Check for updates

Deciphering the genetic architecture of resistance to *Corynespora cassiicola* in soybean (*Glycine max* L.) by integrating genome-wide association mapping and RNA-Seq analysis

Sejal Patel[1]*, Jinesh Patel[1], Kira Bowen[2] and Jenny Koebernick[1]

[1]Department of Crop, Soil and Environmental Sciences, Auburn University, Auburn, AL, United States,
[2]Department of Entomology and Plant Pathology, Auburn University, Auburn, AL, United States

Target spot caused by *Corynespora cassiicola* is a problematic disease in tropical and subtropical soybean (*Glycine max*) growing regions. Although resistant soybean genotypes have been identified, the genetic mechanisms underlying target spot resistance has not yet been studied. To address this knowledge gap, this is the first genome-wide association study (GWAS) conducted using the SoySNP50K array on a panel of 246 soybean accessions, aiming to unravel the genetic architecture of resistance. The results revealed significant associations of 14 and 33 loci with resistance to LIM01 and SSTA *C. cassiicola* isolates, respectively, with six loci demonstrating consistent associations across both isolates. To identify potential candidate genes within GWAS-identified loci, dynamic transcriptome profiling was conducted through RNA-Seq analysis. The analysis involved comparing gene expression patterns between resistant and susceptible genotypes, utilizing leaf tissue collected at different time points after inoculation. Integrating results of GWAS and RNA-Seq analyses identified 238 differentially expressed genes within a 200 kb region encompassing significant quantitative trait loci (QTLs) for disease severity ratings. These genes were involved in defense response to pathogen, innate immune response, chitinase activity, histone H3-K9 methylation, salicylic acid mediated signaling pathway, kinase activity, and biosynthesis of flavonoid, jasmonic acid, phenylpropanoid, and wax. In addition, when combining results from this study with previous GWAS research, 11 colocalized regions associated with disease resistance were identified for biotic and abiotic stress. This finding provides valuable insight into the genetic resources that can be harnessed for future breeding programs aiming to enhance soybean resistance against target spot and other diseases simultaneously.

KEYWORDS

GWAS, target spot, *C. cassiicola*, RNA-Seq, resistance

1 Introduction

Soybean (*Glycine max* L. Merrill) is a principal legume crop that is cultivated in over 50 countries. It provides a valuable source of high-quality oil and protein for human and animal feed. The United States achieved production of approximately 116.3 million metric tons in 2022, representing around 31% of global production (USDA, 2022). However, about 11% of this soybean production is annually constrained by diseases, according to Hartman et al. (2015). The occurrence of target spot, caused by *Corynespora cassiicola*, (Berk & M.A. Curtis) C.T. Wei is a concern for farmers since due to its potential to cause 18 to 32% yield loss (Dixon et al., 2009; Godoy, 2015; Faske, 2016). It was first reported on soybean in 1945 (Olive et al. (1945); it is now considered widespread in the mid-south and the southeast US due to frequent warm and humid conditions that tend to occur during crucial stages of reproductive development.

Disease symptoms initially appear on foliage as lesions with alternating concentric rings of light and dark brown bands encircled by a yellow halo. Severe infections lead to premature senescence of leaves, resulting in yield loss and diminished seed quality (Seaman et al., 1965; Koenning et al., 2006; Godoy, 2015). The foliar pathogen, *C. cassiicola*, belonging to the phylum Ascomycota, has a broad host range, including cotton, tomato, rubber, and cucumber (Blazquez, 1967; Chee, 1988; Barrett et al. 1991; Conner et al., 2013). However, host-specific *C. cassiicola* isolates have also been reported (Dixon et al., 2009; Sumabat et al., 2018).

Current disease control for target spot depends on field management practices and chemical applications. Although fungicides can control target spot, there is a concern over the development of fungicide resistance (Xavier et al., 2013) which may reduce the effectiveness of the fungicide products (Duan et al., 2019; Rondon and Lawrence, 2019). Moreover, fungicidal control of target spot is not always economically or environmentally sustainable. Therefore, development and cultivation of target spot resistant cultivars are needed for economical and enduring strategy for disease management.

Conventional breeding methods for developing resistant soybean cultivars with competitive yield levels involve a lengthy selection process. However, with the availability of abundant molecular markers and affordable genotyping resources, it would be more efficient to identify genomic regions associated with resistance and incorporate them into a marker-assisted breeding program, which would accelerate the selection process and reduce the overall cost of developing new varieties (Neeraja et al., 2007; Hasan et al., 2021).

The soybean cultivar Williams 82 genome was assembled in 2010, spanning 950 Mb distributed in 20 linkage groups (Schmutz et al., 2010). Additionally, reference genomes for two other soybean accessions, 'Zhonghuang 13' (ZH13) and wild soybean 'W05', are available (Shen et al., 2019; Xie et al., 2019). Six cultivated and two wild relatives of soybean were sequenced and analyzed to develop SoySNP50K, a high-throughput SNP assay to facilitate genotyping (Song et al., 2013). This SNP assay has been used to genotype 18,480 *G. max* accessions and 1168 *G. soja* accessions available in the United States Department of Agriculture (USDA) Soybean Germplasm Collection located at the University of Illinois (Song et al., 2015). The SoySNP50K has played a key role in conducting Genome-wide Association Studies (GWAS) and Quantitative Trait Locus (QTL) studies.

A GWAS associates variation across the whole genome with phenotypes and identifies genetic variations of essential traits. A panel of diverse individuals provide a higher mapping resolution than the classic bi-parental QTL mapping method (Korte and Farlow, 2013). The number of recombination events limit QTL mapping, whereas the association panel has accumulated recombination events over several generations, resulting in improved mapping resolution. In soybean, GWAS studies have successfully identified genes associated with resistance to various pathogens including *Sclerotinia sclerotiorum* (Iquira et al., 2015; Wen et al., 2018), *Pythium sylvaticum* (Lin et al., 2020), *Heterodera glycines* Inchinohe (Vuong et al., 2015; Ravelombola et al., 2020), *Fusarium virguliforme* (Bao et al., 2015; Zhang et al., 2015), *Phytophthora sojae* (Li et al., 2016), *Macrophomina phaseolina* (Coser et al., 2017), and *Cercospora sojina* (McDonald et al., 2023).

Target spot is also a concern in other crops such as rubber, cucumber, tomato, and cotton. Transcriptome profiling studies in cucumber and rubber have elucidated genetic mechanisms underlying resistance to *C. cassiicola* (Wang et al., 2018; Roy et al., 2019; Ribeiro et al., 2021). These studies have identified defense-related genes, biosynthesis of plant hormones, transcription factors, Ca^2+ signaling pathways, secondary metabolites, and miRNA involved in the defense response of host plants against *C. cassiicola* infection. In rubber, molecular markers linked to tolerance to *C. cassiicola* infection were identified by screening 104 F_1 clones with 28 SSR markers and performing single marker analysis (Oktavia et al., 2021). Similarly, in cucumber, a combination of SSR and SNP markers were employed to screen a large population of F_2 plants, leading to the fine mapping of a recessive resistance gene, *cca-3*, against *C. cassiicola* (Wen et al., 2015). Further investigation of candidate genes within the mapped region revealed the presence of the *RGA* (Resistance Gene Analog) gene containing a CC-domain, nucleotide-binding domain (NB-ARC), and the Leucine-rich repeat (LRR) domain, which might play a crucial role in hypersensitive responses to attack by the target spot pathogen (Wen et al., 2015). Valuable genomic resources like reference genomes, a diverse germplasm collection, molecular markers, and other genotyping options are available for soybean; however, *C. cassiicola* resistance has not been investigated at the genomic level. Therefore, this study aims to determine genomic regions associated with resistance against *C. cassiicola* and subsequently identify candidate genes located within identified QTL regions.

2 Materials and methods

2.1 Plant material and experiment design

A collection of 246 soybean genotypes from different geographic regions (USA, China, Japan, South Korea, North Korea, Vietnam, Nepal, Pakistan, Taiwan, Australia, and Georgia) were obtained from the National Germplasm Database (GRIN). Two genotypes, 'Council' (PI 587091) and 'Henderson' (PI 665225)

were used as resistant and susceptible checks, respectively (Patel et al., 2022). To identify the genomic regions associated with target spot resistance, greenhouse experiments were conducted using two isolates of *C. cassiicola* namely LIM01 and SSTA (Patel et al., 2022) to pinpoint potential candidate genes. These isolates exhibit variations in their toxin producing cassiicolin gene. LIM01 has *Cas2* while SSTA has *Cas2* and *Cas6*. This approach facilitates the identification of horizontal resistance mechanisms. In brief, infected leaf tissue was collected from soybean fields located in Georgia and Alabama. To ensure the isolation of pure cultures, the collected tissue was rinsed with 70% ethanol, 2% NaOCl, and sterile distilled water. The cleaned leaf tissue was placed on Potato dextrose agar plates supplemented with kanamycin to prevent bacterial growth. The plates were incubated at 28°C to obtain fungal colonies. Single colonies from each isolate were transferred to slant tubes containing V8 juice agar. The inoculated slant tubes were incubated at 28°C for two weeks. In order to preserve the isolates for the long term, the slant tubes were kept in a refrigerator set at 4°C. The PCR method was used to confirm the species and pathovar of isolates by utilizing conserved gene sequences.

Soybean genotypes were arranged in a randomized complete block design (RCBD), with three replications. Plants were grown in polypropylene deepots D40L (Stuewe & Sons, Inc., Tangent, OR) filled with potting mix (PRO-MIX BX, Premier Tech Horticulture, Quakertown, PA). Three seeds per deepot were sown and later thinned to one plant per deepot at five days post-emergence. The experiments with LIM01 and SSTA isolates were conducted in Plant Science Research Center, Auburn University, with a temperature range 24°C to 34°C. Halide bulbs (1000-watt) provided supplemental light to ensure a 14-hour photoperiod. Plants were grown to the V3 toV4 stage before inoculation and were watered as needed.

2.2 Inoculation and phenotyping

Inoculation and disease assessment methods were followed the procedures described by Patel et al. (2022). Each 10-day old *C. cassiicola* culture, grown on V8 agar, was flooded with sterile distilled water and the colony surface was gently rubbed with a glass rod. The resulting conidial suspension was filtered through double layers of sterile cheesecloth and the concentration was adjusted to 50,000 conidia per ml. All leaves on plants were inoculated by applying the conidial suspension to the upper and lower leaf surfaces using a fine mist professional spray bottle. Subsequently, the inoculated plants were transferred to a mist chamber, with mist running two seconds every 10 minutes over period of 72 hours to create high humidity (>90%). Disease was assessed at 14 days post inoculation based on percentage of damaged leaf area per plant. Following the visual diagram, severity rating scale ranged from 0 to 90%, as outlined by Vincelli and Hershman (2011). Genotypes were categorized as resistant if the disease severity was below 25%, moderately susceptible if the severity was between 25 and 50%, and susceptible if severity

exceeded 50%. Descriptive statistics of phenotypic data were performed using PROC GLIMMIX (Generalized linear mixed model) in SAS (version 9.4; SAS Institute Inc., Cary, NC). In this analysis, genotypes, isolates and genotypes×isolates as fixed effect, while replications as random effect.

2.3 Genotyping

The 246 soybean accessions were genotyped using Illumina Infinium SoySNP50K BeadChip (Song et al., 2015). This association panel identified 42,180 SNPs. Individual markers with a minor allele frequency <5% or missing rate >10% were omitted from the further analysis. As a result, a high-quality set of 33,378 SNPs was retained for GWAS analysis.

2.4 Population structure and linkage disequilibrium estimation

The population structure was analyzed through STRUCTURE v.2.3.4 software to determine subpopulations (Pritchard et al., 2000). The analysis used 33,378 SNPs and employed an admixture model with five iterations of 50,000 burn-in and 50,000 Monte Carlo Markov Chain (MCMC) replications for a range of subpopulations from k=1 to k=10. The optimal number of subpopulations was determined by estimating DeltaK using STRUCTURE HARVESTER (Earl and Vonholdt, 2012). Additionally, linkage disequilibrium (LD) between pairs of SNP markers was evaluated using TASSEL 5 (Bradbury et al., 2007) with a sliding window size of 50 SNPs. The average LD decay rate for all the chromosomes was estimated at the distance the squared correlation coefficient (r^2) dropped to half its maximum value in this population (Remington et al., 2001; Huang et al., 2010).

2.5 Genome-wide association study analysis

The mean scores of disease severity ratings and filtered set of SNPs were used to conduct GWAS analysis using GAPIT v3 software. A statistical model FarmCPU (Fixed and random model Circulating Probability Unification) was employed, with PCA (Principle component analysis) and Kinship as covariates (FarmCPU_(PCA+K)) (Liu et al., 2016; Wang and Zhang, 2021). FarmCPU is known to control false positives and false negatives, making a superior model compared to generalized linear model (GLM) and the mixed linear model (MLM), and has been widely used in soybean GWAS studies (Kaler et al., 2018; Chamarthi et al., 2021). The significant threshold for the association between SNPs and target spot severity trait was set at -log10(P) ≥ 3.5. The adjusted P value was calculated by False Discovery Rate (FDR) using the Benjamini-Hochberg method.

2.6 Prediction of candidate genes

To identify potential candidate genes, a search was conducted within a 200 kb genomic region on both sides of the significantly associated SNPs. The selected candidate genes were then annotated using the Glyma.Wm.82.a2 soybean reference genome available in SoyBase (www.soybase.org) (Cheng et al., 2017; Sun et al., 2020).

2.7 Characterization of candidate genes based on RNA-Seq

To identify candidate genes associated with the GWAS-identified loci, we incorporated RNA-Seq data obtained from Patel et al. (2023). Briefly, two resistant genotypes (Bedford and Council) and two susceptible genotypes (Henderson and Pembina) which were inoculated with a *C. cassiicola* isolate at V3-V4 stage and compared to their respective controls (non-inoculated). Total RNA was extracted controls, samples 24 and 48 hours post-infection (hpi) with three biological replications of young leaves using Direct-zol RNA Mini-Prep Kit (Zymo Research, Irvine, CA) according to the manufacturer's instructions. Subsequently, cDNA library construction and 150-bp paired-end sequencing were carried out by Illumina NovaSeq 6000 instrument (Novo gene Bioinformatics Technology Co., Ltd., CA, US). The raw data were collected, and the low-quality reads, adapters, and Poly A sequences were removed by using fastp software v0.23.1 (Chen et al., 2018). The resulting high-quality reads were aligned to the soybean (*Glycine max* (L.) Merr.) Williams 82 reference genome (Schmutz et al., 2010), and transcript quantification was conducted by Salmon ver.0.9.1 (Patro et al., 2017). To identify genes associated with target spot resistance, R-package DESeq2 (Version 1.37.6) (Love et al., 2014) was applied to determine differential expression analysis. Specifically, genes showing significant changes in expression levels between the *C. cassiicola* inoculated samples, and their respective controls were selected. Moreover, the genes with an absolute log2 fold change ≥ 2 (upregulated genes) or ≤ -2 (downregulated genes) and a false discovery rate (FDR) < 0.01 were considered significant DEGs. The gene ontology (GO) enrichment analysis of differentially expressed genes was performed using Shiny Go v 0.76 (Ge et al., 2020) with *Glycine max* set as the background. The enrichment analysis identified function categories of enriched GO terms, which were considered significant when the false discovery rate (FDR)-adjusted P-value was less than or equal to 0.05.

3 Results

3.1 Phenotypic analysis

A total of 246 soybean accessions were evaluated for disease severity to two *C. cassiicola* isolates (LIM01 and SSTA). The phenotypic data obtained from both isolates exhibited a near-normal distribution (Figure 1). For isolate LIM01, disease severity scores ranged from 0 to 80%, with an overall mean of 34.2% (Figure 1A). Isolate SSTA had disease severity scores ranging from 3.3 to 76.6%, with an overall mean of 32.7% (Figure 1B). The disease severity ratings showed a high correlation of (r = 0.85) between the two isolates. Analysis of variance revealed significant differences in genotypes and the genotype by isolate interaction (P <0.0001), while no significant difference was observed between the isolates (P = 0.09) (Table 1). Additionally, 77 germplasm lines were identified with less than 25% disease severity when infected by either isolate, providing a wide range of germplasm lines for developing resistant varieties.

3.2 Distribution of SNP markers and linkage disequilibrium decay

The 33,378 SNPs were distributed across all twenty chromosomes (Figure 2), with an average of 1669 SNPs per chromosome. Chromosome 18 exhibited the highest number of SNPs (2771 SNPs), while chromosome 20 had the lowest number (1106 SNPs). The average distance between the two markers was approximately 28 kb. Chromosome 18 had the smallest inter-marker distance of 21 kb, while chromosome 20 had the largest distance of 43.3 kb.

The linkage disequilibrium (LD) analysis filtered 33,378 SNPs to calculate pairwise r^2 value, which were plotted against the physical distance. The fitted curve (red line) in the graph estimated the rate of LD decay (Figure 3). The blue lines represented half the maximum r^2 value, and the intersection of the green and blue lines indicated the distance in kilobase pair (kb) the maximum r^2 value dropped to half. The estimated LD decay rate reached half of its maximum value at approximately 249.47 kb, which is consistent with previous soybean GWAS studies (Hwang et al., 2014; Dhanapal et al., 2015).

3.3 Population structure

Population structure analysis was performed using STRUCTURE software to determine the number of subpopulations among 246 genotypes. The likelihood value [LnP (D)] was obtained from analysis using the 33,378 SNPs was used to calculate the delta K value, which determined the best-fitted subpopulation. The delta K graph showed two peak values at K=2 and K=6. However, the best-fitted subpopulation group was determined as K=6 because it successfully differentiated soybean accessions collected from different locations (Figure 4A). Under K=6, the 246 accessions were distributed across six groups: group 1 (12 accessions), group 2 (34), group 3 (19), group 4 (87), group 5 (41), and group 6 (53) (Figure 4B). Further, groups 1 and 4 primarily comprised accessions from South Korea. At the same time, China and Vietnam dominated group 2. Accessions from China were predominantly present in group 3, accessions from the US were mainly found in group 5, and accessions from Japan prevailed in group 6 (Supplementary Table 1).
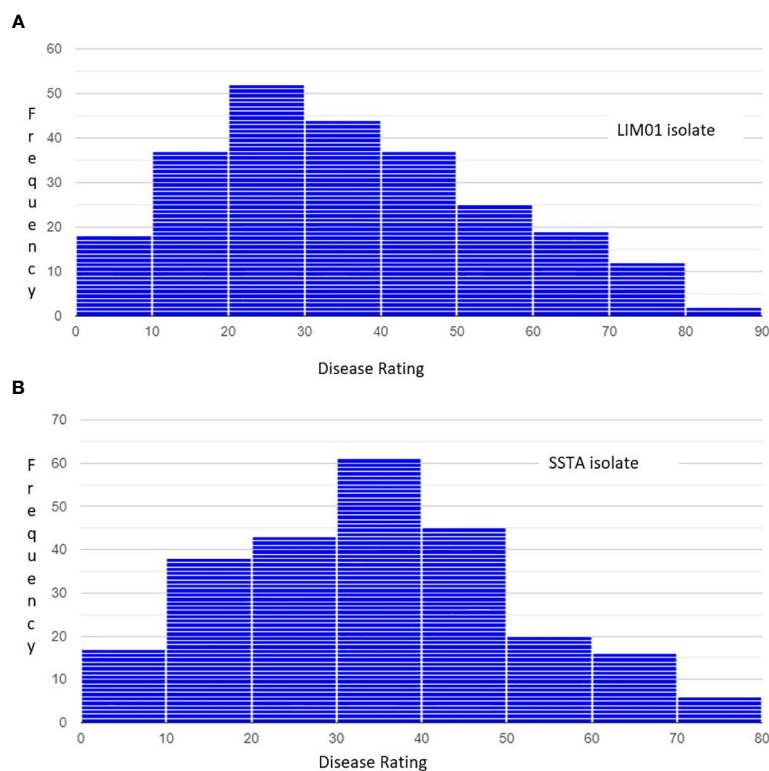
FIGURE 1
Frequency distribution of disease severity ratings among 246 soybean accessions against *Corynespora cassiicola* infection. Severity ratings for check lines included for reference **(A)** LIM01: The severity ratings for Council is 6.05 and for Henderson is 64.38 **(B)** SSTA: The severity ratings for Council is 5.67 and for Henderson is 61.23.

3.4 Genomic regions identified for target spot

A total of 14 marker-trait associations with a significance threshold of $-log10$ (P) $\geq$ 3.5 and P $\leq$ 0.0003 were identified with disease severity ratings in response to the LIM01 isolate. These associations were found across 12 chromosomes of the soybean genome, with chromosomes 3 and 15 harboring two associated SNPs each (Table 2, Figure 5A). Notably, the allelic effect for disease severity ranged from -7.92 (ss715584729, Chr. 3) to 6.5 (ss715637816, Chr. 20). In terms of traits involving disease ratings, a major allele is desirable when the allele effect is positive, whereas a minor allele is preferred for the disease resistance when allele effect is negative.
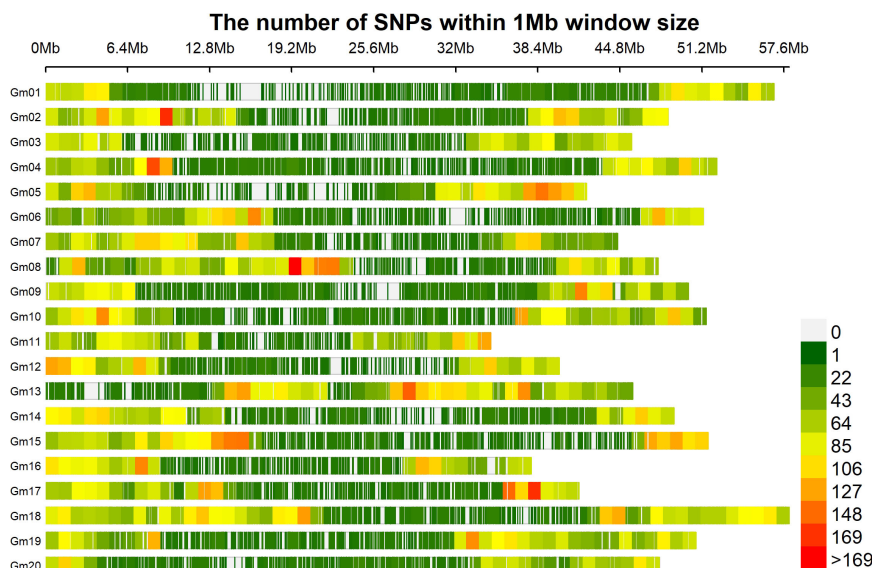
For disease severity ratings in response to the SSTA isolate, a total of 33 marker-trait associations were discovered (Figure 5B). These associations were distributed among 10 chromosomes, with chromosome 10 exhibiting the highest number of SNPs (13) and

chromosome 20 having seven associated SNPs. Additionally, the allelic effect for disease severity ranged from -8.96 (ss715612259, Chr. 12) to 9.98 (ss715632608, Chr. 18).

Common significantly associated SNPs for disease severity ratings for both *C. cassiicola* isolates are important to develop a marker-based selection strategy for target spot. Six SNPs were found to be significantly associated with disease severity ratings for both *C. cassiicola* isolates (Figure 5). Among these SNPs, only one of the SNP (ss715621861 on Chr. 15) exhibited a negative allelic effect, while the remaining five common SNPs showed a positive allele effect for disease severity ratings in response to both isolates (Table 2). Among the 77 lines identified with disease severity below 25% upon infection with either isolate, the majority had favorable alleles for ss715587169 (83.1%), ss715606800 (79.2%), and ss715632608 (98.7%). However, only 39%, 55.8%, and 57.1% had favorable alleles for ss715621861, ss715637816, and ss715616244 respectively.

TABLE 1 Analysis of variance for soybean accessions with two *Corynespora cassiicola* isolates.

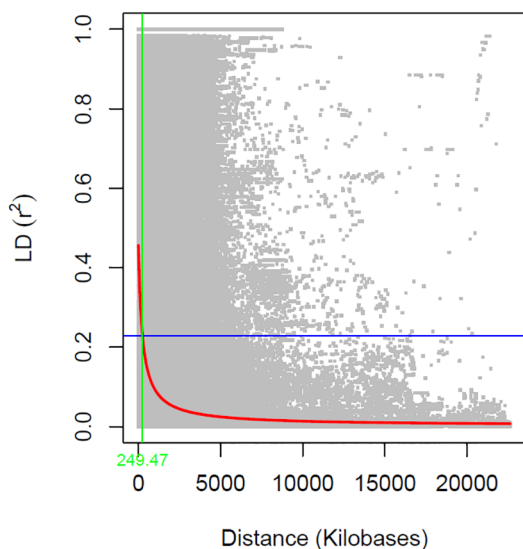| Source of variance	Degree of freedom	F value	*P* Value
Genotype	245	63.08	<0.0001
Isolate	1	9.8	0.0887
Genotype × Isolate	245	3.19	<0.0001

FIGURE 2
Distribution of SNP markers across the soybean genome. Different color represent the SNPs marker contained within 1Mb region. The red area indicates SNPs rich region, and the white area in between indicates the absence of SNP markers.

3.5 Candidate genes associated with target spot resistance

A total of 41 SNPs had significant associations with severity ratings in response to the two isolates. Candidate genes were searched for in the vicinity of 200kb region on either side of the associated SNPs. Several markers were identified on chromosomes 10 and 20 associated with disease severity to SSTA isolate; instead of gene mining for all the SNPs on these chromosomes, only a few

FIGURE 3
Genome-wide linkage disequilibrium (LD) decay rate estimated in 246 soybean accessions. The X- axis represents the genetic distance in Kb (killo bases) and Y-axis represents the squared correlation coefficient (r^2).

markers were selected that can span the entire region. Three markers were selected to cover the 13 SNPs in the region of chromosome 10, while two markers were chosen to cover the seven SNPs in the region of chromosome 20. In total, 615 and 640 genes were found in the vicinity of associated SNPs for disease ratings in response to LIM01 and SSTA isolates, respectively (Supplementary Table 2). Additionally, the six SNPs associated with disease severity ratings for both isolates were found to have 262 genes. Functional annotation of these genes revealed their involvement in many important defense pathways, such as against fungi, and other organisms such as bacteria, viruses, and oomycetes. In addition, chitin binding and chitinase activity, MAPKKK cascade, protein serine/threonine kinase activity, response to ethylene, biosynthesis of lignin, flavonoid, salicylic acid, brassinosteroid, and jasmonic acid (Grant et al., 2010; Morales et al., 2013).

3.6 Integration of information from GWAS and RNA-Seq analysis

To gain further insights into the genetic basis of target spot resistance, a transcriptomic study was conducted for two resistant (Council and Bedford) and two susceptible (Henderson and Pembina) genotypes at different timepoints during *C. cassiicola* following inoculation. RNA-Seq analysis identified differential gene expression patterns, with 6480 being upregulated and 7890 genes being downregulated in at least one of these genotypes (Supplementary Figure 1). The GO and KEGG analyses uncovered the response of the resistant and susceptible genotypes to infection by *C. cassiicola*. GO analysis of upregulated genes in Council enriched the biological and molecular functions process, while Bedford showed that biological processes were enriched. The
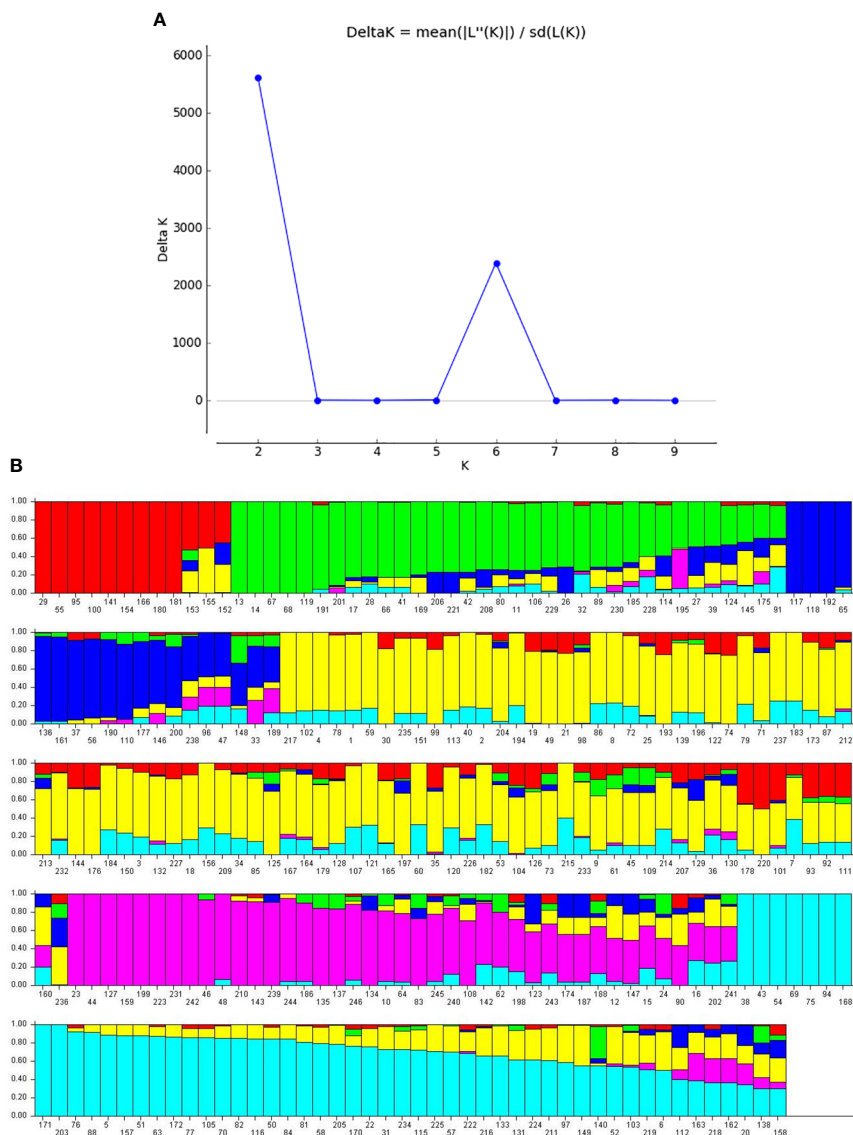
FIGURE 4
Population structure of the 246 soybean accessions. **(A)** Plot of delta k against putative k from 1 to 10. **(B)** Population structure estimates (k=6). Numbers on the X-axis represent germplasm and the membership coefficient displayed on the Y-axis. Accessions sharing the same group are assigned the same color.

Bedford genotype showed enriched biological processes for defense response, response to external biotic stimulus and immune responses, and plant-type cell wall biogenesis (Figure 6A). The upregulated DEGs in Council enriched processes like defense response, hormone-mediated signaling, chitin metabolism/catabolism, and responses to auxin and biotic stimuli (Figure 6B). Regarding molecular function, protein serine/threonine kinases, chitin binding, chitinase, and oxidoreductases. Certain molecular functions, such as protein serine/threonine kinase activity, chitinase activity, and oxidoreductase activity, were also increased (Figure 6C). Furthermore, the KEGG enrichment analysis was applied to the upregulated DEGs that were commonly observed across all four genotypes after being inoculated in response to *C. cassiicola*. The results revealed significant enrichments (with FDR-adjusted p ≤ 0.05) in several key pathways (Figure 6D). These

pathways include metabolic pathways, biosynthesis of secondary metabolites, phenylpropanoid biosynthesis, flavonoid biosynthesis, plant hormone signal transduction, and MAPK signaling pathway.

To integrate the RNA-Seq differentially expressed genes (DEGs) with the candidate genes identified through GWAS analysis, a comprehensive analysis was performed. A total of 131 genes in the vicinity of associated markers were upregulated after *C. cassiicola* inoculation in at least one of the genotypes. Among these genes, 24 were specifically upregulated only in the resistant genotypes, while 31 genes were upregulated only in the susceptible genotypes. Furthermore, 115 genes located near the associated markers showed downregulation after *C. cassiicola* inoculation, with 29 genes being specifically downregulated only in the resistant genotypes and 16 genes downregulated only in the susceptible genotypes (Supplementary Table 3). Notably, eight genes

TABLE 2 The SNPs (single-nucleotide polymorphisms) associated with target spot resistance with each of two *C. cassiicola* isolates.

Isolate	SNP Marker	Chr[a]	Position (bp)	−log10 (*P*-Value)	FDR[b]	Allelic effect[c]	MAF[d]
LIM01	ss715584729	03	2,091,140	6.436	0.003056	-7.915	0.132
	ss715586174	03	40,847,194	4.321	0.176922	-6.19	0.108
	ss715587169	04	15,841,584	7.713	0.000323	5.72	0.246
	ss715590348	05	1,952,710	4.4	0.166151	-4.012	0.421
	ss715594534	06	46,295,298	3.885	0.395674	-3.575	0.423
	ss715606800	10	39,534,982	6.964	0.00121	4.825	0.329
	ss715609438	11	11,087,536	5.593	0.017023	-4.842	0.248
	ss715612322	12	3,373,730	3.725	0.483678	3.445	0.262
	ss715616244	13	40,878,186	3.845	0.397328	4.061	0.272
	ss715621861	15	4,137,385	5.273	0.029659	-5.218	0.236
	ss715622335	15	48,361,099	4.884	0.062293	-4.857	0.173
	ss715624869	16	36,571,566	4.035	0.30799	-3.614	0.374
	ss715632608	18	7,161,394	3.449	0.847014	5.94	0.057
	ss715637816	20	37,162,511	10.039	3.05E-06	6.454	0.264
SSTA	ss715587167	04	15,781,556	4.322	0.173075	-5.389	0.423
	ss715587169	04	15,841,584	4.207	0.173075	5.502	0.246
	ss715592934	06	12,584,572	3.672	0.30788	5.686	0.167
	ss715606800	10	39,534,982	5.943	0.030607	6.255	0.329
	ss715606796	10	39,521,125	5.737	0.030607	6.258	0.301
	ss715606801	10	39,543,852	5.371	0.047377	5.847	0.333
	ss715606793	10	39,500,426	5.205	0.052058	6.288	0.266
	ss715606833	10	39,716,038	4.171	0.173075	5.775	0.248
	ss715606779	10	39,456,010	4.055	0.208483	5.411	0.289
	ss715606822	10	39,670,594	3.734	0.301446	-5.779	0.189
	ss715606823	10	39,682,612	3.636	0.30788	4.904	0.254
	ss715606832	10	39,714,825	3.61	0.30788	4.695	0.321
	ss715606829	10	39,709,662	3.562	0.30788	-5.948	0.167
	ss715606817	10	39,666,999	3.547	0.30788	4.861	0.252
	ss715606824	10	39,687,502	3.544	0.30788	5.279	0.323
	ss715606653	10	38,753,114	3.528	0.309125	-5.303	0.222
	ss715612259	12	33,175,894	4.723	0.090209	-8.969	0.091
	ss715616244	13	40,878,186	4.268	0.173075	6.35	0.272
	ss715615362	13	18,693,392	3.686	0.30788	-6.957	0.14
	ss715621861	15	4,137,385	3.937	0.241394	-6.513	0.236
	ss715621851	15	4,119,316	3.867	0.2519	6.639	0.23
	ss715624393	16	31,349,749	3.722	0.301446	-5.101	0.35
	ss715624785	16	35,789,070	3.568	0.30788	5.876	0.283
	ss715627202	17	37,097,907	3.889	0.2519	-5.52	0.346
	ss715632608	18	7,161,394	4.185	0.173075	9.985	0.057

(Continued)

Isolate	SNP Marker	Chr[a]	Position (bp)	−log10 (P-Value)	FDR[b]	Allelic effect[c]	MAF[d]
	ss715637816	20	37,162,511	4.85	0.090209	5.846	0.264
	ss715637481	20	34,776,314	4.738	0.090209	-5.631	0.419
	ss715637482	20	34,795,864	4.444	0.150129	-5.463	0.415
	ss715637615	20	35,651,832	4.028	0.208483	-4.514	0.498
	ss715637471	20	34,699,114	3.755	0.301446	5.219	0.396
	ss715637820	20	37,173,335	3.594	0.30788	4.947	0.335
	ss715637817	20	37,167,430	3.591	0.30788	4.95	0.327

[a]Chromosome.
[b]False discovery rate (FDR) adjusted p values.
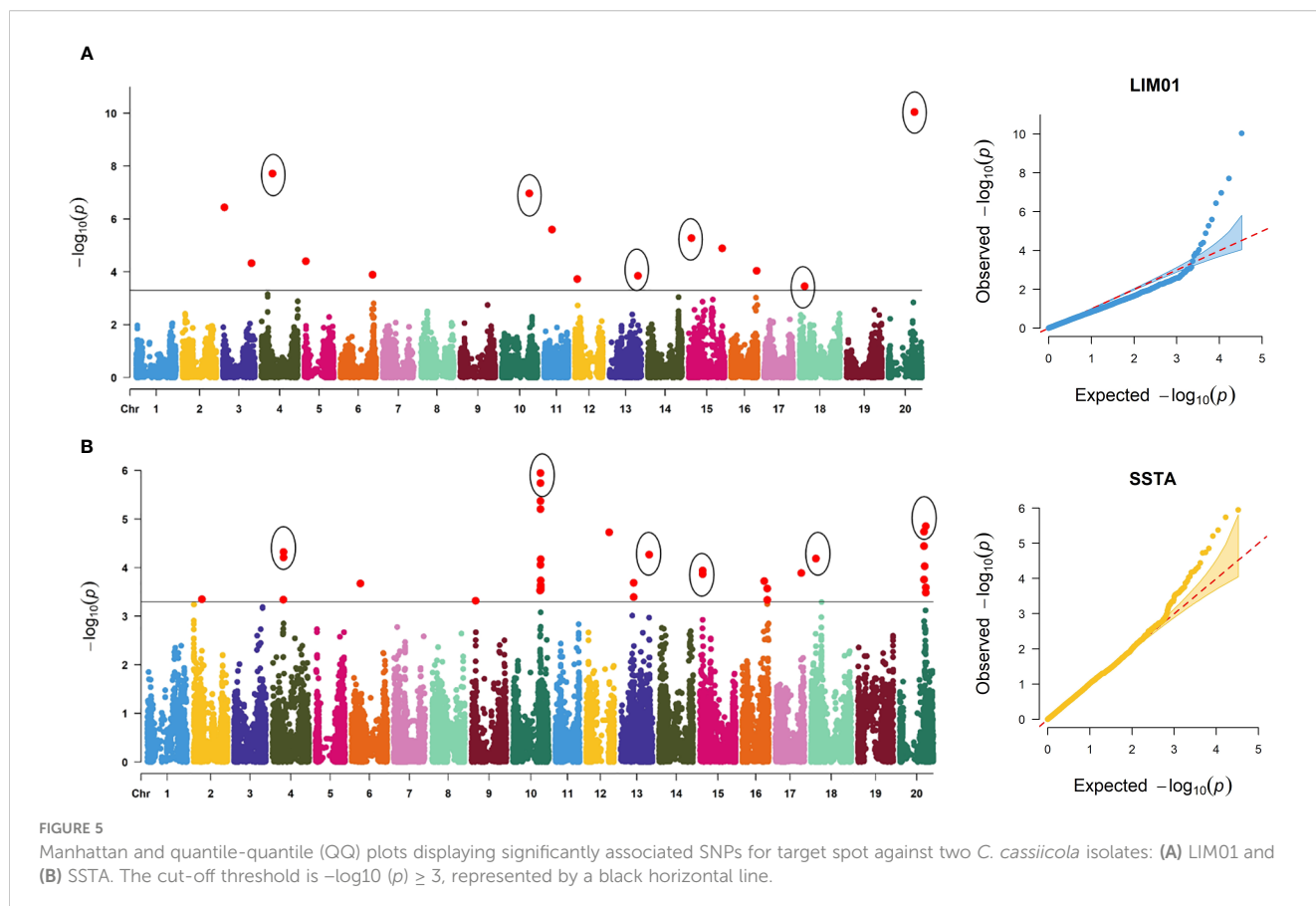[c]Minor allele effect of the SNP.
[d]Minor allele frequency.

exhibited interesting expression patterns after *C. cassiicola* inoculation, showing upregulation in one genotype while being downregulated in another genotype (Supplementary Table 3). These genes may be involved in complex regulatory mechanisms underlying target spot resistance and warrant further investigation.
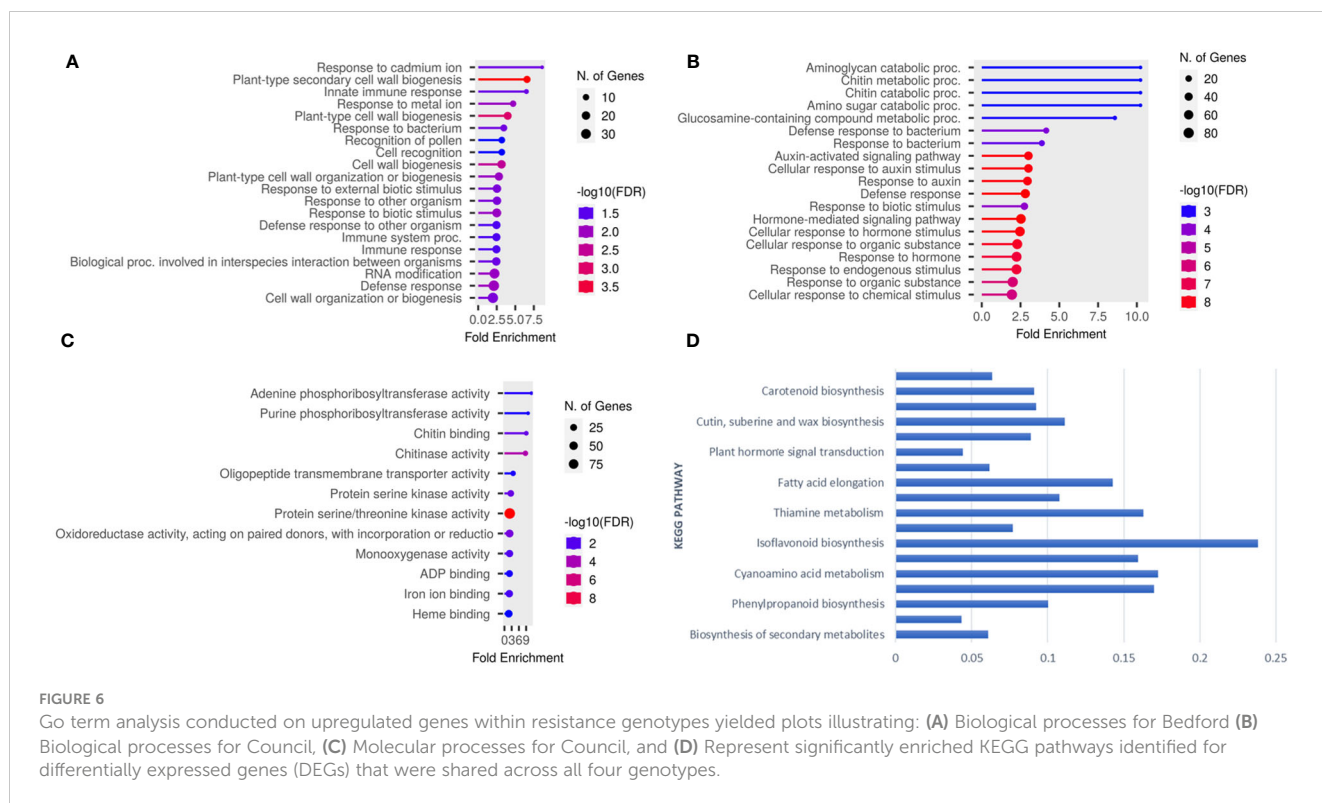
4 Discussion

Target spot is a significant threat to soybean production, particularly in warm and humid regions, and potentially leading to substantial economic losses. With the availability of a well-assembled genome sequence, extensive marker information, and a diverse germplasm collection in soybean (Song et al., 2013), it is possible to explore the relationship between phenotypic variation and genetic information. To the best of our knowledge, this study represents the first attempt to identify SNP markers associated with target spot resistance in soybean.

Disease severity ratings varied greatly among genotypes for each isolate, providing valuable data along with genotypic information for identifying marker-trait associations. In this study, 14 and 33 marker-trait associations were identified for the LIM01 and SSTA

FIGURE 5
Manhattan and quantile-quantile (QQ) plots displaying significantly associated SNPs for target spot against two *C. cassiicola* isolates: **(A)** LIM01 and **(B)** SSTA. The cut-off threshold is −log10 (p) ≥ 3, represented by a black horizontal line.

FIGURE 6
Go term analysis conducted on upregulated genes within resistance genotypes yielded plots illustrating: **(A)** Biological processes for Bedford **(B)** Biological processes for Council, **(C)** Molecular processes for Council, and **(D)** Represent significantly enriched KEGG pathways identified for differentially expressed genes (DEGs) that were shared across all four genotypes.

isolates of *C. cassiicola*, respectively. By examining the candidate genes surrounding the associated SNPs, a total of 993 candidate genes were discovered. To narrow down this list and enhance confidence in identifying target spot defense-associated genes, the GWAS results were scrutinized and combined with RNA-Seq data of soybean gene expression after *C. cassiicola* infection. This integration led to the identification of 238 differentially expressed genes (DEGs) among the candidate genes. Among the 238 genes, several were determined to have a role in the defense response to pathogen attack. These genes were involved in defense response to pathogen, innate immune response, chitinase activity, histone H3-K9 methylation, salicylic acid mediated signaling pathway, kinase activity, and biosynthesis of flavonoid, jasmonic acid, phenylpropanoid, and wax.

Leucine-rich repeat (LRR) domains are commonly found in plant immune receptors and defense genes. LRR-receptor-like kinase (RLK) genes are primarily involved in development and stress responses and play a crucial role in plant defense mechanisms. Flagellin-sensitive 2 (FLS2), a LRR-RLK gene binds bacterial flagellin and initiates pathogen-associated molecular patterns (PAMPs)-triggered immunity (Chinchilla et al., 2007). LRR-RLK genes have also been implicated in resistance to various pathogens, such as *Cercospora sojina* (McDonald et al., 2023) and *Phytophthora sojae* (Wang et al., 2021) in soybean, *Blumeria graminis* in wheat (Hu et al., 2018), *Hyaloperonospora arabidopsidis* in Arabidopsis (Hok et al., 2011) and *Fusarium graminearum* in barley and wheat (Thapa et al., 2018). The RPS2 gene belongs to a LRR Class of disease resistance that plays a vital

role in conferring resistance to a broad spectrum of pathogens (Li et al., 2019). In our study, four LRR-RLK genes and four genes with LRR domains belonging to the disease resistance family were differentially expressed and located near associated SNPs, suggesting their potential involvement in target spot resistance.

RPM1-interacting protein 4 (RIN4) was found near ss715621861 on Chromosome 15, was differentially expressed after *C. cassicola* infection. RIN4 is a plant immunity regulator that plays a crucial role in both PAMP-triggered immunity (PTI) and effector-triggered immunity (ETI) (Zhao et al., 2021). Notably, RIN4 is an inherently disordered protein (IDP) except for the region involved in pathogen-induced post-translational modifications. Functionally, it acts as a central hub protein, binding to multiple proteins and serving as a center for immune signaling pathway formation (Cui et al., 2010; Liu et al., 2011; Oldfield and Dunker, 2014; Toruño et al., 2018).

Plant-specific NAC domain proteins are the class of transcription factors that play crucial role in regulating stress tolerance and defense response against various abiotic and biotic challenges (Oh et al., 2005; Puranik et al., 2012; Yuan et al., 2019). In the narrowed-down candidate genes list, four genes were identified: NAC1, NAC47, and two copies of NAC90. NAC1 gene has been shown to trigger a defense signal upon infection of *Pseudomonas syringae pv. tomato* in tomato (Kud, 2017). NAC090, along with other NAC transcription factors, plays an essential role in regulating accumulation of salicylic acid and leaf senescence (Kim et al., 2018; Zhang et al., 2020). In cotton, both NAC1 and NAC90 are involved in stress response to the fungal pathogen *Verticillium dahlia* and

positively regulate resistance to verticillium wilt (Wang et al., 2016; Bai et al., 2022).

The GWAS analysis identified two copies of the plant U-box protein PUB23 and one copy of PUB25, were differentially expressed after *C. cassiicola* infection in RNA-Seq analysis. PUB23, along with other U-box proteins, has been found to negatively affect tolerance to drought and *Fusarium oxysporum* (Cho et al., 2008; Chen et al., 2014). Similarly, PUB25, along with PUB26, plays a role in the degradation of MYB6 and negatively impacts resistance to *Verticillium dahlia* (Ma et al., 2021). Exploring the structural and functional characteristics of these U-box proteins and their specific impact on *C. cassiicola* resistance would provide valuable insights into the underlying molecular mechanisms.

Durable resistance against multiple pathogens can be achieved through the manipulation of genes encoding a small number of pathogen recognition proteins. It has been observed that several genes and pathways play common roles in response to both abiotic stress and pathogen attacks. For example, the phenylpropanoid pathway is activated in response to pathogen attack as well as iron deficiency chlorosis (IDC) (Zabala et al., 2006; Waters et al., 2018). The RLK, *GmRLK18-1* has been found to confer pleiotropic resistance to both *Fusarium virguliforme* and *Heterodera glycines* (Srour et al., 2012). Similarly, a set of two genes (*GmSNAP18* and *GmSNAP11*), has been shown to provide resistance against both *Heterodera glycines* and *Rotylenchulus reniformis* (Usovsky et al., 2021). A close linkage of favorable alleles and the pleiotropic effect of genes play an essential role in marker-based selection process, enabling the simultaneous improvement of multiple traits. This approach is particularly beneficial for traits controlled by a few genes, such as disease resistance.

By collating the genomic regions associated from the current study and the previous publications assisted in identifying 11 colocalized regions in the genome linked to different biotic and abiotic stress, such as soybean cyst nematode (SCN), reniform nematode (RN), stem and root rot (*Phytophthora sojae*), sclerotinia stem rot (*Sclerotinia sclerotiorum*), sudden death syndrome (*Fusarium virguliforme*), yellow mosaic virus, and IDC (Supplementary Table 4).

One notable region, a block of less than two Mb on Chr.10 which contains 13 marker-trait association (MTA) for target spot, 14 for *Phytophthora sojae*, and one each for RN and SCN. This region encompasses several genes involved in defense mechanisms such as pathogenesis-related 4, LRR-RLK, *Arabidopsis* toxicos en levadura 6 (ATL6), enhanced disease resistance 1 (EDR1), WRKY3, WRKY4, MYB15, pentatricopeptide repeat, and genes belonging to protein kinase family. Pathogenesis-related protein (PR) are encoded by host genes that are selectively activated under pathogen or parasitic attack to provide a long-term resistance. These genes play a pivotal role in systemic acquired resistance in uninfected tissue distal to the original infection site (Sudisha et al., 2012; Jain and Khurana, 2018). Transcription factors like WRKY3, WRKY4 and MYB15 are prominent families involved in large-scale transcriptional reprogramming during plant immune responses against pathogen attacks (Eulgem and Somssich, 2007; Bosamia et al., 2020). For example, *WRKY3* and *WRKY4* has been shown to

enhance plant resistance to necrotrophic pathogens such as *Botrytis cinerea* in *Arabidopsis* (Lai et al., 2008). MYB15 activates lignin biosynthesis genes as part of effector-triggered immune responses, providing resistance to *Pseudomonas syringae* (Kim et al., 2020). In RNA-Seq study, some of these genes showed distinctive expression patterns in the four genotypes upon *C. cassiicola* inoculation (Supplementary Figure 2A).

Another region of interest is a 1.86 Mb block on Chr. 16, which contains two MTAs for target spot, one each for *Sclerotinia sclerotiorum*, IDC, and SCN. A heat map displaying DEGs identified within the 1.86 Mb block. It showcases the gene expression changes when comparing 24 hpi and 48 hpi of *C. cassiicola* with control for all four genotypes (Supplementary Figure 2B). This region contains 17 gene copies of disease resistance family protein/LRR family protein, four copies of disease resistance protein belonging to TIR-NBS-LRR family, four copies of receptor like protein (RLP), three copies of cysteine-rich RLK protein kinase 25, three copies of cytochrome P450 family protein, two copies of LRR-RLK, and two copies of LRR transmembrane protein kinase. These findings suggest the presence of closely linked blocks that confer resistance to various biotic and abiotic stress. The identification of these regions provides valuable insights for developing marker-assisted selection strategies in soybean breeding programs, enabling the development of cultivars with broad-spectrum resistance against multiple pathogens.

5 Conclusion

This study represents the first GWAS conducted in soybean to elucidate the genetic basis of resistance to target spot, a fungal-incited disease. The GWAS analysis identified six common SNPs associated with disease traits in response to two different *C. cassiicola* isolates. Furthermore, the integration of GWAS and RNA-Seq data allowed for the identification of 238 DEGs, several of which were found to be involved in LRR-RLK, LRR protein belonging to disease resistance disease resistance family protein, NAC transcription factor, cytochrome P450, RIN4, Pentatricopeptide repeat (PPR) protein, and plant U-box protein. Moreover, the study revealed the presence of 11 colocalized regions distributed across different chromosomes, which were found to contribute to multiple disease resistances. These regions represent potential QTL harboring genes associated with resistance against various biotic stresses. The identified genetic markers and candidate genes can serve as valuable resources for developing marker-assisted selection strategies, facilitating the development of improved soybean cultivars with enhanced resistance to target spot and potentially other related diseases.

Data availability statement

The data presented in the study are deposited in the National Center for Biotechnology Information (NCBI) repository, accession number is PRJNA1015685. Illumina Infinium SoySNP50K Bead

Chip data is available in the Soybase database (https://www.soybase.org).

Author contributions

SP: Conceptualization, Data curation, Formal Analysis, Methodology, Visualization, Writing – original draft, Writing – review & editing, Investigation, Software. JP: Conceptualization, Formal Analysis, Methodology, Software, Writing – review & editing. KB: Methodology, Writing – review & editing. JK: Conceptualization, Funding acquisition, Methodology, Resources, Supervision, Writing – review & editing.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2023.1255763/full#supplementary-material

References

Bai, S., Niu, Q., Wu, Y., Xu, K., Miao, M., and Mei, J. (2022). Genome-Wide Identification of the NAC Transcription Factors in *Gossypium hirsutum* and Analysis of Their Responses to Verticillium wilt. *Plants (Basel)* 11 (19), 2661. doi: 10.3390/plants11192661

Bao, Y., Kurle, J. E., Anderson, G., and Young, N. D. (2015). Association mapping and genomic prediction for resistance to sudden death syndrome in early maturing soybean germplasm. *Mol. Breed.* 35 (6), 1–14. doi: 10.1007/s11032-015-0324-3

Barrett, R. E., Bartz, J. A., Blazquez, C. H., Correll, J. C., Crill, J. P., Gardner, R., G.,, et al (1991). *Compendium of tomato diseases*. American phytopathological society press.

Blazquez, C. (1967). Corynespora leaf spot of cucumber in *Proceedings of the Florida State horticultural society*, Vol. 80, pp. 177–177. https://scholar.google.com/scholar?

Bosamia, T. C., Dodia, S. M., Mishra, G. P., Ahmad, S., Joshi, B., Thirumalaisamy, P. P., et al. (2020). Unraveling the mechanisms of resistance to *Sclerotium rolfsii* in peanut (*Arachis hypogaea* L.) using comparative RNA-Seq analysis of resistant and susceptible genotypes. *PloS One* 15 (8), e0236823. doi: 10.1371/journal.pone.0236823

Bradbury, P. J., Zhang, Z., Kroon, D. E., Casstevens, T. M., Ramdoss, Y., and Buckler, E. S. (2007). TASSEL: software for association mapping of complex traits in diverse samples. *Bioinformatics* 23 (19), 2633–2635. doi: 10.1093/bioinformatics/btm308

Chamarthi, S. K., Kaler, A. S., Abdel-Haleem, H., Fritschi, F. B., Gillman, J. D., Ray, J. D., et al. (2021). Identification and confirmation of loci associated with canopy wilting in soybean using genome-wide association mapping. *Front. Plant Sci.* 12 (1371). doi: 10.3389/fpls.2021.698116

Chee, K. (1988). Studies on sporulation, pathogenicity and epidemiology of Corynespora cassiicola on Hevea rubber. *J. Nat. Rubber Res.* 3(1), 21–29.

Chen, Y. C., Wong, C. L., Muzzi, F., Vlaardingerbroek, I., Kidd, B. N., and Schenk, P. M. (2014). Root defense analysis against *Fusarium oxysporum* reveals new regulators to confer resistance. *Sci. Rep.* 4 (1), 5584. doi: 10.1038/srep05584

Chen, S., Zhou, Y., Chen, Y., and Gu, J. (2018). fastp: an ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics* 34 (17), i884–i890. doi: 10.1093/bioinformatics/bty560

Cheng, Y., Ma, Q., Ren, H., Xia, Q., Song, E., Tan, Z., et al. (2017). Fine mapping of a Phytophthora-resistance gene *RpsWY* in soybean (*Glycine max* L.) by high-throughput genome-wide sequencing. *Theor. Appl. Genet.* 130 (5), 1041–1051. doi: 10.1007/s00122-017-2869-5

Chinchilla, D., Zipfel, C., Robatzek, S., Kemmerling, B., Nürnberger, T., Jones, J. D., et al. (2007). A flagellin-induced complex of the receptor FLS2 and BAK1 initiates plant defence. *Nature* 448 (7152), 497–500. doi: 10.1038/nature05999

Cho, S. K., Ryu, M. Y., Song, C., Kwak, J. M., and Kim, W. T. (2008). Arabidopsis PUB22 and PUB23 are homologous U-Box E3 ubiquitin ligases that play combinatory roles in response to drought stress. *Plant Cell* 20 (7), 1899–1914. doi: 10.1105/tpc.108.060699

Conner, K., Hagan, A., and Zhang, L. (2013). First report of *Corynespora cassiicola*-incited target spot on cotton in Alabama. *Plant Dis.* 97 (10), 1379–1379. doi: 10.1094/PDIS-02-13-0133-PDN

Coser, S. M., Chowda Reddy, R. V., Zhang, J., Mueller, D. S., Mengistu, A., Wise, K. A., et al. (2017). Genetic architecture of charcoal rot (Macrophomina phaseolina) resistance in soybean revealed using a diverse panel. *Front. Plant Sci.* 8. doi: 10.3389/fpls.2017.01626

Cui, H., Wang, Y., Xue, L., Chu, J., Yan, C., Fu, J., et al. (2010). Pseudomonas syringae effector protein AvrB perturbs Arabidopsis hormone signaling by activating MAP kinase 4. *Cell Host Microbe* 7 (2), 164–175. doi: 10.1016/j.chom.2010.01.009

Dhanapal, A. P., Ray, J. D., Singh, S. K., Hoyos-Villegas, V., Smith, J. R., Purcell, L. C., et al. (2015). Genome-wide association study (GWAS) of carbon isotope ratio (δ 13 C) in diverse soybean [*Glycine max* (L.) Merr.] genotypes. *Theor. Appl. Genet.* 128 (1), 73–91. doi: 10.1007/s00122-014-2413-9

Dixon, L. J., Schlub, R. L., Pernezny, K., and Datnoff, L. (2009). Host specialization and phylogenetic diversity of *Corynespora cassiicola*. *Phytopathology* 99 (9), 1015–1027. doi: 10.1094/PHYTO-99-9-1015

Duan, Y., Xin, W., Lu, F., Li, T., Li, M., Wu, J., et al. (2019). Benzimidazole-and QoI-resistance in Corynespora cassiicola populations from greenhouse-cultivated cucumber: An emerging problem in China. *Pesticide Biochem. Physiol.* 153, 95–105. doi: 10.1016/j.pestbp.2018.11.006

Earl, D. A., and Vonholdt, B. M. (2012). STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conserv. Genet. Resour.* 4 (2), 359–361. doi: 10.1007/s12686-011-9548-7

Eulgem, T., and Somssich, I. E. (2007). Networks of WRKY transcription factors in defense signaling. *Curr. Opin. Plant Biol.* 10 (4), 366–371. doi: 10.1016/j.pbi.2007.04.020

Faske, T. (2016) *Arkansas soybeans: target spot – what do we know*. Available at: https://agfax.com/2016/11/02/arkansas-soybeans-target-spot-what-do-we-know/ (Accessed Nov, 2016).

Ge, S. X., Jung, D., and Yao, R. (2020). ShinyGO: a graphical gene-set enrichment tool for animals and plants. *Bioinformatics* 36 (8), 2628–2629. doi: 10.1093/bioinformatics/btz931

Godoy, C. V. (2015). "Target spot," in *Compendium of soybean diseases and pests*. Eds. G. L. Hartman, J. C. Rupe, E. J. Sikora, L. L. Domier, J. A. Davis and K. L. Steffey (St. Paul, MN: American: Phytopathological Society), 62–63.

Grant, D., Nelson, R. T., Cannon, S. B., and Shoemaker, R. C. (2010). SoyBase, the USDA-ARS soybean genetics and genomics database. *Nucleic Acids Res.* 38 (Database issue), D843–D846. doi: 10.1093/nar/gkp798

Hartman, G. L., Rupe, J. C., Sikora, E. J., Domier, L. L., Davis, J. A., and Steffey, K. L. (2015). *Compendium of soybean diseases and pests* (St. Paul, MN: American Phytopathological Society). p. 56–59.

Hasan, N., Choudhary, S., Naaz, N., Sharma, N., and Laskar, R. A. (2021). Recent advancements in molecular marker-assisted selection and applications in plant breeding programmes. *J. Genet. Eng. Biotechnol.* 19 (1), 128. doi: 10.1186/s43141-021-00231-1

Hok, S., Danchin, E. G., Allasia, V., Panabières, F., Attard, A., and Keller, H. (2011). An Arabidopsis (malectin-like) leucine-rich repeat receptor-like kinase contributes to downy mildew disease. *Plant Cell Environ.* 34 (11), 1944–1957. doi: 10.1111/j.1365-3040.2011.02390.x

Hu, P., Liu, J., Xu, J., Zhou, C., Cao, S., Zhou, W., et al. (2018). A malectin-like/leucine-rich repeat receptor protein kinase gene, RLK-V, regulates powdery mildew resistance in wheat. *Mol. Plant Pathol.* 19 (12), 2561–2574. doi: 10.1111/mpp.12729

Huang, X., Sang, T., Zhao, Q., Feng, Q., Zhao, Y., Li, C., et al. (2010). Genome-wide association studies of 14 agronomic traits in rice landraces. *Nat. Genet.* 42 (11), 961–967. doi: 10.1038/ng.695

Hwang, E.-Y., Song, Q., Jia, G., Specht, J. E., Hyten, D. L., Costa, J., et al. (2014). A genome-wide association study of seed protein and oil content in soybean. *BMC Genomics* 15 (1), 1–12. doi: 10.1186/1471-2164-15-1

Iquira, E., Humira, S., and François, B. (2015). Association mapping of QTLs for sclerotinia stem rot resistance in a collection of soybean plant introductions using a genotyping by sequencing (GBS) approach. *BMC Plant Biol.* 15 (1), 1–12. doi: 10.1186/s12870-014-0408-y

Jain, D., and Khurana, J. P. (2018). Role of pathogenesis-related (PR) proteins in plant defense mechanism. In *Molecular aspects of plant-pathogen interaction* (Springer).

Kaler, A. S., Ray, J. D., Schapaugh, W. T., Davies, M. K., King, C. A., and Purcell, L. C. (2018). Association mapping identifies loci for canopy coverage in diverse soybean genotypes. *Mol. Breed.* 38 (5), 1–12. doi: 10.1007/s11032-018-0810-5

Kim, S. H., Lam, P. Y., Lee, M.-H., Jeon, H. S., Tobimatsu, Y., and Park, O. K. (2020). The Arabidopsis R2R3 MYB transcription factor MYB15 is a key regulator of lignin biosynthesis in effector-triggered immunity. *Front. Plant Sci.* 11, 583153. doi: 10.3389/fpls.2020.583153

Kim, H. J., Park, J.-H., Kim, J., Kim, J. J., Hong, S., Kim, J., et al. (2018). Time-evolving genetic networks reveal a NAC troika that negatively regulates leaf senescence in Arabidopsis. *Proc. Natl. Acad. Sci.* 115 (21), E4930–E4939. doi: 10.1073/pnas.1721523115

Koenning, S., Creswell, T., Dunphy, E., Sikora, E., and Mueller, J. (2006). Increased occurrence of target spot of soybean caused by *Corynespora cassiicola* in the Southeastern United States. *Plant Dis.* 90 (7), 974–974. doi: 10.1094/PD-90-0974C

Korte, A., and Farlow, A. (2013). The advantages and limitations of trait analysis with GWAS: a review. *Plant Methods* 9 (1), 1–9. doi: 10.1186/1746-4811-9-29

Kud, J. (2017). Regulation of the defense-related SlNAC1 transcription factor during defense signaling in tomato. [dissertation]. (University of Idaho).

Lai, Z., Vinod, K. M., Zheng, Z., Fan, B., and Chen, Z. (2008). Roles of arabidopsisWRKY3 and WRKY4 transcription factors in plant responses to pathogens. *BMC Plant Biol.* 8 (1), 68. doi: 10.1186/1471-2229-8-68

Li, L., Guo, N., Niu, J., Wang, Z., Cui, X., Sun, J., et al. (2016). Loci and candidate gene identification for resistance to Phytophthora sojae via association analysis in soybean [Glycine max (L.) Merr.]. *Mol. Genet. Genomics* 291 (3), 1095–1103. doi: 10.1007/s00438-015-1164-x

Li, Z., Huang, J., Wang, Z., Meng, F., Zhang, S., Wu, X., et al. (2019). Overexpression of Arabidopsis nucleotide-binding and leucine-rich repeat genes RPS2 and RPM1 (D505V) confers broad-spectrum disease resistance in rice. *Front. Plant Sci.* 10, 417. doi: 10.3389/fpls.2019.00417

Lin, F., Wani, S. H., Collins, P. J., Wen, Z., Li, W., Zhang, N., et al. (2020). QTL mapping and GWAS for identification of loci conferring partial resistance to *Pythium sylvaticum* in soybean (Glycine max (L.) Merr). *Mol. Breed.* 40 (6), 1–11. doi: 10.1007/s11032-020-01133-9

Liu, J., Elmore, J. M., Lin, Z.-J. D., and Coaker, G. (2011). A receptor-like cytoplasmic kinase phosphorylates the host target RIN4, leading to the activation of a plant innate immune receptor. *Cell Host Microbe* 9 (2), 137–146. doi: 10.1016/j.chom.2011.01.010

Liu, X., Huang, M., Fan, B., Buckler, E. S., and Zhang, Z. (2016). Iterative usage of fixed and random effect models for powerful and efficient genome-wide association studies. *PloS Genet.* 12 (2), e1005767. doi: 10.1371/journal.pgen.1005767

Love, M. I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-Seq data with DESeq2. *Genome Biol.* 15 (12), 550. doi: 10.1186/s13059-014-0550-8

Ma, A., Zhang, D., Wang, G., Wang, K., Li, Z., Gao, Y., et al. (2021). Verticillium dahliae effector VDAL protects MYB6 from degradation by interacting with PUB25 and PUB26 E3 ligases to enhance *Verticillium* wilt resistance. *Plant Cell* 33 (12), 3675–3699. doi: 10.1093/plcell/koab221

McDonald, S. C., Buck, J., Song, Q., and Li, Z. (2023). Genome-wide association study reveals novel loci and a candidate gene for resistance to frogeye leaf spot (*Cercospora sojina*) in soybean. *Mol. Genet. Genomics* 298 (2), 441–454. doi: 10.1007/s00438-022-01986-z

Morales, A. M., O'Rourke, J. A., Van De Mortel, M., Scheider, K. T., Bancroft, T. J., Borém, A., et al. (2013). Transcriptome analyses and virus induced gene silencing identify genes in the Rpp4-mediated Asian soybean rust resistance pathway. *Funct. Plant Biol.* 40 (10), 1029–1047. doi: 10.1071/FP12296

Neeraja, C. N., Maghirang-Rodriguez, R., Pamplona, A., Heuer, S., Collard, B. C. Y., Septiningsih, E. M., et al. (2007). A marker-assisted backcross approach for developing submergence-tolerant rice cultivars. *Theor. Appl. Genet.* 115 (6), 767–776. doi: 10.1007/s00122-007-0607-0

Oh, S.-K., Lee, S., Yu, S. H., and Choi, D. (2005). Expression of a novel NAC domain-containing transcription factor (CaNAC1) is preferentially associated with incompatible interactions between chili pepper and pathogens. *Planta* 222 (5), 876–887. doi: 10.1007/s00425-005-0030-1

Oktavia, F., Sudarsono, and Dinarti, D. (2021). Preliminary QTL detection for Corynespora Leaf Fall disease resistance in rubber plant. *Nusantara Bioscience* 13 (1), 52–60. doi: 10.13057/nusbiosci/n130108

Oldfield, C. J., and Dunker, A. K. (2014). Intrinsically disordered proteins and intrinsically disordered protein regions. *Annu. Rev. Biochem.* 83 (1), 553–584. doi: 10.1146/annurev-biochem-072711-164947

Olive, L. S., Bain, D., and Lefebvre, C. (1945). A leaf spot of Cowpea and Soybean caused by an undescribed species of Helminthosporium. *Phytopathology* 35 (10), 822–831.

Patel, S. J., Bowen, K. L., Patel, J. D., and Koebernick, J. C. (2022). Evaluating target spot (*Corynespora cassiicola*) resistance in soybean (*Glycine max* (L.) Merrill) in a controlled environment. *Crop Prot.* 159, 106018. doi: 10.1016/j.cropro.2022.106018

Patel, S., Patel, J., Silliman, K., Hall, N., Bowen, K., and Koebernick, J. (2023). Comparative Transcriptome Profiling Unfolds a Complex Defense and Secondary Metabolite Networks Imparting *Corynespora cassiicola* Resistance in Soybean (*Glycine max* (L.) Merrill). *Int. J. Mol. Sci.* 24 (13), 10563. doi: 10.3390/ijms241310563

Patro, R., Duggal, G., Love, M. I., Irizarry, R. A., and Kingsford, C. (2017). Salmon provides fast and bias-aware quantification of transcript expression. *Nat. Methods* 14 (4), 417–419. doi: 10.1038/nmeth.4197

Pritchard, J. K., Stephens, M., and Donnelly, P. (2000). Inference of population structure using multilocus genotype data. *Genetics* 155 (2), 945–959. doi: 10.1093/genetics/155.2.945

Puranik, S., Sahu, P. P., Srivastava, P. S., and Prasad, M. (2012). NAC proteins: regulation and role in stress tolerance. *Trends Plant Sci.* 17 (6), 369–381. doi: 10.1016/j.tplants.2012.02.004

Ravelombola, W. S., Qin, J., Shi, A., Nice, L., Bao, Y., Lorenz, A., et al. (2020). Genome-wide association study and genomic selection for tolerance of soybean biomass to soybean cyst nematode infestation. *PloS One* 15 (7), e0235089. doi: 10.1371/journal.pone.0235089

Remington, D. L., Thornsberry, J. M., Matsuoka, Y., Wilson, L. M., Whitt, S. R., Doebley, J., et al. (2001). Structure of linkage disequilibrium and phenotypic associations in the maize genome. *Proc. Natl. Acad. Sci.* 98 (20), 11479–11484. doi: 10.1073/pnas.201394398

Ribeiro, S., Label, P., Garcia, D., Montoro, P., and Pujade-Renaud, V. (2021). Transcriptome profiling in susceptible and tolerant rubber tree clones in response to cassiicolin Cas1, a necrotrophic effector from *Corynespora cassiicola*. *PloS One* 16 (7), e0254541. doi: 10.1371/journal.pone.0254541

Rondon, M. N., and Lawrence, K. S. (2019). *Corynespora cassiicola* isolates from soybean in Alabama detected with G143A mutation in the cytochrome b gene. *Plant Health Prog.* 20 (4), 247–249. doi: 10.1094/PHP-07-19-0046-BR

Roy, C. B., Liu, H., Rajamani, A., and Saha, T. (2019). Transcriptome profiling reveals genetic basis of disease resistance against *Corynespora cassiicola* in rubber tree (Hevea brasiliensis). *Curr. Plant Biol.* 17, 2–16. doi: 10.1016/j.cpb.2019.02.002

Schmutz, J., Cannon, S. B., Schlueter, J., Ma, J., Mitros, T., Nelson, W., et al. (2010). Genome sequence of the palaeopolyploid soybean. *Nature* 463, 178. doi: 10.1038/nature08670

Seaman, W. L., Shoemaker, R. A., and Peterson, E. A. (1965). Pathogenicity of Corynespora cassiicola on soybean. *Can. J. Bot.* 43 (11), 1461–1469. doi: 10.1139/b65-154

Shen, Y., Du, H., Liu, Y., Ni, L., Wang, Z., Liang, C., et al. (2019). Update soybean Zhonghuang 13 genome to a golden reference. *Sci. China Life Sci.* 62, 1257–1260. doi: 10.1007/s11427-019-9822-2

Song, Q., Hyten, D. L., Jia, G., Quigley, C. V., Fickus, E. W., Nelson, R. L., et al. (2013). Development and evaluation of SoySNP50K, a high-density genotyping array for soybean. *PloS One* 8 (1), e54985. doi: 10.1371/journal.pone.0054985

Song, Q., Hyten, D. L., Jia, G., Quigley, C. V., Fickus, E. W., Nelson, R. L., et al. (2015). Fingerprinting soybean germplasm and its utility in genomic research. *G3 Genes|Genomes|Genetics* 5 (10), 1999–2006. doi: 10.1534/g3.115.019000

Srour, A., Afzal, A. J., Blahut-Beatty, L., Hemmati, N., Simmonds, D. H., Li, W., et al. (2012). The receptor like kinase at Rhg1-a/Rfs2 caused pleiotropic resistance to sudden death syndrome and soybean cyst nematode as a transgene by altering signaling responses. *BMC Genomics* 13 (1), 368. doi: 10.1186/1471-2164-13-368

Sudisha, J., Sharathchandra, R. G., Amruthesh, K. N., Kumar, A., and Shetty, H. S. (2012). Pathogenesis related proteins in plant defense response. *Plant defence: Biol. Control*, 379–403.

Sumabat, L. G., Kemerait, R. C. Jr., and Brewer, M. T. (2018). Phylogenetic diversity and host specialization of Corynespora cassiicola responsible for emerging target spot disease of cotton and other crops in the Southeastern United States. *Phytopathology* 108 (7), 892–901. doi: 10.1094/PHYTO-12-17-0407-R

Sun, M., Jing, Y., Zhao, X., Teng, W., Qiu, L., Zheng, H., et al. (2020). Genome-wide association study of partial resistance to sclerotinia stem rot of cultivated soybean based

on the detached leaf method. *PloS One* 15 (5), e0233366. doi: 10.1371/journal.pone.0233366

Thapa, G., Gunupuru, L. R., Hehir, J. G., Kahla, A., Mullins, E., and Doohan, F. M. (2018). A pathogen-responsive leucine rich receptor like kinase contributes to Fusarium resistance in cereals. *Front. Plant Sci.* 9, 867. doi: 10.3389/fpls.2018.00867

Toruño, T. Y., Shen, M., Coaker, G., and Mackey, D. (2018). Regulated disorder: posttranslational modifications control the RIN4 plant immune signaling hub. *Mol. Plant-Microbe Interact.* 32 (1), 56–64. doi: 10.1094/MPMI-07-18-0212-FI

USDA (2021) *World agricultural production*. Available at: https://apps.fas.usda.gov/psdonline/circulars/production.pdf (Accessed August 2023).

USDA (2022). *World agricultural production foreign agriculture service*. Available at: https://apps.fas.usda.gov/psdonline (Accessed August 2023).

Usovsky, M., Lakhssassi, N., Patil, G. B., Vuong, T. D., Piya, S., Hewezi, T., et al. (2021). Dissecting nematode resistance regions in soybean revealed pleiotropic effect of soybean cyst and reniform nematode resistance genes. *Plant Genome* 14 (2), e20083. doi: 10.1002/tpg2.20083

Vincelli, P., and Hershman, D. E. (2011) *Assessing foliar diseases of corn, soybeans, and wheat* (UK cooperative extension service). Available at: https://plantpathology.ca.uky.edu/files/ppfs-misc-06.pdf (Accessed November 2011).

Vuong, T., Sonah, H., Meinhardt, C., Deshmukh, R., Kadam, S., Nelson, R., et al. (2015). Genetic architecture of cyst nematode resistance revealed by genome-wide association study in soybean. *BMC Genomics* 16 (1), 1–13. doi: 10.1186/s12864-015-1811-y

Wang, W., Chen, L., Fengler, K., Bolar, J., Llaca, V., Wang, X., et al. (2021). A giant NLR gene confers broad-spectrum resistance to *Phytophthora sojae* in soybean. *Nat. Commun.* 12 (1), 6263. doi: 10.1038/s41467-021-26554-8

Wang, W., Yuan, Y., Yang, C., Geng, S., Sun, Q., Long, L., et al. (2016). Characterization, expression, and functional analysis of a novel NAC gene associated with resistance to verticillium wilt and abiotic stress in cotton. *G3 (Bethesda)* 6 (12), 3951–3961. doi: 10.1534/g3.116.034512

Wang, J., and Zhang, Z. (2021). GAPIT Version 3: boosting power and accuracy for genomic association and prediction. *Genomics Proteomics Bioinf.* 19 (4), 629–640. doi: 10.1016/j.gpb.2021.08.005

Wang, X., Zhang, D., Cui, N., Yu, Y., Yu, G., and Fan, H. (2018). Transcriptome and miRNA analyses of the response to *Corynespora cassiicola* in cucumber. *Sci. Rep.* 8 (1), 7798. doi: 10.1038/s41598-018-26080-6

Waters, B. M., Amundsen, K., and Graef, G. (2018). Gene expression profiling of iron deficiency chlorosis sensitive and tolerant soybean indicates key roles for phenylpropanoids under alkalinity stress. *Front. Plant Sci.* 9, 10. doi: 10.3389/fpls.2018.00010

Wen, C., Mao, A., Dong, C., Liu, H., Yu, S., Guo, Y. D., et al. (2015). Fine genetic mapping of target leaf spot resistance gene cca-3 in cucumber, *Cucumis sativus L.* *Theor. Appl. Genet.* 128 (12), 2495–2506. doi: 10.1007/s00122-015-2604-z

Wen, Z., Tan, R., Zhang, S., Collins, P. J., Yuan, J., Du, W., et al. (2018). Integrating GWAS and gene expression data for functional characterization of resistance to white mould in soya bean. *Plant Biotechnol. J.* 16 (11), 1825–1835. doi: 10.1111/pbi.12918

Xavier, S. A., Canteri, M. G., Barros, D., and Godoy, C. V. (2013). Sensitivity of Corynespora cassiicola from soybean to carbendazim and prothioconazole. *Trop. Plant Pathol.* 38, 431–435. doi: 10.1590/S1982-56762013005000020

Xie, M., Chung, C. Y.-L., Li, M.-W., Wong, F.-L., Wang, X., Liu, A., et al. (2019). A reference-grade wild soybean genome. *Nat. Commun.* 10 (1), 1–12. doi: 10.1038/s41467-019-09142-9

Yuan, X., Wang, H., Cai, J., Li, D., and Song, F. (2019). NAC transcription factors in plant immunity. *Phytopathol. Res.* 1 (1), 3. doi: 10.1186/s42483-018-0008-0

Zabala, G., Zou, J., Tuteja, J., Gonzalez, D. O., Clough, S. J., and Vodkin, L. O. (2006). Transcriptome changes in the phenylpropanoid pathway of *glycine max* in response to *Pseudomonas syringae* infection. *BMC Plant Biol.* 6 (1), 1–18. doi: 10.1186/1471-2229-6-26

Zhang, J., Singh, A., Mueller, D. S., and Singh, A. K. (2015). Genome-wide association and epistasis studies unravel the genetic architecture of sudden death syndrome resistance in soybean. *Plant J.* 84 (6), 1124–1136. doi: 10.1111/tpj.13069

Zhang, Y., Wang, H.-L., Li, Z., and Guo, H. (2020). Genetic network between leaf senescence and plant immunity: crucial regulatory nodes and new insights. *Plants* 9 (4), 495. doi: 10.3390/plants9040495

Zhao, G., Guo, D., Wang, L., Li, H., Wang, C., and Guo, X. (2021). Functions of RPM1-interacting protein 4 in plant immunity. *Planta* 253 (1), 11. doi: 10.1007/s00425-020-03527-7