

RESEARCH

Open Access



Transcriptome profile of pecan scab resistant and susceptible trees from a pecan provenance collection

Jordan Brungardt¹, Yanina Alarcon^{2,3}, Jason Shiller^{2,4}, Carolyn Young^{2,5*}, Maria J. Monteros^{2,6}, Jennifer J. Randall⁷ and Clive H. Bock^{1*}

Abstract

Pecan scab is a devastating disease that causes damage to pecan (*Carya illinoensis* (Wangenh.) K. Koch) fruit and leaves. The disease is caused by the fungus *Venturia effusa* (G. Winter) and the main management practice for controlling the disease is by application of fungicides at 2-to-3-week intervals throughout the growing season. Besides disease-related yield loss, application of fungicides can result in considerable cost and increases the likelihood of fungicide resistance developing in the pathogen. Resistant cultivars are available for pecan growers; although, in several cases resistance has been overcome as the pathogen adapts to infect resistant hosts. Despite the importance of host resistance in scab management, there is little information regarding the molecular basis of genetic resistance to pecan scab.

The purpose of this study was to elucidate mechanisms of natural pecan scab resistance by analyzing transcripts that are differentially expressed in pecan leaf samples from scab resistant and susceptible trees. The leaf samples were collected from trees in a provenance collection orchard that represents the natural range of pecan in the US and Mexico. Trees in the orchard have been exposed to natural scab infections since planting in 1989, and scab ratings were collected over three seasons. Based on this data, ten susceptible trees and ten resistant trees were selected for analysis. RNA-seq data was collected and analyzed for diseased and non-diseased parts of susceptible trees as well as for resistant trees. A total of 313 genes were found to be differentially expressed when comparing resistant and susceptible trees without disease. For susceptible samples showing scab symptoms, 1,454 genes were identified as differentially expressed compared to non-diseased susceptible samples. Many genes involved in pathogen recognition, defense responses, and signal transduction were up-regulated in diseased samples of susceptible trees, whereas differentially expressed genes in pecan scab resistant samples were generally down-regulated compared to non-diseased susceptible samples.

Our results provide the first account of candidate genes involved in resistance/susceptibility to pecan scab under natural conditions in a pecan orchard. This information can be used to aid pecan breeding programs and development of biotechnology-based approaches for generating pecan cultivars with more durable scab resistance.

Keywords Pecan, Scab, RNA-seq, DEG, Fungal resistance, *Venturia*, *Carya*, Provenance

*Correspondence:

Carolyn Young
cyoung6@ncsu.edu
Clive H. Bock
clive.bock@usda.gov

Full list of author information is available at the end of the article



This is a U.S. Government work and not under copyright protection in the US; foreign copyright protection may apply 2024. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

Introduction

Pecan (*Carya illinoensis* (Wangenh.) K. Koch; *C. illinoensis*) is native to North America, and is one of the most important commercial nut trees in the United States. In a particularly good year, 2017, pecan production for the US was estimated to be worth around \$700 million [1]. The native range of pecan extends from the Mississippi River Watershed in the US to southern Mexico [2]. A provenance collection including germplasm from throughout the native range was established in the late 1980s, and is planted at the USDA-Agriculture Research Service (ARS) – Southeastern Fruit and Tree Nut Research Station (SEFTNRS) at Byron, GA [3]. The provenance collection is a unique genetic resource that can be used to characterize multiple horticulturally important traits including scab resistance [4, 5].

Pecan scab (caused by the plant pathogenic fungus *Venturia effusa* (G. Winter)) is a major constraint to pecan production in the southeastern US. Pecan scab reduces nut quality and yield [6]. This ascomycete pathogen infects leaves, nuts and twigs when weather conditions are conducive, particularly during periods of rainfall, warmth and high humidity [7]. Leaves are susceptible to the pathogen during their expansion [7], whereas nuts are susceptible throughout their development and maturation stages. Until recently *V. effusa* was believed to reproduce solely through the production of asexual conidia, which are considered the primary source of inoculum [8]. However, the recent discovery of a sexual stage of *V. effusa* has been observed through *in vitro* mating of diverse isolates by Charlton et al. [9]. Identification of the sexual stage is consistent with studies that have shown populations of *V. effusa* to be genetically diverse, heterothallic, with both mating types present in equilibrium [10, 11]. Thus, the sexual stage likely plays a role yet to be elucidated in the lifecycle and epidemic development of pecan scab. Although pecan scab may be managed through tree hedging, tree thinning, and use of fungicides, these increase on farm expenses and reduce profitability. The annual cost of fungicide application combined with the inevitable yield loss due to scab is substantial [12]. Moreover, *V. effusa* has developed resistance to several classes of fungicides, which can present challenges to effective management of the disease in production orchards (Reviewed in: [13]). Thus, understanding the mechanisms of resistance to scab that exist in a diverse collection of pecan is highly desirable to guide breeding programs and develop biotechnology approaches for durable scab resistance.

Plant responses to biotic stress are the result of complex processes comprising of pathogen detection and signal transduction mechanisms (Reviewed in: [14, 15]). The first line of defense plants use for sensing an attacking

pathogen is to detect molecules that are common to different types of microbes. These molecules are commonly referred to as pathogen/microbe associated molecular patterns (PAMP/MAMP) and are detected by pathogen recognition receptors (PRR) and kinases with transmembrane domains. This type of immunity is known as pattern triggered immunity (PTI). Pathogens are able to overcome PTI by disrupting the immune response with molecules known as effectors. In return, plants have evolved mechanisms for detecting pathogen effectors in what is known as effector triggered immunity (ETI). Among an array of different types of effector receptors, nucleotide binding leucine-rich repeat domain containing proteins (NB-LRR) are the main source of ETI resistance genes [16]. The ETI resistance genes are highly specific to each plant pathosystem. Short range signal transduction mechanisms involve different chemical pathways typically involving mitogen-activated protein (MAP) kinases [17]. Ultimately, pathogen recognition and signal transduction drive the expression of pathogenesis related genes and defense responses. The epitome of plant defense response is characterized by a hypersensitive response (HR), in which reactive-oxygen species (ROS) are produced followed by programmed cell death. Also, production of chitinases, defensins, thaumatin, and secondary metabolites are additional methods plants use to defend against pathogens [18]. Due to the intricacies present in any particular plant pathosystem, a thorough genetic investigation is required to elucidate mechanisms of resistance. Durable resistance for plants often requires multiple resistance genes, as a compatible interaction between effector and R-gene exhibits a strong negative selection on the pathogen and subsequent effector. Stacking multiple resistance genes has been reported in rice for resistance to bacterial diseases [19], and resistance to fungal disease in apple [20] and sunflower [21]. The relatively long time required for pecan trees to reach maturity and start setting fruit would make stacking R-genes a lengthy endeavor.

Pecan is a perennial, outcrossing plant species with cultivars that are propagated through grafting. Rootstocks are known to be a source of variation for different traits in several fruit trees including apple [22] and citrus [23]. One approach for identifying genes involved with pecan scab resistance/susceptibility, without the added potential of variation from rootstock, is to obtain transcriptome profiles from seed-grown trees like those present in the pecan provenance collection. Such a strategy is able to explore pecan responses to a naturally occurring scab population comprising a diverse population of *V. effusa* pathovars [24]. This in conjunction with the genetic diversity of the provenance collection should provide the framework to identify multiple resistant/susceptibility

mechanisms in the pecan/*V. effusa* pathosystem. Pecan derived from northern germplasm, as opposed to Mexico and Texas germplasm, is generally more resistant to scab. Transcriptome profiling using RNA-Sequencing (RNA-seq) captures the transcriptome profile of an individual at a given time and environmental condition. Comparing transcript profiles between multiple conditions allows one to identify differentially expressed genes (DEG). Differential expression studies have been previously reported for fruit and nut tree species for biotic stresses under field conditions [25–27]. Numerous differential expression studies have been done on *C. illinoensis* in the past, most of which address horticultural questions [28–31]. Only recently has this method been used to examine biotic stress in pecan. Chen et al. examined transcript profiles of three different pecan cultivars with varying levels of resistance to the ascomycete *Pestalotiopsis microspora* [32]. In regards to *V. effusa*, Lovell et al. used RNAseq to find DEGs in pecan after inoculation with *V. effusa* under controlled conditions; however, only a single cultivar-pathotype interaction was studied and there was no evaluation of resistant tree genotypes [33].

For the present study, a subset of pecan trees from the provenance collection that were consistently identified as scab resistant or susceptible during three years of disease assessment in the orchard were used for differential expression analysis [4, 5]. In addition, diseased and disease-free leaves from susceptible trees were collected and analyzed. RNA-Seq of leaf samples enabled transcriptome profiling of the trees for the identification of genes likely involved in scab resistance/susceptibility. Gene ontology (GO) enrichment analysis was used for the identification of functional categories of DEGs and to explore their potential role in the pecan scab pathosystem.

Materials and methods

Plant materials, scab assessments and RNA extraction

Scab evaluations and leaf samples were taken from trees in the *C. illinoensis* provenance collection located at the USDA-ARS SEFTNRS station in Byron, GA. Scab severity was evaluated on leaves during the growing seasons of 2013 and 2014 by Bock et al. [5], as well as on leaves and fruit in 2018 [4]. Based on the scab assessments, ten consistently scab resistant and ten consistently susceptible trees were selected for the current study (Table 1).

For susceptible trees, two sets of samples were taken from each tree, leaves with scab symptoms (S-D = susceptible diseased) and leaves with no scab symptoms (S-ND = susceptible no disease). Only one set of samples (*R* = resistant) was taken from resistant trees as no disease was evident. Leaf tissue (five leaflets per tree) was collected at one-third total leaf expansion from the

twenty trees early in the scab season on April 30th of 2019, when symptoms of scab were easily detected and not confounded by other diseases or pest damage that can occur late in the season. Scissors, sterilized between leaflet samples with 70% ethanol, were used to excise leaflets. Powder free gloves (a new pair for each sample, sterilized with ethanol) were worn during sampling and the samples were placed in RNAase free tubes containing Invitrogen RNeasy Lysis Buffer (ThermoFisher Scientific; MA, US) to store the leaf samples for further processing after sampling. At the time of RNA extraction, RNeasy Lysis Buffer was removed from the samples, the leaf tissue was ground in liquid nitrogen, and RNA extracted using the RNeasy Plant Total RNA Kit (Sigma Aldrich; MO, US). A Bioanalyzer 2100 (Agilent Inc.; CA, US) was used to ensure samples had an RNA integrity number greater than seven.

Pecan RNA-Seq library preparation and differential gene expression determination

RNA-seq libraries were constructed by first polyA purifying 100 ng total RNA with TruSeq® Stranded mRNA Sample Prep kits (Illumina, CA, US) following standard protocols. Sixty individual libraries (two technical replicates of ten S-ND, S-D, and R samples) were uniquely indexed using TruSeq RNA Single Indexes (Illumina, CA, US). The pooled libraries were sequenced using an Illumina NovaSeq 6000, paired-end 150 bp at Novogene Inc. (CA, US). Raw reads can be accessed through the National Center for Biotechnology's Sequence Read Archive (NCBI-SRA) as Bioproject PRJNA992068. The resulting reads were quality trimmed with Trimmomatic 0.38 using the following parameters [34].

```
ILLUMINACLIP:TruSeq3-PE.fa:2:30:10:3:TRUE
HEADCROP:17 LEADING:3 TRAILING:3 SLID-
INGWINDOW:4:24 MINLEN:100
```

The quality of sequence reads was verified with Fastqc 0.11.9 by comparing reads before and after trimming. One resistant replicate was not used for analysis due to low read quality. Reads were aligned to the Phytozome V13, 573 *C. illinoensis* 87MX3-2.11 annotation version 1.1 transcripts and primary coding sequence (CDS) transcripts [35] separately with Bowtie2 v3.4 using default parameters [36]. Sequence alignment/map files were converted to binary alignment/map files with SAMtools 1.10 [37]. Transcript abundance was then quantified using Salmon 0.10.1, correcting for GC bias [38]. Quantified read information was then imported into R [39] for differential analysis with “tximport” using default parameters [40]. Differential analysis was done with DESeq2 [41] using the “ashr” log fold-change shrinkage estimator method [42]. Comparisons made include S-ND leaves vs.

Table 1 Pecan scab severity in 2013 and 2014 on leaves (mean severity) and in 2018 on leaves and nuts (most severe and overall tree severity) for provenance trees at Byron, GA

Provenance-family (mother tree) ^a	Tree ID number	Scab severity						Overall scab response
		2013 (foliar % severity)	2014 (foliar % severity)	2018 (most severe foliar % severity)	2018 (overall tree foliar % severity)	2018 (most severe nut % severity)	2018 (overall tree nut % severity)	
TX2-2	11	12.6	1.2	10	10	60	30	Susceptible
MX4-2	18	5.2	7.3	80	5	70	60	Susceptible
TX3-4	42	3.1	6.6	50	10	90	60	Susceptible
MX1-1	46	3.4	1.4	50	2	70	40	Susceptible
TX2-2	229	3.9	1.3	70	10	100	90	Susceptible
TX5-4	266	2.4	2.2	20	1	70	20	Susceptible
TX2-5	283	2.1	5.4	70	10	90	30	Susceptible
TX3-4	505	2.2	3.5	50	5	100	10	Susceptible
TX5-5	491	1.2	1.3	30	2	60	60	Susceptible
TX3-3	571	2.4	4.5	20	2	100	90	Susceptible
KS2-3	16	0	0	0	0	0	0	Resistant
TX4-2	80	0	0	0	0	0	0	Resistant
MO3-2	116	0	0	0	0	0	0	Resistant
MO1-5	155	0	0	0	0	0	0	Resistant
TN1-1	163	0	0	0	0	0	0	Resistant
MX5-2	268	0	0	0	0	0	0	Resistant
IL1-2	423	0	0	0	0	0	0	Resistant
TN1-1	497	0	0	0	0	0	0	Resistant
KY1-3	560	0	0	0	0	0	0	Resistant
IL1-5	541	0	0	0	0	0	0	Resistant

^a Provenances are indicated as IL1 (Jersey Co., Illinois), KS2 (Cherokee Co., Kansas), KY1 (Webster Co., Kentucky), MO1 (Livingston Co., Missouri), MO3 (Vernon Co., Missouri), MX1 (Santa Catarina, San Luis Potosi, Mexico), MX4 (Ixmiquilpan, Mexico), MX5 (Jaumave, Mexico), TN1 (Lake Co., Tennessee), TX2 (Zavala Co., Texas), TX3 (Kinney Co., Texas), TX4 (Tom Green Co., Texas), TX5 (Val Verde Co., Texas) (Ref. 3)

R leaves as well as S-ND leaves vs. S-D leaves. Genes that had a Benjamini & Hochberg adjusted *p*-value less than 0.05 and an absolute log base two-fold-change (LFC) of one and a half were considered to be differentially expressed.

Venturia effusa reads in pecan samples

To estimate the relative level of *V. effusa* infection in each pecan sample, previously trimmed sample reads were mapped to the *V. effusa* FRT5LL7-albino genome (Genbank:GCA_007735645.1; Ref. [43]) using Bowtie2 v3.4. Information for percent reads mapped to the *V. effusa* genome and those previously mapped to *C. illinoensis* primary CDS transcriptome were retrieved from Bowtie2 log files. Analysis of variance and Tukey’s honestly significant difference (HSD) post-hoc tests were performed with the R base statistics package. Due to high sequence similarity of several *C. illinoensis* and *V. effusa* orthologs, levels of *V. effusa* in samples were also checked with reads mapped to the *V. effusa* high expressing, non-orthologous elongation factor EF-1 alpha transcript ([43]; Protein ID QDS71019.1).

Differential expression post hoc analysis

Subsequent analyses were done with data generated from alignments to the primary CDS transcript information. Genes that were differentially expressed in S-ND vs. R and S-ND vs. S-D analyses were visualized with the R package VennDiagram [44]. Annotation information from the Phytozome V13, 573 *C. illinoensis* 87MX3-2.11 genome version 1.1 (*C. illinoensis* 573 genome; Ref. 35) was used for enrichment analysis of DEGs with Pathview (<https://pathview.uncc.edu/>; Accessed Oct. 2022; [45]), Mapman (Data not shown; [46]), and GOrilla (<http://cbl-gorilla.cs.technion.ac.il/>; Accessed Oct. 2022; [47]). Gene ontology (GO) enrichment was done by identifying DEGs with *Arabidopsis thaliana* (*A. thaliana*) ortholog information as well as GO annotation information from the *C. illinoensis* 573 assembly annotation information. The resulting information was passed into the online GOrilla analysis tool. Unranked gene lists were used with a *p*-value threshold of 10⁻³, evaluating all ontology types. For Pathview analysis, gene set analysis (GSA) was done by first filtering Salmon normalized counts for DEGs with *A. thaliana* ortholog information. Pathview

utilizes Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways to visualize relationships between elements of specific metabolic pathways [48]. Plant pathways of particular interest to interactions with pathogens (KEGG pathways 00999, 04016, 04075, and 4626) were evaluated as well as using the automatic pathway selection feature of Pathview. R-genes were predicted from the *C. illinoensis* 573 genome with DRAGO2-API (DRAGO; [49]) and NLR-annotation v2.1 [50]. Transcription factors were predicted with DeepTFactor [51] and Plant Transcription Factor Database v 5.0 (PlantTFDB; [52]) using default settings.

To verify results generated from the next generation sequencing (NGS) information, expression of several DEGs was checked with quantitative PCR (qPCR) from one-third expanded leaves collected on May 19th, 2023. Tissue was collected from four random trees of the previously identified S-ND and R groups. Tissue was placed into 2 mL microfuge tubes, which were placed directly into liquid nitrogen. Samples were maintained at -80°C until further processing. RNA was extracted using the lignified/polyphenol-rich protocol of the GeneJET Plant RNA purification Mini Kit (ThermoFisher Scientific). Quality of RNA was checked with a Nanodrop spectrophotometer to ensure 260/280⁻¹ and 260/230⁻¹ ratios > 1.8 and integrity of RNA was checked visually via gel-electrophoresis. Potential genomic DNA (gDNA) was removed from samples by treating 1 µg of sample with RNase-free DNase I following manufacturer's recommendations (ThermoFisher Scientific) with the addition of 0.5µL RNasin Ribonuclease Inhibitor (Promega) per reaction. First strand complementary DNA (cDNA) was generated using Avian Myeloblastosis Virus reverse transcriptase and oligo(dT) primers (Promega). Presence/absence of gDNA and cDNA was confirmed using PCR with intron spanning primers (Additional file 1; tab 2 'Primers'). Each sample/gene-target combination of interest was evaluated on a Bio-Rad CFX Opus 96 Real-Time PCR system using three technical replicates. qPCR reactions used PowerUpTM SYBRTM Green master mix (Applied Biosystems; MA, US) and 0.35 µL undiluted cDNA reaction as a template for each quantitative reaction. Primer information used in qPCR reactions is provided (Additional file 1; tab 2 'Primers'). Polyadenylate-binding protein (Caril.16G041200) was chosen as a reference gene due to its small coefficient of variation in expression derived from RNAseq data generated for this study.

Results

RNAseq, mapping, and differential gene expression identification

RNA sequencing resulted in more than 2.55 billion reads being referenced to sample barcodes. Samples had a

mean paired-read count of 43.28 million (s.d. 10.47 mil.). After trimming, approximately 25% of raw reads were removed from each sample. Mapping of RNAseq data was performed separately on two *C. illinoensis* transcript reference assemblies and a *V. effusa* genome. As expected, reads mapped to a much lesser extent to the fungal genome compared to that of the pecan genome assembly (Fig. 1A & B).

In addition, the percentage of *V. effusa* reads per sample was much greater for susceptible diseased leaves (S-D; \bar{x} =1.72%) compared to non-diseased susceptible (S-ND; \bar{x} =0.76%) and resistant leaves (R; \bar{x} =0.79%) (Tukey's HSD p -value < 0.05). In fact, no reads mapped to the non-orthologous *V. effusa* elongation factor alpha gene in R and S-ND samples (Fig. 1C). For the pecan assemblies, the average number of reads mapped to the primary CDS assembly was 69.3%, whereas the average number of reads mapped to the isoform containing assembly was 87.3%. Despite this discrepancy, the composition of resulting DEGs between the two assemblies was very similar (data not shown); therefore, the following analyses were performed on DEGs derived from the primary CDS assembly only. Mapping and DEG analysis against the Phytozome V13 *Carya illinoensis* 588 cv. Pawnee v1.1 (susceptible to pecan scab) primary transcriptome also gave similar results (Ref. 33; data not shown). In addition, the expression pattern of several DEGs was similar in trees sampled at a later time when using qPCR (Fig. 2).

When comparing expression data from S-ND to S-D leaves, 1,454 DEGs were identified (Fig. 3; Additional file 1, tab 3 "DxND_DEGs").

Of these, 475 genes were down-regulated and 979 genes were up-regulated for S-D compared to S-ND leaves. In contrast, fewer genes were differentially expressed when comparing expression data of R to S-ND leaves (Fig. 3; Additional file 1, tab 4 "RxND_DEGs"). Of the 313 DEGs identified for this comparison, sixty-three were up-regulated and 250 were down-regulated in R compared to S-ND leaves. The two sets of DEGs, S-ND by S-D and S-ND by R, had thirty genes in common.

Analysis of differentially expressed genes

Resistance genes (R-genes) are important components of the immune system plants depend on for resistance to pathogens. The program NLR-annotation was able to identify 345 R-genes from the 31,911 annotated 87MX3-2.11 predicted pecan genes. In contrast, DRAGO was more sensitive, predicting 2,410 R-genes from the same gene set (Additional file 1, tab 5 "DRAGO_rGenes"). This difference could be due to DRAGO using more R-gene motifs for prediction. For example, the number of R-genes containing both NBS and LRR sequences in the

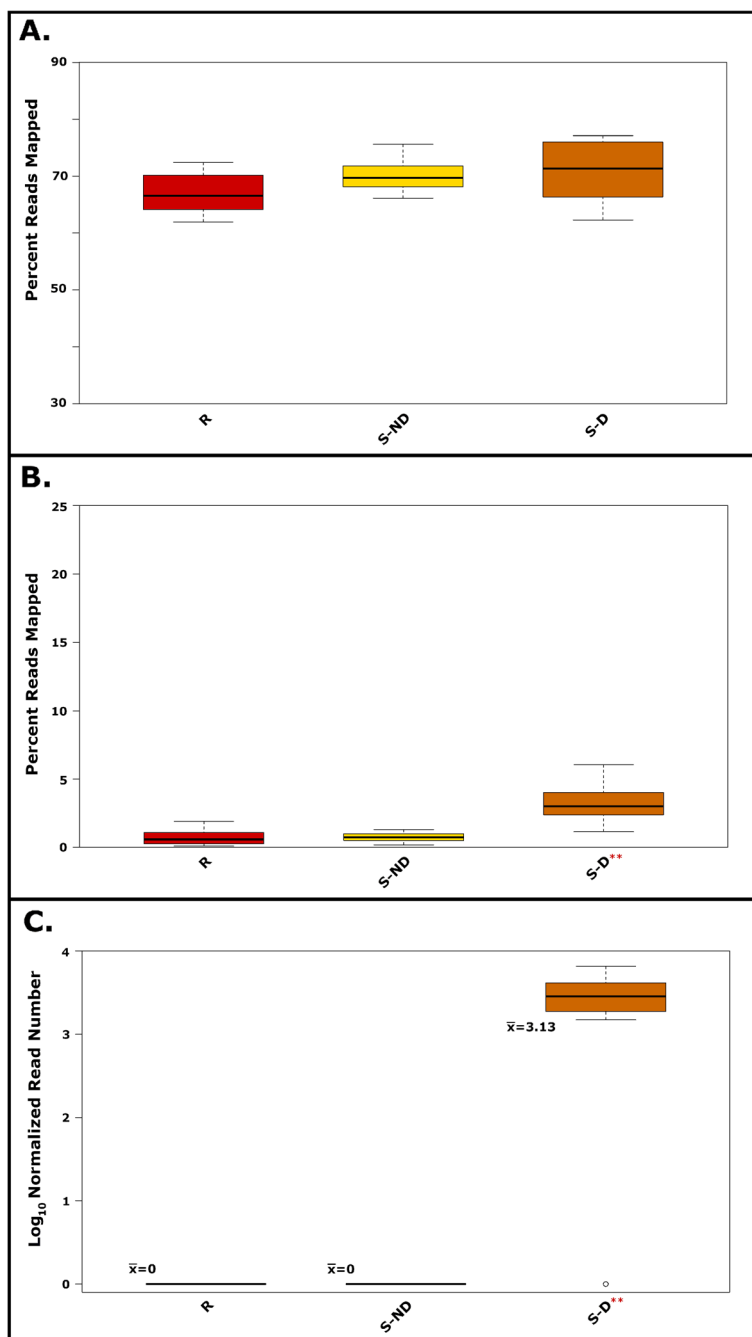


Fig. 1 Boxplots of RNAseq data mapped to either the *Carya illinoensis* transcriptome (A) or *Venturia effusa* genome (B). RNAseq data was generated from susceptible non-diseased leaves (S-ND), diseased susceptible leaves (S-D), and resistant leaves (R). The double asterisk (**) indicates a Tukey's HSD *p*-value < 0.05 for a given treatment. Reads mapped to the high expressing, non-orthologous *Venturia effusa* elongation factor one-alpha are shown (C). Whiskers represent the minimum of either 1.5 times the interquartile range or the extreme of the sample range. Open circles indicate outliers

DRAGO2 R-gene set was much closer to that of the total NLR-annotation set at 218. Subsequent results are based on DRAGO predictions as it gives more detailed information on R-gene types. Of the 313 DEGs identified for the R by S-ND comparison, thirty-eight were identified

as R-genes by DRAGO2, ten of which were up-regulated in R samples (Fig. 4A).

Kinase domain-containing proteins (KIN) were the most up-regulated type of R-gene (four genes), although the same class was more down-regulated (seven genes).

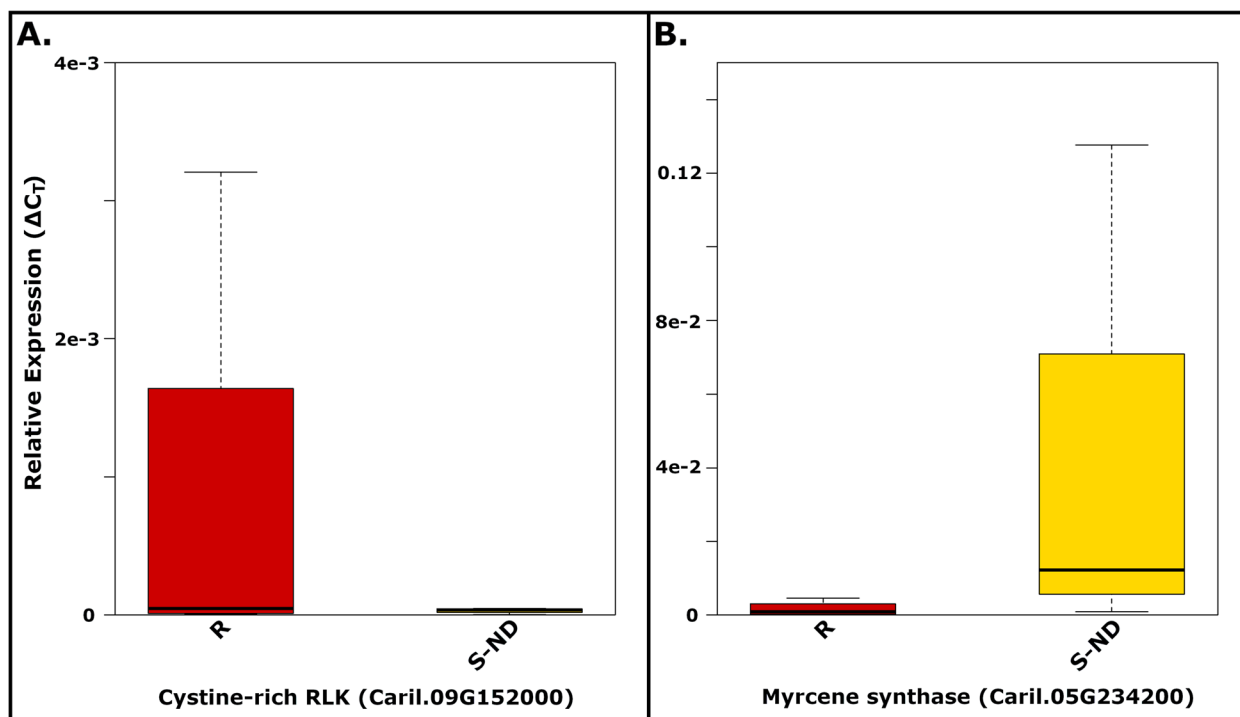


Fig. 2 Quantitative PCR showing relative expression of two genes in pecan leaves collected in 2023 ($n=4$) using polyadenylate-binding protein (Caril.16G041200) as a reference gene. Expression patterns were similar to DEGs identified in RNAseq data from samples collected in 2019, showing up-regulation of a cystine-rich receptor-like kinase (Caril.09G152000, **A**) in resistant leaves (R) compared to susceptible non-diseased leaves (S-ND), while myrcene synthase (Caril.05G234200, **B**) was down-regulated in resistant leaves (R)

The most down-regulated type of R-gene for the R by S-ND comparison were receptor-like proteins (RLP, ten genes). Classes of R-gene containing NB-LRR motifs (NL & TNL) were slightly up-regulated in R samples. For the S-ND by S-D comparison 113 of the DEGs were predicted R-genes, with eighty-three of these being up-regulated for S-D (Fig. 4B). Similar to the R by S-ND comparison, kinase domain containing proteins were the most up-regulated type of R-gene (forty-seven genes) for S-D compared to S-ND; however, relatively few were down-regulated (five genes). The most down-regulated type of R-gene for this comparison were receptor-like kinases (sixteen genes), but this class was up-regulated to the same extent.

Transcription factors (TFs) can regulate the expression of large sets of genes to control complex traits such as pathogen resistance in plants. When evaluating genes annotated for the *C. illinoensis* 87MX3-2.11 genome, PlantTFDB and DeepTFactor were able to predict 2,233 and 2,547 transcription factors respectively. The following results are based on the PlantTFDB predicted TFs, as its output gives more detailed information on TF type (Additional file 1, tab 6 “PlantTFDB_TFs”). Differential expression of TFs in comparisons evaluated reflected the respective overall differential expression for each of the

two comparisons. TFs were generally down-regulated in R compared to S-ND leaves (Fig. 5A), whereas S-D leaves compared to S-ND showed a significant amount of up-regulation in TF expression (Fig. 5B).

The largest discrepancy between the two comparisons was down-regulation in R and up-regulation in S-D of ethylene response factors when compared to S-ND samples. Also, WRKY and MYB type TFs seem to be over-represented in up-regulation for S-D when compared to S-ND.

Enriched gene ontology (GO) analysis and gene set analysis (GSA) are powerful tools for drawing conclusions from large sets of DEGs. For the R by S-ND comparison, fifteen up-regulated genes met requirements for GO analysis and 110 were used for down-regulated DEG analysis. For the S-D by S-ND comparison, 370 up-regulated and 205 down-regulated DEGs were used for GO analysis. GO enrichment was identified for both R by S-ND and S-D by S-ND comparisons. While the S-D by S-ND comparison showed enrichment for each category evaluated (“cellular component”, “molecular function”, and “biological process”) for both down- and up-regulation (Additional file 1, tab 7 “GORilla_DxND”; Additional file 2, Figures 1 & 2); the R by S-ND comparison only displayed enrichment for down-regulation of “biological

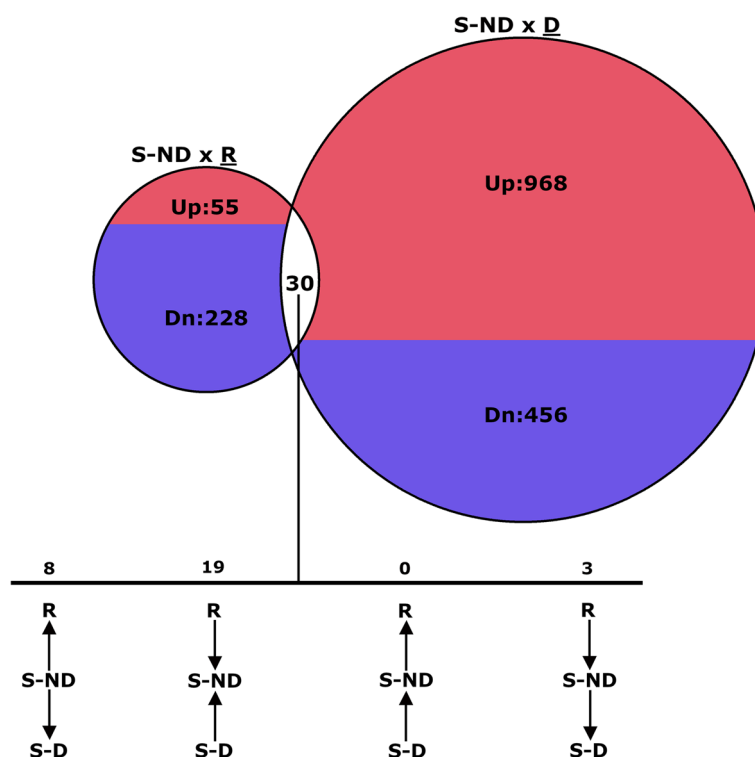


Fig. 3 Venn diagram showing number and fold-change type of differentially expressed genes (DEG) for non-diseased susceptible pecan leaves (S-ND) when compared separately to diseased (S-D) and resistant (R) leaves. Treatments compared are above their respective results. Up- (Up) or down-regulation (Dn) describes expression pattern of underlined treatment group, when compared to S-ND. Fold-change patterns for DEGs shared between comparisons is described in detail below Venn diagram with arrows indicating direction of increasing expression

process” and up-regulation of “molecular function” (Additional file 1, tab 8 “GORilla_RxND”; Additional file 2, Figure 3). The most significantly enriched term for R by S-ND was down-regulation of “hormone catabolic process” in R leaves (GO:0042447; *p*-value 6.99e-5). Surprisingly, other enriched down-regulated GO terms in R samples were “defense response to other organism” and “response to chitin”. “Pyrimidine transmembrane transporter activity” was the only enriched term for up-regulation in R samples. The most significantly enriched term for the S-D by S-ND comparison was down-regulation of “cellular component extracellular region” in S-D leaves (GO: GO:0005576; *p*-value 1.38e-8). Surprisingly, the same term had a similar significance for up-regulation in S-D (*p*-value 1.77e-8). Other terms that were found to be enriched for down-regulation in S-D include “endopeptidase activity” and “response to auxin”. GO terms up-regulated in S-D include a myriad of biotic stress responses (excluding “response to fungi”), “ethylene-activated signaling pathway”, and “transcription factor activity”.

Gene set analysis with Pathview shows more detailed information on regulation of genes within biochemical pathways compared to GO analysis. For the R by S-ND comparison, the Pathview auto-select function identified

three pathways that were enriched for differential expression. These pathways were “RNA degradation” (KEGG pathway 03018), “plant hormone signal transduction” (KEGG pathway 04075), and “plant-pathogen interaction” (KEGG pathway 04626). All pathway terms for this comparison were generally down-regulated in R compared to S-ND. DEGs in the plant-pathogen interactions pathway included WRKY25 and WRKY33 type transcription factors as well as a calcium-binding protein (CML40) (Fig. 6A).

DEGs in the plant-pathogen interaction pathway for S-D were generally up-regulated compared to S-ND samples (Fig. 6B). Pathview detected an element of the Ccr4-NOT complex to be down-regulated in the “RNA degradation” pathway as well as elements of the brassinosteroid and jasmonic acid (JA) pathways in the “plant hormone signaling” pathway. Other elements down-regulated in R samples that were not auto-selected by Pathview include genes responsible for camalexin synthesis, ethylene synthesis, defense responses, and wounding responses. All original Pathview output files can be found in Additional file 2, Figures 4-12.

Unlike the R by S-ND comparison, Pathview showed a mixture of down- and up-regulated enriched gene sets for the S-ND by S-D comparison. The auto-select feature

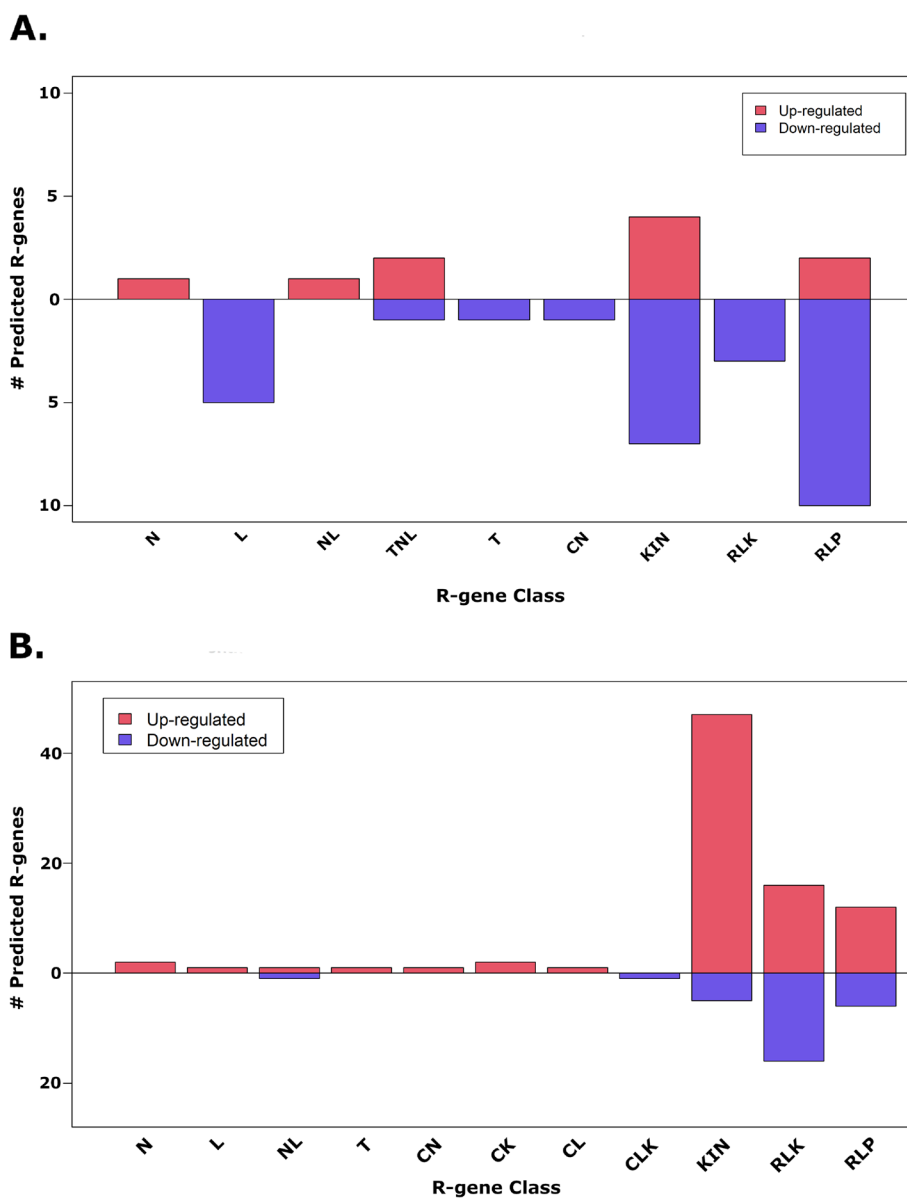


Fig. 4 Differentially expressed predicted R-genes for pecan scab related treatments. Regulation type explains fold change for (A) resistant leaves (R), and (B) diseased leaves (S-D) when compared to susceptible non-diseased leaves (S-ND). Abbreviations: N=nucleotide binding site (NBS), L=leucine rich-repeat (LRR), NL=NBS-LRR, TNL=TIR-NB-LRR, T=toll-interleukin receptor (TIR), CN=coiled-coil (CC)-NBS, CK=CC-kinase, CL=CC-LRR, CLK=CC-LRR-kinase, KIN=kinase, RLK=receptor-like kinase, RLP=receptor like protein

of Pathview reported two pathways, “plant hormone signaling” and the “starch/sucrose metabolism” pathway (KEGG pathway 00500). For the “plant hormone signaling” pathway many genes were down-regulated for auxin signaling in S-D compared to S-ND, up-regulated for gibberellin, ethylene, and jasmonic acid signaling, and perturbed for salicylic acid signaling. For the “starch/sucrose metabolism” pathway, genes were up-regulated in S-D for gene products responsible for the biosynthesis of trehalose, D-glucose-6P, and D-fructose. In contrast, many

genes responsible for the production of D-glucose and its intermediates were down-regulated. When examining plant specific pathways using Pathview, genes responsible for plant defenses were consistently up-regulated in S-D leaves. Descriptors for pathways belonging to these genes include “disease resistance”, “hypersensitive response (HR)”, “defense-related gene induction”, and “late defense responses for pathogens”. In addition to this, elements responsible for suppression of plant HR and suppression of defense responses were down-regulated.

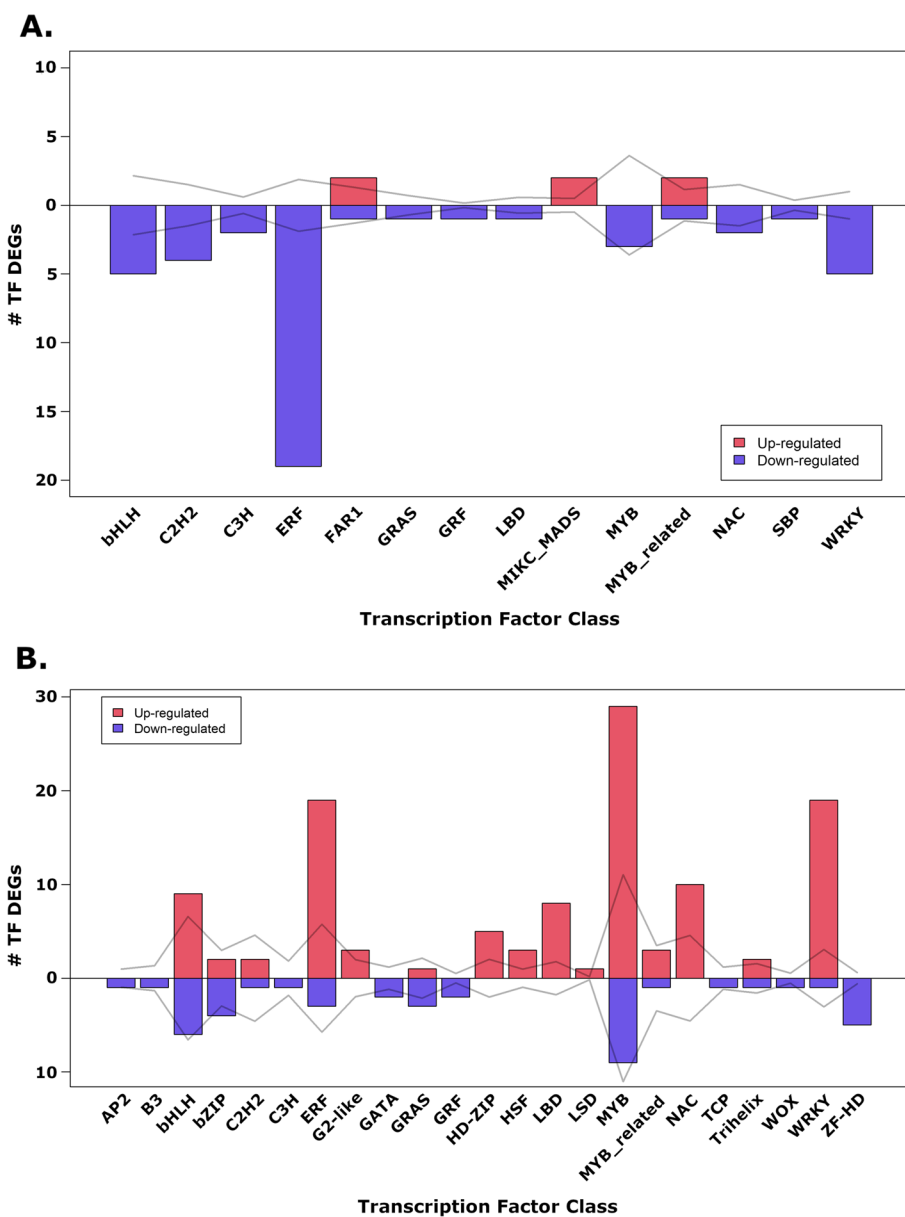


Fig. 5 Differentially expressed (DE) transcription factors (TFs) for different pecan scab treatments. Regulation type explains fold change for (A) resistant leaves (R), and (B) susceptible diseased leaves (S-D) when compared to non-diseased susceptible leaves (S-ND). Line graph shows number of expected DE TFs if chosen at random from all *Carya illinoensis* TFs as predicted with PlantTFDB v5.0 (1/2 DE TF * ratio *C. illinoensis* TF class)

Discussion

Venturia effusa is a formidable fungal pathogen of pecan. Genetic resistance to this phytopathogen occurs naturally in some pecan trees and is complicated by the presence of different pathovars. Breeding for resistance can be slow due to the prolonged time required for pecan to begin flowering and setting fruit. The research presented here aims to decipher which genetic elements are responsible for resistance/susceptibility to *V. effusa* on the basis

of gene expression. To meet this goal, the current study leverages the comprehensive genetic diversity present in the pecan provenance collection at Byron, GA, along with years of pecan scab phenotyping data against diverse natural populations of *V. effusa*. The trees evaluated in this study cover a range of the genetic diversity of *C. illinoensis* as they represent several of the provenances in the collection (Additional file 2, Figure 13), which itself accounts for 94% of the genetic diversity of the species [3].

PLANT-PATHOGEN INTERACTION

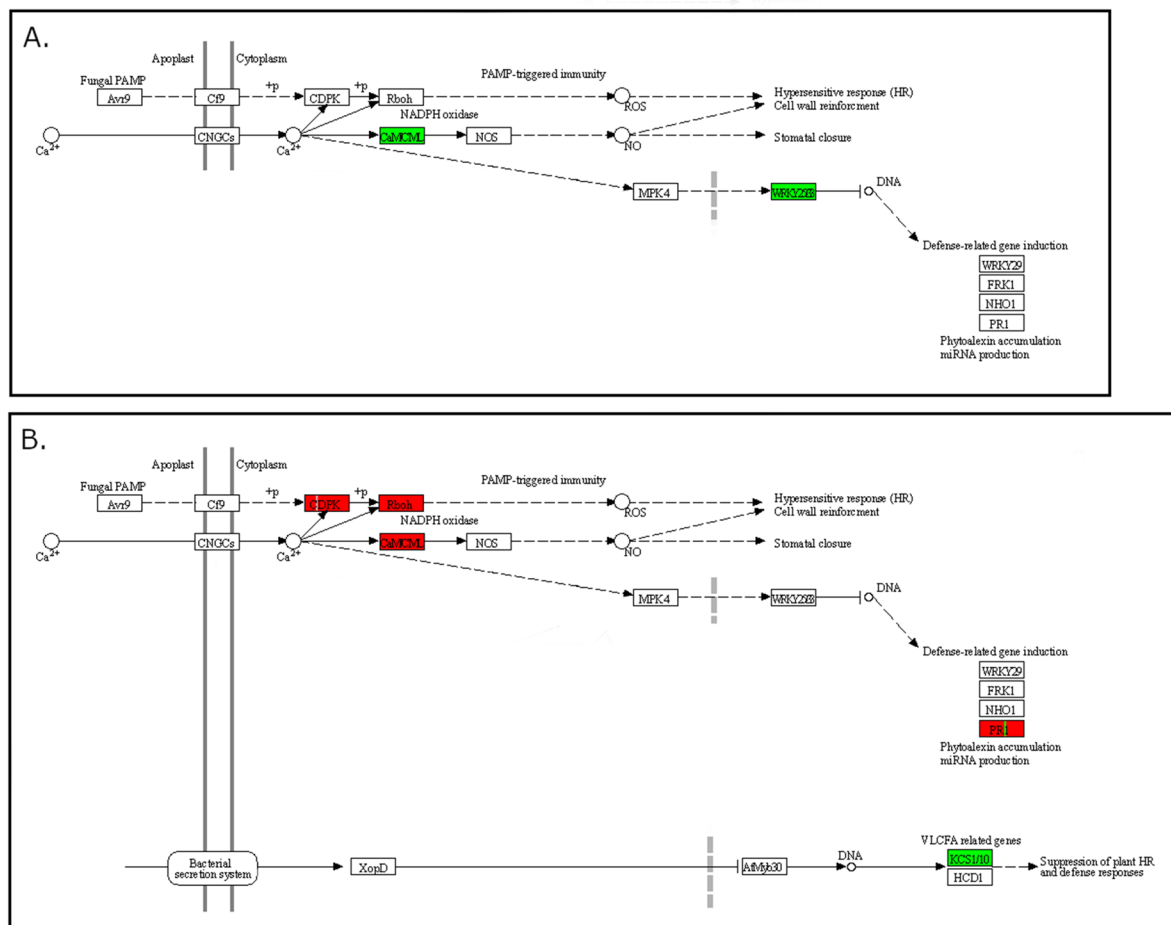
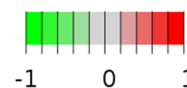


Fig. 6 Differentially expressed genes (DEGs) in the “plant-pathogen interaction” KEGG pathway 04626 for (A) pecan (*Carya illinoensis*) trees resistant to *Venturia effusa* (R) as opposed to susceptible non-diseased trees (S-ND) and (B) susceptible diseased leaves (S-D) compared to susceptible non-diseased (S-ND) leaves [48]. Significant DEGs have a Benjamini & Hochberg adjusted *p*-value < 0.05 and an absolute log2 fold change > 1.5. Genes in the figure with no expression information are not differentially expressed to a significant extent or *C. illinoensis* DEGs have no *Arabidopsis thaliana* ortholog information. Pathway element in the figure shows expression pattern for each individual in the case group (R or S-D) compared to S-ND

Each of the conditions examined, S-ND, S-D, and R leaves, showed unique expression patterns. These expression patterns are essential for gaining insight into different aspects of the pecan scab pathosystem. By comparing R and S-ND samples one might elucidate factors responsible for preventing or allowing initial infection by *V. effusa*. Whereas comparing S-ND to S-D, potential targets of fungal effectors may be identified, among other genes active in the infection process. The comparison R to S-D was not evaluated, as it confounds diseased state and basal expression type factors. The larger proportion of down-regulated DEGs in R samples could indicate

susceptibility factors as a causal agent of resistance/susceptibility to pecan scab. When examining the distribution of differentially expressed R-genes within the 87MX3-2.11 genome, one locus contained eight down-regulated leucine rich-repeat type predicted R-genes (Fig. 7, Additional file 1; tab 9 “SusQTL”).

If population level resistance to pecan scab is dependent on R-gene mediated resistance however, a single up-regulated R-gene could be responsible for resistance despite the general down-regulation of R-genes in resistant samples. Regardless, expression of each of the up-regulated DEG predicted R-genes in



Fig. 7 Gene expression at predicted susceptibility quantitative trait loci (QTL) and flanking regions in scab resistant leaves of pecan (*Carya illinoensis* 87MX3-2.11). Red gene names indicate genes that are significantly down-regulated. Eight of the nine differentially expressed genes (DEGs) are predicted leucine rich repeat R-genes. Expression of most proximal 5' and 3' genes also shown

resistant samples is present in S-ND samples. Their presence in susceptible samples should negate them as possible sources of resistance, at least at the level of gene expression. This is because R-gene mediated resistance in plants typically shows a well-defined dominance-recessive phenotype.

A number of other essential components of plant defense pathways are also down-regulated in resistant trees. Mitogen-activated protein kinases (MAPK) are intermediate elements of the defense pathway, carrying signals generated by R-genes to the nucleus resulting in the expression of a myriad of defense genes (Reviewed in: [53]). Genes set analysis was able to identify MPK9 as being down-regulated in the R group. As a further point of confusion, despite WRKY33 being shown to be essential for resistance to fungal pathogens in *A. thaliana* [54], WRKY33 is down-regulated in scab resistant pecan. Regardless of the down-regulation for these upstream genes of the defense pathway, there was no significant difference in expression of known defense genes such as PR1 in resistant trees compared to susceptible trees.

Potentially less obvious components of the plant immune system are also down-regulated in resistant pecan trees. A component of the CCR4-CAF1-Not1 cytoplasmic deadenylase complex, CNOT7/8, is down-regulated. In Arabidopsis, over-expression of another component of this complex led to increased susceptibility to the necrotrophic fungal pathogen *Botrytis cinerea* [55]. Of the few genes up-regulated in resistant pecan samples, gene ontology enrichment identified up-regulation of uracil transmembrane transporter activity. Rapidly growing plant tissues, like the expanding leaves required for *V. effusa* infection, are a sink for a variety of biomolecules. Since uracil is an intermediate for important enzyme cofactors, signaling molecules, and secondary metabolites, the up-regulation of uracil transporter activity in resistant plants could allow them to respond more quickly to *V. effusa* in a tailored fashion. Data analyzed here could be compared to a closer related plant species

with a similar pathogen such as apple (*Malus domestica*) and apple scab (*Venturia inaequalis*); although, there are obvious limitations with non-model systems. For example, *A. thaliana* has forty-two terms of its KEGG plant-pathogen interaction pathway defined while apple only has thirty-seven terms defined.

In stark contrast to the R by S-ND comparison, DEGs in S-D leaf samples were largely up-regulated when compared to S-ND leaf samples. Enriched gene ontology analysis identified similar terms compared to the cultivar-pathovar level analysis done by Lovell et al. [33]. These terms include down-regulation of “extracellular region” (GO:0005576), down-regulation of “response to auxin” (GO:0009733), and up-regulation of “DNA-binding transcription factor activity” (GO:0003700). Expression patterns for components of plant defense pathways clearly show recognition of the pathogen by susceptible plants. For example, two fungal/ethylene inducible genes PR-1 and ChiB were consistently up-regulated [56, 57]. Also, jasmonic acid/ethylene signaling pathways are used by plants for resistance to fungal necrotrophs [58, 59]. Genes in both of these pathways were identified to be significantly up-regulated in susceptible diseased samples. Resolution of the current study is not detailed enough to determine if these defenses are responsible for disease susceptibility or are successful mechanisms for preventing further spread of infection to healthy plant tissues. Deliberately triggering host defenses is a well-documented infection strategy for necrotrophic fungal phytopathogens [60, 61]. If on the other hand these defenses are responsible for stopping the spread of the pathogen to healthy tissues, they represent prime candidates for overexpression lines for decreasing disease severity to *V. effusa*. RNA-seq experiments using controlled inoculations, time-course series, and laser-capture microdissection are better suited to answer these questions.

Dissimilar transcriptome profiles are the underlying cause for cellular, tissue, and structural differences found within the same individual. The correlation between scab

susceptibility in fruit and leaf tissue is high [4]. Given this, genes and pathways responsible for resistance to diverse populations of *V. effusa* could be further refined by comparing fruit whole transcriptome profiles of the trees in this study to their counterpart leaf data. Evaluating the proteomics and metabolomics in these trees will also help give a better understanding of the intricacies of the pecan/scab pathosystem. Because S-ND and S-D samples were collected as pairs from each susceptible tree, these samples could be analyzed as paired-data. Analysis performed using this method has been shown to have increased power at detecting differentially expressed genes that have lower average expression and shorter transcript length [62]. Since *V. effusa* was present on the same tree that S-ND samples were collected from, long-range transport and signaling of phytohormones could skew S-ND samples toward an expression profile more closely resembling S-D. This has the potential to produce artifacts in both the S-ND by S-D and R by S-ND comparisons made in this study. To mitigate for this, a combination of fungicides, metabolomics, and time-course sampling could be performed in future studies. Despite the uncertainties present in the S-ND treatment group, controlled inoculations of susceptible pecan done by Lovell et al. (2021) gave similar gene ontology enrichment results as those obtained for the S-ND by S-D comparison done for the present study. This would suggest that the S-ND samples do indeed reflect basal expression of susceptible plants. Analysis of basal expression for resistant and susceptible plants is often overlooked despite the known importance of basal defenses and susceptibility factors (Reviewed in: [63, 64]). Including this type of analysis when examining other pathosystems will help draw parallels and explain the importance of these factors in pathogen resistance/susceptibility. Furthermore, a large portion of the predicted *C. illinoensis* genes have missing *A. thaliana* and GO descriptions (34.8% and 38.0% respectively for the Phytozome 573 v1.1 annotation). As our understanding of gene functions continues to grow, gene ontology and gene set analysis will become more powerful at shedding light on this important pecan pathosystem.

Conclusions

The present study shows differentially expressed genes of scab resistant pecan trees to be generally down-regulated while susceptible pecan trees have many up-regulated defense related pathways upon infection. This makes a compelling case for susceptibility as the basis for compatibility between the fungus *V. effusa* and pecan. Differentially expressed genes identified by this study will aid further investigation in the *V. effusa*/pecan interaction and provide a means for ultimately developing durable resistance to this formidable pathogen.

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s12864-024-10010-0>.

Additional file 1. Differentially expressed genes plus other supporting information. Differentially expressed genes including annotation information, predicted *Carya illinoensis* transcription factors, predicted *Carya illinoensis* pathogen resistance genes, GOrrilla gene ontology enrichment information, primer information, and susceptibility quantitative trait loci information.

Additional file 2: Figure 1. Gene ontology enrichment results from GOrrilla reporting significantly down-regulated ontology types in susceptible diseased samples (S-D), compared to susceptible non-diseased samples (S-ND). Significance of GO term is indicated with the color corresponding to the *P*-value color scale at the top of the figure. **Figure 2.** Gene ontology enrichment results from GOrrilla reporting significantly up-regulated ontology types in susceptible diseased samples (S-D), compared to susceptible non-diseased samples (S-ND). Significance of GO term is indicated with the color corresponding to the *P*-value color scale at the top of the figure. **Figure 3.** Gene ontology enrichment results from GOrrilla reporting down-regulated & up-regulated ontology types in resistant samples (R), compared to susceptible non-diseased samples (S-ND). Significance of GO term is indicated with the color corresponding to the *P*-value color scale at the top of the figure. Ontology pathway-regulation types not shown were not reported by GOrrilla. **Figure 4.** Kegg pathway (00500) analysis via Pathview for differentially expressed genes (DEG; absolute \log_2 fold change > 1.5, Benjamini & Hochberg adjusted *p*-value less than 0.05) in susceptible-diseased samples (S-D), compared to susceptible non-diseased samples (S-ND). Color represents expression pattern. Genes without expression information were either not differentially expressed or *Carya illinoensis* DEGs had no corresponding *Arabidopsis thaliana* ortholog information. Expression pattern for each case sample (S-D) is represented in each gene box. Reproduction of KEGG pathway granted by Kanehisa Laboratories under open access license. **Figure 5.** Kegg pathway (00999) analysis via Pathview for differentially expressed genes (DEG; absolute \log_2 fold change > 1.5, Benjamini & Hochberg adjusted *p*-value less than 0.05) in susceptible-diseased samples (S-D), compared to susceptible non-diseased samples (S-ND). Color represents expression pattern. Genes without expression information were either not differentially expressed or *Carya illinoensis* DEGs had no corresponding *Arabidopsis thaliana* ortholog information. Expression pattern for each case sample (S-D) is represented in each gene box. Reproduction of KEGG pathway granted by Kanehisa Laboratories under open access license. **Figure 6.** Kegg pathway (04016) analysis via Pathview for differentially expressed genes (DEG; absolute \log_2 fold change > 1.5, Benjamini & Hochberg adjusted *p*-value less than 0.05) in susceptible-diseased samples (S-D), compared to susceptible non-diseased samples (S-ND). Color represents expression pattern. Genes without expression information were either not differentially expressed or *Carya illinoensis* DEGs had no corresponding *Arabidopsis thaliana* ortholog information. Expression pattern for each case sample (S-D) is represented in each gene box. Reproduction of KEGG pathway granted by Kanehisa Laboratories under open access license. **Figure 7.** Kegg pathway (04075) analysis via Pathview for differentially expressed genes (DEG; absolute \log_2 fold change > 1.5, Benjamini & Hochberg adjusted *p*-value less than 0.05) in susceptible-diseased samples (S-D), compared to susceptible non-diseased samples (S-ND). Color represents expression pattern. Genes without expression information were either not differentially expressed or *Carya illinoensis* DEGs had no corresponding *Arabidopsis thaliana* ortholog information. Expression pattern for each case sample (S-D) is represented in each gene box. Reproduction of KEGG pathway granted by Kanehisa Laboratories under open access license. **Figure 8.** Original Kegg pathway (04626) analysis via Pathview for differentially expressed genes (DEG; absolute \log_2 fold change > 1.5, Benjamini & Hochberg adjusted *p*-value less than 0.05) in susceptible-diseased samples (S-D), compared to susceptible non-diseased samples (S-ND). Color represents expression pattern. Genes without expression information were either not differentially expressed or *Carya illinoensis* DEGs had no corresponding *Arabidopsis thaliana* ortholog information. Expression pattern for each case sample (R) is represented in each gene box. Reproduction of KEGG pathway granted

by Kanehisa Laboratories under open access license. **Figure 9.** Kegg pathway (03018) analysis via Pathview for differentially expressed genes (DEG; absolute \log_2 fold change > 1.5, Benjamini & Hochberg adjusted p -value less than 0.05) in resistant samples (R), compared to susceptible non-diseased samples (S-ND). Color represents expression pattern. Genes without expression information were either not differentially expressed or *Carya illinoensis* DEGs had no corresponding *Arabidopsis thaliana* ortholog information. Expression pattern for each case sample (R) is represented in each gene box. Reproduction of KEGG pathway granted by Kanehisa Laboratories under open access license. **Figure 10.** Kegg pathway (04016) analysis via Pathview for differentially expressed genes (DEG; absolute \log_2 fold change > 1.5, Benjamini & Hochberg adjusted p -value less than 0.05) in resistant samples (R), compared to susceptible non-diseased samples (S-ND). Color represents expression pattern. Genes without expression information were either not differentially expressed or *Carya illinoensis* DEGs had no corresponding *Arabidopsis thaliana* ortholog information. Expression pattern for each case sample (R) is represented in each gene box. Reproduction of KEGG pathway granted by Kanehisa Laboratories under open access license. **Figure 11.** Kegg pathway (04075) analysis via Pathview for differentially expressed genes (DEG; absolute \log_2 fold change > 1.5, Benjamini & Hochberg adjusted p -value less than 0.05) in resistant samples (R), compared to susceptible non-diseased samples (S-ND). Color represents expression pattern. Genes without expression information were either not differentially expressed or *Carya illinoensis* DEGs had no corresponding *Arabidopsis thaliana* ortholog information. Expression pattern for each case sample (R) is represented in each gene box. Reproduction of KEGG pathway granted by Kanehisa Laboratories under open access license. **Figure 12.** Original Kegg pathway (04626) analysis via Pathview for differentially expressed genes (DEG; absolute \log_2 fold change > 1.5, Benjamini & Hochberg adjusted p -value less than 0.05) in resistant samples (R), compared to susceptible non-diseased samples (S-ND). Color represents expression pattern. Genes without expression information were either not differentially expressed or *Carya illinoensis* DEGs had no corresponding *Arabidopsis thaliana* ortholog information. Expression pattern for each case sample (R) is represented in each gene box. Reproduction of KEGG pathway granted by Kanehisa Laboratories under open access license. **Figure 13.** Principal component analysis showing the genetic diversity of 835 trees of the pecan provenance collection in Byron, GA using genotype by sequencing (data not shown). Scab susceptible and resistant individuals used in the present study are shown as well as Mexico and Texas genotype clustering. Cultivars developed from northern provenance trees (not Mexico or Texas germplasms) are generally considered to be more resistant to pecan scab.

Acknowledgments

The authors would like to thank USDA-ARS technician Minling Zhang for help with sample collection. From the Noble Research Institute, we thank Nick Krom for preliminary data analysis, Guifen Li and Taegun Kwon for help with sample prep, and Yuhong Tang for coordinating sequencing of the samples. Integral high performance computing resources were provided by the SCINet project of the USDA Agricultural Research Service, ARS project number 0500-00093-001-00-D.

This article reports the results of research only. Mention of a trademark or proprietary product is solely for the purpose of providing specific information and does not constitute a guarantee or warranty of the product by the U.S. Department of Agriculture and does not imply its approval to the exclusion of other products that may also be suitable.

Authors' contributions

JJB, experimental design, sample collection and processing for qPCR, data analysis, final manuscript write-up. YA, sample processing, data analysis, preliminary manuscript write-up. CHB, experimental design, sample collection, manuscript editing, funding acquisition. JS, data analysis, manuscript editing. MJM, experimental design, sample collection, funding acquisition, manuscript editing. JJR, overall project coordination, funding acquisition, manuscript editing. CY, experimental design, data curation, data analysis, manuscript editing. All authors have agreed to the final manuscript.

Funding

JJB and CHB were funded by the USDA-ARS projects 6042-21220-013-000-D and 6042-21220-014-000-D. MJM and YA were funded by NIFA-Specialty Crops Research Initiative "Coordinated Development of Genetic Tools for Pecan" (Grant no. 2016-51181-25408), The Oklahoma Department of Agriculture, Food and Forestry, and the Noble Research Institute. JJR was funded by the NIFA-Specialty Crops Research Initiative "Coordinated Development of Genetic Tools for Pecan" (Grant no. 2016-51181-25408) and NIFA-Specialty Crops Research Initiative "Trees for the future: Coordinated development of genetic resources and tools to accelerate breeding of geographic and climate adapted pecan trees" (Grant no. 2022-51181-38332).

Availability of data and materials

The dataset supporting the conclusion of this article are available in the National Center for Biotechnology's Sequence Read Archive (NCBI-SRA) repository, as Bioproject PRJNA992068.

Declarations

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

Authors declare that they have no competing interests.

Author details

¹USDA-ARS, Byron, GA, USA. ²Noble Research Institute, Ardmore, OK, USA. ³University of Texas Southwestern, Dallas, TX, USA. ⁴The New Zealand Institute for Plant and Food Research, Auckland, New Zealand. ⁵Entomology and Plant Pathology, NC State University, Raleigh, NC, USA. ⁶Bayer Crop Science, Chesterfield, MO, USA. ⁷New Mexico State University, Las Cruces, NM, USA.

Received: 29 August 2023 Accepted: 12 January 2024

Published online: 15 February 2024

References

- USDA-NASS. Statistics by subject: Pecans, Utilized, In Shell - Production, Measured in \$ [Available from: https://www.nass.usda.gov/Statistics_by_Subject/result.php?EB7A7D83-A799-32A3-836F-0C54A8513C38§or=CROPS&group=FRUIT%20%26%20TREE%20NUTS&comm=PECANS].
- Grauke LJ, Payne JA, Wood BW. North American pecans: a provenance study. *Ann Rep Northern Nut Growers Assoc.* 1989;80:124–31.
- Rüter B, Hamrick JL, Wood BW. Genetic diversity within provenance and cultivar germplasm collections versus natural populations of pecan (*Carya illinoensis*). *J Heredity.* 1999;90(5):521–8.
- Bock CH, Alarcon Y, Conner PJ, Young CA, Randall JJ, Pisani C, et al. Foliage and fruit susceptibility of a pecan provenance collection to scab, caused by *Venturia effusa*. *CABI Agric Biosci.* 2020;1(1):19.
- Bock CH, Grauke LJ, Conner P, Burrell SL, Hotchkiss MW, Boykin D, et al. Scab susceptibility of a provenance collection of pecan in three different seasons in the Southeastern United States. *Plant Dis.* 2016;100(9):1937–45.
- Adams JR, Bailey ME, Burkhead CE, Cook HT, Vermeer J, Wilson RC, et al. Losses in Agriculture. In: Agriculture USDo, editor. 1965.
- Gottwald TR. Influence of temperature, leaf wetness period, leaf age, and spore concentration on infection of pecan leaves by conidia of *Cladosporium caryigenum*. *Phytopathology.* 1985;75(2):190–4.
- Demaree JB. Pecan scab with special reference to sources of the early spring infections. 1924:p. 321–30. plates-USDA.
- Charlton ND, Yi M, Bock CH, Zhang M, Young CA. First description of the sexual stage of *Venturia effusa*, causal agent of pecan scab. *Mycologia.* 2020;112(4):711–21.

10. Bock CH, Hotchkiss MW, Young CA, Charlton ND, Chakradhar M, Stevenson KL, et al. Population genetic structure of *Venturia effusa*, cause of pecan scab, in the Southeastern United States. *Phytopathology*. 2017;107(5):607–19.
11. Young CA, Bock CH, Charlton ND, Mattupalli C, Krom N, Bowen JK, et al. Evidence for sexual reproduction: identification, frequency, and spatial distribution of *Venturia effusa* (pecan scab) mating type idiomorphs. *Phytopathology*. 2018;108(7):837–46.
12. Bock CH, Wells L, Hotchkiss MW. Effect of tractor speed and spray application volume on severity of scab and fruit weight at different heights in the canopy of tall pecan trees. *Plant Dis*. 2021;105(12):3909–24.
13. Standish JR, Breneman TB, Bock CH, Stevenson KL. Fungicide resistance in *Venturia effusa*, cause of pecan scab: current status and practical implications. *Phytopathology*. 2020;111(2):244–52.
14. Andersen EJ, Ali S, Byamukama E, Yen Y, Nepal MP. Disease resistance mechanisms in plants. *Genes (Basel)*. 2018;9(7).
15. Muthamilarasan M, Prasad M. Plant innate immunity: an updated insight into defense mechanism. *J Biosci*. 2013;38(2):433–49.
16. Jones JDG, Dangl JL. The plant immune system. *Nature*. 2006;444(7117):323–9.
17. Asai T, Tena G, Plotnikova J, Willmann MR, Chiu W-L, Gomez-Gomez L, et al. MAP kinase signalling cascade in Arabidopsis innate immunity. *Nature*. 2002;415(6875):977–83.
18. Selitrennikoff CP. Antifungal proteins. *Appl Environ Microbiol*. 2001;67(7):2883–94.
19. Pradhan SK, Nayak DK, Mohanty S, Behera L, Barik SR, Pandit E, et al. Pyramiding of three bacterial blight resistance genes for broad-spectrum resistance in deepwater rice variety, Jalmagna. *Rice (NY)*. 2015;8(1):51.
20. Baumgartner IO, Patocchi A, Frey J, Peil A, Kellerhals M. Breeding elite lines of apple carrying pyramided homozygous resistance genes against apple scab and resistance against powdery mildew and fire blight. *Plant Mol Biol Report*. 2015;33:1573–83.
21. Qi L, Ma G. Marker-assisted gene pyramiding and the reliability of using SNP markers located in the recombination suppressed regions of sunflower (*Helianthus annuus* L.). *Genes (Basel)*. 2019;11(1).
22. Jensen PJ, Rytter J, Detwiler EA, Travis JW, McNellis TW. Rootstock effects on gene expression patterns in apple tree scions. *Plant Mol Biol*. 2003;53(4):493–511.
23. Benjamin G, Tietel Z, Porat R. Effects of rootstock/scion combinations on the flavor of citrus fruit. *J Agric Food Chem*. 2013;61(47):11286–94.
24. Conner PJ. A detached leaf technique for studying race-specific resistance to *Cladosporium caryigenum* in pecan. *J Am Soc Hortic Sci*. 2002;127(5):781–5.
25. Martinelli F, Uratsu SL, Albrecht U, Reagan RL, Phu ML, Britton M, et al. Transcriptome profiling of citrus fruit response to Huanglongbing disease. *PLoS One*. 2012;7(5):e38039.
26. Chakraborty S, Britton M, Martínez-García PJ, Dandekar AM. Deep RNA-Seq profile reveals biodiversity, plant-microbe interactions and a large family of NBS-LRR resistance genes in walnut (*Juglans regia*) tissues. *AMB Express*. 2016;6(1):12.
27. Acquadro A, Torello Marinoni D, Sartor C, Dini F, Macchio M, Botta R. Transcriptome characterization and expression profiling in chestnut cultivars resistant or susceptible to the gall wasp *Dryocosmus kuriphilus*. *Mol Genet Genomics*. 2020;295(1):107–20.
28. Thompson MY, Randall JJ, VanLeeuwen D, Heerema RJ. Differential expression of key floral initiation genes in response to plant growth regulator application and alternate bearing in pecan. *J Am Soc Hortic Sci*. 2021;146(3):206–14.
29. Xu Z, Ni J, Shah FA, Wang Q, Wang Z, Wu L, et al. Transcriptome analysis of pecan seeds at different developing stages and identification of key genes involved in lipid metabolism. *PLoS One*. 2018;13(4):e0195913.
30. Wang M, Xi D, Chen Y, Zhu C, Zhao Y, Geng G. Morphological characterization and transcriptome analysis of pistillate flowering in pecan (*Carya illinoensis*). *Scientia Horticulturae*. 2019;257:108674.
31. Mattison CP, Rai R, Settledge RE, Hinchliffe DJ, Madison C, Bland JM, et al. RNA-Seq analysis of developing pecan (*Carya illinoensis*) embryos reveals parallel expression patterns among allergen and lipid metabolism genes. *J Agric Food Chem*. 2017;65(7):1443–55.
32. Chen Y, Zhang S, Zhao Y, Mo Z, Wang W, Zhu C. Transcriptomic analysis to unravel potential pathways and genes involved in pecan (*Carya illinoensis*) resistance to *Pestalotiopsis microspora*. *Int J Mol Sci*. 2022;23(19).
33. Lovell JT, Bentley NB, Bhattarai G, Jenkins JW, Sreedasyam A, Alarcon Y, et al. Four chromosome scale genomes and a pan-genome annotation to accelerate pecan tree breeding. *Nat Commun*. 2021;12(1):4125.
34. Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics*. 2014;30(15):2114–20.
35. Randall JJ, Schmutz J. *Carya illinoensis* cv. 87MX3-2.11. *Phytozome* V13; 2021.
36. Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. *Nat Methods*. 2012;9(4):357–9.
37. Danecek P, Bonfield JK, Liddle J, Marshall J, Ohan V, Pollard MO, et al. Twelve years of SAMtools and BCFtools. *GigaScience*. 2021;10(2):1–4.
38. Patro R, Duggal G, Love MI, Irizarry RA, Kingsford C. Salmon provides fast and bias-aware quantification of transcript expression. *Nat Methods*. 2017;14(4):417–9.
39. R Development Core Team. R: A language and environment for statistical computing. Vienna: R Foundation for Statistical Computing 2013.
40. Soneson C, Love MI, Robinson MD. Differential analyses for RNA-seq: transcript-level estimates improve gene-level inferences. *F1000Res*. 2015;4:1521.
41. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol*. 2014;15(12):550.
42. Stephens M. False discovery rates: a new deal. *Biostatistics*. 2017;18(2):275–94.
43. Winter DJ, Charlton ND, Krom N, Shiller J, Bock CH, Cox MP, et al. Chromosome-level reference genome of *Venturia effusa*, causative agent of pecan scab. *Mol Plant Microbe Interact*. 2020;33(2):149–52.
44. Chen H, Boutros PC. VennDiagram: a package for the generation of highly-customizable Venn and Euler diagrams in R. *BMC Bioinformatics*. 2011;12(1):35.
45. Luo W, Pant G, Bhavnasi YK, Blanchard SG Jr, Brouwer C. Pathview Web: user friendly pathway visualization and data integration. *Nucleic Acids Res*. 2017;45(W1):W501–8.
46. Thimm O, Blasing O, Gibon Y, Nagel A, Meyer S, Krüger P, et al. MAPMAN: a user-driven tool to display genomics data sets onto diagrams of metabolic pathways and other biological processes. *Plant J*. 2004;37:914–39.
47. Eden E, Navon R, Steinfeld I, Lipson D, Yakhini Z. GOrilla: a tool for discovery and visualization of enriched GO terms in ranked gene lists. *BMC Bioinformatics*. 2009;10(1):48.
48. Kanehisa M, Furumichi M, Sato Y, Kawashima M, Ishiguro-Watanabe M. KEGG for taxonomy-based analysis of pathways and genomes. *Nucleic Acids Res*. 2023;51(D1):D587–d92.
49. Osuna-Cruz CM, Paytuy-Gallart A, Di Donato A, Sundesha V, Andolfo G, Aiese Cigliano R, et al. PRGdb 3.0: a comprehensive platform for prediction and analysis of plant disease resistance genes. *Nucleic Acids Res*. 2018;46(D1):D1197–d201.
50. Steuernagel B, Witek K, Krattinger SG, Ramirez-Gonzalez RH, Schoonbeek H-J, Yu G, et al. The NLR-Annotator tool enables annotation of the intracellular immune receptor Repertoire. *Plant Physiol*. 2020;183(2):468–82.
51. Kim GB, Gao Y, Palsson BO, Lee SY. DeepTFactor: A deep learning-based tool for the prediction of transcription factors. *Proc Natl Acad Sci*. 2021;118(2):e2021171118.
52. Jin J, Tian F, Yang D-C, Meng Y-Q, Kong L, Luo J, et al. PlantTFDB 4.0: toward a central hub for transcription factors and regulatory interactions in plants. *Nucleic Acids Res*. 2017;45(D1):D1040–5.
53. Meng X, Zhang S. MAPK cascades in plant disease resistance signaling. *Annu Rev Phytopathol*. 2013;51:245–66.
54. Zheng Z, Qamar SA, Chen Z, Mengiste T. Arabidopsis WRKY33 transcription factor is required for resistance to necrotrophic fungal pathogens. *Plant J*. 2006;48(4):592–605.
55. Kwon T, Young-Byoung Y, Nam J. Overexpression of AtCAF1, CCR4-associated factor 1 homologue in *Arabidopsis thaliana*, negatively regulates wounding-mediated disease resistance. *J Plant Biotechnol*. 2011;38:278–84.
56. Mitsuhashi I, Iwai T, Seo S, Yanagawa Y, Kawahigasi H, Hirose S, et al. Characteristic expression of twelve rice PR1 family genes in response to pathogen infection, wounding, and defense-related signal compounds (12/1/180). *Mol Genet Genomics*. 2008;279(4):415–27.

57. Samac DA, Hironaka CM, Yallaly PE, Shah DM. Isolation and characterization of the genes encoding basic and acidic chitinase in *Arabidopsis thaliana*. *Plant Physiol.* 1990;93(3):907–14.
58. Penninckx IAMA, Eggermont K, Terras FRG, Thomma BPHJ, De Samblanx GW, Buchala A, et al. Pathogen-induced systemic activation of a plant defensin gene in *Arabidopsis* follows a salicylic acid-independent pathway. *Plant Cell.* 1996;8:2309–23.
59. Berrocal-Lobo M, Molina A. Ethylene response factor 1 mediates *Arabidopsis* resistance to the soilborne fungus *Fusarium oxysporum*. *Mol Plant Microbe Interact.* 2004;17(7):763–70.
60. El Oirdi M, Bouarab K. Plant signalling components EDS1 and SGT1 enhance disease caused by the necrotrophic pathogen *Botrytis cinerea*. *N Phytologist.* 2007;175(1):131–9.
61. Lorang J, Kidarsa T, Bradford CS, Gilbert B, Curtis M, Tzeng SC, et al. Tricking the guard: exploiting plant defense for disease susceptibility. *Science.* 2012;338(6107):659–62.
62. Chung LM, Ferguson JP, Zheng W, Qian F, Bruno V, Montgomery RR, et al. Differential expression analysis for paired RNA-seq data. *BMC Bioinformatics.* 2013;14(1):110.
63. He Q, McLellan H, Boevink PC, Birch PRJ. All roads lead to susceptibility: the many modes of action of fungal and oomycete intracellular effectors. *Plant Commun.* 2020;1(4):100050.
64. Niks RE, Marcel TC. Nonhost and basal resistance: how to explain specificity? *New Phytologist.* 2009;182(4):817–28.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.