PROJECT TITLE: Mitigating Needle Blight: A Growing Economic Threat to Pine Forests

PROJECT LOCATION: Auburn University

PROJECT LEAD & PRINCIPAL INVESTIGATOR: Dr. Lori G. Eckhardt, Professor & Director, Integrated Forest Pathology and Entomology, College of Forestry Wildlife and Environment, Auburn University, eckhalg@auburn.edu

CO-PIs:

Dr. Lana Narine (Auburn University, Remote Sensing and Modeling)

Dr. Janna Willoughby (Auburn University, Genetics)

Dr. Joseph Fan (Auburn University, Statistician)

J. Ryan Mitchell (Auburn University, Extension)

COOPERATORS:

Dr. Annakay Newell (Auburn University, Pathology)

Dr. Brian Via (Auburn University, Wood Products)

Dr. Iris Erramuspe (Auburn University, Wood Chemistry)

Austin Reese (Alabama Forestry Commission, Forest Health Coordinator)

Clarissa Balbalian (Mississippi State University Extension Diagnostics Lab, Plant Pathologist)

Alan Wilson (Rayonier, R&D Coordinator)

Dr. David Wilkinson (Manulife Investment Management, Silviculture Program Mgr)

Kristopher Bradley (Regions, Vice President & Forester)

Jenny Leblanc (Weyerhaeuser, Forester)

Dr. Emily Carter (USFS-R, Soil Scientist)

USFS POINT OF CONTACT: Dr. Rabiu Olatinwo, Research Plant Pathologist, Southern Research Station, Pineville, rabiu.o.olatinwo@usda.gov

LONG-TERM GOALS AND OBJECTIVES:

The long-term goals of this work are to (1) determine the distribution and movement of the needle pathogens, (2) understand the disease cycle and the environmental factors that drive the emergence and distribution of the needle pathogens, and (3) determine if the appearance is due to more aggressive strains of the pathogens.

JUSTIFICATION:

Pine forests and industrial wood plantations in the southeastern U.S. are crucial for the region's economic sustainability. In 2020, Alabama forestry sales of forest products and related sectors totaled more than $11 billion. The sustainability and profitability of these pine forests and industrial wood plantations rely on optimal tree growth. However, the continued introduction of non-native insect pests and fungal pathogens, as well as the movement of native forest pests into forest ecosystems, can result in significant economic impacts. Costs associated with damage caused by non-native pests and pathogens within forests throughout the U.S. in 2000 were estimated to be valued at approximately $4.2 billion annually. Consequently, insect pests and fungal diseases are a great concern to the forest industry.

There has been an increase of reports throughout Alabama and the southeastern U.S. of a suite of needle blight pathogens over the past ten years. This problem may not only occur on a large regional scale but also on isolated acreages, which is vital as the majority of the seven million acres of pines in Alabama are privately owned. With over one-third of the counties in Alabama currently affected, it is estimated that a 50% needle blight infection rate in Alabama's susceptible loblolly pine trees could result in economic losses of $2 billion. An investment in mitigating forest pests, such as those associated with needle blight requires adaptive management geared to prevention and remediation that provide economically sound solutions.

PREDICTED PROJECT OUTCOMES:

This work is not meant to replace any research that is currently underway focused on developing solutions, but rather is aimed at determining the actual impacts on productivity and biological cause(s) of needle blight so that landowners and forest managers may more precisely predict future timber revenues from affected stands and adjust management activities accordingly.

Results from the project shall provide:

1. A collection of factors to account for losses (tree death as well as predicted growth losses) from brown spot needle blight in loblolly in productivity models.
2. An improved understanding of the interactive effect of fungal infection, stand environment, and tree physiology on loblolly pine sustainability which is required for developing remedial actions and productivity models for trees and stands already affected.

3. The levels of infection that are acceptable (minimal growth loss and low probability of mortality) and those that fall above the damage thresholds.
4. An understanding of tree-level infection levels.
5. An understanding of the genetic variability of the fungus and how it is related to infection level and severity.
6. A screening protocol for testing seedlings to find families tolerant to the pathogen.
7. Distribution and movement of the pathogen across the southeast.

The knowledge produced by this project will be used to develop best management practices for areas affected by needle blight. It shall also help direct future research actions, especially when little is known regarding the impact of the pests/pathogens associated with loblolly pine in the southeastern United States.

APPROACH:

Our approach to this integrated project will consist of four research and one extension components. Research component one is specifically designed to determine strain aggressiveness and develop a screening protocol which addresses Objective 2 and 3. Research component two is specifically designed to determine the distribution and infection level per county across Alabama and what environmental factors drive its emergence and exacerbate the pathogen addressing portions of Objectives 1 and 2. Research component three is specifically designed to use remote sensing to detect the pathogen and determine environmental factors (i.e. Hurricanes) that facilitate its distribution which addresses portions of Objectives 1 and 2. Research component four is specifically designed to understand the genetic diversity and relatedness in and between *L. acicola, Pinus* hosts and other needle pathogens associated with BSNB infection and mortality, which addresses Objectives 3.

<u>Plot Selection and Needle Collection</u>:

The forest health dynamics lab will collaborate with multiple agencies, companies, and private landowners to select *Pinus taeda* sites which possess trees showing symptoms of brown spot needle blight, from around the State of Alabama. Eighteen 1/6 acre plots (Forest Health Monitoring protocols) will be established across the State to cover the six physioregions (Southern Coastal

Plain, Southeastern Plains, Piedmont, Ridge and Valley, Southwestern Appalachians, and Interior Plateau) and land ownership (government, industry, private). One central plot and three same sized sub-plots will be established at each site. The subplots will be located 120 m away from the central plot at bearings of 120, 240, and 360 degrees (Dunn, 1999). These plots will provide the needle samples (Components 1, 2, and 4), data (Component 2) and ground truthed check plots (Component 3). Site history will be assessed by obtaining historical records to describe land use history (agriculture or forestry), past management (prescribed fire, thinning, herbicide treatment, etc.), and other events (hurricane, wind, ice, fire, animal and any types of damage). Plot data collected will include basal area (pine and total), slope and aspect, and convexity. Five trees on the center plot will be selected to collect needle samples and vigor data. Dendrometer bands will be installed and trees tagged with an aluminum tag and nail. Needles will be harvested with hand tools or a .22 magnum caliber rifle with a scope (depending on tree height) and transported to Auburn University. Samples will be stored a 4 °C until processed. Disease verification will be completed as in Component 1.

Component 1: Inoculation protocol development for *Lecanosticta acicola* to develop a screening method to determine strain aggressiveness and seedling tolerance

Fungal Isolation and Identification:

Conidiomata formed on the needles will be aseptically excised, rolled onto 2% Dothistroma Sporulating Media (DSM: 5 g yeast extract, 20 g malt extract and 15 g agar per litre of distilled water) with 100 mg/L streptomycin in order to release conidia from the conidiomata as described by Barnes et al. (2004). The isolated conidiomata were incubated for one to two days at 23 °C. The plates will be examined using a dissection microscope and single germinating conidia will be selected and replated onto 2% DSM. The single conidial isolates will be grown for 4–6 wk. on a natural day light cycle, at 23 °C. Isolates of *Lecanosticta acicola* will be confirmed by amplifying extracted DNA with PCR and then sequencing using species-specific primers.

Propagation and Inoculation Development Study:

Temperature requirements for growth in culture will be studied on representative isolates selected for each of the novel species. Four by four-millimeter blocks of each culture will be plated, in triplicate, onto the centers of 2% MEA plates per temperature (10, 15, 20, 25, and 30 °C) and

incubated in darkness. The diameters of each colony will be recorded weekly along perpendicular axes for 4 wk. The best performing isolates will be selected for the inoculation study.

Conidial suspension with varying concentrations of spores will be prepared by harvesting conidia from 5- week-old cultures of the selected isolates and sprayed onto the needles of 1-year-old healthy seedlings. Non-inoculated seedlings of the same age will serve as a control. Inoculated and non-inoculated plants will be kept in a humid greenhouse. They will be inspected for leaf spot symptoms daily. The best performing spore suspension will be selected an inoculation study will be performed.

Koch Postulate & Inoculation Study:

Approximately 700 pine seedlings, comprising four *Pinus* species (longleaf, loblolly, shortleaf, and slash), will be obtained from the Southern Forest Nursery Management Cooperative at Auburn University. All seedlings will be potted in individual one-gallon pots filled with ProMix BX® peat-based potting mix. Seedlings will be potted soon after lifting in early December. Seedlings will be placed in a greenhouse, restricting natural precipitation while maintaining adequate light and natural outdoor temperature. Treatments will be imposed approximately three months after planting. Pots will be placed on raised benches and arranged randomly. Seedlings will be adequately watered equally when needed throughout the study. Treatments will be imposed approximately three months after planting. Treatments will be imposed on all seedlings in early March and will be randomly assigned to seedlings within each species. Prior to imposing treatments, dead or dying seedlings will be culled from each of the treatment groups, leaving approximately 50 seedlings within each treatment group per species.

Seedlings will be monitored daily until symptoms appear and then biweekly for the rest of the growing season. At the culmination of the study, seedlings will be recorded as alive or dead and rated for symptom severity.

Outputs:

One Ph.D. degree will be granted. Undergraduate student workers and undergraduate research fellows will be trained. Several manuscripts will be submitted for publication in refereed journals. A summary of research findings will be published on the project, forest health dynamics lab and

forest health cooperative websites. Technical transfer will also occur through presentations at conferences, landowner meetings and training sessions. Resistance screening protocol for seedling to *Lecanosticta acicola* will be developed and published.

Component 2: Environmental factors that drive the emergence and severity of infection from *Lecanosticta acicola* across Alabama

Tree Vigor Assessment:

Tree assessment and vigor measurements will be taken to determine the impact of the differing conditions on the pine trees. Diameter at breast height (DBH), tree height, tree age, growth increment (5 and 10 year) and crown ratings will be collected on each sample tree at plot installation. The current annual increment of BA will be calculated annually. Vegetation density for the understory will be calculated (Eckhardt, 2003). Resin will be sampled from all measurement trees from which foliage is collected. These measurements provide a measure of site conditions, stand density and influence of external stresses.

Crown evaluations quantitatively assess current tree conditions and provide an integrated measure of site conditions, stand density and influence of external stresses. Crown rating for severity will happen yearly as follows:

1. less than one-third of the crown infected (< 1/3),
2. one-third to two-thirds of the crown infected (1/3 to 2/3)
3. more than two-thirds of the crown infected (> 2/3)

LAI estimation will be determined seasonally four times per year to resolve potential relationships that arise during primary periods of (1) fascicle expansion and minimal resource stress (April-May), (2) when two complements of foliage are present, the potential for resource stress is maximized, and premature fascicle senescence is possible (July-August), (3) fascicle senescence of one complement of foliage occurs naturally (October-November), and (4) the dormant season when one complement of foliage is present (January – February) using UAS lidar or UAS multispectral imagery. A crown rating survey of all measurement trees in each plot will be conducted when the leaf area is measured in July-August. The crown rating survey will consist of live crown ratio, crown light exposure, crown position, crown density, crown dieback, and foliage transparency which are Forest Health Monitoring (FHM) crown/ damage indicators to describe

relative tree health (USDA 2001). Crown evaluations quantitatively assess current tree conditions and provide an integrated measure of site conditions, stand density and influence of external stresses. Trees with high scores for live crown ratio, density and diameter and low scores for dieback and foliage transparency have increased potential for carbon fixation, nutrient storage and survival and reproduction (USDA, 2001).

Foliar nutrition will be measured in July-August in the second year of the study. One upper crown branch per measurement tree will be shot out of tree crowns. Fully mature needles from the most recently mature flush of the previous year and the current year's first flush fascicles will be pooled by plot and treatment. Foliage will be processed by standard methods and analyzed for macro- and micronutrient levels. Total phenolics and total soluble sugars will be determined from foliage seasonally (four times per year) by wet chemistry, and Fourier Transform Near-Infrared (FT-NIR) spectra of the samples will be acquired. NIR spectroscopy will be coupled with chemometric algorithms for predicting the chemical composition of the samples based on the statistic models.

To analyze *L. aciaola* disease severity on shoot and needle length, *L. acicola* infected trees were assessed at the end of the growing season to ensure maximum retention of needle and shoot growth. Since whorl height has an effect on the shoot and needle lengths due to light availability, two whorl heights will be chosen. From the base of the tree, the first whorl height is at 2-5 m and the second at 5-8 m. From each tree, nine to eleven side apical shoots will be assessed at each whorl height. Shoot length and whorl length will be measured in cm. A total of 10 random fascicles from each shoot will be measured to obtain the average needle length (cm).

Soil Data:

Soils within the study area will be described and classified according to the following classification system(s): FHP method and USDA Soil Taxonomy. Soil samples will be collected to determine physical and chemical properties by removing a sufficient number of soil cores approximately 5 cm in diameter and 60 cm in length from each replication and sub-sectioned into 10 cm lengths for analysis. Dry soil weight data and associated moisture determinations will be used to calculate the bulk density (BD) and gravimetric soil moisture content (SMC_g) (Grossman and Reinsch, 2002). If warranted, particle size analysis will be conducted (Gee and Or, 2002). Soil chemical analyses

will be conducted either as individual samples or as a composited sample and include soil pH in a 1:1 soil:water suspension and 1M KCL solution, exchangeable cations (Calcium (Ca), Magnesium (Mg), Potassium (K), Sodium (Na) determined in 1M NH4 Acetate, pH7 extract, exchangeable aluminum (Al) and manganese (Mn) determined in 1M KCl extracts, exchangeable sulphur (S) by 1M NH4Cl extract, extractable phosphorus(P) by Bray 1, Total Carbon (TC) and Total Nitrogen (TN) by combustion, inorganic nitrogen by extraction, Organic Carbon (OC) by combustion according to FHM Methods of Analysis (2001). Effective cation exchange capacity (ECEC) will be computed from the sum of exchangeable Ca, Mg, K, Na, and Al, while effective exchangeable base content (EB) will be calculated as the sum of exchangeable Ca, Mg, K, and Na.

Climate Data:

Climate data available online in National Climatic Data Center will be collected to obtain regional historical weather data and station history. Daily summary observations of temperature, precipitation and relative humidity around infected stands will be collected. Since preceding year temperature, moisture and precipitation have been shown to be the best indicators of following years defoliation (Munck & Burns, 2012; Wyka et al., 2017), daily maximum and minimum temperature, the sum of seasonal precipitation in the year preceding scoring defoliation ratings of infected trees will be obtained from NOAA online data. Counties with available weather stations will be identified and distance measured from the infected stands using an interactive mapping tool. Weather data will be collected within a 10 miles radius of infected stands (if available). To collect relative humidity data, POWER Data Access Viewers will be used. Data points will be selected based on the longitude and latitude of the infected stand. Missing data will be adjusted, maximum and minimum temperature will be averaged, and the sum of precipitation will be estimated.

Data analysis:

The analysis of variance (ANOVA) and Duncan's test will be conducted to analyze the severity of infection between the six physiographic regions and three land ownerships using the *aov* function and *duncan.test* function in the *R: agricolae* package (Team, 2018). The generalized linear mixed-effect model will be used to analyze the influence of individual-tree characteristics, soil and climate data on the severity of infection. We will first use the *regsubsets* function in the

R: leaps package to select the best set of predictors. Then, we run the generalized linear mixed-effect model. The linear mixed-effect model will be analyzed by the R package *lme4*. The coefficient of variation (R^2) for fixed and random effects will be calculated using the function *r.squaredGLMM* of the R package MuMIn.

With the above selected best models, we will plot the relationships between response variables (i.e., the severity measures of infection) and significantly influential explanatory variables (i.e., soil and climate data, individual-tree level characteristics) using the code of Searle and Chen (Searle and Chen, 2020). According to the selected best model, these plots generate the fitting line and the 95% confidence interval lines of environmental factors.

Outputs:

Two Ph.D. degrees will be granted. Undergraduate student workers and undergraduate research fellows will be trained. Several manuscripts will be submitted for publication in refereed journals. A summary of research findings will be published on the project, forest health dynamics lab and forest health cooperative websites. Technical transfer will also occur through presentations at conferences, landowner meetings and training sessions. Distribution maps of *Lecanosticta acicola* in Alabama across the southeast will be produced. This work will also serve to better understanding the environmental factors which affect its emergence, distribution and severity.

Component 3: Detection and movement of *Lecanosticta acicola* with remote sensing

Data:

Multi-source remote sensing data (from multispectral and lidar sensors) will be leveraged to detect and construct a framework monitoring *Lecanosticta acicola* across landscapes. Changes in vegetation spectral response as a result of discoloration or chlorosis, allow for detection using multispectral imagery and imagery-derived products, particularly visible and infrared wavelengths and indices like the normalized vegetation index (Cotrozzi 2021) and red-edge metrics. In addition, since defoliation represents a structural symptom, there is potential for the utilization of light detection and ranging (lidar) for extracting canopy structural information to aid mapping efforts. The increasing availability of airborne lidar from the U.S. Geological

Survey (USGS) 3D Elevation Program (3DEP) (Lukas and Stoker 2016) and high temporal and spatial resolution satellite imagery from the ESA's Sentinel-2 program, present a unique opportunity to achieve detailed, finer-scale observations for analyses. Key benefits of using Sentinel-2 EO data are the availability of imagery with high temporal frequency (5-10 days), the number of spectral bands and bandwidths for computing a broad range of vegetation indices, and high spatial resolutions (10-20 m). Airborne lidar data from 3DEP and Sentinel-2 image will be used to extract vegetation structural and phenological information to support detection and distribution mapping. At finer scales (tree and plot-level), spectral, textural, and structural properties will be collected from unmanned aerial vehicle (UAV)-borne sensors to acquire high-resolution multispectral and hyperspectral imagery, as well as lidar point clouds.

Processing:

UAV-borne data (validated with ground-based data) will be used for tree-level infection detection, while satellite, airborne data will be analyzed for broad-scale detection and distribution mapping. UAV-derived imagery and UAV-based high-density lidar collected from selected sites with *Pinus taeda* and *Pinus palustris* around the State of Alabama, where pine needles show symptoms of brown spots or bands, will be examined. A local maxima algorithm will be adopted to map individual trees from UAV lidar-derived canopy height models and assess structural attributes (tree height, diameter at breast height, aboveground biomass). A suite of spectral indices indicative of vegetation health (e.g. red-edge chlorophyll index, normalized difference vegetation index, plant senescence reflectance index, pigment specific normalized difference) will be extracted from UAV imagery, and accuracy assessments on derived products will be carried out using field data. UAV-based outputs will be used to identify specific spectral and structural attributes of infected trees and serve as fine-scale, individual pine tree infection reference datasets for landscape level mapping with airborne and satellite data. Specifically, these fine-resolution datasets will be used to; (1) delineate trees species of interest, (2) calibrate airborne lidar-derived metrics, (3) determine criteria for broader-scale remote sensing detection, and (4) assess classification performance with airborne and spaceborne data. Recent acquisitions of airborne lidar data from 3DEP will be used to quantify forest structural indicators to provide landscape-level information and facilitate broad coverage mapping based on satellite-derived parameters. Airborne lidar point clouds from 3DEP, will be filtered to remove noise and returns

will be normalized using lidar-derived digital elevation models (DEMs). Gridded canopy products consisting of height percentiles, canopy density, and canopy cover, will be generated from aboveground-level data, at multiple spatial resolutions (10m, 20m, 30m). Vegetation structural parameters identified as significant to discriminating diseased trees will be extrapolated using spaceborne-derived parameters with machine learning methods. Vegetation structural products will be combined with a suite of spectral, textural and moisture indices extracted from Sentinel-2 bands (e.g. Normalized Difference Vegetation Index (NDVI), Enhanced Vegetation Index (EVI), and simple ratio and stacked and species information (two-stage mapping with UAV-derived imagery and satellite data) for image classification.

Data analysis:

The following classification and machine learning modeling techniques will be applied using combined spectral, vegetation structural information and ancillary data (topographic and climatic factors); Random Forests (RF), Stochastic Gradient Boosting (SGB), Artificial Neural Networks (ANNs), maximum likelihood, and support vector machine (SVM). Independent test sets for conducting accuracy assessments will be based on ground truth and UAV-borne information and the impact of each data source on detection accuracies will be assessed.

Outputs:

One Ph.D. degree will be granted. Undergraduate student workers and undergraduate research fellows will be trained. A summary of research findings will be published on the project, lab and forest health cooperative websites. Technical transfer will also occur through presentations at conferences, landowner meetings and training sessions, and we anticipate several publications in high-quality, peer-reviewed journals on several topics, including; (1) criteria for detection of *Lecanosticta acicola* infection using remote sensing, (2) spatially explicit distribution and movement maps of *Lecanosticta acicola* across the southeast where the pathogen is present, (3) UAV-borne lidar remote sensing of *Lecanosticta acicola*, (4) multi-scale approaches for assessing *Lecanosticta acicola* based on UAV, airborne and satellite data, and (5) machine learning methods for mapping *Lecanosticta acicola* This work will also serve to better understand environmental factors that facilitate its spread and contribute to development of a framework for monitoring *Lecanosticta acicola*.

Component 4: Genetic diversity of *Lecanosticta acicola*, pathogen origins, and invasion history

Tissue Collection and Isolation:

Pine needles, showing symptoms of brown spots or bands, will be collected from *Pinus taeda* and *Pinus palustris* native to North America. For each tree species, we will collect 250 samples in Louisiana, Mississippi, Alabama, Georgia, and Florida for a total of 2,500 needle samples. Conidiomata formed on the needles will be aseptically excised, rolled onto 2% Dothistroma Sporulating Media (DSM: 5 g yeast extract, 20 g malt extract and 15 g agar per litre of distilled water) with 100 mg/L streptomycin in order to release conidia from the conidiomata as described by Barnes et al. (2004). The isolated conidiomata will be incubated for one to two days at 23 °C. The plates will be examined using a dissection microscope and single germinating conidia selected and replated onto 2% DSM. The single conidial isolates will be grown for 4–6 wk. on a natural day light cycle, at 23 °C. Fungal tissue will be scraped from the surface of the cultures on 2% DSM with a sterile scalpel blade and lyophilized. The freeze-dried mycelium will be homogenized using a morter and pestal with liquid N and approximately 20 ng of the crushed mycelium will be used as starting material for DNA extractions. DNA will be extracted using a Zymo Research ZR Fungal/Bacterial DNA MiniPrep™ kit and eluted to a final volume of 50 µl. DNA will be stored at − 20 °C.

Tree host tissue will be collected from *P. taeda* and *P. palustris*, sampling one representative individual per species in each of six physioregions (n=12). To maintain high DNA integrity, each tissue sample will be frozen immediately. DNA will be extracted using a Qiagen DNeasy Plant Pro extraction kit and stored at -20° C.

DNA Sequencing:

Fungal samples

Following DNA isolation, we will use whole genome sequencing to efficiently characterize the genomes of our fungal samples. Specifically, we will uniquely barcode and sequence each sample to ~25x coverage using the MinIon whole molecule sequencing platform (flowcell R9.4.1). This platform supports on-the-fly assessment and screening of sequencing data using their propriety software MinKNOW. Following sequencing, we will export our data to a standard format (fastq)

and archive raw data for backup on servers with daily imaging. Using Auburn's supercomputer Easely, we will then remove low quality reads to maximize our data quality (Homburger et al. 2019) and map the high quality reads to the existing reference genome (NCBI accession GCA_002441625.1) using BWA (Li and Durbin 2009). We will also annotate the genome sequences using eggNOG mapper and existing fungal genome data and annotations (Huerta-Cepas et al. 2017). Genome variants (i.e. SNPs) will be identified and called using the program GATK (Van der Auwera and O'Connor 2020; Zanti et al. 2021).

Host samples

A reference genome assembly is already available for *Pinus taeda* (NCBI accession GCA_000404065.3) but not for *P. palustris*. We will create a reference assembly for *P. palustris* of comparable quality to *P. taeda* to support our inquiries into pathogen origins and invasion history. We will extract high-molecular weight DNA from one *P. palustris* sample using a modified nuclei extraction protocol (Zhang et al. 2012). We will sequence the extracted genomic DNA to a depth of 25X on a PacBio Sequel IIe to generate HiFi long reads (≥13 kb in length). The reference assembly will be generated using hifiasm (Cheng et al. 2021) and its quality assessed with BUSCO in "genome" mode (Manni et al. 2021). We will submit the final assembly to the NCBI Eukaryotic Genome Annotation Pipeline for repeat masking and gene model prediction and gene annotation (Thibaud-Nissen et al. 2013).

For the remaining eleven *Pinus* samples, we will perform whole genome resequencing on an Illumina NovaSeq platform, using five S4 lanes to generate 2x150 bp paired-end reads. This approach will produce ~15X coverage for each sample, which is sufficient to confidently call genotypes at each variable location in the genome. Low-quality bases and sequencing adapter sequences will be removed from reads using TrimGalore (Krueger et al. 2019). Reads passing quality control will be aligned to the appropriate reference genome (i.e., *P. taeda* or *P. palustris*) using the BWA-MEM algorithm implemented in BWA (Li and Durbin 2009). Genome variants will be called using GATK (Van der Auwera and O'Connor 2020; Zanti et al. 2021). All raw and cleaned genome sequences generated in this study will be made freely available by depositing fastq files and associated metadata into GenBank, a site hosted by the National Center for Biotechnology Information (NCBI; http://www.ncbi.nlm.nih.gov/genbank/).

Data Analysis:

We will first mine the fungal genome data by extracting sequences for each of the commonly used species identification loci: internal transcribed spacers (ITS) 1 and 2 (White et al. 1990), translation elongation factor 1-α gene (*TEF*1) (Carbone and Kohn 1999; O'Donnell et al. 1998), Beta-tubulin-2 gene region (O'Donnell and Cigelnik 1997 ; Stukenbrock et al. 2012; Glass and Donaldson 1995), Beta-tubulin-1 (Glass and Donaldson 1995), RNA polymerase II second largest subunit (*RPB*2) (Sung et al. 2007; Liu et al. 1999), and the guanine nucleotide-binding protein subunit beta (*MS204*) (Fourie et al. 2015). We will use these gene sequences to positively identify species in our fungal sequences prior to additional downstream analyses. Genomic location for each gene will be determined by searching the reference genome assembly and by blasting amplification primer sequences to the reference assembly (Camacho et al. 2008).

To understand the origins and evolutionary patterns of *L. acicola*, we will construct a phylogenomic trees using two sets of gene sequences. First, we will use a 264 protein dataset (Galindo et al. 2019) garnered from the genomes sequenced here as well as outgroup sequences from closely related species that are available via GenBank and the Joint Genome Institute (http://www.jgi.doe.gov) to understand the recent expansion and evolution of *L. acicola* in loblolly and longleaf pine. Second, we will use a phylogenomic dataset of 53 highly conserved protein sequences (Capella-Gutiérrez et al. 2012) across 2 representative individuals sequenced here and all closely related fungal species with existing data in GenBank and the Joint Genome Institute (currently 59 species but this is constantly increasing; Galindo et al. 2021). In each dataset, gene sequences will be aligned with MAFFT (Katoh and Standley 2013) and then concatenated. From these concatenated sequences, single gene trees will be created with FastTree (Price et al. 2007) and phylogenetic trees will be reconstructed using the Bayesian inference program PhyloBayes (Lartillot et al. 2009) as well as the maximum likelihood program IQ-Tree (Nguyen et al. 2015). From these outputs, alternative tree topologies will be compared using Mesquite (Mesquite Project Team 2014). From these two sets of trees, we will identify the relationship between disparate strains of *L. acicola* across the southeast, providing insight into the geographic patterns of spread and how *L. acicola* is related to other, similar fungal species.

We are also interested in the coevolutionary history shared between *L. acicola* and each of the host *Pinus* species because this may provide insight into the disparate pathologies in the two tree species. We will investigate the degree to which genetic elements have been horizontally transferred between *L. acicola* and each of the *Pinus* species. We will begin by assembling a pangenome for *L. acicola* using an iterative mapping and assembly approach following Golicz et al. (2016). Briefly, reads are mapped to the annotated reference genome, unmapped reads are assembled into inferred contiguous sequences, and these new contigs are added to the updated reference sequence. We will use the final reference pangenome as our set of query sequences to search for *L. acicola* genetic sequences in each of the 12 *Pinus* genomes. We will tabulate the number and total length of *L. acicola* genetic elements (i.e., transposable elements, protein-coding sequences) represented in each *Pinus* genome (Bergman and Quesneville 2007). Although identifying genetic element transfer events cannot directly quantify the length of time two species have co-occurred, we will be able to rank our *Pinus* samples according to degree of transfer to infer the relative duration of association between the host and pathogen. In addition, we will examine our fungal pangenome for additional transposable elements found in the *Pinus* species as an additional signal of the differences in coevolutionary history between *Pinus* and *L. acicola* (Caspi and Pachter 2005). This analysis will complement the phylogenetic approaches described above to describe the relationships among *L. acicola* in each of the six sampled subregions. In addition, this will provide insight into the relative length of time each *Pinus* species has been in contact with *L. acicola*, perhaps explaining the disparate outcomes for infection in the two species.

In addition to coevolutionary history, we are also interested in how *L. acicola* has evolved to permit increased infection in loblolly pine. We will evaluate this by examining functional diversity of our sequenced *L. acicola* genomes, using the pangenome constructed above. We will identify *L. acicola* genes that have either high sequence divergence between loblolly and longleaf tissue origins, high sequence variation within each of the sampled tree origins, or a combination of both parameters. Sequence divergence will be assessed by computing gene-by-gene divergence using the Kimura 2-parameter estimator (Beaumont et al. 1998). Sequence variation will be quantified by comparing the number SNPs found within each gene first within origin tree species and then between origin tree species. We will also assess our pangenome for structural variants, as these are also known to influence functional diversity (Hickey et al. 2020). If genome annotation

permits, we will assess these variants to determine if the identified SNPs have resulted in changes to expressed proteins structures (i.e. non-synonymous SNPs), providing a better mechanistic understanding of the functional genome changes in *L. acicola*. Similarly, we will assess how gene variants differ among the sampled habitats (as outlined and refined in other components) to identify any interactions between habitat and gene variants that may influence spread or virulence in particular areas.

Outputs:

One Ph.D. degree will be granted and one postdoc will be trained. Undergraduate student workers and undergraduate research fellows will be trained. A summary of research findings resulting from these analyses will be published on the lab and forest health cooperative websites. Technical transfer and descriptions of *L. acicola* genetic diversity, origins, and invasion history will also occur through presentations at conferences, landowner meetings, and training sessions. In addition, we anticipate the following topics will be the focus of publications in high quality, peer reviewed journals: 1) *L. acicola* genomic diversity and its distribution across the southeast; 2) *L. acicola* phylogeography and occurrence within the southeast landscape; 3) *L. acicola* relationship to other fungal pathogens; 4) *Pinus taeda* reference genome assembly; 5) transposable element distribution and relatedness in and between *L. acicola* and *Pinus* host species; 6) *L. acicola* pangenome construction and identification of functional variation associated with each host *Pinus* species and ecosystems; 7) *L. acicola* copy number variants and associations with host and ecosystems. Combined, these will summarize our insights into the genetic diversity of *L. acicola* in the southeast and its origins, including the distribution of *L. acicola* across the southeast.

Component 5: Extension and Outreach Activities

A strong extension and outreach component involving the Alabama Cooperative Extension System (ACES), Alabama Forestry Commission (AFC), and Auburn University Forest Health Cooperative (AUFHC) will parallel the research components throughout the duration of the study. The primary objectives of the extension component are (1) to "train the trainers" including ACES, AFC, and other educators on needle blight issues across the southeast; (2) to conduct landowner and land manager training through "hands on" field days and workshops; (3) to provide web based and social media (ie. Twitter) information updates through multiple

formats including fact sheets and videos on a dedicated project website; (4) create popular press articles for magazines including *Alabama Treasured Forests*; and (5) to hold a Needle Blight Conference at the end of project. These activities will be coordinated on parallel tracks with research activities from the beginning of the project.

Train the Trainer workshops:

Extension and outreach efforts will begin with workshops to engage and inform educators including regional and county ACES Extension agents and AFC personnel in order that they can take part in outreach efforts. Stakeholder input from the Auburn University Forest Health Cooperative indicates a critical need for training of educators on these interrelated issues. These workshops will include information regarding the objectives of the proposed project, needle blight identification, ecology, and control, and management strategies on these ecosystem services. Exchange of information regarding continuing and new stakeholder interests and needs will be encouraged. Discussion regarding potential locations for demonstration sites will be initiated. Additionally, we will initiate a Microsoft Sharepoint file sharing website that will provide an online calendar of all planned events, a discussion forum, and all relevant data and documents.

Field days, workshops, and seminars

Three field days each year will be developed and coordinated based upon locations of the study plots. We will attempt to utilize a cluster approach in establishing demonstration sites regionally to allow for field tours that participants can easily drive between sites within a cluster. This will allow for multiple site visits within a single tour. At each field day, we will utilize instruction and group discussion facilitated at each demonstration site. Educational materials will be provided to participants for learning reinforcement. Seminars will be conducted at various meetings throughout the state where travel distance may limit clientele from participating in the field days. These may include local extension meetings, NRCS sponsored meetings, and others.

Videos

Short, three-to-five-minute videos demonstrating multiple applied aspects of the project will

be created. All videos will be shot in the field at the research sites. These videos, produced using Flip UltraHD camcorders, will be posted on Auburn University's YouTube channel, Alabama Cooperative Extension System web pages and others.

Forest health outreach

We will provide timely updates on needle blight concerns on the Forest Health Cooperative web page and the Alabama Cooperative Extension System web pages. Extension services such as diagnostics for needle blight fungi, site visits and phone/email questions and answers will be available to landowners. Updates will be given at Forest Health Cooperative meetings and will be incorporated into the annual short course "Important Forest Health Issues: Disease, Insects and Invasives." Forest Health issues dealing with the management of needle blight will be incorporated into field days and workshops.

Timely Reports and web updates:
We will provide research and extension updates to the ACES group of trainers through Sharepoint and postings.

Fact sheets

We will create two new ACES fact sheets on needle blight. These facts sheets will be based on the findings from research components one through four and will be based on two years of data at each research site.

Alabama Forestry Commission "Treasured Forest" magazine article: We will coauthor a short article for Alabama Treasured Forest magazine, which is read by over 14,000 subscribers. This is one of the best outlets for widespread dissemination among Alabama forest landowners.