

Jun Qin^{1, 5}, Qijian Song², Ainong Shi³, Song Li⁴, Mengchen Zhang⁵, Bo Zhang^{4*}

¹ Department of Crop and Soil Environmental Sciences, Virginia Tech, Blacksburg, VA 24061, USA, ² National Soybean Improvement Center Shijiazhuang Sub-Center, North China Key Laboratory of Biology and Genetic Improvement of Soybean Ministry of Agriculture, Cereal & Oil Crop Institute, Hebei Academy of Agricultural and Forestry Sciences, Shijiazhuang, Hebei 050031, P.R. China, ³ USDA, Agricultural Research Service, Soybean Genomics and Improvement Lab, Beltsville, MD 20705, USA, ⁴ Department of Horticulture, University of Arkansas, Fayetteville, AR 72701, USA

In Press in PLoS ONE

ABSTRACT

Phytophthora sojae, an oomycete pathogen of soybean, causes stem and root rot, resulting in annual economic loss up to \$2 billion worldwide. Varieties with *P. sojae* resistance are environmental friendly to effectively reduce disease damages. In order to improve the resistance of *P. sojae* and broaden the genetic diversity in Southern soybean cultivars and germplasm in the U.S., we established a *P. sojae* resistance gene pool that has high genetic diversity, and explored genomic regions underlying the host resistance to *P. sojae* races 1, 3, 7, 17 and 25. A soybean germplasm panel from maturity groups (MGs) IV and V including 189 accessions originated from 10 countries were used in this study. The panel had a high genetic diversity compared to the 6,749 accessions from MGs IV and V in USDA Soybean Germplasm Collection. Based on disease evaluation dataset of these accessions inoculated with *P. sojae* races 1, 3, 7, 17 and 25, which are publically available, five accessions in this panel were resistant to all races. Genome-wide association analysis identified a total of 32 significant SNPs, which were clustered in resistance-associated genomic regions, among those, ss715619920 was only 3 kb away from the gene Glyma.14g087500, a subtilisin protease. Gene expression analysis showed that the gene was down-regulated more than 4 fold (log2 fold > 2.2) in response to *P. sojae* infection. The identified molecular markers and genomic regions that are associated with the disease resistance in this gene pool will greatly assist the U.S. Southern soybean breeders in developing elite varieties with broad genetic background and *P. sojae* resistance.

BACKGROUND

- Phytophthora root rot, caused by a fungus *Phytophthora sojae* (*P. sojae*), is a major disease of soybean (*Glycine max*), especially in the areas where soybeans have been cultivated for many years. Yield loss can be substantial, and even entire fields may be destroyed.
- P. sojae* follows a gene-for-gene interaction and is hypothesized to have coevolved with its host plant soybean. At least 13 resistance *Rps* genes are known in soybean, and different races of *P. sojae* are able to overcome these resistance genes, both singly and in combination. Linkage mapping and association mapping have been both conducted in *P. sojae* research to locate *Rps* genes and quantitative trait loci (QTLs).
- MGs IV and V soybeans are adapted to and widely grown in the Southern states such as Virginia, Arkansas and Tennessee. However, most *Rps* genes were identified in soybeans from MGs 000 to IV [1], and not many *P. sojae* resistant donors are evaluated for the adaptation to the South, which made it difficult to implement *P. sojae* resistant genes into the Southern soybean germplasm.

OBJECTIVES

- Establish a *P. sojae* resistance gene pool that is high in genetic diversity to be employed in Southern soybean breeding programs
- Explore *P. sojae* resistance gene regions underlying the host resistance to *P. sojae* races 1, 3, 7, 17 and 25 for molecular breeding selection.

MATERIALS

- Highly diverse 189 soybean accessions from MGs IV and V based on 52,041 SNPs data [2], originally collected from 10 countries (Table 1).

Table 1. The accession numbers from each country.

Country	Number	Country	Number
China	141	Uganda	2
Japan	16	Nepal	1
South Korea	15	Morocco	1
United States	8	South Africa	1
India	2	Vietnam	1

METHODS

- The responses of 189 soybean accessions to races 1, 3, 7, 17 and 25 of *P. sojae* were evaluated by USDA-ARS germplasm curation program, and the protocol and data are available at www.ars-grin.gov.
- The STRUCTURE and MEGA softwares were used to analyze population structure and genetic diversity.
- Association mapping of *P. sojae* resistance to each of the *P. sojae* races 1, 3, 7, 17, and 25 was conducted separately based on three different models in TASSEL: SMR, GLM-Q, and MLM-Q+K.
- Published gene expression data generated by RNA sequencing were downloaded from GEO database (GSE48524). Gene expression analysis was performed according to the method in the original publication [3].

RESULTS AND DISCUSSIONS

Phenotypic variation

- The ratio of resistant accessions to susceptible accessions: race 1, 81/108 (0.75), race 3, 48/71 (0.68), race 7, 27/63 (0.43), race 17, 25/56 (0.45), and race 25, 21/65 (0.32).
- Five accessions showed resistance to all five *P. sojae* races: PI 567780B, PI 567781, PI 587612A, and PI 588021A from China, and PI 378682C from Japan.
- Ten accessions resistant to four *P. Sojae* races, 11 resistant to three races, 32 resistant to two races, and 39 resistant to only one race.
- Huang et al. (2016) also observed a total of 168 (75%) soybean accessions showed resistance to more than one *P. sojae* race and suggested that abundant resistant resources existed [4].

Population structure

- The 189 accessions were divided into two sub-populations (clusters), Q1 (red) and Q2 (green) based on STRUCTURE analysis as the maximal delta K value was observed when K=2 (Fig. 1A and 1B). The phylogenetic tree was constructed by neighbor-joining (NJ) based on genetic distance obtained from MEGA 6 (Fig. 1C)
- The correlation coefficient of 0.227 between clusters (Q1/Q2) and MGs (IV/V) ($P = 0.0017$) indicated that maturity group may affect the clustering of all accessions.

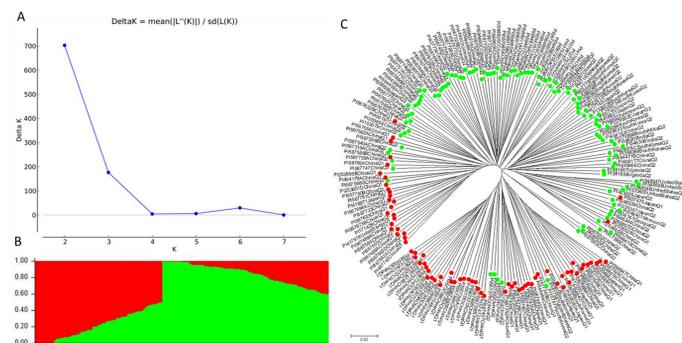


Fig.1 Model-based population structure for the soybean panel: (A) Delta K values for different number of populations assumed (K) in the STRUCTURE analysis, (B) Classification of two populations using STRUCTURE 2.3.4., and (C) phylogenetic tree constructed by neighbor-joining (NJ) based on genetic distance obtained from MEGA 6. The distribution of the accessions to different populations is indicated by the color code (Q1: red and Q2: green), consistent in the figures (B) and (C).

Association analysis and SNP markers identification

- A total of 32 SNPs ($p < 0.001$, $LOD \geq 3$) were identified to be associated with the resistance to at least one of the five tested *P. sojae* races 1, 3, 7, 17, and 25.
- They were grouped into thirteen significant regions covering the genome size from 1.5 to 46 kb (Table 2).

- Genome-wide association analysis of the *P. sojae* resistance in the Chinese soybean mini core collection showed that a total of 14 markers were significantly associated with *Phytophthora* resistance[4], but not all genomic regions of these SNPs were consistent with the regions reported in this study.

Table 2. SNP markers associated with resistance to *P. sojae* using three statistical models, SMR, GLM (Q), and MLM (Q+K).

SNP marker	SMR (LOD)	GLM (LOD)	MLM (LOD)	Trait
ss715585768	5.12	5.82	3.57	chr3-race1&7
ss715614941	4.01	4.35	3.76	chr13-race1
ss715614943	4.18	4.46	4.03	chr13-race1
ss715631615	4.40	4.01	3.43	chr18-race1
Ss715596704	5.02	4.06	3.30	chr7-race3
ss715588855	5.90	4.72	4.02	chr4-race7
ss715588864	5.21	4.04	3.72	chr4-race7
ss715588866	4.13	3.52	3.28	chr4-race7
ss715613726	4.96	9.22	3.12	chr13-race7
ss715617288	4.39	3.31	3.40	chr13-race7
ss715585114	3.67	4.41	3.70	chr3a-race17
ss715586846	6.90	6.02	4.18	chr3b-race17
ss715606847	5.88	5.17	3.24	chr10-race25
ss715606865	5.88	5.17	3.24	chr10-race25
ss715619920	6.73	6.45	5.55	chr14-race25
ss715619926	7.49	7.49	6.05	chr14-race25
ss715631508	5.62	5.03	3.18	chr18-race25

Validation of SNP markers for *P. sojae* resistance

- For each SNP, we searched for the closest gene that was differentially expressed and we found 13 unique genes (Part of them were shown in Table 3).
- Glyma.14g087500, a subtilisin protease, was down-regulated more than 4 fold (log2 fold > 2.2) in response to *P. sojae* infection. This gene was only 3,000 base pairs away from ss715619920. Subtilisin proteases were well known regulators of plant-pathogen responses [5].
- Glyma.13g176600, a MAC/Perforin domain-containing protein, was up-regulated more than 4 fold in response to *P. sojae* infection. This gene was located 24 kb from ss715614943. Perforin domain-containing proteins were well known in their role in immune responses [6].

Table 3. Expression analysis of differentially expressed genes closely-linked to significant SNPs ($P < 0.001$) associated with *P. sojae* races.

SNP marker	SNPs type	Gene Name	distance to SNP (bp)	logFC	logCPM	PValue	Lod	FDR	Gene Annotation
ss715614941	A/G	Glyma.13g176600	30,005	2.66	5.51	7.39E-05	4.13	0.01	MAC/Perforin domain-containing protein
ss715614943	A/C	Glyma.13g176600	24,733	2.66	5.51	7.39E-05	4.13	0.01	MAC/Perforin domain-containing protein
ss715631615	A/G	Glyma.18g261900	194,102	2.88	4.02	5.25E-04	3.28	0.05	Predicted small molecule transporter
ss715596704	C/T	Glyma.07g023300	1,586,568	1.97	6.62	5.66E-04	3.25	0.04	WRKY DNA-binding protein
ss715586336	A/G	Glyma.03g042700	923,940	2.81	5.17	6.19E-05	4.21	0.01	WRKY DNA-binding protein
ss715588829	A/G	Glyma.04g190700	934,146	-2.76	5.64	7.67E-05	4.12	0.01	Cysteine proteinases superfamily protein
ss715588855	A/G	Glyma.04g190700	1,169,785	-2.76	5.64	7.67E-05	4.12	0.01	Cysteine proteinases superfamily protein
ss715606865	C/T	Glyma.10g161500	174,369	6.49	5.18	1.02E-10	9.99	0.00	No annotation
ss715619920	A/G	Glyma.14g087500	3,014	-2.21	6.89	6.04E-05	4.22	0.01	Subtilase family protein
ss715619924	G/T	Glyma.14g087500	17,098	-2.21	6.89	6.04E-05	4.22	0.01	Subtilase family protein
ss715619926	A/G	Glyma.14g087500	25,344	-2.21	6.89	6.04E-05	4.22	0.01	Subtilase family protein

CONCLUSIONS

- We explored the most efficient way to utilize the publically available datasets for the purpose of improving genetic diversity and improving traits.
- The findings will help Southern soybean breeders to improve elite varieties with broad genetic diversity and with *P. sojae* resistance.

REFERENCES

- Dorrance AE, Schmitthenner AF. Plant Dis. 2000;84:1303–1308. doi:10.1094/PDIS.2000.84.12.1303.
- Song Q, Hyten DL, Jia G, Quigley CV, Fickus EW, Nelson RL, et al. G3: Genes|Genomes|Genetics. 2015;5(10):1999–2006.
- Desgroux A, L'Anthoëne V, Roux-Duparque M, Rivière J-P, Aubert G, Tayeh N, et al. BMC genomics. 2016;17(1):1.
- Huang J, Guo N, Li Y, Sun J, Hu G, Zhang H, et al. BMC genetics. 2016;17(1):85.
- Ilgueiredo A, Monteiro F, Sebastiana M. Plant Sci. 2014;5:1-4.
- Rosado CJ, Kondos S, Bull TE, Kuiper MJ, Law RH, Buckle AM, et al. Cell Microbiol. 2008;10(9):1765-74.