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

Wellington Muchero[a], Kelsey L. Sondreli[b], Jin-Gui Chen[a], Breeanna R. Urbanowicz[c], Jin Zhang[a], Vasanth Singan[d], Yongil Yang[a], Robert S. Brueggeman[e], Juan Franco-Coronado[e], Nivi Abraham[e], Jeong-Yeh Yang[c], Kelley W. Moremen[c], Alexandra J. Weisberg[b], Jeff H. Chang[b], Erika Lindquist[d], Kerrie Barry[d], Priya Ranjan[a], Sara Jawdy[a], Jeremy Schmutz[d,f], Gerald A. Tuskan[a,d], and Jared M. LeBoldus[b,e,g,1]

[a]Biosciences Division, Oak Ridge National Laboratory, Oak Ridge, TN 37831; [b]Department of Botany and Plant Pathology, Oregon State University, Corvallis, OR 97331; [c]Complex Carbohydrate Research Center, University of Georgia, Athens, GA 30602; [d]Joint Genome Institute, US Department of Energy, Walnut Creek, CA 94598; [e]Department of Plant Pathology, North Dakota State University, Fargo, ND 58102; [f]HudsonAlpha Institute for Biotechnology, Huntsville, AL 35806; and [g]Forest Engineering, Resources, and Management, Oregon State University, Corvallis, OR 97331

Edited by Ronald R. Sederoff, North Carolina State University, Raleigh, NC, and approved September 24, 2018 (received for review March 14, 2018)

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*.

septoria canker | invasive disease | association mapping | disease resistance | *Populus trichocarpa*

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 counteradaption, 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 sympatries (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

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 reagents/analytic 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.

Data deposition: The sequences reported in this paper have been deposited in the National Center for Biotechnology Information Sequence Read Archive (accession nos. SRX2502320–SRX2502340).

[1]To whom correspondence should be addressed. Email: jared.leboldus@science.oregonstate.edu.

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 paralogous leucine-rich receptor-like proteins (RLPs) (Potri.005G012100, *P* value = 1.56E-38; Potri.003G028200, *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 lectin domain(s) as C-, G-, and L-type (15, 16). Lectins 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 nonsynonymous and deleterious mutations (i.e., early translation termination, frameshifts, 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× upregulation 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

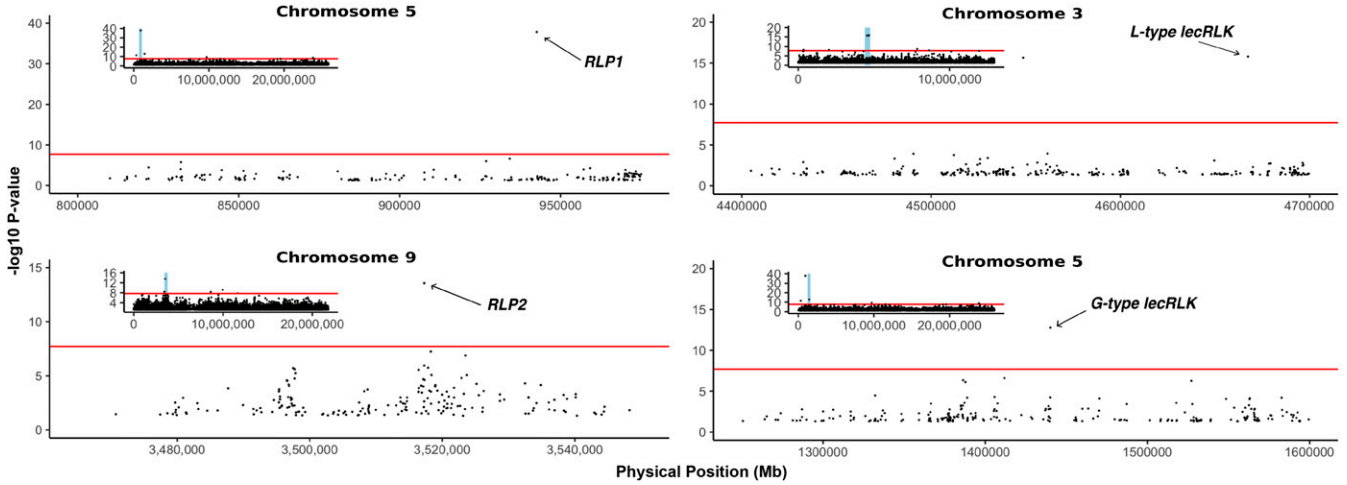

Fig. 1. Summary of phenotyping and allele analysis identifying resistance- and susceptibility-associated loci in the *P. trichocarpa* association-mapping population inoculated with the fungal pathogen *S. musiva*. The Manhattan plot of *P. trichocarpa* chromosome 5 region depicts significant associations 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 association 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.

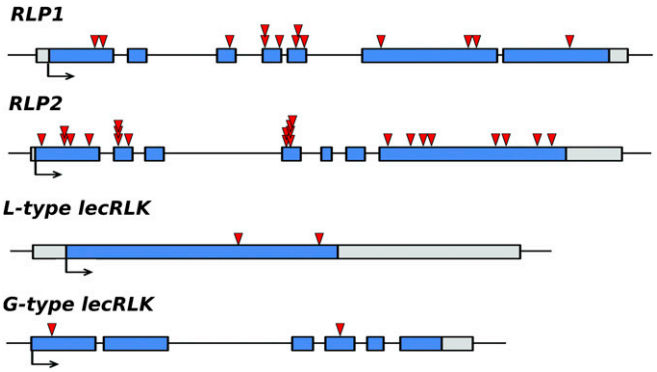

Fig. 2. Positions of high-impact mutations (premature stop codons, frame shifts, and splice-site mutations) are indicated by red arrowheads in the three resistance-associated loci (*RLP1*, *RLP2*, and *L-type lecRLK*) and the susceptibility-associated locus (*G-type lecRLK*). The blue boxes represent 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.

genotype, regardless of the times compared. In the susceptible genotype, the *G-type lecRLK*, associated with susceptibility (Fig. 1), was expressed at each examined time point. In the resistant genotype, expression of the *G-type lecRLK* was barely above the detectable threshold (Fig. 3). The changes in expression of six genes commonly used as markers for transcriptional reprogramming during host resistance (*SI Appendix*, Fig. S6) were also compared between resistant and susceptible genotypes at 0, 24, and 72 hpi. All six of the marker genes peaked at 24 hpi in the resistant genotype. In the susceptible genotype, the six markers were expressed at statistically similar levels. The pattern of expression of all six marker genes is consistent with defense-response signaling in plants described in the literature (*SI Appendix*, Fig. S6) (19–22).

We performed overexpression analysis using *Populus* leaf mesophyll protoplasts and evaluated the transcriptional response of defense marker genes. As predicted, defense-signaling genes were induced in response to *L-type lecRLK* overexpression (*SI Appendix*, Fig. S7*A*), 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 H_2O), 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. S7*B*) (22).

Samples in the RNA-sequencing (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

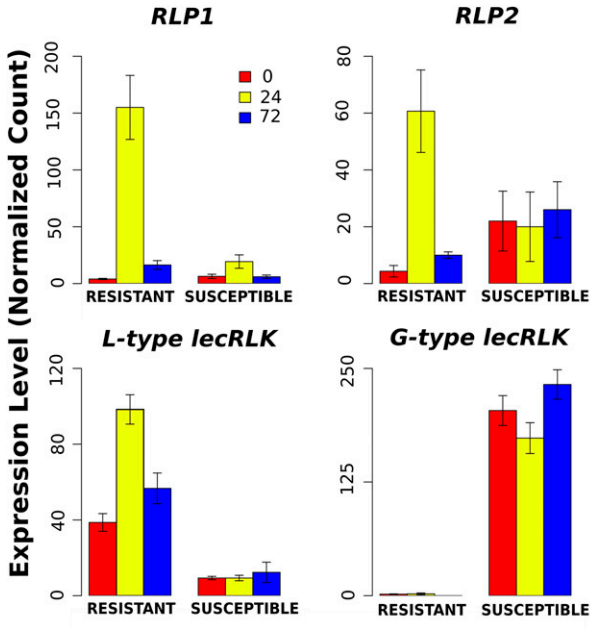

Fig. 3. Comparison of normalized gene counts of the four loci with the strongest associations to resistant (BESC-22) (*RLP1*, *RLP2*, and *L-type lecRLK*) and susceptible (BESC-801) (*G-type lecRLK*) interactions between *P. trichocarpa* and *S. musiva* across three time points (0, 24, and 72 hpi). Error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

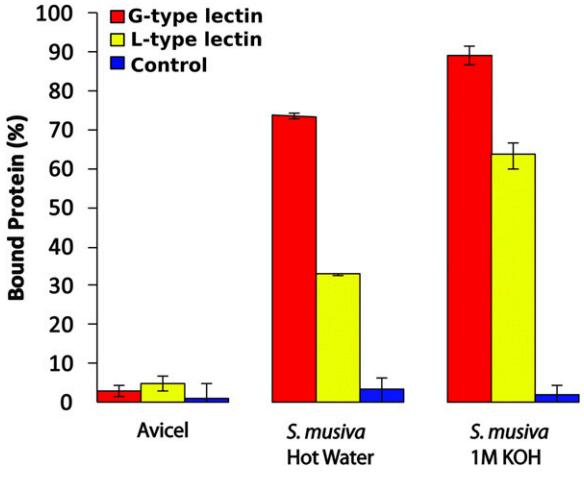

Fig. 4. Lectin-binding assays with the G-type and L-type lectin domains. Shown are lectin-binding assays of GFP-L-type (amino acids 30–283) and GFP-G-type (amino acids 36–318) lectins to sequentially extracted *S. musiva* cell walls. GFP-ESK1Δ44–133 was used as a control. GFP fluorescence was used to quantify the total percent of bound proteins using the depletion method. Values are the means ± the SD of triplicate reactions.

GENETICS

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 conservation of a locus implicated in susceptibility (*G-type lecRLK*) results in population-wide susceptibility of *P. trichocarpa* to the allopatric 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 *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 cuttings, measuring 10 cm in 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.) measuring 3.8-cm in diameter and 21-cm deep filled with growing medium (SunGro Professional Mix #8; SunGro Horticulture Ltd.) amended with 12 g of Nutricote slow release fertilizer (15-9-12) (N-P-K) (7.0% NH_3-N, 8.0% NO_3-N, 9.0% P_2O_5, 12.0% K2O, 1.0% Mg, 2.3% S, 0.02% B, 0.05% Cu, 0.45% 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 day/night temperature regime of 20 °C/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 (Jr. 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 (Sylvania; 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 spore suspension was collected with a pipette (7). The spore suspensions were individually bulked from the three isolates at a concentration of 10^6 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.

Phenotyping. 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 phenotyping. The order of individuals selected for phenotyping 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 SNPeFF analysis 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/. 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 6.1×10^{-09} (0.05/8,253,066) was used to determine significance. Deviation of P values from expectation was evaluated using quantile–quantile (QQ) plots with lambda (λ) 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 rather than whole-tree inoculations were conducted. Three lenticels on each plant were 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 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 RNA-seq library(s) were generated and quantified using qPCR. Sequencing was performed on an Illumina HiSeq 2500 (150mer paired-end sequencing). Raw fastq file reads were filtered and trimmed using the JGI QC pipeline. Using BBDuk (https://sourceforge.net/projects/bbmap/), raw 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 of 100 bp. Using TopHat v2.1.0 (30) and Bowtie2 v2.2.3 (32), 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://phytozome.jgi.doe.gov/). Reads were mapped with settings "-r 0 -i 36 -I 1000 -p 4" and "-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: G-RLK1-36F, 5′-AACTTGTACTTTCAAGGCCAGTCTCTCTCTGCAAGC-3′/G-RLK1-192R, 5′-ACAAGAAAGCTGGGTC**CTA**ACCTGGTGCAGGATCTT-3′ and L-RLK2-30F, 5′-AACTTGTACTTTCAAGGCCACTTCATCTATCATGG-3′/L-RLK2-281, 5′-ACAAGAAAGCTGGGTC**CTA**AGGCAACTTTGACACATC-3′. The control protein was a noncatalytic peptide fragment of *Arabidopsis* ESK1 (amino acids 44–133) and was amplified from *Arabidopsis* cDNA using the following gene specific primer pairs: ESK1-44F, 5′-AACTTGTACTTTCAAGGCGTGGAATTGCCGCCG-3′/ESK1133R, 5′-ACAAGAAAGCTGGGTC**CTA**CGAACGGGAAATGATAC-3′. Underlined sequences indicate the partial attB adapter sequences appended to the primers for the first round of PCR amplification, and bold sequences denote the inserted STOP codon. A second set of universal primers, attB_Adapter-F, 5′- GGGGACAAGTTTGTACAAAAAAGCAGGCTCTGAAAACTTGTACTTTCAAGGC-3′/attB_Adapter-R, 5′-GGGGACCACTTTGTACAAGAAAGCTGGGTC-3′, was used to complete the attB recombination site and append a tobacco etch virus (TEV) protease cleavage site (24). The attB-PCR product was cloned into pDONR221 (Life Technologies) using Gateway BP 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). The resulting expression constructs (His-GFP-G-type lecRLK$^{\Delta 36–192}$ and His-GFP-L-type lecRLK$^{\Delta 30–281}$) encode fusion proteins comprised of an N-terminal signal sequence, an 8xHis 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).

Expression and Purification of His-GFP-G-Type lecRLK36-192 and His-GFP-L-Type lecRLK30-281. Recombinant expression was performed by transient transfection of suspension cultured HEK293-F cells (FreeStyle 293-F cells; Thermo Fisher Scientific) in a humidified CO_2 platform shaker incubator at 37 °C with 80% humidity. The HEK293-F cells were maintained in Freestyle 293 expression medium (Thermo Fisher Scientific), and transfection with plasmid DNA using polyethyleneimine as transfection reagent (linear 25-kDa polyethyleneimine; Polysciences, Inc.) was performed as previously described (22, 23). After 24 h, the cell cultures were diluted 1:1 with fresh medium supplemented with valproic acid (2.2 mM final concentration), and protein production was continued for an additional 4–5 d at 37 °C. The cell culture was harvested, clarified by sequential centrifugation at 241 × *g* for 10 min and 2,054 × *g* for 20 min, and passed through a 0.45-μM filter (Millipore).

All chromatography experiments were carried out on an ÄKTA FPLC System (GE Healthcare). The medium was adjusted to contain Hepes (50 mM, pH 7.2), sodium chloride (400 mM), and imidazole (20 mM) before column loading. Small-scale purification of His8-GFP–tagged enzymes secreted into the culture medium by HEK293 cells was performed using HisTrap HP columns (GE Healthcare). To eliminate the possibility of protein contamination, purification of each enzyme was carried out on an individual 1-mL HisTrap HP column. Before use, a blank run was performed on each new column to remove any weakly bound Ni^{2+} ions. Adjusted medium was loaded onto HisTrap HP columns (GE Healthcare) equilibrated with buffer A [50 mM Hepes (pH 7.2), 0.4 M sodium chloride, and 20 mM imidazole]. The columns were washed and eluted with a step gradient consisting of five column volumes per condition of buffer A to buffer B [50 mM Hepes (pH 7.2), 0.4 M sodium chloride, and 500 mM imidazole]. These consisted of three sequential wash steps of 0%, 10%, and 20% buffer B, followed by two elution steps of 60% and 100% buffer B. Fractions containing GFP fluorescence (60% buffer B elution) were collected and pooled. Protein purity was assessed by SDS-PAGE. Purified His-GFP-G-type lecRLK36-192 and His-GFP-L-type lecRLK30-281 were concentrated to ∼1.5 mg/mL using a 30-kDa–cutoff Amicon Ultra centrifugal filter device (Merck Millipore, www.emdmillipore.com/US/en?bd=1) and dialyzed (3,500 molecular weight cutoff) into binding buffer without divalent metals [75 mM Hepes·HCl (pH 6.8), 150 mM NaCl] in the presence of Chelex 100 Molecular Biology Grade Resin (1 g/L; Bio-Rad) and used directly for binding experiments. Protein concentrations were determined with the Pierce BCA Protein Assay Kit (Thermo Fisher Scientific) and BSA standards.

Growth of *S. musiva* in Liquid Culture. Sporulating 1-wk-old *S. musiva* cultures growing on solid K-V8 medium [180 mL/L V8 juice, 2 g/L CaCO3, 2% (vol/vol) agar] were rinsed with 1 mL of sterile double-distilled water, and the conidia were dislodged with an inoculating loop. For the inoculation of the liquid cultures, 200-μL aliquots of the spore suspensions were pipetted into 100 mL of liquid K-V8 medium in 250-mL Erlenmeyer flasks. The cultures were incubated at ambient temperature in darkness for 5 d. During the incubation, the cultures were constantly agitated at 150 rpm with an orbital platform shaker (Innova 2100; New Brunswick). To harvest the mycelium, the cultures were filtered with Miracloth (EMD Millipore). The harvested mycelium was rinsed with 50 mL of double-distilled water and squeezed dry by pressing the mycelium inside the Miracloth between stacks of paper towels. Finally, 50-mg mycelium samples were collected and lyophilized for lectin-binding assays.

Analysis of Lectin Binding to *S. musiva* Cell Walls. To evaluate the ability of recombinant plant lectins to bind to *S. musiva*, cell walls from cultured fungi were sequentially extracted with cold water, hot water, and aqueous KOH as described in ref. 33, with minor modifications. Briefly, freeze-dried fungal mycelium was resuspended in cold water (100 mL/g) containing sodium azide (0.02%) and was extensively homogenized using a Polytron homogenizer (Brinkmann Instruments) in a cold room at 4 °C. The homogenate was centrifuged at 15,302 × *g* for 15 min, and the pellet was washed extensively with cold water. The debris containing the cell walls was resuspended in hot water containing sodium azide (0.02%), homogenized again, and incubated at 60 °C overnight in a shaking incubator (250 rpm). The pellets were collected again by centrifugation, treated with hot water for 1 h, and centrifuged again. This was repeated another two times. The washed pellets were resuspended in 1 M KOH containing sodium borohydride (1%) and were incubated overnight at 30 °C. Next, residues were pelleted again and washed extensively with water. A portion of the hot water and KOH insoluble *S. musiva* cell walls was collected, washed extensively with acetone, and air-dried under vacuum.

Lectin-binding assays were carried out based on the methods of Lim et al. (34) with minor modifications. Microcrystalline cellulose (Avicel PH-101; Sigma-Aldrich) was used as a control substrate for all binding assays. For lectin pull-down assays, 2 mg of each dry substrate was carefully weighed into tubes. Then 250 μL of protein (50 μg/mL) in lectin-binding buffer [75 mM Hepes·HCl (pH 6.8), 150 mM NaCl, 5 mM $MnCl_2$, 5 mM $CaCl_2$, 1 mg/mL BSA] was added, and samples were incubated for 2 h at room temperature with end-over-end rotation. Samples were centrifuged (13,523 × *g*, 5 min), and 100 μL of the supernatants containing the unbound proteins was assayed for GFP fluorescence (excitation, 415 nm; emission, 550 nm). The percent of bound enzyme was calculated by the depletion method (35).

In Vivo Overexpression of *L-Type lecRLK* and *G-Type lecRLK* in *Populus* Protoplasts. For protoplast transfection, protoplasts from *P. tremula* × *P. alba* clone INRA 717-1-B4 were isolated and subsequently transfected, as previously described (36). For overexpression, 10 μg of *L-type lecRLK* constructs with a 35S promoter and vector control were transfected into 100 μL of protoplasts. After 12 h incubation, protoplasts were collected by 2-min centrifugation at 2,000 × *g* and were frozen in liquid nitrogen for the qRT-PCR experiment.

Generation of Transgenic *Populus* Hairy Roots. To generate binary vectors of *G-type lecRLK* for hairy roots transformation, the cDNA sequence was first cloned into pENTR/D TOPO vectors and then into the pGWB402omega binary vector by LR recombination reaction. The binary vector was transformed into *Agrobacterium rhizogenes* strain ARqua1 by electroporation, and hairy roots were generated by transforming *P. tremula* × *P. alba* clone INRA 717-1-B4 with *A. rhizogenes* (37). Hairy-root cultures were inoculated with *S. musiva* in a manner similar to that described above. Briefly, each plate was sprayed with a suspension of 1×10^6 spores/mL of *S. musiva* isolate MN-14. The mock-inoculated roots were sprayed with sterile distilled water. After a 24-h incubation period samples were flash-frozen in liquid nitrogen for RNA extraction and the qRT-PCR experiment.

RNA Extraction and qRT-PCR. RNA was extracted from protoplasts and hairy-roots samples using the Sigma plant RNA extraction kit. cDNA synthesis was performed using DNase-free total RNA (1.5 μg), oligo dT primers, and

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 GCCTTCGTGGTGGTTATTAAGC

UBQ10b_R TCCAACAATGGCCAGTAAACAC

BAK1a_F TGGCATCCTGATGAGAACAG

BAK1a_R AAAGGTCCAAACCACTTACGC

BAK1b_F GGAGATGGCATTTGTGAAGG

BAK1b_R GCTCGAAAGATGACCAATCC

WRKY40_F CATGGATGTCTTTCCCTCTTG

WRKY40_R TTCTCTTTCTGCCTGTGTTCC

WRKY70a_F ACTATCATCAAGCAGGGAAAGG

WRKY70a_R TTCTGGAGGCGAATTTGAAG

WRKY70b_F GAATCTGCTGATTTCGATGATG

WRKY70b_R AGGCGGAAATTACAAAGAAGC

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 2012-34103-19771 (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 National Institute of Food and Agriculture, US Department of Agriculture Award 2014-51181-22384 (to J.H.C.). J.-Y.Y. and K.W.M. were supported by NIH Grants P41GM103390 and P01GM107012. We would also 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. Burdon JJ, Thrall PH (2009) Coevolution of plants and their pathogens in natural habitats. *Science* 324:755–756.
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. Cobb RC, Filipe JAN, Meentemeyer RK, Gilligan CA, Rizzo DM (2012) Ecosystem transformation by emerging infectious disease: Loss of large tanoak from California forests. *J Ecol* 100:712–722.
6. Harvell CD, et al. (2002) Climate warming and disease risks for terrestrial and marine biota. *Science* 296:2158–2162.
7. Qin R, Stanosz GR, LeBoldus JM (2014) A nonwounding greenhouse screening protocol for prediction of field resistance of hybrid poplar to Septoria canker. *Plant Dis* 98:1106–1111.
8. Callan B, Leal I, Foord BM, Dennis JJ, van Oosten C (2007) *Septoria musiva* isolated from cankered stems in hybrid poplar stool beds, Fraser Valley, British Columbia. *N Am Fungi* 2:1–9.
9. Feau N, Mottet MJ, Périnet P, Hamelin RC, Bernier L (2010) Recent advances related to poplar leaf spot and canker caused by *Septoria musiva*. *Can J Plant Pathol* 32:122–134.
10. Evans LM, et al. (2014) Population genomics of *Populus trichocarpa* identifies signatures of selection and adaptive trait associations. *Nat Genet* 46:1089–1096.
11. Xing Y, Frei U, Schejbel B, Asp T, Lübberstedt T (2007) Nucleotide diversity and linkage disequilibrium in 11 expressed resistance candidate genes in *Lolium perenne*. *BMC Plant Biol* 7:43.
12. Liebrand TW, et al. (2013) Receptor-like kinase SOBIR1/EVR interacts with receptor-like proteins in plant immunity against fungal infection. *Proc Natl Acad Sci USA* 110: 10010–10015, and erratum (2013) 110:13228.
13. Chen X, et al. (2006) A B-lectin receptor kinase gene conferring rice blast resistance. *Plant J* 46:794–804.
14. Navarro-Gochicoa MT, et al. (2003) Characterization of four lectin-like receptor kinases expressed in roots of Medicago truncatula. Structure, location, regulation of expression, and potential role in the symbiosis with Sinorhizobium meliloti. *Plant Physiol* 133:1893–1910.
15. Vaid N, Pandey PK, Tuteja N (2012) Genome-wide analysis of lectin receptor-like kinase family from *Arabidopsis* and rice. *Plant Mol Biol* 80:365–388.
16. Yang Y, et al. (2016) Genome-wide analysis of lectin receptor-like kinases in *Populus*. *BMC Genomics* 17:699.
17. Lannoo N, Van Damme EJM (2014) Lectin domains at the frontiers of plant defense. *Front Plant Sci* 5:397.
18. 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.
19. Chinchilla D, Shan L, He P, de Vries S, Kemmerling B (2009) One for all: The receptor-associated kinase BAK1. *Trends Plant Sci* 14:535–541.
20. 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.
21. Zhang Y, Goritschnig S, Dong X, Li X (2003) 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.
22. Zhang Y, et al. (2010) *Arabidopsis* snc2-1D activates receptor-like protein-mediated immunity transduced through WRKY70. *Plant Cell* 22:3153–3163.
23. Urbanowicz BR, Peña MJ, Moniz HA, Moremen KW, York WS (2014) Two *Arabidopsis* proteins synthesize acetylated xylan in vitro. *Plant J* 80:197–206.
24. Meng L, et al. (2013) Enzymatic basis for N-glycan sialylation: Structure of rat α2,6-sialyltransferase (ST6GAL1) reveals conserved and unique features for glycan sialylation. *J Biol Chem* 288:34680–34698.
25. Gilbert HJ, Knox JP, Boraston AB (2013) Advances in understanding the molecular basis of plant cell wall polysaccharide recognition by carbohydrate-binding modules. *Curr Opin Struct Biol* 23:669–677.
26. Marcus SE, et al. (2008) Pectic homogalacturonan masks abundant sets of xyloglucan epitopes in plant cell walls. *BMC Plant Biol* 8:60.
27. Kang HM, et al. (2010) Variance component model to account for sample structure in genome-wide association studies. *Nat Genet* 42:348–354.
28. Slavov GT, et al. (2012) Genome resequencing reveals multiscale geographic structure and extensive linkage disequilibrium in the forest tree *Populus trichocarpa*. *New Phytol* 196:713–725.
29. Ohm RA, et al. (2012) Diverse lifestyles and strategies of plant pathogenesis encoded in the genomes of eighteen Dothideomycetes fungi. *PLoS Pathog* 8:e1003037.
30. Trapnell C, et al. (2012) Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. *Nat Protoc* 7:562–578.
31. Joshi NA, Fass JN (2011) Sickle: A Sliding-Window, Adaptive, Quality-Based Trimming Tool for FastQ Files (V1.33). Available at https://github.com/najoshi/sickle. Accessed October 9, 2018.
32. Langmead B, Salzberg SL (2012) Fast gapped-read alignment with Bowtie 2. *Nat Methods* 9:357–359.
33. de Lourdes Corradi da Silva M, et al. (2008) Structural characterization of the cell wall D-glucans isolated from the mycelium of *Botryosphaeria rhodina* MAMB-05. *Carbohydr Res* 343:793–798.
34. Lim BL, Wang JY, Holmskov U, Hoppe HJ, Reid KB (1994) Expression of the carbohydrate recognition domain of lung surfactant protein D and demonstration of its binding to lipopolysaccharides of gram-negative bacteria. *Biochem Biophys Res Commun* 202:1674–1680.
35. Chundawat SP, et al. (2011) Restructuring the crystalline cellulose hydrogen bond network enhances its depolymerization rate. *J Am Chem Soc* 133:11163–11174.
36. Guo J, et al. (2012) Highly efficient isolation of *Populus* mesophyll protoplasts and its application in transient expression assays. *PLoS One* 7:e44908.
37. Yoshida K, Ma D, Constabel CP (2015) The MYB182 protein down-regulates proanthocyanidin and anthocyanin biosynthesis in poplar by repressing both structural and regulatory flavonoid genes. *Plant Physiol* 167:693–710.