



Full-Length Transcriptional Analysis of the Same Soybean Genotype With Compatible and Incompatible Reactions to *Heterodera glycines* Reveals Nematode Infection Activating Plant Defense Response

Minghui Huang^{1†}, Ye Jiang^{1,2†}, Ruifeng Qin^{1,2}, Dan Jiang^{1,2}, Doudou Chang^{1,2}, Zhongyan Tian², Chunjie Li^{1*‡} and Congli Wang^{1*‡}

OPEN ACCESS

Edited by:

Madan K. Bhattacharyya,
Iowa State University, United States

Reviewed by:

Peijian Cao,
Zhengzhou Tobacco Research
Institute of CNTC, China
Yan Li,
Sichuan Agricultural University, China

*Correspondence:

Chunjie Li
lichunjie@iga.ac.cn
Congli Wang
wangcongli@iga.ac.cn

[†]These authors have contributed
equally to this work and share first
authorship

[‡]These authors share last authorship

Specialty section:

This article was submitted to
Technical Advances in Plant Science,
a section of the journal
Frontiers in Plant Science

Received: 31 January 2022

Accepted: 22 March 2022

Published: 18 May 2022

Citation:

Huang M, Jiang Y, Qin R, Jiang D,
Chang D, Tian Z, Li C and Wang C
(2022) Full-Length Transcriptional
Analysis of the Same Soybean
Genotype With Compatible
and Incompatible Reactions
to *Heterodera glycines* Reveals
Nematode Infection Activating Plant
Defense Response.
Front. Plant Sci. 13:866322.
doi: 10.3389/fpls.2022.866322

¹ Key Laboratory of Soybean Molecular Design Breeding, Northeast Institute of Geography and Agroecology, Chinese Academy of Sciences, Harbin, China, ² Heilongjiang Academy of Agricultural Sciences, Daqing, China

Full-length transcriptome sequencing with long reads is a powerful tool to analyze transcriptional and post-transcriptional events; however, it has not been applied on soybean (*Glycine max*). Here, a comparative full-length transcriptome analysis was performed on soybean genotype 09-138 infected with soybean cyst nematode (SCN, *Heterodera glycines*) race 4 (SCN4, incompatible reaction) and race 5 (SCN5, compatible reaction) using Oxford Nanopore Technology. Each of 9 full-length samples collected 8 days post inoculation with/without nematodes generated an average of 6.1 GB of clean data and a total of 65,038 transcript sequences. After redundant transcripts were removed, 1,117 novel genes and 41,096 novel transcripts were identified. By analyzing the sequence structure of the novel transcripts, a total of 28,759 complete open reading frame (ORF) sequences, 5,337 transcription factors, 288 long non-coding RNAs, and 40,090 novel transcripts with function annotation were predicted. Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) enrichment analyses of differentially expressed genes (DEGs) revealed that growth hormone, auxin-activated signaling pathway and multidimensional cell growth, and phenylpropanoid biosynthesis pathway were enriched by infection with both nematode races. More DEGs associated with stress response elements, plant-hormone signaling transduction pathway, and plant–pathogen interaction pathway with more upregulation were found in the incompatible reaction with SCN4 infection, and more DEGs with more upregulation involved in cell wall modification and carbohydrate bioprocess were detected in the compatible reaction with SCN5 infection when compared with each other. Among them, overlapping DEGs with a quantitative difference was triggered. The combination of protein–protein interaction with DEGs for the first time indicated that nematode infection activated the interactions between transcription factor WRKY and VQ (valine-glutamine motif) to contribute to soybean defense. The knowledge of the SCN–soybean interaction mechanism as a model will present more understanding of other plant–nematode interactions.

Keywords: full-length transcriptome sequencing, *glycine max*, *Heterodera glycines*, incompatible and compatible response, VQ–WRKY interaction

INTRODUCTION

Soybean cyst nematode (SCN, *Heterodera glycines* Ichinohe) is one of the most economically important diseases in soybean (*Glycine max* L. Merrill) worldwide. The latest statistical analysis of estimated cumulative soybean economic losses due to diseases in 28 states within the United States from 1996 to 2016 indicated that SCN accounted for 23.2% of the total losses (US\$ 73,535 per hectare) which was top-ranked, far greater than the second disease charcoal rot (10%) (Bandara et al., 2020). In China, annual economic losses in soybean could reach up to US\$ 120 million (Li Y. et al., 2011). SCN is a soil-borne sedentary parasitic nematode in the major host soybean. In the only free-moving stage of SCN, a second-stage juvenile (J2) hatches from an egg (first-stage juvenile inside the egg) to search for host roots through signals released from plant roots, moves to root tips, and penetrates roots using a stylet. The J2 moves inside the root, establishes a feeding site, and reprograms host root cells by directing gland secretions into plants that can dissolve cell walls and fuse the protoplast of neighboring cells, and eventually forms a unique feeding structure called syncytium as a nutrient source for nematode development, consequently suppressing plant growth and affecting yield (Niblack et al., 2006; Mitchum and Baum, 2008).

Host-plant resistance combined with crop rotation is the most effective way to control SCN. More than 300 quantitative trait loci (QTLs) associated with SCN resistance were mapped to 20 chromosomes (chr) (soybase.org), but only two major resistance genes, *rhg1* on chromosome (chr) 18 and *Rhg4* on chr 8, were cloned and characterized (Cook et al., 2012, 2014; Liu et al., 2012, 2017). The presence of multiple minor QTLs contributing to resistance in either resistant or susceptible soybean plants makes resistance breeding more difficult than expected (Huang et al., 2021). Furthermore, shifts in virulence have caused a decrease or loss of resistance because of the long-term planting of a single source of resistance varieties, e.g., 90% of resistance sources derived from the PI88788 background in the United States and mainly Peking resistance sources in China (Mitchum et al., 2007; Niblack et al., 2008; Acharya et al., 2016; Hua et al., 2018; Huang et al., 2022). Furthermore, Peking and PI88788 display resistance only to some SCN races or HG Types. The lack of broad resistance sources and the presence of multiple SCN races or HG types in the field result in SCN spreading widely and quickly. Thus, understanding the molecular mechanisms of SCN infection and plant resistance will gain more insights to develop new control strategies, including engineering important candidate genes to increase resistance.

Plants have evolved to develop two layers of pathogen defense immune systems: the first layer is pathogen- or microbe-associated molecular pattern (PAMP/MAMP) triggered immunity (PTI); the second layer is effector-triggered immunity (ETI), which fits species-specific disease resistance (Dangl and Jones, 2001; Eitas and Dangl, 2010; Monaghan and Zipfel, 2012). The plant cell wall surface contains pattern recognition receptors (PRRs) that can detect pathogen or microbe structures to activate PTI, and phytohormones as defense-related signaling

molecules, such as salicylic acid (SA), jasmonate acid (JA), ethylene (ET), induce plant to produce pathogenesis-related (PR) proteins, e.g., β -1,3-glucanases, peroxidase, oxidase-like, thionin, and proteinase inhibitor, etc. (Sels et al., 2008). ETI is initiated by intracellular nucleotide-binding protein domain leucine-rich repeat proteins (NLRs) to generate a hypersensitive response (HR) with local cell death (Eitas and Dangl, 2010). PTI and ETI are activated by two distinct classes of receptors in early signaling, but recent evidence suggests that PTI and ETI can crosstalk in downstream response, although how they contribute to immunity with quantitative and/or qualitative outputs is still undefined (Naveed et al., 2020; Yuan et al., 2021). Transcriptome analysis of different pathosystems denotes that both compatible and incompatible interactions are able to trigger an overlapping change of gene expression but with quantitative differences (Mine et al., 2018; Yuan et al., 2021).

Studies have been performed on compatible and incompatible interactions between SCN and soybean roots based on microarray and RNA-seq transcriptome analyses (Ithal et al., 2007; Klink et al., 2007; Puthoff et al., 2007; Klink and Matthews, 2009; Kandoth et al., 2011; Li X. et al., 2011; Li et al., 2012, 2018; Mazarei et al., 2011; Wan et al., 2015; Zhang et al., 2017; Kang et al., 2018; Neupane et al., 2019; Song et al., 2019; Jiang et al., 2020; Miraeiz et al., 2020). All sequence annotations revealed that SCN infection can induce or suppress gene expression either in susceptible or resistant cultivars, and a series of defense genes (*PPRs* and *NLRs*), *MAPK* (mitogen-activated protein kinase) signaling cascade, *WRKY* and *MYB* transcription factors (TFs), heat shock protein (*HSP*) genes, *PR* genes, and phenylpropanoid metabolism genes have been identified but with variance depending on SCN race/HG type and plant type.

RNA-sequencing (RNA-seq) based on high-throughput next-generation sequencing (NGS) (e.g., Illumina) has been used widely to measure differential gene expression because it is a cost-effective and advanced technology (Finotello and Di Camillo, 2015). However, RNA-seq requires fragmentation of RNA or cDNA to generate short reads when preparing samples, which diminishes the information from original full-length transcripts; thus, it is harder to obtain post/co-transcriptional processing events that are responsible for producing a mature RNA molecule that can leave the nucleus to function in the cell by chemical structure alteration of the RNA primary transcript (Kiss, 2001). Advanced full-length transcriptome sequencing with longer reads circumvents these challenges. Currently, PacBio and Oxford Nanopore (ONT) are the most popular 3rd-generation full-length sequencing technologies that can provide more complex transcription and reveal the real structure of sequences during transcription, such as alternative splicing (AS), alternative polyadenylation (APA), and long non-coding RNA (lncRNA) and gene fusion, which can increase the complexity of the transcriptome and proteome. Compared with PacBio sequencing, ONT uses ion current blockades to directly sequence more long-native DNA or full-length RNA molecules (Cui et al., 2020; Xie et al., 2021). AS, one of the important steps in post-transcriptional modification, can recognize and eliminate intronic regions of a precursor messenger RNA (pre-mRNA) to generate multiple mRNAs to regulate gene expression that consequently promotes

proteome diversity. AS plays key roles not only in plant growth and development but also in response to biotic or abiotic stimuli or adaptation (Matsukura et al., 2010; Severing et al., 2012; Syed et al., 2012; Mandadi and Scholthof, 2015; Wang et al., 2018; Bedre et al., 2019; Martín et al., 2021). APA can produce multiple mRNA polyadenylation isoforms through pre-mRNA endonucleolytic cleavage and poly(A) tail addition at the 3' cleavage site end of a nascent transcript to change the length of untranslated regions (UTRs) or coding regions that may influence mRNA stability, translation efficiency, subcellular localization, or gene function gain or loss. Consequently, all these changes will result in various plant physiological and biochemical processes (Yeh and Yong, 2016; Sadek et al., 2019; Zhang et al., 2020; Tu et al., 2021); for example, APA participates in cell wall modification, root hair development, DNA repair, and gene regulation in response to abiotic and biotic stresses (Cao et al., 2019; Ye et al., 2019; Yan et al., 2021). LncRNAs longer than 200 nucleotides are epigenetic regulators that regulate gene expression by interacting with mRNAs, DNAs, proteins, and miRNAs to participate in biological processes such as plant growth and development and biotic and abiotic stress responses (Budak et al., 2020; Yu et al., 2020; Tu et al., 2021; Urquiaga et al., 2021).

The application of the full-length transcriptome sequencing technique in plants is limited compared with second-generation sequencing because of its higher cost. Currently, full-length transcriptome sequencing for inferring and improving gene models and identifying novel genes has been reported on rice, wheat, maize, cotton, pecan, poplar, and others but not on soybean (Clavijo et al., 2017; Wang et al., 2018; Zhang et al., 2019; Zhao et al., 2019; Li C. et al., 2020; Yang et al., 2021).

Our previous study has demonstrated that the soybean breeding line 09-138, developed in northeast China, carries the *rhg1-a* and *Rhg4-b* loci and has resistance to SCN race 4 (SCN4, HG type 1.2.3.5.6.7) but has susceptibility to SCN race 5 (SCN5, HG type 2.5.7), which is different from the Peking (*rhg1-a* + *Rhg4-a*) and PI88788 (*rhg1-b* + *Rhg4-b*) resistance backgrounds (Hua et al., 2018; Huang et al., 2022). We postulated that differences in response to the two SCN races should be related to distinct transcriptional responses in the early stages of nematode infection as described above. To test this, line 09-138 was inoculated with the two SCN races, and nematode development was observed inside the roots. Comparative full-length transcriptome analysis of soybean breeding lines 09-138 infected with SCN4 (resistance response) and SCN5 (susceptibility response) was performed using the ONT technology. We investigated AS and APA events and identified lncRNA and transcription factors in detected novel transcripts. Then, differentially expressed genes (DEGs) and transcripts (DETs) were analyzed, and DEGs associated with stress response elements were explored. Enriched DEG-GO (Gene Ontology) terms and DEG-KEGG (Kyoto Encyclopedia of Genes and Genomes) pathways were compared between resistant and susceptible responses. Protein–protein interactions were predicted, and a defense mode was established. Finally, the expression of DEGs was validated by quantitative RT-PCR. Comparisons between compatible and incompatible responses

will provide insight into the resistance mechanism and identify candidate defense or resistance genes for further study.

MATERIALS AND METHODS

Plant Materials and Nematode Culture

The soybean breeding line 09-138 was developed by the Heilongjiang Academy of Agricultural Sciences (Hua et al., 2018). SCN race 4 (SCN4, HG type 1.2.3.5.6.7) and SCN race 5 (SCN5, HG type 2.5.7) were originally collected from the field and cultured from single cysts for more than 5 generations on the susceptible soybean variety Dongsheng1 in a greenhouse with 16 h of light and 8 h of darkness at 23–28°C, and then identified by race test and HG type indicator lines (Hua et al., 2018). Every other year, nematode HG types/races were reconfirmed without virulence change with time.

A nematode inoculum was prepared according to the method described by Huang et al. (2022). Plant root tissue and soil were collected 35–40 days after inoculation and put into a 2 L-beaker. The mixture was stirred vigorously with a glass rod for approximately 1 min and then precipitated for 10 s. The fluid of the supernatant was gently poured into 75/25 µm nested sieves. The mixture of cysts and root debris on the top of the sieves was rubbed with a rubber stopper to release eggs. Eggs on 25-µm sieve were rinsed with a high-pressure water faucet for 1 min and then with sterile water before collection. The collected eggs were then transferred onto six to eight layers of tissue paper supported by a metal screen on a hatching dish containing 3 mM ZnSO₄ in sterile water for hatch at 28°C. J2s were then collected for inoculation after 3–4 days.

Nematode Inoculation, Root Staining, and Root Preparation for RNA Extraction

Seeds of 09-138 were sterilized by soaking in 0.5% sodium hypochlorite for 20 min and rinsed with sterile water three times. Two seeds were sown in a black plastic pot (8 cm diameter × 12 cm depth) filled with autoclaved soil and sand at a ratio of 1:1. After 4 days, two seedlings were thinned to one in each pot. An eight-day seedling was inoculated with a 1-ml suspension containing 2,000 J2 of SCN4 or SCN5. Seedlings were inoculated with 1 ml water as control. The plants were maintained in a growth chamber at a 16-/8-h day/night regime, 28°C day/22°C night, and 50% relative humidity.

Collected roots were stained 3, 6, 8, 10, and 12 days after inoculation with acid fuchsin (Byrd et al., 1983). Nematode development inside the roots was observed, and roots and nematodes were photographed under an Olympus SZX16 dissecting microscope using the Cellsens Standard image software (Olympus Corporation, Japan).

To collect roots for RNA extraction, plant roots 8 days after inoculation were washed and rinsed thoroughly with water. Three roots from each treatment were wrapped together with aluminum foil as one replication (one sample). Three replications were made for each treatment. A total of 9 samples with SCN4- and SCN5-infection and control were collected. Immediately, each prepared sample was put in liquid nitrogen to freeze it and was kept at -80°C for RNA extraction and sequencing.

RNA Extraction, cDNA Library Construction, and Nanopore Sequencing

Total RNA was extracted using RNAeasy Plant Mini Kit (Qiagen, United States), and RNase-free DNase (Qiagen) was used to remove DNA contamination in the total RNA. The concentration, purity, and integrity of the extracted RNA were measured with 1% agar gel (Thermo Fisher Scientific, United States) and Agilent 2100 Bioanalyzer (Agilent Technologies, United States). cDNA library construction started with 1 µg total RNA using a cDNA-PCR sequencing kit (SQK-PCS109) provided by Oxford Nanopore Technologies (ONT, Inc., United Kingdom) following the manufacturer's instructions. Final cDNA libraries were added to FLO-MIN109 flow-cells and run on the PromethION platform at Biomarker Technology Company (Beijing, China) for sequencing. Experimental processes including sample quality testing, library building, library-quality testing, and library sequencing were performed in accordance with standard procedures provided by ONT.

Raw Data Processing to Obtain Full-Length Transcriptome

Raw reads were subject to filtering with an average read quality score ≤ 7 , and read length ≤ 500 bp, and ribosomal RNA mapped to the rRNA database was also discarded. Primer sequences on both ends of clean reads were searched to determine full-length, non-chimeric (FLNC) sequences. Detected FLNC transcripts were then mapped to soybean reference genome Williams 82.a2.v1¹ with minimap2 (Li, 2018) to obtain FLNC clusters, and pinfish² was applied to polish each cluster to attain consensus isoforms. All mapped reads were further collapsed using the cDNA_Cupcake package with a mini-identity of 90% and mini-coverage of 85%. When the redundant transcripts were collapsed, 5' difference was not considered. The achieved transcripts were compared with known transcripts of reference genome Williams 82.a2.v1 utilizing gffcompare, and novel transcripts were identified in order to make supplementary genome annotation. Gene boundaries were modified, and transcripts with expressed levels ≤ 1 were filtered.

Structure Analysis: Identification of Alternative Polyadenylation, Fusion Transcript, Alternative Splicing Events, and Microsatellite Markers

Alternative polyadenylation was identified through further analysis of FLNC by Transcriptome Analysis Pipeline from Isoform Sequencing (TAPIS) (Foissac and Sammeth, 2007). Multiple Expectation Maximization for Motif Elicitation (MEME) (Bailey et al., 2006) was used to analyze the 50-bp sequence upstream of the poly A site to detect FL motifs. The consensus sequence before remove-redundant analysis was used for fusion transcript analysis. Fusion transcript was defined under the following conditions: aligned to 2 or more sites; each site covers at least

5% of a transcript with a 1-bp minimum alignment length; the total length covers more than 95% of the total length of transcripts with at least 10k bp distance between the two sites.

Alternative splicing indicates the process of pre-mRNA treatment. Gene transcription generates pre-mRNAs with many splicing methods. Five types of alternative splicing events (3' splice site, 5' splice site, exon skipping, intron retention, and mutually exclusive exon) of transcripts were examined by employing the Astalavista software based on the alignment results of individual samples to the reference genome (Foissac and Sammeth, 2007). The MicroSatellite (MISA, an identification tool) software was used for SSR analysis, and transcripts below 500 bp were discarded.

Coding Sequence Prediction, Long Non-coding RNA Identification, and Transcription Factor Detection From Novel Transcripts

Coding sequences (CDSs) were predicted with TransDecoder (v3.0.0; Haas et al., 2013) based on the ORF. LncRNA does not code for protein. Therefore, LncRNA in novel transcripts was predicted whether it had a coding potential by protein domain analysis including all four methods, Coding Potential Calculator (CPC) (Kong et al., 2007), Coding-Non-Coding Index (CNCI) (Sun et al., 2013), Coding Potential Assessment Tool (CPAT) (Wang et al., 2013) and Protein family (Pfam) (Finn et al., 2014). LncRNA target genes were predicted using two methods: first, depending on the location relationship between the differentially expressed lncRNA and adjacent mRNA (within 100k bp distance) expressed differentially; second, according to complementary base pairing between lncRNA and mRNA using the lncTAR tool (Li et al., 2015). TFs were detected with iTAK (Zheng et al., 2016).

Quantification of Transcript/Gene Expression Levels and Differential Expression Analysis

Mapped full-length reads with > 5 match quality were chosen for quantification. Transcript or gene expression levels were measured in counts per million (CPM) (Zhou et al., 2014) and calculated by the following:

$$\text{CPM} = (\text{read number matched the transcript}) / (\text{total read number matched referenced transcriptome}) \times 10^6$$

The differential expression among the treatments was analyzed with DESeq2 (Anders and Huber, 2010) depending on a negative binary distribution model, consequently gaining DEGs or DETs. False discovery rate (FDR) was adjusted and controlled with the method of Benjamini and Hochberg (1995), and DEGs or DETs with \log_2 fold change (FC) ≥ 2 and FDR < 0.01 were chosen. A heat map for DEGs in each group was developed using the pheatmap package in R (Version 1.0.12³).

¹<https://phytozome.jgi.doe.gov/>

²<https://github.com/nanoporetech/pinfish>

³<https://CRAN.R-project.org/package=pheatmap>

Functional Annotation and Enrichment Analysis of Differentially Expressed Genes/Transcripts

Functional annotation of the genes/transcripts was conducted by blasting with databases including NR (NCBI non-redundant protein sequences) (Deng et al., 2006), Swissprot (Apweiler et al., 2004), GO (Ashburner et al., 2000), Clusters of Orthologous Groups (COG) (Tatusov et al., 2000), euKaryotic Ortholog Groups (KOG) (Koonin et al., 2004), Pfam (Kanehisa et al., 2004), and KEGG (Mckenna et al., 2010).

Gene ontology enrichment analysis of DEGs or DETs was conducted using the Goseq R package-based Wallenius non-central hypergeometric distribution (Young et al., 2010). KEGG pathway enrichment analysis of DEGs or DETs was subject to the KEGG Orthology Based Annotation System (KOBAS) software (Mao et al., 2005). Protein–protein interactions (PPIs) for all detected DEGs were predicted using the STRING database⁴ and were visualized in Cytoscape (Shannon et al., 2003).

Validation of Differentially Expressed Genes by Real-Time Quantitative Reverse Transcription-PCR

Samples leftover from full-length transcriptome sequencing was subject to qRT-PCR for verification of DEGs. Primers were designed by using the Primer Premier 5 software (Lalitha, 2000) and synthesized by Comate Bioscience Company Limited (Changchun, China). A total of 1 µg of treated RNA was used to synthesize the first-strand cDNA using FastKing gDNA Dispelling RT SuperMix (TIANGEN, China). PCR reactions were conducted in LightCycler[®] 480 System (Roche Life Science, United States) with ChamQ Universal SYBR qPCR Master Mix (Vazyme Biotech Co., Ltd., Nanjing, China) following the manufacturer's protocol. PCR was carried out in a 20-µl volume containing 100 ng cDNA (2 µl). PCR was performed as follows: initial denaturing for 10 min at 95°C, followed by 40 two-step cycles of 95°C for 10 s and then at 60°C for 1 min. The relative expression of tested genes was calculated with the $-\Delta\Delta C_t$ method using ACTIN as a control. Three independent biological replicates and three technical repetitions were performed for all the experiments. The primers used are listed in **Supplementary Table 1**. The correlation between transcriptome sequencing and qRT-PCR was completed in Excel 2016, and independent-samples *t*-test was performed using SPSS 17.0.

RESULTS

Differential Development of *Heterodera glycines* Race 4 and Race 5 Inside Roots of Lines 09-138 in Early Stages

Various development stages of the two races, SCN4 (**Figure 1A**) and SCN5 (**Figure 1B**), were observed in roots. Soybean roots were penetrated by both nematode races, and there was no

obvious difference in nematode size on day 3 days (**Figure 1**). On day 6, most of SCN5 developed to the J3 stage but SCN4 remained in the J2 stage. Nematodes had developed from J3 to late J4 stages in roots infected with SCN5 at 8, 10, and 12 days (**Figure 1B**) when compared with J2, J3, or a few early J4 stages in roots infected with SCN4 (**Figure 1A**), confirming 09-138 is resistant to SCN4 (cyst number/plant, $13 \pm SE 2.7$; female index, FI = 10) and susceptible to SCN5 (cyst number/plant, $119 \pm SE 8.16$; FI = 40) (Huang et al., 2022). More brown spots (hypersensitive response) around some nematode feeding sites were observed in resistant roots with SCN4 than in those with SCN5 (**Figure 1A**).

Nanopore Full-Length Transcriptome Sequence Statistics and Redundant Remove of Transcript

Since an obvious difference in nematode development on day 8 was observed in 09-138, root samples from day 8 were used for full-length sequencing. An average of 6.1 Gbp of clean data for nine cDNA libraries was obtained with a range from 5.73 to 6.82 Gbp (**Supplementary Table 2**). The average N50 length was 1,287 bp with a mean length of 1,142 bp, an average max length of 11,276 bp (9,575–13,426 bp), and a mean quality value of 11 (Q11) (**Supplementary Table 2**). After rRNA was filtered out, an average of 5,296,377 clean reads (4,681,541–5,888,283) and an average of 4,258,467 FLNC read numbers (3,767,343–4,641,061) were generated with an average of 80.4% FLNC ratio (**Table 1**).

Finally, 65,038 redundant-removed transcript sequences containing 92,079,818 bp with an N50 length of 1,665 bp, mean length of 1,415 bp, and max-length of 7,285 bp were obtained through a merging of the consistent sequences (**Supplementary Figure 1**). After alignment with the reference genome and filtering with known annotations of the reference genome, 1,117 novel genes and 41,096 novel transcripts were identified. The consistent sequence of each sample was used for AS analysis.

Structure Analysis of Alternative Polyadenylation Events, Transcript Fusion, Alternative Splicing Events, and SSR Prediction Revealed Soybean Structure Variation in Response to *Heterodera glycines* Race 4 and Race 5 Infections

Full-length sequencing can accurately identify the structure of transcripts. APA analysis based on FLNC displayed a total of 214,760 transcripts with various numbers of poly A sites. Among the three treatments, the greatest portion of 23.8% transcripts contained > 5 poly A sites, followed by 1 poly A site (21.5%), and the least with 5 poly A sites (9.6%) (**Figure 2A**). The number of transcripts in each poly A site distribution showed no significant difference ($P > 0.05$) among the three treatments (**Figure 2B**). In a more dissecting way, we found that SCN4 infection made soybean generate a significantly greater ($P < 0.05$) mean number ($840 \pm SE 26$) of >10 poly A sites than SCN5 infection ($734 \pm SE 12$) and the control ($698 \pm SE 26$), indicating APA with more poly A sites might be involved in the

⁴<http://string-db.org>

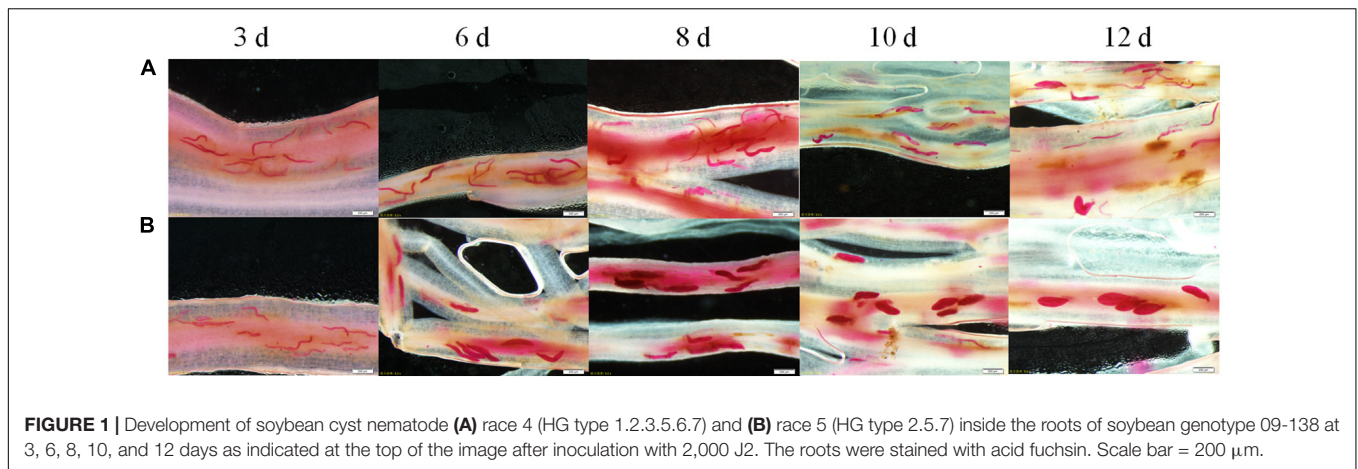


FIGURE 1 | Development of soybean cyst nematode (A) race 4 (HG type 1.2.3.5.6.7) and (B) race 5 (HG type 2.5.7) inside the roots of soybean genotype 09-138 at 3, 6, 8, 10, and 12 days as indicated at the top of the image after inoculation with 2,000 J2. The roots were stained with acid fuchsin. Scale bar = 200 μ m.

TABLE 1 | Number of clean reads and full-length reads, and full-length percentage.

Treatment ^a	Number of clean reads (except rRNA) ^b	Number of full-length reads ^c	Full-length percentage(%) ^d
CK-1	5,592,903	4,437,097	79.33
CK-2	5,683,524	4,641,061	81.66
CK-3	5,399,629	4,311,944	79.86
SCN4-1	5,459,017	4,409,582	80.78
SCN4-2	5,888,283	4,631,268	78.65
SCN4-3	5,184,003	4,146,135	79.98
SCN5-1	4,727,932	3,785,053	80.06
SCN5-2	5,050,558	4,196,724	83.09
SCN5-3	4,681,541	3,767,343	80.47
Mean	5,296,377	4,258,467	80.43

^aCK-1, CK-2, and CK-3 represent control with water for three biological replications on soybean breeding line 09-138; Similarly, SCN4-1, SCN4-2, and SCN4-3 represent root samples infected with soybean cyst nematode race 4 (HG type 1.2.3.5.6.7) for three replications; and SCN5-1, SCN5-2 and SCN5-3 represent SCN race 5 for three replications. ^bNumber of clean reads (except rRNA): number of clean read sequences after filtering rRNA. ^cNumber of full-length reads: number of full-length sequences. ^dFull-length percentage (%): percentage of number of full-length reads compared with clean reads.

incompatible reaction. Enrichment in T and A were detected in the upstream and downstream of the 50-bp position, respectively (Figure 2C). Three GC-rich APA motifs (CAGGGG, GGCTGC, and GGCCGC) were identified in the 50-bp sequence upstream of the poly A sites (Figure 2D). Fusion transcript screening before the remove-redundant analysis revealed 2-12 fusion transcripts for each sample. Similarly, SCN4 infection generated a total of 21 fusion transcripts, while only 14 fusion transcripts were found for SCN5 and the control (Supplementary Table 3), suggesting that fusion transcripts might be participating in defense response.

The Astaravista analysis of novel transcripts demonstrated that all the three treatments had no significant difference in the average number of AS events, $1,925 \pm \text{SE } 54.6$, $2,265 \pm \text{SE } 145.8$; and $2,287 \pm \text{SE } 130.1$ for the control, SCN4, and SCN5 treatments, respectively (Figure 2E), but that the AS events varied among the treatments. Based on the average percentage of AS events among the 9 libraries, there were 40.11% intron retentions, 26.28% alternative 3' splice sites, 18.81% exon skipping, 14.34% alternative 5' splice sites, and 0.46% mutually exclusive exons (Figure 2F). The numbers of AS events at alternative 3' splice sites in both nematode-infected treatments were significantly lower ($P < 0.05$) than those in the control, while the numbers

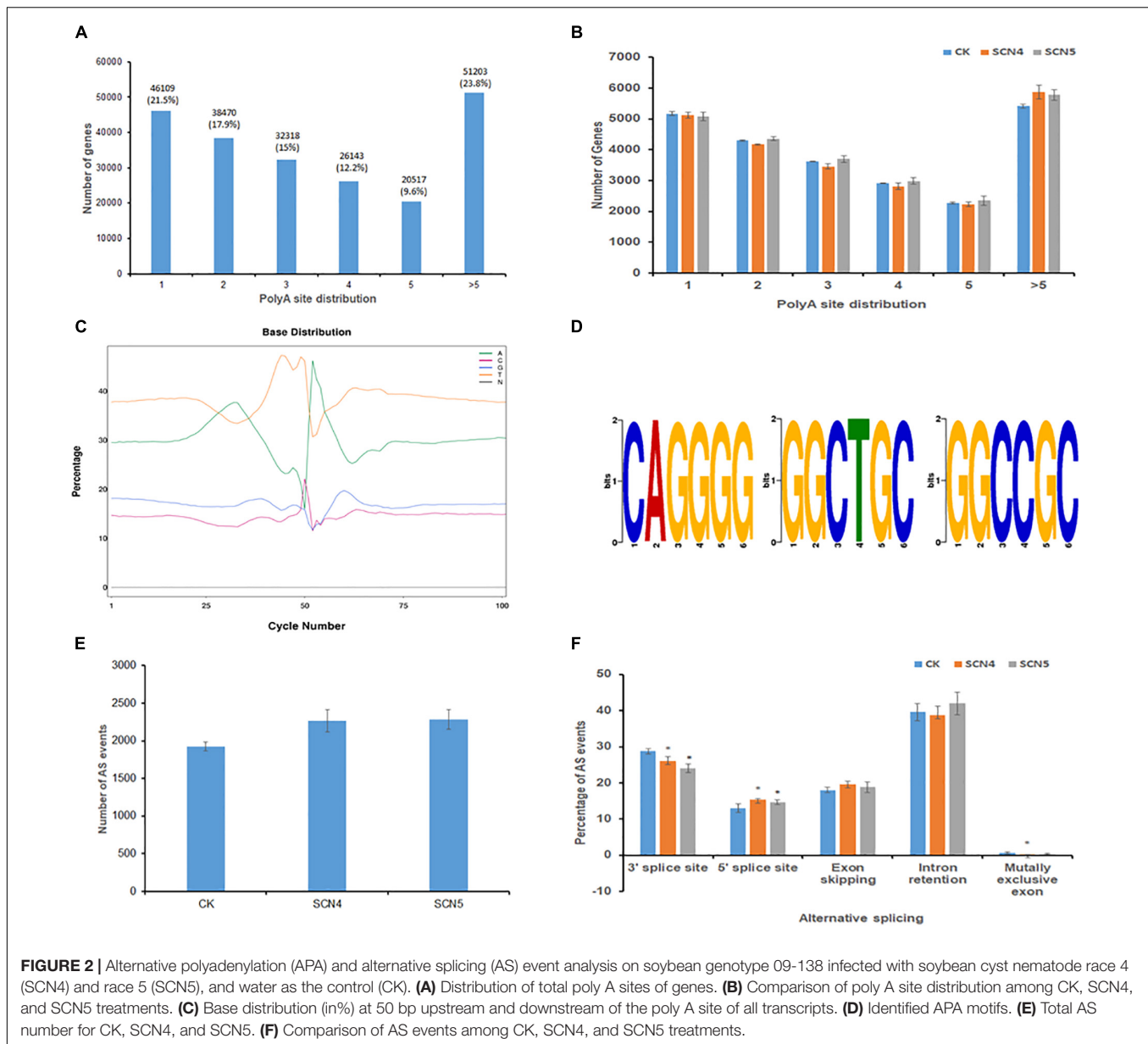
at 5' splice sites were significantly greater ($P < 0.05$) than those in the control.

A total of 22,486 SSRs were identified from 61,900 novel transcripts containing a total of 90,765,603 bp (Supplementary Table 4). Three types of SSR with mono-, di-, and trinucleotide accounted for 98% of total SSRs. There were 1,718 SSRs present in compound formation (hybrid microsatellite, a distance of two SSRs < 100 bp) (Supplementary Figure 2 and Supplementary Table 4).

Thus, all the tested structure variations among the control, SCN4, and SCN5 treatments demonstrated that post-transcriptional modification might be involved in nematode resistance or susceptibility in soybean.

Coding Sequences Prediction of Novel Transcripts

A total number of 37,469 ORFs were obtained, including 28,759 complete ORFs. Approximately, 38 and 43% of all predicted CDS coding protein lengths were within 0-100 and 100-200 aa, respectively (Figure 3A). The corresponding complete CDS protein lengths are shown in Figure 3A.



Long Non-coding RNA Associated With Defense Responsive Transcription Factors

Long Non-coding RNAs associated with plant growth, development, and stress responses have garnered widespread attention (Szcześniak et al., 2015; Urquiaga et al., 2021). In total, 288 lncRNAs were predicted by CPC, CNC, CPAT, and Pfam protein domain analyses; 24 and 90 of them were unique to the CPAT and CPC methods, respectively (**Figure 3B**); 47.6 (137) and 42.4% (122) of them were classified as sense_lincRNA and lincRNA (long intergenic non-coding RNA), respectively (**Figure 3C**). Antisense-lincRNA and intronic-lincRNA accounted for 7.3 (21) and 2.8% (8), respectively (**Figure 3C**). The number of cis-targeted genes (275) regulated by these lncRNAs was

nearly four times greater than those (85) of trans-targeted genes. A lncRNA-Pfam protein domain analysis exhibited 164 transcripts matched to the Pfam database, including defense-responsive transcription factors WRKY, TIR, EF-hand, AUX/IAA, zf-RVT (zinc finger in the reverse transcript), zf-CCHC, BoA (bacterial stress-induced morphogen protein) and others, suggesting that lncRNAs may be involved in nematode stress response in soybean. Interestingly, 69 of 164 (42%) Pfam domains were extensin-like protein repeats from two novel transcripts, ONT.12398 (chr 9) and ONT. 12400 (chr 9), and 3 domains were extensin-like regions in ONT.14325 (chr 10). Extensins are a family of hydroxyproline-rich glycoproteins (HRGPs) in plant cell walls that can mediate resistance to viruses, bacteria, fungi, and nematodes (Deepak et al., 2010; Hirao et al., 2012). Especially, extensins identified in SCN syncytium

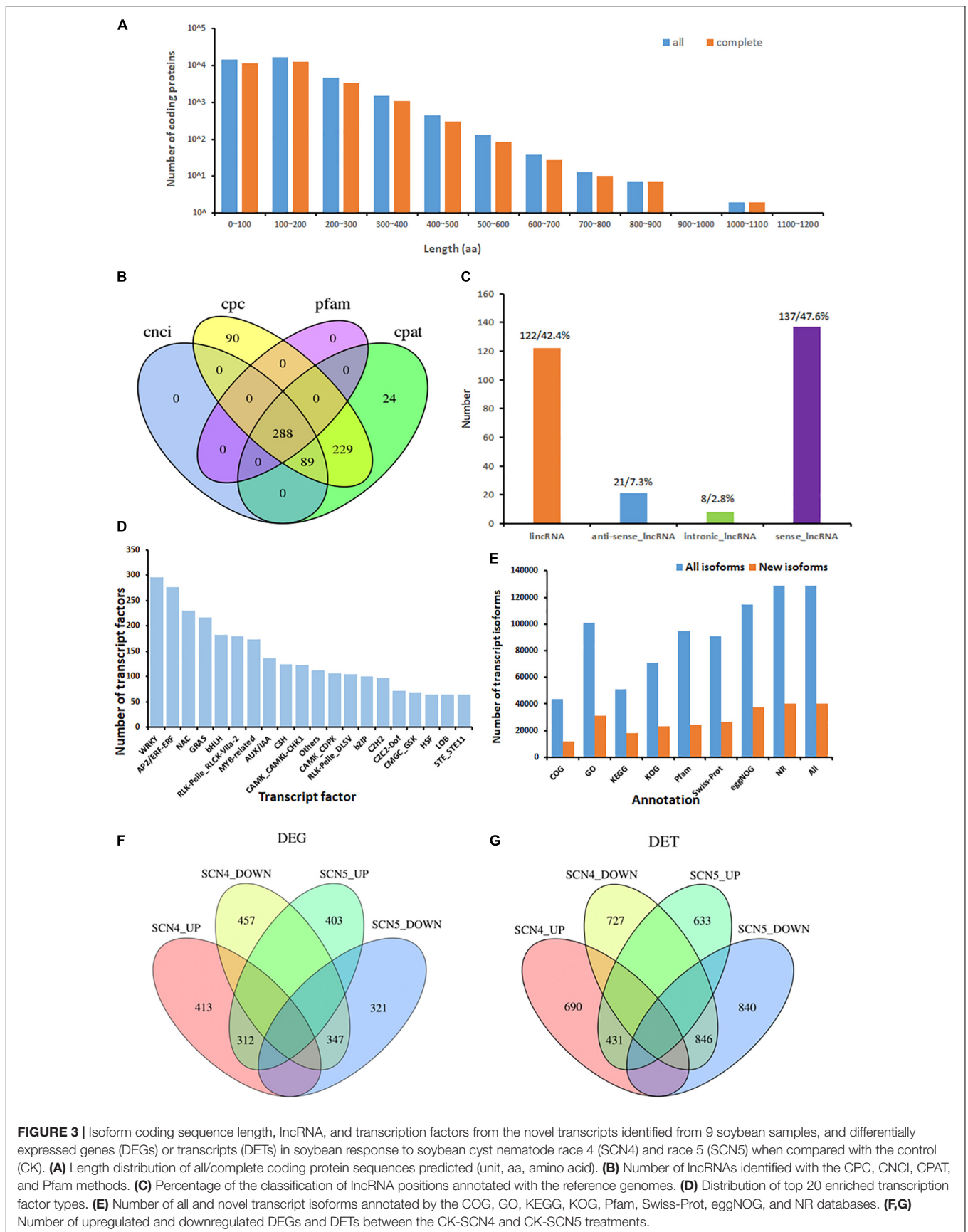


FIGURE 3 | Isoform coding sequence length, lncRNA, and transcription factors from the novel transcripts identified from 9 soybean samples, and differentially expressed genes (DEGs) or transcripts (DETs) in soybean response to soybean cyst nematode race 4 (SCN4) and race 5 (SCN5) when compared with the control (CK). **(A)** Length distribution of all/complete coding protein sequences predicted (unit, aa, amino acid). **(B)** Number of lncRNAs identified with the CPC, CNCI, CPAT, and Pfam methods. **(C)** Percentage of the classification of lncRNA positions annotated with the reference genomes. **(D)** Distribution of top 20 enriched transcription factor types. **(E)** Number of all and novel transcript isoforms annotated by the COG, GO, KEGG, KOG, Pfam, Swiss-Prot, eggNOG, and NR databases. **(F,G)** Number of upregulated and downregulated DEGs and DETs annotated by the CK-SCN4 and CK-SCN5 treatments.

formation (Ithal et al., 2007; Matsye et al., 2011) and nematode-related lncRNA co-expression with HRGP potentially mediating SCN infection (Khoei et al., 2021) consistently confirmed that altered cell wall composition with extensins may play roles in disease defense.

Transcription Factor Prediction and Function Analysis of the Novel Transcripts

There were 5,337 TFs predicted, including 176 families of TFs (transcription factor, transcription regulator, and protein kinase) with a number ranging from 1 to 296 (**Supplementary Table 5**). The three treatments, the control, SCN4 and SCN5, contained 1,804, 1,787, and 1,746 TFs, respectively. The most abundant TFs were WRKY (296), AP2/ERF-ERF (APETALA2/Ethylene Responsive Factors, 276), NAC (NAM, ATAF, and CUC, 230), and GRAS (GAI, RGA, and SCA, 217) (**Figure 3D**).

Function annotation of transcripts obtained from the alternative splicing analysis exhibited 128,648 known and 40,090 novel transcript isoforms (**Supplementary Table 6**) and 56,865 known and 859 novel genes (**Supplementary Table 7**). The number of annotated transcripts in all and new isoforms is displayed in **Figure 3E**. Among them, NR annotation of species distribution indicated that 86.8, 8.9, 0.95, and 0.4% of the transcripts were aligned to *G. max*, *G. soja*, *Phaseolus vulgaris*, and *Medicago truncatula*, respectively. The GO-enrichment analysis showed that 13,483 out of 100,994 (13.5%) *G. max* isoforms and 4,350 out of 31,102 (14.0%) novel isoforms were associated with response to stimulus (**Supplementary Figures 3A,B**).

Transcript Expression, Differentially Expressed Genes, and Differentially Expressed Transcripts

The CPM density distribution of transcript expressions of samples is displayed in **Supplementary Figure 4A**, and the dispersion degree of expression level distribution in a sample is shown with a boxplot in **Supplementary Figure 4B**. The evaluation of the Pearson correlation coefficient among the biological replicates indicated that the *r* range was from 0.682 to 0.963 (**Supplementary Figure 4C**). A clear separation of the control treatment from the SCN4- and SCN5- infected treatments is shown in principal component analysis (PCA) (**Supplementary Figure 4D**).

In total, 2,255 DEGs and 4,167 DETs were identified for all three comparisons; the upregulated and downregulated DEGs or DETs are listed in **Table 2**, and volcano plots for the DEGs and DETs are shown in **Supplementary Figure 5**. There were no overlapping DEGs (**Figure 3F**) or DETs (**Figure 3G**) found between CK-SCN4-up and CK-SCN5-down and vice versa. The annotated numbers of DEGs and DETs with functional annotation database are listed in **Supplementary Table 8**. Only 7 DEGs were found between the SCN4- and SCN5- treatments, including 3 DEGs identified in CK-SCN5, *Glyma.04G061500* (protein serine/threonine activity, GO:0004674), *Glyma.20G205700* (serine-type endopeptidase

inhibitor activity, GO:0004867), and *Glyma.U029900* (function unknown, plasma membrane, GO:0005886), and 4 DEGs were identified in CK-SCN4, *Glyma.03G120700* (calmodulin binding, GO:0005516), *Glyma.10G093900*, *Glyma.11G099300* (a structural constituent of ribosome, GO:0003735), and *Glyma.19G125300* (calmodulin binding, GO:0005516).

Differentially Expressed Genes-Gene Ontology Annotation and Enrichment Analysis

By GO annotation analysis, even though only 7 DEGs were found between the SCN4- and SCN5- infected soybean roots, the SCN4-treated roots displayed 986 DEGs, while the SCN5-treated roots showed 913 DEGs when each was compared with the control. The top GO classifications for both CK-SCN4 and CK-SCN5 treatments in BP, MF, and CC are displayed in **Supplementary Figures 6A,B**, respectively. The top three enriched BPs for both treatments were a response to growth hormone, auxin-activated signaling pathway, and multidimensional cell growth. The top three CCs were plant-type cell wall and intracellular membrane-bound organelle for CK-SCN4, cell part, intracellular part, and intracellular membrane-bound organelle for CK-SCN5. The top two enriched MFs for both CK-SCN4 and CK-SCN5 were quercetin 3-O-glucosyltransferase (GTF, EC:2.4) activity and coniferyl-alcohol GTF activity, and the third MF type was pectinesterase (PE, EC:3.1.1.11) activity for CK-SCN4 and quercetin 4'-O-GTF activity for CK-SCN5. Various numbers of enriched DEGs for BP, MF, and CC were observed between CK-SCN4 and CK-SCN5 (**Figure 4A**). Especially, the numbers of SREs (arrowhead pointing in **Figure 4A**) in CK-SCN4 are greater than those in CK-SCN5, e.g., response to stimulus (SCN4/SCN5, 613/565), signaling (225/181), immune system process (121/84), detoxification (41/32), and binding (582/493) (**Figure 4A**).

Heterodera glycines Race 4 Infection-Induced More Stress Response Elements With More Upregulation Than H. glycines Race 5 Infection

In further analysis with BP annotation of SREs, a greater number of upregulated DEGs than that of downregulated ones was found in both CK-SCN4 and CK-SCN5 (**Figure 4B** and **Supplementary Table 9**). For example, the total number of the up-/down-regulated DEGs added together for each detected SRE was 8,780/6,662 for CK-SCN4 and 7,113/5,932 for CK-SCN5, indicating that more SRE genes were induced in the incompatible reaction than in the compatible reaction (**Figure 4** and **Supplementary Table 9**). The BP response to stimulus included all kinds of biotic and abiotic stresses or stimuli (**Supplementary Table 9**), e.g., pathogens or pests, organic or inorganic chemicals, hormones, temperature, endogenous or external stimulus, defense response, cell death, hypersensitive response, immune response, gene silencing, and MAP kinases (**Figure 4B**).

Since the response to growth hormone was the first top enriched GO-BP for both treatments mentioned above, the response to each hormone was inspected and compared.

TABLE 2 | Number of differentially expressed genes (DEGs) and transcripts (DETs) between un-infected (CK) and *Heterodera glycines*-infected treatments (SCN4 and SCN5), and between SCN5 and SCN4.

	DEG			DET		
	Total number	Up-regulated	Down-regulated	Total number	Up-regulated	Down-regulated
CKvsSCN4	1,533	726	807	2,694	1,121	1,573
CKvsSCN5	1,383	715	668	2,750	1,064	1,686
SCN5vsSCN4	7	6	1	8	7	1

SCN, soybean cyst nematode; vs.: versus.

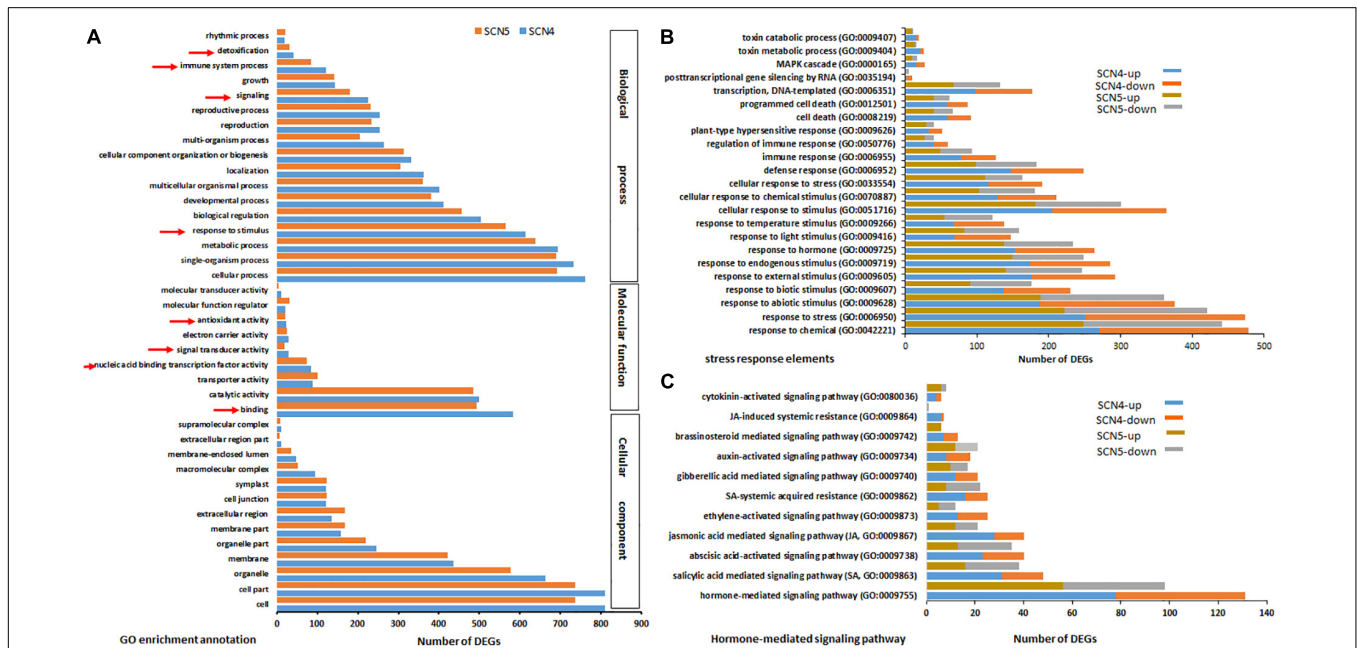


FIGURE 4 | Top GO-enriched annotation of the differentially expressed genes (DEGs) (A) and comparison of DEGs with up-/downregulation associated with stress response elements (B) and hormones (C) in soybean 09-138 infected by soybean cyst nematode race 4 (SCN4) and race 5 (SCN5) when compared with the control (CK). The red arrowheads in A indicate stress response elements.

Different responses to eight types of hormones were identified in both resistant and susceptible reactions compared with the control (Figure 4C and Supplementary Table 9). Although the hormones responded to infection with both nematode races, a greater number of upregulated DEGs than downregulated ones were found in the resistant reaction to SCN4 in SA- (up/down, 31/17), JA- (28/12), abscisic acid- (ABA, 23/17), ET(13/12), gibberellic acid- (GA, 12/9), brassinosteroid- (BR, 7/6), and cytokinin- (CTK, 4/2) mediated signaling pathways and SA- (16/9) and JA- (6/1) induced systemic resistance except for auxin-activated signaling pathway (IAA, 8/10) (Figure 4C). On the contrary, in the susceptible CK-SCN5 reaction, a lower number of upregulated than downregulated DEGs were found in SA- (16/22), ABA- (13/22), and ET- (5/7) mediated signaling pathways, and SA- (8/14) or JA- (0/1) induced systemic resistance. These results suggested that more SREs with more upregulation were activated by the incompatible reaction with the SCN4 infection when compared with the SCN5 infection.

***Heterodera glycines* Race 5 Infection Induced More Differentially Expressed Genes With More Upregulated Genes Associated With Cell Wall Modification and Carbohydrate Biological Process Than *H. glycines* Race 4 Infection When Each Was Compared With the Control**

Unsurprisingly, 23 (17 up/6 down) and 2 (2 up/0 down) DEGs for response to nematodes and syncytium formation were identified in the resistant soybean reaction to SCN4, respectively, while 34 (24 up/10 down) and 9 (9 up/0 down) DEGs for response to nematodes and syncytium formation were found in the susceptible reaction to SCN5, respectively. SCN syncytium formation requires plant cell wall modification, while cell wall modification involves multidimensional cell growth (SCN4/SCN5 DEGs: 38/48) which was the third BP-enriched element in both treatments. Therefore, all annotations involved in the cell wall were collected together, and a total of 28/26 GOs were annotated for CK-SCN4/CK-SCN5 (Figure 5A). A greater number of DEGs in CK-SCN5 than in CK-SCN4 was detected

in cell wall organization or biogenesis (SCN5/SCN4: 172/159), cell wall modification (51/27), plant-type cell wall modification (27/15), cell wall modification involved in multidimensional cell growth (12/1), plant-type cell wall loosening (11/2), cell wall thickening (11/9), and callose deposition in the cell wall (8/5).

Syncytium formation requires carbohydrates as a nutrition sink for nematode feeding, and sugar transporters are active in syncytia through inter- and intracellular transport processes (Hofmann et al., 2007, 2009; Hofmann and Grundle, 2008). The greater DEG number associated with syncytium formation in CK-SCN5 than in CK-SCN4 denoted the difference in the carbohydrate metabolic process. Thus, the biological processes of carbohydrates and corresponding three major groups (sugars, oligosaccharides, and polysaccharides), including metabolic, biosynthetic, catabolic and transport processes, were examined and compared between CK-SCN4 and CK-SCN5 (Figure 5B). In total, 178, 94, 57, and 13 DEG in CK-SCN4 were involved in carbohydrate metabolic, biosynthetic, catabolic, and transport processes, respectively; 203, 114, 60, and 22 DEGs in CK-SCN5 were associated with the four processes, respectively. For all carbohydrate-associated biological processes, 739 DEGs with 317 up-/422 downregulation in CK-SCN4 and 870 DEGs with 467 up-/403 downregulation in CK-SCN5 were found (Figure 5B). Polysaccharides are major components of carbohydrate metabolic, biosynthetic, and catabolic processes. The lower number of upregulated DEGs (130) than downregulated DEGs (195) for CK-SCN4 was detected in all the groups of carbohydrate metabolic and biosynthetic processes; for CK-SCN5, the greater number of upregulated genes than downregulated genes were only found in polysaccharide metabolic (up/down: 80/51), and biosynthetic (up/down: 55/30) processes and monosaccharide biosynthetic (7/5) process (Figure 5B). Polysaccharides and monosaccharides were a major carbohydrate catabolic process, but polysaccharides displayed more upregulation than downregulation, and monosaccharides showed more downregulation in both CK-SCN4 and CK-SCN5 (Figure 5B). The transport process of carbohydrates demonstrated that oligosaccharides and monosaccharides were major transport formats; only 1 upregulated DEG for polysaccharide transport was detected in CK-SCN5 but not in CK-SCN4. These results suggested that the compatible reaction induced more upregulated expression in the carbohydrate biological process than the incompatible reaction.

Kyoto Encyclopedia of Genes and Genomes Pathway Enrichment Analysis

The KEGG annotation showed that 273 and 292 unigenes were involved in 96 pathways in CK-SCN4 and 91 pathways in CK-SCN5, respectively (Supplementary Table 10). Two treatments shared 83 pathways; 13 and 8 unique pathways were detected in CK-SCN4 and CK-SCN5, respectively, e.g., SNARE interactions in vesicular transport (3 DEGs) and oxidative phosphorylation (4 DEGs) only in CK-SCN4 and N-Glycan biosynthesis (4 DEGs) and fatty acid degradation (1 DEG) only in CK-SCN5.

Phenylpropanoid biosynthesis was the top enriched KEGG pathway for both CK-SCN4 (Figure 6A) and CK-SCN5

(Figure 6B). The following top 5 enriched KEGG pathways were photosynthesis-antenna proteins, caffeine metabolism, cutin, suberin and wax biosynthesis, and AGE-RAGE signaling pathway in diabetic complications for CK-SCN4 (Figure 6A), and starch and sucrose metabolism, photosynthesis-antenna proteins, nitrogen metabolism, and AGE-RAGE signaling pathway in diabetic complications for CK-SCN5 (Figure 6B). Among the top 20 enriched KEGG pathways, glycine, serine and threonine metabolism, monoterpene biosynthesis, other glycan degradation, phagosome, plant hormone signal transduction, pyrimidine metabolism, and ubiquitin-mediated proteolysis were unique for CK-SCN4; five metabolisms (nitrogen, galactose, cysteine and methionine, ether lipid, and inositol phosphate), and valine, leucine, and isoleucine degradation were unique for CK-SCN5 (Figure 6B). Interestingly, the KEGG pathway circadian rhythm-plant was enriched in the top 8th for CK-SCN4 and in the top 14th for CK-SCN5 (Figures 6A,B).

Phenylpropanoid Biosynthesis Pathway Was Enriched by Both Nematode Infections

Both SCN4 and SCN5 infections induced phenylpropanoid biosynthesis, the top enriched KEGG pathway. Phenylpropanoid compounds are rich sources of metabolism in plants, providing precursors of lignin synthesis (Yao et al., 2021), while lignin is abundant in the cell wall, and it plays vital roles in plant defense in local or systematic required resistance (Dixon et al., 2002). Enzymatic reactions are considered key steps in the biosynthesis of major classes of phenylpropanoid compounds. Thirty-five KEGG orthologies of peroxidase (K00430, EC:1.11.1.7), the richest enzyme annotated in this pathway, were detected in either up- or down-regulation for both CK-SCN4 (up/down: 24/4) and CK-SCN5 (up/down: 23/6); 10 out of 35 *peroxidase* genes were unique for CK-SCN5, and 7 were unique for CK-SCN4 (Figures 7A,B). The peroxidase enzyme plays a role in the final step of the biosynthesis of lignin to produce p-hydroxyphenyl lignin, guaiacyl lignin, 5-hydroxyguaiacyl lignin, and syringyl lignin (Figure 7B). The gene expression levels of *peroxidase 3* (*Glyma.03G208200*, *Glyma.10G022500*) were increased by up to 5.2- to 7.3-fold after nematode infection when compared to the control. The SCN4 infection induced two downregulated DEGs: one gene, *Glyma.15G002600*, encoded enzyme shikimate O-hydroxycinnamoyltransferase (HST, EC:2.3.1.133), which catalyzes the reaction of two substrates, 4-coumaroyl-CoA (4-CCoA) and shikimate, to produce two products, CoA and 4-coumaroylshikimate (CSH); the other DEG, *Glyma.03G070300*, encoded serine carboxypeptidase-like 11 (SCPL11, EC:2.3.1.91), which can form sinapoylcholine (sinapine) and plays vital roles in abiotic and biotic stress responses, and growth and development as well (Xu et al., 2021). The SCN5 infection induced the downregulation of DEG *Glyma.04G227700* encoding caffeic acid 3-O-methyltransferase (COMT)-like (EC:2.1.1.68), a key enzyme regulating lignin synthesis (Trabucco et al., 2013) but not for the SCN4 infection. One gene (*Glyma.15G031300*) encoding the enzyme cyanogenic beta-glucosidase 13-like (EC:3.2.1.21) involved in cell wall lignification increased the expression level 5.87-fold by SCN5 infection (Figure 7A). Both nematode infections activated the upregulation of the

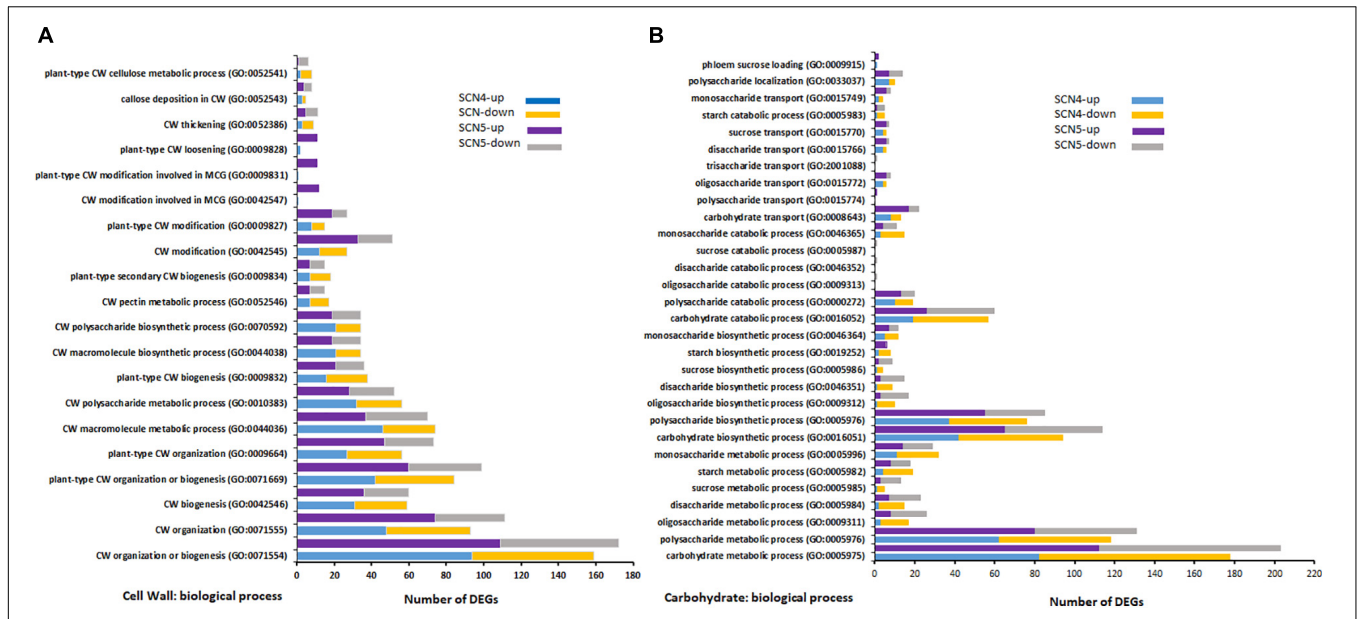


FIGURE 5 | Top GO-enriched differentially expressed up-/down-regulated genes (DEGs) associated with cell wall (CW) modification (A) and carbohydrate biological process (B) on soybean breeding line 09-138 infected by SCN4 (soybean cyst nematode race 4) and SCN5 when compared with the control. The x-axis displays DEG number, and the y-axis represents the GO pathway.

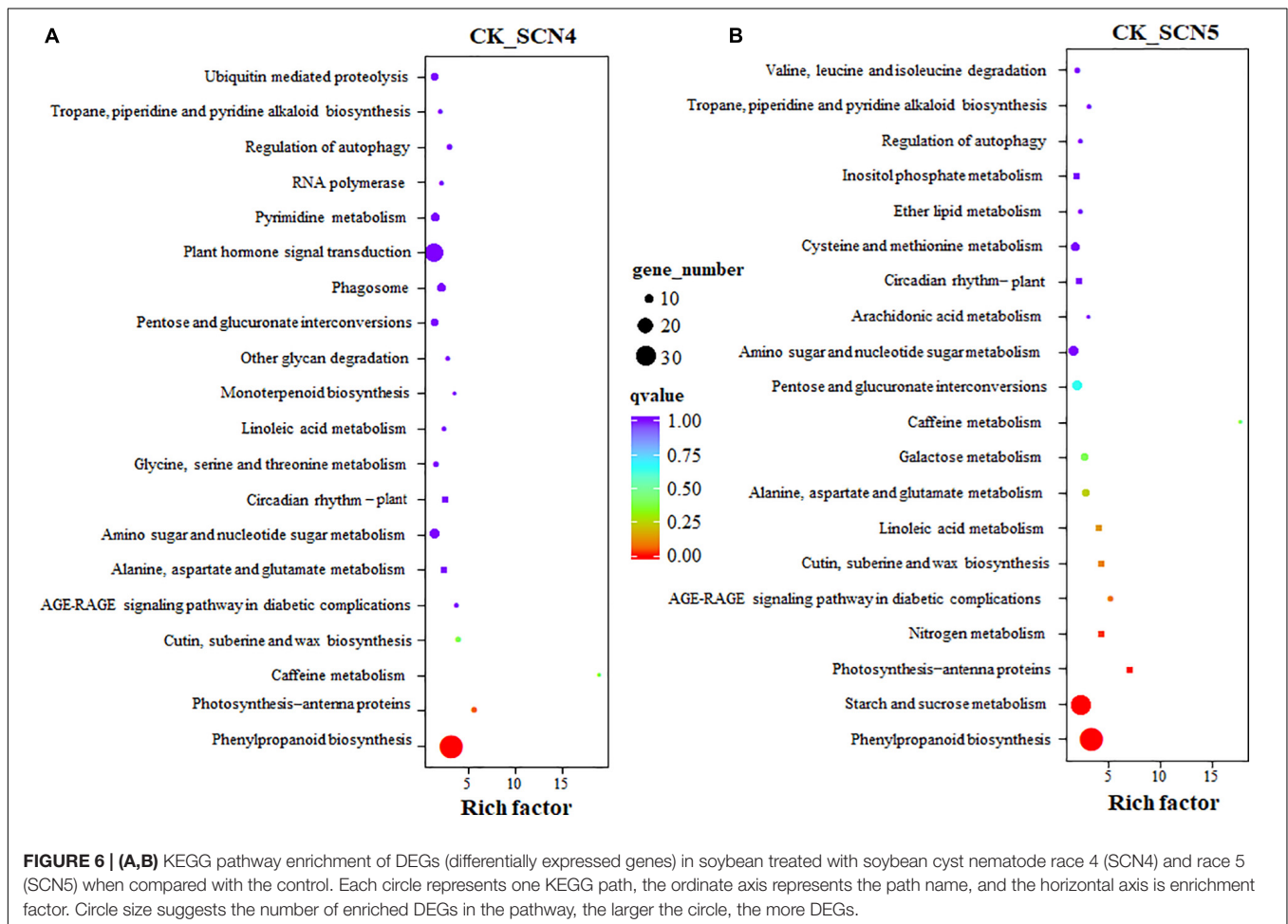


FIGURE 6 | (A,B) KEGG pathway enrichment of DEGs (differentially expressed genes) in soybean treated with soybean cyst nematode race 4 (SCN4) and race 5 (SCN5) when compared with the control. Each circle represents one KEGG path, the ordinate axis represents the path name, and the horizontal axis is enrichment factor. Circle size suggests the number of enriched DEGs in the pathway, the larger the circle, the more DEGs.

DEG *Glyma.09G201200* encoding enzyme cinnamyl-alcohol dehydrogenase (CAD, EC:1.1.1.195) to catalyze the final step of monolignol biosynthesis, including p-coumaryl alcohol (H), coniferyl alcohol (G), and sinapyl alcohol (S), the main component of lignin (Figures 7A,B). These results indicated the expression variation of peroxidase enzyme, HST, SCPL, and other enzymes associated with lignin synthesis after nematode infection resulted in compatible and incompatible reactions by regulating plant cell wall modification.

Plant Hormone Signal Transduction, Plant–Pathogen Interaction, and MYB Transcription Factors Contributed to Defense Response

A similar number of DEGs were identified in both the plant hormone signal transduction pathway (SCN4/SCN5: 26/25) (Figures 7C,D) and the plant–pathogen interaction pathway (SCN4/SCN5: 12/14) (Figures 7E,F) but with various expressions in levels or gene types. As illustrated in Figures 7C,D, both nematode infections positively or negatively regulated the expression of the hormones (SA, JA, ET, GA, ABA, IAA, and CTK) but with qualitative and quantitative differences in soybean (Figure 7D). For example, the expression level of *PR1* (pathogenesis-related protein 1, *Glyma.15G062400*), which is involved in multiple pathways not only plant hormone signal transduction (SA) but also MAPK signaling and plant–pathogen interaction, was increased by more than 5-fold in soybean by the SCN4 (6.5-fold) and SCN5 (5.2-fold) infections when compared with the control. In the auxin pathway, which is linked to cell enlargement and plant growth, *AUX1* and *AUX/IAA* were upregulated, e.g., auxin transporter-like protein 4 (*Glyma.03G063900*) was positively regulated by both the SCN4 (4.1-fold) and SCN5 (4.5-fold) infections, and auxin-responsive gene *SAUR* (small auxin upregulated RNA) was both up- and downregulated (Figures 7C,D). Two transcription factors PIF3 (phytochrome-interacting factor 3)-like encoded by *Glyma.19G224700* and *Glyma.20G091200*, were identified in branch GA pathway as well as in circadian rhythm plant pathway, with the former being downregulated by both nematode infections and the latter being downregulated by the SCN5 infection. Three genes (*Glyma.09G066500*, *Glyma.19G069200*, and *Glyma.11G018000*) encoding PP2C (protein phosphatase 2C) in the ABA pathway were upregulated; the SCN4 infection induced an approximately 3-fold increase in the expression level of the three genes, and the SCN5 infection increased by 2.2- to 2.8-fold the expression level of the first two genes. *JAZ* (*jasmonate* zim domain) encoded by *Glyma.17G043700* in JA associated with stress response was detected only in CK-SCN5 with downregulation but not in CK-SCN4. On the contrary, DEGs encoding ETR (ethylene response sensor 2), EBF1 (EIN3-binding F-box protein 1), and ERF1 (ethylene-responsive transcription factor 1) in the ET pathway linked to fruit ripening and senescence were found only in resistant response to SCN4 but not to SCN5 when compared with the control (Figures 7C,D), indicating that the ET pathway played a role in defense response, and that JA expression was inhibited in susceptible response.

In the plant-pathogen interaction pathway, CERK1 (chitin elicitor receptor kinase 1, *Glyma.02G270800*) classified as PRRs in PTI, CDPK (calcium-dependent protein kinase), Rboh (respiratory burst oxidase) triggering reactive oxygen species (ROS) burst correlated to disease HR, CaM/CML (calmodulin/calmodulin-like protein) associated with HR, cell wall reinforcement and stomatal closure, and *PR1* (*Glyma.15G062400*) linked to phytoalexin accumulation and miRNA production, were identified in both CK-SCN4 and CK-SCN5 but with various expression levels; downstream WRKY transcription factor *AtWRKY33* (*GmWRKY15*, *Glyma.02G232600*) and *Pti6* (PR genes transcriptional activator 6, *Glyma.02G236800*) connected to defense-related gene induction were significantly detected only in CK-SCN4 (Figures 7E,F). One gene, *Glyma.01G068000*, encoding HSP90 was detected with downregulation in CK-SCN4, but two other genes, *Glyma.09G131500* (*hsp83*) and *Glyma.16G178800* (*hsp83*), were negatively regulated with significance in CK-SCN5, suggesting that PTI was triggered initially by both nematode infections and that later, ETI was activated by the SCN4 infection.

In the KEGG analysis, 20 MYB or MYB-like transcription factors were detected; of these, 12 were found only in CK-SCN4 and 3 only in CK-SCN5 treatments, denoting that MYB transcription factors might regulate defense response. Three DEGs (*Glyma.03G261800*, *Glyma.16G017400*, and *Glyma.19G260900*) were associated with MYB-related transcription factor LHY involved in plant circadian rhythm as well, were negatively regulated after both SCN4 and SCN5 infections.

Starch and Sucrose Metabolism Pathway Regulated Susceptible Response to *Heterodera glycines* Race 5 and Chitinase I Was Upregulated in Defense Response to *H. glycines* Race 4

The top second KEGG enriched starch and sucrose metabolism pathway in CK-SCN5 treatment had 40 DEGs including 5 sucrose synthases, 11 pectinesterase-like or pectinesterase/pectinesterase inhibitors, 5 beta-glucosidase, and 4 alphas, alpha-trehalose-phosphate synthase [UDP-forming] and others, which are involved in plant growth and cell wall modification or cell wall lignification; these genes are necessary for syncytium formation. Of these, 8/9 and 16/19 up-/downregulated genes were identified for CK-SCN4 and CK-SCN5, respectively (Supplementary Figure 7A), e.g., *Glyma.15G223500* (pectinesterase/pectinesterase inhibitor 47) with a 5-fold increase in expression level after SCN5 infection, while upregulated *Glyma.03G216000* (pectinesterase/pectinesterase inhibitors 20) had 3.3-fold increased expression level in SCN4-infected roots, indicating that these DEGs played roles in the compatible or incompatible reaction. In the amino sugar and nucleotide pathway, 17/3 up-/down regulated DEGs were found, two of which had more than 4-fold significantly increased expression level, *Glyma.02G042500* encoding chitinase class I precursor and *Glyma.18G120700* encoding hevacine-A-like protein in SCN4-infected roots (Supplementary Figure 7B), suggesting that the two genes might contribute to defense response with SCN4 infection.

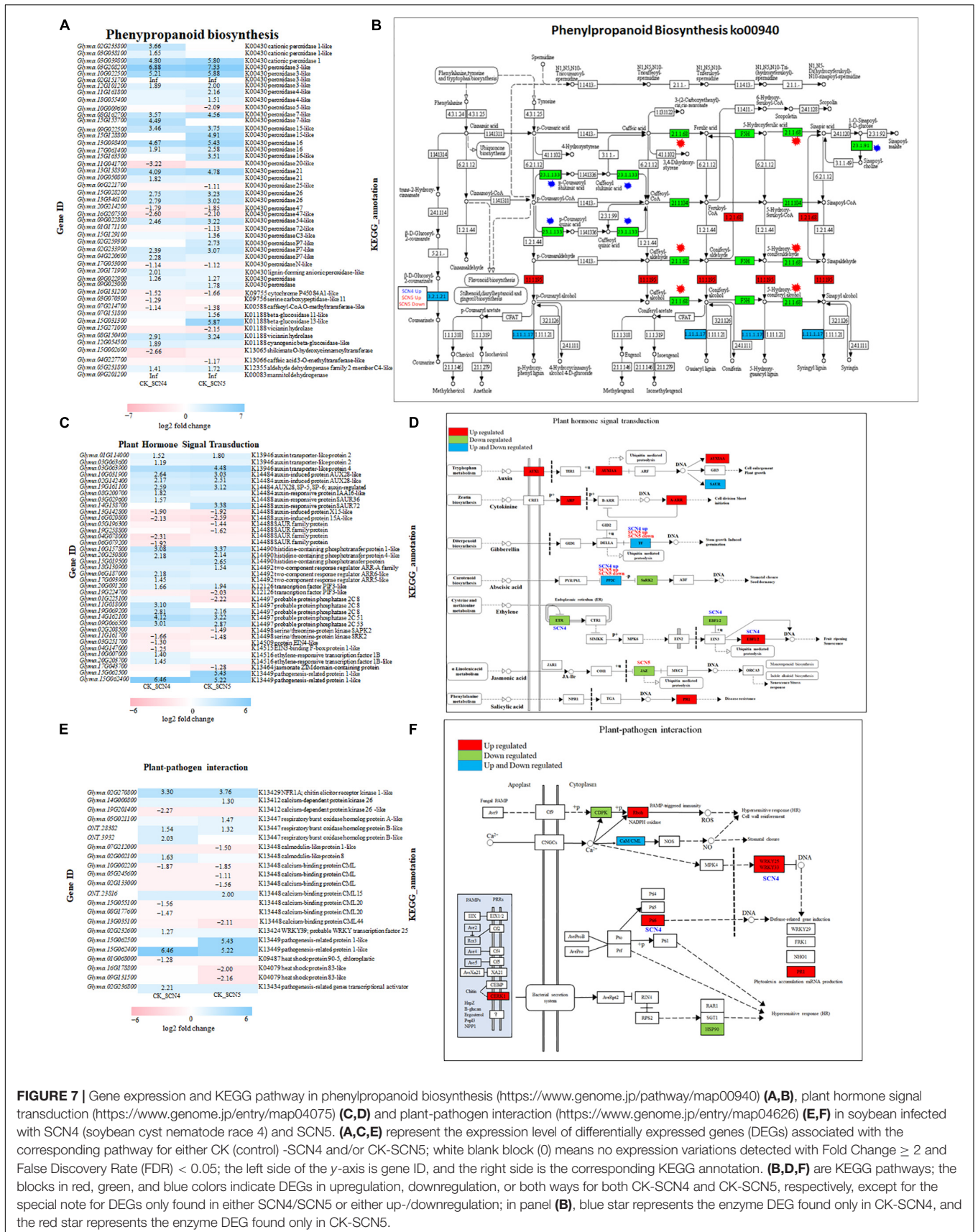


FIGURE 7 | Gene expression and KEGG pathway in phenylpropanoid biosynthesis (<https://www.genome.jp/pathway/map00940>) (A,B), plant hormone signal transduction (<https://www.genome.jp/entry/map04075>) (C,D) and plant-pathogen interaction (<https://www.genome.jp/entry/map04626>) (E,F) in soybean infected with SCN4 (soybean cyst nematode race 4) and SCN5. (A,C,E) represent the expression level of differentially expressed genes (DEGs) associated with the corresponding pathway for either CK (control) -SCN4 and/or CK-SCN5; white blank block (0) means no expression variations detected with Fold Change ≥ 2 and False Discovery Rate (FDR) < 0.05 ; the left side of the y-axis is gene ID, and the right side is the corresponding KEGG annotation. (B,D,F) are KEGG pathways; the blocks in red, green, and blue colors indicate DEGs in upregulation, downregulation, or both ways for both CK-SCN4 and CK-SCN5, respectively, except for the special note for DEGs only found in either SCN4/SCN5 or either up-/downregulation; in panel (B), blue star represents the enzyme DEG found only in CK-SCN4, and the red star represents the enzyme DEG found only in CK-SCN5.

Protein–Protein Interaction Analysis Revealed Valine-Glutamine-WRKY Interactions and Cytokinin Two-Component Response Regulator ARR-Glutaredoxins-BolA Interactions Might Be Involved in Defense Response

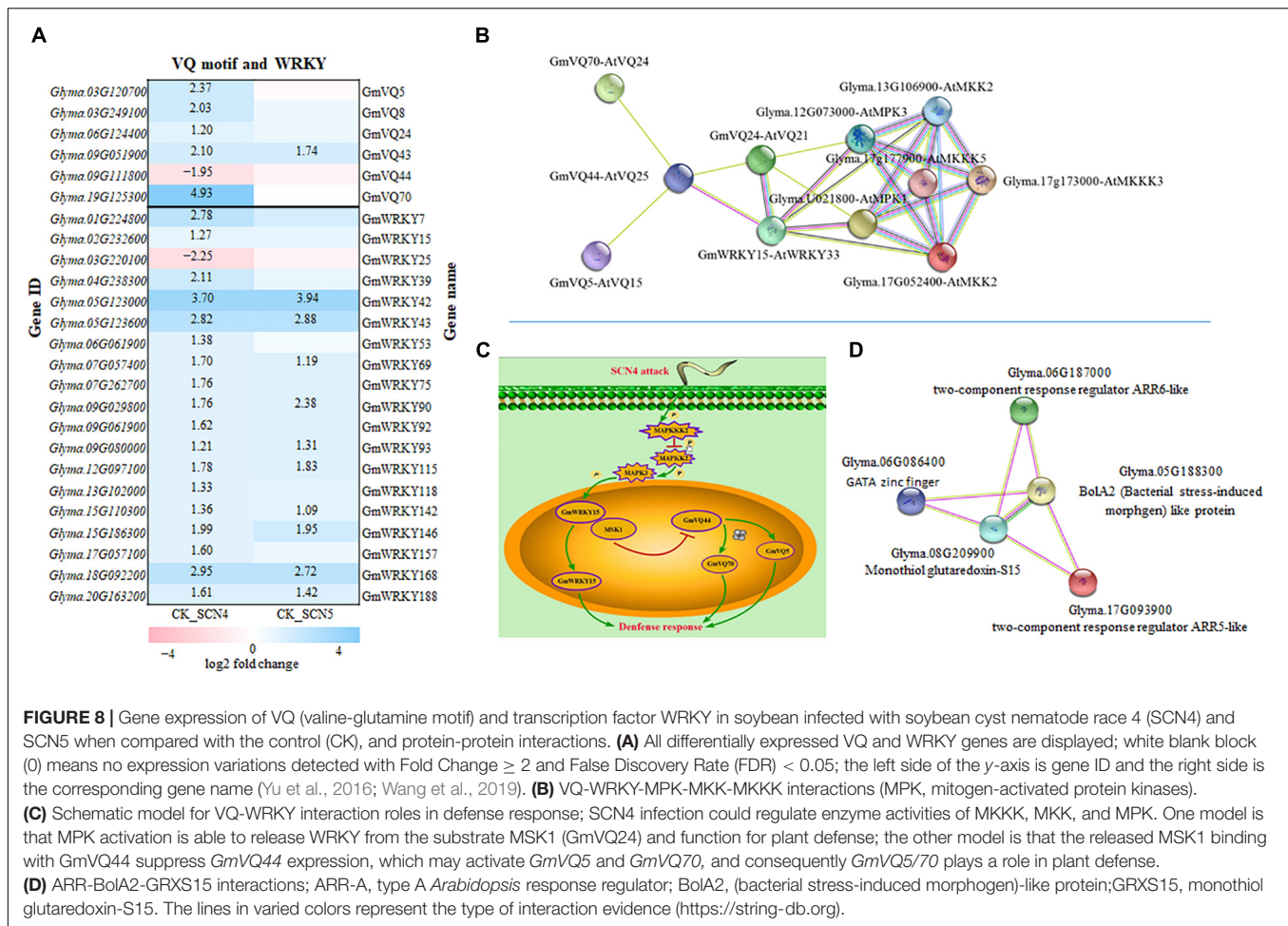
The PPI analysis among all the 2,255 DEGs indicated a ratio of 7,063/2,632 (2.7 ×) of protein-protein networks for CK-SCN4/CK-SCN5, including 866/194 activation (4.5 ×), 2715/1381 (2 ×) binding, 781/280 (2.8 ×) catalysis, 401/104 (3.9 ×) expression, 462/96 (4.8 ×) inhibition, 874/166 (5.3 ×) ptmod (post translational modification), and 964/411 (2.3 ×) reaction, indicating soybean resistant response to the SCN4 infection activated more protein–protein interactions than a susceptible response to the SCN5 infection, especially in post translational modification, inhibition, and activation.

Two genes, *Glyma.03G120700* (*GmVQ5*) and *Glyma.19G125300* (*GmVQ70*), were identified as significant DEGs with a 2.7–3.2-fold increase in CK-SCN4 than in CK-SCN5, encode calmodulin binding protein (CaMBP) (GO:0005516) carrying conserved VQ (Valine-Glutamine) motif (FxxhVQxhTG), which is capable of interacting with the transcription factor WRKY DNA-binding domain to play vital roles in stress response (Wang et al., 2014). Thus, the interactions between VQ and WRKY proteins encoded by DEGs were explored based on the Pfam domain; 6 (*GmVQ5*, *GmVQ8*, *GmVQ24*, *GmVQ43*, *GmVQ44*, and *GmVQ70*) out of 26 VQs and 19 out of 74 WRKYs annotated were detected in either CK-SCN4, CK-SCN5, or SCN5-SCN4 (Figure 8A). Interestingly, *GmWRKY15*, encoded by the upregulated *Glyma.02G232600*, homologous with *Arabidopsis* *AtWRKY33* (WRKYGQK/WRKYGQK) in the plant–pathogen interaction pathway, was identified in CK-SCN4, which could interact with *GmVQ24* (*AtVQ21*, *Glyma.06G124400*), and both were linked to *GmVQ44* (*AtVQ25*, calcium-binding protein, *Glyma.09G111800*) with downregulation, which could bind with the two homologous CaM-binding proteins *GmVQ5* and *GmVQ70* (Figure 8B). VQ and WRKY numbers were designed according to Yu et al. (2016), Wang et al. (2019), respectively.

AtWRKY33 can interact with a VQ protein called mitogen-activated protein (MAP) kinase substrate1 (MSK1, *AtVQ21*), a substrate of MPK4 (MAP kinase 4), which can be activated in the presence of pathogen attack or flagellin, and then *AtWRKY33* is released to induce the expression of phytoalexin deficient 3 (*PAD3*) in the nucleus, and thereby to increase defense response (Andreasson et al., 2005; Cheng et al., 2012). In addition, MPK3 and MPK6 can also interact with *AtWRKY33* together, leading to an increase in phytoalexin-related gene expression (Alves et al., 2014). The overexpression of *AtWRKY33* resulting in decreased susceptibility to beet cyst nematode (*Heterodera schachtii*) (Ali et al., 2013) supports that *AtWRKY33* may be involved in soybean cyst nematode resistance. To find the kinase of the substrate *GmVQ24* (MSK1) in the annotation data, all expressed MAP kinases were searched

and no homologous *AtMPK4* was found; interestingly, only two homologous MPK3s encoded by *Glyma.U021800* and *Glyma.12G073000*, each with a 1-fold increase in expression level in CK-SCN4, were identified to interact with *GmWRKY15* and *GmVQ24*. Further searching of other DEGs encoding MAPKs exhibited that two genes encoding MAP kinase kinase 2 (MPKK2) had a 1.5- to 1.8-fold reduction in the expression level only in CK-SCN4 (Figure 8B). Two more MPKKs encoded by *Glyma.17G173000* and *Glyma.17G177900*, which could bind to MPK3 and MPKK2, were found with a 1.2-fold increase of gene expression in CK-SCN4 (data not shown). All these results demonstrated that MPKKK/MPKK/MPK/WRKY-VQ-CaMBPVQ might work together in soybean to defend against SCN4 infection. Based on these data, a model with a VQ–WRKY interaction in response to SCN4 infection was established with two possibilities, one is *GmWRKY33* release by MPK3 activation or any other kinase to induce phytoalexin production, which activates plant defense as described in *Arabidopsis* (Alves et al., 2014); the other is *GmVQ44* binding to *GmVQ24* (MSK1) to suppress *GmVQ44* expression, which results in high expression of CaMBP *GmVQ5/70* to induce SCN resistance (Figure 8C).

Another interesting protein–protein interaction was detected in the plant hormone signal branch cytokinin pathway with two upregulated CK-SCN4 DEG genes (Figure 8D), *Glyma.06G187000* (*ARR6*, $\text{Log}_2\text{FC} = 2.2$) and *Glyma.17G093900* (*ARR5*, $\text{Log}_2\text{FC} = 1.5$), which are classified as two-component response regulator ARR-A (type A *Arabidopsis* response regulator) family composed of an inner membrane-spanning histidine kinase and a cytoplasmic response regulator to allow organisms to sense and respond to changes in environmental stimuli (Stock et al., 2000). In CK-SCN4, the *ARR5*-like and *ARR6*-like proteins were capable of interacting with a mitochondrial chaperone, monothiol glutaredoxin-S15 (*GRXS15*), and *BolA2*. *GRXS15* encoded by the downregulated DEG *Glyma.08G209900* ($\text{Log}_2\text{FC} = -1.8$), and *BolA2* encoded by a downregulated DEG *Glyma.05G188300* ($\text{Log}_2\text{FC} = -1.6$) could connect with the transcription factor GATA zinc finger (*Glyma.06G086400*, $\text{Log}_2\text{FC} = 1.7$) (Figure 8C). Mitochondria *GRAX15* was reported to play key roles in iron-sulfur protein maturation in the plant (Moseler et al., 2015) and a transcription regulator *BolA2* containing a helix-turn-helix (HTH) motif for nucleic acid binding (Kasai et al., 2004). *AtBola3*–*GRXS17* interaction plays a key role in suppressing abiotic tolerance (Cheng et al., 2011; Qin et al., 2015). In a similar manner, inhibition of the gene expression of both *BolA2* and *GRAX15* with SCN4 infection resulting in resistance denoted that *BolA2*–*GRAX15* interaction may negatively regulate soybean resistance to SCN4. The function of *BolA2*–*GRAX15* interaction has not been reported yet, but it is known that *BolA2* is nucleon-cytoplasmic and interacts with *GRAX15* (Couturier et al., 2014). Furthermore, *BolA2* was linked to the lncRNA detected above. Therefore, the interactions of *ARR5/6*–*GRAX15*–*BolA2*–*GATA* might indicate a new complex interaction partaking in plant defense in response to SCN4 infection.



Validation of Differentially Expressed Genes by QRT-PCR Assay

The expression levels of 23 DEGs obtained from full-length transcriptome sequencing were compared with those obtained from qRT-PCR. The expression correlation (R^2) between full-length-seq and qRT-PCR was up to 0.6958 (Figure 9A), and the comparison of relative expression levels between SCN4 and SCN5 demonstrated that the qRT-PCR data almost matched with the full-length-seq data (Figure 9B). The relative expression levels of four WRKY and three VQ genes are also displayed in Figure 9C.

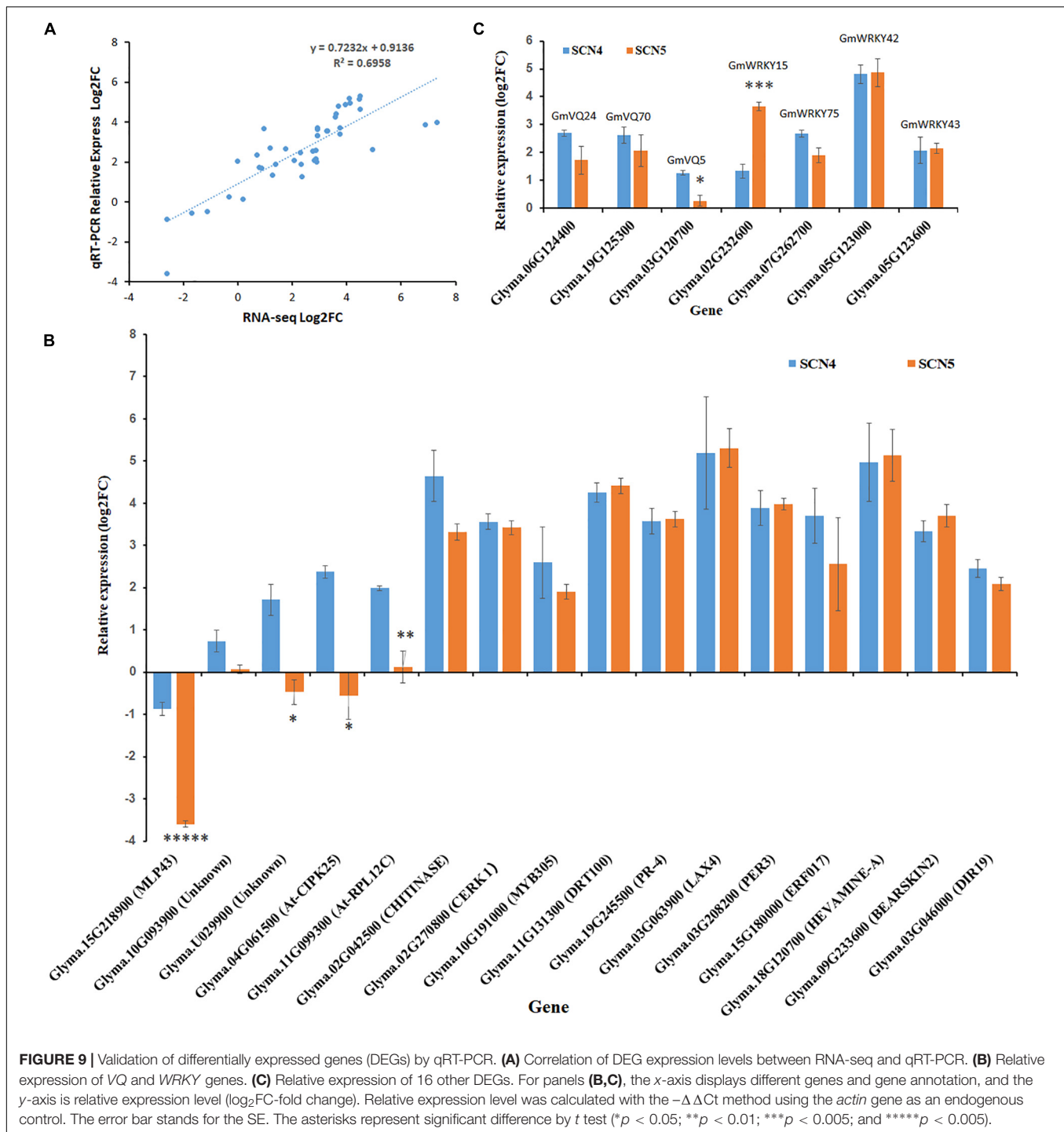
DISCUSSION

Full-Length Transcriptome Analysis: A Powerful Tool to Analyze Transcriptional and Post-transcriptional Regulation

In this study, full-length transcriptome analysis on soybean was first conducted to compare responses of the same soybean genotype incompatible and compatible to different *H. glycines* races. Obviously, the advantages of full-length transcriptome sequencing include high-cost performance, high throughput, no

GC specificity, and base bias, long sequencing reads, accurate quantification at the transcriptome level, accurate identification of structure characteristics (e.g., AS, fusion gene, APA) without breaking sequences or gene structures, and DEG/DET analysis at once. Compared to NGS, ONT needs less read numbers to cover the same amount of transcripts. For example, an average of 3–5 \times greater number of clean reads was obtained by NGS (Zhang et al., 2017; Neupane et al., 2019; Miraeiz et al., 2020) when compared with the average number of clear reads in this study. In addition, the identified novel genes can generate new AS events.

The change in numbers of AS events at 5' splicing sites or at 3' splicing sites after both nematode infections and various transcripts among the three treatments indicated that the nematode infections did trigger differential AS events to generate multiple transcripts that might result in plant defense or susceptibility. A few cases have been reported that AS regulates plant stress responses and/or adaptations (Bedre et al., 2019; John et al., 2021; Martín et al., 2021). For instance, the tobacco gene *N* resistant to tobacco mosaic virus (TMV) can be alternatively spliced to produce two transcripts that are required to contribute to complete resistance to TMV (Dinesh-Kumar and Baker, 2000). AS regulates ABA and light signaling pathways to coordinate plant growth and stress response (Thirupathi, 2020).



Interestingly, we found that more abundant APA events per gene were identified with > 5 poly A sites (23.8%) than those in 1-5 poly A sites in soybean, while 2 poly A sites were found as the most abundant APA event distribution in sorghum (Chakrabarti et al., 2020), and 1 poly A site for the tree species *Liriodendron chinense* in the magnolia family (Tu et al., 2021). The poly A site distribution in plants can be shifted by abiotic stresses (Yan et al., 2021), and novel stress-specific cis-elements in intronic

poly A sites may contribute to abiotic stresses (Chakrabarti et al., 2020). In this study, nematode infection increased the > 5 -poly A site distribution than control, especially, the > 10 poly A sites in the incompatible reaction further confirmed that APA may be involved in nematode stress response. Generally, stress-responsive APA plays a key role in regulating abiotic stress and biotic stress through crucial stress-responsive genes and pathways (Ye et al., 2019; Chakrabarti et al., 2020;

Yan et al., 2021). For instance, Ye et al. (2019) found that APA in rice responds to heat stress tolerance as a negative regulator, to Cd stress by regulating DNA repair and cell wall formation, and to disease stress (bacterial blight, rice stripe virus, and rice blast) by regulating chlorophyll metabolism. Comparisons of APA sites linked with annotated genes or pathways between control and nematode-infected treatments will uncover the APA function in soybean associated with nematode resistance or susceptibility. Furthermore, the difference in the fusion transcripts and lncRNAs among the treatments demonstrates that full-length transcriptome sequencing is a powerful tool to analyze post-transcriptional modification.

Plant Hormone Signal Transduction and Plant–Pathogen Interaction Implicating Plant Defense Response

Both the GO and KEGG enrichment analyses consistently confirmed that stress response elements and associated pathways (plant hormone signal transduction and plant–pathogen interaction) contribute to plant innate immune response to SCN. Surprisingly, when all the DEGs in these groups/pathways identified in this study were compared with those listed by Zhang et al. (2017), Miraeiz et al. (2020), only 1–3 overlapping genes were found; the only DEG, *Glyma.11G207000*, encoding leucine-rich repeat-containing protein was found in all three studies. The difference might be caused by various SCN–soybean interaction systems and be affected by inoculation time, genotypes, nematode types, and possible nematode inoculation density. For example, a transcriptome analysis was conducted as early as 8 h post inoculation with HG type 0 on soybean Peking, *G. soja* PI 468916, Fayette, and Williams 82 by Miraeiz et al. (2020), and later sedentary phase at 3, 5 and 8 days (pooled samples) with HG type 2.5.7 on two *G. soja* genotypes by Zhang et al. (2017), and 8 days with HG Type 2.5.7 and HG Type 1.2.3.5.6.7 on the same genotype 09-138 in this study. Most transcriptome profiling studies are examined within 2–10 days since SCN feeding site establishment generally required 48 h after inoculation, and syncytium formation and syncytium collapse happens within 2–10 days after inoculation in resistance response (Klink et al., 2007; Kandoth et al., 2011).

Although uniquely differential expression genes to SCN4 or SCN5 infection were identified on 09-138, most DEGs were expressed in both resistant and susceptible reactions with only a small difference (Figure 7), indicating qualitative and quantitative traits. For example, in the plant–pathogen interaction pathway, upstream PPR-CERK1 and downstream Ca^{2+} -dependent signal components CDPK, Rboh (triggering reactive oxygen burst, ROS), CaM/CML, HSP, and PR1 were all identified in both compatible and incompatible roots, and downstream WRKYs and Pti6 with upregulation were only detected in the incompatible response (Figures 7E,F), while these components are involved in PTI and/or ETI activity. In the plant-hormone pathway, all phytohormones except for brassinosteroid were activated with 1–4 key DEGs, while the phytohormone networks of the JA, ET, and SA signaling pathways are required for PTI and ETI as well (Cui et al., 2015).

Naveed et al. (2020) found that both resistant and susceptible plants showed almost identical transcriptome responses, but that the resistant plants achieved high-amplitude transcriptional reprogramming several hours earlier than the susceptible plants through a defense phytohormone signaling network. Here, we only detected one time point in the later infection stage, and a time series examination will reveal more about how the pathways work together to defend against nematode attacks. The PTI-ETI continuum concept with crosstalk between PTI-ETI signal components effectively activating plant immune responses has the increasing attention of biologists (Mine et al., 2018; Naveed et al., 2020; Yuan et al., 2021), and this concept may also be applicable for SCN–soybean interaction based on the identified signal components in these pathways. In addition, surprisingly, even though 09-138 contained the Peking-*rhg1a* locus, there were not any DEGs found in that region, suggesting an alternative resistance mechanism in 09-138.

Transcription Factor Functional Protein–Protein Interactions in Soybean Defense Response

Transcription factors as transcriptional regulators function by binding to the promoter region of target genes and regulating plant response to environmental stress, e.g., altering the expression of cascades of defense genes (Chen et al., 2002; Alves et al., 2014). The novel transcripts identified exhibited more than 5,000 TFs, with the top 5 being WRKY, AP2/ERF-ERF, NAC, GRAS, and bHLH. WRKY TFs, as the largest family of transcriptional regulators, were consistently identified in almost all transcriptome analyses of SCN infection (Ithal et al., 2007; Kandoth et al., 2011; Mazarei et al., 2011; Wan et al., 2015; Zhang et al., 2017; Song et al., 2019; Jiang et al., 2020; Miraeiz et al., 2020). The WRKY family, associated with stress responses in soybean, has been examined, e.g., in response to soybean rust disease (*Phakopsora pachyrhizi*) (Bencke-Malato et al., 2014), salt stress (Yu et al., 2016), dehydration and salt stress (Song et al., 2016), and soybean cyst nematode (Yang et al., 2017). However, 19 WRKY-DEGs to SCN4 or SCN5 identified in this study were not found in previous studies for SCN infection (Yang et al., 2017; Zhang et al., 2017; Miraeiz et al., 2020), indicating that these identified WRKY factors may be specific to the line 09-138 or the nematode races used in this study.

WRKY factors in plants are divided into five groups (I, IIa + IIb, IIc, IID + IIE, and III), and the interaction between the WRKY domain and its partner domain (e.g., VQ) is involved in signaling, transcription, and other important biological processes (Chi et al., 2013; Alves et al., 2014). The WRKY–VQ interaction has been well-studied in *Arabidopsis*. As mentioned above, AtWRKY33 (group I)/MSK1 interaction activates plant defense gene expression (Andreasson et al., 2005; Qiu et al., 2008; Alves et al., 2014). AtWRKY25, a homolog of AtWRKY33, is also able to interact with MSK1 and MPK4 in the absence of a pathogen (Cheng et al., 2012). Additionally, AtWRKY33 is able to interact with two other VQ proteins, AtSIB1 (sigma factor binding protein 1, AtVQ23) and AtSIB2 (AtVQ16), in the nucleus to activate

resistance to a necrotrophic pathogen, *Botrytis cinerea* (Lai et al., 2011). In this study, the lack of homologous MPK4 indicates a possible alternate interaction mechanism participating in nematode defense response. In the soybean–pest interaction system, only one GmVQ58 was identified as a negative regulator for soybean resistance to the common cutworm (*Spodoptera litura* Fabricius), and GmVQ58 could interact with GmWRKY32 (Li X. et al., 2020).

The CaM family is composed of ubiquitous Ca²⁺-binding proteins or calcium sensor proteins, which can play a key role in cellular signaling cascades coupling various environmental stimuli (Zeng et al., 2015). For instance, AtCaMBP25/AtVQ15 was found as a negative effector regulating osmotic stress tolerance during seed germination and seedling growth (Perruc et al., 2004). CaM-mediated signaling can regulate ROS homeostasis directly and indirectly (Zeng et al., 2015). Here, we found DEGs *CaM/CML* in the plant-pathogen pathway and two *CaMBP*GmVQ5/VQ70s, which directly or indirectly interact with other VQs and WRKYs to form complex GmVQ5/70-GmVQ44-GmWRKY15-GmVQ24 (MSK1) (Figure 8B). Only the incompatible reaction could activate GmVQ5/70, GmWRKY15, and GmVQ24, and suppress the connector GmVQ44, indicating that VQ–WRKY interactions partake in plant defense response to SCN4. Additionally, the high expression level of *GmWRKY15* increasing SCN resistance is consistent with that the homologous *AtWRKY33* overexpression decreasing *H. schachtii* susceptibility (Ali et al., 2013).

Top Enriched Phenylpropanoid Biosynthesis Pathway Involved in Both Resistant and Susceptible Response

The phenylpropanoid pathway has been considered a ubiquitous defense response against pathogens including nematodes, and this pathway is always enriched in SCN–soybean interaction (Edens et al., 1995; Dixon et al., 2002; Zhang et al., 2017; Li et al., 2018; Singh et al., 2019; Miraeiz et al., 2020). Approximately 75% of DEGs in this pathway were peroxidase, and more than 80% of the enriched peroxidase genes could be induced after nematode infection either in susceptible or in resistant response, which is matched with previous reports (Miraeiz et al., 2020), suggesting that peroxidase plays a central role in response to nematode attack. Peroxidase contributes to plant defense by modifying the cell wall composed of lignin, suberin, feruloylated polysaccharides, and HPRG (extensins), enhancing ROS production and phytoalexin productions (Pandey et al., 2017). The key enzymes linked to lignin biosynthesis and uniquely expressed in the incompatible (e.g., HST, *Glyma.15G002600*) or compatible (e.g., beta-glucosidase, *Glyma.15G031300*) reaction will be potentially targeted genes for further studies to elucidate plant defense or pathogenesis.

Carbohydrate Biological Process and Cell Wall Modification Roles in Soybean Cyst Nematode Susceptibility

SCN5 infection-induced more DEGs with more upregulated genes in cell wall modification and carbohydrate metabolic

process than SCN4 infection but with some quantitatively overlapping DEGs, indicating unique and accumulative DEGs in carbohydrate biological process including metabolism, biosynthesis, catalyst and transport work together to make SCN susceptibility or resistance. Miraeiz et al. (2020) also reported that susceptible W82 showed less responsive genes, more downregulated genes in carbohydrate metabolism and transport proteins, and some overlapping genes with the resistant genotype 8 h post inoculation. Nevertheless, again, there are less common DEGs found between 8 h post inoculation (Miraeiz et al., 2020) and 8 days (this study), a beta-glucosidase was downregulated at 8 h but upregulated at 8 days in this study, suggesting spatially and temporally differential expression. In the starch and sucrose metabolism pathway, sucrose synthase, hexokinase, beta-amylase, and polygalacturonase were suppressed more, and beta-glucosidase, beta-fructofuranosidase, and galacturonosyltransferase were induced more in the compatible interaction than in the incompatible interaction, proving the complexity for nematode pathogenicity. Hofmann et al. (2007) demonstrated that sucrose supply to *H. schachtii*-induced syncytia depends on the apoplasmic pathway in the early stage during syncytium formation and on the symplasmic pathway in the later stage when syncytia are linked to the phloem. The change of a series of enzymes during a metabolic process may explain various enzyme expression patterns in different SCN–soybean interaction systems both spatially and temporally.

CONCLUSION

In conclusion, this study represents for the first time a full-length transcriptome sequencing comparison between compatible and incompatible reactions on the same soybean genome to *H. glycine*. Stress response elements, plant pathogen interaction pathway, plant hormone signaling transduction pathway, and transcription factors contributed to plant immune response. The related genes associated with cell wall modification and carbohydrate metabolism played critical roles in nematode susceptibility. The phenylpropanoid biosynthesis pathway was enriched by two nematode infections. For the first time, a model of WRKY–VQ interaction resulting in plant defense response to nematode infection in an incompatible reaction was established. The identified AS events, APA, and lncRNA will provide insights into the function of post transcriptional modification during plant–nematode interaction. The knowledge of SCN–soybean interaction will aid us to understand the evolution of resistance and susceptibility, and further functional studies will help to explore new control strategies against nematodes.

DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in online repositories. The name of the repository and accession number can be found below: NCBI; PRJNA803218.

AUTHOR CONTRIBUTIONS

CW and CL conceived and designed the study, and wrote, reviewed, and edited the manuscript. MH, YJ, RQ, DJ, and DC performed the laboratory work and conducted the data analysis. MH and YJ wrote the original draft. ZT was a curator of plant material. CW and MH validated the data and analysis. All authors read and approved the final manuscript.

FUNDING

This study was supported by the Strategic Priority Research Program of the Chinese Academy of Sciences (XDA24010307) and the National Natural Science Foundation of China (31772139) to CW.

ACKNOWLEDGMENTS

We thank Philip A Roberts (University of California, Riverside, CA, United States) for the helpful comments on the manuscript.

REFERENCES

- Acharya, K., Tande, C., and Byamukama, E. (2016). Determination of *Heterodera glycines* virulence phenotypes occurring in South Dakota. *Plant Dis.* 100, 2281–2286. doi: 10.1094/PDIS-04-16-0572-RE
- Ali, M. A., Abbas, A., Kreil, D. P., and Bohlmann, H. (2013). Overexpression of the transcription factor RAP2.6 leads to enhanced callose deposition in syncytia and enhanced resistance against the beet cyst nematode *Heterodera schachtii* in *Arabidopsis* roots. *BMC Plant Biol.* 13:47. doi: 10.1186/1471-2229-13-47
- Alves, M. S., Dadalto, S. P., Gonçalves, A. B., de Souza, G. B., Barros, V. A., and Fietto, L. G. (2014). Transcription factor functional protein-protein interactions. *Proteomes* 2, 85–106. doi: 10.3390/proteomes2010085
- Anders, S., and Huber, W. (2010). Differential expression analysis for sequence count data. *Genome Biol.* 11:R106. doi: 10.1186/gb-2010-11-10-r106
- Andreasson, E., Jenkins, T., Brodersen, P., Thorgrimsen, S., Petersen, N. H., Zhu, S., et al. (2005). The MAP kinase substrate MSK1 is a regulator of plant defense responses. *EMBO J.* 24, 2579–2589. doi: 10.1038/sj.emboj.7600737
- Apweiler, R., Bairoch, A., Wu, C. H., Barker, W. C., Boeckmann, B., Ferro, S., et al. (2004). UniProt: the Universal Protein Knowledgebase. *Nucleic Acids Res.* 32, D115–D119. doi: 10.1093/nar/gkh131
- Ashburner, M., Ball, C. A., Blake, J. A., Botstein, D., Butler, H., Cherry, J. M., et al. (2000). Gene ontology: tool for the unification of biology. *Nat. Genet.* 25, 25–29. doi: 10.1038/75556
- Bailey, T. L., Williams, N., Misleh, C., and Li, W. W. (2006). MEME: discovering and analyzing DNA and protein sequence motifs. *Nucleic Acids Res.* 34, 369–373. doi: 10.1093/nar/gkl198
- Bandara, A. Y., Weerasooriya, D. K., Bradley, C. A., Allen, T. W., and Esker, P. D. (2020). Dissecting the economic impact of soybean diseases in the United States over two decades. *PLoS One* 15:e0231141. doi: 10.1371/journal.pone.0231141
- Bedre, R., Irigoyen, S., Schaker, P. D. C., Monteiro-Vitorello, C. B., Da Silva, J. A., and Mandadi, K. K. (2019). Genome-wide alternative splicing landscapes modulated by biotrophic sugarcane smut pathogen. *Sci. Rep.* 9:8876. doi: 10.1038/s41598-019-45184-1
- Bencke-Malato, M., Cabreira, C., Wiebke-Strohm, B., Bucker-Neto, L., Mancini, E., Osorio, M. B., et al. (2014). Genome-wide annotation of the soybean WRKY family and functional characterization of genes involved in response to *Phakopsora pachyrhizi* infection. *BMC Plant Biol.* 14:236. doi: 10.1186/s12870-014-0236-0

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fpls.2022.866322/full#supplementary-material>

Supplementary Table S1 | Primer sequences for qRT-PCR.

Supplementary Table S2 | Statistics of clean data.

Supplementary Table S3 | All fusion gene list for different samples.

Supplementary Table S4 | Summary of SSR analysis.

Supplementary Table S5 | All identified transcription factors.

Supplementary Table S6 | Novel isoform functional annotation.

Supplementary Table S7 | Novel gene functional annotation.

Supplementary Table S8 | Number of annotated DEGs and DETs.

Supplementary Table S9 | Top DEG-GO annotation of stress response element comparison between CK vs SCN4 and CK vs SCN5.

Supplementary Table S10 | Top DEG-KEGG pathway comparison between CK vs SCN4 and CK vs SCN5.

- Benjamini, Y., and Hochberg, Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J. R. Stat. Soc. Series B. Stat. Methodol.* 57, 289–300. doi: 10.1111/j.2517-6161.1995.tb02031.x
- Budak, H., Kaya, S. B., and Cagirici, H. B. (2020). Long non-coding RNA in plants in the era of reference sequences. *Front. Plant Sci.* 11:276. doi: 10.3389/fpls.2020.00276
- Byrd, D. W., Kirkpatrick, T., and Barker, K. R. (1983). An improved technique for clearing and staining plant tissues for detection of nematodes. *J. Nematol.* 15, 142–143. doi: 10.1007/11527503_12
- Cao, J., Ye, C., Hao, G., Dabney-Smith, C., Hunt, A. G., and Li, Q. Q. (2019). Root hair single cell type specific profiles of gene expression and alternative polyadenylation under cadmium stress. *Front. Plant Sci.* 10:589. doi: 10.3389/fpls.2019.00589
- Chakrabarti, M., de Lorenzo, L., Abdel-Ghany, S. E., Reddy, A., and Hunt, A. G. (2020). Wide-ranging transcriptome remodelling mediated by alternative polyadenylation in response to abiotic stresses in Sorghum. *Plant J.* 102, 916–930. doi: 10.1111/tpj.14671
- Chen, W., Provart, N. J., Glazebrook, J., Katagiri, F., Chang, H. S., Eulgem, T., et al. (2002). Expression profile matrix of *Arabidopsis* transcription factor genes suggests their putative functions in response to environmental stresses. *Plant Cell* 14, 559–574. doi: 10.1105/tpc.010410
- Cheng, N. H., Liu, J. Z., Liu, X., Wu, Q., Thompson, S. M., Lin, J., et al. (2011). *Arabidopsis* monothiol glutaredoxin, AtGRXS17, is critical for temperature-dependent postembryonic growth and development via modulating auxin response. *J. Biol. Chem.* 286, 20398–20406. doi: 10.1074/jbc.M110.201707
- Cheng, Y., Zhou, Y., Yang, Y., Chi, Y. J., Zhou, J., Chen, J. Y., et al. (2012). Structural and functional analysis of VQ motif-containing proteins in *Arabidopsis* as interacting proteins of WRKY transcription factors. *Plant Physiol.* 159, 810–825. doi: 10.1104/pp.112.196816
- Chi, Y., Yang, Y., Zhou, Y., Zhou, J., Fan, B., Yu, J. Q., et al. (2013). Protein-protein interactions in the regulation of WRKY transcription factors. *Mol. Plant* 6, 287–300. doi: 10.1093/mp/sst026
- Clavijo, B. J., Venturini, L., Schudoma, C., Accinelli, G. G., Kaithakottil, G., Wright, J., et al. (2017). An improved assembly and annotation of the allohexaploid wheat genome identifies complete families of agronomic genes and provides genomic evidence for chromosomal translocations. *Genome Res.* 27, 885–896. doi: 10.1101/gr.217117.116
- Cook, D. E., Bayless, A. M., Wang, K., Guo, X., Song, Q., Jiang, J., et al. (2014). Distinct copy number, coding sequence, and locus methylation patterns

- underlie *Rhg1*-mediated soybean resistance to soybean cyst nematode. *Plant Physiol.* 165, 630–647. doi: 10.1104/pp.114.235952
- Cook, D. E., Lee, T. G., Guo, X., Melito, S., Wang, K., Bayless, A. M., et al. (2012). Copy number variation of multiple genes at *Rhg1* mediates nematode resistance in soybean. *Science* 338, 1206–1209. doi: 10.1126/science.1228746
- Couturier, J., Wu, H. C., Dhalleine, T., Pégeot, H., Sudre, D., Gualberto, J. M., et al. (2014). Monothiol glutaredoxin–BoLA interactions: redox control of *Arabidopsis thaliana* BoLA2 and SufE1. *Mol. Plant* 7, 187–205. doi: 10.1093/mp/sst156
- Cui, H., Tsuda, K., and Parker, J. E. (2015). Effector-triggered immunity: from pathogen perception to robust defense. *Annu. Rev. Plant Biol.* 66, 487–511. doi: 10.1146/annurev-arplant-050213-040012
- Cui, J., Shen, N., Lu, Z., Xu, G., Wang, Y., and Jin, B. (2020). Analysis and comprehensive comparison of PacBio and nanopore-based RNA sequencing of the *Arabidopsis* transcriptome. *Plant Methods* 16:85. doi: 10.1186/s13007-020-00629-x
- Dangl, J. L., and Jones, J. D. G. (2001). Plant pathogens and integrated defence responses to infection. *Nature* 411, 826–833. doi: 10.1038/35081161
- Deepak, S., Shailasree, S., Kini, R. K., Muck, A., Mithöfer, A., and Shetty, S. H. (2010). Hydroxyproline-rich glycoproteins and plant defense. *J. Phytopathol.* 158, 585–593. doi: 10.1111/j.1439-0434.2010.01669.x
- Deng, Y. Y., Li, J. Q., Wu, S. F., Zhu, Y. P., Chen, Y. W., and He, F. C. (2006). Integrated NR database in protein annotation system and its localization. *Comput. Eng.* 32, 71–74. doi: 10.1109/INFOCOM.2006.241
- Dinesh-Kumar, S. P., and Baker, B. J. (2000). Alternatively spliced *N* resistance gene transcripts: their possible role in tobacco mosaic virus resistance. *Proc. Natl. Acad. Sci. U.S.A.* 97, 1908–1913. doi: 10.1073/pnas.020367497
- Dixon, R. A., Achnine, L., Kota, P., Liu, C. J., Reddy, M. S. S., and Wang, L. (2002). The phenylpropanoid pathway and plant defence—a genomics perspective. *Mol. Plant Pathol.* 3, 371–390. doi: 10.1046/j.1364-3703.2002.00131.x
- Edens, R. M., Anand, S. C., and Bolla, R. I. (1995). Enzymes of the phenylpropanoid pathway in soybean infected with *Meloidogyne incognita* or *Heterodera glycines*. *J. Nematol.* 27, 292–303. doi: 10.1006/jjpa.1995.1090
- Eitas, T. K., and Dangl, J. L. (2010). NB-LRR proteins: pairs, pieces, perception, partners, and pathways. *Curr. Opin. Plant Biol.* 13, 472–477. doi: 10.1016/j.pbi.2010.04.007
- Finn, R. D., Bateman, A., Clements, J., Coggill, P., Eberhardt, R. Y., Eddy, S. R., et al. (2014). Pfam: the protein families database. *Nucleic Acids Res.* 42, D222–D230. doi: 10.1093/nar/gkt1223
- Finotello, F., and Di Camillo, B. (2015). Measuring differential gene expression with RNA-seq: challenges and strategies for data analysis. *Brief. Funct. Genomics* 14, 130–142. doi: 10.1093/bfpg/elu035
- Foissac, S., and Sammeth, M. (2007). ASTALAVISTA: dynamic and flexible analysis of alternative splicing events in custom gene datasets. *Nucleic Acids Res.* 35, W297–W299. doi: 10.1093/nar/gkm311
- Haas, B. J., Papanicolaou, A., Yassour, M., Grabherr, M., Blood, P. D., Bowden, J., et al. (2013). De novo transcript sequence reconstruction from RNA-sequencing the trinity platform for reference generation and analysis. *Nat. Protoc.* 8, 1494–1512. doi: 10.1038/nprot.2013.084
- Hirao, T., Fukatsu, E., and Watanabe, A. (2012). Characterization of resistance to pine wood nematode infection in *Pinus thunbergii* using suppression subtractive hybridization. *BMC Plant Biol.* 12:13. doi: 10.1186/1471-2229-12-13
- Hofmann, J., and Grundler, F. M. W. (2008). Starch as a sugar reservoir for nematode-induced syncytia. *Plant Signal. Behav.* 3, 961–962. doi: 10.4161/psb.6075
- Hofmann, J., Hess, P. H., Szakasits, D., Blöchl, A., Wiczorek, K., Daxböck-Horvath, S., et al. (2009). Diversity and activity of sugar transporters in nematode-induced root syncytia. *J. Exp. Bot.* 60, 3085–3095. doi: 10.1093/jxb/erp138
- Hofmann, J., Wiczorek, K., Blöchl, A., and Grundler, F. M. W. (2007). Sucrose supply to nematode-induced syncytia depends on the apoplasmic and symplasmic pathways. *J. Exp. Bot.* 58, 1591–1601. doi: 10.1093/jxb/erl285
- Hua, C., Li, C. J., Hu, Y. F., Mao, Y. Z., You, J., Wang, M. Z., et al. (2018). Identification of HG types of soybean cyst nematode *Heterodera glycines* and resistance screening on soybean genotypes in Northeast China. *J. Nematol.* 50, 41–50. doi: 10.21307/jofnem-2018-007
- Huang, M., Qin, R., Li, C., Liu, C., Jiang, Y., Yu, J., et al. (2021). Transgressive resistance to *Heterodera glycines* in chromosome segment substitution lines derived from susceptible soybean parents. *Plant Genome* 14:e20091. doi: 10.1002/tpg2.20091
- Huang, M., Qin, R., Li, C., Wang, M., Jiang, Y., Yu, J., et al. (2022). Response of soybean genotypes from Northeast China to *Heterodera glycines* races 4 and 5, and characterisation of *rhg1* and *Rhg4* genes for soybean resistance. *Nematology* 24, 333–345. doi: 10.1163/15685411-bja10134
- Ithal, N., Recknor, J., Nettleton, D., Hearne, L., Maier, T., Baum, T. J., et al. (2007). Parallel genome-wide expression profiling of host and pathogen during soybean cyst nematode infection of soybean. *Mol. Plant Microbe Interact.* 20, 293–305. doi: 10.1094/MPMI-20-3-0293
- Jiang, H. P., Bu, F. S., Tian, L. Z., Sun, Q. X., Bao, D. F., Zhao, X., et al. (2020). RNA-seq-based identification of potential resistance mechanism against the soybean cyst nematode (*Heterodera glycines*) HG Type 0 in soybean (*Glycine max*) cv. Dongnong L-204. *Crop Pasture Sci.* 71, 539–551. doi: 10.1071/CP20060
- John, S., Olas, J. J., and Mueller-Roeber, B. (2021). Regulation of alternative splicing in response to temperature variation in plants. *J. Exp. Bot.* 72, 6150–6163. doi: 10.1093/jxb/erab232
- Kandath, P. K., Ithal, N., Recknor, J., Maier, T., Nettleton, D., Baum, T. J., et al. (2011). The Soybean *Rhg1* locus for resistance to the soybean cyst nematode *Heterodera glycines* regulates the expression of a large number of stress- and defense-related genes in degenerating feeding cells. *Plant Physiol.* 155, 1960–1975. doi: 10.1104/pp.110.167536
- Kanehisa, M., Goto, S., Kawashima, S., Okuno, Y., and Hattori, M. (2004). The KEGG resource for deciphering the genome. *Nucleic Acids Res.* 32, D277–D280. doi: 10.1093/nar/gkh063
- Kang, W., Zhu, X., Wang, Y., Chen, L., and Duan, Y. (2018). Transcriptomic and metabolomic analyses reveal that bacteria promote plant defense during infection of soybean cyst nematode in soybean. *BMC Plant Biol.* 18:86. doi: 10.1186/s12870-018-1302-9
- Kasai, T., Inoue, M., Koshiba, S., Yabuki, T., Aoki, M., Nunokawa, E., et al. (2004). Solution structure of a BoLA-like protein from *Mus musculus*. *Protein Sci.* 13, 545–548. doi: 10.1110/ps.03401004
- Khoei, M. A., Karimi, M., Karamian, R., Amini, S., and Soorni, A. (2021). Identification of the complex interplay between nematode-related lncRNAs and their target genes in *Glycine max* L. *Front. Plant Sci.* 12:779597. doi: 10.3389/fpls.2021.779597
- Kiss, T. (2001). Small nucleolar RNA-guided post-transcriptional modification of cellular RNAs. *EMBO J.* 20, 3617–3622. doi: 10.1093/emboj/20.14.3617
- Klink, V. P., and Matthews, B. F. (2009). Emerging approaches to broaden resistance of soybean to soybean cyst nematode as supported by gene expression studies. *Plant Physiol.* 151, 1017–1022. doi: 10.1104/pp.109.144006
- Klink, V. P., Overall, C. C., Alkharouf, N. W., MacDonald, M. H., and Matthews, B. F. (2007). A time-course comparative microarray analysis of an incompatible and compatible response by *Glycine max* (soybean) to *Heterodera glycines* (soybean cyst nematode) infection. *Planta* 226, 1423–1447. doi: 10.1007/s00425-007-0581-4
- Kong, L., Zhang, Y., Ye, Z., Liu, X., Zhao, S., Wei, L., et al. (2007). CPC: assess the protein-coding potential of transcripts using sequence features and support vector machine. *Nucleic Acids Res.* 36, W345–W349. doi: 10.1093/nar/gkm391
- Koonin, E. V., Fedorova, N. D., Jackson, J. D., Jacobs, A. R., Krylov, D. M., Makarova, K. S., et al. (2004). A comprehensive evolutionary classification of proteins encoded in complete eukaryotic genomes. *Genome Biol.* 5:R7. doi: 10.1186/gb-2004-5-2-r7
- Lai, Z., Li, Y., Wang, F., Cheng, Y., Fan, B., Yu, J. Q., et al. (2011). *Arabidopsis* sigma factor binding proteins are activators of the WRKY33 transcription factor in plant defense. *Plant Cell* 23, 3824–3841. doi: 10.1105/tpc.111.090571
- Lalitha, S. (2000). Primer premier 5. *Biotech. Softw. Internet Rep.* 1, 270–272. doi: 10.1089/152791600459894
- Li, C., Xiang, X., Huang, Y., Zhou, Y., An, D., Dong, J., et al. (2020). Long-read sequencing reveals genomic structural variations that underlie creation of quality protein maize. *Nat. Commun.* 11:17. doi: 10.1038/s41467-019-14023-2
- Li, X., Qin, R., Du, Q., Cai, L., Hu, D., Du, H., et al. (2020). Knockdown of *GmVQ58* encoding a VQ motif-containing protein enhances soybean resistance to the common cutworm (*Spodoptera litura* Fabricius). *J. Exp. Bot.* 71, 3198–3210. doi: 10.1093/jxb/eraa095
- Li, H. (2018). Minimap2: pairwise alignment for nucleotide sequences. *Bioinformatics* 34, 3094–3100. doi: 10.1093/bioinformatics/bty191

- Li, J. W., Ma, W., Zeng, P., Wang, Z. Y., Geng, B., Yang, J. C., et al. (2015). LncTar: a tool for predicting the RNA targets of long noncoding RNAs. *Brief. Bioinform.* 16:806. doi: 10.1093/bib/bbu048
- Li, S., Chen, Y., Zhu, X., Wang, Y., Jung, K. H., Chen, L., et al. (2018). The transcriptomic changes of Huipizhi Heidou (*Glycine max*), a nematode-resistant black soybean during *Heterodera glycines* race 3 infection. *J. Plant Physiol.* 220, 96–104. doi: 10.1016/j.jplph.2017.11.001
- Li, X., Wang, X., Zhang, S., Liu, D., Duan, Y., and Dong, W. (2012). Identification of soybean microRNAs involved in soybean cyst nematode infection by deep sequencing. *PLoS One* 7:e39650. doi: 10.1371/journal.pone.0039650
- Li, Y., Qi, X., Chang, R., and Qiu, L. (2011). "Evaluation and utilization of soybean germplasm for resistance to cyst nematode in China," in *Soybean-Molecular Aspects of Breeding*, ed. A. Sudaric (London: InTech Open), 373–396. doi: 10.5772/14379
- Li, X., Wang, X., Zhang, S., Liu, D., Duan, Y., and Dong, W. (2011). Comparative profiling of the transcriptional response to soybean cyst nematode infection of soybean roots by deep sequencing. *Chin. Sci. Bull.* 56:1904. doi: 10.1007/s11434-011-4510-3
- Liu, S., Kandoth, P. K., Lakhssassi, N., Kang, J., Colantonio, V., and Heinz, R. (2017). The soybean *GmSNAP18* gene underlies two types of resistance to soybean cyst nematode. *Nat. Commun.* 8:14822. doi: 10.1038/ncomms14822
- Liu, S., Kandoth, P. K., Warren, S. D., Yeckel, G., Heinz, R., and Alden, J. (2012). A soybean cyst nematode resistance gene points to a new mechanism of plant resistance to pathogens. *Nature* 492, 256–260. doi: 10.1038/nature11651
- Mandadi, K. K., and Scholthof, K. B. (2015). Genome-wide analysis of alternative splicing landscapes modulated during plant-virus interactions in *Brachypodium distachyon*. *Plant Cell* 27, 71–85. doi: 10.1105/tpc.114.133991
- Mao, X. Z., Cai, T., Olyarchuk, J. G., and Wei, L. P. (2005). Automated genome annotation and pathway identification using the KEGG Orthology (KO) as a controlled vocabulary. *Bioinformatics* 21, 3787–3793. doi: 10.1093/bioinformatics/bti430
- Martín, G., Márquez, Y., Mantica, F., Duque, P., and Irimia, M. (2021). Alternative splicing landscapes in *Arabidopsis thaliana* across tissues and stress conditions highlight major functional differences with animals. *Genome Biol.* 22:35. doi: 10.1186/s13059-020-02258-y
- Matsukura, S., Mizoi, J., Yoshida, T., Todaka, D., Ito, Y., Maruyama, K., et al. (2010). Comprehensive analysis of rice *DREB2*-type genes that encode transcription factors involved in the expression of abiotic stress-responsive genes. *Mol. Genet. Genom.* 283, 185–196. doi: 10.1007/s00438-009-0506-y
- Matsye, P. D., Kumar, R., Hosseini, P., Jones, C. M., Tremblay, A., Alkharouf, N. W., et al. (2011). Mapping cell fate decisions that occur during soybean defense responses. *Plant Mol. Biol.* 77, 513–528. doi: 10.1007/s11103-011-9828-3
- Mazarei, M., Liu, W., Al-Ahmad, H., Arelli, P. R., Pantalone, V. R., and Stewart, C. N. Jr. (2011). Gene expression profiling of resistant and susceptible soybean lines infected with soybean cyst nematode. *Theor. Appl. Genet.* 123, 1193–1206. doi: 10.1007/s00122-011-1659-8
- Mckenna, A., Hanna, M., Banks, E., Sivachenko, A., Cibulskis, K., Kernytzky, A., et al. (2010). The Genome analysis toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res.* 20, 1297–1303. doi: 10.1101/gr.107524.110
- Mine, A., Seyfferth, C., Kracher, B., Berens, M. L., Becker, D., and Tsuda, K. (2018). The defense phytohormone signaling network enables rapid, high-amplitude transcriptional reprogramming during effector-triggered immunity. *Plant Cell* 30, 1199–1219. doi: 10.1105/tpc.17.00970
- Miraeiz, E., Chaiprom, U., Afsharifard, A., Karegar, A., Drnevich, J. M., and Hudson, M. E. (2020). Early transcriptional responses to soybean cyst nematode HG Type 0 show genetic differences among resistant and susceptible soybeans. *Theor. Appl. Genet.* 133, 87–102. doi: 10.1007/s00122-019-03442-w
- Mitchum, M. G., and Baum, T. J. (2008). "Genomics of the soybean cyst nematode-soybean interaction," in *Genetics and Genomics of Soybean*, ed. G. Stacey (New York, NY: Springer), 321–341. doi: 10.1007/978-0-387-72299-3_17
- Mitchum, M. G., Wrather, J. A., Heinz, R. D., Shannon, J. G., and Danekas, G. (2007). Variability in distribution and virulence phenotypes of *Heterodera glycines* in Missouri during 2005. *Plant Dis.* 91, 1473–1476. doi: 10.1094/PDIS-91-11-1473
- Monaghan, J., and Zipfel, C. (2012). Plant pattern recognition receptor complexes at the plasma membrane. *Curr. Opin. Plant Biol.* 15, 349–357. doi: 10.1016/j.pbi.2012.05.006
- Mosler, A., Aller, I., Wagner, S., Nietzel, T., Przybyla-Toscano, J., Mühlhoff, U., et al. (2015). The mitochondrial monothiol glutaredoxin S15 is essential for iron-sulfur protein maturation in *Arabidopsis thaliana*. *Proc. Natl. Acad. Sci. U.S.A.* 112, 13735–13740. doi: 10.1073/pnas.1510835112
- Naveed, Z. A., Wei, X., Chen, J., Mubeen, H., and Ali, G. S. (2020). The PTI to ETI continuum in phytophthora-plant interactions. *Front. Plant Sci.* 11:593905. doi: 10.3389/fpls.2020.593905
- Neupane, S., Mathew, F. M., Varenhorst, A. J., and Nepal, M. P. (2019). Transcriptome profiling of interaction effects of soybean cyst nematodes and soybean aphids on soybean. *Sci. Data* 6:133. doi: 10.1038/s41597-019-0140-4
- Niblack, T. L., Colgrove, A. L., Colgrove, K., and Bond, J. P. (2008). Shift in virulence of soybean cyst nematode is associated with use of resistance from PI 88788. *Plant Health Prog.* 9:29. doi: 10.1094/PHP-2008-0118-01-RS
- Niblack, T. L., Lambert, K. N., and Tylka, G. L. (2006). A model plant pathogen from the kingdom animalia: *Heterodera glycines*, the soybean cyst nematode. *Annu. Rev. Phytopathol.* 44, 283–303. doi: 10.1146/annurev.phyto.43.040204.140218
- Pandey, V. P., Awasthi, M., Singh, S., Tiwari, S., and Dwivedi, U. N. (2017). A comprehensive review on function and application of plant peroxidases. *Anal. Biochem.* 6:308. doi: 10.4172/2161-1009.100030
- Perruc, E., Charpentreau, M., Ramirez, B. C., Jauneau, A., Galaud, J. P., Ranjeva, R., et al. (2004). A novel calmodulin-binding protein functions as a negative regulator of osmotic stress tolerance in *Arabidopsis thaliana* seedlings. *Plant J.* 38, 410–420. doi: 10.1111/j.1365-313X.2004.02062.x
- Puthoff, D. P., Ehrenfried, M. L., Vinyard, B. T., and Tucker, M. L. (2007). GeneChip profiling of transcriptional responses to soybean cyst nematode, *Heterodera glycines*, colonization of soybean roots. *J. Exp. Bot.* 58, 3407–3418. doi: 10.1093/jxb/erm211
- Qin, L., Wang, M., Zuo, J., Feng, X., Liang, X., Wu, Z., et al. (2015). Cytosolic BoLA plays a repressive role in the tolerance against excess iron and MV-induced oxidative stress in plants. *PLoS One* 10:e0124887. doi: 10.1371/journal.pone.0124887
- Qiu, J. L., Fiil, B. K., Petersen, K., Nielsen, H. B., Botanga, C. J., Thorgrimsen, S., et al. (2008). *Arabidopsis* MAP kinase 4 regulates gene expression through transcription factor release in the nucleus. *EMBO J.* 27, 2214–2221. doi: 10.1038/emboj.2008.147
- Sadek, J., Omer, A., Hall, D., Ashour, K., and Gallouzi, I. E. (2019). Alternative polyadenylation and the stress response. *Wiley Interdiscip. Rev. RNA* 10:e1540. doi: 10.1002/wrna.1540
- Sels, J., Mathys, J., De Coninck, B. M., Cammue, B. P., and De Bolle, M. F. (2008). Plant pathogenesis-related (PR) proteins: a focus on PR peptides. *Plant Physiol. Biochem.* 46, 941–950. doi: 10.1016/j.plaphy.2008.06.011
- Severing, E. I., Van Dijk, A. D., Morabito, G., Busscher-Lange, J., Immink, R. G., and Van Ham, R. C. (2012). Predicting the impact of alternative splicing on plant MADS domain protein function. *PLoS One* 7:e30524. doi: 10.1371/journal.pone.0030524
- Shannon, P., Markiel, A., Ozier, O., Baliga, N. S., Wang, J. T., Ramage, D., et al. (2003). Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res.* 13, 2498–2504. doi: 10.1101/gr.1239303
- Singh, R. R., Chinnasri, B., De Smet, L., Haec, A., Demeestere, K., Van Cutsem, P., et al. (2019). Systemic defense activation by COS-OGA in rice against root-knot nematodes depends on stimulation of the phenylpropanoid pathway. *Plant Physiol. Biochem.* 142, 202–210. doi: 10.1016/j.plaphy.2019.07.003
- Song, H., Wang, P., Hou, L., Zhao, S., Zhao, C., Xia, H., et al. (2016). Global analysis of *WRKY* genes and their response to dehydration and salt stress in soybean. *Front. Plant Sci.* 7:9. doi: 10.3389/fpls.2016.00009
- Song, W., Qi, N., Liang, C., Duan, F., and Zhao, H. (2019). Soybean root transcriptome profiling reveals a nonhost resistant response during *Heterodera glycines* infection. *PLoS One* 14:e0217130. doi: 10.1371/journal.pone.0217130
- Stock, A. M., Robinson, V. L., and Goudreau, P. N. (2000). Two-component signal transduction. *Annu. Rev. Biochem.* 69, 183–215. doi: 10.1146/annurev.biochem.69.1.183

- Sun, L., Luo, H., Bu, D., Zhao, G., Yu, K., Zhang, C., et al. (2013). Utilizing sequence intrinsic composition to classify protein-coding and long non-coding transcripts. *Nucleic Acids Res.* 41:e166. doi: 10.1093/nar/gkt646
- Syed, N. H., Kalyana, M., Marquez, Y., Barta, A., and Brown, J. W. (2012). Alternative splicing in plants—coming of age. *Trends Plant Sci.* 17, 616–623. doi: 10.1016/j.tplants.2012.06.001
- Szczeniak, M. W., Rosikiewicz, W., and Makalowska, I. (2015). CANTATAdb: a collection of plant long non-coding RNAs. *Plant Cell Physiol.* 57:e8. doi: 10.1093/pcp/pcv201
- Tatusov, R. L., Galperin, M. Y., Natale, D. A., and Koonin, E. V. (2000). The COG database: a tool for genome scale analysis of protein functions and evolution. *Nucleic Acids Res.* 28, 33–36. doi: 10.1093/nar/28.1.33
- Thirupathi, D. (2020). SPLICE in the seeds: integration of abscisic acid and light signaling in *Arabidopsis*. *Plant Physiol.* 183, 445–446. doi: 10.1104/pp.20.00361
- Trabucco, G. M., Matos, D. A., Lee, S. J., Saathoff, A. J., and Hazen, S. P. (2013). Functional characterization of cinnamyl alcohol dehydrogenase and caffeic acid O-methyltransferase in *Brachypodium distachyon*. *BMC Biotechnol.* 13:61. doi: 10.1186/1472-6750-13-61
- Tu, Z., Shen, Y., Wen, S., Liu, H., Wei, L., and Li, H. (2021). A Tissue-Specific landscape of alternative polyadenylation, lncRNAs, TFs, and gene co-expression networks in *Liriodendron chinense*. *Front. Plant Sci.* 12:705321. doi: 10.3389/fpls.2021.705321
- Urquiaga, M. C. O., Thiebaut, F., Hemery, A. S., and Ferreira, P. C. G. (2021). From trash to luxury: the potential role of plant lncRNA in DNA methylation during abiotic stress. *Front. Plant Sci.* 11:603246. doi: 10.3389/fpls.2020.603246
- Wan, J., Vuong, T., Jiao, Y., Joshi, T., Zhang, H., Xu, D., et al. (2015). Whole-genome gene expression profiling revealed genes and pathways potentially involved in regulating interactions of soybean with cyst nematode (*Heterodera glycines* Ichinohe). *BMC Genomics* 16:148. doi: 10.1186/s12864-015-1316-8
- Wang, L., Park, H. J., Dasari, S., Wang, S., Kocher, J. P., and Li, W. (2013). CPAT: coding-potential assessment tool using an alignment-free logistic regression model. *Nucleic Acids Res.* 41:e74. doi: 10.1093/nar/gkt006
- Wang, M., Wang, P., Liang, F., Ye, Z., Li, J., Shen, C., et al. (2018). A global survey of alternative splicing in allopolyploid cotton: landscape, complexity and regulation. *New Phytol.* 217, 163–178. doi: 10.1111/nph.14762
- Wang, X., Zhang, H., Sun, G., Jin, Y., and Qiu, L. (2014). Identification of active VQ motif-containing genes and the expression patterns under low nitrogen treatment in soybean. *Gene* 543, 237–243. doi: 10.1016/j.gene.2014.04.012
- Wang, Y., Jiang, Z., Li, Z., Zhao, Y., Tan, W., Liu, Z., et al. (2019). Genome-wide identification and expression analysis of the VQ gene family in soybean (*Glycine max*). *PeerJ* 7:e7509. doi: 10.7717/peerj.7509
- Xie, S., Leung, A. W., Zheng, Z., Zhang, D., Xiao, C., Luo, R., et al. (2021). Applications and potentials of nanopore sequencing in the (epi)genome and (epi)transcriptome era. *Innovation* 2:100153. doi: 10.1016/j.xinn.2021.100153
- Xu, X., Zhang, L., Zhao, W., Fu, L., Han, Y., Wang, K., et al. (2021). Genome-wide analysis of the serine carboxypeptidase-like protein family in *Triticum aestivum* reveals *TaSCPL184-6D* is involved in abiotic stress response. *BMC Genomics* 22:350. doi: 10.1186/s12864-021-07647-6
- Yan, C., Wang, Y., Lyu, T., Hu, Z., Ye, N., Liu, W., et al. (2021). Alternative polyadenylation in response to temperature stress contributes to gene regulation in *Populus trichocarpa*. *BMC Genomics* 22:53. doi: 10.1186/s12864-020-07353-9
- Yang, J., Lv, W., Shao, L., Fu, Y., Liu, H., Yang, C., et al. (2021). PacBio and Illumina RNA sequencing identify alternative splicing events in response to cold stress in two poplar species. *Front. Plant Sci.* 12:737004. doi: 10.3389/fpls.2021.737004
- Yang, Y., Zhou, Y., Chi, Y., Fan, B., and Chen, Z. (2017). Characterization of soybean WRKY gene family and identification of soybean WRKY genes that promote resistance to soybean cyst nematode. *Sci. Rep.* 7:17804. doi: 10.1038/s41598-017-18235-8
- Yao, T., Feng, K., Xie, M., Barros, J., Tschaplinski, T. J., Tuskan, G. A., et al. (2021). Phylogenetic occurrence of the phenylpropanoid pathway and lignin biosynthesis in plants. *Front. Plant Sci.* 12:704697. doi: 10.3389/fpls.2021.704697
- Ye, C., Zhou, Q., Wu, X., Ji, G., and Li, Q. Q. (2019). Genome-wide alternative polyadenylation dynamics in response to biotic and abiotic stresses in rice. *Ecotoxicol. Environ. Saf.* 183:109485. doi: 10.1016/j.ecoenv.2019.109485
- Yeh, H. S., and Yong, J. (2016). Alternative polyadenylation of mRNAs: 3'-untranslated region matters in gene expression. *Mol. Cells* 39, 281–285. doi: 10.14348/molcells.2016.0035
- Young, M. D., Wakefield, M. J., and Smyth, G. K. (2010). Gene ontology analysis for RNA-seq: accounting for selection bias. *Genome Biol.* 11:R14. doi: 10.1186/gb-2010-11-2-r14
- Yu, Y., Wang, N., Hu, R., and Xiang, F. (2016). Genome-wide identification of soybean WRKY transcription factors in response to salt stress. *Springerplus* 5:920. doi: 10.1186/s40064-016-2647-x
- Yu, Y., Zhou, Y. F., Feng, Y. Z., He, H., Lian, J. P., Yang, Y. W., et al. (2020). Transcriptional landscape of pathogen-responsive lncRNAs in rice unveils the role of ALEX1 in jasmonate pathway and disease resistance. *Plant Biotechnol. J.* 18, 679–690. doi: 10.1111/pbi.13234
- Yuan, M., Ngou, B., Ding, P., and Xin, X. F. (2021). PTI-ETI crosstalk: an integrative view of plant immunity. *Curr. Opin. Plant Biol.* 62:102030. doi: 10.1016/j.pbi.2021.102030
- Zeng, H., Xu, L., Singh, A., Wang, H., Du, L., and Poovaiah, B. W. (2015). Involvement of calmodulin and calmodulin-like proteins in plant responses to abiotic stresses. *Front. Plant Sci.* 6:600. doi: 10.3389/fpls.2015.00600
- Zhang, G., Sun, M., Wang, J., Lei, M., Li, C., and Zhao, D. (2019). PacBio full-length cDNA sequencing integrated with RNA-seq reads drastically improves the discovery of splicing transcripts in rice. *Plant J.* 97, 296–305. doi: 10.1111/tpl.14120
- Zhang, H. Y., Kjettrup-Lovelace, S., Li, C. B., Luo, Y., Chen, L. P., and Song, B. H. (2017). Comparative RNA-seq analysis uncovers a complex regulatory network for soybean cyst nematode resistance in wild soybean (*Glycine soja*). *Sci. Rep.* 7:9699. doi: 10.1038/s41598-017-09945-0
- Zhang, J., Gu, H., Dai, H., Zhang, Z., and Miao, M. (2020). Alternative polyadenylation of the stacyose synthase gene mediates source-sink regulation in cucumber. *J. Plant Physiol.* 245:153111. doi: 10.1016/j.jplph.2019.153111
- Zhao, L., Zhang, H., Kohnen, M. V., Prasad, K., Gu, L., and Reddy, A. (2019). Analysis of transcriptome and epitranscriptome in plants using PacBio iso-Seq and Nanopore-based direct RNA sequencing. *Front. Genet.* 10:253. doi: 10.3389/fgene.2019.00253
- Zheng, Y., Jiao, C., Sun, H., Rosli, H. G., Pombo, M. A., Zhang, P., et al. (2016). iTAK: a program for genome-wide prediction and classification of plant transcription factors, transcriptional regulators, and protein kinases. *Mol. Plant* 9, 1667–1670. doi: 10.1016/j.molp.2016.09.014
- Zhou, X., Lindsay, H., and Robinson, M. D. (2014). Robustly detecting differential expression in RNA sequencing data using observation weights. *Nucleic Acids Res.* 42:e91. doi: 10.1093/nar/gku310

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.

Copyright © 2022 Huang, Jiang, Qin, Jiang, Chang, Tian, Li and Wang. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.