Mol Breeding (2017) 37: 30
DOI 10.1007/s11032-017-0634-8

Genome-wide association study (GWAS) of salt tolerance in worldwide soybean germplasm lines

A. Zeng · P. Chen · K. Korth · F. Hancock · A. Pereira · K. Brye · C. Wu · A. Shi

Received: 9 August 2016 / Accepted: 6 February 2017 / Published online: 6 March 2017

Abstract Salt is a severe abiotic stress causing soybean yield loss in saline soils and irrigated fields. Marker-assisted selection (MAS) is a powerful genomic tool for improving the efficiency of breeding salt-tolerant soybean varieties. The objectives of this study were to uncover novel single-nucleotide polymorphisms (SNPs) and quantitative trait loci (QTLs) associated with salt tolerance and to confirm the previously identified genomic regions and SNPs for salt tolerance. A total of 283 diverse soybean plant introductions (PIs) were screened for salt tolerance in the greenhouse based on leaf chloride concentrations and leaf chlorophyll concentrations after 12–18 days of 120-mM NaCl treatment. A total of 33,009 SNPs across 283 genotypes from the Illumina Infinium SoySNP50K BeadChip database were employed in the association analysis with leaf chloride concentrations and leaf chlorophyll concentrations. Genome-wide association mapping showed that 45 SNPs representing nine genomic regions on chromosomes (Chr.) 2, 3, 7, 8, 10, 13, 14, 16, and 20 were significantly associated with both leaf chloride concentrations and leaf chlorophyll concentrations in 2014, 2015, and combined years. A total of 31 SNPs on Chr. 3 were mapped at or near the previously reported major salt tolerance QTL. The significant SNP on Chr. 2 was also in proximity to the previously reported SNP for salt tolerance. The other significant SNPs represent seven putative novel QTLs for salt tolerance. The significant SNP markers on Chr. 2, 3, 14, 16, and 20, which were identified in both general linear model and mixed linear model, were highly recommended for MAS in breeding salt-tolerant soybean varieties.

Keywords Single-nucleotide polymorphisms · Genome-wide association study · Marker-assisted selection · Leaf chloride concentration · Leaf chlorophyll concentration · Salt tolerance

Electronic supplementary material The online version of this article (doi:10.1007/s11032-017-0634-8) contains supplementary material, which is available to authorized users.

A. Zeng · P. Chen (✉) · A. Pereira · K. Brye · C. Wu
Department of Crop, Soil, and Environmental Sciences, University of Arkansas, Fayetteville, AR 72701, USA
e-mail: chenpe@missouri.edu

K. Korth
Department of Plant Pathology, University of Arkansas, Fayetteville, AR 72701, USA

F. Hancock
Former Monsanto Soybean Breeder, 2711 Blacks Ferry Road, Pocahontas, AR 72455, USA

A. Shi
Department of Horticulture, University of Arkansas, Fayetteville, AR 72701, USA

Present Address:
P. Chen
Fisher Delta Research Center, University of Missouri, 147 State Hwy T, Portageville, MO 63873, USA

Introduction

Soybean (*Glycine max*) is grown globally mainly for its protein and oil. However, global salinization of the

arable land caused more than 20% of soybean yield reduction (Beecher 1994; Blumwald and Grover 2006; Katerji et al. 2003). Therefore, maximizing soybean yield potential depends on increases in salt tolerance to some extent. Soybean germplasms range widely in their response to salt stress (Shao et al. 1986). Salt-tolerant soybean lines (chloride excluder) accumulate less chloride in the leaves than salt-sensitive lines (chloride includer) (Lee et al. 2004; Ledesma et al. 2016), whereas chloride excluder has higher leaf chlorophyll concentration than chloride includer under salt stress (Patil et al. 2016). Evaluation of salt tolerance in the greenhouse is time-consuming, labor-intensive, and costly (Valencia et al. 2008; Ledesma et al. 2016), and selection of salt-tolerant lines in the field is not accurate since the salt concentration varies in the field. Bi-parental quantitative trait locus (QTL) mapping has been implemented to reveal the mechanisms for salt tolerance in soybean, and simple sequence repeat (SSR) markers have been reported significantly associated with salt tolerance (Lee et al. 2004; Hamwieh et al. 2011).

Compared to bi-parental QTL mapping, genome-wide association mapping provides more precise location of QTLs. The application of genome-wide association studies (GWAS) is beneficial from the advance of next-generation sequencing (NGS). GWAS has been carried out to identify markers associated with iron deficiency chlorosis (Mamidi et al. 2011), chlorophyll concentration (Hao et al. 2012), seed protein and oil content (Hwang et al. 2014), resistance to sudden death syndrome (SDS) (Wen et al. 2014, 2015), grain yield, lodging, seed coat color, pubescence, flower color (Wen et al. 2015), flowering time, maturity dates, and plant height in soybean (Wen et al. 2015; Zhang et al. 2015).

An important factor to consider in the application of GWAS is the extent of linkage disequilibrium (LD). LD refers to the degree of non-random association of alleles at different loci. The structure of LD across the genome determines the resolution of association mapping (Zhu et al. 2008). The average LD (R^2) decayed to 0.2 within 360 k base pairs (kb) in euchromatic region while decayed to 0.2 within 9600 kb in heterochromatic region of soybean (Hwang et al. 2014). A high marker density is required for the regions with low LD for GWAS (Hwang et al. 2014). The other problem confronted by GWAS is the potential spurious association caused by population structure and familiar relatedness. To control the false association error rate, general linear model (GLM) considering population structure (Pritchard and Donnelly 2001) has been initially employed. Mixed linear model (MLM)-based methods such as unified mixed model (Yu et al. 2006) and compressed MLM (Zhang et al. 2010) have been used to correct for genetic relatedness and population structure.

Although salt tolerance in soybean has been studied using various germplasms including domesticated soybean (*G. max*) and wild soybean (*Glycine soja*), the major soybean QTL conferring salt tolerance has been consistently mapped on chromosome (Chr.) 3 (Lee et al. 2004; Hamwieh and Xu 2008; Hamwieh et al. 2011; Guan et al. 2014; Qi et al. 2014; Concibido et al. 2015). Thus, a worldwide collection of diverse soybean plant introductions (PIs) included in GWAS provides a potential broad genetic basis for underlying the mechanism of salt tolerance. The availability of SoySNP50K has paved the way for identification of single-nucleotide polymorphism (SNP) markers associated with salt tolerance in soybean using GWAS. The GWAS of salt tolerance at vegetative 1 (V1) stage (one set of unfolded trifoliolate leaves) was initially carried out by Huang (2013) using 192 diverse soybean germplasm lines, of which 61% were originated from USA. A total of 62 SNP markers on Chr. 2, 3, 5, 6, 8, and 18 from SoySNP50K data were significantly associated with leaf scorch score (LSS) at V1 stage under salt stress (Huang 2013). Kan et al. (2015) conducted GWAS on an association mapping panel consisting of 191 soybean landraces for salt tolerance at germination stage. One SNP on Chr. 9 and seven SNPs on Chr. 2, 3, 9, 12, and 13 were significantly associated with germination index ratio and germination rate ratio, respectively (Kan et al. 2015). Patil et al. (2016) performed GWAS on a panel of 106 soybean lines for salt tolerance at V2 stage (two sets of unfolded trifoliolate leaves). A total of 19 and 11 SNPs on Chr. 3 from SoySNP50K data were associated with LSS and leaf chlorophyll concentrations which were expressed as soil plant analysis development (SPAD) value, respectively (Patil et al. 2016). In this study, we collected 283 PI lines distributed in 29 countries worldwide to provide a wide genetic basis for GWAS. Two salt tolerance trait indicators, leaf chloride concentrations and leaf chlorophyll concentrations, were utilized for salt tolerance evaluation at V1 stage. The objectives of this study were to uncover novel genomic regions and SNP markers for salt tolerance and to confirm the previously identified genomic regions and SNP markers associated with salt tolerance by GWAS.

Materials and methods

Plant materials and evaluation of salt tolerance in greenhouse

A total of 283 PIs (Supplement Table 1) were obtained from the USDA Soybean Germplasm Collection. Maturity group (MG) of these PIs ranged from 000 to VIII. Forty-nine percent of PIs were from USA, Russia, China, Germany, and Bulgaria; the rest were from other 24 different countries. Of the PIs, 283 were planted in a randomized complete block design with two replications in December 2014 and June 2015, in greenhouse (25 ± 2 °C, 14-h photoperiod) of Rosen Center at University of Arkansas. Chloride excluder (salt-tolerant) cultivar Osage (Chen et al. 2007) and chloride includer (salt-sensitive) cultivar Dare (Brim 1966) were used as checks. For each line, ten seeds were sown in a 3.5-in. plastic pot (Plasticflowerpots.net, Lake Worth, FI), containing approximately 300 g loamy sand (Kibler, AR) as the growth medium. Soil particle analysis based on a 2-h hydrometer method described by Arshad et al. (1996) showed that the sandy loam consists of 66% sand, 26% clay, and 8% silt. Pots were placed in trays (17 3/4″ × 25 1/2″ × 1″; US Plastic Corp., Lima, OH) for watering and salt treatment. At VC stage (unifoliolate leaves unrolled), plants were thinned to five per pot. Seedlings were fertilized once per week by Miracle-Gro®, All Purpose Plant Food (The Scotts Miracle-Gro Company, Marysville, OH). Salt treatment using 120 mM NaCl solution was initiated at V1 stage (one set of unfolded trifoliolate leaves). Salt treatments were performed for 2 h per day and continued until the checks were showing contrasting leaf symptom. It took 12–18 days. After the last day of treatment, the measurements of leaf chlorophyll concentrations were conducted on the top secondary fully expanded leaves for three times. Leaf chlorophyll concentrations, expressed as SPAD value, were measured by a chlorophyll meter (Konica Minolta SPAD-502). Subsequently, the plant leaves in each pot were harvested and were dried in a forage dryer. The leaf samples were ground to fine powder using a coffee bean grinder (Krups®, Shelton, CT). The fine powder sample was used for chloride extraction and quantified by inductively coupled plasma–optical emission spectroscopy (ICP-OES) as described by Wheal and Lyndon (2010). The chloride concentration was converted into mg kg^{-1}.

Analysis of variance and heritability

The descriptive statistics for the leaf chloride concentration and leaf chlorophyll concentration of the populations and checks were obtained from JMP 9.0 (SAS Institute Inc. 1989). Broad-sense heritability (H^2) of leaf chloride concentration and leaf chlorophyll concentration were calculated using the following equation (Greene et al. 2008): $H^2 = \frac{\sigma_g^2}{\sigma_g^2 + \frac{\sigma_{gxe}^2}{e} + \frac{\sigma^2}{re}}$, where σ_g^2 is the genetic variance, σ_{gxe}^2 is the genotype by year interaction, σ^2 is the error variance, r is the number of replications, and e is the number of years. Analysis of variance (ANOVA) and the estimation of variance components were performed using the PROC GLM procedure of SAS 9.4 (SAS Institute 2004), where genotype was considered as a random effect and replication nested with year was used as a random effect.

Genotyping and quality control

Genotypic data of ~42,509 SNPs for a total of 283 soybean genotypes were obtained from the Illumina Infinium SoySNP50K BeadChip database (Song et al. 2013). A total of 290 SNPs, which were not assigned to any chromosome, were excluded from further analysis. Markers with missing rate >2% and minor allele frequency (MAF) <5% were excluded from further analysis. Subsequently, a total of 33,009 SNPs (Supplement Table 2) with MAF ≥5% across 283 genotypes were employed in the association analysis with leaf chloride concentration and leaf chlorophyll concentration.

Linkage disequilibrium estimation

Pairwise LD between markers was calculated as squared correlation coefficient (R^2) of alleles using 33,009 SNPs. The calculation of LD was based on 10 M base pairs (Mb) windows using R package synbreed (Wimmer et al. 2012). The R^2 was calculated separately for euchromatic and heterochromatic regions in each chromosome because of the substantial difference in recombination rate between these two regions. The physical lengths of euchromatic and heterochromatic regions (Supplement Table 2) were obtained from Gmax 1.01 reference genome (Grant et al. 2010). The LD decay rate of the population, defined as the chromosomal distance where the LD decays to half of its maximum

value (Huang et al. 2010), was calculated using R script developed by Marroni et al. (2011).

Population structure

All the 33,009 SNP markers were sorted by chromosome and physical distance, and one SNP marker was selected every ten SNP markers based on the physical distance; 3301 out of 33,009 SNP markers were used to infer the population structure by STRUCTURE 2.3.4 (Pritchard et al. 2000). The hypothetical number of subpopulations (K) was set from 1 to 10, and five independent iterations were performed for each K. Admixture and allele frequency correlated models were used. The burn-in iteration was 10,000, followed by 25,000 Markov chain Monte Carlo (MCMC) replications. The optimum value of K was determined by plotting the rate of change in the log probability of data (ΔK) against the successive K values (Evanno et al. 2005). The K value was considered to be optimum while ΔK reaches the maximum. The population structure (Q matrix) was generated as the STRUCTURE result. Missing genotypes were imputed by k nearest neighbors with Euclidean distance, and the kinship matrix (K) was calculated by centered-IBS method (Endelman and Jannink 2012) using TASSEL 5.0.

Genome-wide association analysis

GLM considering population structure (Q matrix) and MLM accounting for population structure (Q matrix) and kinship (K matrix) were implemented in the TASSEL 5.0 (Bradbury et al. 2007) for genome-wide association analysis. For the GLM analysis, the equation was $y = \mu + X\alpha + P\beta + e$; for MLM analysis, the equation was $y = \mu + X\alpha + P\beta + Zu + e$, where y is $N \times 1$ vector of best liner unbiased predictors (BLUPs) of genetic effect (N is the population size), μ is the overall mean value, X is the incidence matrix relating to the plant introduction lines to the marker effect α, P is the incidence matrix relating to the plant introduction lines to population structure effect β, Z is the incidence matrix relating to the plant introduction lines to kinship effect u, and e is the random error term. For GLM with Q matrix model, 10,000 permutation runs were conducted to find out the significant SNP markers associated with leaf chloride concentration and leaf chlorophyll concentration (Bradbury et al. 2007). The SNPs with $-\log10\,(P) > 4.1$ or $p < 7.9 \times 10^{-5}$ in GLM and the SNPs with $-\log10\,(P) > 2.1$ or $p < 7.9 \times 10^{-3}$ in MLM were considered to be significant. For MLM, optimum compression level was used in TASSEL 5.0 (Bradbury et al. 2007). Manhattan plots of $-\log10\,(P)$ values for each SNP vs. chromosomal position were generated as the TASSEL results.

Results

Phenotypic data

The leaf chloride concentrations of 283 PIs ranged from 21,985 to 106,399 with an average of 64,056 mg kg^{-1} in 2014 and ranged from 6295 to 83,350 with an average of 35,730 mg kg^{-1} in 2015 (Supplement Table 3 and Supplement Fig. 1). The average leaf chloride concentration of chloride excluder Osage was 32,636 and 7383 mg kg^{-1} in 2014 and 2015, respectively (Supplement Table 3). In contrast, the average leaf chloride concentration of chloride includer Dare was 89,209 and 47,525 mg kg^{-1} in 2014 and 2015, respectively (Supplement Table 3). Leaf chlorophyll concentrations were significantly negatively correlated with leaf chloride concentrations under salt stress. The chloride excluder Osage had higher leaf chlorophyll concentrations than chloride includer Dare (Supplement Table 3). The leaf chlorophyll concentrations of 283 PIs ranged from 16.4 to 43.7 with an average of 29.9 in 2014 and ranged from 15.6 to 44.9 with an average of 32.3 in 2015 (Supplement Table 3 and Supplement Fig. 2). Both genotype and year variances were significant. However, both leaf chloride concentrations and leaf chlorophyll concentrations from two years were highly correlated, and most of the genotypes in the population ranked similarly between two years, as reflected by the insignificantly genotype × year variance components. As a result, relatively high broad-sense heritability (H^2) estimates (0.76 for leaf chloride concentrationt, 0.65 for leaf chlorophyll concentration) with the year as the environment factor were obtained based on variance components.

Distribution of SNP markers, linkage disequilibrium, and population structure

A total of 33,009 SNPs were employed for GWAS analysis of salt tolerance traits, resulting in a marker density of 59 SNPs per Mb in euchromatic region and

18 SNPs per Mb in heterochromatic region (Supplement Table 2). MAF of SNPs ranged from 0.05 to 0.50 with an average of 0.30 (Supplement Fig. 3). The LD decayed at 348 and 4838 kb in euchromatic region and heterochromatic region, respectively (Supplement Table 4). STRUCTURE analysis indicated that the calculated ΔK reached the maximum while $K = 3$, suggesting that three subpopulations contain all PIs with the greatest possibility (Supplement Fig. 4 and Fig. 1). Significant divergence among subpopulations and average distance among populations in the same population were obtained (Supplement Table 5). None of the subpopulations had PI lines exclusively from one country (Supplement Table 5 and Supplement Fig. 4).

Genome-wide association analysis

For GLM, a total of 45 SNPs distributed on Chr. 2, 3, 7, 8, 10, 13, 14, 16, and 20 were significantly associated with both leaf chloride concentration and leaf chlorophyll concentration in 2014, 2015, and combined years (Table 1). The GWAS analysis based on the average phenotypic data over years indicated that the major alleles on Chr. 2 and Chr. 7 decreased the leaf chloride concentration by up to 17,492 mg kg^{-1} and increased the leaf chlorophyll concentrations by up to 6.6; meanwhile, three major alleles on Chr. 3 decreased the chloride concentration by up to 14,809 mg kg^{-1} and increased the leaf chlorophyll concentration by up to 4.6. On the other hand, other 36 major alleles on Chr. 3, 8, 10, 13, 14, 16, and 20 increased the leaf chloride concentration by up to 26,345 mg kg^{-1} and decreased leaf chlorophyll by up to 6.9. Overall, the significant SNPs associated with salt tolerance explained 8–52% of phenotypic variation for leaf chloride concentrations and 8–42% of phenotypic variation in leaf chlorophyll concentrations. For MLM, a total of 47 SNPs on Chr. 3 were significantly associated with both leaf chloride concentrations and leaf chlorophyll concentrations in 2014, 2015, and combined years. Among those 47 markers, 27 significant SNPs on Chr. 3 which were stable across years and traits were detected in both GLM and MLM (Table 2). The SNP markers ss715581136 on Chr. 2 and ss715637438 on Chr. 20, which were significantly associated with both leaf chloride concentrations and leaf chlorophyll concentrations in 2014, 2015, and combined years in GLM (Table 1), were also detected to be significantly associated with leaf chloride concentrations in 2014, 2015, and combined years in MLM (Table 2 and Fig. 2). Also, the SNP markers ss715618712 on Chr. 14 and ss715624611 on Chr. 16, which were significantly associated with both leaf chloride concentrations and leaf chlorophyll concentrations in 2014, 2015, and combined years in GLM (Table 1), were also identified to be significantly associated with leaf chlorophyll concentrations in 2014, 2015, and combined years in MLM (Table 2 and Fig. 3).

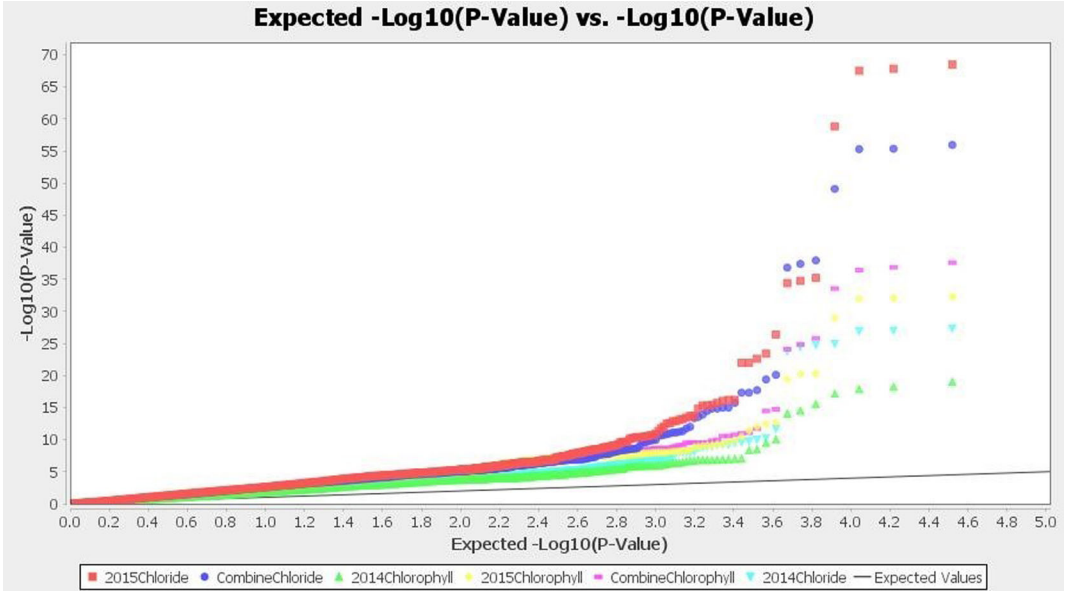

Fig. 1 QQ plot for GLM analysis (expected −log10 (*P* value) plot against −log10 (*P* value)

Table 1 List of SNPs significantly associated with both leaf chloride concentrations (mg kg^{-1}) and leaf chlorophyll concentrations over two years based on GLM ($-\log10\ P \geq 4.1$, $p < -7.9 \times 10^{-5}$)

SNP ID	Chr[a]	Position (bp)	Major[b]/minor[c]	MAF[d]	Year	Leaf chloride concentration			Leaf chlorophyll concentration (SPAD value)		
						−log10 *P*	R^2	Diff[f]	−log10 *P*	R^2	Diff
ss715581136	2	13,098,947	T/C	0.10	2014	5.5	7.2	−16,729.2	5.8	7.8	5.2
					2015	7.4	9.8	−18,253.8	8.9	11.3	8.0
					Com[e]	7.8	10.2	−17,491.5	10.0	13.0	6.6
ss715585865	3	37,902,931	C/T	0.25	2014	6.0	7.8	14,138.3	4.3	5.6	−3.5
					2015	13.1	17.0	16,814.3	5.8	7.2	−4.3
					Com	10.9	14.1	15,476.3	6.8	8.7	−3.9
ss715585870	3	37,984,044	A/G	0.20	2014	6.9	9.0	15,738.7	4.3	5.5	−3.8
					2015	13.4	17.4	18,487.0	5.3	6.7	−5.0
					Com	11.8	15.2	17,112.8	6.5	8.3	−4.4
ss715585874	3	38,037,948	T/C	0.33	2014	9.3	12.1	15,284.8	5.1	6.6	−3.2
					2015	16.2	20.7	15,661.9	5.6	7.1	−3.2
					Com	15.0	19.2	15,473.3	7.3	9.3	−3.2
ss715585876	3	38,045,898	C/T	0.32	2014	8.8	11.6	14,935.9	4.8	6.3	−3.1
					2015	15.7	20.1	15,468.0	5.4	6.8	−3.2
					Com	14.5	18.5	15,201.9	7.0	8.9	−3.2
ss715585877	3	38,047,240	A/G	0.32	2014	8.0	10.8	14,299.1	4.4	5.8	−3.0
					2015	14.8	19.6	14,986.4	4.9	6.3	−3.0
					Com	13.4	17.8	14,642.7	6.3	8.3	−3.0
ss715585878	3	38,048,300	C/T	0.33	2014	9.3	12.1	15,284.8	5.1	6.6	−3.2
					2015	16.2	20.7	15,661.9	5.6	7.1	−3.2
					Com	15.0	19.2	15,473.3	7.3	9.3	−3.2
ss715585883	3	38,103,441	C/A	0.25	2014	11.7	15.3	18,764.7	6.4	8.4	−4.1
					2015	22.6	28.0	20,931.6	9.9	12.7	−5.5
					Com	20.1	25.1	19,848.2	11.1	14.3	−4.8
ss715585889	3	38,139,117	T/C	0.33	2014	9.6	12.7	14,771.1	10.1	13.4	−4.4
					2015	15.3	19.9	15,475.0	6.3	8.1	−3.7
					Com	14.8	19.2	15,123.0	11.0	14.2	−4.0
ss715585890	3	38,140,216	A/G	0.33	2014	9.5	12.5	14,597.5	9.5	12.7	−4.2
					2015	15.4	19.8	15,459.0	6.2	7.8	−3.7
					Com	14.8	19.0	15,028.3	10.6	13.6	−4.0
ss715585896	3	38,183,067	A/C	0.50	2014	10.3	13.5	15,343.0	6.6	8.8	−3.5
					2015	16.0	20.6	14,840.0	7.3	9.3	−3.5
					Com	15.7	20.1	15,091.5	9.5	12.3	−3.5
ss715585899	3	38,197,459	C/A	0.31	2014	5.2	6.7	−12,815.7	4.8	6.2	3.4
					2015	8.5	11.1	−12,659.8	4.6	5.7	3.3
					Com	8.1	10.4	−12,737.7	6.3	8.0	3.3
ss715585930	3	38,480,306	A/G	0.28	2014	24.7	30.5	24,001.8	17.1	22.4	−6.0
					2015	58.8	57.3	27,417.7	29.0	34.0	−7.8
					Com	49.1	50.9	25,709.8	33.5	38.2	−6.9
ss715585934	3	38,532,213	C/A	0.40	2014	24.9	30.8	21,744.5	15.6	20.4	−5.0
					2015	35.2	40.5	20,225.9	20.3	25.1	−5.7
					Com	38.0	42.6	20,985.2	25.8	30.9	−5.3

Table 1 (continued)

SNP ID	Chr[a]	Position (bp)	Major[b]/minor[c]	MAF[d]	Year	Leaf chloride concentration			Leaf chlorophyll concentration (SPAD value)		
						−log10 *P*	R^2	Diff[f]	−log10 *P*	R^2	Diff
ss715585936	3	38,543,691	A/G	0.40	2014	24.4	30.5	21,478.3	14.1	18.8	−4.8
					2015	34.7	40.3	20,079.4	19.4	24.5	−5.5
					Com	37.4	42.4	20,778.9	24.1	29.6	−5.2
ss715585937	3	38,551,155	G/A	0.30	2014	27.0	33.1	24,640.9	19.0	24.8	−6.1
					2015	67.5	63.1	27,710.6	32.0	37.3	−7.8
					Com	55.3	55.5	26,175.8	37.6	42.1	−7.0
ss715585942	3	38,571,016	G/A	0.39	2014	23.7	29.8	21,304.1	14.5	19.3	−4.8
					2015	34.4	40.2	20,226.1	20.2	25.5	−5.8
					Com	36.8	41.9	20,765.1	24.8	30.2	−5.3
ss715585943	3	38,579,634	G/A	0.30	2014	27.3	33.1	24,674.3	18.2	23.7	−6.0
					2015	68.5	62.6	28,016.8	32.3	37.0	−7.8
					Com	55.9	55.3	26,345.6	36.9	41.1	−6.9
ss715585944	3	38,583,144	T/C	0.37	2014	9.8	12.9	−15,400.2	8.4	11.2	4.0
					2015	12.5	16.4	−13,890.1	12.5	16.0	5.0
					Com	13.6	17.6	−14,645.1	14.5	18.5	4.5
ss715585945	3	38,585,840	G/A	0.37	2014	10.0	13.2	−15,504.1	8.5	11.3	4.0
					2015	12.9	16.8	−14,113.0	12.7	16.2	5.1
					Com	13.9	17.9	−14,808.5	14.7	18.7	4.6
ss715585948	3	38,591,888	T/C	0.30	2014	26.9	32.8	24,535.2	17.9	23.3	−6.0
					2015	67.9	62.2	27,983.9	32.1	36.9	−7.9
					Com	55.4	54.9	26,259.6	36.4	40.7	−6.9
ss715585953	3	38,649,377	A/G	0.31	2014	9.2	12.1	16,241.4	6.7	8.9	−4.1
					2015	26.4	31.9	21,247.7	10.5	13.4	−5.5
					Com	19.4	24.3	18,744.6	11.7	15.1	−4.8
ss715586057	3	39,303,211	G/A	0.14	2014	4.7	6.0	10,449.0	6.0	7.9	−3.6
					2015	10.6	13.9	13,989.5	6.0	7.7	−3.8
					Com	8.6	11.3	12,219.2	8.1	10.6	−3.7
ss715586063	3	39,357,229	C/T	0.15	2014	5.7	7.4	6,821.1	6.7	9.0	−6.0
					2015	13.2	17.2	18,646.9	6.1	7.8	1.6
					Com	10.7	13.9	12,734.0	8.7	11.2	−2.2
ss715586070	3	39,403,852	G/T	0.14	2014	5.2	6.7	10,704.0	7.0	9.3	−3.9
					2015	11.8	15.4	14,607.8	5.2	6.5	−3.2
					Com	9.5	12.5	12,655.9	8.1	10.5	−3.5
ss715586082	3	39,510,751	G/T	0.12	2014	5.3	6.8	10,660.5	6.1	8.1	−3.7
					2015	12.6	16.5	16,298.9	5.7	7.3	−3.9
					Com	10.0	13.0	13,479.7	8.0	10.4	−3.8
ss715586086	3	39,600,402	T/C	0.12	2014	4.5	5.7	9,411.5	5.1	6.6	−3.2
					2015	9.2	12.1	13,254.8	4.6	5.7	−3.3
					Com	7.8	10.1	11,333.1	6.5	8.3	−3.2
ss715586087	3	39,601,253	G/A	0.12	2014	4.5	5.6	9,380.0	5.0	6.5	−3.2
					2015	9.1	11.9	13,173.0	4.5	5.6	−3.3
					Com	7.7	10.0	11,276.5	6.4	8.1	−3.2
ss715586094	3	39,693,677	A/G	0.12	2014	4.5	5.6	9,380.0	5.0	6.5	−3.2

Table 1 (continued)

SNP ID	Chr[a]	Position (bp)	Major[b]/minor[c]	MAF[d]	Year	Leaf chloride concentration			Leaf chlorophyll concentration (SPAD value)		
						−log10 *P*	R^2	Diff[f]	−log10 *P*	R^2	Diff
					2015	9.1	11.9	13,173.0	4.5	5.6	−3.3
					Com	7.7	10.0	11,276.5	6.4	8.1	−3.2
ss715586095	3	39,708,387	A/G	0.12	2014	4.4	5.6	9,275.2	4.7	6.2	−3.1
					2015	8.6	11.2	12,737.2	4.5	5.6	−3.3
					Com	7.4	9.6	11,006.2	6.2	8.0	−3.2
ss715586102	3	39,759,678	T/C	0.13	2014	4.5	5.7	9,365.6	5.8	7.7	−3.4
					2015	10.9	14.3	14,022.6	5.8	7.3	−3.5
					Com	8.5	11.1	11,694.1	7.9	10.1	−3.4
ss715586154	3	40,440,832	G/A	0.05	2014	5.0	6.5	18,987.6	6.0	8.1	−6.5
					2015	8.0	10.6	21,370.8	4.4	5.6	−5.6
					Com	7.6	10.1	20,179.2	7.0	9.2	−6.0
ss715597790	7	38,606,673	G/A	0.16	2014	4.1	5.2	−12,669.9	4.2	5.3	3.9
					2015	4.4	5.5	−12,399.1	7.2	9.3	6.0
					Com	5.1	6.5	−12,534.5	7.7	9.9	5.0
ss715597794	7	38,623,703	C/T	0.16	2014	4.3	5.4	−12,727.0	4.2	5.4	3.9
					2015	4.5	5.6	−12,470.2	7.5	9.6	6.0
					Com	5.3	6.7	−12,598.6	7.9	10.2	5.0
ss715597821	7	38,800,856	T/C	0.16	2014	5.1	6.5	−13,710.6	4.9	6.3	4.2
					2015	6.9	9.0	−15,024.7	11.4	14.7	7.3
					Com	7.1	9.2	−14,367.7	10.8	14.0	5.7
ss715601563	8	37,018,844	T/C	0.36	2014	4.2	5.2	8,032.1	4.5	5.8	−2.8
					2015	6.5	8.4	10,377.4	4.9	6.1	−3.6
					Com	6.3	8.0	9,204.8	6.4	8.1	−3.2
ss715607372	10	44,402,011	C/T	0.48	2014	5.1	6.6	3,883.4	4.5	5.8	−1.3
					2015	6.2	8.2	5,139.0	4.5	5.7	−1.5
					Com	6.8	8.8	4,511.2	6.1	7.8	−1.4
ss715607376	10	44,420,445	C/T	0.48	2014	5.0	6.5	3,530.5	5.0	6.6	−1.5
					2015	6.4	8.4	4,978.3	4.7	5.9	−1.5
					Com	6.8	8.9	4,254.4	6.5	8.5	−1.5
ss715616115	13	39,607,602	C/A	0.28	2014	4.8	6.1	−8,511.2	4.7	6.0	2.7
					2015	10.1	13.2	−12,224.4	6.2	7.8	3.6
					Com	8.4	10.9	−10,367.8	7.3	9.4	3.2
ss715618138	14	2,684,853	C/A	0.22	2014	6.8	9.0	14,149.3	5.5	7.2	−3.6
					2015	7.2	9.5	11,753.0	6.5	8.3	−3.7
					Com	8.5	11.2	12,951.2	8.1	10.6	−3.7
ss715618149	14	2,714,940	T/C	0.23	2014	6.1	8.0	13,056.8	4.6	5.9	−3.2
					2015	6.5	8.5	10,662.8	5.3	6.7	−3.1
					Com	7.6	10.0	11,859.8	6.7	8.7	−3.1
ss715618712	14	4,123,566	G/T	0.12	2014	6.5	8.5	15,631.8	6.4	8.4	−4.6
					2015	4.6	5.7	10,919.0	6.4	8.1	−4.9
					Com	6.7	8.7	13,275.4	8.7	11.2	−4.8
ss715618731	14	4,161,326	G/A	0.28	2014	8.0	10.5	13,499.2	6.5	8.6	−5.2
					2015	5.3	6.8	6,285.6	5.1	6.4	−3.3

Table 1 (continued)

SNP ID	Chr[a]	Position (bp)	Major[b]/minor[c]	MAF[d]	Year	Leaf chloride concentration			Leaf chlorophyll concentration (SPAD value)		
						−log10 *P*	R^2	Diff[f]	−log10 *P*	R^2	Diff
					Com	8.1	10.5	9,892.4	7.8	10.1	−4.3
ss715624611	16	33,415,484	G/A	0.13	2014	4.8	6.1	12,311.8	6.1	8.0	−4.0
					2015	6.2	8.0	11,121.3	7.0	8.9	−3.9
					Com	6.5	8.4	11,716.5	8.9	11.5	−4.0
ss715637438	20	34,422,775	G/A	0.30	2014	5.7	7.5	−4,835.5	4.2	5.4	0.8
					2015	5.8	7.6	−3,335.0	5.5	7.0	0.9
					Com	7.0	9.2	−4,085.3	6.5	8.4	0.9

[a] Chromosome

[b] Major allele

[c] Minor allele

[d] Minor allele frequency

[e] Pooled data from 2014 and 2015

[f] Difference is calculated using the following formula: average major allele effect − average minor allele effect; positive sign in difference of leaf chloride concentration indicates that minor allele is beneficial for salt tolerance, while negative sign in difference of leaf chloride concentration indicates that major allele is beneficial for salt tolerance. Positive sign in difference of leaf chlorophyll concentration indicates that major allele is beneficial for salt tolerance, while negative sign in difference of leaf chlorophyll concentration indicates that minor allele is beneficial for salt tolerance)

Discussion

The increases in NaCl concentrations were significantly associated with leaf chloride concentrations (Ledesma et al. 2016) and leaf chlorophyll concentrations (Lenis et al. 2011). Both leaf chloride concentrations and leaf chlorophyll concentrations, which are more objectively than visually scoring of leaf scorch, were used as the indicators of salt tolerance in this study. Leaf chloride concentrations were negatively correlated with leaf chlorophyll concentrations under salt stress (Patil et al. 2016). The significant negative correlation between leaf chloride concentrations and leaf chlorophyll concentrations in this study indicated that a reliable phenotypic data was generated. An important feature of GWAS is the broad genetic variation in the mapping population, which consisted of diverse germplasm resources. In order to capture the possible alleles relating to salt tolerance, 283 plant introduction lines were collected from 29 different countries. The mapping population showed a wide range of leaf chloride concentrations and leaf chlorophyll concentrations (Supplement Table 3), indicating that salt tolerance is controlled by a few quantitative trait loci. Variation of genotype × year interaction was not significant for leaf chloride concentrations and leaf chlorophyll concentrations, as expected for controlled environments in the greenhouse. However, the variation of year effect was significant. It is likely because pots were exposed to the salt treatment for different durations in 2014 and 2015, respectively, since the treatments were continued until the checks started to show the contrasting foliar symptoms in each year, which continued 18 and 12 days for 2014 and 2015, respectively. Significant differences among genotypes were present for both indicators of salt tolerance.

The mapping resolution and marker density required for an effective association analysis were determined by LD decay distance (Zhu et al. 2008). To obtain high-resolution mapping, a short LD decay distance requires greater number of markers (Zhu et al. 2008). The LD decay distances in euchromatic and heterochromatic regions, which were 348 and 4838 kb, respectively (Supplement Table 4), were similar to those reported by Zhang et al. (2016). The 33,009 SNPs gave an average marker distance of 16.9 and 55.6 kb in euchromatic and heterochromatic regions, respectively, which were much lower than the LD decay distance. Therefore, high density of SNPs in this study provided a robust genotypic data for the association analysis.

Table 2 List of common significant SNPs between GLM and MLM

Chromosome	Number of common SNPs	List of common SNPs
2	1	ss715581136
		ss715585874
		ss715585876
		ss715585877
		ss715585878
		ss715585883
		ss715585889
		ss715585890
		ss715585896
		ss715585930
		ss715585934
		ss715585936
		ss715585937
		ss715585942
3	27	ss715585943
		ss715585944
		ss715585945
		ss715585948
		ss715585953
		ss715586057
		ss715586063
		ss715586070
		ss715586082
		ss715586086
		ss715586087
		ss715586094
		ss715586095
		ss715586102
14	1	ss715618712
16	1	ss715624611
20	1	ss715637438

In GWAS analysis, population structure and relative kinship may cause spurious association between traits and markers (Yu et al. 2006). GLM corrects for population structure while MLM takes both population structure and familiar relatedness into account. Both GLM and MLM control the genomic inflation effectively and have been widely used in GWAS of soybean traits (Wen et al. 2014; Wen et al. 2015; Zhang et al. 2016). In this study, three subpopulations as suggested by STRUCTURE analysis were used to remove the spurious association caused by population structure. MLM was also conducted at the optimum compression level since compressed MLM has been demonstrated to be more powerful and effective in association studies (Zhang et al. 2010).

A major soybean QTL conferring salt tolerance has been consistently mapped on Chr. 3 (Lee et al. 2004; Hamwieh et al. 2011; Huang 2013; Guan et al. 2014; Concibido et al. 2015; Patil et al. 2016). In our study, the major QTL on Chr. 3 was confirmed and narrowed down to a region of 1.18 Mb, which was flanked by SNP markers ss715585943 and ss715586102. Moreover, the salt tolerance gene Glyma03g32900 (40,623,066–40,634,451) reported by Guan et al. (2014) and Patil et al. (2016) was also located near the SNP marker ss715586154 (40,440,832). A total of 18 out of 31 significant SNPs for both leaf chloride concentrations and leaf chlorophyll concentrations on Chr. 3 can be considered as major SNPs with explanation of phenotypic variation greater than 10%. SNPs significantly associated with both leaf chloride concentrations and leaf chlorophyll concentrations over years were also detected on Chr. 2, 7, 8, 10, 13, 14, 16, and 20 in GLM. Huang (2013) reported that three SNPs on Chr. 2 and two SNPs on Chr. 8 were significantly associated with leaf scorch score under salt stress. The significant SNP on Chr. 2 identified in our study is around 1.4 Mb away from those detected by Huang (2013). However, the significant SNP on Chr. 8 in our study is about 18 Mb away from those reported by Huang (2013). The SNP ss715581136 on Chr. 2 and ss715618138 and ss715618731 on Chr. 14 can be also considered as major SNPs since they explained greater than 10% of phenotypic variation for both leaf chloride concentrations and leaf chlorophyll concentrations (Table 1). The major alleles for 80% of significant SNPs in our population contributed to the increase of leaf chloride concentration in soybean under 120-mM NaCl treatment, which was in agreement with that soybeans were generally sensitive to salt stress (Launchli 1984). Minor alleles of most of the significant SNPs on Chr. 3 contributed to the decrease of leaf chloride concentrations under salt stress, which was in agreement with previously reported results (Huang 2013). However, minor alleles of three significant SNPs on Chr. 3, three SNPs on Chr. 7, and one SNP on Chr. 13 accounted for the increase of the leaf chloride concentrations under

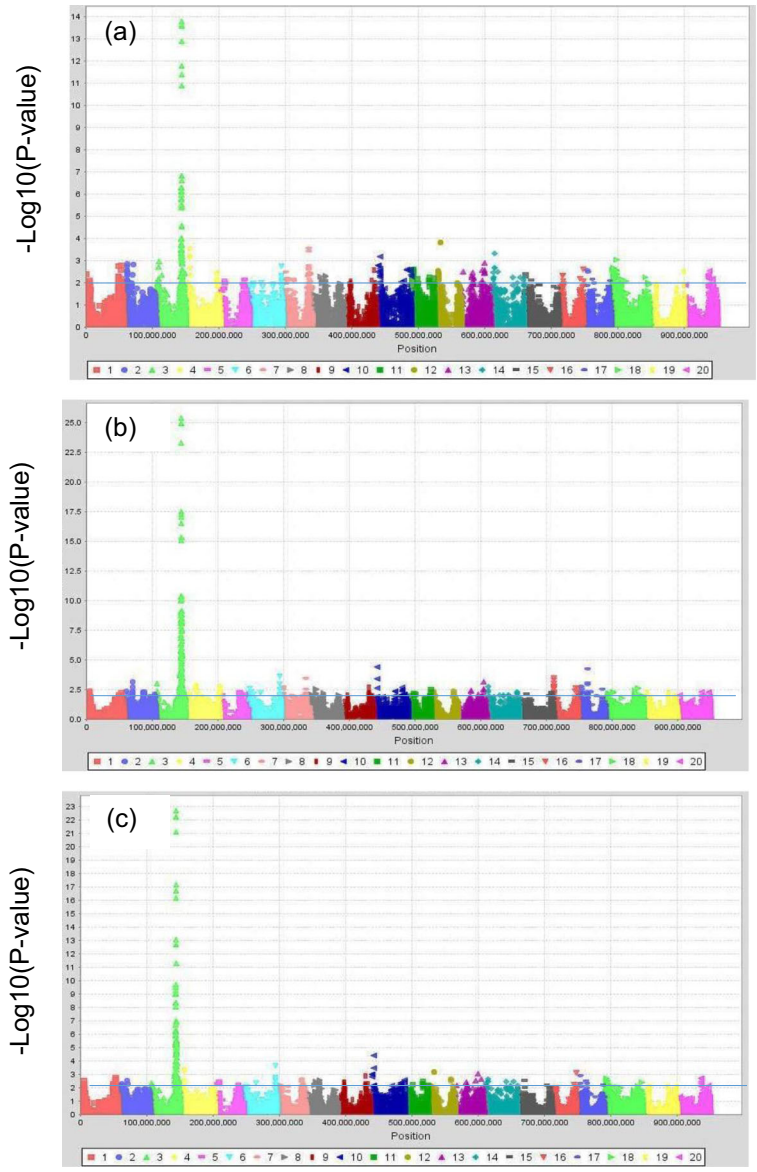

Fig. 2 Manhattan plot of −log10(*P*) values for each SNP against chromosomal position for leaf chloride concentrations in 2014 (**a**), 2015 (**b**), combined chloride over 2014 and 2015 (**c**)

salt stress in our study (Table 1). Overall, the significant SNP markers on Chr. 2, 3, 14, 16, and 20, which were identified in both GLM and MLM (Table 2), are highly recommended for marker-assisted selection in breeding salt-tolerant soybean lines. Moreover, we found out that the SNPs and QTLs for salt tolerance at germination stage (Kan et al. 2015) were probably different from those identified at vegetative stages (V1 and V2) of soybean, because none of the SNPs or QTLs for salt tolerance at germination stage (Kan et al. 2015) was confirmed in our study or previously reported studies (Lee et al. 2004; Huang 2013; Concibido et al. 2015; Patil et al. 2016). Although one salt tolerance SNP at germination stage has been also identified on Chr. 3 (Kan et al. 2015), the physical position of this SNP was more than 35 Mb away from the previously reported major salt tolerance QTL at vegetative stages (Lee et al. 2004; Huang 2013; Concibido et al. 2015). In addition, the salt tolerance QTL on Chr. 13 indicated by current GWAS analysis was different from the QTL on Chr.13 indicated by bi-parental QTL mapping analysis (Zeng 2016); the QTL on Chr.15 identified in the

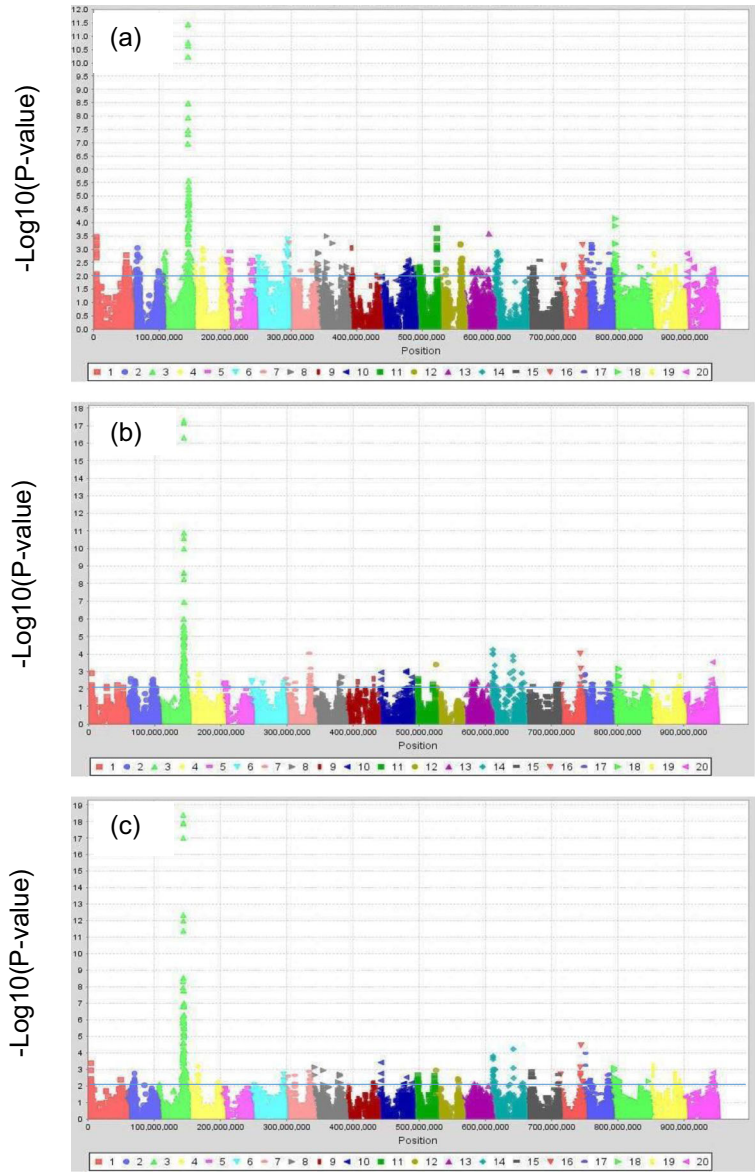

Fig. 3 Manhattan plot of −log10(*P*) values for each SNP against chromosomal position for leaf chlorophyll concentrations (SPAD values) in 2014 (**a**), 2015 (**b**), and combined chlorophyll over 2014 and 2015 (**c**)

bi-parental QTL mapping analysis (Zeng 2016) was not confirmed by current GWAS analysis.

In summary, a genome-wide association analysis was conducted using high-density SNP markers and two indicators of salt tolerance trait in a mapping population consisting of diverse germplasm lines worldwide. With the implementation of GLM and MLM, the major salt tolerance QTL for vegetative stage was confirmed on Chr. 3; a minor salt tolerance QTL for vegetative stage was confirmed on Chr. 2; and seven novel salt tolerance QTLs for vegetative stage were identified on Chr. 7, 8, 10, 13, 14, 16, and 20. The newly identified significant SNP markers for salt tolerance at vegetative stage will benefit the breeders in developing salt-tolerant varieties by assisting in parent line selection, trait introgression, and evaluation of germplasm.

Acknowledgements This work was supported by Arkansas Soybean Promotion Board and United Soybean Board. I would also like to thank the team members in soybean breeding and genetics program at the University of Arkansas for helping me with the greenhouse experiments.

References

Arshad MA, Lowery B, Grossman B (1996) Physical tests for monitoring soil quality. Methods for assessing soil quality, methodsforasses, pp.123–141

Beecher HG (1994) Effects of saline irrigation water on soybean yield and soil salinity in the Murrumbidgee Valley. Aust J Exp Agr 34:85–91

Blumwald E, Grover A (2006) Salt tolerance. In: Halford NG (ed) Current and future uses of genetically modified crops. Wiley, UK, pp 206–224

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

Brim CA (1966) Dare Soybeans1 (Reg. No. 50). Crop Sci 6:95

Chen P, Sneller CH, Mozzoni LA, Rupe JC (2007) Registration of 'Osage' soybean. J Plant Regist 1:89–92

Concibido VC, Husic I, Lafaver N, Vallee BL, Narvel J, Yates J and Ye XH (2015) Molecular markers associated with chloride tolerant soybeans. U.S. Patent Application 14/424, 454, filed august 28, 2013

Endelman JB, Jannink JL (2012) Shrinkage estimation of the realized relationship matrix. G3 2:1405–1413

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

Grant D, Nelson RT, Cannon SB Shoemaker RC (2010) SoyBase, the USDA-ARS soybean genetics and genomics database. Nucl Acids Res D843–846

Greene NV, Kenworthy KE, Quesenberry KH, Unruh JB, Sartain JB (2008) Variation and heritability estimates of common carpetgrass. Crop Sci 48:2017–2025

Guan R, Qu Y, Guo Y, Yu L, Liu Y, Jiang J, Chen J, Ren Y, Liu G, Tian L, Jin L (2014) Salinity tolerance in soybean is modulated by natural variation in GmSALT3. Plant J 80:937–950

Hamwieh A, Xu DH (2008) Conserved salt tolerance quantitative trait locus (QTL) in wild and cultivated soybeans. Breed Sci 58:355–359

Hamwieh A, Tuyen DD, Cong H, Benitez ER, Takahashi R, Xu DH (2011) Identification and validation of a major QTL for salt tolerance in soybean. Euphytica 179:451–459

Hao D, Chao M, Yin Z, Yu D (2012) Genome-wide association analysis detecting significant single nucleotide polymorphisms for chlorophyll and chlorophyll fluorescence parameters in soybean (*Glycine max*) landraces. Euphytica 186: 919–931

Huang L (2013) Genome-wide association mapping identifies QTLS and candidate genes for salt tolerance in soybean. Thesis (M.S.). University of Arkansas, Fayetteville

Huang X, Wei X, Sang T, Zhao Q, Feng Q, Zhao Y, Li C, Zhu C, Lu T, Zhang Z, Li M (2010) Genome-wide association studies of 14 agronomic traits in rice landraces. Nat Genet 42: 961–967

Hwang EY, Song Q, Jia G, Specht JE, Hyten DL, Costa J, Cregan PB (2014) A genome-wide association study of seed protein and oil content in soybean. BMC Genomics 15:1

Kan G, Zhang W, Yang W, Ma D, Zhang D, Hao D, Hu Z, Yu D (2015) Association mapping of soybean seed germination under salt stress. Mol Gen Genomics 290:2147–2162

Katerji N, Hoorn JW, Hamdy A, Mastrorilli M (2003) Salinity effect on crop development and yield, analysis of salt tolerance according to several classification methods. Agr. Water Manage 62:37–66

Launchli A (1984) Salt exclusion: an adaptation of legumes for crops and pastures under saline conditions. In: Staples RC, Toeniessen GH (eds) Salinity tolerance in plants. Strategies for crop improvement. Wiley, New York, pp 171–187

Ledesma F, Lopez C, Ortiz D, Chen P, Korth KL, Ishibashi T, Zeng A, Orazaly M, Florez-Palacios L (2016) A simple greenhouse method for screening salt tolerance in soybean. Crop Sci 56:585–594

Lee GJ, Boerma HR, Villagarcia MR, Zhou X, Carter TE, Li Z, Gibba MO (2004) A major QTL conditioning salt tolerance in S-100 soybean and descendent cultivars. Theor Appl Genet 109:1610–1619

Lenis JM, Ellersieck M, Blevins DG, Sleper DA, Nguyen HT, Dunn D, Lee JD, Shannon JG (2011) Differences in ion accumulation and salt tolerance among glycine accessions. J Agron Crop Sci 197:302–310

Mamidi S, Chikara S, Goos RJ, Hyten DL, Moghaddam SM, Cregan PB, McClean PE (2011) Genome-wide association analysis identifies candidate genes associated with iron deficiency chlorosis in soybean. Plant Genome 4:154–164

Marroni F, Pinosio S, Zaina G, Fogolari F, Felice N, Cattonaro F, Morgante M (2011) Nucleotide diversity and linkage disequilibrium in *Populus nigra* cinnamyl alcohol dehydrogenase (CAD4) gene. Tree Genet Genomes 7:1011–1023

Patil G, Do T, Vuong TD, Valliyodan B, Lee JD, Chaudhary J, Shannon JG, Nguyen HT (2016) Genomic-assisted haplotype analysis and the development of high-throughput SNP markers for salinity tolerance in soybean. Sci Rep 6:19199

Pritchard JK, Donnelly P (2001) Case-control studies of association in structured or admixed populations. Theor Popul Biol 60:227–237

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

Qi X, Li MW, Xie M, Liu X, Ni M, Shao G, Song C, Yim AKY, Tao Y, Wong FL, Isobe S (2014) Identification of a novel salt tolerance gene in wild soybean by whole-genome sequencing. Nat Commun 5

SAS Institute (2004) The SAS system for windows. Release 9.4. SAS Institute, Cary

SAS Institute Inc. (1989-2007) JMP® Version 9.0., Cary, NC

Shao GH, Song JZ, Liu HL (1986) Preliminary studies on the evaluation of salt tolerance in soybean varieties. Acta Agron Sin 6:30–35

Song Q, Hyten DL, Jia G, Quigley CV, Fickus EW, Nelson RL, Cregan PB (2013) Development and evaluation of SoySNP50K, a high-density genotyping array for soybean. PLoS One 8:e54985

Valencia R, Chen P, Ishibashi T, Conatser M (2008) A rapid and effective method for screening salt tolerance in soybean. Crop Sci 48:1773–1779

Wen Z, Tan R, Yuan J, Bales C, Du W, Zhang S, Chilvers MI, Schmidt C, Song Q, Cregan PB, Wang D (2014) Genome-wide association mapping of quantitative resistance to sudden death syndrome in soybean. BMC Genomics 15:1

Wen ZX, Boyse JF, Song Q, Cregan PB, Wang D (2015) Genomic consequences of selection and genome-wide association mapping in soybean. BMC Genomics 16:671

Wheal MS, Lyndon TP (2010) Chloride analysis of botanical samples by ICP-OES. J Anal At Spectrom 25:1946–1952

Wimmer V, Albrecht T, Auinger HJ, Schon CC (2012) Synbreed: a framework for the analysis of genomic prediction data using R. Bioinformatics 28:2086–2087

Yu J, Pressoir G, Briggs WH, Bi IV, Yamasaki M, Doebley JF, McMullen MD, Gaut BS, Nielsen DM, Holland JB, Kresovich S (2006) A unified mixed-model method for association mapping that accounts for multiple levels of relatedness. Nat Genet 38:203–208

Zeng A (2016) Quantitative trait loci mapping, genome-wide association analysis, and gene expression of salt tolerance in soybean. Dissertation (Ph.D.), University of Arkansas

Zhang Z, Ersoz E, Lai CQ, 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

Zhang J, Song Q, Cregan PB, Nelson RL, Wang X, Wu J, Jiang G (2015) Genome-wide association study for flowering time, maturity dates and plant height in early maturing soybean (*Glycine max*) germplasm. BMC Genomics 16:217

Zhang J, Song Q, Cregan PB, Jiang GL (2016) Genome-wide association study, genomic prediction and marker-assisted selection for seed weight in soybean (*Glycine max*). Theor Appl Genet 129:117–130

Zhu C, Core M, Buckler E, Yu J (2008) Status and prospects of association mapping in plants. Plant Genome 1:5–20