously associated with the genetic control of delayed canopy wilting (Abdel-Haleem et al. 2012; Hwang et al. 2016; Kaler et al. 2017).

Of particular note, the region which contained our QTL for the SNP marker ss250269376 has been identified in three previously reported studies (Abdel-Haleem et al. 2012; Hwang et al. 2016; Kaler et al. 2017); the SNP that we detected is 3090 bp away from the SNP (ss715635661) detected by Kaler et al. (2017) that was associated with delayed canopy wilting. In terms of our association panel, it was notable that cultivars with allele G for the ss250269376 position had significantly higher DT-GR values than those with allele A across all accessions (P value < 0.01) (Fig. 5a). Further analysis indicated that the distribution of allele G was also uneven across the three geographical subpopulations; the frequency of the favorable allele G was only 5.98% in SG 2 (mainly from Heilongjiang province), 12.35% in SG 3 (mainly from Jilin province), but rose to fully 82.93% in SG 1 (mainly from Hsu region) (Fig. 5b). Seed hardness is known to be closely related with drought tolerance (Liu 2009). Thus, it was interesting that the region near

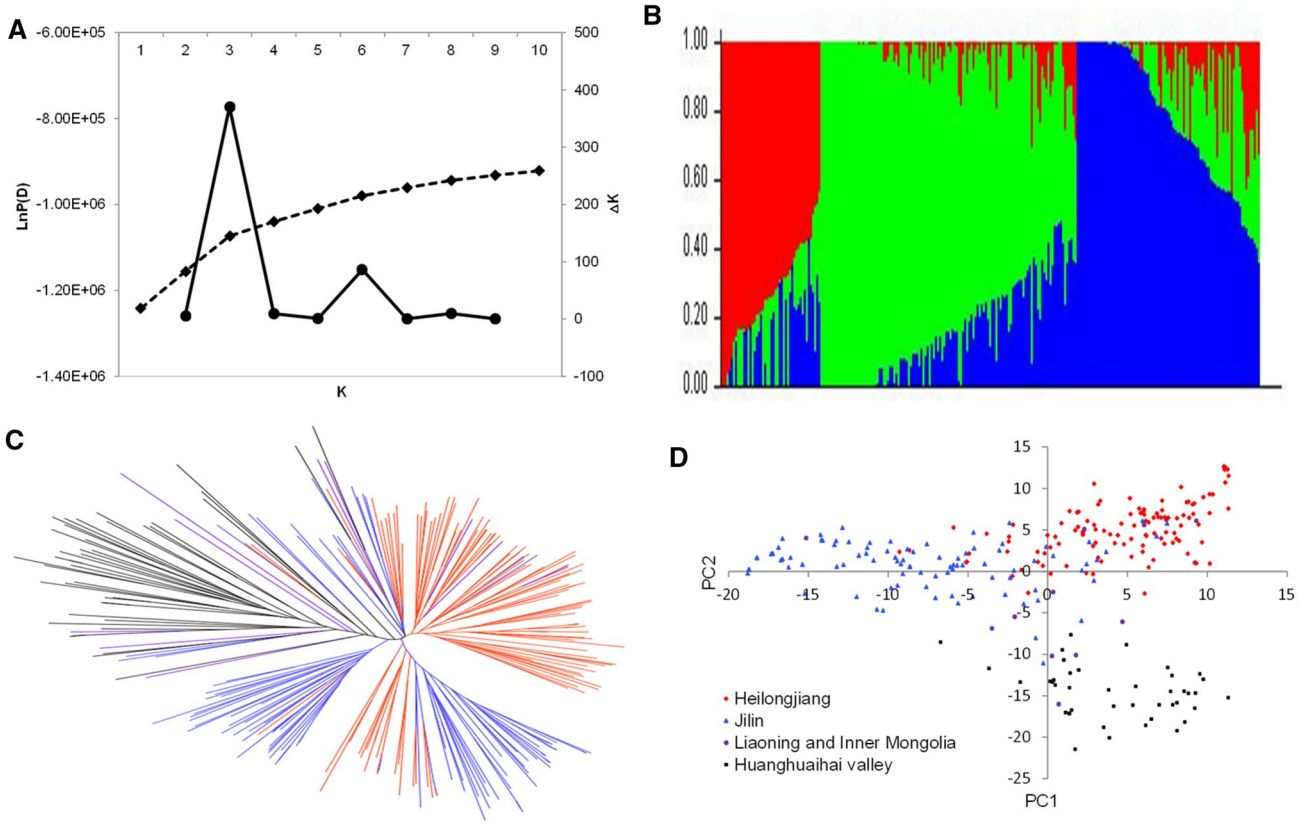

Fig. 2 Genetic structure and relatedness of the 259 soybean cultivars. **a** Evolution of the natural logarithm probability of the data against K and the magnitude of Δk for each K value; **b** Clustering for $K = 3$ for the entire panel of soybean cultivars. Each individual is represented by a vertical bar, partitioned into colored segments with the length of each segment representing the proportion of the individual's genome from groups when $k = 3$; **c** Neighbor-joining tree constructed using SNP data; soybean accessions from Heilongjiang are shown in red; those from Jilin are shown in blue; those from Liaoning and Inner Mongolia are shown in purple; and those from the Huanghuaihai valley region are shown in black. **d** Principal component analysis for the entire panel of soybean cultivars

Table 5 SNPs significantly associated with the three calculated drought-tolerance indices

Marker	Chr	Position	Associated traits (R^2%)	$-\log_{10}(P)$	Alleles and allelic effect	Reported QTLs/genes
ss245388356	4	16,902,743	DT-RW (4.99)	2.6990	C (0.0476)	
ss245627013	5	8,209,039	DT-GR (7.02)	3.3790	T (0.1404)	
ss245905576	6	16,437,297	DT-GR (4.97)	2.6383	C (0.5007)	Seed yield (Du et al. 2009a), seed protein (Mao et al. 2013), Stem strength (Chen et al. 2011)
ss246144068	7	746,249	DT-RW (5.34)	2.8268	C (0.0085)	Seed sucrose (Maughan et al. 2000), plant height (Yao et al. 2015)
ss246991464	9	38,525,177	DT-RL (5.30)	2.7670	G (0.0263)	Canopy wilt (Kaler et al. 2017)
ss247449682	11	16,289,406	DT-GR (5.48), DT-RW (4.80)	2.9469	T (0.5326); T (0.0450)	Canopy wilt (Kaler et al. 2017)
ss247772178	12	31,934,231	DT-GR (4.91)	2.6326	A (0.5698)	
ss247780010	12	33,115,103	DT-GR (5.57)	2.9208	G (0.0752)	
ss248097753	13	31,586,157	DT-GR (6.25)	3.3443	G (0.2081)	Seed weight (Hyten et al. 2004; Eskandari et al. 2013a; Teng et al. 2009), seed protein (Hyten et al. 2004; Eskandari et al. 2013a), seed oil (Eskandari et al. 2013b)
ss248468442	14	44,520,959	DT-GR (6.84)	3.6765	C (−0.1277)	Canopy wilt (Abdel-Haleem et al. 2012), seed weight (Liu et al. 2011), seed oil (Eskandari et al. 2013b)
ss249472124	17	39,325,657	DT-GR (6.72), DT-RL (5.18)	3.5559	C (0.2742); C (0.0390)	Seed yield (Kabelka et al. 2004), seed protein (Tajuddin et al. 2003), seed oil (Liang et al. 2010; Mao et al. 2013; Han et al. 2015)
ss249791040	18	47,660,490	DT-GR (5.23)	2.7144	T (−0.1576)	Pod number (Zhang et al. 2007), seed protein (Lu et al. 2012; Jun et al. 2008)
ss250269376	19	47,214,600	DT-GR (6.63)	3.4927	G (0.2933)	mqCanopy wilt (Abdel-Haleem et al. 2012; Hwang et al. 2016), canopy wilt (Kaler et al. 2017), seed weight (Csanadi et al. 2001; Funatsuki et al. 2005), seed number (Funatsuki et al. 2005), pod number (Yang et al. 2013), seed oil (Hyten et al. 2004), SCN (Guo et al. 2005), seed hardness (Zhang et al. 2008)
ss250603455	20	46,154,637	DT-GR (5.11)	2.7305	C (0.2772)	Seed yield (Yuan et al. 2002; Kabelka et al. 2004), Seed weight (Han et al. 2012)
ss250606162	20	46,546,046	DT-RL (6.35), DT-RW (7.88)	3.3085	A (0.0309); A (0.0154)	Canopy wilt (Kaler et al. 2017)

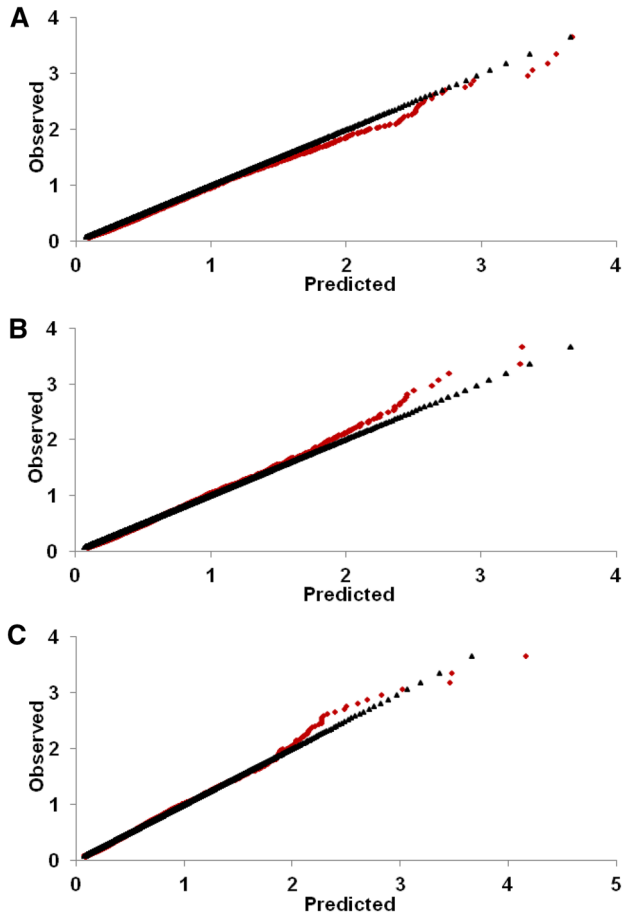

Fig. 3 Quantile–quantile plots from GWAS across the three drought-tolerance indices: DT-GR (**a**), DT-RL (**b**), and DT-RW (**c**)

ss250269376 detected in the present study was previously found to be associated with seed hardness (Zhang et al. 2008; Table 5).

It is interesting to note that the QTLs which we detected for the drought indices are genetically linked to markers associated with seed weight and yield-related traits on Chrs. 6, 13, 14, 17, 18, 19, and 20 (Table 5) (Csanadi et al. 2001; Du et al. 2009a; Eskandari et al. 2013a; Funatsuki et al. 2005; Han et al. 2012; Hyten et al. 2004; Kabelka et al. 2004; Liu et al. 2011; Teng et al. 2009; Yang et al. 2013; Yuan et al. 2002; Zhang et al. 2007). Furthermore, the genomic regions for six of our QTLs have been previously associated with protein content and oil content, including ss245905576 on Chr. 6, ss248097753 on Chr. 13, ss248468442 on Chr. 14, ss249472124 on Chr. 17, ss249791040 on Chr. 18, and ss250269376 on 19 (Eskandari et al. 2013a, b; Han et al. 2015; Hyten et al. 2004; Jun et al. 2008; Liang et al. 2010; Lu et al. 2012; Mao et al. 2013; Tajuddin et al. 2003). One possible implication of these findings is that drought stress signaling may influence protein and oil concentrations of seeds (Dornbos and Mullen 1992) or, alternatively, that the genes controlling these traits may simply be located in similar chromosomal regions. Identification of more trade-associated markers in such regions could help to delineate the specific functional contributions of particular loci. It is even possible that marker-assisted selection could be used to simultaneously improve both drought tolerance and protein or oil content traits. The other four drought-associated QTLs that we identified have not been reported in the previous studies.

The soybase database (https://www.soybase.org/) was used to identify candidate genes hit directly by the SNPs of our detected QTLs or in nearby genes; a total of 22 candidate genes were thusly identified (Table S6). Of note, none of the SNPs was direct hits in the exons of genes; all 15 of the SNPs were located in intergenic regions or in regulatory regions. Functional annotations for the 22 genes were obtained from Soybase, which revealed that our candidate genes potentially encode transcription factors, protein kinases, auxin response factors and transporters, and proteins relating to the E3 ubiquitin system. Notably, three genes (*Glyma.04G126000*, *Glyma.12G174000,* and *Glyma.11G167100*) are apparently homologs of genes from other plants that have serine/threonine-protein kinase or phosphatase functions related to tolerance stress in potato (Bai et al. 2017), Arabidopsis (Huang et al. 2018), wheat (Ghorbel et al. 2017), rice (Liao et al. 2016), and *Vigna aconitifolia* (Tiwari et al. 2018). Such genes are known to be regulated by the well-characterized drought-related plant hormone ABA (Hou et al. 2016). Work in other plant species has also shown that ABA impacts drought-related plant regulatory networks that includes the type of transcription factors putatively encoded by our hits *Glyma.06g188400* and *Glyma.17g239700* (MYB), and *Glyma.20g231800* (bHLH99) (Chander et al. 2018; Chen et al. 2015; Gao et al. 2016; Park et al. 2018; Takahashi et al. 2016; Wang et al. 2017; Xu et al. 2018; Zhao et al. 2018). Finally, we were also intrigued by a putative E3 ubiquitin-protein ligase gene among our hits, *Glyma.20g231900*. Many previous studies have reported drought-tolerance-related functions for E3 ligase proteins (Ding et al. 2015; Serrano et al. 2018).

In conclusion, we here used GWAS to examine a germplasm association panel comprising 259 soybean released Chinese cultivars for drought-associated traits based on phenotypic data from drought-treated germinating soybean seeds. The analysis was based on a total of 4,616 SNPs, and 15 SNP-trait associations were identified by GWAS, among which three SNPs were significantly associated with two of the drought-tolerance indices. Practically, subsequent applied research based on our study research will

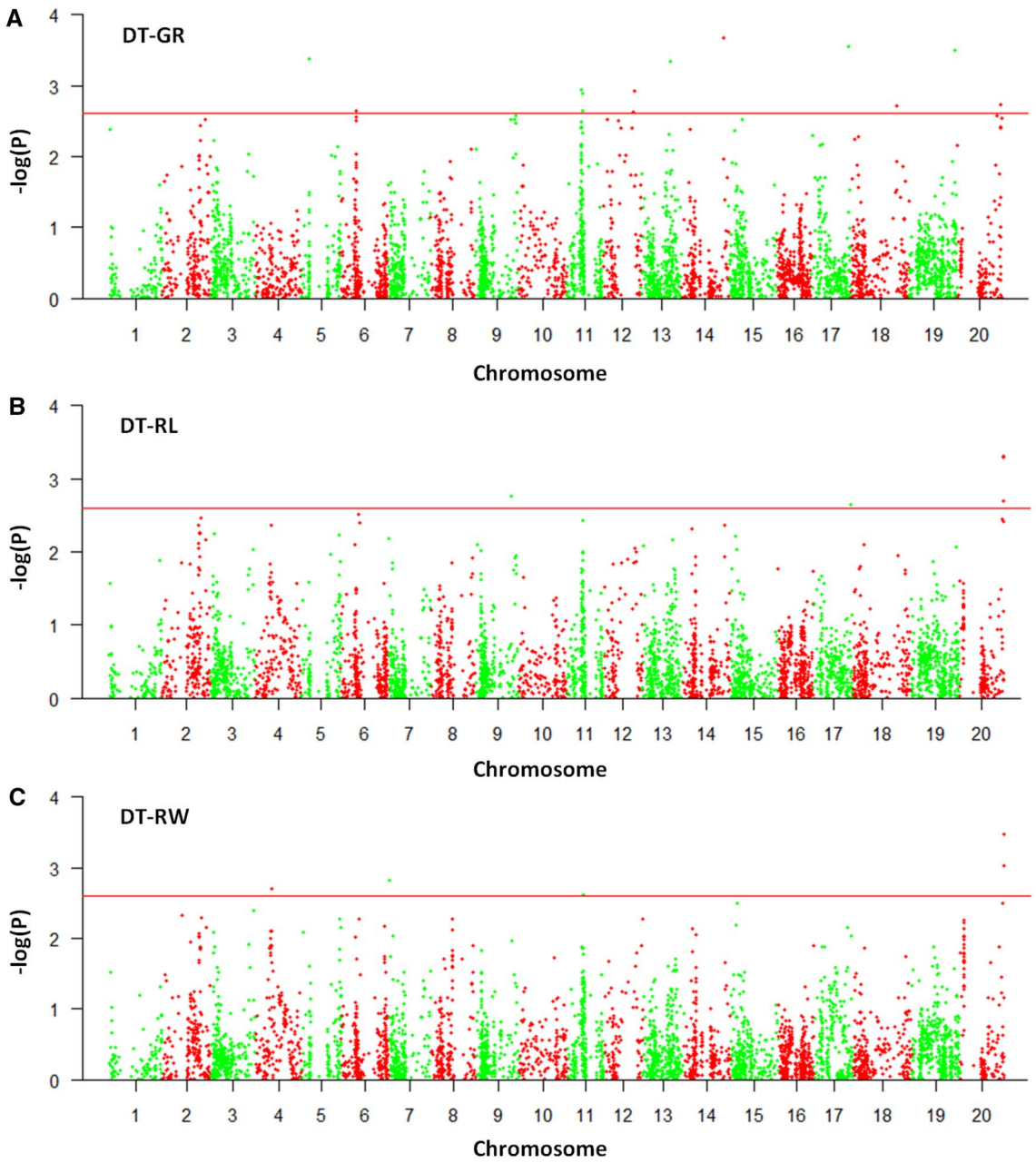

Fig. 4 Manhattan plot from the $Q+K$ model across the three drought-tolerance indices: DT-GR (**a**), DT-RL (**b**), and DT-RW (**c**)

seek to confirm the functional contributions of the identified QTLs in different genetic backgrounds and in different environments. Scientifically, the loci detected in the present study lay the foundation for deepening our understanding of the genetic basis of soybean drought tolerance during the germination stage of plant growth. Collectively, our work will facilitate future molecular breeding of soybean for improved drought tolerance, which will help to ensure stable high yields in increasingly variable climatological conditions.

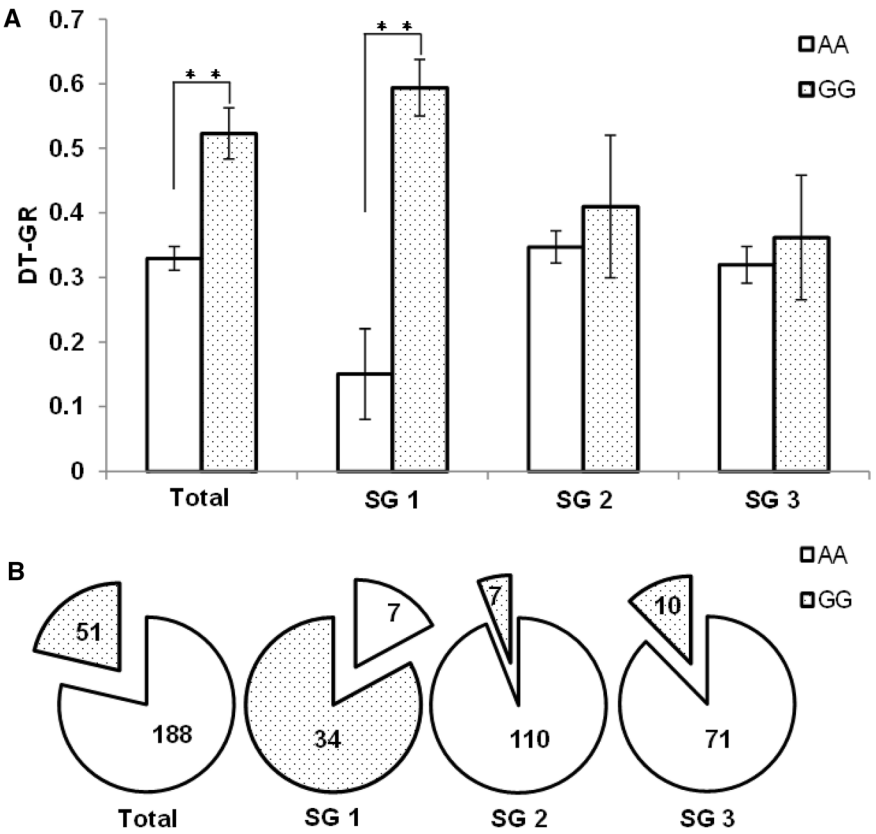

Fig. 5 Diagrams depicting the genetic effects of bi-allele variation at position ss250269376 in relation to **a** DT-GR across all accessions and geographical subpopulations inferred by STRUCTURE, **b** allele frequencies in geographical subpopulations inferred by STRUCTURE. SG 1, 2, and 3 are structure groups 1, 2, and 3, inferred by STRUCTURE analysis. *P* values of 0.01 are shown as **

Acknowledgements This work was supported by the National Key R & D Program for Crop Breeding (Grant No. 2016YFD0100304), the Development of Novel Elite Soybean Cultivars and Lines with High Oil Content (Grant No. Z161100000916005-06), the Crop Germplasm Resources Protection (Grant No. 2017NWB036-5), the Platform of National Crop Germplasm Resources of China (Grant Nos. 2018-004 and 2017-004), the Agricultural Science and Technology Innovation Program of the Chinese Academy of Agricultural Sciences (CAAS), and the Improvement of Soybean Abiotic Stress Tolerance to Address the Climate Change (Grant No. PJ0121092018).

Author contributions ZL, XQ, and LQ conceived and designed the experiments. ZL, ZG, YZ, XW, and HR performed the experiments. ZL, HL, ZW, BKW, YL, LY, HG, DW, XQ, and LQ analyzed data and wrote the manuscript. All authors read and approved the manuscript.

Compliance with ethical standards

Conflict of interest The authors have declared that no competing interests exist.

Ethical approval This article does not contain any studies with human participants or animals performed by any of authors.

References

Abdel-Haleem H, Lee GJ, Boerma RH (2011) Identification of QTL for increased fibrous roots in soybean. Theor Appl Genet 122:935–946

Abdel-Haleem H, Carter TE, Purcell LC, King CA, Ries LL, Chen P, Schapaugh W, Sinclair TR, Boerma HR (2012) Mapping of quantitative trait loci for canopy-wilting trait in soybean (*Glycine max* L. Merr). Theor Appl Genet 125:837–846

Akond M, Liu S, Schoener L, Anderson JA, Kantartzi SK, Meksem K, Song Q, Wang D, Wen Z, Lightfoot DA, Kassem MA (2013) SNP-based genetic linkage map of soybean using the SoySNP6K Illumina Infinium BeadChip genotyping array. J Plant Genome Sci 1:80–89

Bai J, Mao J, Yang H, Khan A, Fan A, Liu S, Zhang J, Wang D, Gao H, Zhang J (2017) Sucrose non-ferment 1 related protein kinase 2 (SnRK 2) genes could mediate the stress responses in potato (*Solanum tuberosum* L.). BMC Genet 18:41

Boyer JS (1982) Plant productivity and environment. Science 218:443–448

Bradbury PJ, Zhang Z, Kroon DE, Casstevens TM, Ramdoss Y, Buckler ES (2007) TASSEL: Software for association mapping of complex traits in diverse samples. Bioinformatics 23:2633–2635

Cattivelli L, Rizza F, Badeck FW, Mazzucotelli E, Mastrangelo AM, Francia E, Mare C, Tondelli A, Stanca AM (2008) Drought tolerance improvement in crop plants: an integrated view from breeding to genomics. Field Crop Res 115:1–14

Chander S, Almeida DM, Serra TS, Jardim-Messeder D, Barros PM, Lourenço TF, Figueiredo DD, Margis-Pinheiro M, Costa JM, Oliveira MM, Saibo NJ (2018) *OsICE1* transcription factor improves photosynthetic performance and reduces grain losses in rice plants subjected to drought. Environ Exp Bot 150:88–98

Charlson D, Bhatnagar S, King C, Ray J, Sneller C, Carter T, Purcell L (2009) Polygenic inheritance of canopy wilting in soybean [*Glycine max* (L.) Merr.]. Theor Appl Genet 119:587–594

Chen H, Shan Z, Sha A, Wu B, Yang Z, Chen S, Zhou R, Zhou X (2011) Quantitative trait loci analysis of stem strength and related traits in soybean. Euphytica 179:485–497

Chen T, Li W, Hu X, Guo J, Liu A, Zhang B (2015) A cotton MYB transcription factor, *GbMYB5*, is positively involved in plant adaptive response to drought stress. Plant Cell Physiol 56:917–929

Condon AG, Richards RA, Rebetzke GJ, Farquhar GD (2004) Breeding for high water-use efficiency. J Exp Bot 55:2447–2460

Csanadi G, Vollmann J, Stift G, Lelley T (2001) Seed quality QTLs identified in a molecular map of early maturing soybean. Theor Appl Genet 103:912–919

Dias FG, Borges ACN, Viana AAB, Mesquita RO, Romano E, Grossi de Sa MF, Nepomuceno AL, Loureiro ME, Ferreira MA (2012) Expression analysis in response to drought stress in soybean: shedding light on the regulation of metabolic pathway genes. Genet Mol Biol 35:222–232

Ding S, Zhang B, Qin F (2015) Arabidopsis *RZFP34/CHYR1*, a ubiquitin E3 ligase, regulates stomatal movement and drought tolerance via SnRK2. 6-mediated phosphorylation. Plant Cell 27:3228–3244

Dogan E, Kirnak H, Copur O (2007) Deficit irrigations during soybean reproductive stages and CROPGRO-soybean simulations under semi-arid climatic conditions. Field Crops Res 103:154–159

Dornbos DL, Mullen RE (1992) Soybean seed protein and oil contents and fatty acid composition adjustments by drought and temperature. J Am Oil Chem Soc 69:228–231

Du W, Fu S, Yu D (2009a) Genetic analysis for the leaf pubescence density and water status traits in soybean [*Glycine max* (L.) Merr.]. Plant Breed 128:259–265

Du W, Wang M, Fu S, Yu D (2009b) Mapping QTL for seed yield and drought susceptibility index in soybean (*Glycine max* L.) across different environments. J Genet Genomics 36:721–731

Du W, Yu D, Fu S (2009c) Detection of quantitative trait loci for yield and drought tolerance traits in soybean using a recombinant inbred line population. J Integr Plant Biol 51:868–878

Eskandari M, Cober E, Rajcan I (2013a) Genetic control of soybean seed oil: II. QTL and genes that increase oil concentration without decreasing protein or with increased seed yield. Theor Appl Genet 126:1677–1687

Eskandari M, Cober E, Rajcan I (2013b) Genetic control of soybean seed oil: I. QTL and genes associated with seed oil concentration in RIL populations derived from crossing moderately high-oil parents. Theor Appl Genet 126:403–495

Evanno G, Regnaut S, Goudet J (2005) Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. Mol Ecol 14:2611–2620

Excoffier L, Smouse PE, Quattro JM (1992) Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. Genetics 131:479–491

Excoffier L, Laval G, Schneider S (2005) Arlequin (version 3.0): an integrated software package for population genetics data analysis. Evol Bioinform 1:47–50

Frederick JR, Camp CR, Bauer PJ (2001) Drought-stress effects on branch and mainstem seed yield and yield components of determinate soybean. Crop Sci 41:759–763

Funatsuki H, Kawaguchi K, Matsuba S, Sato Y, Ishimoto M (2005) Mapping of QTL associated with chilling tolerance during reproductive growth in soybean. Theor Appl Genet 111:851–861

Gao F, Yao H, Zhao H, Zhou J, Luo X, Huang Y, Li C, Chen H, Wu Q (2016) Tartary buckwheat *FtMYB10* encodes an R2R3-MYB transcription factor that acts as a novel negative regulator of salt and drought response in transgenic *Arabidopsis*. Plant Physiol Bioch 109:387–396

Ghorbel M, Cotelle V, Ebel C, Zaidi I, Ormancey M, Galaud JP, Hanin M (2017) Regulation of the wheat MAP kinase phosphatase 1 by 14-3-3 proteins. Plant Sci 257:37–47

Guo B, Sleper DA, Arelli PR, Shannon JG, Nguyen HT (2005) Identification of QTLs associated with resistance to soybean cyst nematode races 2, 3 and 5 in soybean PI 90763. Theor Appl Genet 103:1167–1173

Han Y, Li D, Zhu D, Li H, Li X, Teng W, Li W (2012) QTL analysis of soybean seed weight across multi-genetic backgrounds and environments. Theor Appl Genet 125:671–683

Han Y, Teng W, Wang Y, Zhao X, Wu L, Li D, Li W (2015) Unconditional and conditional QTL underlying the genetic interrelationships between soybean seed isoflavone, and protein or oil contents. Plant Breed 134:300–309

Hou YJ, Zhu Y, Wang P, Zhao Y, Xie S, Batelli G, Wang B, Duan CG, Wang X, Xing L, Lei M, Yan J, Zhu X, Zhu JK (2016) Type one protein phosphatase 1 and its regulatory protein inhibitor 2 negatively regulate ABA signaling. PLoS Genet 12:e1005835

Huang X, Hou L, Meng J, You H, Li Z, Gong Z, Yang S, Shi Y (2018) The antagonistic action of abscisic acid and cytokinin signaling mediates drought stress response in *Arabidopsis*. Mol Plant 11:970–982

Hufstetler EV, Boerma HR, Carter TE, Earl HJ (2007) Genotypic variation for three physiological traits affecting drought tolerance in soybean. Crop Sci 47:25–35

Hwang S, King CA, Ray JD, Cregan PB, Chen P, Carter TE, Li Z, Abdel-Haleem H, Matson KW, Schapaugh W, Purcell LC (2015) Confirmation of delayed canopy wilting QTLs from multiple soybean mapping populations. Theor Appl Genet 128:2047–2065

Hwang S, King CA, Chen P, Ray JD, Cregan PB, Carter TE, Li Z, Abdel-Haleem H, Matson KW, Schapaugh W, Purcell LC (2016) Meta-analysis to refine map position and reduce confidence intervals for delayed-canopy-wilting QTLs in soybean. Mol Breeding 36:91

Hyten DL, Pantalone VR, Sams CE, Saxton AM, Landau-Ellis D, Stefaniak TR, Schmidt ME (2004) Seed quality QTL in a prominent soybean population. Theor Appl Genet 109:552–561

Jun TH, Van K, Kim MY, Lee SH, Walker DR (2008) Association analysis using SSR markers to find QTL for seed protein content in soybean. Euphytica 162:179–191

Kabelka EA, Diers BW, Fehr WR, LeRoy AR, Baianu IC, You T, Neece DJ, Nelson RL (2004) Putative alleles for increased yield from soybean plant introductions. Crop Sci 44:784–791

Kaler AS, Ray JD, Schapaugh WT, King CA, Purcell LC (2017) Genome-wide association mapping of canopy wilting in diverse soybean genotypes. Theor Appl Genet 130:2203–2217

Kisha TJ, Sneller CH, Diers BW (1997) Relationship between genetic distance among parents and genetic variance in populations of soybean. Crop Sci 37:1317–1325

Liang H, Yu Y, Wang S, Lian Y, Wang T, Wei Y, Gong P, Liu X, Fang X, Zhang M (2010) QTL mapping of isoflavone, oil and protein contents in soybean (*Glycine max* L. Merr.). Ag Sci China 9:1108–1116

Liao YD, Lin KH, Chen CC, Chiang CM (2016) *Oryza sativa* protein phosphatase 1a (OsPP1a) involved in salt stress tolerance in transgenic rice. Mol Breeding 36:22

Liu X (2009) Drought. In: Lam HM et al (ed) Research on tolerance to stresses in Chinese soyben. China Agriculture Press, Beijing

Liu KJ, Muse SV (2005) PowerMarker: an integrated analysis environment for genetic marker analysis. Bioinformatics 21:2128–2129

Liu W, Kim M, Van K, Lee Y, Li H, Liu X, Lee S (2011) QTL identification of yield-related traits and their association with flowering and maturity in soybean. J Crop Sci Biotech 14:65–70

Liu Z, Li H, Wen Z, Fan X, Li Y, Guan R, Guo Y, Wang S, Wang D, Qiu LJ (2017) Comparison of genetic diversity between Chinese and American soybean (*Glycine max* (L.)) accessions revealed by high-density SNPs. Front Plant Sci 8:11. https://doi.org/10.3389/fpls.2017.02014

Loiselle BA, Sork VL, Nason J, Graham C (1995) Spatial genetic structure of a tropical understory shrub, *PSYCHOTRIA OFFICINALIS* (Rubiaceae). Am J Bot 82:1420–1425

Lu W, Wen Z, Li H, Yuan D, Li J, Zhang H, Huang Z, Cui S, Du W (2012) Identification of the quantitative trait loci (QTL) underlying water soluble protein content in soybean. Theor Appl Gen 126:425–433

Mao T, Jiang Z, Han Y, Teng W, Zhao X, Li W (2013) Identification of quantitative trait loci underlying seed protein and oil contents of soybean across multi-genetic backgrounds and environments. Plant Breed 132:630–641

Maughan PJ, Maroof MAS, Buss GR (2000) Identification of quantitative trait loci controlling sucrose content in soybean (*Glycine max*). Mol Breeding 6:105–111

Mederski HJ, Jeffers DL (1973) Yield response of soybean varieties grown at two soil moisture levels. Agron J 65:410–412

Meng L, Li H, Zhang L, Wang J (2015) QTL IciMapping: Integrated software for genetic linkage map construction and quantitative trait locus mapping in biparental populations. Crop J 3:269–283

Mian MAR, Bailey MA, Ashley DA, Wells R, Carter TE, Parrott WA, Boerma HR (1996) Molecular markers associated with water use efficiency and leaf ash in soybean. Crop Sci 36:1252–1257

Mian MAR, Ashley DA, Boerma HR (1998) An additional QTL for water use efficiency in soybean. Crop Sci 38:390–393

Mittler R, Blumwald E (2010) Genetic engineering for modern agriculture: challenges and perspectives. Annu Rev Plant Biol 61:443–462

Mohammadi PP, Moieni A, Hiraga S, Komatsu S (2012) Organ specific proteomic analysis of drought-stressed soybean seedlings. J Proteomics 75:1906–1923

Morison JIL, Baker NR, Mullineaux PM, Davies WJ (2008) Improving water use in crop production. Philos Trans R Soc Biol Sci 363:639–658

Oya T, Nepomuceno AL, Farias JRB, Tobita S, Ito O (2004) Drought tolerance characteristics of Brazilian soybean cultivars—evaluation and characterization of drought tolerance of various Brazilian soybean cultivars in the field. Plant Prod Sci 7:129–137

Park DY, Shim Y, Gi E, Lee BD, An G, Kang K, Paek NC (2018) The MYB-related transcription factor RADIALIS-LIKE3 (OsRL3) functions in ABA-induced leaf senescence and salt sensitivity in rice. Environ Exp Bot 156:86–95

Patterson NJ, Price AL, Reich D (2006) Population structure and eigenanalysis. PLoS Genet 2:e190

Pennisi E (2008) The blue revolution, drop by drop, gene by gene. Science 320:171–173

Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. Genetics 155:945–959

Ribaut JM, Jiang C, Gonzalez-de-Leon D, Edmeades GO, Hoisington DA (1997) Identification of quantitative trait loci under drought conditions in tropical maize. 2. Yield components and marker assisted selection strategies. Theor Appl Genet 94:887–896

Rönnberg-Wästljung AC, Glynn C, Weih M (2005) QTL analyses of drought tolerance and growth for a *Salix dasyclados* × *Salix viminalis* hybrid in contrasting water regimes. Theor Appl Genet 110:537–549

SAS Institute Inc (2007) SAS/STAT software: Version 9.1.3. SAS Institute, Cary, NC

Serrano I, Campos L, Rivas S (2018) Roles of E3 ubiquitin-ligases in nuclear protein homeostasis during plant stress responses. Front Plant Sci 9:139

Sharma KK, Lavanya M (2002) Recent developments in transgenics for abiotic stress in legumes of the semi-arid tropics. In: JIRCAS Working Report. pp 61–73

Shu Y, Zhou Y, Shi X, Hu N, Shao Q, Du J (2015) Screening of appropriate PEG-6000 concentration for the identification of soybean drought tolerance at germination stage. Soybean Sci 1:56–59

Specht JE, Williams JH, Weidenbenner CJ (1986) Differential responses of soybean genotypes subjected to a seasonal soil water gradient. Crop Sci 26:922–934

Specht JE, Hume DJ, Kumudini SV (1999) Soybean yield potential—a genetic and physiological perspective. Crop Sci 39:1560–1570

Specht JE, Chase K, Macrander M, Graef GL, Chung J, Markwell JP, Germann M, Orf JH, Lark KG (2001) Soybean response to water: a QTL analysis of drought tolerance. Crop Sci 41:493–509

Tajuddin T, Watanabe S, Yamanaka N, Harada K (2003) Analysis of quantitative trait loci for protein and lipid contents in soybean seeds using recombinant inbred lines. Breed Sci 53:133–140

Takahashi Y, Kinoshita T, Matsumoto M, Shimazaki KI (2016) Inhibition of the Arabidopsis bHLH transcription factor by monomerization through abscisic acid-induced phosphorylation. Plant J 87:559–567

Teng W, Han Y, Du Y, Sun D, Zhang Z, Qiu L, Sun G, Li W (2009) QTL analyses of seed weight during the development of soybean (Glycine max L Merr). Heredity 102:372380

Thabet SG, Moursi YS, Karam MA, Graner A, Alqudah AM (2018) Genetic basis of drought tolerance during seed germination in barley. PLoS ONE 13:e0206682

Tiwari B, Kalim S, Tyagi N, Kumari R, Bangar P, Barman P, Kumar S, Gaikwad A, Bhat KV (2018) Identification of genes associated with stress tolerance in moth bean [*Vigna aconitifolia* (Jacq.) Marechal], a stress hardy crop. Physiol Mol Biol Pla 24:551–561

Tuberosa R, Salvi S (2006) Genomics-based approaches to improve drought tolerance of crops. Trends Plant Sci 11:405–412

Wang N, Zhang W, Qin M, Li S, Qiao M, Liu Z, Xiang F (2017) Drought tolerance conferred in soybean (*Glycine max*. L) by GmMYB84, a novel R2R3-MYB transcription factor. Plant Cell Physiol 58:1764–1776

Wen Z, Tan R, Yuan J, Bales C, Du W, Zhang S, Chilvers MI, Schmidt C, Song Q, Cregan PB, Wang D (2014) Genome-wide association mapping of quantitative resistance to sudden death syndrome in soybean. BMC Genomics 15:809

Xiong LM, Wang RG, Mao GH, Koczan JM (2006) Identification of drought tolerance determinants by genetic analysis of root response to drought stress and abscisic acid. Plant Physiol 142:1065–1074

Xu Z, Ge Y, Zhang W, Zhao Y, Yang G (2018) The walnut *JrVHAG1* gene is involved in cadmium stress response through ABA-signal pathway and MYB transcription regulation. BMC Plant Biol 18:19

Yang Z, Xin D, Liu C, Jiang H, Han X, Sun Y, Qi Z, Hu G, Chen Q (2013) Identification of QTLs for seed and pod traits in soybean and analysis for additive effects and epistatic effects of QTLs among multiple environments. Mol Genet Genomics 288:651–667

Yao D, Liu Z, Zhang J, Liu S, Qu J, Guan S, Pan L, Wang D, Liu J, Wang P (2015) Analysis of quantitative trait loci for main plant traits in soybean. Genet Mol Res 14:6101–6109

Yuan J, Njiti VN, Meksem K, Iqbal MJ, Triwitayakorn K, Kassem MA, Davis GT, Schmidt ME, Lightfoot DA (2002) Quantitative trait loci in two soybean recombinant inbred line populations segregating for yield and disease resistance. Crop Sci 42:271–277

Zhang D, Cheng H, Wang H, Zhang H, Liu C, Yu D (2007) Identification of genomic regions determining flower and pod numbers development in soybean (*Glycine max* L.). J Genet Genom 37:545–556

Zhang B, Chen P, Chen C, Wang D, Shi A, Hou A, Ishibashi T (2008) Quantitative trait loci mapping of seed hardness in soybean. Crop Sci 48:1341–1349

Zhao Y, Cheng X, Liu X, Wu H, Bi H, Xu H (2018) The wheat MYB transcription factor TaMYB^{31} is involved in drought stress responses in Arabidopsis. Front Plant Sci 9:1426