Title
Association mapping, transcriptomics, and transient expression identify candidate genes mediating plant-pathogen interactions in a tree.

Permalink
https://escholarship.org/uc/item/2dh1x4dz

Journal
Proceedings of the National Academy of Sciences of the United States of America, 115(45)

ISSN
0027-8424

Authors
Muchero, Wellington
Sondreli, Kelsey L
Chen, Jin-Gui
et al.

Publication Date
2018-11-01

DOI
10.1073/pnas.1804428115

Peer reviewed
Association mapping, transcriptomics, and transient expression identify candidate genes mediating plant–pathogen interactions in a tree

Wellington Muchero, Kelsey L. Sondrelli, Jin-Gui Chen, Breeanna R. Urbanowicz, Jin Zhang, Vasanth Singan, Yongil Yang, Robert S. Brueggeman, Juan Franco-Coronado, Nivi Abraham, Jeong-Yeh Yang, Kelley W. Moremen, Alexandra J. Weisberg, Jeff H. Chang, Erika Lindquist, Kerrie Barry, Priya Ranjan, Sara Jawdy, Jeremy Schmutz, Gerald A. Tuskan, and Jared M. LeBoldus

Invasive microbes causing diseases such as sudden oak death negatively affect ecosystems and economies around the world. The deployment of resistant genotypes for combating introduced diseases typically relies on breeding programs that can take decades to complete. To demonstrate how this process can be accelerated, we employed a genome-wide association mapping of ca. 1,000 resequenced *Populus trichocarpa* trees individually challenged with *Sphaerulina musiva*, an invasive fungal pathogen. Among significant associations, three loci associated with resistance were identified and predicted to encode one putative membrane-bound L-type receptor-like kinase and two receptor-like proteins. A susceptibility-associated locus was predicted to encode a putative G-type D-mannose-binding receptor-like kinase. Multiple lines of evidence, including allele analysis, transcriptomics, binding assays, and overexpression, support the hypothesized function of these candidate genes in the *P. trichocarpa* response to *S. musiva*.

Host–pathogen coevolution has been described for many species and is the major focus of research on innate immunity in plant and animal systems (1). In what is commonly referred to as a coevolutionary “arms race,” models predict adaptation and counteradaptation, whereby host and pathogen genomes undergo complementary changes to thwart or facilitate infection, respectively (2). Because of the focus on coevolved hosts and microbes, no models exist that predict the mechanism(s) by which exotic pathogens counter innate immune responses and infect naive hosts. Diseases that exemplify such naive pathosystems include chestnut blight (3), white pine blister rust (4), and sudden oak death (5). These examples highlight the catastrophic consequences of exotic pathogens when most host genotypes are susceptible to the introduced microbe. As the host disappears, ecosystem structure and function are perturbed, resulting in declines in forest health (5). This is particularly problematic in an age in which global trade and climate change are permanently altering species distributions, resulting in new host–pathogen symmetries (6). It is unclear whether current models (1) of host–parasite interactions sufficiently describe the interactions between plants and exotic pathogens.

We developed a rapid phenotyping platform to identify loci associated with *Populus trichocarpa* response to *Sphaerulina musiva* (7) to characterize the genetic mechanism(s) underlying host–pathogen compatibility in the absence of coevolution (3). In eastern North America, the fungal pathogen *S. musiva* is endemic in natural stands of *Populus* where it has coevolved with its host, *Populus deltoides*, causing leaf-spot disease. *S. musiva* was recently introduced to western North America (8) where interactions with the naive host *P. trichocarpa* cause severe stem-girdling cankers leading to premature crown death (9). It is predicted that either (i) as a naive host, *P. trichocarpa* will lack immunity to *S. musiva* or (ii) the pathogen will suppress the host’s immune response.

Results and Discussion

In a greenhouse experiment, 5,405 plants from a population of 1,081 distinct *P. trichocarpa* genotypes were planted. Three to five of the planted cuttings from each genotype successfully rooted. Three or four trees from each genotype were used in the subsequent genome-wide association studies (GWAS), and any extra trees were discarded. At 3 wk postinoculation a total of 3,404 trees were characterized for phenotypic responses to *S. musiva* (Fig. 1). The broad sense heritability of this trait was estimated to be 0.35 (Dataset S1, Table S4). Phenotypes were correlated to 8,253,066 SNPs and insertion/deletions (indels) (10). The combined rapid phenotyping, GWAS, and allele analysis was completed within 5 mo of planting the *Populus* trees. A total of 96 polymorphisms encompassing 73 candidate genes were identified (SI Appendix, Fig. S1 and Dataset S1, Table S1). Of these, 29 were identified in all three datasets. Of these, 10 genes and 12 SNPs were observed in at least two of the independent datasets. Of these, nine genes were observed in all datasets. This is the first GWAS for adaptation to an invasive pathogen and demonstrates the potential of genomics to rapidly identify candidate genes associated with resistance and susceptibility to introduced plant diseases. The unprecedented speed and accuracy with which the candidate genes can be identified in woody trees demonstrates the potential of genomics to mitigate the impacts of invasive diseases on forest health.

Significance

International trade has resulted in the introduction of plant diseases into natural ecosystems around the world. These introductions have potentially catastrophic impacts on ecosystem structure and function. Leveraging genomic tools, natural variation within a tree species, and a high-throughput phenotyping platform, we present a framework that can be broadly applied to rapidly identify candidate genes associated with resistance and susceptibility to introduced plant diseases. The unprecedented speed and accuracy with which the candidate genes can be identified in woody trees demonstrates the potential of genomics to mitigate the impacts of invasive diseases on forest health.

Author contributions: W.M., J.-G.C., B.R.U., R.S.B., J.F.-C., G.A.T., and J.M.L. designed research; W.M., K.L.S., B.R.U., J.F.-C., N.A., K.W.M., S.J., and J.M.L. performed research; B.R.U., E.L., K.B., and J.S. contributed new reagent/analytical tools; W.M., K.L.S., J.-G.C., B.R.U., J.Z., V.S., Y.Y., J.-Y.Y., A.J.W., J.H.C., E.L., K.B., P.R., J.S., and J.M.L. analyzed data; and W.M., K.L.S., J.-G.C., B.R.U., A.J.W., J.H.C., G.A.T., and J.M.L. wrote the paper. The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

This open access article is distributed under Creative Commons Attribution-NonCommercial-NoDerivatives License 4.0 (CC BY-NC-ND).

Data deposition: The sequences reported in this paper have been deposited in the National Center for Biotechnology Information Sequence Read Archive (accession nos. SRR2502320–SRR2502340). This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1804428115/-/DCSupplemental.

Published online October 18, 2018.
these, nine genomic intervals exceeded the Bonferroni-corrected significance threshold in at least two replicates (Dataset S1, Table S8). Six of the top eight associations involved SNPs located in genes predicted to encode proteins with domains common to receptors. These included two paralogue-rich-leucine receptor-like proteins (RLPs) (Potri.005G012100, P value = 1.56E-38; Potri.005G028200, P value = 2.78E-14), an L-type lectin receptor-like protein kinase (L-type lecRLK; Potri.009G036300, P value = 2.15E-16), and a G-type lectin receptor-like protein kinase (G-type lecRLK; Potri.005G018000, P value = 1.161E-13) (Figs. 1 and 2). Analyses of allelic effect direction suggested that the two RLPs and L-type lecRLK are associated with resistance, whereas the G-type lecRLK is associated with susceptibility. Pairwise linkage disequilibrium (LD) for all four candidate loci decayed rapidly, falling below $R^2 = 0.10$ within 50 bp (Dataset S1, Table S2). A similar rate of LD decay has been reported for R-genes in other plant species (11).

RLPs interact with receptor-like protein kinases (RLKs) to perceive a ligand signal and trigger downstream protein phosphorylation cascades (12). For example, the RLPs (CF-4 or Ve1) interact with an RLK (SOBIR1/EVR) in tomato to trigger immune signaling in response to *Cladosporium fulvum* and *Verticillium dahliae* pathogens, respectively (12). The two RLPs of *P. trichocarpa* contain an extracellular leucine-rich-repeat domain, a transmembrane domain, and a short cytoplasmic tail but lack a kinase domain. Lectin receptor kinases (LecRLKs) have also been implicated in mediating resistance to fungal pathogens (12, 13) or facilitating symbiosis (14). They are typically comprised of an extracellular N-terminal lectin domain, a transmembrane region, and a cytosolic kinase domain. LecRLKs have been classified into three types in *Arabidopsis*, rice, and *Populus* based on the linked lec domain(s) as C-, G-, and L-type (15, 16). LecRLKs are proteins that possess at least one noncatalytic domain that can selectively recognize and reversibly bind to specific carbohydrate structures (17). We hypothesize that the functional receptors (RLPs and LecRLKs) in *P. trichocarpa* genotypes are necessary to perceive *S. musiva*-derived ligand(s) and mediate the interaction.

We used SNPEff (effects of SNPs) analysis (18) to examine the population-wide occurrence of inferred nonconsensual and deleterious mutations (i.e., early translation termination, frame-shifts, and changes in splice-site acceptor and/or splice-site donor sequences) (Fig. 2, SI Appendix, Figs. S2–S5, and Dataset S1, Table S3). Specifically, the occurrence of predicted high-impact mutations across protein-coding regions revealed that RLP1 and RLP2 had the highest mutation frequencies, 32% and 43%, respectively. The L-type lecRLK had a mutation frequency of 10%. Mutation frequency in this context refers to the number of mutations predicted to have a high impact on protein translation as a function protein length. Only two high-impact mutations were found at low frequency, 1.5% and 8.0%, in the population for the locus Potri.005G018000 (G-type lecRLK) associated with susceptibility. The first is a premature stop codon at position 1,441,171 bp (G→A) on chromosome 5 that is predicted to truncate the protein to ∼5% of its length. The second is a frame shift at position 1,443,941 bp (AGGG→AGG), which is predicted to result in a premature stop codon truncating the protein to ∼75% of its length. As expected, the individuals with these rare alleles were more resistant to the pathogen.

Transcriptome changes of resistant (BESC-22) and susceptible (BESC-801) genotypes were determined at 0, 24, and 72 h postinoculation (hpi) with *S. musiva*. The BESC-22 genotype was chosen for carrying functional alleles of the resistance-associated loci (RLP1, RLP2, and the L-type lecRLK) and a defective allele of the susceptibility-associated locus (G-type lecRLK). In contrast, BESC-801 was selected for carrying a functional allele of the susceptibility-associated locus (G-type lecRLK) and defective alleles of the resistance-associated loci (RLP1, RLP2, and the L-type lecRLK). Comparisons were made within genotypes (at different time points) and between genotypes (at the same time points). In total, 4,686 genes were differentially expressed between the 0- and 24-hpi time points in the resistant genotype (Dataset S1, Table S5) compared with 76 in the susceptible genotype (Dataset S1, Table S6). Additionally, 16 of the 62 GWAS candidates exhibited differential expression (Dataset S1, Tables S7 and S8). PFAM domain-enrichment analysis, comparing responses of resistant genotypes to the responses of susceptible genotypes, revealed major protein families associated with innate immunity responses with a ≥2× up-regulation in the resistant genotype.

The two RLPs and the L-type lecRLK, associated with resistance (Fig. 1), peaked in expression at 24 hpi (Fig. 3). In contrast, the three genes did not exhibit changes in expression in the susceptible

**Fig. 1.** Summary of phenotyping and allele analysis identifying resistance- and susceptibility-associated loci in the *P. trichocarpa* association-mapping population inoculated with the fungal pathogen *S. musiva*. The Manhattan plot of *P. trichocarpa* chromosome 5 region depicts significant associations of receptor-like protein 1 (RLP1). The Inset depicts the entire length of *P. trichocarpa* chromosome 5. The Manhattan plot of *P. trichocarpa* chromosome 9 region depicts significant associations of receptor-like protein 2 (RLP2). The Inset depicts the entire length of *P. trichocarpa* chromosome 9. The Manhattan plot of *P. trichocarpa* chromosome 3 region depicts significant associations of L-type lectin receptor-like kinase (L-type lecRLK). The Inset depicts the entire length of *P. trichocarpa* chromosome 3. Manhattan plot of *P. trichocarpa* chromosome 5 depicting significant associations of a G-type lectin receptor-like kinase (G-type lecRLK) with susceptibility. Each black dot on the Manhattan plots corresponds to a marker, its level of significance, and its physical position on the chromosome (in megabases). The red line represents the Bonferroni-corrected significance threshold, and the blue squares in the Insets represent the region of each chromosome immediately surrounding the significant associations that are visualized in the local Manhattan plot.
Positions of high-impact mutations (premature stop codons, frame shifts, and splice-site mutations) are indicated by red arrowheads in the three resistance-associated loci (RLP1, RLP2, and L-type lecRLK) and the susceptibility-associated locus (G-type lecRLK). The blue boxes represent the exons, the black lines represent introns, the gray boxes represent UTRs, and the black arrows represent the 5’ start position of the coding region.

Fig. 2. Positions of high-impact mutations (premature stop codons, frame shifts, and splice-site mutations) are indicated by red arrowheads in the three resistance-associated loci (RLP1, RLP2, and L-type lecRLK) and the susceptibility-associated locus (G-type lecRLK). The blue boxes represent the exons, the black lines represent introns, the gray boxes represent UTRs, and the black arrows represent the 5’ start position of the coding region.

Fig. 3. Comparison of normalized gene counts of the four loci with the strongest associations to resistant (BESC-22) (RLP1, RLP2, and L-type lecRLK) and susceptible (BESC-801) (G-type lecRLK) interactions between P. trichocarpa and S. musiva across three time points (0, 24, and 72 hpi). Error bars represent the SEM for the three biological replicates.

Fig. 4. Lectin-binding assays with the G-type and L-type lectin domains. Shown are lectin-binding assays of GFP-L-type (amino acids 30–283) and G-type (amino acids 36–318) lecRLKs as a fusion to “superfolder” GFP in HEK293 cells (SI Appendix, Fig. S8) (23, 24). The expressed proteins were purified and subsequently incubated with cell wall fractions of S. musiva. Microcrystalline cellulose was used as a binding substrate control, and a noncatalytic fragment of Arabidopsis ERK1 was used as a protein control in all experiments. The G-type and L-type lectin domains specifically bound to cell wall preparations of S. musiva but not to the controls, indicating specificity for fungal cell wall carbohydrates or proteoglycans (Fig. 4). The G-type lectin bound a larger proportion of the cell wall fractions than the L-type lectin regardless of treatment. Interestingly, binding of the L-type lectin to S. musiva increased significantly after the walls were treated with KOH, indicating that recognition of the ligand is restricted by either alkaline-extractable cell wall components or esterification (25, 26). Very few LecRLKs have been functionally characterized. Ligand identification has been challenging, due to difficulties in expressing and purifying high-quality, functional preparations of these highly glycosylated eukaryotic proteins (SI Appendix, Fig. S8).

In summary, we identified genes predicted to encode receptors that were significantly associated with resistance and susceptibility of defense marker genes. As predicted, defense-signaling genes were induced in response to L-type lecRLK overexpression (SI Appendix, Fig. S7A), demonstrating that the BESC-22 allele was functional as a defense inducer in the protoplast assays. We generated Populus hairy roots overexpressing the G-type lecRLK; a comparison of transcriptional responses in control, mock-inoculated (sterile distilled H2O), and S. musiva-inoculated hairy roots demonstrated that the BESC-801 allele in the presence of S. musiva triggered transcriptional repression of major defense modulators (SI Appendix, Fig. S7B) (22).

Samples in the RNA-seq experiments contained both host and pathogen transcripts. To exploit this, we examined transcriptome changes of the pathogen, which was challenging given that pathogen biomass, RNA, and read counts are low during the initial 24 hpi, resulting in low statistical power. Nonetheless, in S. musiva we identified 16 differentially expressed genes in the resistant interaction and 44 differentially expressed genes in the susceptible interaction at 24 hpi relative to 0 hpi (Dataset S1, Table S9). These genes are likely involved in mediating interactions with host plants and potentially influencing the host responses described above.

Finally, we expressed the N-terminal lectin domains of the L-type (amino acids 30–283) and G-type (amino acids 36–318) lecRLKs as a fusion to “superfolder” GFP in HEK293 cells (SI Appendix, Fig. S8) (23, 24). The expressed proteins were purified and subsequently incubated with cell wall fractions of S. musiva. Microcrystalline cellulose was used as a binding substrate control, and a noncatalytic fragment of Arabidopsis ERK1 was used as a protein control in all experiments. The G-type and L-type lectin domains specifically bound to cell wall preparations of S. musiva but not to the controls, indicating specificity for fungal cell wall carbohydrates or proteoglycans (Fig. 4). The G-type lectin bound a larger proportion of the cell wall fractions than the L-type lectin regardless of treatment. Interestingly, binding of the L-type lectin to S. musiva increased significantly after the walls were treated with KOH, indicating that recognition of the ligand is restricted by either alkaline-extractable cell wall components or esterification (25, 26). Very few LecRLKs have been functionally characterized. Ligand identification has been challenging, due to difficulties in expressing and purifying high-quality, functional preparations of these highly glycosylated eukaryotic proteins (SI Appendix, Fig. S8).

In summary, we identified genes predicted to encode receptors that were significantly associated with resistance and susceptibility of defense marker genes. As predicted, defense-signaling genes were induced in response to L-type lecRLK overexpression (SI Appendix, Fig. S7A), demonstrating that the BESC-22 allele was functional as a defense inducer in the protoplast assays. We generated Populus hairy roots overexpressing the G-type lecRLK; a comparison of transcriptional responses in control, mock-inoculated (sterile distilled H2O), and S. musiva-inoculated hairy roots demonstrated that the BESC-801 allele in the presence of S. musiva triggered transcriptional repression of major defense modulators (SI Appendix, Fig. S7B) (22).

Samples in the RNA-seq experiments contained both host and pathogen transcripts. To exploit this, we examined transcriptome changes of the pathogen, which was challenging given that pathogen biomass, RNA, and read counts are low during the initial 24 hpi, resulting in low statistical power. Nonetheless, in S. musiva we identified 16 differentially expressed genes in the resistant interaction and 44 differentially expressed genes in the susceptible interaction at 24 hpi relative to 0 hpi (Dataset S1, Table S9). These genes are likely involved in mediating interactions with host plants and potentially influencing the host responses described above.
to *S. musiva*. The population-wide allele analysis revealed that in the sampled population the loci associated with resistance harbor many high-impact mutations, potentially impairing the ability of genotypes to recognize *S. musiva* and initiate an immune response. Furthermore, the loss of function in genes encoding putative immunity receptors (RLPs and *L*-type lecRLK) in parallel with the contrasting expression responses of lox (L-type lectin) domains bind to the allopatic pathogen *S. musiva*. The genes associated with host–pathogen interactions exhibited contrasting expression responses between resistant and susceptible genotypes. Biochemical analysis demonstrated that both the G-type and *L*-type lectin domains bind to *S. musiva* cell walls. The associations and gene-expression profiles are predictive of the resistance/susceptibility phenotype. These can be further tested when *Populus* transgenic plants with combinatorial overexpression or silencing of these four loci become available. As such, the use of high-resolution phenotyping and host resequencing across the species range enabled the identification of candidate loci associated with *P. trichocarpa* response to *S. musiva*. Once confirmed in transgenic plants, these loci can be incorporated into future breeding efforts that include marker-based selection of parents and progeny resistant to Septoria stem canker to potentially accelerate the mitigation of disease in native ecosystems.

### Materials and Methods

**Plant Material.** Plant material from 1,081 *P. trichocarpa* genotypes, originally collected from wild populations in California, Oregon, Washington, and British Columbia, were planted in a stool bed at the Oregon State University Research Farm in Corvallis, OR (10). During January 2014, dormant branch cuttings were collected and sent to the North Dakota State University’s Agricultural experiment station research greenhouse complex in Fargo, ND. For each genotype, branches were cut into 10 cm length, with at least one bud. Cuttings were soaked in distilled water for 48 h, planted in cone-tainers (Ray Leach SC10 Super Cone-tainers; Stuewe and Sons, Inc.) in a 3.8-cm diameter and 21 cm deep filled with growing medium (SunGro Professional Mix #8; SunGro Horticulture Ltd.) amended with 2 g of Nutriote slow release fertilizer (15:12:19 (N-P-K) (7.0% NH₄₂N, 8.0% NO₃-N, 9.0% PO₄₂O₁₂, 12.0% K₂O, 1.0% Mg, 2.3% S, 0.02% B, 0.05% Cu, 0.05% Fe, 0.23% chelated Fe, 0.06% Mn, 0.02% Mo, 0.05% Zn; Scotts Osmocote Plus; Scotts Company Ltd.). The cuttings were planted so that the uppermost bud remained above the surface of the growing medium. Plants were grown in a greenhouse with a night/daytime temperature regime of 20/16°C and an 18-h photoperiod supplemented with 600 W high-pressure sodium lamps. Slow-release fertilizer was added weekly with 15-30-15 (N-P-K) Jack’s fertilizer (Ir. Peters, Inc.) at 200 ppm for 2 mo to promote root growth, and plants subsequently were fertilized with 20-20-20 (N-P-K) liquid fertilizer (Scotts Peters Professional; Scotts Company, Ltd.) once a week. Plants were watered as needed.

**Pathogen Culture.** Three isolates of *S. musiva* (MN-12, MN-14, and MN-20) collected from three separate trees near Garfield, MN, were chosen for inoculation, based on preliminary virulence testing, and were transferred from storage (−80 °C) onto *K*-V8 [180 mL V8 juice (Campbell Soup Company), 2 g calcium carbonate, 20 g agar, and 820 mL deionized water] growth medium, sealed with Parafilm (Structure Probe, Inc.). Petri plates were placed on a light bench under full-spectrum fluorescent bulbs (Syvania; Osram GmbH) at room temperature until sporulation was observed. Following sporulation, five 5-mm plugs were transferred onto another *K*-V8 plate and grown for 14 d under continuous light. There were total of 200 plates for each isolate.

**Inoculation for GWAS.** The experimental design was a randomized complete-block design with four blocks. Plants were inoculated when they reached a minimum height of 30 cm (~54 d after planting). Plates containing isolates were unsealed, and ~1 mL of deionized water was added to the plate. Rubbing the medium surface with an inoculation loop dislodged the spores, and the spor suspension was collected with a pipette (7). The spor suspensions were individually bulked from the three isolates at a concentration of 10⁶ spores/mL for each isolate. Plants were taken out of the greenhouse, and their heights were measured before inoculation. They were sprayed with a high-pressure, low-volume gravity-fed air-spray gun (Central Pneumatic; Harbor Freight Tools) at 20 psi until the entire leaf and stem surface was wet (15 mL) and were placed into a black plastic bag for 48 h. Following incubation plants were placed on the greenhouse bench for 3 wk.

### Phytotyping

At 3 wk postinoculation, phenotypic responses were characterized by measuring the height of each tree. Subsequently, the number of cankers was counted, and digital images were acquired. This information was analyzed providing a range of phenotypes: (i) number of cankers; (ii) number of cankers/cm, and (iii) disease severity based on digital imagery. Initially the number of cankers and number of cankers/cm were used for the GWAS phytotyping. The order of individuals selected for phytotyping from each block was done randomly. Broad-sense heritability was estimated from mean squares estimates derived from an ANOVA analysis based on 426 genotypes with all four replicates.

**GWAS Analysis.** Whole-genome resequencing, SNP/indel calling, and SNP/eff for the 545 individuals of this *Populus* GWAS population were previously described (10). In this study, we used the same sequencing and analytical pipelines to incorporate an additional 337 genotypes. The resulting SNP and indel dataset is available at [https://bioenergycenter.org/besc/gwas](https://bioenergycenter.org/besc/gwas). To assess genetic control, we used the EMMA algorithm in EMMAX software (University of Michigan) with kinship as the correction factor for genetic background effects (27) to compute genotype-to-phenotype associations using 8,253,066 SNP variants with minor allele frequencies >0.05 identified from whole-genome resequencing (10). Four independent replicates of absolute canker numbers and number of cankers/cm were used as phenotypes. A $P$ value threshold of $6.1 \times 10^{-9}$ ($0.05/8,253,066$) was used to determine significant GWAS $P$ values from combined datasets. One SNP was identified using quantile–quantile (QQ) plots with lambda ($\lambda$) as the test statistic. Pairwise LD around the four candidate receptors was established using SNPs 5 kb upstream and downstream of the position with the lowest $P$ value.

**RNA-seq Experiment.** The resistant genotype BESC-22 and the susceptible genotype BESC-801 were selected based on the results from the GWAS analysis described above. The experimental design was a randomized complete-block design with three blocks. Each plant-by-time point combination occurred once per block. Each plant was inoculated at three points. Following mRNA extraction, the samples from the three inoculation points were pooled. Each pool was considered a biological replicate for the RNA-seq experiment.

Inoculum was prepared in an identical manner to that described above. However, to ensure that only tissue exposed to the fungal pathogen was used for transcriptome sequencing, position-based inoculations at the lenticels represented the whole-tree inoculum were conducted. The inoculation point was located on the underside of a lenticel 15 mm from the cambium. Each plant was inoculated with a 5-mm plug of sporulating mycelium from isolate MN-14 and wrapped in Parafilm. At the time of sample collection tissue from all three lenticels was sampled. Approximately 100 mg of symptomatic tissue from each inoculation point was harvested, placed in a MP Biomedicals Lysing Matrix tube, and flash-frozen in liquid nitrogen. The frozen samples were placed in a 1.5-cm deep filled with growing medium (SunGro Professional Mix #8; SunGro Horticulture). The cuttings were placed in a Bead Beater homogenizer (BioSpec Products) and ground to a fine powder. The mRNA from each sample was isolated using the Dynabeads mRNA DIRECT Kit following the manufacturer’s protocol with the additional steps of adding Ambion Plant Isolation Aid to the lysis buffer as well as a chloroform cleanup step after centrifuging the lysate.

Stranded mRNA-seq library(s) were generated and quantified using qPCR. Sequencing was performed on an Illumina HiSeq 2500 (150mer paired-end sequencing). Raw fastq files were filtered and trimmed using the JGI QTrim Tool (28). Using BBduk sequences were aligned to the reference genome with bowtie2 using settings “-v 0 -i 3 -o 136 -i 1000 -d 4 -G”. Reads were evaluated for sequence artifacts by kmer matching (kmer = 25) allowing one mismatch, and detected artifacts were trimmed from the 3’ end of the reads. RNA spike-in reads, PhIX reads, and reads containing any Ns were removed. Quality trimming was performed using the Phred trimming method set at Q6. Following trimming, reads under the length threshold (minimum length 25 bases or one-third of the original read length, whichever was longer) were removed. Raw reads from each library were aligned to the reference genome (28, 29) using TopHat (30). Only reads that mapped uniquely to one locus were counted. FeatureCounts (30) was used to generate raw gene counts. DESeq2 (v1.2.10) (30) was subsequently used to determine which genes were differentially expressed between pairs of conditions. The parameters used to call a gene between conditions was determined at a $P$ value <0.05. RNA-seq differential expression analysis for *S. musiva* was performed using the Tuxedo suite pipeline (31). Illumina short paired reads were trimmed for quality using Sickle (31) set with a minimum quality score cutoff of 30 and a minimum read length 100 bp. Using TopHat v2.1.0 (30) and Bowtie2 v2.2.3 (2), we aligned trimmed reads for each sample replicate to combined assembly contigs from *S. musiva* strain SO2202 (GenBank accession no. GCA_000320565.2) and *P. trichocarpa* (GenBank accession no. GCF_000002775.3; https://phytopathogen.jgi-doe.gov). Reads were mapped with settings “+ 0 136 -i 1000 -d 4 -G” with combined gene annotations from the *S. musiva* and *P. trichocarpa* reference genomes. *S. musiva* contigs and mapped reads were extracted using SAMTools v0.1.18. Transcript isoforms for each of the sample replicates were individually
assembled and quantified using Cufflinks v2.2.1 (30) guided by the S. musiva reference genome and gene annotations. Transcripts assembled from each alignment were merged using Cuffmerge (30).

Differential gene-expression analysis was performed using Cuffdiff (30). Time-series comparisons were performed for the resistant interaction between BESC-22 and S. musiva at 24 and 72 hpi and the susceptible interaction with BESC-801 and S. musiva at 24 and 72 hpi, with three replicates per time point. These analyses excluded 0 hpi due to low sequencing depth for S. musiva. Differential expression analyses were also performed comparing gene expression at 24 and 72 hpi between the resistant and susceptible interactions.

**Generation of Constructs for Protein Expression.** The predicted lectin domains of G-type lecRLK and L-type lecRLK were cloned (23). Briefly, to create Gateway entry clones, truncated coding regions of G-type lecRLK (amino acids 36-192) and L-type lecRLK (amino acids 30-281) were amplified from P. trichocarpa cDNA using the following gene-specific primer pairs: GRLK30-36F, 5′-AACATGTTACTTCTGCAAGCGTTC-3′ and LRLK2-30F, 5′-AACTTCTGTCAGAATGCT-3′. These consisted of three sequential wash steps of 0%, 10%, and 20% sodium chloride, and 20 mM imidazole. The columns were washed and eluted with a 1 M filter (Millipore).

**Expression and Purification of His-GFP-G-Type lecRLK36-192 and His-GFP-L-Type lecRLK30-281.** Recombinant expression of G-type lecRLK (pGEn2-EXP-G-type lecRLK36-192) and L-type lecRLK (pGEn2-EXP-L-type lecRLK30-281), the entry clone was recombined into a Gateway-adapted version of the pGEn2 mamalian expression vector (pGEn2-DEST) (25) using Gateway LR Clonase II Enzyme Mix (Life Technologies) to create entry clones. To generate expression clones of G-type lecRLK (pGEN2-EXP-G-type lecRLK36-192) and L-type lecRLK (pGEN2-EXP-L-type lecRLK30-281), the entry clones were recombined into a Gateway-adapted version of the pGEn2 mammalian expression vector (pGEn2-DEST) (25) using Gateway LR Clonase II Enzyme Mix (Life Technologies) to create expression constructs (H-GFP-type lecRLK36-192 and His-GFP-type lecRLK30-281) encode fusion proteins comprised of an N-terminal signal sequence, an 6×His tag, an AviTag recognition site, the superfolder GFP (sfGFP) coding region, the recognition sequence of the TEV protease, and the indicated lectin domains. For transfection, plasmids were purified using the PureLink HiPure Plasmid Filter Maxiprep Kit (Life Technologies).
RevertAid Reverse Transcriptase (Thermo Fisher). qRT-PCR was performed using 3 ng cDNA, 250 nM gene-specific primers, and iTaq Universal SYBR Green Supermix (Bio-Rad). Gene expression was calculated by the 2ΔΔCt method using UBQ10b as the internal control.

Primers.

The following primers were used:

UBQ10b_F GCCCTCTGTTGTTATTAAGC
UBQ10b_R TCCCAAATGTTGACGTAACAC
BAKa1_F TGGCATCCTGATGAAGACAG
BAKa1_R AAGGATCTCAACACTTACGC
BAKb_F GGAGATGGCATTGGAGAAC
BAKb_R GCCGAAAAAGTACGAAATCC
WRKY40_F CATGGATGTCTTTCTCCCTTG
WRKY40_R TTCTCTTCTGCTTCGTTCC
WRKY70a_F ACTATCATACAAGCGAGAAAGG
WRKY70a_R TTCTGAGGCGAATTTGAAG

ACKNOWLEDGMENTS. This study was supported by Department of Energy (DOE) Office of Science, Office of Biological and Environmental Research (BER) Grant DE-SC0018196, US Department of Agriculture Grant 2012-34033-19971 (to J.M.L.), the Plant–Microbe Interfaces Scientific Focus Area in the DOE BER Genomic Science Program, and the DOE Center for Bioenergy Innovation Project. The Center for Bioenergy Innovation is a US DOE Bioenergy Research Center supported by the Office of BER in the DOE Office of Science. Oak Ridge National Laboratory is managed by UT-Battelle, LLC for the US DOE under Contract DE-AC05-00OR22725. The work conducted by the US DOE Joint Genome Institute is supported by the Office of Science of the US DOE under Contract DE-AC02-05CH11231. This work was also supported by Natural Science Foundation of China, Department of Agriculture of Agriculture Innovation Project. The Center for Bioenergy Innovation is a US DOE Bioenergy Research Center supported by the Office of BER in the DOE Office of Science. Oak Ridge National Laboratory is managed by UT-Battelle, LLC for the US DOE under Contract DE-AC05-00OR22725. The work conducted by the US DOE Joint Genome Institute is supported by the Office of Science of the US DOE under Contract DE-AC02-05CH11231. This work was also supported by National Key Research and Development Program of China. This research was supported by the US DOE Joint Genome Institute is supported by the Office of Science of the US DOE under Contract DE-AC02-05CH11231. This work was also supported by National Key Research and Development Program of China. The authors would like to acknowledge the DOE-funded Center for Plant and Microbial Complex Carbohydrates Grant DE-SC0015662 (DE-FG02-93ER20097) for equipment support. This research used resources of the Compute and Data Environment for Science (CADES) at the Oak Ridge National Laboratory.

1. Jones JD, Dangl JL (2006) The plant immune system. Nature 444:323–329.
2. Schaller GG, Klessig DF (2000) Evolution of plants and their pathogens in natural habitats. Science 294:575–576.
3. Anagnostakis SL (1987) Chestnut blight: The classical problem of an introduced pathogen. Mycologia 79:23–37.
4. Kinloch BB, Jr (2003) White pine blister rust in North America: Past and prognosis. Phytopathology 93:1044–1047.
5. Cingolani P, et al. (2012) A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of Drosophila melanogaster strain w1118; iso-2; iso-3. Fly (Austin) 6:80–92.
6. Chinchilla D, Shan L, He P, de Vries S, Kemmerling B (2009) One for all: The receptor-like kinase family from Arabidopsis and rice. Plant Mol Biol 60:356–388.
7. Xu X, Chen C, Fan B, Chen Z (2006) Physical and functional interactions between pathogen-induced Arabidopsis WRKY18, WRKY40, and WRKY60 transcription factors. Plant Cell 18:1310–1326.
8. Zhang X, et al. (2012) A gain-of-function mutation in a plant disease resistance gene leads to constitutive activation of downstream signal transduction pathways in suppressor of npr1-1, constitutive 1. Plant Cell 15:2636–2646.
9. Zhang Y, et al. (2010) Arabidopsis sn_o2-10 activates receptor-like protein-mediated immunity transduced through WRKY70. Plant Cell 22:3152–3163.
10. Xu X, Chen C, Fan B, Chen Z (2006) Physical and functional interactions between pathogen-induced Arabidopsis WRKY18, WRKY40, and WRKY60 transcription factors. Plant Cell 18:1310–1326.