



# Association Mapping of Genomic Loci Controlling Salt Tolerance in a Diverse Panel of Soybean Germplasm Accessions

Tuyen Do<sup>1</sup>, Tri D. Vuong<sup>1</sup>, Scotty Smothers<sup>2</sup>, Gunvant Patil<sup>1</sup>, James G. Shannon<sup>2</sup>, Henry T. Nguyen<sup>1</sup>

<sup>1</sup>Division of Plant Sciences and National Center for Soybean Biotechnology (NCSB), University of Missouri, Columbia, MO 65211, USA

<sup>2</sup>Division of Plant Sciences and NCSB, University of Missouri, Portageville, MO 63873, USA.

## ABSTRACT

Salinity is one of the major abiotic stresses that adversely affects productivity and quality in several crops including soybean [*Glycine max* (L.) Merr.]. Conventional genetic studies have commonly utilized bi-parental mapping populations to identify and characterize genomic region(s) associated with tolerance to salinity. Even though this strategy has successfully mapped major loci on different chromosomes (Chr.), it has limitations to capture allelic diversity and genomic resolution in parental lines. In this study, a genome-wide association study (GWAS) was performed to detect and map genomic regions for salt tolerance in a diverse panel of over 400 exotic soybean plant introductions (PIs), consisting of both cultivated (*Glycine max*) and wild soybean (*Glycine soja*). Over 37,000 single nucleotide polymorphism (SNP) markers with minor allele frequency (MAF)  $\leq 0.5$ , generated by the SoySNP50K iSelect BeadChip, were utilized for analysis. The GWAS identified a major locus on Chr. 3 ( $p$ -value =  $8.08E-34$ ), which was consistently mapped at the same genomic position as the known locus of previous studies. A second minor locus was identified and mapped on Chr. 5 ( $p$ -value =  $5.39E-06$ ). It was speculated that it could be a novel locus for salinity tolerance in this panel of exotic germplasm. Several candidate genes along with significantly associated SNPs will be helpful to reveal the molecular mechanism involved in salt tolerance and also useful for marker-assisted selection in soybean molecular breeding programs.

## MATERIALS

### Plant materials

A diverse panel of 350 cultivated soybean accessions (*Glycine max*) and 83 wild soybean accessions (*Glycine soja*) were selected from the United Soybean Board whole genome re-sequencing project (Phase I).

### Genotypic data

Over 45,000 SNP markers of the soybean genome generated in the SoySNP50K iSelect BeadChip were accessed from the soybean database (<http://www.soybase.org>). Of these, 35,270 SNPs were selected for GWAS after excluding SNPs with more than 20% missing data and a minor allele frequency less than 5%.

## METHODS

### Phenotyping for salt tolerance

Soybean lines of the panel and the checks were evaluated for salt tolerance under greenhouse conditions following a previously described method (Lee *et al.*, 2008). Briefly, soybean seedlings grown in cone-trainers to V2 growth stage were exposed to salt water (120 mM NaCl) (Fig. 1). Individual soybean plants were visually rated for salt sensitivity or tolerance when the sensitive parent, cultivar Hutcheson, showed severe leaf scorch (about 2 weeks after treatment with salt water). The greenhouse assays were replicated twice.



Figure 1: Screening salt tolerance of soybean lines in a greenhouse

Leaf scorch score (LSS) was recorded using a 1-5 scale (Fig. 2), where 1 = no apparent chlorosis; 2 = slight (25% of the leaves showed chlorosis); 3 = moderate (50% of the leaves showed chlorosis and some necrosis); 4 = severe chlorosis (75% of the leaves showed chlorosis and severe necrosis); and 5 = dead (leaves showed severe necrosis and were withered).



Figure 2: A 1-5 scale was utilized to assess leaf scorch

### Genome-wide association study (GWAS)

All GWAS analyses were performed using SNP & Variation Suite (SVS 8 - (Golden Helix) and the Genomic Association and Prediction Integrated Tool (GAPIT) (Lipka *et al.*, 2012).

## RESULTS AND DISCUSSION

### Frequency distribution of leaf scorching score (LSS) in soybean germplasm accessions

The distribution of LSS of soybean germplasm accessions (*G. max* and *G. soja*) showed that most soybean accessions evaluated were sensitive to salt (Fig. 3).

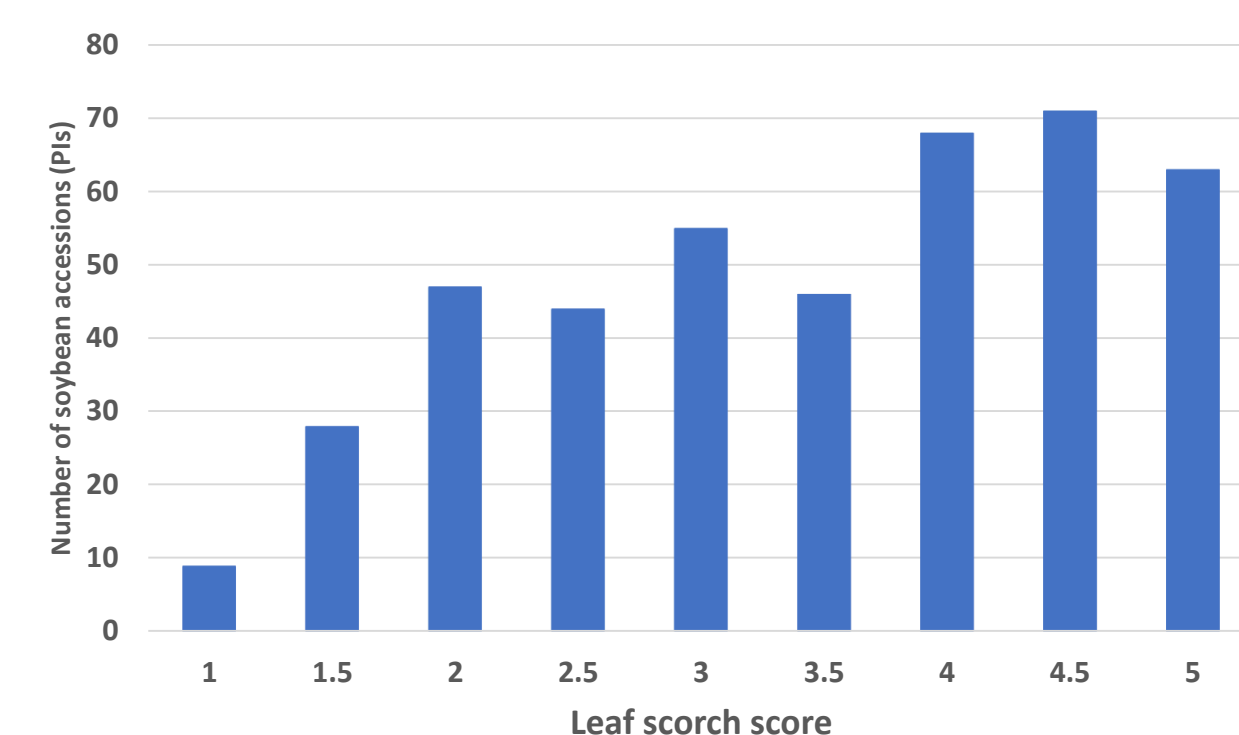


Figure 3: Frequency distribution of LSS in soybean germplasm accessions

### Phylogeny of soybean accessions

A neighbor-joining (NJ) tree for a diverse set of over 400 soybean accessions was constructed based on Nei's genetic distance with 35,270 SNPs and a graphical visualization of phylogenetic tree was made (Fig. 4). The resulting NJ tree showed *G. max* accessions and *G. soja* accessions were clustered into different groups and diversity of subgroups.

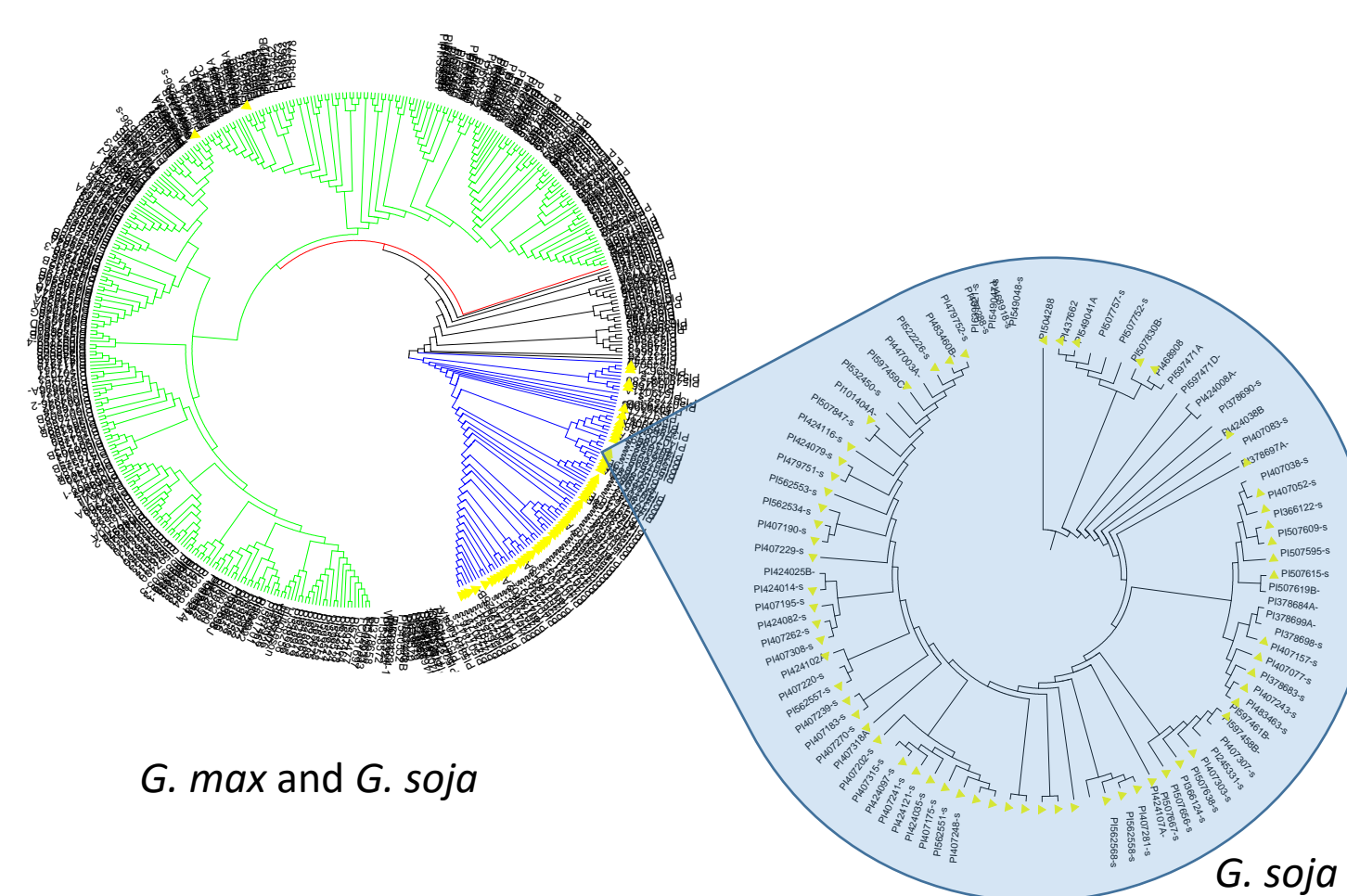


Figure 4: Phylogenetic tree of soybean germplasm accessions was constructed using the SoySNP50K iSelect BeadChip data set

### Population structure and kinship of soybean accessions

Principal component analysis (PCA) showed dispersed genotypes among different components suggesting very diverse genetic backgrounds (Fig. 5A). Kinship analysis showed a clustering pattern of the PIs similar to the NJ tree. The PCA and Kinship information were further utilized for the association analysis (Fig. 5B).

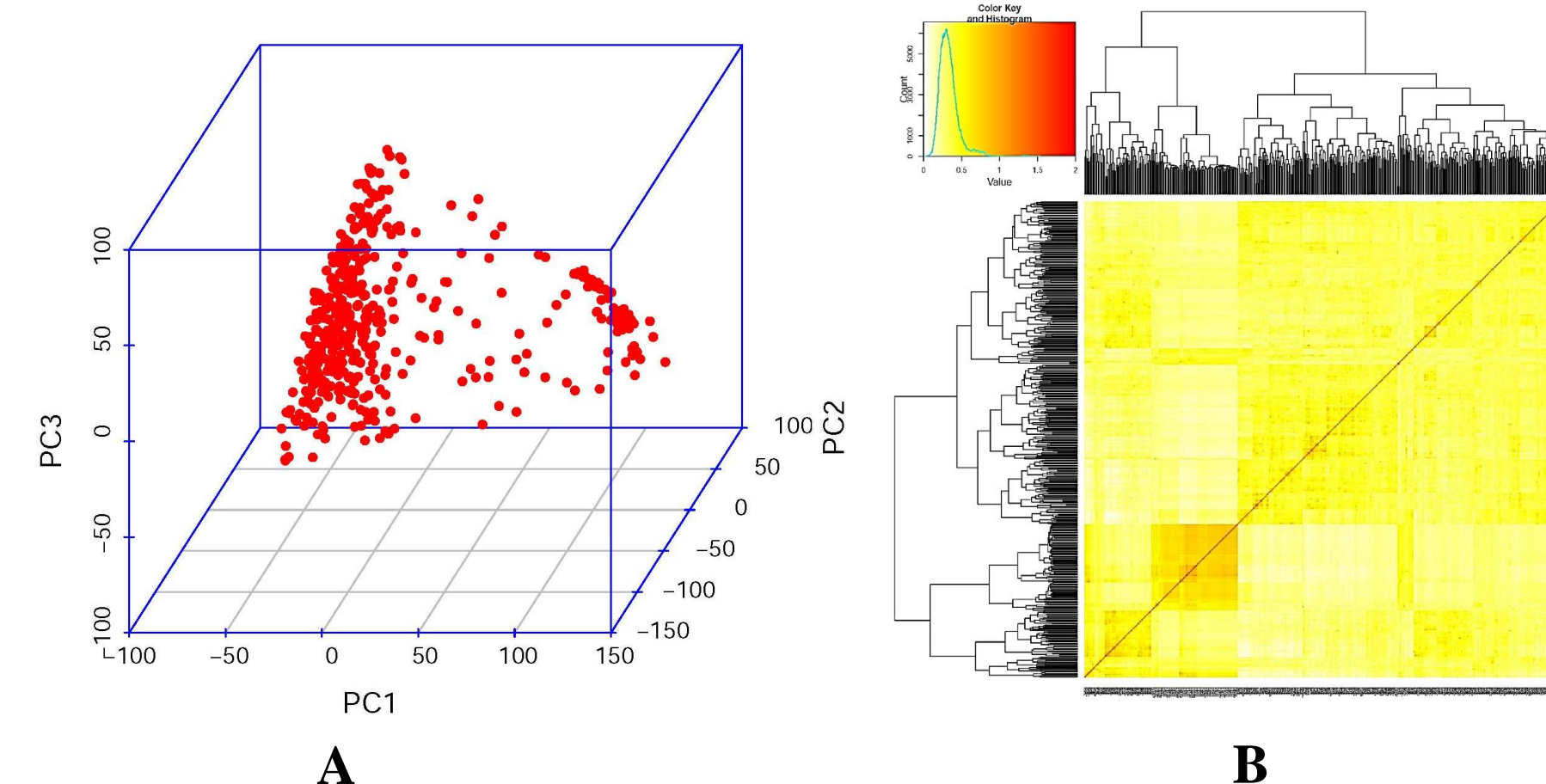


Figure 5: Principal component analysis (A) and kinship matrix (B) of soybean germplasm accessions using the SoySNP50K iSelect BeadChip data set

### Genome-wide association study (GWAS)

The association of SNPs distributed throughout 20 soybean chromosomes and salt tolerance is presented in a Manhattan plot (Fig. 6). A total of 81 significant SNPs were identified for salt tolerance. Of these, five selected SNPs were shown in Table 1. Most of significant SNPs were located from 37779807 to 39594385 on Chr. 3 (Fig. 7). R-square values of Model tested with SNPs range from 0.21 to 0.51 with a highest  $p$  value of  $6.27E-05$ . A second significant SNP locus was identified on Chr. 5 with a  $p$  value of  $5.39E-06$ . Results of the present study are in agreement with previous studies, which determined a major genomic locus for salt tolerance on Chr. 3 (Lee *et al.*, 2004; Hamwieh and Xu, 2008; Hamwieh *et al.* 2011; Ha *et al.*, 2013; Qi *et al.*, 2014).

Table 1: Five of 81 most significant SNPs were identified on Chrs. 3 and 5 for salt tolerance

No.	SNP	Chr.	Position	P value	MAF	nobs	R <sup>2</sup> of Model without SNP	R <sup>2</sup> of Model with SNP
1	ss715585943	3	38579634	8.08E-34	0.356	433	0.179	0.507
2	ss715585952	3	38644367	7.30E-15	0.498	433	0.179	0.301
3	Ss715585950	3	38605485	1.55E-07	0.169	433	0.179	0.232
4	Ss715586085	3	39594385	6.27E-05	0.136	433	0.179	0.209
5	Ss715590236	5	25357327	5.39E-06	0.143	433	0.179	0.218

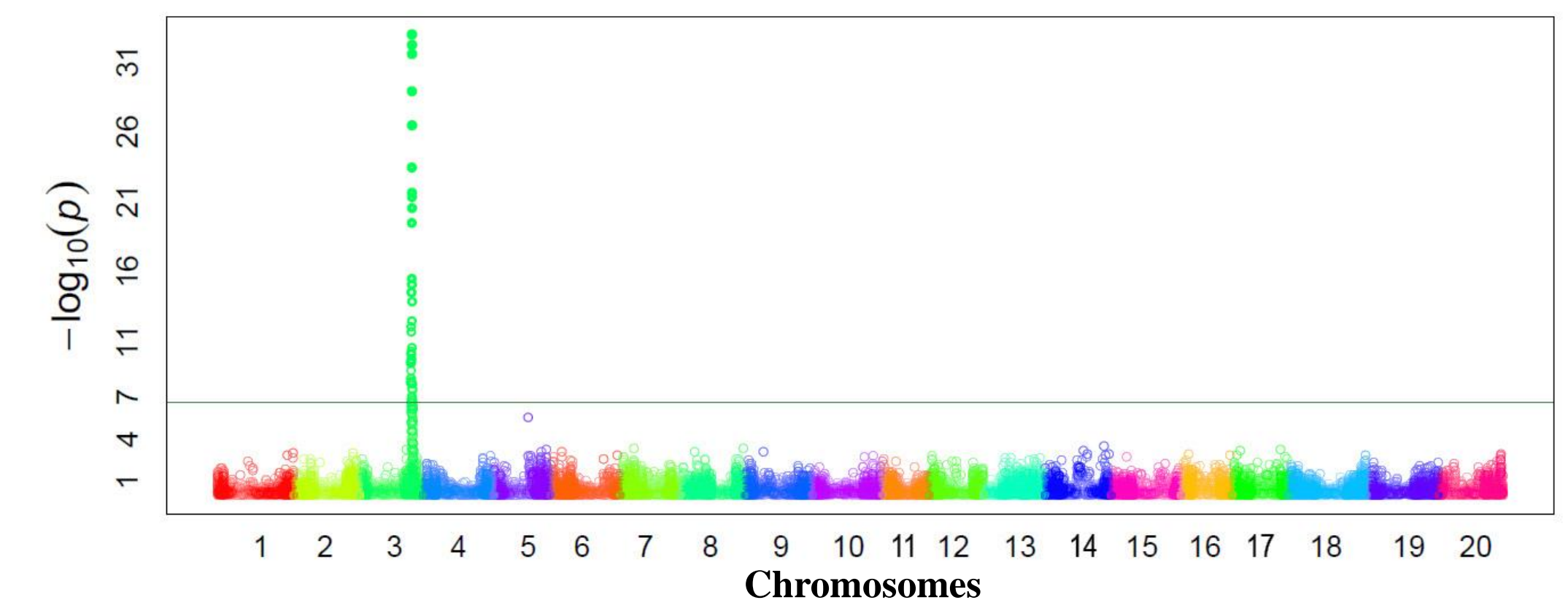


Figure 6: Association of SNPs distributed throughout 20 soybean chromosomes and salt tolerance in soybean by GWAS

The association map of significant SNPs associated with salt tolerance and distribution of gene models on Chr. 3 (Fig. 7).

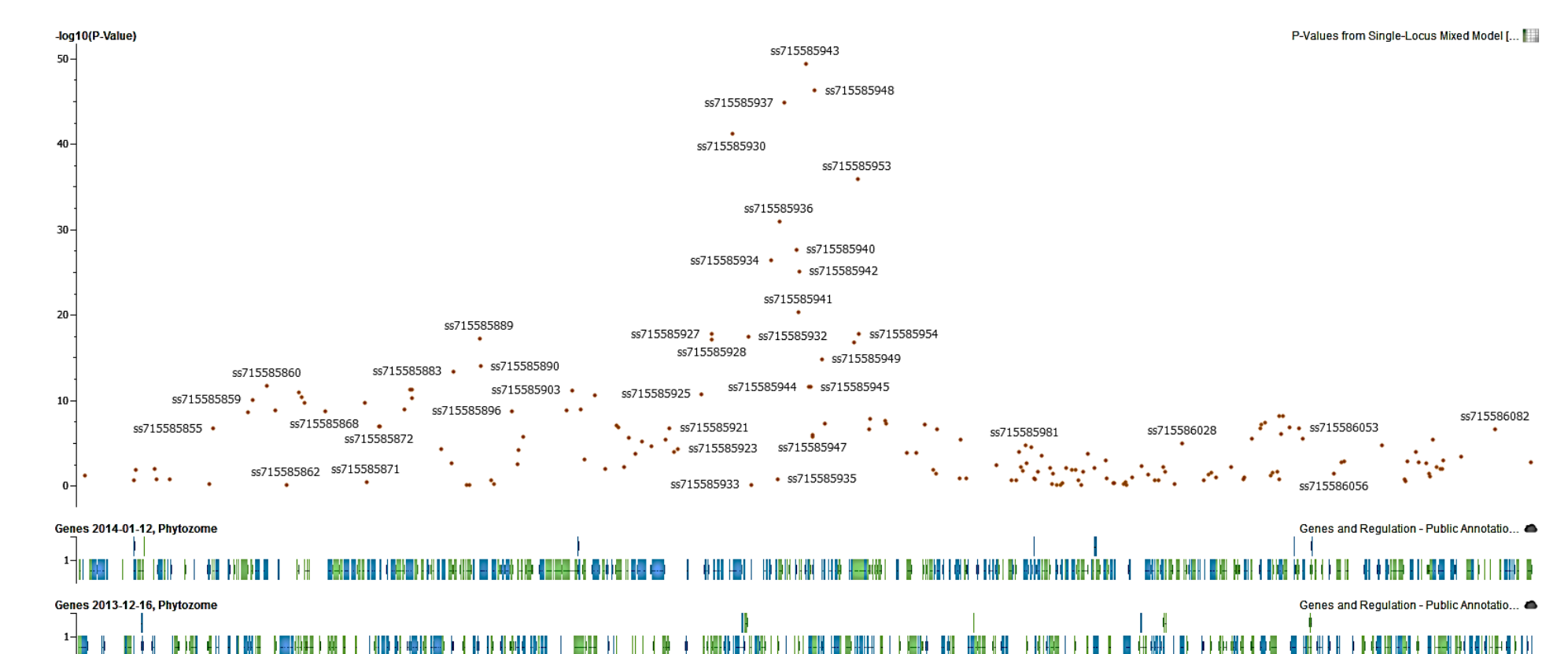


Figure 7: Distribution of significant SNPs associated with salt tolerance on Chr. 3 visualized association analysis

The major gene controlling salt tolerance in soybean, *Glyma03g32900* (*GmCHX1*) was identified on Chr. 3 (Qi *et al.*, 2014; Guan *et al.*, 2014). In this study, two significant SNPs flanking this gene were ss715585952 and ss715585950 (Fig. 8).

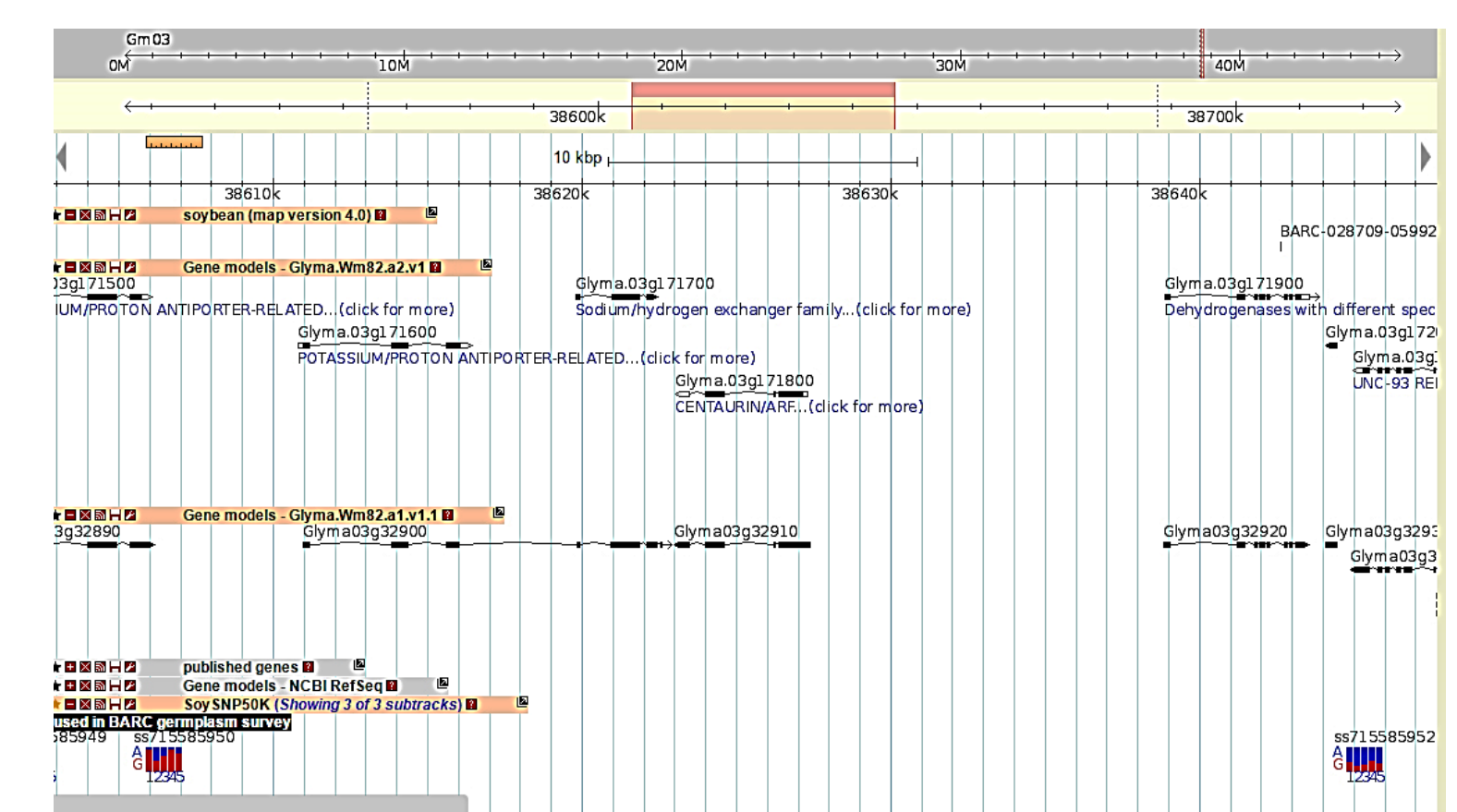


Figure 8: Flanking SNP markers of gene controlled salt tolerance

## CONCLUSIONS

- Among soybean accessions evaluated in this study, five *G. max* accessions and four *G. soja* accessions were identified to be new salt tolerant sources, which can be valuable for salt tolerance improvement in soybean.
- The GWAS identified a major locus significantly associated with salt tolerance located on Chr. 3 ( $p$ -value =  $8.08E-34$ ). The finding in our study was in agreement with many previous studies.
- A second minor locus was identified and mapped on Chr. 5 ( $p$ -value =  $5.39E-06$ ); however, additional investigation needs to be conducted to confirm this locus.

## FUTURE WORK

- The most significant SNP for salt tolerance was mapped on chromosome 3 but relatively distant from the gene identified in previous studies. Thus, these new sources may have different alleles for tolerance which needs to be confirmed.
- A second locus on Chr. 5 showed a significant but low R-square value for salt tolerance and the effects of this allele for tolerance needs further study.
- Thus, additional phenotyping of this diverse set of soybean accessions and association analysis conducted with sequence data from the USB whole genome sequencing project is necessary.



## ACKNOWLEDGMENTS



This research was supported in part by the United Soybean Board, the Missouri Soybean Merchandising Council, and the National Center for Soybean Biotechnology, University of Missouri. The authors would like to thank Drs. Perry B. Cregan and Qijian Song, Soybean Genomics and Improvement Lab, USDA-ARS, Beltsville, MD 20705, for the development of SoySNP50K iSelect BeadChip used in this study.