INTERNATIONAL JOURNAL OF AGRICULTURE & BIOLOGY
ISSN Print: 1560–8530; ISSN Online: 1814–9596
18F–184/2019/21–4–810–818
DOI: 10.17957/IJAB/15.0961
http://www.fspublishers.org

Full Length Article

Genome-Wide Association Analysis for the Genetic Basis of Seven Traits Associated with Corn Grain Moisture and Ear Threshing Rate

Xinqi Wang[1], Xiaoci Ren[1], Nan Mei[2], Yanning Xie[1], Xiangyu Yang[1], Xiaojie Wang[3] and Rengui Zhao[1*]
[1]*College of Agronomy, Jilin Agricultural University, Changchun, 130118, China*
[2]*College of Agronomy, Shenyang Agricultural University, Shenyang, 110000, China*
[3]*Dandong Academy of Agricultural Sciences, Dandong, 118100, China*
[*]For correspondence: zhaorengui@sina.com; dr_ebadi2000@yahoo.com

Abstract

Maize is an important food and economic crop around the world. Grain moisture is an important trait for maize harvest and generally easier to thresh low-moisture grains. In this study, eight agronomic traits, including cob hardness, cob lignin layer thickness, cob percentage of lignin layer, cob diameter, bract length, bract number, thickness and compaction, were selected which associated with maize grain moisture and threshing rate based on field investigation and analyzed seven traits in 80 elite inbred lines planted in two environments (Changchun, Jilin Province and Dandong, Liaoning Province) in 2015 and 2016. Total 76 single nucleotide polymorphism (SNPs) and 41 candidate genes significantly associated with the above seven agronomic traits ($P < 1.0 \times 10^{-6}$) through a genome-wide association study involving 1,490,007 high-quality SNPs were identified. Finally, RT-PCR expression of 19 of the 41 candidate genes in samples with extreme traits was observed. 10 candidate genes were up-regulated expressed in high value samples and nine candidate genes down-regulated expressed in low value samples. This study highlights the genetic architecture of seven traits in corn cobs and bracts which identified potential target genes associated with maize grain moisture and ear threshing rate. © 2019 Friends Science Publishers

Keywords: Corn (*Zea mays*); Cob; Bract; SNP; GWAS; Genetic architecture; Threshing efficiency

Introduction

Genome-wide association study (GWAS) (Cantor *et al*., 2010) is a method to detect a genome-wide set of genetic variants (*i.e*., single-nucleotide polymorphism, copy number variations, deletions and insertions) in natural populations to see whether a particular variant is associated with a distinct trait, which then can be used in the identification of specific single-nucleotide polymorphisms (SNPs) or candidate genes controlling a phenotypic trait based on the level of linkage disequilibrium (LD) among genes (Christoforou *et al*., 2012). In the past few years, GWAS has been performed in multiple lines of maize and identified numerous SNPs or genes associated with different valuable traits (Xiao *et al*., 2017).

Threshing is a key step in maize production. Breeding easy-threshing maize lines promotes mechanization (LI *et al*., 2007) and increases the income of growers and operators. High-grain moisture levels at harvest largely influence maize threshing and increase the degree of potential grain damage (Petkevichius *et al*., 2008). Based on the importance of grain moisture in maize threshing and storage, various approaches have been performed to elucidate the genetic basis of grain moisture. Six quantitative trait locus (QTLs) located on five chromosomes for grain moisture at harvest were identified in a double haploid population consisting of 240 lines using a quantitative genetics approach (Srivastava *et al*., 2017). Field grain drying rate influenced the grain moisture at harvest. GWAS among 80 elite maize inbred lines identified 19 significant SNP markers that are associated with field grain drying rate, and concluded that the field grain drying rate is mainly controlled by broad-sense heritability (0.76) (Dai *et al*., 2017), which coincides with grain moisture at harvest (0.71) (Song *et al*., 2017).

The majority of GWAS investigations have focused on locating SNPs or genes associated with grain moisture or other quantitative traits, whereas studies on identifying agronomic traits associated with grain moisture and elucidating the genetic basis of these traits are limited. In field experiments, it was found that some bract traits (*i.e*., bract compaction) increases grain moisture by limiting the grain drying rate, which in turn influences maize grain threshing efficiency. Furthermore, some cob traits also affect maize threshing rate; for example, a fragile cob will introduce more impurities in the mechanized threshing process. In the present study, eight traits of cobs and bracts from 80 elite inbred lines from Jilin Province, China were collected and then 76 SNP markers associated with these traits were identified. This research also highlights the genetic basis of specific traits associated with maize grain moisture at harvest and provides potential genetic material that may be used in breeding easy-threshing lines.

To cite this paper: Wang, X., X. Ren, N. Mei, Y. Xie, X. Yang, X. Wang and R. Zhao, 2019. Genome-wide association analysis for the genetic basis of seven traits associated with corn grain moisture and ear threshing rate. *Intl. J. Agric. Biol*., 21: 810–818

Material and Methods

Field Experiments and Phenotyping

All 80 maize elite inbred lines (detailed information as described by Dai *et al.* (2017) of the association panel were planted in two environments (Dandong, Liaoning Province and Changchun, Jilin Province) in 2015 and 2016. Samples from Dandong in 2015 and 2016 were labeled as DD15 and DD16, respectively. While from Changchun in 2015 and 2016 were labeled as CC15 and CC16, respectively. All lines were planted in three-row plots with three replicates using a complete randomized block design at each sample origin.

Eight traits, which included four cob traits (*i.e.*, hardness, lignin layer thickness, percentage of lignin layer, and diameter) and four bract traits (*i.e.*, length, number, thickness and compaction) were recorded for each accession from the four origins (CC15, CC16, DD15 and DD16). Cob hardness was the average value of puncture strength of three cob parts from top to bottom. The thickness of cob lignin layer was calculated as follows: where *clt* is the thickness of the cob lignin layer, *a* is the width of cob lignin layer diameter plus pith diameter, *b* is pith diameter and the average value of the cob lignin layer thickness of three parts from top to bottom was used in further analyses. The percentage of the cob lignin layer was measured from the lignin layer diameter over the cob diameter. The cob diameter was the average value of three samples. Bract length was calculated from the sum of each bract length over the number of bracts. Bract number was the average value of all complete number of bracts from three samples. Bract thickness was measured from the bract thickness in each cob unit area, which was the average value of three parts from the top to the bottom. Bract compaction was determined (Ma *et al*., 2015) as the bract length to the ear length.

Genotyping and SNP Discovery

Genomic DNA was extracted from fresh young leaves using the cetyltrimethylammonium bromide (CTAB) method. DNA purity was quantified using agarose gel electrophoresis and NanoDrop ($OD_{260/280}$ was calculated), and DNA concentrations were determined using Qubit. The libraries were constructed using a TruSeq Library Construction Kit with random DNA fragments which produced by digestion with restriction enzymes. The samples were pooled per plate and PCR amplified. Each library was sequenced on an Illumina HiSeq 2000 platform using paired-end sequencing.

The adaptor and low-quality reads of raw Illumina fastq files were filtered by using fastq-filter in FASTQ-Toolkit (v.2.0.0; Illumina Basespace Labs) with following criteria: 1) The percentage of N bases > 10%; 2) The percentage of low-quality base (≤ 5) higher than 50%. The clean reads were mapped to the reference genome RefGen_v3 of the maize inbred line B37 (Schnable *et al*., 2009) using BWA (Li and Durbin, 2009) with setting options of "mem –t 4 –k 32 –M" and duplicate reads were removed using SAMTOOLS (option: rmdup) (Li *et al*., 2009). The initial SNPs were identified using SAMTOOLS with the default option, then filtered using the following criteria: 1) Individual quality value > 5; 2) Population quality value > 20; 3) Individual depth range: 10–1,000; 4) Population depth range: 80–16,000; 5) Miss rate > 0.10; 6) Minor allele frequency (MAF) > 0.05. Finally, all SNPs that passed the above criteria were annotated using ANNOVAR (Wang *et al*., 2017) and used in the subsequent analyses.

Genome-wide Association Study

All high-quality SNPs was used in GWAS to identify significant SNP markers and candidate genes associated with seven target traits using Farm CPU (Liu *et al*., 2016) based on the average decay distance of LD calculated using the parameter r^2 ($r^2 = 0.1$) with PLINK (Purcell *et al*., 2007). GAPIT (Lipka *et al*., 2012) was used in the kinship matrix, and the first five principal components were used in Farm CPU analysis. The SNPs significantly associated with seven traits were determined at the Bonferroni-corrected threshold $-\log_{10}(P) > 6.00$. Candidate genes associated with target traits were identified in independent genomic positions and determined from obtaining these SNPs with LD along the genome. Finally, Maize Genetics and Genomics Database (http://www.maizedb.org) (Polacco and Coe, 2002) and the U.S. National Center for Biotechnology Information (http://www.ncbi.nlm.nih.gov/) were used to highlight details on the candidate genes.

Relative Expression Analysis of Candidate Genes using RT-PCR

RNA was extracted using TaKaRa MiniBEST Universal Total RNA Extraction Kit. RNA concentration and 260/280 nm ratios were determined with a NanoDrop ND-1000 spectrophotometer (NanoDrop Technologies, Wilmington, DE, USA). PCR primers were designed using Primer Express 2.0 software (Applied Biosystems, Foster City, CA, USA). All RT-PCR analyses of each sample were run in triplicate, and the relative expression of each gene in each trait was measured relative to the housekeeping gene maize translation elongation factor (7) (EF-1α) using the $2^{-\Delta\Delta Ct}$ method (Schmittgen and Livak, 2008).

Statistical Analysis

Descriptive analysis of data collected from two locations and two years was performed using the R statistical package. Variations among genotypes for all phenotypic traits were analyzed using the mean, range standard deviation (SD) and coefficient of variation. Correlations

among multiple traits were characterized using Pearson correlation coefficient (Benesty *et al*., 2009) and the corresponding coefficient was calculated using the R statistical package. The broad-sense heritability of each trait was calculated using lme4 (Bates *et al*., 2014) in the R statistical package.

Results

Phenotypic Evaluation

The frequency distribution of cob hardness widely varied in maize (Fig. 1). Variations ranged from 0.98 to 2.34 (mean ± SD = 1.39 ± 0.28) for bract compaction, 14.1 to 34.5 (mean ± SD = 20.85 ± 3.01) for bract length, 5.33 to 15.33 (mean ± SD = 8.51 ± 1.74) for bract number, 0.22 to 1.32 (mean ± SD = 0.73 ± 0.19) for bract thickness, 1.87 to 3.52 (mean ± SD = 2.54 ± 0.32) for cob diameter, 68 to 676.57 (mean ± SD = 232.21 ± 97.68) for cob hardness, 0.19 to 0.54 (mean ± SD = 0.36 ± 0.07) for cob lignin layer percentage and 0.23 to 0.71 (mean ± SD = 0.45 ± 0.09) for cob lignin layer thickness (Table 1). The broad-sense heritability of these eight traits ranged from 0.78 to 0.90 (Table 1).

Frequency distribution is obtained from the statistics of each trait at 15CC, 15DD, 15CC and 16DD. A is cob hardness, B is cob diameter, C is cob lignin layer thickness, and D is cob lignin layer percentage, E pertains to bract compaction, F is bract length, G is bract number, H is bract thickness. The X axis indicates the measured value and the Y axis represents the relative frequency.

The field grain drying rate (GDR) and grain moisture (GM) values of 80 elite maize inbred lines were obtained from Dai *et al.* (2017). Correlation analysis based on Pearson correlation coefficient showed that four traits (*i.e*., bract thickness, cob hardness, cob diameter, and bract length) shared a significant positive correlation with grain moisture, a significant negative correlation cob lignin layer percentage with grain moisture, and cob diameter with a significant negative correlation with field grain drying rate (Table 2).

Pearson correlation coefficient was also calculated after comparing eight traits with each other. Bract thickness shared a significant negative correlation with cob lignin layer percentage, and a significant positive correlation with bract compaction, length, number and cob hardness. The more bract number, the longer and thicker bract, the thicker and harder cob and the lower cob lignin layer percentage; bract compaction was significantly positively correlated with bract length and bract thickness.

For the cob character, cob diameter shared a significant positive correlation with bract number, cob hardness, and cob lignin layer thickness; The more harder cob, the more and thicker bract, the thicker cob and cob lignin layer; cob lignin layer thickness was significantly positively correlated with cob diameter, cob hardness and cob lignin layer percentage; cob lignin layer percentage

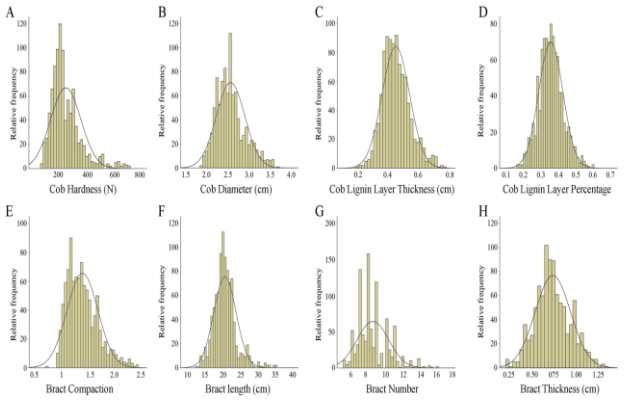

Fig. 1: Frequency distribution of eight traits in samples from four origins

shared a significant positive correlation with cob lignin layer thickness and negative correlated with bract number and cob diameter (Table 3).

SNP Marker Statistics

Initially, a total of 34,872,961 putative SNPs using SAMTOOLS and retained 1,490,007 SNP markers for further analyses after removing the SNPs with special criteria as method description.

Marker-trait Associations for Seven Target Traits in Corn Bracts and Cobs

Based on the extent of LD estimation, the average decay distance across all chromosomes was approximately 10 kb, with $r^2 = 0.1$ (Dai *et al*., 2017). GWAS conducted using a total of 80 maize inbred lines and 1,490,007 SNPs detected a total of 76 significant SNPs above the Bonferroni-corrected threshold -log (P) score of 6.00 that were distributed among 10 maize chromosomes. The Manhattan plots and quantile-quantile (Q-Q) plots for illustrating observed associations between SNPs and seven target traits in corn cobs and bracts were compared to the expected associations (Fig. 2A-F).

Each dot represents an SNP, a-1, b-1, c-1 represents Manhattan plot showing candidate genes identified from significant SNPs, which used dashed horizontal line to depict the Bonferroni-adjusted significance threshold (1.0×10^{-6}); a-2, b-2, c-2 represents Quantile-quantile plot. A a for cob diameter in CC16; A b for cob hardness in CC15; A c for cob hardness in CC16; B a for cob hardness in DD15; B b for cob hardness in DD16; B c for cob lignin layer thickness in CC16; C a for bract compaction in CC16; C b for bract compaction in DD15; C c for bract compaction in DD16; D a for bract thickness in CC16; D b for bract length in CC15; D c for bract length in CC16; E a for bract length in DD15; E b for bract length in DD16; E c for bract number in CC15; F a for bract number in CC16; F b for bract number in DD15; F c for bract number in DD16.

Table 1: Descriptive statistics from ANOVA of eight traits of cobs and bracts in samples from four origins

Trait	Year	Location	Mean	SD	CV%	Range		F-value	Heritability
						Min	Max		
Cob lignin layer thickness (cm)	2016	Dandong	0.45	0.08	0.18	0.28	0.70	9.287**	0.90
	2016	Changchun	0.44	0.09	0.20	0.23	0.71	24.87**	
	2015	Dandong	0.44	0.09	0.19	0.26	0.68	19.67**	
	2015	Changchun	0.45	0.09	0.21	0.27	0.70	416.5**	
	ALL	ALL	0.45	0.09	0.20	0.23	0.71	29.51**	
Cob lignin layer percentage (%)	2016	Dandong	0.36	0.06	0.16	0.23	0.54	6.392**	0.87
	2016	Changchun	0.36	0.07	0.18	0.21	0.53	5.032**	
	2015	Dandong	0.36	0.07	0.19	0.20	0.54	38.99**	
	2015	Changchun	0.35	0.07	0.20	0.19	0.54	76.89**	
	ALL	ALL	0.36	0.07	0.18	0.19	0.54	20.26**	
Cob hardness (N)	2016	Dandong	229.78	81.44	0.35	76.12	515.19	2.089**	0.87
	2016	Changchun	228.98	99.20	0.43	73.73	552.31	21.87**	
	2015	Dandong	234.51	94.92	0.40	76.31	638.71	8.568**	
	2015	Changchun	235.59	112.47	0.48	68	676.57	15.92**	
	ALL	ALL	232.21	97.68	0.42	68	676.57	20.89**	
Cob diameter (cm)	2016	Dandong	2.54	0.29	0.11	1.91	3.51	5.994**	0.89
	2016	Changchun	2.51	0.33	0.13	1.93	3.36	26.36**	
	2015	Dandong	2.50	0.31	0.13	1.93	3.42	37.000**	
	2015	Changchun	2.60	0.34	0.13	1.87	3.52	45.62**	
	ALL	ALL	2.54	0.32	0.13	1.87	3.52	28.07**	
Bract thickness (cm)	2016	Dandong	0.73	0.15	0.21	0.39	1.05	5.889**	0.78
	2016	Changchun	0.74	0.17	0.23	0.22	1.23	835.8**	
	2015	Dandong	0.76	0.24	0.32	0.32	1.09	10.99**	
	2015	Changchun	0.71	0.19	0.26	0.31	1.32	370.8**	
	ALL	ALL	0.73	0.19	0.26	0.22	1.32	16.2**	
Bract number	2016	Dandong	8.77	1.77	0.20	5.72	14.09	7.095***	0.85
	2016	Changchun	8.08	1.44	0.18	5.67	11.33	18.92***	
	2015	Dandong	8.11	1.40	0.17	5.50	11.38	24.83***	
	2015	Changchun	9.09	2.05	0.23	5.33	15.33	24.1***	
	ALL	ALL	8.51	1.74	0.20	5.33	15.33	18.26***	
Bract length (cm)	2016	Dandong	21.14	2.41	0.11	16.17	30.62	4.722***	0.89
	2016	Changchun	20.61	3.30	0.16	14.1	34.5	116.6***	
	2015	Dandong	20.47	3.45	0.17	14.30	33.68	156.9***	
	2015	Changchun	21.20	2.69	0.13	15.7	30.4	16.05***	
	ALL	ALL	20.85	3.01	0.14	14.1	34.5	24.41***	
Bract compaction	2016	Dandong	1.25	0.18	0.14	0.99	1.90	8.582***	0.79
	2016	Changchun	1.56	0.30	0.19	1.01	2.31	320.7***	
	2015	Dandong	1.53	0.29	0.19	1.01	2.34	80.62***	
	2015	Changchun	1.23	0.18	0.14	0.98	1.85	7.071***	
	ALL	ALL	1.39	0.28	0.20	0.98	2.34	12.14***	

Table 2: Pearson correlation coefficient between GM, FDR and eight traits (Only contained significant results)

Trait1	Trait2	Label2	Pearson Correlation Coefficient	P-value
GM	bract_thickness	DD16	0.26	0.02
GM	bract_thickness	DD15	0.30	0.01
GM	bract_thickness	CC16	0.31	0.00
GM	cob_lignin_layer_percentage	CC15	-0.23	0.04
GM	cob_lignin_layer_percentage	DD15	-0.24	0.03
GM	cob_lignin_layer_percentage	CC16	-0.25	0.03
GM	cob_hardness	DD16	0.24	0.03
GM	cob_hardness	CC15	0.31	0.01
GM	cob_hardness	DD15	0.28	0.01
GM	cob_hardness	CC16	0.33	0.00
FDR	cob_diameter	CC15	-0.24	0.03
GM	cob_diameter	CC15	0.32	0.00
GM	bract_length	DD16	0.23	0.04

For three cob traits, eight SNPs distributed across chromosomes 1–4, 6, 8 and 10 were significantly associated with maize cob diameter ($P < 1.0 \times 10^{-6}$) and detected in CC16, none of the SNPs were identified in other locations (CC15, DD15 and DD16).

Six SNPs located in intergenic regions, one SNP was located within an intronic region and another was situated downstream of coding gene. The SNP marker (G/A) at the physical position 31,527,060 on chromosome 10 was the most significantly associated with cob diameter

Table 3: Pearson Correlation coefficient among eight traits from four locations

	Cob hardness	Cob diameter	Cob lignin layer thickness	Cob lignin layer percentage	Bract compaction	Bract length	Bract number	Bract thickness
Cob hardness		**0.00**	**0.00**	0.61	0.83	0.59	**0.02**	**0.02**
Cob diameter	0.39		**0.00**	**0.00**	0.20	0.05	**0.00**	0.38
Cob lignin layer thickness	0.28	0.41		**0.00**	0.81	0.3458	0.75	0.38
Cob lignin layer percentage	0.03	-0.17	0.70		0.23	0.4630	**0.01**	0.63
Bract compaction	0.01	-0.07	0.01	0.07		**0.00**	0.22	**0.03**
Bract length	-0.03	0.11	0.05	0.04	0.23		0.28	**0.00**
Bract number	0.13	0.26	0.02	-0.14	-0.07	0.06		**0.00**
Bract thickness	0.13	0.05	-0.05	-0.03	0.12	0.39	0.16	

Note: The lower left corner is the Pearson Correlation Coefficient, and the upper right corner is the *P*-value

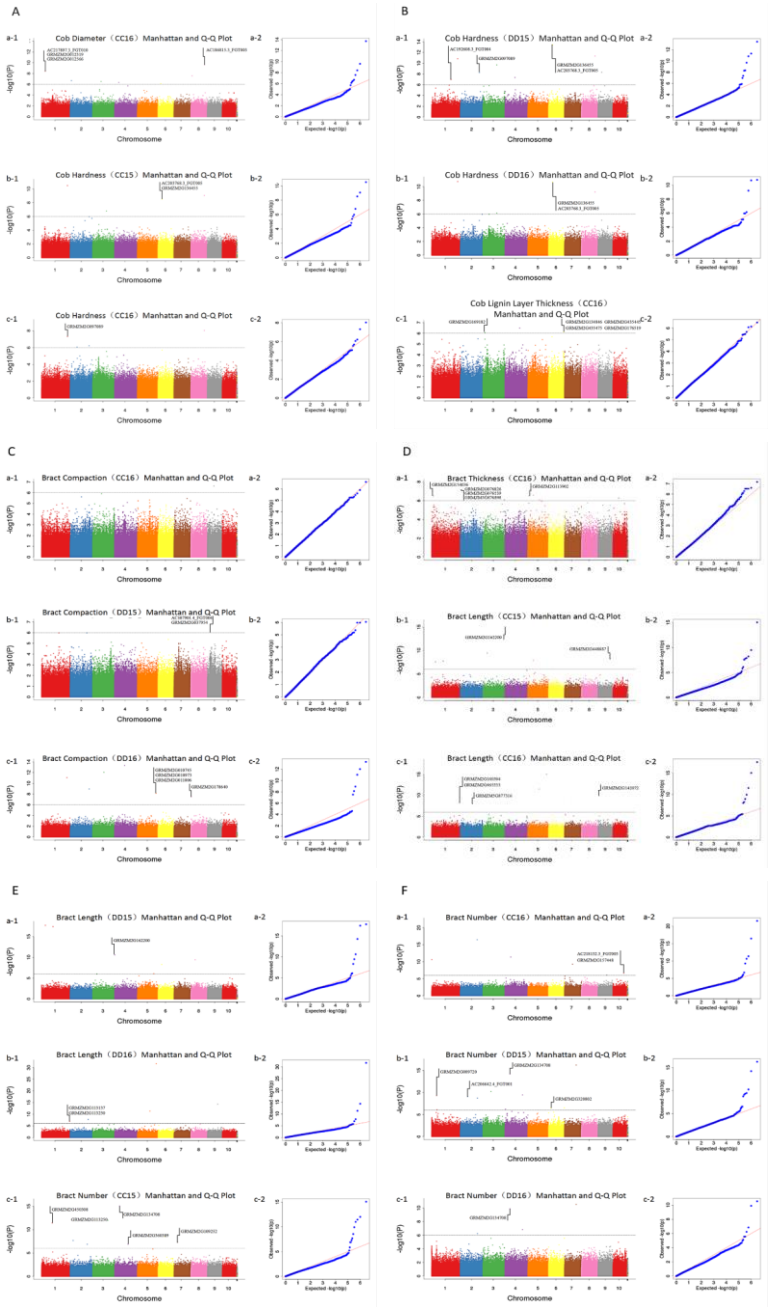

Fig. 2: Manhattan plot and Q-Q plot from the GWAS result for seven traits

as indicated by the lowest *P* value (1.87×10^{-14}), was located in an intergenic region. Four candidate genes, namely, AC217897.3_FG010, GRMZM2G012319_T02, GRMZM2G012566_T02 and AC186815.3_FGT003, were identified among eight significant SNP markers. Nine SNPs were distributed on chromosomes 1–4, 6, 8–9 and significantly associated with maize cob hardness ($P < 1.0 \times 10^{-6}$) were detected in four locations, namely, CC15, CC16, DD15, and DD16. The SNP marker at the physical position 42,892,298 on chromosome 6 was most significantly associated with cob hardness as indicated by the lowest *P* value (3.61×10^{-14}), was identified in three locations (CC15, DD16 and DD15), as well as in exons of two candidate genes (AC203768.3_FGT005 and GRMZM2G136455_T01). Among these SNPs, four candidate genes identified were designated as AC203768.3_FGT005, GRMZM2G136455_T01, GRMZM2G097089_T01 and AC192608.3_FGT004. Three SNPs and five candidate genes located on chromosomes 3, 4 and 6 significantly associated with cob lignin layer thickness were identified in CC16 and none of the SNPs were detected in another location (CC15, DD16 and DD15). The SNP marker at the physical position 154,738,873 on chromosome 4 was the most significantly associated with cob lignin layer thickness as indicated by the lowest *P* value (3.36×10^{-7}), was located in intergenic regions of two genes, and named as GRMZM2G102681_T01 (22,641 bp) and GRMZM2G138407_T01 (85,019 bp).

In terms of bract traits, eight SNP markers were located on chromosomes 1–4, 5, 8, 9 and significantly associated with maize bract compaction ($P < 1.0 \times 10^{-6}$). The SNP marker at the physical position 105,294,960 on chromosome 4 was most significantly associated with bract compaction as indicated by a lowest *P* value (4.65×10^{-14}). Six candidate genes associated with three SNP markers were identified among abovementioned SNPs, namely, GRMZM2G010745_T01, GRMZM2G010973_T01, GRMZM2G011006_T01, GRMZM2G178640_T01, AC187901.4_FGT004 and GRMZM2G037954_T01. Eight SNPs distributed on chromosomes 1–3, 5 and 10, were associated with bract thickness as indicated by a P value lower than 1.0×10^{-6}. The SNP marker at the physical position 8,905,357 on chromosome 2 was most significantly associated with bract thickness as indicated by the lowest *P* value (6.96×10^{-8}) and was located in the exon of three candidate genes named as GRMZM2G076539_T01, GRMZM2G076826_T03, and GRMZM5G876898_T01. Two other candidate genes were also identified among the above mentioned SNP markers and designated as GRMZM2G113902_T01 and GRMZM2G154036_T01. Nineteen SNPs, with *P* values lower than 1.0×10^{-6}, were significantly associated with bract length. The SNP marker at the physical position 197,952,318 on chromosome 5 was most significantly associated with bract length as indicated by the lowest *P* value (2.39×10^{-32}). Among these SNP markers, eight candidate genes were identified and named as GRMZM2G160304_T05, GRMZM2G465553_T01, GRMZM5G877316_T03, GRMZM2G142072_T01, GRMZM2G162200_T01, GRMZM2G448687_T02, GRMZM2G113137_T01, GRMZM2G113250_T01 and GRMZM2G162200_T01. Twenty-one SNPs distributed on chromosomes 1–4, 6, 7, 9, and 10, were significantly associated with bract number as indicated by a *P* value $< 1.0 \times 10^{-6}$. The SNP marker at the physical position 176,535,862 was most significantly associated with bract number as indicated by the lowest P value (3.48×10^{-17}). Nine candidate genes, hereby named as AC218152.3_FGT005, GRMZM2G157448_T01, GRMZM2G450308_T01, GRMZM2G134708_T01, GRMZM2G360389_T01, GRMZM2G109252_T01, GRMZM2G089720_T01, AC206642.4_FGT001, and GRMZM2G320802_T01 were identified among the abovementioned SNPs. In sum, a total of 76 SNPs and 41 candidate genes identified were significantly associated with seven cob and bract traits.

Relative Quantification of Partial Candidate Gene Expression

The relative quantification of 19 genes randomly selected from all candidate genes were analyzed using RT-PCR.

For cob diameter, the relative expression level of three genes (AC217897.3_FGT010, GRMZM2G012319_T02, and GRMZM2G012566_T02) was analyzed, which indicated that these were expressed at higher levels in samples with larger cob diameters. For cob hardness, two candidate genes (AC203768.3_FGT005 and GRMZM2G136455_T01) identified in three sample locations (CC15, DD16 and DD15) were analyzed, which indicated that their relative expression levels increased in multiple samples from high cob hardness to low cob hardness. Five candidate genes significantly associated with cob lignin layer thickness were also analyzed, which showed that three candidate genes (GRMZM2G134846_T01, GRMZM2G176519_T01 and GRMZM2G435445_T01) associated with SNP marker (G/A) at the physical position 158,670,479 of chromosome 6 were expressed at higher levels in samples with thicker cob lignin layers, GRMZM2G435475_T01 and GRMZM2G169182_T02 were expressed at lower levels (Table 4).

The relative expression levels of 9 candidate genes associated with four corn bract traits were analyzed in samples with high and low trait values. For bract compaction, the expression of GRMZM2G010745_T01 was down-regulated in samples with high bract compaction values and the expression of GRMZM2G010973_T01 and GRMZM2G011006_T01 was up-regulated in same samples. For bract thickness, the expression of all three selected candidate genes (GRMZM2G076539_T01 and GRMZM5G876898_T01) was down-regulated in samples with thicker bracts. In terms of bract length, the expression

Table 4: Relative expression levels and Pearson Correlation Coefficient of 19 candidate genes

Gene name	High value	Low value	Pearson Correlation Coefficient
AC217897.3_FGT010	3.91 ± 0.66	1 ± 0.10	0.46
GRMZM2G012319_T02	2.49 ± 0.15	1 ± 0.01	0.53*
GRMZM2G012566_T02	47.41 ±3.91	1 ± 0.07	0.47
GRMZM2G134846_T01	32.50 ±1.11	1 ± 0.18	0.41
GRMZM2G176519_T01	3.38 ± 0.23	1 ± 0.13	0.46
GRMZM2G435445_T01	1.31 ± 0.07	0.97 ± 0.13	0.49
GRMZM2G435475_T01	1 ± 0.07	1.36 ± 0.15	-0.55
GRMZM2G169182_T02	1 ± 0.16	1.29 ± 0.10	-0.26
AC203768.3_FGT005	1 ± 0.05	12.93 ± 1.73	0.73**
GRMZM2G136455_T01	1 ± 0.08	6.12 ± 0.65	0.69*
GRMZM2G162200_T01	3.19 ± 0.14	1 ± 0.04	0.40
GRMZM5G877316_T03	1 ± 0.03	2.42 ± 0.12	0.47*
GRMZM2G142072_T01	1 ± 0.06	2.46 ± 0.10	0.26
GRMZM2G076539_T01	1 ± 0.14	1.88 ± 0.03	0.09
GRMZM5G876898_T01	1 ± 0.04	1.54 ± 0.06	-0.45
GRMZM2G134708_T01	7.22 ± 0.69	1 ± 0.04	0.49*
GRMZM2G010745_T01	1 ± 0.04	1.40 ± 0.04	-0.15
GRMZM2G010973_T01	3.07 ± 0.07	1 ± 0.04	0.39
GRMZM2G011006_T01	1.74 ± 0.02	1 ± 0.04	0.29

Note: * located on Pearson Correlation Coefficient is *P*-value of t-test between relative expression level and phenotype (0.01 < *P*-value < 0.05 marked *, *P*-value < 0.01 marked **)

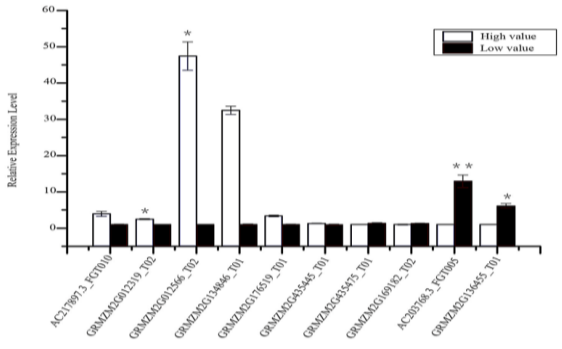

Fig. 3: Relative expression analysis of 10 candidate genes significantly associated with three cob traits

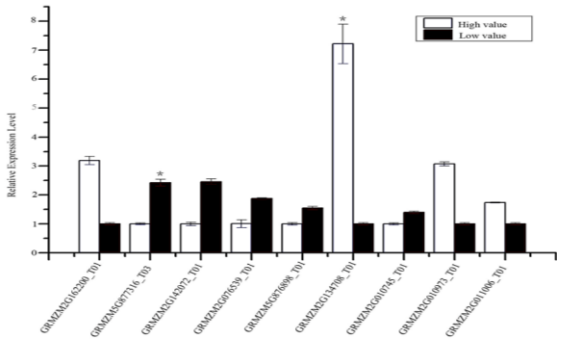

Fig. 4: Relative expression analysis of 9 candidate genes significantly associated with four bract traits

of two candidate genes (GRMZM5G877316_T03 and GRMZM2G142072_T01) was down-regulated in samples with longer bracts and the expression of one candidate gene (GRMZM2G162200_T01) was up-regulated in same samples, GRMZM5G877316_T03 was associated with bract length. Only one candidate gene (GRMZM2G134708_T01) associated with bract number was analyzed and up-regulated in samples with higher number of bracts (Table 4).

The relative expression levels of all analyzed candidate genes and the Pearson Correlation Coefficient are between relative expression levels and phenotype are shown in Fig. 3 (three corn cob traits) Fig. 4 (four corn bract traits).

The X-axis indicates the names of the candidate genes, and the Y-axis indicates $2^{-\triangle\triangle Ct}$, which is comparable to the log-transformed, normalized mRNA abundance. High value and Low value were two extreme trait labels; High value means sample with a high trait value, and Low value means sample with a low trait value. Each group of bars represents the relative expression level of a gene between high and low value. * located on histogram is P-value of t-test between relative expression level and phenotype (0.01< *P*-value < 0.05 marked *, *P*-value < 0.01 marked **).

The X-axis indicates the names of the candidate genes, and the Y-axis indicates $2^{-\triangle\triangle Ct}$, which is comparable to the log-transformed, normalized mRNA abundance. High value and Low value were two extreme trait labels; High value means sample with a high trait value, and Low value means sample with a low trait value. Each group of bars represents the relative expression level of a gene between high and low value. * located on histogram is *P*-value of *t*-test between relative expression level and phenotype (0.01 < *P*-value < 0.05 marked *, *P*-value < 0.01 marked **).

Discussion

China is one of the top three maize-producing country of the world, contributing 24% of the total maize production (Ranum *et al*., 2014). For farmers, threshing is an important step for maize harvesting (Basappa *et al*., 2007), and mechanized threshing decreases the cost of maize threshing by increasing threshing efficiency (Ajaib, 2014). Grain moisture is one of the most important factors for the development of mechanized threshing (Petkevichius *et al*., 2008) and high-grain moisture levels in maize increase the cost of maize threshing by increasing grain damage and decreasing threshing efficiency (Ajaib, 2014).

Field grain dry rate pertains to the rate of grain moisture reduction within the period of physiological maturity to harvest, is an essential determinant of grain moisture at harvest. High moisture at harvest may lead to mildew and yield loss (Jayas and White, 2003). Several candidate genes and SNP markers significantly associated with maize field grain dry rate were identified in various GWASs (Sala *et al*., 2006; Liu *et al*., 2010; Dai *et al*., 2017), which facilitates the elucidation of the genetic basis of this trait among different maize populations. However, efforts in investigating other agronomic traits associated with maize grain moisture and ear threshing are limited, despite the association of numerous agronomic traits with

maize grain moisture and ear threshing.

The present study selected 80 elite inbred lines propagated in northeast China, where maize is mainly produced, and determined eight traits of cobs and bracts associated with grain moisture and threshing of maize in field experiments. According to Dai *et al.* (2017), these inbred lines could be divided into three subpopulations, 19 SNP markers and 6 candidate genes associated with field grain drying rate identified through GWAS. Correlation analyses showed that bract thickness, bract length, cob diameter and cob hardness shared a significant positive correlation with grain moisture, which means that the high values of these traits would increase maize grain moisture, and thus should be considered in further breeding analysis. Cob lignin layer percentage shared a significant negative correlation with grain moisture, which means that maize lines with low cob lignin layer percentages may have low grain moisture and thus would be easy to threshing. The cob diameter also had a significantly negative correlation with field grain dry rate, which coincided with the significantly positive correlation between cob diameter and grain moisture. Cob diameter is also associated with grain production (Bavec and Bavec, 2002), and thus it is important to find a balance between grain moisture and grain production when cob diameter is the trait of interest in breeding.

A total of 76 SNP markers and 41 candidate genes significantly associated with seven traits were identified in this study by using Farm CPU (Wang *et al*., 2017), which uses fixed and random effect models for GWAS. Many studies (Dai *et al*., 2017; Meng *et al*., 2017) showed that this software could detect 50 more QTNs compared with mixed linear model (MLM) under 10% FDR, especially samples with limited populations. However, homologs for seven candidate genes that were significantly associated with cob hardness (AC192608.3_FGT004) and four traits of bract (GRMZM2G154036_T01, GRMZM2G162200_T01, AC218152.3_FGT005, GRMZM2G157448_T01, GRMZM2G178640_T01, and AC187901.4_FGT004) were not identified.

For the three cob traits, two candidate genes, AC203768.3_FGT005 and GRMZM2G136455_T01 significantly associated with cob hardness were identified in three sample origins (CC15, DD16 and DD15). The candidate gene AC203768.3_FGT005 predicted to encode a peptidase family protein differentially expressed between the NKD1 and NKD2 mutation type and the wild-type, thereby suggesting that endosperm gene transcripts are regulated by NKD1 and NKD2 (Gontarek *et al*., 2016). The candidate gene GRMZM2G136455_T01 predicted to encode a conserved myosin tail-binding protein is thought to be involved in Golgi vesicle-mediated transport (Hashimoto *et al*., 2008) and is a gene with *trans*-acting QTLs (Thatcher *et al*., 2014). Li and Thatcher (2015) described that this is as a unique gene that could be utilized to improve specific agronomic traits in maize (Li and Thatcher, 2015). Relative expression analyses showed that both of the genes were up-regulated in samples with high and low cob hardness. The candidate genes GRMZM2G012319_T02 significantly associated with cob diameter were identified in CC16 and predicted encodes a SH3 domain-containing protein. This plays an important role in cell membrane formation (Ahn *et al*., 2017). GRMZM2G176519_T01 predicted to encode CBL-interacting serine/threonine-protein kinase 12 was significantly associated with cob lignin layer thickness and identified in CC16 that could enhances abiotic stress tolerance in plants (Abdula *et al*., 2016).

For the four bract traits, the candidate gene GRMZM2G134708_T01 significantly associated with bract number, was identified in CC15, DD15 and DD16, and predicted to encode probable mono-dehydroascorbate reductase, cytoplasmic isoform 4. It may be involved in the physiological activities of plants against stress (Leterrier *et al*., 2005). This gene has been identified as a member of the maize glutathione-ascorbate cycle gene family, which differentially responds to abiotic stress (Liu *et al*., 2012). Nan *et al.* (2011) described this as a leptotene/zygotene transition gene in maize, which is important for meiotic progression.

Conclusion

Seven traits of cobs and bracts collected from 80 elite inbred lines were analyzed through GWAS. Correlation analyses showed that several target traits shared significant correlations with grain moisture, which in turn suggests that these can be used as target traits for breeding easy-threshing lines. It was also identified several SNP markers and candidate genes were significantly associated with particular traits. Relative expression analysis of partial candidate genes showed that several candidate genes are differentially expressed between high and low-trait value samples. This serves as the first report that aimed to elucidate the genetic basis of agronomic traits that are associated with grain threshing and provides valuable guidance for further molecular breeding.

Acknowledgments

The “948” Project of the Ministry of Agriculture (2013-Z47) and breeding grant of Jilin Province Ministry of Agriculture (2016013) supported this study.

References

Abdula, S.E., H.J. Lee, H. Ryu, K.K. Kang, I. Nou, M.E. Sorrellsand and Y.G. Cho, 2016. Overexpression of BrCIPK1 gene enhances abiotic stress tolerance by increasing proline biosynthesis in rice. *Plant Mol. Biol. Rep.*, 34: 501–511

Ahn, G., H. Kim, D.H. Kim, H. Hanh, Y. Yoon, I. Singaram, K.J. Wijesinghe, K.A. Johnson, X. Zhuang, Z. Liang, R.V. Stahelin, L. Jiang, W. Cho, B.H. Kang and I. Hwang, 2017. SH3P2 plays a crucial role at the step of membrane tubulation during cell plate formation in plants. *Plant Cell*, 29: 1388–1405

Ajaib, S., 2014. Study of different methods of maize harvesting and threshing in Hoshiarpur district of Punjab. *Intl. J. Agric. Biol. Eng.*, 7: 267–270

Basappa, G., J. Deshmanya and B. Patil, 2007. Post-harvest losses of maize crop in karnataka-an economic analysis. *Karnat. J. Agric. Sci.*, 20: 69

Bates, D., M. Maechler, B. Bolker and S. Walker, 2014. lme4: Linear mixed-effects models using Eigen and S4. *R Pack. Vers.*, 1: 1–23

Bavec, F. and M. Bavec, 2002. Effects of plant population on leaf area index, cob characteristics and grain yield of early maturing maize cultivars (FAO 100–400). *Eur. J. Agron.*, 16: 151–159

Benesty, J., J. Chen, Y. Huang and I. Cohen, 2009. *Pearson Correlation Coefficient*, pp: 1–4. In Noise reduction in speech processing. Springer

Cantor, R.M., K. Lange and J.S. Sinsheimer, 2010. Prioritizing GWAS results: A review of statistical methods and recommendations for their application. *Amer. J. Human Genet.*, 86: 6

Christoforou, A., M. Dondrup, M. Mattingsdal, M. Mattheisen, S. Giddaluru, M.M. Nöthen, M. Rietschel, S. Cichon, S. Djurovic, O.A. Andreassen, I. Jonassen, V.M. Steen, P. Puntervoll and S.L. Hellard, 2012. Linkage-disequilibrium-based binning affects the interpretation of GWASs. *Amer. J. Human Genet.*, 90: 727–733

Dai, L., L. Wu, Q. Dong, Z. Zhang, N. Wu, Y. Song, S. Lu and P.W. Wang, 2017. Genome-wide association study of field grain drying rate after physiological maturity based on a resequencing approach in elite maize germplasm. *Euphytica*, 213: 182

Gontarek, B.C., A.K. Neelakandan, H. Wu and P.W. Becraft, 2016. NKD transcription factors are central regulators of maize endosperm development. *Plant Cell*, 28: 1–67

Hashimoto, K., H. Igarashi, S. Mano, C. Takenaka, T. Shiina, M. Yamaguchi, T. Demura, M. Nishimura, T. Shimmen and E. Yokota, 2008. An isoform of Arabidopsis myosin XI interacts with small GTPases in its C-terminal tail region. *J. Exp. Bot.*, 59: 3523–3531

Jayas, D.S. and N.D.G. White, 2003. Storage and drying of grain in Canada: low cost approaches. *Food Contr.*, 14: 255–261

Leterrier, M., F.J. Corpas, J.B. Barroso, L.M. Sandalio and A. Luis, 2005. Peroxisomal monodehydroascorbate reductase genomic clone characterization and functional analysis under environmental stress conditions. *Plant Physiol.*, 138: 2111–2123

Li, B. and S. Thatcher, 2015. *Novel transcripts and uses thereof for improvement of agronomic characteristics in crop plants.*United States Patent Application Publication, 14/628, 469

Li, H. and R. Durbin, 2009. Fast and accurate short read alignment with Burrows–Wheeler transform. *Bioinformatics*, 25: 1754–1760

Li, H., B. Handsaker, A. Wysoker, T. Fennell, J. Ruan, N. Homer, G. Marth, G. Abecasis and R. Durbin, 2009. 1000 Genome Project Data Processing Subgroup. The sequence alignment/map format and SAMtools. *Bioinformatics*, 25: 2078–2079

LI, X.P., L.X. Gao and F. MA, 2007. Experimental research of seed corn on the threshing property. *J. Agric. Mechan. Res.*, 2: 49

Lipka, A.E., F. Tian, Q. Wang, J. Peiffer, M. Li, P.J. Bradbury, M.A. Gore, E.S. Buckler and Z. Zhang, 2012. GAPIT: genome association and prediction integrated tool. *Bioinformatics*, 28: 2397–2399

Liu, Y.J., Y. Yuan, Y.Y. Liu, Y. Liu, J.J. Fu, J. Zheng and G.Y. Wang, 2012. Gene families of maize glutathione–ascorbate redox cycle respond differently to abiotic stresses. *J. Plant Physiol.*, 169: 183–192

Liu, X., M. Huang, B. Fan, E.S. Buckler and Z. Zhang, 2016. Iterative usage of fixed and random effect models for powerful and efficient genome-wide association studies. *PLoS Genet.*, 12: e1005767

Liu, X., Z. Wang, X. Wang, T. Li and L. Zhang, 2010. Primary mapping of QTL for dehydration rate of maize kernel after physiological maturing. *Acta Agron. Sin.*, 36: 47–52

Ma, Z., Y. Dong, D. Qiao, C. Hu, F. Deng and Y. Li, 2015. Analysis on bract traits of different maize hybrids. *J. Henan Agric. Sci.*, 44: 15–18

Meng, Q., K. Wang, X. Liu, H. Zhou, L. Xu, Z. Wang and M. Fang, 2017. Identification of growth trait related genes in a Yorkshire purebred pig population by genome-wide association studies. *Asian Aust. J. Anim. Sci.*, 30: 462

Nan, G.L., A. Ronceret, R.C. Wang, J.F. Fernandes, W.Z. Cande and V. Walbot, 2011. Global transcriptome analysis of two ameiotic1 alleles in maize anthers: defining steps in meiotic entry and progression through prophase I. *BMC Plant Biol.*, 11: 120

Petkevichius. S., L. Shpokas and H.D. Kutzbach, 2008. Investigation of the maize ear threshing process. *Biosci. Eng.*, 99: 532–539

Polacco, M. and E. Coe, 2002. *MaizeDB: The Maize Genome Database*, pp: 151–162. Bioinformatics Databases System

Purcell, S., B. Neale, K. Todd-Brown, L. Thomas, M.A. Ferreira, D. Bender, J. Maller, P. Sklar, P.I. Bakker, M.J. Dalyba and P.C. Sham, 2007. PLINK: a tool set for whole-genome association and population-based linkage analyses. *Amer. J. Human Genet.*, 81: 559–575

Ranum, P., J.P. Peña-Rosas and M.N. Garcia-Casal, 2014. Global maize production, utilization, and consumption. *Ann. N. Y. Acad. Sci.*, 1312: 105–112

Sala, R.G., F.H. Andrade, E.L. Camadro and J.C. Cerono, 2006. Quantitative trait loci for grain moisture at harvest and field grain drying rate in maize (*Zea mays* L.). *Theor. Appl. Genet.*, 112: 462–471

Schnable, P.S., D. Ware, R.S. Fulton, J.C. Stein, F. Wei, S. Pasternak and P. Minx, 2009. The B73 maize genome: complexity, diversity, and dynamics. *Science*, 326: 1112–1115

Schmittgen, T.D. and K.J. Livak, 2008. Analyzing real-time PCR data by the comparative CT method. *Nat. Protoc.*, 3: 1101–1108

Song, W., Z. Shi, J. Xing, M. Duan, A. Su, C. Li, R. Zhang, Y. Zhao, M. Luo, J. Wang and J. Zhao, 2017. Molecular mapping of quantitative trait loci for grain moisture at harvest in maize. *Plant Breed.*, 136: 28–32

Srivastava, S., G. Brychkova, D. Yarmolinsky, A. Soltabayeva, T. Samani and M. Sagi, 2017. Aldehyde Oxidase 4 plays a critical role in delaying silique senescence by catalyzing aldehyde detoxification. *Plant Physiol.*, 173: 1977

Thatcher, S.R., W. Zhou, A. Leonard, B.B. Wang, M. Beatty, G. Zastrow-Hayes, X. Zhao, A. Baumgarten and B. Li, 2014. Genome-wide analysis of alternative splicing in *Zea mays*: landscape and genetic regulation. *Plant Cell*, 26: 3472–3487

Wang, K., M. Li and H. Hakonarson, 2017. ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. *Nucl. Acids Res.*, 38: 164

Xiao, Y., H. Liu, L. Wu, M. Warburton and J. Yan, 2017. Genome-wide association studies in maize: praise and stargaze. *Mol. Plant*, 10: 359–374

(Received 30 August 2018; Accepted 20 November 2018)