Theor Appl Genet
DOI 10.1007/s00122-013-2158-x

ORIGINAL PAPER

Genome-wide association analysis for nine agronomic traits in maize under well-watered and water-stressed conditions

Yadong Xue · Marilyn L. Warburton · Mark Sawkins · Xuehai Zhang · Tim Setter · Yunbi Xu · Pichet Grudloyma · James Gethi · Jean-Marcel Ribaut · Wanchen Li · Xiaobo Zhang · Yonglian Zheng · Jianbing Yan

Received: 1 March 2013 / Accepted: 12 July 2013

Abstract Drought can cause severe reduction in maize production, and strongly threatens crop yields. To dissect this complex trait and identify superior alleles, 350 tropical and subtropical maize inbred lines were genotyped using a 1536-SNP array developed from drought-related genes and an array of 56,110 random SNPs. The inbred lines were crossed with a common tester, CML312, and the test-crosses were phenotyped for nine traits under well-watered and water-stressed conditions in seven environments. Using genome-wide association mapping with correction for population structure, 42 associated SNPs ($P \leq 2.25 \times 10^{-6}$ 0.1/N) were identified, located in 33 genes for 126 trait × environment × treatment combinations. Of these genes, three were co-localized to drought-related QTL regions. Gene GRMZM2G125777 was strongly associated with ear relative position, hundred kernel weight and timing of male and female flowering, and encodes NAC domain-containing protein 2, a transcription factor expressed in different tissues. These results provide some good information for understanding the genetic basis for drought tolerance and further studies on identified candidate genes should illuminate mechanisms of drought tolerance and provide tools for designing drought-tolerant maize cultivars tailored to different environmental scenarios.

Communicated by T. Lubberstedt.

Electronic supplementary material The online version of this article (doi:10.1007/s00122-013-2158-x) contains supplementary material, which is available to authorized users.

Introduction

Maize (*Zea mays* L.) is a staple for humans and other animals, and serves as raw material for production of starch, oil, protein, food sweeteners, and alcohol. Drought is the most serious environmental stress obstructing maize production and greatly reduces crop yields. It has been estimated that drought causes yield reductions ranging from 9.3 to 35.1 % in China (Wang and Li 2010), and even a transient

Y. Xue · X. Zhang · X. Zhang · Y. Zheng · J. Yan (✉)
National Key Laboratory of Crop Genetic Improvement, Huazhong Agricultural University, Wuhan 430070, China
e-mail: yjianbing@mail.hzau.edu.cn

M. L. Warburton
USDA-ARS Corn Host Plant Research Resistance Unit, Mississippi State University, Box 9555, Starkville, MS 39762, USA

M. Sawkins · J.-M. Ribaut
Generation Challenge Program, Apdo. Postal 6-641, 06600 Mexico, D.F., Mexico

T. Setter
Department of Crop and Soil Sciences, Cornell University, Ithaca, NY 14853, USA

Y. Xu
International Maize and Wheat Improvement Center, Apdo. Postal 6-641, 06600 Mexico, D.F., Mexico

P. Grudloyma
Nakhon Sawan Field Crops Research Center (NSFCRC), Takfa, Nakhon Sawan 60190, Thailand

J. Gethi
Kenya Agricultural Research Institute (KARI), RRC Mtwapa, P.O. Box 16, Mtwapa, Kenya

W. Li
Maize Research Institute, Sichuan Agricultural University, Ya'an, China

Published online: 25 July 2013

drought may cause yield losses of 5–15.5 % annually in some area of the central US Corn Belt (Campos et al. 2004). Predicted long-term effects of global warming include increased drought conditions over much of the world (Cook et al. 2007). Therefore, drought-tolerant maize is urgently needed to maintain sufficient global production (Bruce et al. 2002). However, it is clear that the genetic mechanisms of drought tolerance are very complex, and successful genetic improvement programs of drought tolerance require information regarding genotype, environment, and genotype by environment (G × E) interactions (Yue et al. 2005; Bänziger and Araus 2007). Characterization of functional genes or markers closely linked to genes related to drought tolerance is a key step towards genomics-assisted plant breeding.

Grain yield (GY) under water stress is the primary trait used to assess the degree of drought tolerance in many crops, including maize. Correlated secondary traits, such as anthesis-silking interval (ASI), grain yield components, and plant height, are generally easier to measure than yield and show a higher heritability, and thus may be more suitable for improving maize selection response to water-stressed conditions (Bänziger et al. 2000; Setter 2012; Edmeades et al. 2000). ASI is commonly used as a selection criterion for drought-tolerant maize genotypes, as it has been shown to be highly correlated with grain yield under water-stressed conditions (Gutierrez-Rodriguez et al. 1998). At the International Maize and Wheat Improvement Center (CIMMYT), yield under water-stressed conditions was increased 3 % per cycle over eight cycles using this approach (Bolaños and Edmeades 1993). In addition, a limited water supply often postpones developmental stages in maize, reducing plant and ear height, and thus available photosynthate for grain production, resulting in drastic reduction of yield (Sari-Gorla et al. 1999; Lopes et al. 2011); thus, selection for taller plants and higher ears can enhance yield under drought conditions. Therefore, GY, along with these secondary traits, could be used to study the effects of drought stress on plants and to identify underlying functional genes (Messina et al. 2010).

Quantitative trait locus (QTL) mapping in biparental populations has been used to identify regions of the maize genome likely responsible for changes in morphological, metabolic, and enzymatic traits related to drought response (Ribaut et al. 2004; Welcker et al. 2007). Clustering of drought-related QTL including GY, OP (osmotic potential), LA (leaf surface area), and PH, consistent with pleiotropy or linkage, has been observed in several studies (Guo et al. 2008; Nikolic et al. 2011; Rahman et al. 2011). Mapping QTL associated with metabolic traits that are significantly correlated with drought tolerance has also uncovered QTL for carbohydrate and abscisic acid (ABA) content accumulation during stress (Tuberosa et al. 2002; Capelle et al. 2010). Biparental mapping cannot explore the full extent of allelic diversity that would be present in diverse germplasm. The limited resolution (10–20 cM; Holland 2007) associated with this technique necessitates further fine mapping to isolate the possible candidate gene(s).

Association analysis based on linkage disequilibrium has several advantages over biparental QTL mapping and is a useful tool to identify superior alleles for complex traits (Remington et al. 2001; Newton-Cheh and Hirschhorn 2005; Zhu et al. 2008; Myles et al. 2009; Rosenberg et al. 2010; Yan et al. 2011). Drought tolerance has also been subjected to association mapping to identify several SNPs associated with genes controlling the ABA pathway (Setter et al. 2010) and numerous SNPs directly associated with drought tolerance (Hao et al. 2011; Lu et al. 2010). The previous drought studies were run with a low marker density, which would be unable to capture the global genetic diversity for drought.

In this study, a commercial maize SNP50 array (Ganal et al. 2011) was used to perform genome-wide association analysis with a diverse panel of 350 maize genotypes phenotyped under well-watered and water-stressed conditions. The objective of this study was to identify candidate genes significantly associated with grain yield and related phenotypic traits under drought stress. In addition to the information gained on drought tolerance, this study may serve as a model for genetic dissection of other complex quantitative traits.

Materials and methods

Plant material

The association-mapping panel was comprised of 350 maize inbred lines selected to represent a wide range of diversity that could all be grown in a common environment, i.e., the tropical and subtropical field conditions typical of the developing world. The germplasm used in this experiment includes known drought-tolerant lines, known drought-susceptible lines (Setter et al. 2010), and others not previously tested (Online Resource 1). All inbred lines were testcrossed to a common tester, CML312, which has a good general combining ability and moderate drought susceptibility. CML312 was selected from CIMMYT Population 500, which is a subtropical, white dent population of intermediate maturity. It was generated by combining germplasm from DeKalb and Northrup King commercial hybrids, South Asia Pop.-3 × Suwan 1, Pools 32 and 20, and Populations 42 and 44. The backgrounds of the latter four pools and populations are very diverse, containing germplasm from South Asia, North, South, and Central America, the Caribbean, Europe, and the Middle East. Although it displays good general combining ability with most inbred lines tested to date, it displays specific combining ability with CML313, 314, 315, and 321. Of these, only CML321 is included in the current drought association-mapping panel. In a study of genetic relationships

between CIMMYT derived inbred lines by Xia et al. (2005), CML312 was most closely related to CML311 (not included in the present study), and fairly equidistant to all other lines in the study. By crossing to CML312, the effects of the genes coming from the other inbred parent could be measured, and would not be masked by the effects of specific combining ability or narrow sense heterosis (except perhaps in the testcross with CML321). In addition, the weak drought tolerance coming from CML312 should not mask the tolerance coming from the other parent; nor should it contribute to poor performance in the hybrids, making tolerance impossible to measure. Genotypes were assigned to one of three precocity groups, and planting dates of each precocity group were staggered, so that all lines would experience the same drought stress during flowering time.

Phenotyping of hybrids

All hybrids were grown over 2 years (2006–2007 and 2007–2008 dry seasons) in four locations in Tlaltizapan, Mexico, Sichuan, China (2006–2007 only), Nairobi, Kenya, and Nakhon Sawan, Thailand, typical of arid climates in developing countries. Fields were planted with two replications in an alpha lattice design, using 5-m rows per plot and two seeds per hole, and plants were thinned after emergence. Water stress was applied at flowering of each maturity group, and continued through grain filling. The following traits under well-watered (WW) and water-stressed (WS) conditions were measured on five plants and averaged over the plot: grain yield (GY); hundred kernel weight (HKW); kernel number (KNO); ear height (EH); plant height (PH); relative ear position (EPO); female flowering, (days to silk DTS); male flowering, (days to anthesis DTA); and anthesis-silking interval (ASI). The correlation between the two replications within each location and treatment was calculated to estimate repeatability and confidence in the data. The difference in yield between WW and WS materials was calculated to ensure that plants subjected to drought treatment were under more stressful conditions.

Phenotypic data analysis

Repeatability (w^2) for each trait was calculated across environments, where an environment was defined as each field × year combination. Repeatabilities were computed as follows:

$$w^2 = \sigma_G^2 / \left[\sigma_G^2 + (\sigma_{GE}^2/n) + \sigma_e^2/(nr)\right]$$

where σ_G^2 is the genotypic variance, σ_{GE}^2 is the genotype × environment variance, σ_e^2 is the error variance, n is the number of environments, and r is the number of replications. The estimates of σ_G^2, σ_{GE}^2, and σ_e^2 were analyzed by analysis of variance (ANOVA) using the *aov* function in R (version 2.14.1, R Foundation for Statistical Computing, http://www.r-project.org/). Correlation coefficients were obtained using Pearson's statistic as implemented in the *cor* procedure in R. Best linear unbiased prediction (BLUP) was done by fitting the following mixed linear model in R package "lme4" for estimation of breeding values of all testers:

$$Y = (1|\text{LINE}) + (1|\text{ENV}) + (1|\text{REP\%in\%LINE} : \text{ENV}) + (1|\text{LINE} : \text{ENV})$$

where Y is trait data, the parentheses indicate random effects, "1|" means groups, and ":" means interactions. LINE refers to all testcrosses used, ENV to environments, each of which is a combination of years and locations, and REP to the replications in one ENV.

Genotyping and filtering

In this study, the MaizeSNP50 BeadChip maize array (Illumina) was used to genotype all lines. SNP content and selection criteria are described in Ganal et al. (2011). Briefly, 56,110 markers that were evenly spaced along the B73 reference sequence (AGPv2) were used to cover the whole maize genome. These markers were selected from several published and unpublished sources. Of 350 maize lines, 333 were successfully genotyped. After removal of unsuccessful calls, SNPs that were missing in more than 20 % of the tested lines, and SNPs exhibiting >20 % heterozygosity (unexpected in inbred lines), usable information was reported for 50,989 SNPs. Because SNPs with low MAF (minor allelic frequency) in the preliminary data analysis can often result in false positive associations, markers with MAFs <5 % were excluded from the association analysis, leaving 43,990 SNPs. The association panel had been previously genotyped with the SNP1536 chip (Setter et al. 2010), and these two data sets were compared to confirm the quality of the current genotyping assays. Five SNPs and 15 lines for which the differences in SNPs were >10 % between the two data sets were removed from the analysis. A final total of 44,314 markers from the two chips scored in 318 lines were considered sufficiently robust and consistent to be employed in the present analysis.

Association analysis

The linkage disequilibrium measurement parameter r^2 was used to estimate LD between SNPs on each chromosome via the software package TASSEL3.0 (Bradbury et al. 2007). A principal components analysis (PCA) was used to correct for population stratification (Price et al. 2006; Zhang et al. 2009b, 2010), and a kinship matrix was calculated using the Loiselle algorithm (Loiselle et al. 1995) to determine relatedness among individuals.

Zhao et al. (2007) demonstrated that the PCA method provides results comparable with alternative approaches such as the STRUCTURE algorithm. The SNP data set from 318 inbred lines was analyzed with the GWAS tool GAPIT (Genome Association and Prediction Integrated Tool-R package) (Lipka et al. 2012). The PC matrix was generated automatically by setting GAPIT parameters pca.total to 6. The general linear model (GLM) with PCs for population structure control (QQ plots showed over fitting of PC + K model, data not shown) was used for genome-wide association mapping with 44,314 SNPs (MAF > 0.05) on data of each environment. A Bonferroni-corrected threshold probability based on individual tests was calculated to correct for multiple comparisons, using 1/*N*, where *N* is the number of individual trait-SNP combinations tested.

Results

Analysis of phenotypes under WW and WS

The repeatability values (w^2) of the nine target traits and their average phenotypic performance based on BLUP values across the seven environments are shown in Table 1. Non-significant differences were observed between the means for DTA under WW and WS conditions using *F* test, most likely because drought stress was imposed just prior to flowering. Average ASI was three times higher under WS than under WW conditions, which was expected because ASI reflects susceptibility to drought (Welcker et al. 2007). The repeatability values for DTA and DTS were high under both WW and WS, while w^2 for ASI was lower (58 % under WW and 47 % under WS). WS decreased GY by 54 % on average, which indicated that severe drought conditions had been experienced. Average repeatability of GY was 61 % in the WW environment and 54 % under WS conditions. The results for KNO were analogous to those of GY except that KNO repeatability was very low under WS. HKW repeatability was also significantly decreased, in WS compared with WW conditions, and repeatability was higher for this trait than for other yield-related traits. PH and EH were lower under WS conditions because plants were unable to fully develop under drought. The upper part of the plant (above the ear) was more affected by drought, and this was reflected in much lower WS/WW ratio for PH than for EH. Although the difference in EPO between WW and WS was significant, the difference itself (WS/WW = 1.1) was not as great as for PH or EH.

Correlations

Under WW conditions, no phenotypic correlations were observed between GY and flowering traits. A negative correlation ($P \leq 0.01$) between ASI and GY (−0.425) was observed under WS conditions, as well as between DTS and KNO (−0.389) (Table 2). There was a correlation

Table 1 Trait performance based on the BLUP value across seven environments

Category	Trait	Treatment	Average ± SD	Range	WS/WW	Repeatability
Plant architecture	PH	WW	219.9 ± 8.48	184.7–240.60	0.79**	0.840
		WS	173.5 ± 4.41	156.2–185.22		0.522
	EH	WW	107.6 ± 8.57	85.2–133.91	0.85**	0.897
		WS	92.0 ± 5.36	76.4–112.04		0.744
	EPO	WW	0.5 ± 0.03	0.4–0.55	1.1**	0.899
		WS	0.5 ± 0.03	0.5–0.68		0.823
Flowering time	DTA	WW	70.4 ± 1.96	65.4–75.02	0.99	0.906
		WS	70.3 ± 2.08	64.6–74.90		0.897
	DTS	WW	72.5 ± 2.15	67.2–77.99	1.05**	0.915
		WS	76.4 ± 2.05	70.7–82.32		0.781
	ASI	WW	2.0 ± 0.50	0.6–3.43	3.11**	0.577
		WS	6.2 ± 0.68	4.5–8.04		0.466
Yield components	GY	WW	2.7 ± 0.15	2.2–3.04	0.46**	0.606
		WS	1.2 ± 0.12	0.9–1.59		0.544
	KNO	WW	1,022.2 ± 39.72	884.0–1,116.15	0.45**	0.49
		WS	457.7 ± 17.44	416.0–523.91		0.161
	HKW	WW	33.4 ± 1.95	28.5–39.44	0.85**	0.844
		WS	28.4 ± 1.63	23.5–34.55		0.699

PH plant height, *EH* ear height, *EPO* relative ear position, *DTA* days to anthesis, male flowering, *DTS* days to silk, female flowering, *ASI* anthesis-silking interval, *GY* grain yield, *KNO* kernel number, *HKW* hundred kernels weight

*, ** Significant at $P = 0.05, 0.01$, respectively

Table 2 Correlation matrix for all traits based on BLUP value across seven environments under water stress (WS, under diagonal) and well-watered (WW, above diagonal)

Trait	PH	EH	EPO	GY	KNO	HKW	DTA	DTS	ASI
PH	1	0.77**	0.36**	0.40**	0.33**	0.08	0.48**	0.46**	0.03
EH	0.66**	1	0.87**	0.30**	0.35**	−0.04	0.60**	0.55**	−0.07
EPO	0.002	0.74**	1	0.14*	0.25**	−0.11*	0.51**	0.45**	−0.12*
GY	0.26**	0.21**	0.04	1	0.45**	0.30**	0.08	0.06	−0.03
KNO	0.14**	0.02	−0.12*	0.37**	1	−0.44**	0.27**	0.25**	−0.02
HKW	0.17**	0.10	−0.01	0.43**	−0.23**	1	−0.10	−0.10	0.04
DTA	0.25**	0.47**	0.42**	−0.24**	−0.19**	−0.01	1	0.94**	−0.04
DTS	0.11*	0.29**	0.31**	−0.44**	−0.39**	−0.02	0.83**	1	0.28**
ASI	−0.18**	−0.22**	−0.11*	−0.43**	−0.42**	−0.01	−0.06	0.50**	1

PH plant height, *EH* ear height, *EPO* relative ear position, *DTA* days to anthesis, male flowering, *DTS* days to silk, female flowering, *ASI* anthesis-silking interval, *GY* grain yield, *KNO* kernel number, *HKW* hundred kernels weight

*, ** Significant at $P \leq 0.05, 0.01$, respectively

between GY and HKW, and it was higher than the correlation between GY and KNO under WS, but lower under WW while correlations between PH and EH with GY were higher under both conditions.

Linkage disequilibrium

All 44,314 SNPs with a MAF > 0.05 were loaded into Plink 1.07 (Purcell et al. 2007) to calculate genome-wide LD in this panel of tropical/subtropical maize germplasm. A rapid decline in LD was observed with increasing physical distance on all chromosomes (Online Resource 2), but the decay rate varied over different chromosomes: equilibrium was reached within 30–45 kb on chromosome 1, 50–60 kb on chromosomes 4 and 5, and 80–150 kb on the rest of the chromosomes. The mean LD decay across all chromosomes was 80–100 kb ($r^2 = 0.1$). If we relax the cut-off to $r^2 = 0.2$, the mean LD decay was very rapid reach to gene level as 5 kb.

Genome-wide association studies

A total of 126 trait × environment × treatment combinations were analyzed, resulting from measurements of nine traits in seven environments and two watering treatments. Population structure was controlled with the PC matrix. Previous work with this same panel indicated that population structure was relatively modest, with five PC axes explaining approximately 7.2 % of the variation in population structure (Setter et al. 2010). In the current study, population structure was re-estimated using a larger number of SNPs, and a similar result was obtained, with six PC axes explaining about 6.9 % of the variation. Quantile–quantile plots (Fig. 1 and Online Resource 3) showed that the GLM model with six PC axes effectively accounted for

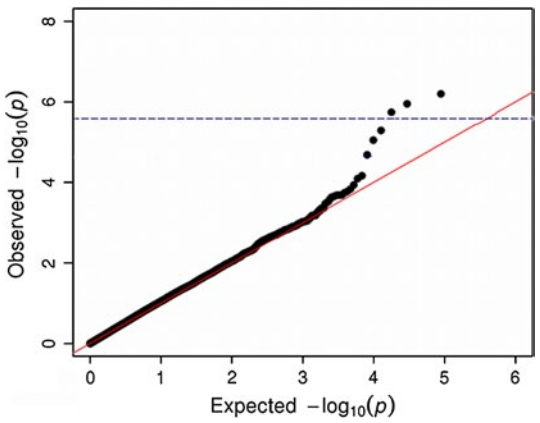

Fig. 1 Quantile–quantile plot for associations with ear height in water-stressed condition in Thailand in 2008. *P* values are shown on a −log10 scale and the dashed horizontal line indicate Bonferroni-corrected threshold 0.1/*N* (color figure online)

population substructure of all traits. A total of 51 associations involving 42 SNPs, located in 33 genes, were identified at the $P < 2.25 \times 10^{-6}$ (0.1/*N*) level (Online Resource 4). Seven of these genes were associated with more than one trait. These results are consistent with the quantitative nature of drought tolerance, which is known to be controlled by a large number of genes with small effects. Detailed association results are presented in Table 3 and Online Resource 5. More associations were identified under WW than WS for most traits. With respect to plant architecture traits, 4, 10, and 11 significant SNPs/treatments were detected for PH, EH, and EPO, respectively, across each environment; for flowering time traits, 4 associations were detected for DTA, 2 for DTS, and 5 for ASI; and for yield-related traits, 6, 2, and 3 associations were detected for GY, KNO, and HKW, respectively.

Table 3 Significant SNP-trait associations for nine agronomic traits

Trait	C07	K07	K08	M07	M08	T07	T08	Total	SNPs	Genes
PH										
WW	0	0	0	0	0	1	1	4	4	4
WS	0	0	0	1	0	0	1			
EH										
WW	1	2	0	0	0	2	3	12	10	9
WS	1	0	0	0	0	0	3			
EPO										
WW	2	3	0	1	0	2	0	12	11	7
WS	1	0	0	0	2	1	0			
DTA										
WW	0	1	0	0	0	0	0	4	4	4
WS	1	2	0	0	0	0	0			
DTS										
WW	0	0	0	0	1	0	2	3	2	2
WS	0	0	0	0	0	0	0			
ASI										
WW	0	2	1	0	0	2	0	5	5	5
WS	0	0	0	0	0	0	0			
GY										
WW	0	3	0	0	0	1	1	6	6	5
WS	0	1	0	0	0	0	0			
KNO										
WW	0	0	0	0	1	0	1	2	2	2
WS	0	0	0	0	0	0	0			
HKW										
WW	0	0	0	0	0	1	1	3	3	3
WS	1	0	0	0	0	0	0			
Total								51	42	33

PH plant height, *EH* ear height, *EPO* relative ear position, *DTA* days to anthesis, male flowering, *DTS* days to silk, female flowering, *ASI* anthesis-silking interval, *GY* grain yield, *KNO* kernel number, *HKW* hundred kernel weight, *C* China, *K* Kenya, *M* Mexico, *T* Thailand, SNPs numbers of different, *SNPs* detected, *Genes* numbers of distinct genes detected

Numbers of significant SNP-trait associations of measured traits at $p = 2.25\text{e}{-06}$ from column C07 to column total

Discussion

Maize yields are most damaged by water shortage during flowering and early kernel development (Setter et al. 2001). In our study, withdrawal of irrigation before these important stages led to the expected severe water-stressed condition, and decreased PH and EH substantially, delayed DTS and increased ASI, and reduced KNO, and thus GY (Table 2). HKW was less affected by drought because plants, as part of their drought response, were able to distribute resources to pollinated kernels to ensure the optimal development of at least some of them, even though fewer kernels would be pollinated under drought. Higher correlations between PH and EH with GY under WW than under WS conditions show that tall plants have a larger photosynthetically active leaf area and more stem reserves (Sari-Gorla et al. 1999; Lopes et al. 2011) and, thus, better capacity for grain filling than short plants.

When accurate grain yield prediction under water deficit during flowering time can be carried out using ASI, selection gain is increased and phenotyping cost is decreased in breeding programs (Bolaños and Edmeades 1993; Edmeades et al. 1999; Monneveux et al. 2006; Ribaut et al. 2009; Barker et al. 2010). In the current study, ASI was only correlated with GY under water stress, confirming the predictive power of ASI on drought tolerance and also indicated that late silking is unfavorable for kernel set and grain yield under water deficit. Strong selection on elite germplasm, however, has narrowed the genotypic variation of this trait (Monneveux et al. 2008). In the present study, we have presented drought-related information collected for a diverse range of germplasm in replicated field experiments across four locations over 2 years. The availability of this data set not only advances our understanding of drought tolerance in maize, but also may be used by the maize community as a resource to deepen the drought resistance gene pool.

Genome-wide association mapping of drought tolerance

The association panel used in this study was comprised mainly of tropical and subtropical lines, and is part of a

larger association resource genetically characterized by Yan et al. (2009), who estimated average LD decay to be 5–10 kb. In the current study, LD calculated based on 44,314 SNPs and the subset panel was higher, averaging 80–100 kb across the whole genome. The current panel lacks the diversity from the temperate lines included in the larger panel, which would have uncovered more recombinations and thus smaller linkage blocks. The higher LD in the current study suggests the need for fewer SNP markers to adequately cover the genome.

Large-scale genome-wide association analyses have been employed to dissect complex traits (Buckler et al. 2009; Brown et al. 2011; Weng et al. 2011; Riedelsheimer et al. 2012; Li et al. 2013) and have proven to be an effective way to find candidate genes for these traits. In this study, 42 SNPs were associated with one or more phenotyped traits, and some of these SNPs were located within consensus QTL found by a previous meta-analysis of drought tolerance in maize (Online Resource 5) (Li et al. 2010). Using a candidate gene association strategy, several SNPs in previous studies were found to be associated with metabolite (Setter et al. 2010) or agronomic (Lu et al. 2010; Hao et al. 2011) traits under water stress; however, none of those SNPs were detected in our study. This may be due to the low correlations between metabolite and agronomic traits (Schauer et al. 2006; Fu et al. 2009; Sulpice et al. 2010; Kooke and Keurentjes 2011) or the different populations under study. In this study, the testcrosses used and the different environments in which they were measured are further reasons that no common SNPs were detected compared with the study of Setter et al. (2010). A more recent genome-wide association study on metabolite traits in maize leaves under WS conditions by Riedelsheimer et al. (2012) found several significant associations; these included one SNP, PZB00083.3, associated with gama-aminobutyric acid ($P = 2.05 \times 10^{-5} < 0.1/N$) and also with an agronomic trait (GY_TS8) under WS in our study. The gene identified (a phytochrome receptor gene) is worthy of further investigation in drought-stressed maize ears.

Many of the 33 candidate genes identified in this study were associated with plant architecture or flowering time traits that had low correlation with GY under either WW or WS. Other studies have shown that these traits are related to drought tolerance and can be used to improve maize yield under drought (Zhang et al. 2009a; Saeed et al. 2011; Tardieu 2011; Lopes et al. 2011). Finding direct strong correlations between grain yield and secondary traits such as PH, EH, root growth, and root system architecture are often difficult; even though such traits may be associated with drought tolerance, they may have opposite influences on drought tolerance at different levels of drought severity. SNPs and genes found to be associated with specific drought environments may be used to design lines with enhanced drought tolerance under the same conditions (Tardieu 2011). Although impressive outcomes obtained in one drought scenario may have a limited application to other drought prone environments, the results of this study are still a valuable resource for improving drought tolerance of maize in the tested areas. In addition, SNPs and genes may highlight the important underlying drought-tolerance mechanisms, and may provide information for dissecting the genetic basis of drought tolerance and further molecular breeding.

Detailed analysis of candidate genes

A meta-analysis of published QTL related to drought tolerance identified over 400 QTL and 79 discrete consensus genomic regions (or mQTL) related to drought tolerance in maize (Li et al. 2010). Physical coordinates of the consensus genomic regions based on the B73 maize reference sequence version 2 overlapped with 3 of 33 associated genes identified in our study, which were located in three mQTL regions. For further examination, we chose annotated genes from within the consensus regions, and which were associated with flowering time or yield-associated traits in WS or in both WS and WW conditions. Of the 33 most significant genes, seven were associated with multiple traits. The highest number of simultaneous trait associations was seen with the gene GRMZM2G125777, which harbored the SNP PZE-104036909. This gene was associated with four different traits (EPO, DTA, DTS, and HKW), and encodes NAC domain-containing protein 2, which is expressed in all associated tissues (Sekhon et al. 2011).

Two genes, GRMZM2G140082 and GRMZM2G313643, were associated with grain yield-related traits measured under WS environment (HKW_CS7 and GY_KS7, respectively) and are plausible candidate genes for drought tolerance. GRMZM2G140082 and GRMZM2G313643 encode proteins with the same function, tyrosine-protein kinase. This kind of protein is involved in growth, proliferation, dead, survival, etc. (Li and Hristova 2006; Sharma et al. 2009). The two genes, which functioned in ear, may have the potential to improve GY under stress. Further knowledge about how these two genes work together, and with other genes associated in this and other studies, will aid targeted efforts to increase maize grain yield under drought stress.

Given that ASI is strongly correlated with GY under water stress (Bänziger et al. 2000), genes related to ASI are of particular interest. Two genes, GRMZM2G137961 and GRMZM2G119079, associated with ASI_KW7, were identified in this study (Online Resource 5). GRMZM2G119079 encodes a protein tyrosine phosphatase (PTPase), which functions in cell cycle control and cell

growth, proliferation, differentiation, and transformation (Fordham-Skelton et al. 1999; Corellou et al. 2000; Ghelis 2011). GRMZM2G137961 encodes a member of Acyl-CoA *N*-acyltransferase superfamily that is involved in Krebs cycle, fatty acid metabolism, etc. It was highly expressed in mature tassel (Sekhon et al. 2011). There are three more genes associated with ASI. These genes are good candidates for future study, because regulation of cell proliferation and growth under drought conditions are factors that might determine ASI and drought tolerance (Ribaut et al. 2009; Setter et al. 2010; Lorković 2009).

Most of the remaining genes reported in Online Resource 5 were associated with traits pertaining to plant architecture. Although plant architecture traits have a low correlation with GY and drought tolerance, a full understanding the mechanism of these traits would assist breeders in the development of ideotypic maize phenotypes for high yield under WW as well as WS conditions.

The GWAS has been successfully used at a high resolution to uncover associations involving complex traits from field-grown maize inbred lines with genetic variants (Buckler et al. 2009; Tian et al. 2011; Riedelsheimer et al. 2012; Li et al. 2013). With the development of next-generation sequencing technology, the cost of sequencing has dropped dramatically, and the large number of markers needed to adequately cover the maize genome for GWAS is now available at a reasonable cost. Although gene level resolution can now be achieved in some diverse lines using GWAS (Yan et al. 2009), the technique still only provides a statistical (i.e., indirect) clue for connecting traits with their associated genomic sequences. It is therefore necessary to validate the uncovered associations using biological evidence obtained from approaches such as transgenic or mutation techniques (RNAi, antisense methods, or production of knock-out mutants for inducing loss-of-function point mutations in candidate genes) or the slower creation of near isogenic lines for comparative analysis of pairs of maize lines with and without the genes of interest.

Acknowledgments This research was supported by the Generation Challenge Program and the National Hi-Tech Research and Development Program of China (2012AA10A307) and the National Natural Science Foundation of China (31101156).

Conflict of interest The authors declare that they have no conflict of interest.

References

Bänziger M, Araus JL (2007) Recent advances in breeding maize for drought and salinity stress tolerance. In: Jenks MA, Hasegawa PM, Jain SM (eds) Advances in molecular breeding toward drought and salt tolerant crops. Springer Netherlands, Dordrecht, pp 587–601

Bänziger M, Edmeades GO, Beck D, Bellon M (2000) Breeding for drought and nitrogen stress tolerance in maize: from theory to practice. CIMMYT, Mexico, D.F.

Barker T, Campos H, Cooper M, Dolan D, Edmeades G, Habben J, Schussler J, Wright D, Zinselmeier C (2010) Improving drought tolerance in maize. In: Janick J (ed) Plant breeding reviews. Wiley, Oxford, pp 173–253

Bolaños J, Edmeades GO (1993) Eight cycles of selection for drought tolerance in lowland tropical maize. I. Responses in grain yield, biomass, and radiation utilization. Field Crops Res 31:233–252. doi:10.1016/0378-4290(93)90064-T

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. doi:10.1093/bioinformatics/btm308

Brown PJ, Upadyayula N, Mahone GS, Tian F, Bradbury PJ, Myles S, Holland JB, Flint-Garcia S, McMullen MD, Buckler ES, Rocheford TR (2011) Distinct genetic architectures for male and female inflorescence traits of maize (J Flint, Ed.). PLoS Genet 7:e1002383. doi:10.1371/journal.pgen.1002383

Bruce WB, Edmeades GO, Barker TC (2002) Molecular and physiological approaches to maize improvement for drought tolerance. J Exp Bot 53:13–25. doi:10.1093/jexbot/53.366.13

Buckler ES, Holland JB, Bradbury PJ et al (2009) The genetic architecture of maize flowering time. Science 325:714–718. doi:10.1126/science.1174276

Campos H, Cooper M, Habben JE, Edmeades GO, Schussler JR (2004) Improving drought tolerance in maize: a view from industry. Field Crops Res 90:19–34. doi:10.1016/j.fcr.2004.07.003

Capelle V, Remoué C, Moreau L, Reyss A, Mahé A, Massonneau A, Falque M, Charcosset A, Thévenot C, Rogowsky P, Coursol S, Prioul JL (2010) QTLs and candidate genes for desiccation and abscisic acid content in maize kernels. BMC Plant Biol 10:2. doi:10.1186/1471-2229-10-2

Cook ER, Seager R, Cane MA, Stahle DW (2007) North American drought: reconstructions, causes, and consequences. Earth Sci Rev 81:93–134. doi:10.1016/j.earscirev.2006.12.002

Corellou F, Potin P, Brownlee C, Kloareg B, Bouget FY (2000) Inhibition of the establishment of zygotic polarity by protein tyrosine kinase inhibitors leads to an alteration of embryo pattern in Fucus. Dev Biol 219:165–182. doi:10.1006/dbio.1999.9603

Edmeades GO, Bolaños J, Chapman SC, Lafitte HR, Bänziger M (1999) Selection improves drought tolerance in tropical maize populations: I. Gains in Biomass, Grain Yield, and Harvest Index. Crop Sci 39:1306–1315. doi:10.2135/cropsci1999.3951306x

Edmeades GO, Bolaños J, Elings A, Ribaut JM, Bänziger M, Westgate ME (2000) The role and regulation of the anthesis-silking interval in maize. In: Otegui ME, Slafer GA (eds) Physiological bases for maize improvement. Food Products Press, New York, pp 75–111

Fordham-Skelton AP, Skipsey M, Eveans IM, Edwards R, Gatehouse JA (1999) Higher plant tyrosine-specific protein phosphatases (PTPs) contain novel amino-terminal domains: expression during embryogenesis. Plant Mol Biol 39:593–605. doi:10.1023/A:1006170902271

Fu JY, Keurentjes JJB, Bouwmeester H, America T, Verstappen FWA, Ward JL, Beale MH, Vos RCH, Dijkstra M, Scheltema RA, Johannes F, Koornneef M, Vreugdenhil D, Breitling R, Jansen RC (2009) System-wide molecular evidence for phenotypic buffering in Arabidopsis. Nat Genet 41:166–167. doi:10.1038/ng.308

Ganal MW, Durstewitz G, Polley A et al (2011) A large maize (*Zea mays* L.) SNP genotyping array: development and germplasm genotyping, and genetic mapping to compare with the B73 reference genome. PLoS One 6:e28334. doi:10.1371/journal.pone.0028334

Ghelis T (2011) Signal processing by protein tyrosine phosphorylation in plants. Plant Signal Behav 6:942–951. doi:10.4161/psb.6.7.15261

Guo J, Su G, Zhang J, Wang G (2008) Genetic analysis and QTL mapping of maize yield and associate agronomic traits under semi-arid land condition. Afr J Biotechnol 7:1829–1838

Gutierrez-Rodriguez M, Miguel-Chavez RS, Larque-Saavedra A (1998) Physiological aspects in Tuxpeno maize with improved drought tolerance. Maydica 43:137–141

Hao Z, Li X, Xie C, Weng J, Li M, Zhang D, Liang X, Liu L, Liu S, Zhang S (2011) Identification of functional genetic variations underlying drought tolerance in maize using SNP markers. J Integr Plant Biol 53:641–652. doi:10.1111/j.1744-7909.2011.01051.x

Holland JB (2007) Genetic architecture of complex traits in plants. Curr Opin Plant Biol 10:156–161. doi:10.1016/j.pbi.2007.01.003

Kooke R, Keurentjes JJB (2011) Multi-dimensional regulation of metabolic networks shaping plant development and performance. J Exp Bot 63:3353–3365. doi:10.1093/jxb/err373

Li E, Hristova K (2006) Role of receptor tyrosine kinase transmembrane domains in cell signaling and human pathologies. Biochemistry 45:6241–6251. doi:10.1021/bi060609y

Li WJ, Liu Z, Shi Y, Song Y, Wang T, Xu C, Li Y (2010) Detection of consensus genomic region of QTLs relevant to drought-tolerance in maize by QTL Meta-analysis and bioinformatics approach. Acta Agronomica Sinica 36:1457–1467. doi:10.1016/S1875-2780(09)60072-9

Li H, Peng Z, Yang X, Wang W, Fu J, Wang J, Han Y, Chai Y, Guo T, Yang N, Liu J, Warburton ML, Cheng Y, Hao X, Zhang P, Zhao J, Liu Y, Wang G, Li J, Yan J (2013) Genome-wide association study dissects the genetic architecture of oil biosynthesis in maize kernels. Nat Genet 45:43–50. doi:10.1038/ng.2484

Lipka AE, Tian F, Wang Q, Peiffer J, Li M, Bradbury PJ, Gore M, Buckler ES, Zhang Z (2012) GAPIT: genome association and prediction integrated tool. Bioinformatics 28:2397–2399. doi:10.1093/bioinformatics/bts444

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. doi:10.2307/2445869

Lopes MS, Araus JL, van Heerden PDR, Foyer CH (2011) Enhancing drought tolerance in C4 crops. J Exp Bot 62:3135–3153. doi:10.1093/jxb/err105

Lorković ZJ (2009) Role of plant RNA-binding proteins in development, stress response and genome organization. Trends Plant Sci 14:229–236. doi:10.1016/j.tplants.2009.01.007

Lu Y, Zhang S, Shah T, Xie C, Hao Z, Li X, Farkhari M, Ribaut JM, Cao M, Rong T, Xu Y (2010) Joint linkage–linkage disequilibrium mapping is a powerful approach to detecting quantitative trait loci underlying drought tolerance in maize. Proc Natl Acad Sci USA 107:19585–19590. doi:10.1073/pnas.1006105107

Messina CD, Podlich D, Dong Z, Samples M, Cooper M (2010) Yield-trait performance landscapes: from theory to application in breeding maize for drought tolerance. J Exp Bot 62:855–868. doi:10.1093/jxb/erq329

Monneveux P, Sánchez C, Beck D, Edmeades GO (2006) Drought tolerance improvement in tropical maize source populations. Crop Sci 46:180–191. doi:10.2135/cropsci2005.04-0034

Monneveux P, Sanchez C, Tiessen A (2008) Future progress in drought tolerance in maize needs new secondary traits and cross combinations. J Agric Sci 146:287–300. doi:10.1017/S0021859608007818

Myles S, Peiffer J, Brown PJ, Ersoz ES, Zhang Z, Costich DE, Buckler ES (2009) Association mapping: critical considerations shift from genotyping to experimental design. Plant Cell 21:2194–2202. doi:10.1105/tpc.109.068437

Newton-Cheh C, Hirschhorn JN (2005) Genetic association studies of complex traits: design and analysis issues. Mutat Res 573:54–69. doi:10.1016/j.mrfmmm.2005.01.006

Nikolic A, Andjelkovic V, Dodig D, Ignjatovic-Micic D (2011) Quantitative trait loci for yield and morphological traits in maize under drought stress. Genetika 43:263–276. doi:10.2298/GENSR1102263N

Price AL, Patterson NJ, Plenge RM, Weinblatt ME, Shadick NA, Reich D (2006) Principal components analysis corrects for stratification in genome-wide association studies. Nat Genet 38:904–909. doi:10.1038/ng1847

Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MAR, Bender D, Maller J, Sklar P, de Bakker PIW, Daly MJ, Sham PC (2007) PLINK: a toolset for whole-genome association and population-based linkage analysis. Am J Hum Genet 81:559–575. doi:10.1086/519795

Rahman H, Pekic S, Lazic-Jancic V, Quarrie SA, Shah SMA, Pervez A, Shah MM (2011) Molecular mapping of quantitative trait loci for drought tolerance in maize plants. Genet Mol Res 10:889–901. doi:10.4238/vol10-2gmr1139

Remington DL, Thornsberry JM, Matsuoka Y, Wilson LM, Whitt SR, Doebley J, Kresovich S, Goodman MM, Buckler ES (2001) Structure of linkage disequilibrium and phenotypic associations. PNAS 98:11479–11484. doi:10.1073/pnas.201394398

Ribaut JM, Bänziger M, Setter T, Edmeades G, Hoisington D (2004) Genetic dissection of drought tolerance in maize: a case study. In: Nguyen H, Blum A (eds) Physiology and biotechnology integration for plant breeding. Marcel Dekker Inc., New York, pp 571–611

Ribaut JM, Betran J, Monneveux P, Setter T (2009) Drought tolerance in maize. Handbook of maize: its biology. Springer, New York, pp 311–344

Riedelsheimer C, Lisec J, Czedik-Eysenberg A, Sulpice R, Flis A, Grieder C, Altmann T, Stitt M, Willmitzer L, Melchinger AE (2012) Genome-wide association mapping of leaf metabolic profiles for dissecting complex traits in maize. Proc Natl Acad Sci 109:8872–8877. doi:10.1073/pnas.1120813109

Rosenberg NA, Huang L, Jewett EM, Szpiech ZA, Jankovic I, Boehnke M (2010) Genome-wide association studies in diverse populations. Nat Rev Genet 11:356–366. doi:10.1038/nrg2760

Saeed M, Guo WZ, Ullah I, Tabbasam N, Zafar Y, Rahman MU, Zhang TZ (2011) QTL mapping for physiology, yield and plant architecture traits in cotton (*Gossypium hirsutum* L.) grown under well-watered versus water-stress conditions. Electron J Biotechnol 14:3. doi:10.2225/vol14-issue3-fulltext-3

Sari-Gorla M, Krajewski P, Di Fonzo N, Villa M, Frova C (1999) Genetic analysis of drought tolerance in maize by molecular markers. II. Plant height and flowering. Theor Appl Genet 99:289–295. doi:10.1007/s001220051234

Schauer N, Semel Y, Roessner U, Gur A, Balbo I, Carrari F, Pleban T, Perez-Melis A, Bruedigam C, Kopka J, Willmitzer L, Zamir D, Fernie AR (2006) Comprehensive metabolic profiling and phenotyping of interspecific introgression lines for tomato improvement. Nat Biotechnol 24:447–454. doi:10.1038/nbt1192

Sekhon RS, Lin H, Childs KL, Hansey CN, Buell CR, de Leon N, Kaeppler SM (2011) Genome-wide atlas of transcription during maize development. Plant J 66:553–563. doi:10.1111/j.1365-313X.2011.04527.x

Setter TL (2012) Analysis of constituents for phenotyping drought tolerance in crop improvement. Front Physiol 3:180. doi:10.3389/fphys.2012.00180

Setter TL, Flannigan BA, Melkonian J (2001) Loss of kernel set due to water deficit and shade in maize. Crop Sci 41:1530–1540. doi:10.2135/cropsci2001.4151530x

Setter TL, Yan J, Warburton M, Ribaut JM, Xu Y, Sawkins M, Buckler ES, Zhang Z, Gore MA (2010) Genetic association mapping identifies single nucleotide polymorphisms in genes that affect abscisic acid levels in maize floral tissues during drought. J Exp Bot 62:701–716. doi:10.1093/jxb/erq308

Sharma PS, Sharma R, Tyagi T (2009) Receptor tyrosine kinase inhibitors as potent weapons in war against cancers. Curr Pharm Des 15:758–776. doi:10.2174/128161209787582219

Sulpice R, Trenkamp S, Steinfath M, Usadel B, Gibon Y, Witucha-Wall H, Pyl ET, Tschoep H, Steinhauser MC, Guenther M, Hoehne M, Rohwer JM, Altmann T, Fernie AR, Stitt M (2010) Network analysis of enzyme activities and metabolite levels and their relationship to biomass in a large panel of Arabidopsis accessions. Plant Cell 22:2872–2893. doi:10.1105/tpc.110.076653

Tardieu F (2011) Any trait or trait-related allele can confer drought tolerance: just design the right drought scenario. J Exp Bot 63:25–31. doi:10.1093/jxb/err269

Tian F, Bradbury PJ, Brown PJ, Hung H, Sun Q, Flint-Garcia S, Rocheford TR, McMullen MD, Holland JB, Buckler ES (2011) Genome-wide association study of leaf architecture in the maize nested association mapping population. Nat Genet 43:159–162. doi:10.1038/ng.746

Tuberosa R, Salvi S, Sanguineti MC, Landi P, Maccaferri M, Conti S (2002) Mapping QTLs regulating morpho-physiological traits and yield: case studies, shortcomings and perspectives in drought-stressed maize. Ann Bot 89:941–963. doi:10.1093/aob/mcf134

Wang C, Li S (2010) Assessment of limiting factors and techniques prioritization for maize production in China. Scientia Agricultrea Sinica 43:1136–1146

Welcker C, Boussuge B, Bencivenni C, Ribaut JM, Tardieu F (2007) Are source and sink strengths genetically linked in maize plants subjected to water deficit? A QTL study of the responses of leaf growth and of anthesis-silking interval to water deficit. J Exp Bot 58:339–349. doi:10.1093/jxb/erl227

Weng J, Xie C, Hao Z, Wang J, Liu C, Li M, Zhang D, Bai L, Zhang S, Li X (2011) Genome-wide association study identifies candidate genes that affect plant height in chinese elite maize (*Zea mays* L.) inbred lines. PLoS One 6:e29229. doi:10.1371/journal.pone.0029229

Xia XC, Reif JC, Melchinger AE, Frisch M, Hoisington D, Pixley BK, Warlburton ML (2005) Genetic diversity among CIMMYT maize inbred lines investigated with SSR markers II. Subtropical, tropical midaltitude, and highland maize inbred lines and their relationships with elite U.S. and European maize. Crop Sci 45:2573–2582. doi:10.2135/cropsci2005.0246

Yan J, Shah T, Warburton ML, Buckler ES, McMullen MD, Crouch J (2009) Genetic characterization and linkage disequilibrium estimation of a global maize collection Using SNP Markers. PLoS One 4:e8451. doi:10.1371/journal.pone.0008451

Yan J, Warburton M, Crouch J (2011) Association mapping for enhancing maize (*Zea mays* L.) genetic improvement. Crop Sci 51:433–449. doi:10.2135/cropsci2010.04.0233

Yue B, Xue W, Xiong L, Yu X, Luo L, Cui K, Jin D, Xing Y, Zhang Q (2005) Genetic basis of drought resistance at reproductive stage in rice: separation of drought tolerance from drought avoidance. Genetics 172:1213–1228. doi:10.1534/genetics.105.045062

Zhang SW, Li CH, Cao J, Zhang YC, Zhang SQ, Xia YF, Sun DY, Sun Y (2009a) Altered architecture and enhanced drought tolerance in rice via the down-regulation of indole-3-acetic acid by TLD1/OsGH3.13 activation. Plant Physiol 151:1889–1901. doi:10.1104/pp.109.146803

Zhang Z, Buckler ES, Casstevens TM, Bradbury PJ (2009b) Software engineering the mixed model for genome-wide association studies on large samples. Brief Bioinforma 10:664–675. doi:10.1093/bib/bbp050

Zhang Z, Ersoz E, Lai C-Q, Todhunter RJ, Tiwari HK, Gore MA, Bradbury PJ, Yu J, Arnett DK, Ordovas JM, Buckler ES (2010) Mixed linear model approach adapted for genome-wide association studies. Nat Genet 42:355–360. doi:10.1038/ng.546

Zhao K, Aranzana MJ, Kim S, Lister C, Shindo C, Tang C, Toomajian C, Zheng H, Dean C, Marjoram P, Nordborg M (2007) An Arabidopsis example of association mapping in structured samples. PLoS Genet 3:e4. doi:10.1371/journal.pgen.0030004

Zhu C, Gore M, Buckler ES, Yu J (2008) Status and prospects of association mapping in plants. Plant Genome 1:5–20. doi:10.3835/plantgenome2008.02.0089