ABSTRACT

KOHLWAY IV, WILLIAM HENRY. Development of Genomic Tools for Insights into Phytophthora Root Rot Resistance in Trojan Fir. (Under the direction of Ross Whetten).

The oomycete, *Phytophthora cinnamomi* Rands, causes root rot disease on a broad range of fir and pine species used as Christmas trees, leading to more than $US 6 million in losses annually. One of the most valuable Christmas tree species, Fraser fir (*Abies fraseri* [Pursch] Poir.) has no documented innate immunity to Phytophthora. However an exotic fir species, Trojan fir (*Abies equi-trojani* Aschers. et Sint), has previously shown varying amounts of resistance to Phytophthora root rot (PRR). There are currently few genomic resources available for Christmas tree fir species so this study will create Trojan fir genomic resources for future research and help to identify the genetic basis of Phytophthora resistance in fir species.

Trojan fir seedlings germinated from a single open-pollenated family were inoculated with a single strain of *Phytophthora cinnamomi*. The root tip from each seedling was harvested for RNA extraction and the seedling was transplanted to sterile medium. Mortality of the transplanted seedlings was observed over 16 weeks. The extracted RNA samples were pooled and sequenced in groups based on experimental phenotype of mortality due to Phytophthora root rot along with non-inoculated control seedlings. A reference root transcriptome was assembled and used to identify genes differentially expressed between susceptible and resistant seedlings. In conjunction with the root RNA study, DNA markers were generated using Genotyping-by-Sequencing of 96 individual megagametophytic tissues collected from the seedlings as they germinated. Those markers were used to create a linkage map of Trojan fir and identify markers associated with PRR resistance.

Development of Genomic Tools for Insights into Phytophthora Root Rot Resistance in Trojan Fir

by
William Henry Kohlway IV

A thesis submitted to the Graduate Faculty of
North Carolina State University
in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

Functional Genomics

Raleigh, North Carolina

2022

APPROVED BY:

Ross Whetten Committee Chair	John Frampton
Dahlia Nielsen	David Reif

BIOGRAPHY

William Henry Kohlway IV was born in New Windsor, Maryland on June 10, 1992 to Marisol and William Kohlway III. After attending Nash-Rocky Mount Early College High School in 2007, he found his passion in biology and pursued a BS in Microbiology at North Carolina State University. During that time, he was introduced to Tree Improvement and Christmas Tree Genetics as an undergraduate research assistant and after working with loblolly pine and Fraser fir, he quickly became interested in forestry. After graduation, a desire to continue working with Christmas trees and the problems they face led to a Masters of Functional Genomics. Will currently lives in Raleigh, NC where he enjoys local beer and the company of good friends.

ACKNOWLEDGMENTS

A lot of commendable people have assisted me in the completion of my degree. First of all, I would like to thank Dr. Ross Whetten for being such a driving force in my academic and research career as well as providing much-needed mentorship these last few years. I would also like to recognize Dr. John Frampton; his instruction has guided my study of Christmas trees as well as my academic career. Dr. Nielsen and Dr. Reif are also acknowledged for their service on my committee and as teachers. I would like to thank AnneMargaret Braham for her indispensable aid with any and all work in the field and greenhouse. Jill Wegrzyn from UCONN was also especially crucial in collaborating on bioinformatics processes. Finally, thank you to my friends and family for keeping me grounded and sane most of the time. I am grateful to the agencies providing funding for this research. Primary support was from the Specialty Crop Block Grant Program of the North Carolina Department of Agriculture & Consumer Services. The USDA Specialty Crops Research Initiative Program and the North Carolina Agriculture Research Service via the Christmas Tree Genetics Program provided additional funding.

TABLE OF CONTENTS

LIST OF TABLES

CHAPTER 4

LIST OF FIGURES

CHAPTER 1

1.1 Christmas Tree Industry

1.1.1 Brief history of Christmas tree industry

The traditional harvest and indoor ornamental display of evergreens originated as a 16th century Germanic celebration of the winter solstice. The practice would later be adapted to the Christian holiday, Christmas, and the ornamental evergreens became known as Christmas trees (Chastagner and Benson 2000). By the 1800s, German immigrants had spread the Christmas tree tradition to America. Fir (*Abies* spp.), pine (*Pinus* spp.) and other coniferous trees were selected, felled, and carried home from local forests by celebrating families. By the 1850s, urban families desired Christmas trees of their own, but lacked local forests to gather from, so that in 1851 the first Christmas tree market in the USA was formed in New York (Albers and Davis 1997). Close to a million spruce and fir trees were collected from neighboring natural stands to stock the New York market and other markets that sprang up.

By 1901, more commercial, agricultural means of Christmas tree production developed in response to the increased demand for the ornamental trees. 25,000 Norway spruce (*Picea abies* (L.) H. Karst.) were planted in New Jersey by M. V. McGalliard with the intent to be sold as Christmas trees (Albers and Davis 1997). Many other farmers followed suit, utilizing land unsuited to other more popular crops. Subsequently, new techniques were developed to further enhance aesthetic traits; crown shape, branch density, and size. Additionally, due to the advancement of transportation, a more diverse selection of tree species became viable options for use as Christmas trees. (Chastagner and Benson 2000). The U.S. Christmas tree industry currently includes over 15,000 farmers in multiple parts of

the country who produce over 33.2 million trees annually, based on a 2013 survey (NCTA 2013).

1.1.2 North Carolina Christmas tree industry and Fraser fir

Fraser fir (*Abies fraseri* [Pursch] Poir.) is a conifer species found in geographically isolated pockets in the southern Appalachian Mountains at elevations above 1,300 m (Busing et al. 1993). Fraser fir was found especially suitable and lucrative for Christmas trees due to their dark green foliage, color, fragrance, and extended post-harvest needle retention (Chastagner and Benson 2000). In 1950, a portion of Roan Highlands in the Pisgah National forest was opened to public for Fraser fir harvest. However, as demand for the more desirable Fraser fir increased, the North Carolina Forest Service agreed to produce Fraser fir seedlings. To hasten production, these were originally gathered as wildings from natural stands, primarily near Mount Rogers, VA. Following the production of low-quality Fraser fir at two environmentally-unsuited nurseries, a seed orchard and seedling production site was established in a more suitable location in Avery County by the NC Forest Service in 1968 (Sidebottom 2011). Currently generating over $100 million in annual profits, the Christmas tree industry is a major part of North Carolina's economy, especially for the mountainous, western regions (Napier and Sidebottom 2011).

1.1.3 North Carolina Christmas tree industry and Phytophthora

Historically, the presence of the pathogen, *Phytophthora cinnamomi* Rands, in North Carolina predates the Christmas tree industry. A map published by Crandall et al. in 1945 included details of the observed range of root rot caused by *P. cinnamomi* on American

chestnut (*Castanea dentata* [Marsh.] Borkh.). Furthermore, it is hypothesized that *P. cinnamomi* was originally introduced into the United States as early as the 1800s on exotic plants for use in estate gardens (Crandall et al. 1945). Phytophthora root rot (PRR) of Fraser fir was first reported in 1963 on seedlings in a nursery in Penrose, NC, and Kuhlman and Hendrix (1963) identified *P. cinnamomi* as the causal organism. Root rot of fir is a disease that affects a tree of any age as *Phytophthora* attacks the root system, and because of the destruction of the roots and stem, the foliage of the fir yellows and eventually reddens as the tree perishes. Additionally, *Phytophthora* can remain in the soil surrounding infected trees and can continue to infect any subsequently planted susceptible hosts.

By 1995, the growing incidence of PRR at local nurseries led many growers to turn to out-of-state nurseries for their seedling planting stock. However, a 2016 survey discovered an increase of Phytophthora species diversity in Christmas tree nurseries in the Southern Appalachian region, with a notable rise in the incidence of *Phytophthora cryptogea* Pethybr. & Laff. which comprised 23.1% of the identified isolates (Pettersson et al., 2017).
The damage caused by *Phytophthora* has been estimated to be as much as $9 million (Chastagner and Benson 2000). Additionally, the current means of control against Phytophthora are costly; chemicals like metalaxyl (Benson and Grand 2000) and the grafting of susceptible species to root rot resistant rootstock (Hinesley and Frampton 2002). A cheaper alternative for Christmas tree growers with Phytophthora problems would be to employ root rot resistant trees.

1.2 Phytophthora

1.2.1 Phytophthora general information

The eukaryotic genus *Phytophthora* contains over 100 species of soil-borne oomycetes, or water-molds, which share hyphal and nutrient acquisition characteristics with true fungi. This similarity led to *Phytophthora* and other oomycetes originally being misclassified as fungi. Eventually with the rise of new molecular techniques, *Phytophthora* was differentiated from true fungi and characterized by a cellulose-based cell wall, two unique tubular flagella with mastigonemes on the zoospore, lysine synthesis pathways, and a diploid somatic thallus (Hardham 1994). Figure 1.1 shows the current location of the *Phytophthora* genus within the Eukaryotic phylogenetic tree, grouped with other heterokont algae (named for their distinctive two different flagella) (Hardham 2005). *Phytophthora* have versatile lifestyles and can live as saprophytes or as plant pathogens, and the genus includes infamous pathogens such as *Phytophthora infestans*, the causal agent of the Great Irish Famine of the 1840s. The most recent phylogeny of the *Phytophthora* genus was created using 11 loci, seven nuclear and four mitochondrial genes (cox2, nad9, rps10 and secY), and divided the genus into 10 clades (Figure 1.2) (Martin et al., 2014).

Phytophthora species have been spread world-wide through contaminated plant stock, but different regions have their own combinations of species adapted to their specific climates and present hosts. *Phytophthora* are difficult to control due in part to their ability to survive deep in the soil; their cellulose based cell wall, which prevents many of the standard fungicides from being effective, and their ability to form thickly walled chlamydospores (Hardham 2005).

1.2.2 Phytophthora cinnamomi

As previously described, *P. cinnamomi* is the major *Phytophthora* species culpable for Phytophthora root rot (PRR) of Fraser fir in North Carolina. *P. cinnamomi* Rands has caused numerous problems throughout the world both economically and ecologically, due to its large host range and the severity of the diseases it causes (Hardham 2005). The host range of *P. cinnamomi* contains approximately 950 species including firs, chestnuts, and many other herbaceous and woody plants (Zentmyer 1980). *P. cinnamomi* thrives under warm, rainy conditions where it spreads asexually using motile zoospores carried in water through wet soil and swimming towards susceptible plant roots. *P. cinnamomi* is heterothallic, but even in cases where both A1 and A2 mating types are present they have been observed to prefer to grow asexually (Hardham 2005). Of the two *P. cinnamomi* mating types, A2 is the most prevalent in North America (Benson and Grand 2000).

1.2.3 Phytophthora cinnamomi infection cycle

The infection cycle of *Phytophthora cinnamomi* can be partitioned into the following stages: sporulation, zoospore production and mobilization, encystment and penetration, and finally colonization (Figure 2). In the first stage, a combination of bacteria and abundant nutrients stimulates *P. cinnamomi* to asexually sporulate (Chee & Newhook, 1966), producing multinucleated sporangia at hyphal tips. During sporulation, the expression of over 200 genes encompassing ribosomal, metabolic, structural, and regulatory genes are modified (Hardham, 2005). There are 20-30 nuclei that are localized to the sporangial wall which are destined to become the motile zoospores.

Maturation and release of the zoospores from the sporangia (zoosporogenesis) is triggered by external stimuli, generally temperature change, which activates a signaling cascade to increase cytoplasmic calcium ion and pH. During zoosporogenesis, the sporangia are subdivided into uni-nucleate compartments while flagella are constructed from each nuclei's associated basal bodies. Other important organelles like mitochondria are then polarized to the sub-compartments. Finally, a large water expulsion vacuole forms in each future zoospore, and the sporangia begins to cleave open near the sporangial nuclear pore through the fusion microtubule-delivered electron dense vesicle. As the sporangia are cleaved, the plasma membrane is partitioned into the maturing zoospores, and the numerous kidney bean-shaped zoospores are 'squeezed' out of the cleaved sporangia in an osmotic wave as gels and other solutes build up.

The released zoospores then travel toward prospective host roots along chemical and electrical gradients. The two different flagella of the zoospores are used to swim short distances towards a host: the posterior flagella, the longer of the two, is used as a rudder, while the mastigoneme (flagellar hairs) containing anterior flagella coverts the sinusoidal waves of both flagella into forward locomotion (Hardham 2005). Flagellar control is mediated by calcium binding proteins, calmodulin and centrin, which regulate calcium ion levels. However, large distance travel is generally through external forces, like moving water from rainfall.

Once the zoospore is close enough to a host's root system, the zoospore turns so that the ventral side of the zoospore is towards the root as it begins to encyst. During encystment, the cell membrane of the zoospore undergoes rapid change (1-2 minutes) as peripheral cisternae fuses into it. Additionally, dorsal vesicles within the zoospore secrete a mucosal

compound over the surface of the encysting zoospore to prevent desiccation while the root facing ventral vesicles form a docking pad via the adhesive Vsv protein (Hardham 2005). After 30 minutes the adhered cyst germinates.

During germination, the germ tube grows out from the ventral side of the cyst and either directly penetrates the root epidermal cells along (or through) the perpendicular anticlinal/radial wall, or along the root surface for better penetration sites. Once the germ tube has penetrated the root, *P. cinnamomi* acts in the typical necrotrophic fashion; as pectin and other cell wall component-degrading enzymes like polygalacturonases are produced and secreted, the pathogen colonizes and devours the host roots. In a study by Götesson et al, 19 such *P. cinnamomi* polygalacturonase genes were sequenced and functionally validated (2002). The number and diversity of the identified polygalacturonases grant flexibility to interactions with various plant cell wall compositions and play a large role in the success of the pathogen across a broad host range. *P. cinnamomi* also secretes two elicitins, alpha- and beta-cinnamomin. These elicitins are sterol carrier proteins that may induce the host plant's defenses to activate a hypersensitive response which leads to necrosis of plant cells, providing more nutrients for the necrotrophic *P. cinnamomi* (Hardham, 2005). As the hyphae grow throughout the host roots, new sporangia can be produced as early as two days post-infection. As *Phytophthora* uses up the last of the host cell nutrients, either new sporangia are produced to move on to the next host or, if conditions are unfavorable, long-term survival chlamydospores are produced.

1.3 Host-pathogen interaction

1.3.1 Phytophthora root rot resistance screens

Family-based screenings for Phytophthora root rot (PRR) resistance have historically been used to identify and rank resistance within different conifer species. Screens for resistance have identified a broad range of susceptibility and resistance amongst popularly used Christmas tree and related species. There are several *Phytophthora* inoculation methods but the two most widely-used methods for inoculation are: *Phytophthora* colonized rice grains (or other seed) inserted directly into individually containerized seedlings, or by medium with equally distributed ground *Phytophthora*-colonized rice grain pellets (Holmes, et al. 1994). Both methods introduce harsher inoculum loads than what would naturally occur, ensuring that inoculated survivors are truly resistant. Another inoculation method is submerging roots in a zoospore suspension, though zoospore production is more difficult than the aforementioned methods.

Fraser fir, the most desirable Christmas tree produced in North Carolina, is extremely susceptible to Phytophthora root rot and consequently has been included as the susceptible control in several studies. A native Japanese species, momi fir (*A. firma* Sieb. et Zucc.) has consistently shown extremely high resistance to a number of *Phytophthora* species, but is not well suited for use as a Christmas tree due to its early bud break, suffering extensive frost damage most years. However, grafting susceptible fir species like Fraser fir onto momi fir rootstock grafts allows for the grafted seedling to be grown in Phytophthora infested sites (Hinesley and Frampton 2002). A large-scale resistance screen of seedlings from 32 different *Abies* species led to the identification of a few fir species that displayed some levels of root rot resistance (Frampton and Benson 2012). Two of the resistant species identified in the

2012 study were then further studied in 2013. Both native to Turkey, Turkish (*Abies bornmuelleriana* Mattf.) and Trojan fir (*Abies equi-trojani* Aschers. et Sint), have shown both family and geographically based variation in resistance to PRR with the eastern-most Turkish fir families most resistant to root rot (Frampton, Isik, and Benson 2013). Many of these studies have been carried out with only a single A2 isolate of *P. cinnamomi* known to be highly aggressive on Fraser fir, 23ss04. However, in a recent resistance screen of Turkish and Trojan fir families with 6 different isolates of *P. cinnamomi* and 2 isolates of P. *cryptogea*, the same PRR resistance rankings among the tested fir species were produced as in previous single-isolate screens (Kohlway et al. 2017).

1.3.2 Molecular response to Phytophthora in woody plants

While little molecular based research has been done in fir species, the interaction between *Phytophthora* and other woody plant species has been studied (Oßwald et al. 2014). The molecular response between hosts has been characterized in both susceptible and resistant hosts. Additionally, the intermediate ranges of resistance with different rates of disease progression within Turkish and Trojan fir families, points to complex quantitative rather than qualitative resistance mechanisms within firs (Kohlway et al. 2017).
In susceptible hosts, *Phytophthora* hyphae penetrate the root cortex and grow within the pericycle. The invasion of the root damages the phloem as *Phytophthora* effectors like pectinase and other cell wall degrading polygalacturonases are released by the growing hyphae. Two major elicitins produced by *P. cinnamomi* are the alpha- and beta-cinnamomin which, along with other elicitins and effectors, target host defense responses (Hardham 2005). *Phytophthora* can utilize these elicitins and effectors to evade the susceptible host

defenses. For instance, during infection with *P. cinnamomi*, the level of secondary defensive compounds, phenylalanine ammonium lyase (PAL) activity, lignin and phenolic synthesis of susceptible *Eucalyptus marginata* [Donn ex Sm.] were no different that the non-inoculated controls (Cahill and McComb, 1992). A family of cytolytic effectors, necrosis and ethylene-inducing peptide 1 (Nep1)-like proteins (NLPs), can contribute to the host cellular instability and pathogen virulence (Dong et al. 2012). The reduction in the ability for the host to absorb and transport water is not only due to the physical damage to the roots. In susceptible individuals, abscisic acid (ABA) and other cytokinin hormonal changes elicited by *Phytophthora* toxins amplify the decrease in host water utilization. The hormonal changes also cause the stomata to close reducing photosynthesis. The combination of these effects contribute to the death of the susceptible hosts.

The mechanisms for PRR resistance in woody plants rely on recognition of invading *Phytophthora* through pattern recognition receptors, the activation of pathogenesis-related (PR) genes, and the production of secondary metabolites against the invading pathogen. Effectors produced by *Phytophthora* can be detected by a number of coiled-coil (CC) or Toll/interleukin-1 (TIR) domain containing intracellular receptors like the TIR-NBS-LRR, QsRPc in cork oak (Oh and Hansen, 2007). These receptors induce signaling cascades (through MAPKs) that activate the transcription of PR genes and produce secondary metabolites (Oßwald et al. 2014). *Phytophthora cinnamomi* resistant Eucalyptus trees produced phenolic and lignin compounds through PAL activity, whereas the susceptible hosts were not able activate PAL (Cahill and McComb, 1992). The lignin content of resistant Eucalyptus roots was increased by 54% in the form of cell wall deposited-lignin. Lignin strengthens the cell wall of the roots, protecting the plant from degrading enzyme

Phytophthora effectors and inhibiting hyphal penetration. Additionally, invading *Phytophthora* can be sequestered through the rapid production and localization of callose layers around infected regions. Pathogenesis-related genes like cinnamoyl alcohol dehydrogenase1 (CAD1) have been proposed to detoxify *Phytophthora* aromatic aldehyde effectors (Coelho et al. 2006).

An additional example of resistance is the case of LecRK-1.9 in *Arabidopsis thaliana* (L.) Heynh. This R-gene provides resistance to a member of the largest *Phytophthora* effector family, RXLR-dEER (Bouwmeester et al., 2011). The specific RXLR effector studied was IPI-O, and specifically the cell-attachment RGD motif (Arg-Gly-Asp) portion of the effector can lead to the degradation of the connection between host cell walls and plasma membranes. Legume-like lectin receptor kinase 1.9 (LecRk-1.9) directly inhibits IPI-O and maintains root cell wall/membrane cohesion, providing enhanced resistance to *Phytophthora* (Bouwmeester et al., 2011). Non-functional LecRL-1.9 mutants were fully susceptible to *Phytophthora* infection, with symptomatic collapsed cortexes.

In Port Orford cedar (*Chamaecyparis lawsoniana* [A. Murray] Parl.), susceptible individuals had *Phytophthora* hyphae spread through the vascular system as well as hyphae within the root cortex 48 hours post root inoculation via zoospore suspension (Oh and Hansen, 2007). However, when the resistant roots were inoculated, fewer cysts were formed on the root surface, and many that had encysted, did not germinate. However, the cysts that did germinate carried out the invasion of the root cortex in the same fashion as in the susceptible roots. Additionally, after 24 hours the hyphae within the resistant root had not penetrated very far within the cortex compared with the susceptible counterpart. After a total of 72 hours, the hyphae within the resistant roots had still not penetrated the vascular system.

By comparing the two types of roots, a number of host cell-wall differences were present in the resistant roots. For instance, the cortical walls were thicker in the resistant roots and electron-dense compounds produced in the cortical cells inhibited encysted *Phytophthora* penetration of the roots (Oh and Hansen, 2007).

1.4 Research Goals

The goal for this research project is twofold; first to develop genomic tools and molecular techniques for use in these and future studies, and second, to use these tools to further the understanding of the host-pathogen interaction between *Phytophthora* and fir trees. To simplify and consolidate phenotypic and genetic data, a single open pollinated family of Trojan fir, originating from the Kazdağı provenance in Turkey was used for each study.

The first study, described in Chapter 2, is a novel assembly of a Trojan fir genetic linkage map using DNA markers developed via genotyping-by-sequencing (GBS) of megagametophytic tissue. Additionally, Trojan fir DNA markers found in association with PRR were identified. The megagametophyte, a haploid maternally derived tissue, provides a valuable tool for genomic mapping in conifer genomes. The haploid state of the megagametophyte DNA ameliorates historic issues in genetic mapping in conifers, that is, numerous multi-copy genes. Formation of a genetic linkage map will have far reaching applications in not only this study but in future endeavors as well.

In parallel with the DNA based work described above, chapters 3 and 4 utilize RNA from *Phytophthora cinnamomi* challenged seedling root tips collected from a PRR resistance screen to identify differences in molecular response between susceptible and resistant hosts.

Chapter 3 describes a novel RNA sample preparation using highly-necrotic, *Phytophthora*-infested Trojan fir seedling roots. This new technique enabled the recovery of usable RNA samples from otherwise unusable, damaged tissue as well as a metagenomic view of the microbial community associated with resistant and susceptible fir roots. Validation of these tools will allow for the collection of a more robust dataset in future projects not only under experimental conditions, but from diseased individuals in actual Christmas tree farms.

Finally, chapter 4 describes a more traditional RNA sequencing approach to identify early vs late gene responses of Trojan fir roots to *Phytophthora*. Two time points, 48 hours and 96 hours post-challenge, were used to identify transcriptional responses at different stages of PRR disease progression between susceptible and resistant roots. Additionally, the combined RNA datasets from both chapters 3 and 4 were used to assemble a transcriptome for Trojan fir roots. The information gained from this study will help expand the current understanding of the host-pathogen interaction between firs and *Phytophthora*.

This project pioneers the study of genetic responses to *Phytophthora* in Trojan fir specifically and firs in general. The creation of a Trojan fir linkage map and transcriptome provides an invaluable bioinformatic foundation for future research. Additionally, the identification of genetic responses that provide survival to *Phytophthora* in these unnaturally harsh experimental conditions and genetic markers in association with those genotypes will help guide future breeding efforts to provide *Phytophthora* root rot resistant Christmas tree planting stock to affected farmers.

1.5 Figures

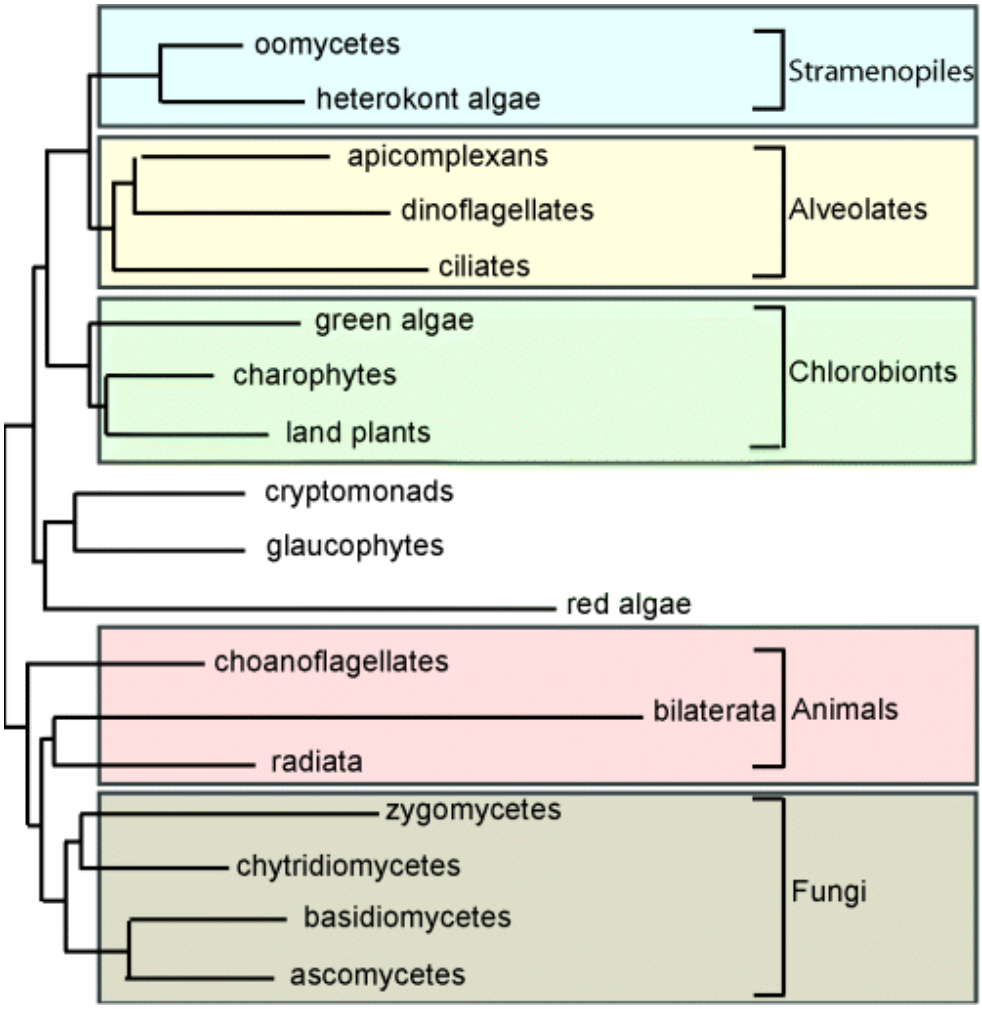


Figure 1.1: Evolutionary Phylogeny of Oomycetes

Phylogenetic tree of Eukaryotes showing oomycete similarity to heterokont algae rather than fungi (Hardham 2005).

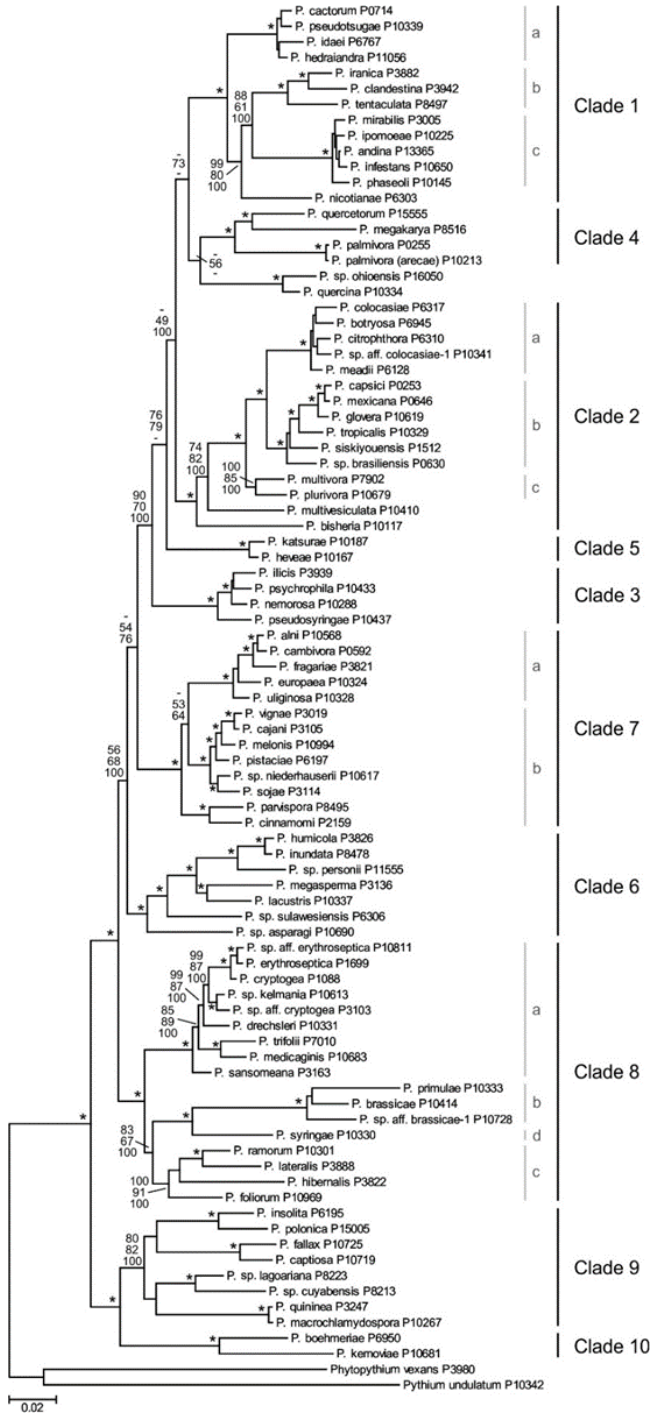


Figure 1.2: Phylogeny of the *Phytophthora* genus

The most recent phylogeny created using 11 loci, seven nuclear and four mitochondrial genes:cox2, nad9, rps10 and secY (Martin et al., 2014).

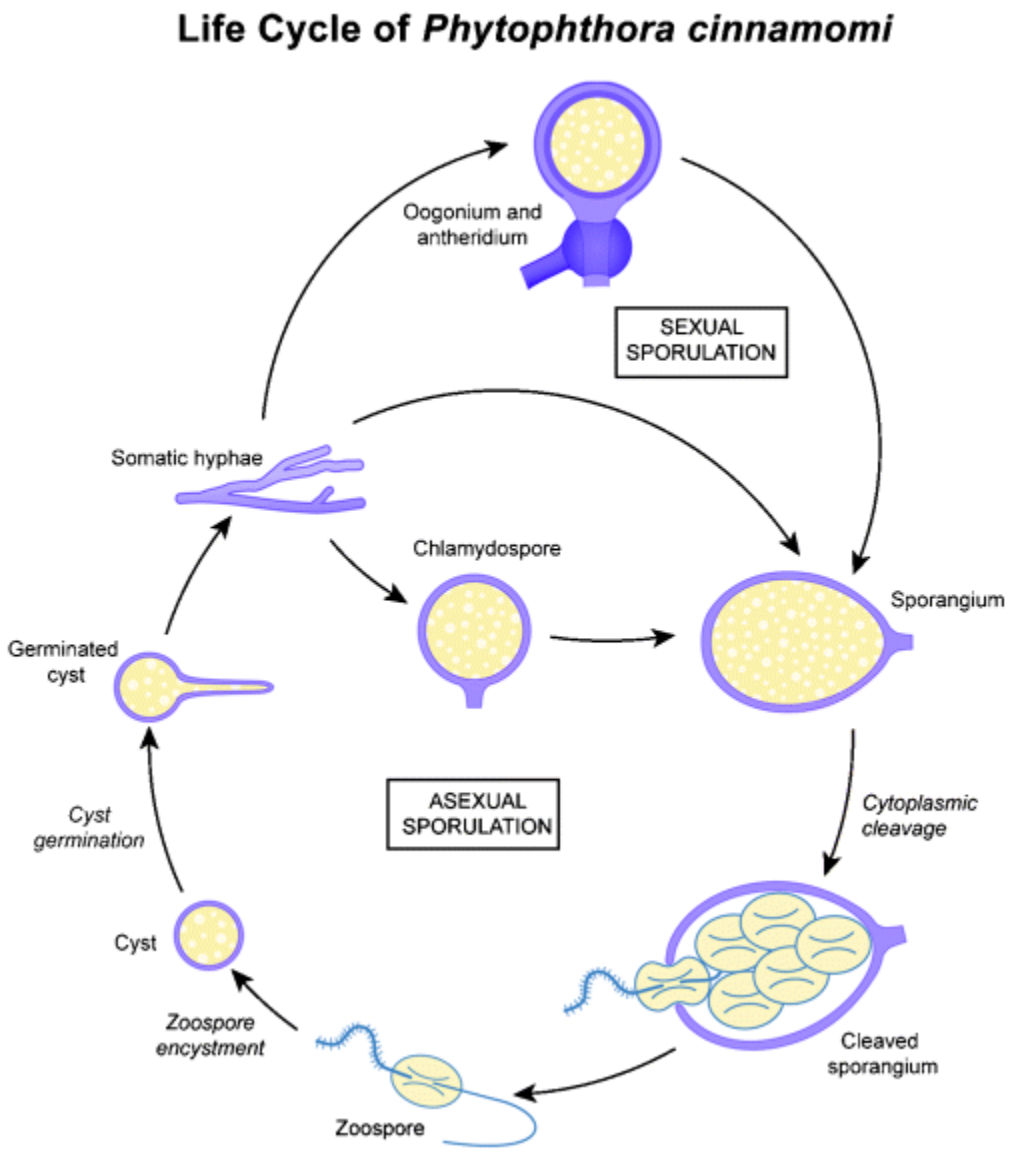


Figure 1.3 Life cycle of *Phytophthora cinnamomi* Rands

Sexual and Asexual life cycle of *Phytophthora cinnamomi* Rands. Spread of *P. cinnamomi* is believed to be primarily asexual (Hardham 2005).

1.6 Literature Cited

Albers HH, Davis AK (1997). The wonderful world of Christmas trees. Mid-Prairie Books, Parkersburg, 100p

Benson DM, Grand LF (2000). Incidence of Phytophthora Root Rot of Fraser Fir in North Carolina and Sensitivity of Isolates of *Phytophthora cinnamomi* to Metalaxyl. *Plant Disease*, *84*(6), 661–664. https://doi.org/10.1094/PDIS.2000.84.6.661

Bouwmeester K, de Sain, M., Weide, R., Gouget, A., Klamer, S., Canut, H., & Govers, F. (2011). The Lectin Receptor Kinase LecRK-I.9 Is a Novel Phytophthora Resistance Component and a Potential Host Target for a RXLR Effector. *PLoS Pathogens*, *7*(3), e1001327. https://doi.org/10.1371/journal.ppat.1001327

Cahill, D. M., & McComb, J. A. (1992). A comparison of changes in phenylalanine ammonia- lyase activity, lignin and phenolic synthesis in the roots of *Eucalyptus calophylla* (field resistant) and *E. marginata* (susceptible) when infected with *Phytophthora cinnamomi*, 18. http://dx.doi.org/10.1016/0885-5765(92)90014-m

Chastagner GA, Benson DM (2000) The Christmas tree: traditions, production, and diseases. *Online Plant Health Progress*. doi:10.1094/PHP-2000-1013-01-RV

Chee, K.-H., & Newhook, F. J. (1966). Relationship of Micro-Organisms to sporulation of *Phytophthora cinnamomi* Rands. *New Zealand Journal of Agricultural Research*, *9*(1), 32–43. https://doi.org/10.1080/00288233.1966.10418115

Coelho, A. C., Horta, M., Neves, D., & Cravador, A. (2006). Involvement of a cinnamyl alcohol dehydrogenase of Quercus suber in the defence response to infection by *Phytophthora cinnamomi*. *Physiological and Molecular Plant Pathology*, *69*(1–3), 62–72. https://doi.org/10.1016/j.pmpp.2007.01.001

Crandall, B. S., G.F. Gravatt, & M.M. Ryan. (1945). Root disease of *Castanea* species and some coniferous and broadleaf nursery stocks, caused by *Phytophthora cinnamomi*. *Phytopathology*, *35*.

Dong, S., Kong, G., Qutob, D., Yu, X., Tang, J., Kang, J., … Wang, Y. (2012). The NLP Toxin Family in *Phytophthora sojae* Includes Rapidly Evolving Groups That Lack Necrosis-Inducing Activity. *Molecular Plant-Microbe Interactions*, 25(7), 896–909. https://doi.org/10.1094/MPMI-01-12-0023-R

Frampton, J., & Benson, D. M. (2012). Seedling resistance to *Phytophthora cinnamomi* in the genus *Abies*. *Annals of Forest Science*, *69*(7), 805–812. https://doi.org/10.1007/s13595-012-0205-4

Frampton J, Fikret I, Benson DM (2013) Genetic variation in resistance to *Phytophthora cinnamomi* in seedlings of two Turkish *Abies* species. *Tree Genetics & Genomes*. 9:53-63

Götesson, A., Marshall, J. S., Jones, D. A., & Hardham, A. R. (2002). Characterization and Evolutionary Analysis of a Large Polygalacturonase Gene Family in the Oomycete Plant

Pathogen *Phytophthora cinnamomi*. Molecular Plant-Microbe Interactions, 15(9), 907–921. https://doi.org/10.1094/MPMI.2002.15.9.907

Hardham, A. R. (2005). *Phytophthora cinnamomi*. *Molecular Plant Pathology*, *6*(6), 589–604. https://doi.org/10.1111/j.1364-3703.2005.00308.x

Hinesley, E., & Frampton, J. (2002). Grafting Fraser Fir onto Rootstocks of Selected *Abies* Species. *HortScience*, *37*, 4.

Holmes, K. A. (1994). Evaluation of *Phytophthora parasitica* var. *nicotianae* for biocontrol of *Phytophthora parasitica* on *Catharanthus roseus*. *Plant Disease*, *v. 78*(2), 193–199. https://doi.org/10.1094/PD-78-0193

Kohlway, W. H., Whetten, R. W., Benson, D. M., & Frampton, J. (2017). Response of Turkish and Trojan fir to *Phytophthora cinnamomi* and *P. cryptogea*. *Scandinavian Journal of Forest Research*, *32*(5), 406–411. https://doi.org/10.1080/02827581.2017.1280076

Kuhlman, E. G., & Hendrix, F. F. (1963). Phytophthora root rot of Fraser Fir. *Plant Disease Reporter*, *47*(6), 552–553.

Napier, A.S.; Sidebottom, J.R. (2011). North Carolina Christmas trees by the numbers. https://christmastrees.ces.ncsu.edu/christmastrees-nc-christmas-trees-by-the-numbers/. (April 2016

Martin, F. N., Blair, J. E., & Coffey, M. D. (2014). A combined mitochondrial and nuclear multilocus phylogeny of the genus *Phytophthora*. *Fungal Genetics and Biology*, *66*, 19–32. https://doi.org/10.1016/j.fgb.2014.02.006

McKeever, K. M., & Chastagner, G. A. (2016). A Survey of *Phytophthora* spp. Associated with *Abies* in U.S. Christmas Tree Farms. *Plant Disease*, *100*(6), 1161–1169. https://doi.org/10.1094/PDIS-08-15-0939-RE

Oßwald, W., Fleischmann, F., Rigling, D., Coelho, A. C., Cravador, A., Diez, J., … Werres, S. (2014). Strategies of attack and defence in woody plant- *Phytophthora* interactions. *Forest Pathology*, 44(3), 169–190. https://doi.org/10.1111/efp.12096

Oh, E., & Hansen, E. M. (2007). Histopathology of Infection and Colonization of Susceptible and Resistant Port-Orford-Cedar by *Phytophthora lateralis*. *Phytopathology*, *97*(6), 684–693. https://doi.org/10.1094/PHYTO-97-6-0684

Pettersson, M., Frampton, J., Rönnberg, J., Shew, H. D., Benson, D. M., Kohlway, W. H., … Cubeta, M. A. (2017). Increased diversity of *Phytophthora* species in Fraser fir Christmas tree plantations in the Southern Appalachians. *Scandinavian Journal of Forest Research*, *32*(5), 412–420. https://doi.org/10.1080/02827581.2016.1265144

Pettersson, M., Frampton, J., & Sidebottom, J. (2017). Influence of Phytophthora Root Rot on Planting Trends of Fraser Fir Christmas Trees in the Southern Appalachian Mountains, *Tree Planters Notes 60*(1), 8.

Quesada-Ocampo, L. M., Fulbright, D. W., & Hausbeck, M. K. (2009). Susceptibility of Fraser Fir to *Phytophthora capsici*. *Plant Disease*, *93*(2), 135–141. https://doi.org/10.1094/PDIS-93-2-0135

Sidebottom, J.R. (2011). History of the North Carolina Christmas tree industry. https://christmastrees.ces.ncsu.edu/christmastrees-history-of-the-north-carolina-christmas-tree-industry/ (December 2017).

Zentmyer, G. A. (1988). Origin and distribution of four species of Phytophthora. *Transactions of the British Mycological Society*, *91*(3), 367–378. https://doi.org/10.1016/S0007-1536(88)80111-6

CHAPTER 2

2.1 Introduction

Trojan fir (*Abies equi-trojani*), like other conifers, is far from a model organism. Many conifers have massive, complex genomes, such as the 16.8 Gbp (2n=24 chromosomes) Fraser fir genome (Auckland et al. 2001) for instance, with a high fraction of their genome made up of transcriptionally repressed non-genic transposable elements. These non-genic, highly methylated regions of the genome are likely to be distinct from regions of the genome that contain active genes, including those associated with disease resistance, and markers in these non-genic regions are unlikely to be useful for genetic analysis of resistance. Linkage disequilibrium (LD) is also low in conifers and decays rapidly (Brown et al, 2004). This makes genetic marker comparisons among different conifer populations difficult.

Wheat and barley have large and complex genomes with high portions of repeated genetic elements. Unlike most fir species, both wheat and barley have decent genetic maps, with described microsatellite and restriction fragment length polymorphism (RFLP) markers. However, high-throughput genotyping using RFLP and microsatellite markers is a labor-intensive endeavor.

The difficulty of genomic marker based studies in conifers in open-pollinated (OP) natural stands has been explored in white spruce (*Picea glauca* (Moench) Voss) (Beaulieu et al. 2014). The main problem when dealing with OP families is that there are numerous paternal sources of DNA from pollen, and little if any information is typically available about the pool of possible pollen parents. The large number of random paternal haplotypes detected among the progeny confounds the population structure of the family. Most offspring in an OP family are half-sibs, meaning on average the offspring share 25% of the genome in

common, but some pairs of offspring may not share any haplotypes in common with each other. Depending on genetic diversity in the parental generation and the number of segregating alleles, there is likely to be a large range of genotypic and phenotypic combinations. With such a scrambled mixture of progeny genotypes, and without knowledge of the maternal genotype, it is hard to discern genetic markers that are significantly associated with a trait such as disease resistance. Beaulieu et al. (2014) were able to overcome these difficulties with a genotyping chip that allowed for the rapid genotyping of 7338 single-nucleotide polymorphism (SNP) markers, which covered the entire genome of white spruce and included markers linked to a number of quantitative trait loci (QTL). Using markers mapped to a reference genome sequence, they were able to parse out the individual relatedness for use in the prediction of breeding values for genomic selection. However, in Trojan fir this method is not possible because there is no reference genome or genetic map.

To minimize costs and labor, an alternative high-throughput genotyping technique, two-enzyme Genotyping-by-Sequencing (GBS), was developed by Poland et al (2012). The two-enzyme GBS protocol employs methylation-sensitive restriction enzymes, effectively enriching fragments from regions of the genome that are transcriptionally active. The reduced pool of selectable fragments forms a reproducible set of markers that can be sequenced with less coverage required per individual, reducing the cost of genotyping per individual and increasing the throughput. A previous GBS-based study in an OP Trojan fir family was carried (Kohlway, 2015) with goal of creating genetic markers associated with Phytophthora root rot resistance (PRR). 205 out of 117,345 segregating GBS markers were identified as significantly associated with PRR using a contingency test. However, due to the

amount of genetic noise from the paternally derived alleles of the sample, it was not possible to create a linkage map.

To describe linkage maps a few important terms must be defined, the first of which is genetic linkage (linkage) itself. Linkage is a measure of genetic proximity between two loci within a genome. Genetic distance is defined by the rate of recombination between the two loci in the genome and is measured in centiMorgans (cM), where 1cM is equivalent to 1% recombination between the two loci. Because there are multiple separate chromosomes that make up the genome of an individual, which are individually assorted to form gametes, the maximum observed recombination frequency between two loci should be 50%, or 50cM, representing loci on two separate chromosomes. However, genetic distances can be calculated as larger than 50cM when using estimations generated from pedigreed data.

DNA sequence variants such as single-nucleotide polymorphisms (SNPS) or microsatellites can be utilized as markers to identify specific loci on chromosomes. By comparing the recombination rate between sets of genetic markers, their location and order can be assembled into linkage groups which are, in the ideal case, equivalent to separate chromosomes.

Despite the many disadvantages conifer species pose for genetic research, they do possess a unique tissue, the megagametophyte, which has many beneficial traits. Megagametophytes are maternally derived haploid tissue that produce the female gametes during reproduction and persist in the mature seed as the nutritive tissue to feed the embryo during seed germination. Because the megagametophyte contains a single recombinant maternal haplotype the DNA markers generated can be verified as single-copy versus multi-copy, an important distinction for effective DNA markers in conifer breeding.

Genetic tools greatly benefit breeders in several ways. The use of DNA and RNA based genetic markers to track pedigrees, and important traits within them, has long been a major tool of the plant and animal breeder. These tools are especially important when the species of interest has a long generation time, like firs and other conifer species. In linkage analyses, marker inheritance patterns from pedigrees are used to test for coinheritance of chromosomal regions with a trait of interest. There are two different methods that can be used to do this: parametric and non-parametric.

The parametric, model-based linkage analyses were the first to be developed. Parametric models test if a pedigree's inheritance model fits a specified genetic model for a trait. The major principle behind these studies is logarithm of the odds (LOD) score of a marker which is calculated by the likelihood of a specific chromosomal region (designated by the marker) being linked to the trait-controlling gene(s) as opposed to the loci being un-linked. The likelihood ratio is calculated by specifying a genetic model through the penetrance (the proportion of individuals with a specific allele/genotype that display the associated phenotype) and allele frequency at the trait locus,

One linkage analysis approach uses positive result in a parametric linkage analysis test, generally LOD score of above 3, which represents 1000 to 1 odds that the marker in question is in close genetic proximity to the trait locus: and is thus linked and inherited together. There are few different methods generally used for computing the LOD scores: the Elston-Stewart algorithm and the Lander-Green algorithm. The Elston-Stewart algorithm is best suited for large, extensive pedigrees with few markers genotyped per individual. Conversely, the Lander-Green algorithm is apt for computing LODs for data sets with small pedigrees but genotyped with numerous markers.

Parametric linkage studies have been used to identify many simple Mendialian disease loci. However, there are some limitations in using the traditional parametric linkage analysis. One limitation is that the genetic model for the trait/disease has to be specified, with known (or estimated) allele frequencies and allele penetrance parameters. This is not always known, especially for complex traits and diseases, and an incorrect estimate of these parameters can greatly affect the linkage analyses.

In nonparametric linkage analyses, the inheritance pattern of a loci in a pedigree is tested for a deviation from independent sorting. One nonparametric analysis is the affected-sib model which postulates that affected sibling pairs will share more loci that are linked to the disease locus Identical-by-Descent (IBD) than the norm. Two alleles of a locus between siblings are IBD if they have inherited the same allele without recombination from a common ancestor, such as a parent. Under random assortment, the expected probability that two full-siblings (siblings that have the same father and mother) share a locus IBD is: ¼ for 0 alleles IBD, ½ for 1 allele IBD, and ¼ for 2 alleles IBD. Pearson chi-square goodness of fit tests can be used to compare the observed number of sibling-paired IBD alleles to the expected number of alleles IBD under the null hypothesis (no linkage). These tests can be extended to more distant family relationships as long as the proper expected IBD sharing probabilities are specified for the null hypothesis.

For both parametric and nonparametric methods the physical size of the linked region is dependent on the density of the markers across the genome and the rate of recombination observed between them in the pedigree. This means that, especially in small human pedigrees with relatively few recombination events, the trait/disease locus can be 10s-100s of Mb in physical size with no easy way of identifying the molecular basis for the trait.

The main goal of this project is to expand upon the genomic resources available for Trojan fir via the production of a linkage map. The hope is that this will help with screening for genetic markers associated with Phytophthora root rot resistance in future studies. Development of such markers could allow numerous Trojan fir trees to be screened for Phytophthora resistance with a cheap, fast, and repeatable method, such as multiplex-PCR. Additionally, the genetic map will help to direct future full genome sequencing projects.

2.2 Materials and Methods

2.2.1 Trojan fir megagametophyte production

250 Trojan fir seed collected from natural stands in the Kazdağı provenance of Turkey for a previous study (Frampton et al. 2013) were surface sterilized with 10% bleach and stratified at 4^{o}C for one month. The cold-stratified seed were then planted in a vermiculite-filled plastic bin at a depth of 2.5 cm. The sowed vermiculite bed was then placed under a fluorescent light and kept damp for 3 weeks to promote germination and growth. Megagametophytic tissue was collected as soon as the Trojan fir seedlings sprouted above the surface by sliding the seed including the megagametophyte off the cotyledons. The megagametophyte was then separated from the seed along with the remnants of the seed coat. The completely isolated megagametophytic tissue was then transferred to a Qiagen 96-deepwell plate preloaded with a small (3 mm) metallic bead. An additional bead was then placed on top of the megagametophyte and the sample was capped and stored at -80^{o}C. The seedlings germinated in this study were used for the seedling Phytophthora root rot (PRR) screen in chapter 4 where the megagametophyte-identified seedlings were phenotyped for PRR resistance.

2.2.2 GBS library preparation and sequencing

DNA was extracted from the megagametophyte following the method of Matallana et al. (2019). The frozen megagametophytic tissue was ground to a powder in a Qiagen mixer mill using two cycles of 30hz shaking for 1 minute each, with rotation of the plate on the mill between cycles, to ensure equal and total grinding of the tissue. DNA was purified from the powdered samples using a protocol modified from an insect DNA extraction protocol (Ivanova et al. 2006). Extracted DNA was quantified using a NanoDrop microspectrophotometer. The sequencing libraries were prepared following a modified two-enzyme genotyping-by-sequencing procedure described by Poland et al. (2012). The GBS libraries were prepared using 700 ng of extracted genomic DNA from each sample. The restriction endonucleases, *Pst1*-HF and *Msp1*, were used to fragment the genomic DNA and create ends to which 96 different barcoded adapters were ligated to identify samples. The 96 libraries were then pooled and size-selected for 300bp +/- 36bp by Pippen Prep™ (Sage Science). The libraries were then redistributed into 8 separate PCR reactions. Each PCR ran for 16 cycles with one of two different index primers, (NEBNext™ index 6 and 12) to ensure nucleotide balance during the index read step. Each amplified library was then combined in equimolar quantities and sent for sequencing on two lanes of 100bp single-end reads on an Illumina Hi-Seq 2500.

2.2.3 Sequence analysis

The sequences were analyzed using several software programs. First, Flexbar (Dodt, et al. 2012) was used to separate and trim reads from each individual based on the adapter barcodes. The Tassel-4.0 standalone Genotyping by Sequencing pipeline (Glaubitz et al.

2014) converted the raw sequences into sequence-tagged loci, or tags, using the MergedTagsByTaxa plugin, with a read-count threshold of at least five reads per locus required to declare a genotype. The clusterfast option of Vsearch (Rognes et al. 2016) was used to cluster reads to putative bi-allelic markers by screening with 63/64 basepair matching or 98% identity. Any two-member cluster that had both SNPs present within the same haploid individual was removed as paralogous sequence. The resulting tags were imported into R to be filtered for segregating tags, or tags that have a presence/absence profile of 25-75%. Another GBS analysis tool, STACKS (version 2.3), was used in parallel to the Tassel pipeline to take advantage of the full length of the sequencing reads. The STACKS de novo script was run with a read-count threshold of at least five reads per locus in at least 5 samples. The Hierfstat R package (Goudet 2005) was used to generate a kinship matrix from both the Kaz24 trojan fir megagametophyte GBS sequence and single RNA genotyped Fraser fir megagametophyte family (FF51) with a fully verified chain of custody (unpublished) following the genome-agnostic protocol of Voichek and Weigel (2020). Principal components analysis (PCA) of the raw sequencing reads was also carried out using R.

2.2.4 Linkage map assembly and QTL mapping of Phytophthora root rot resistance

The assembled markers from 63 putative progeny of the Kaz24 seed parent were combined into a single reference file as a mock Kaz24 parent genotype. The populations command with the default settings from Stacks was used to output a doubled-haploid mapping cross using the mock Kaz24 tree and the megagametophyte samples as its progeny.

The potential bi-allelic markers identified as candidate genetic markers in the Kaz24 63-sample mapping population were used to generate a genetic map with the MSTmap

function in ASmap with the following settings: kosambi method, Maximum Likelihood, p value= 10^{-8}. The assembled genetic map was screened for sites linked to Phytophthora root rot resistance with the binary model of rQTL R package (Broman et al. 2003) with kosambi map function and error probability set to 0.001. Significant LOD thresholds for linkage was determined using 10,000 permutations.

Independent association of segregating Kaz24 megagametophyte tags with the *Phytophthora* resistant phenotype was determined using the coin R package (Hothorn et al. 2008). The coin R package runs a contingency test that compares the presence/absence pattern of each marker across all the individuals compared to the survival/death phenotype. The significance level was determined by permutation, with 1000 permutations of randomized genotype-phenotype pairings, to preserve the correlation structure among tag genotypes but break any correlation between genotype and phenotype. The significantly associated tags were then searched for sequence similarity with described genes via BLAST server on the NCBI website.

2.3 Results

2.3.1 Results of sequencing and Genotyping-by-sequencing

The two lanes of 100bp Hi-seq resulted in 331 million raw sequencing reads, of which 202 million reads passed fastQC filtering with quality scores above 30. Tassel 4.0 assembled those reads into 173,250 marker tags. Clustering with vsearch for two member groups of tags that differ by a single base reduced the total down to 10,875.

The results of the principal component analysis of raw sequencing reads is shown in figure 2.1, principal component one explains 23.67% of the total variation and principal

component two explains 9.04% of the variation. The kinship matrix generated by Hierfstat is displayed in figure 2.2. There appear to be four distinct genetic groupings. Table 2.1 displays the results of the PRR screen of the seedlings the megagametophytic tissue was harvested from.

2.3.2 Genetic map construction and PRR Linkage and association mapping

ASmap produced a genetic map with 30 LG larger than 15 markers. Figures 2.3 and 2.4 display the graphical representation of the linkage groups and QTL mapping respectively and table 2.2 is a table of the most significant markers on the QTL map. Many of the markers collapsed to LG1, an issue common with MSTmap assemblies of samples with a high degree of missing data (Monroe et al. 2017).

Based on independent marker contingency testing, 26 tags out of the 124,159 tags segregating within the Kaz24 family were significantly associated with the Phytophthora root rot resistance phenotype at an experiment-wide false discovery rate of 0.05 (table 2.3).

2.4 Discussion

2.4.1 Genotyping-by-sequencing as an effective marker generation technique

Conifers make for difficult genomic research subjects, with genomes that can be larger than 70Gb, and complex genomic landscape heavily interspersed with non-genic transposable elements (De La Torre et al. 2014). The Trojan fir genome is estimated to be ~16 Gb, with a haploid chromosome number of 12, similar in size to Caucasian fir (*Abies nordmanniana*)(Kaya et al. 2008).

GBS of diploid needle foliage from Kaz24 seedlings generated 185,814 tags using Tassel software (Kohlway, 2015), of which 117,345 were putatively segregating. In this study of megagametophyte tissue, the same software generated 173,240 tags of which 124,159 were segregating. Only 63% of the total GBS tags overlap between the two data sets. The discrepancy between the two data sets can be explained by the difference in tissue types. Recent studies have shown that even though mature megagametophytes in conifers are mostly lipid and starch storage tissues (Vuosku et al. 2015), they do have unique gene expression profiles (Cervantes et al. 2021). In *Pinus sylvestris,* 21% of 7,171 genes were uniquely expressed in megagametophyte tissue relative to the bud, embryo, needle, and phloem tissues (Cervantes et al. 2021). Where there is differential gene expression, there is likely also differential DNA methylation. Both of the restriction endonucleases used in this study are sensitive to cytosine methylation, so the amplified restriction fragments could be different from tissues with differing methylation patterns. Additionally, the megagametophytic tissue is haploid and contains only the maternal genetic material whereas the first study used diploid needle foliage tissue. The needle foliage libraries contained not only a recombinant maternal haplotype, but an unknown number of pollen parent haplotypes. There are likely tags unique to the pollen parents that are not present in the megagametophytic tissue.

In this study, GBS generated markers identified an unexpected amount of genetic structure, likely owing to foreign seed contamination. The unknown genotypes likely were introduced at some point during the subsequent processing of the seed from cones collected in Turkey. The presence of “alien” seed or pollen within a breeding program is a well-known issue with mis-identified or mislabeled material as a common problem among populations. In

one study 10% of documented relationships were proven incorrect with molecular markers and a pine breeding program had 5% of their crosses mislabeled (Richardson et al 2005). Burdon and Wilcox sum up the issue perfectly: “The misidentification rates being thus detected are embarrassing-but often not admitted in print” (2006). Early genotyping work using restriction fragment length polymorph (RFLP) tags has been used to confirm pedigrees in research and commercial field plots but are costly and time consuming. GBS represents a rapid, high-throughput alternative capable of genotyping many individuals at a reduced cost (Poland et al. 2012). Furthermore, the markers generated by GBS can be used to develop bait for genotyping assays to reduce cost and time for genotyping future genetic cohorts (Rasheed et al. 2017).

2.4.2 Genotyping-by-sequencing for use in genetic map construction in conifers

The major advantage of using megagametophytic tissue is that the tissue represents recombinant haplotypes of the maternal parent. This feature is a two-fold boon as not only does it allow for the screening of pseudogene copies, with enough sampling, the maternal genotype can be reassembled. The collection of new seed from fir natural stands has often been done by collection of mature open pollenated cones. There is no historical genetic information of the maternal Kaz24 tree, so only other method for identifying the maternal genotype would be to return to the site of the tree, a difficult or costly endeavor for internationally sourced samples. With the removal of the foreign genetic components, the total number of individuals was reduced from the original set of 96 samples down to 63. Furthermore, aggressive marker screening required to remove multicopy genes further reduced the pool of markers. The best conditions combined with the pseudo-parent generated

with STACKs and ASMap produced a genetic map with of 30 linkage groups, more than would be expected for a true to chromosome scale map.

Contemporary studies have effectively used reduced marker linkage mapping to study Fusiform rust resistance in loblolly pine (*Pinus taeda*) (Lauer and Isik 2021). The study was able to create a high-resolution linkage map using two controlled crossed mapping populations of 1000 full-sibling cohorts with fully genotyped parents. Similarly large, designed mapping populations would likely be required to resolve a Trojan fir linkage map.

2.4.3 Phytophthora root rot linkage and association mapping in Trojan fir megagametophytes

Throughout the QTL analyses, no combination of parameters and settings were able to pick up a significant signal, which in part is due to the small sample size, but this result suggests that it is unlikely that Phytophthora root rot resistance is due to a single gene. Even with a sample size as small as 63, if a single resistance gene conferred resistance to *Phytophthora*, it is likely that it would have been picked up. PRR resistance in Kaz24 has been shown to be segregating. This adds evidence for a more quantitative, multi-gene resistance host-pathogen interaction in Trojan fir, and potential other fir species. Indeed, resistance to *Phytophthora capsici* in commercial squash (*Cucurbita pepo*) has also been shown to be controlled by multiple loci (Michael et al. 2021) with at least 5 different sites along the genome linked to resistance to *Phytophthora*.

Even though the megagametophyte GBS markers failed to generate a complete linkage map, the markers generated can be individually tested for association to PRR resistance. Of the 26 tags, only three had any sequence similarity to genes of known function.

Tag.130336: chinesis cation/calcium exchanger 5, Tag.13145: *Abies alba* NADH dehydrogenase subunit 4L, and Tag.253578: *Abies grandis* pinene synthase gene. Tag.13145 is most notable here as NADH dehydrogenases have been previously linked to *Phytophthora* resistance (Coelho et al. 2006).

2.5 Conclusion

The foreign seed contamination, the abundance of missing data in GBS libraries, and the complexity of the Trojan fir genome compounded together to prevent the realization of the original goal of chromosome scale linkage groups. This study may provide a cautionary tale to future conifer genomic researchers and projects, as the constraints of conifer genomics should not be underestimated.

The rapid development of sequencing technologies has outpaced much of the utility in using GBS for the generation of linkage maps, at least in genomes as large and complex as conifers. If this study was repeated in 2022, a mixture of full genome shotgun sequencing and long read sequencing through either PacBio or Nanopore systems could be used. The longer sequencing reads those systems create would help to bridge non-genic repeatable elements that confound short-read sequencing assemblies.

The megagametophyte GBS tags generated here still represent and useful genomic tool for future studies. The markers associated with PRR resistance within the Kaz24 family will guide future progeny selection and controlled crosses. Additionally, the filtered markers represent a pool of potential non-pseudogene sites within the Trojan fir genome. These sites would make great targets for identification of SNPs for a genotyping chip assay. With a

reduced set of target markers, the cost to genotype each individual would go down, allowing for more total genotyped samples in a cost-constrained mapping population.

2.6 Tables

Table 2.1: Phytophthora root rot resistance phenotype of megagametophyte samples

Phytophthora root rot response phenotype for *Phytophthora cinnamomi* inoculated seedlings from which the megagametophytic tissue was collected: R= Resistant; S= Susceptible. A count of the days each seedling was alive post Phytophthora inoculation during the 16-week trial: a value of 112 days corresponds to the seedling surviving beyond the screening period.

Sample	PRR Resistance	Days alive Post inoculation
M01	R	112
M02	S	14
M03	R	112
M04	S	18
M05	S	10
M06	R	112
M07	S	18
M08	R	112
M09	S	28
M10	S	14
M11	S	52
M12	S	32
M13	S	34
M14	S	16
M15	R	112
M16	S	14

Table 2.1 (continued).

M17	**R**	**112**
M18	S	20
M19	S	10
M20	S	28
M21	R	112
M22	S	70
M23	R	112
M24	R	112
M25	S	10
M26	R	112
M27	R	112
M28	S	28
M29	R	112
M30	S	10
M31	S	18
M32	S	24
M33	R	112
M34	R	112
M35	R	112
M36	S	20
M37	S	36
M38	S	16

Table 2.1 (continued).

M39	**S**	**28**
M40	S	18
M41	S	26
M42	S	30
M43	S	28
M44	S	70
M45	R	112
M46	R	112
M47	R	112
M48	S	24
M49	R	112
M50	S	24
M51	S	18
M52	R	112
M53	R	112
M54	S	14
M55	R	112
M56	R	112
M57	R	112
M58	S	20
M59	R	112
M60	S	24

Table 2.1 (continued).

M61	**R**	**112**
M62	S	30
M63	S	30
M64	S	30
M65	R	112
M66	S	18
M67	S	20
M68	S	22
M69	R	112
M70	S	28
M71	R	112
M72	S	26
M73	S	44
M74	S	14
M75	S	70
M76	R	112
M77	R	112
M78	R	112
M79	S	22
M80	R	112
M81	S	12
M82	R	112

Table 2.1 (continued).

M83	**R**	**112**
M84	R	112
M85	S	24
M86	S	34
M87	S	56
M88	S	22
M89	R	112
M90	S	22
M91	S	38
M92	R	112
M93	R	112
M94	R	112
M95	S	22
M96	S	10

Table 2.2: PRR resistance QTL mapping
A list of markers from rQTL linkage mapping against PRR response phenotype with linkage group (LG) assignment and position in cM along the LG. Log-odds score (LoD) for each marker for linkage to PRR resistance.

Marker	LG	Position	LoD
cL.1.loc3422	L.1	3422	1.7974
31035	L.275	18.6	1.5174
32895	L.45	87.4	1.3889
487	L.31	382.1	1.3852
32306	L.56	164.5	1.3621
37506	L.19	86.6	1.1672
2027	L.54	93.7	0.9995
39838	L.139	140.4	0.9816
27141	L.52	142.3	0.9462
13240	L.36	85.9	0.9435
cL.34.loc230	L.34	230	0.9283
127502	L.32	180.8	0.9203
1414	L.62	141.3	0.8961
72005	L.43	415	0.8714
655	L.35	131.6	0.8551
cL.25.loc101	L.25	101	0.7517
cL.99.loc204	L.99	204	0.7383
2951	L.123	0	0.7327
258073	L.112	206.7	0.7198
2886	L.63	118.4	0.6473
cL.97.loc207	L.97	207	0.6364
33958	L.28	216.4	0.6276
19795	L.67	167	0.5958
207983	L.96	152.4	0.5577
cL.13.loc77	L.13	77	0.4985
3456	L.49	99.2	0.4682
cL.14.loc66	L.14	66	0.4519
73019	L.44	17.8	0.4383
5483	L.90	17.3	0.4004
18723	L.211	103.1	0.3878
9141	L.30	357	0.3489
cL.78.loc306	L.78	306	0.3152
cL.21.loc49	L.21	49	0.3059
40615	L.76	0	0.2877
380	L.24	163.1	0.2776
102265	L.66	262.8	0.2682
117152	L.17	24.9	0.2665

Table 2.2 (continued).

24844	L.7	363.7	0.2534
103489	L.144	192.4	0.2518
36693	L.108	121.2	0.2506
cL.8.loc110	L.8	110	0.2388
4313	L.57	123.7	0.2354
26895	L.27	101.8	0.2329
cL.65.loc54	L.65	54	0.1986
115031	L.170	20.8	0.1498
4333	L.145	0	0.134
31550	L.55	17.9	0.1007
113457	L.9	151	0.0814

Table 2.3: Kaz24 Megagametophyte tags associated with PRR resistance

Significant marker full sequences.

Marker	Sequence
tag.165217	TGCAGCGGCTGTGCCAGTCGATGCGGGAATAATTTACGCCACGGACTGCGAGAGCAGCTCGGCT
tag.69350	TGCAGATCGATTCAGACCTGTGTGATCCATCGTGGTTTTTTTTCGTTGTGCTTCCTCAGCATGT
tag.17706	TGCAGAAGCTACGATTGTGAGCACCGCCGCAATCACGACCAGAACCGAGCTCTTTCGGGCGCTC
tag.208237	TGCAGGAATAAGTTACGGAATTCTTAAAAAAGATATTGATTTATTAATATTTTAACTGGTGGAA
tag.82764	TGCAGATTGCTTTTACCCGTCACTGTCTGGTATCTAAGCTGCCATTATTTCGATCCACAAAATT
tag.278640	TGCAGGTCATGGGCGATCTGGGACTGGAGAATGCCTGGCAAGGCGAAACTAATTTCTGGGGCAG
tag.130336	TGCAGCCAGGGCCACGAAGCCCACGACGAATGCGGAGACGAAGGTGCCCGCGCTGAGAATGGCC
tag.360447	TGCAGTTTGTCTCCACGGTGCCCGCCAAAGGGGAAACCTGGCTGTACAGCAACGTGCCCTATAA
tag.16213	TGCAGAAGCACATCAAGAGGACCTTACAAAACGCTTTTTAAAACACTAGGGCTGATAGTCTTGC
tag.121996	TGCAGCATCTGGGGCTACGCGAGTACTTCGATGTGATTGTGGGGGCCGATGACGTCACCCGCCA
tag.152740	TGCAGCGATACGGGCATGTCGCCGATCTCGTCGAGGAACAGCGTGCCCCGATGCGCGGTTTCGA
tag.77541	TGCAGATGGTGAGATAAGCAAGTACAAAGCTCGACTTGTCGCAAAAGGATACTCACAGGTACAT
tag.64430	TGCAGATATCTTTTATTTAGACTTAACATTATTTGTTGATGGCAGGGCAGACACTATGCATGAT
tag.253754	TGCAGGGAGCCTGCACAACTTGCCCATCTCAATTTTTTTTTCCCACTGTGACGCAGGGATTTAT
tag.263563	TGCAGGGGTCTAAAATTGCCTTGATTCTCGGAATATCAAGCCATTGTAAATTTACCAGTTTTAA
tag.131451	TGCAGCCATCCGCAACTATGCTTCCGCAACTGATGGTAGATCCACTGACTCTGCTGTCCCTCTA
tag.184938	TGCAGCTGATCACCAACATATGCATCTTTTTGCCCCATACGAACTATCACACTATTGTGCATTG
tag.106252	TGCAGCAGATCCCGTCAGTGATCAACGGCAGCCTGGTATCCGATCTGATCGTATACCTCGGTAG
tag.267527	TGCAGGGTTTCCTGCACCGCGTTGAAGCGGGGGGCGAGGCGCATGAACAGCAGCAGCATCGTGA
tag.50976	TGCAGAGCTCCCCACTAGCGGGTTCCAACGGACTGTAACACCTGCATTCATGCAGTTCCATAAC
tag.253578	TGCAGGGAGAGCCTACAATGCGGTTCCCAGGGAATGGACCCCAGTACCTGTGGGCACCGAAGTG
tag.154027	TGCAGCGCAAGCACGATGTTCTCCCGTATGGAAAGCTCCGCAATGATGCCTTCCTTCTTCCTGT
tag.348480	TGCAGTTCTAGTGCCGCTCCACCTTTCGACCCCACAACGTCTTCCTCATATAAATCATTAGCCT
tag.122268	TGCAGCATCTTTCACTCTTTGTAAAGATGAGACTTTGTTTATACAGAAATACGATATTAAAACA

Table 2.3 (continued).

tag.157936	TGCAGCGCCTTGGGCCGCTTGATCATATATCGCTCCCCAGTGGTGTTTTCTTGGAGCTCATTGC
tag.203598	TGCAGGAACAATGAGAACTGCTCTTCTTCCAGCCCCTGCAAAATGTCGTGTTCCTGGGTCAGCC

2.7 Figures

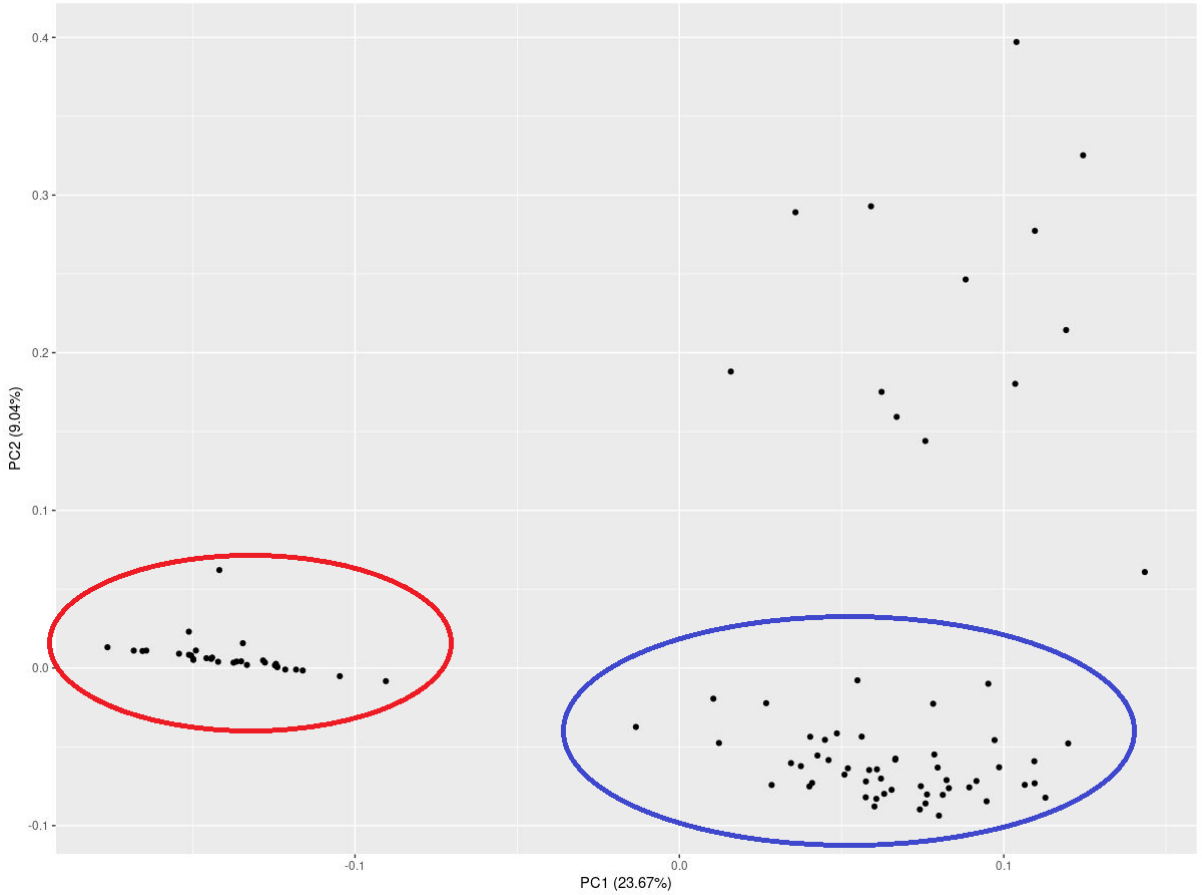
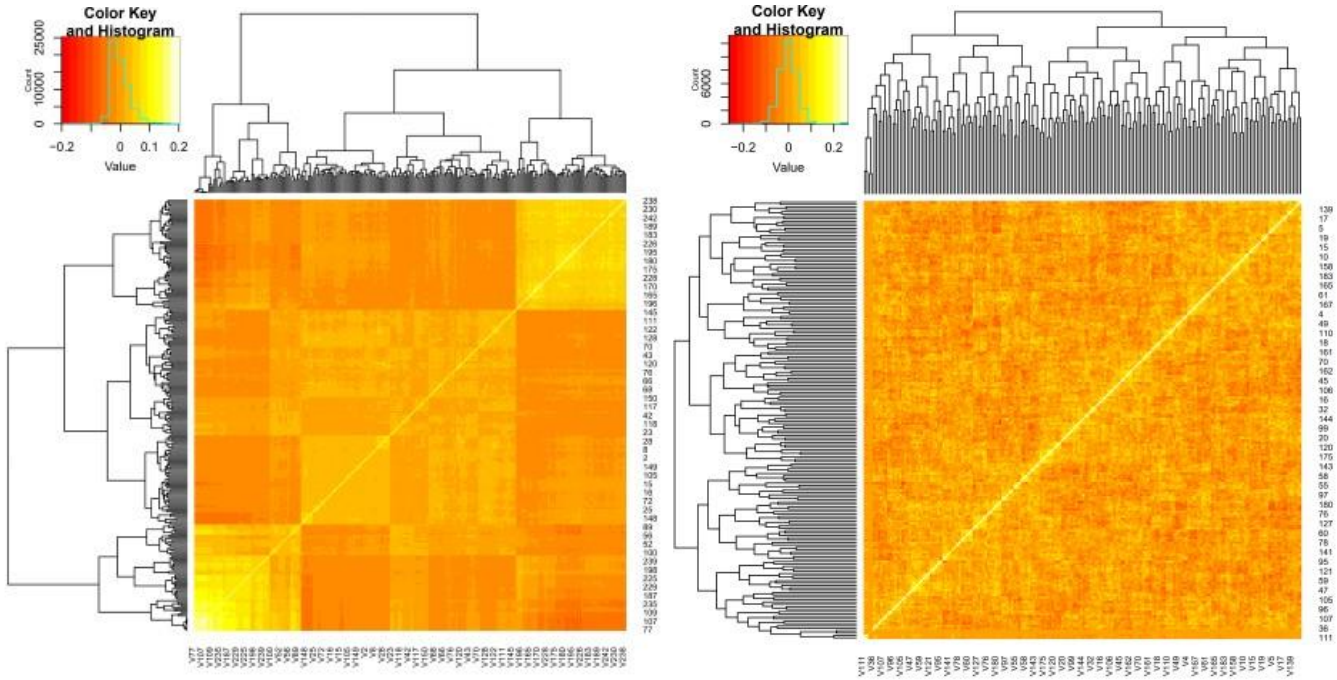


Figure 2.1: PCA of Raw Megagametophyte sequencing reads

Results of the Principal component analysis of the raw sequencing reads per individual. The individuals representing the two expected Kaz24 haplotypes are circled in red and blue.



Kaz24 Trojan Fir Megagametophytes

FF51 Fraser Fir Megagametophytes

Figure 2.2: Comparisons of kinship matrices from two single family analyses

There is notable genetic structure present in the Kaz24 trojan fir megagamtophytes not present in the single Fraser fir megagametophyte family.

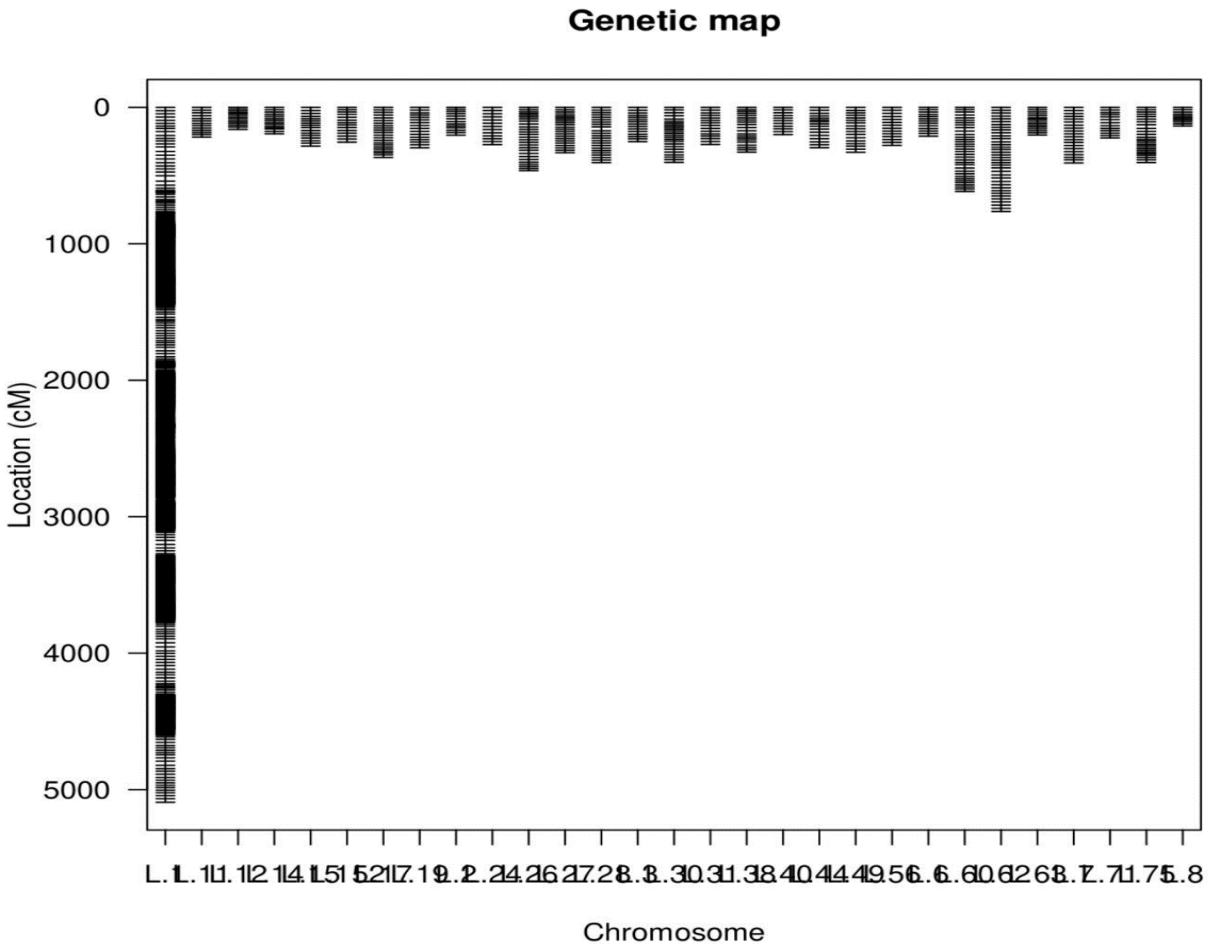


Figure 2.3: Full Kaz24 megagametophyte genetic map

Kaz24 Trojan fir genetic map with 30 linkage groups.

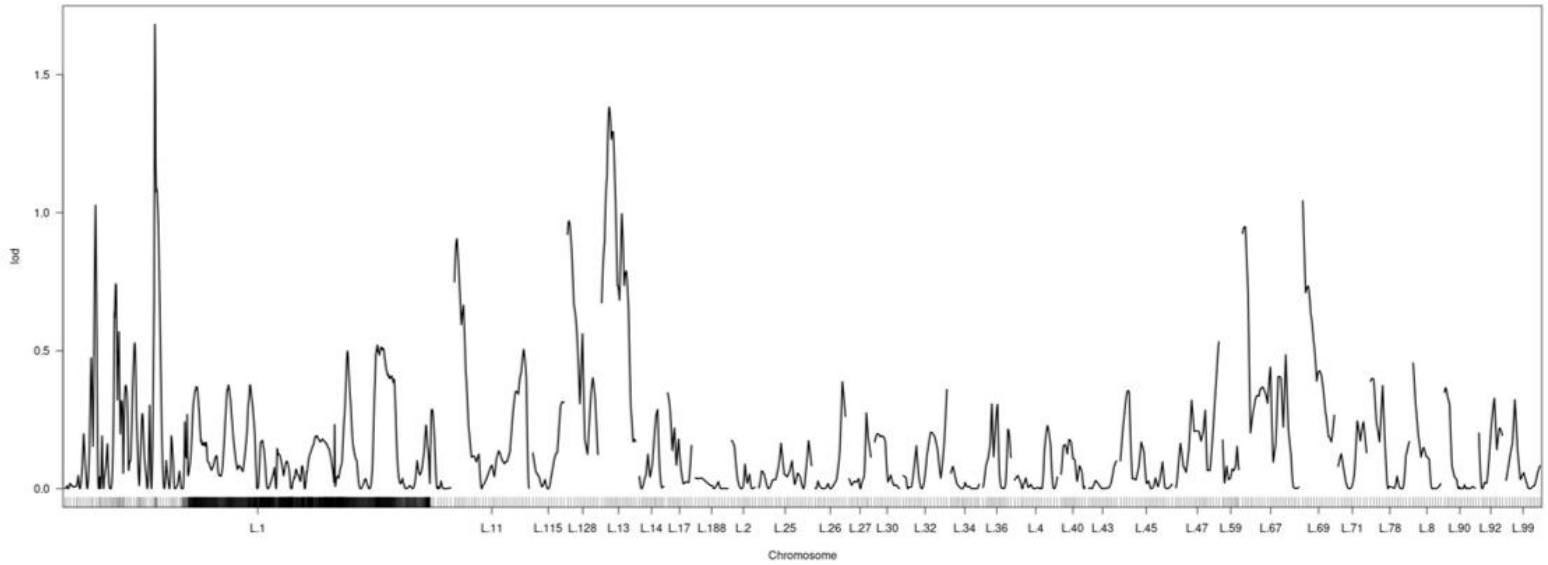


Figure 2.4: Phytophthora root rot resistance QTL mapping in assembled Kaz24 megagametophyte genetic map

Results of the linkage mapping of GBS markers with Phytophthora root rot resistance. No test marker had significant linkage to PRR resistance.

2.8 Literature cited

Auckland, L. D., Johnston, J. S., Price, H. J., & Bridgwater, F. E. (2001). Stability of nuclear DNA content among divergent and isolated populations of Fraser fir. Canadian Journal of Botany; Ottawa, 79(11), 1375–1378.

Beaulieu, J., Doerksen, T., Clément, S., MacKay, J., & Bousquet, J. (2014). Accuracy of genomic selection models in a large population of open-pollinated families in white spruce. Heredity, 113(4), 343–352. https://doi.org/10.1038/hdy.2014.36

Bernhardsson, C., Vidalis, A., Wang, X., Scofield, D. G., Schiffthaler, B., Baison, J., Street, N. R., García-Gil, M. R., & Ingvarsson, P. K. (2019). An Ultra-Dense Haploid Genetic Map for Evaluating the Highly Fragmented Genome Assembly of Norway Spruce (Picea abies). G3 Genes|Genomes|Genetics, 9(5), 1623–1632. https://doi.org/10.1534/g3.118.200840

Broman, K. W., Wu, H., Sen, S., & Churchill, G. A. (2003). R/qtl: QTL mapping in experimental crosses. Bioinformatics, 19(7), 889–890. https://doi.org/10.1093/bioinformatics/btg112

Brown, G. R., Gill, G. P., Kuntz, R. J., Langley, C. H., Neale, D. B., & Clegg, M. T. (2004). Nucleotide Diversity and Linkage Disequilibrium in Loblolly Pine. Proceedings of the National Academy of Sciences of the United States of America, 101(42), 15255–15260.

Burdon, R. D., & Wilcox, P. L. (2007). Population management: Potential impacts of advances in genomics. New Forests, 34(2), 187–206. https://doi.org/10.1007/s11056-007-9047-6

Cervantes, S., Vuosku, J., & Pyhäjärvi, T. (2021). Atlas of tissue-specific and tissue-preferential gene expression in ecologically and economically significant conifer Pinus sylvestris. PeerJ, 9, e11781. https://doi.org/10.7717/peerj.11781

Coelho, A. C., Horta, M., Neves, D., & Cravador, A. (2006). Involvement of a cinnamyl alcohol dehydrogenase of Quercus suber in the defence response to infection by *Phytophthora cinnamomi*. *Physiological and Molecular Plant Pathology*, *69*(1–3), 62–72. https://doi.org/10.1016/j.pmpp.2007.01.001

Covarrubias-Pazaran, G., Diaz-Garcia, L., Schlautman, B., Deutsch, J., Salazar, W., Hernandez-Ochoa, M., Grygleski, E., Steffan, S., Iorizzo, M., Polashock, J., Vorsa, N., & Zalapa, J. (2016). Exploiting genotyping by sequencing to characterize the genomic structure of the American cranberry through high-density linkage mapping. BMC Genomics, 17(1), 451. https://doi.org/10.1186/s12864-016-2802-3

De La Torre, A. R., Birol, I., Bousquet, J., Ingvarsson, P. K., Jansson, S., Jones, S. J., Keeling, C. I., MacKay, J., Nilsson, O., Ritland, K., Street, N., Yanchuk, A., Zerbe, P., & Bohlmann, J. (2014). Insights into conifer giga-genomes. Plant physiology, 166(4), 1724–1732.

Dodt, M., Roehr, J., Ahmed, R., & Dieterich, C. (2012). FLEXBAR—Flexible Barcode and Adapter Processing for Next-Generation Sequencing Platforms. Biology, 1(3), 895–905. https://doi.org/10.3390/biology1030895

Frampton, J., Isik, F. & Benson, D.M. (2013). Genetic variation in resistance to Phytophthora cinnamomi in seedlings of two Turkish Abies species. Tree Genetics & Genomes, vol 9, pp. 53–63.

Glaubitz, J. C., Casstevens, T. M., Lu, F., Harriman, J., Elshire, R. J., Sun, Q., & Buckler, E. S. (2014). TASSEL-GBS: A High Capacity Genotyping by Sequencing Analysis Pipeline. PLoS One; San Francisco, 9(2), e90346. http://dx.doi.org.prox.lib.ncsu.edu/10.1371/journal.pone.0090346

Goudet, J. (2005). Hierfstat, a package for r to compute and test hierarchical F-statistics. Molecular Ecology Notes, 5(1), 184–186. https://doi.org/10.1111/j.1471-8286.2004.00828.x

Ivanova, N. V., Dewaard, J. R., & Hebert, P. D. N. (2006). An inexpensive, automation-friendly protocol for recovering high-quality DNA: TECHNICAL NOTE. Molecular Ecology Notes, 6(4), 998–1002. https://doi.org/10.1111/j.1471-8286.2006.01428.x

Kaya, Zeki & Skaggs, A. & Neale, D.B.. (2008). Genetic differentiation of Abies equi-trojani (Asch. & Sint. ex Boiss) Mattf. populations from Kazdağı, Turkey and the genetic relationship between Turkish firs belonging to the Abies nordmanniana Spach complex. Turkish Journal of Botany. 32. 1-10.

Kohlway, W.H. (2015). Mapping Phytophthora Root Rot Resistance in Fir Species through Genotyping by Sequencing. [Master Dissertation, North Carolina State University]

Kumar, S., & Richardson, T. E. (2005). Inferring relatedness and heritability using molecular markers in radiata pine. Molecular Breeding, 15(1), 55–64. https://doi.org/10.1007/s11032-004-2059-4

Lauer E, Isik F. (2021). Major QTL confer race-nonspecific resistance in the co-evolved Cronartium quercuum f. sp. fusiforme–Pinus taeda pathosystem. Heredity. 1–12.

Matallana Ramirez, L., Kurt, Y., Kohlway, W., Frampton, J., & Whetten, R. (2019). Easy and Inexpensive Nucleic Acid Extraction Protocol for Recalcitrant Species https://doi.org/10.17504/protocols.io.5pwg5pe

Monroe, John & Allen, Zachariah & Tanger, Paul & Mullen, Jack & Lovell, John & Moyers, Brook & Whitley, Darrell & Mckay, John. (2017). TSPmap, a tool making use of traveling salesperson problem solvers in the efficient and accurate construction of high-density genetic linkage maps. BioData Mining. 10. 10.1186/s13040-017-0158-0.

Nyholt, D. R. (2000). All LODs Are Not Created Equal. Am. J. Hum. Genet., 7.

Poland, J. A., Brown, P. J., Sorrells, M. E., & Jannink, J.-L. (2012). Development of High-Density Genetic Maps for Barley and Wheat Using a Novel Two-Enzyme Genotyping-by-Sequencing Approach. PLoS One; San Francisco, 7(2), e32253. http://dx.doi.org.prox.lib.ncsu.edu/10.1371/journal.pone.0032253

Przybylski, P., Kowalczyk, J., Odrzykoski, I., & Matras, J. (2019). Identyfing alien genotypes and their consequences for genetic variationin clonal seed orchards of Pinus sylvestris L. Dendrobiology, 81, 40–46. https://doi.org/10.12657/denbio.081.005

Rognes, T., Flouri, T., Nichols, B., Quince, C., & Mahé, F. (2016). VSEARCH: A versatile open source tool for metagenomics. PeerJ, 4, e2584. https://doi.org/10.7717/peerj.2584

Voichek, Y., & Weigel, D. (2020). Identifying genetic variants underlying phenotypic variation in plants without complete genomes. Nature Genetics, 52(5), 534–540. https://doi.org/10.1038/s41588-020-0612-7

Vuosku, J. (2015). Expression of catalase and retinoblastoma-related protein genes associates with cell death processes in Scots pine zygotic embryogenesis. 13.

Wu, Y., Bhat, P. R., Close, T. J., & Lonardi, S. (2008). Efficient and Accurate Construction of Genetic Linkage Maps from the Minimum Spanning Tree of a Graph. PLoS Genetics, 4(10), e1000212. https://doi.org/10.1371/journal.pgen.1000212

Zoubarev, A., Hamer, K. M., Keshav, K. D., McCarthy, E. L., Santos, J. R. C., Van Rossum, T., McDonald, C., Hall, A., Wan, X., Lim, R., Gillis, J., & Pavlidis, P. (2012). Gemma: A resource for the reuse, sharing and meta-analysis of expression profiling data. Bioinformatics, 28(17), 2272–2273. https://doi.org/10.1093/bioinformatics/bts430

CHAPTER 3

3.1 Introduction

RNA sequencing, or RNA-seq, has been a vital tool for the study of numerous genetic traits and diseases. The goal of RNA-seq experiments is to obtain DNA sequences corresponding to messenger RNAs (mRNAs) expressed in the tissue sample of interest. To prepare RNA-seq libraries, total cellular RNA is first extracted from a tissue sample. Because ribosomal RNA (rRNA) makes up a large proportion of the total cellular RNA (typically greater than 80%), and are generally not expected to be useful in mapping traits such as disease resistance, they are screened out of the extracted RNA in one of two ways: either using oligo-dT affinity materials to select for Poly-A tracts on polyadenylated mRNAs (Poly-A selection)(Aviv and Leder, 1972), or by targeted removal with oligonucleotide probes complementary to the rRNA sequences (ribosomal RNA depletion)(Morlan et al. 2012). Poly-A selection has the advantage that no prior sequence knowledge is required for removal of rRNAs, but the disadvantage that it will filter out other important non-polyadenylated RNAs (like miRNA or siRNA). The selection stage is precisely when the quality of the RNA is most important, and high-quality RNA is required for good RNAseq libraries using oligo-dT. High-quality RNA can be described as RNA extracted from a tissue with mostly intact, full-length RNA molecules (Schroeder et al. 2006). Conversely, low quality RNA is comprised of sheared, broken, or partially degraded fragments of the RNA, and is generally common in samples isolated from necrotic tissues or mishandled library preparations. In fragmented, low-quality RNA the poly-A tail is often separated from most of the rest of the mRNA molecule, preventing the isolation of full-length mRNA molecules. After removal of ribosomal RNA by poly-A selection or rRNA depletion, the selected RNA is then

fragmented, and with random priming, reverse transcribed into complementary DNA (cDNA). The cDNA is then prepared for sequencing based on the specifications of the sequencing platform used. Once the cDNA fragments have been sequenced, they can be reassembled back into full length transcript contigs either using a reference transcriptome or standard sequence assembly methods if no reference is available.

It's important to note that transcriptional profiles are variable between different tissue types; for instance, photosynthetic foliage will have a different expression profile than the roots of the same plant. Furthermore, even single cells from within a tissue can display differences in expression (Moreno-Risueno et al. 2015). When searching for potential genes involved in an immune response it is essential to select the cells in which a transcriptional response to a pest is hypothesized to occur. In the case of Phytophthora root rot of fir, the literature points to the root tips as the zone likely to respond (Hardham 2005). However, the study of root tissue can be difficult to collect from established conifers, especially the actively growing apical tips. In natural stands, it requires digging down to symptomatic root tissue in non-sterile settings, collecting root samples which must then be immediately frozen in either liquid nitrogen or dry ice and kept frozen until they can be processed in the lab. An additional consequence of field research is that the tissue quality collected has many opportunities to degrade throughout the collection process even in ideal circumstances. RNA is not as stable as DNA and degrades quickly in necrotic or mishandled tissue samples due to the exposure to cellular and environmental RNAses. *Phytophthora cinnamomi* Rands, and other members of the *Phytophthora* genus exhibit necrotrophic parasitism in which the pathogen first kills the susceptible host's tissue and then acquires nutrients from the lysed

cells (Hardham 2005). Therefore, the extraction of high-quality host RNA from infested root samples is often a difficult endeavor.

The use of ribosomal RNA depletion methods makes possible construction of RNA-seq libraries from degraded RNA samples (Adiconis et al. 2013). This approach often allows analysis of gene expression patterns in tissue samples from which isolation of high-quality RNA is very difficult (Morlan et al. 2012). The collection of infested root samples from Fraser fir in Christmas tree, or a replicated facsimile under field lab conditions, are such conditions.

In this chapter we describe a library preparation technique novel to conifer tissue to save what would otherwise be un-usable, necrotic tissues samples collected from PRR diseased root tissue and gain a glimpse into the microbial communities associated with different disease phenotypes in infected roots.

3.2 Materials and Methods

3.2.1 Phytophthora preparation

Phytophthora cinnamomi isolate 23ss04 (mating type 2) was chosen as the inoculum for this study. The isolate was collected from infested Fraser fir in the North Carolina Mountains and has been used as the primary *P. cinnamomi* inoculum in previous PRR resistance screens due to its high aggression and correlation to resistance profiles of different isolates and *Phytophthora* species (Frampton et al., 2013; Kohlway et al., 2017).

The *P. cinnamomi* culture was roused from hempseed-water cold storage (Jeffers et al., 2007) by plating colonized agar plugs onto a PARPH media plate, a *Phytophthora*-selective growth medium, and incubating in the dark at room temperature. The rice grain

inoculum method was prepared following Holmes and Benson's specifications (1994). Briefly, four Erlenmeyer flasks filled with 25g of long-grain rice in 17ml of diH_2O were autoclaved twice, with 24 hours between each sterilization. Three small (2 cm diameter) *P. cinnamomi*-colonized agar plugs were transferred to each flask after the rice had cooled to room temperature following the second round of sterilization. The flasks were then sealed and incubated in the dark at room temperature for two weeks, with daily perturbations to facilitate full coverage of the rice grains by *Phytophthora*. The *Phytophthora*-colonized rice grains were then mixed into 0.014158 m^3 of potting media comprised of peat, vermiculite, and perlite at a 1:1:1 ratio and placed into a 50.8x35.56x15.24 cm container and stored on a bench in a greenhouse. The *Phytophthora*-colonized potting medium was incubated for two weeks to allow for full colonization of the medium.

3.2.2 Trojan fir seedling production and Inoculation

300 Trojan fir seed collected in 2008 from natural stands in the Kazdağı provenance of Turkey for a previous study (Frampton et al. 2012), were cold-stratified for one month. The stratified seed were then buried 2.54-cm-deep into a vermiculite filled container kept indoor at room temperature under a fluorescent lamp. The germinating seed were watered daily for two weeks. After two weeks, the seedlings had grown to about 6 cm in length with unbranched root systems. The two-week-old seedlings were then transferred to a greenhouse and transplanted into a box of Phytophthora-colonized potting medium, a prepared as described in section 3.2.1 above. 20 seedlings were transplanted into a sterile potting medium to serve as non-inoculated controls.

Each seedling was gently removed from the *Phytophthora*-ridden medium after 8 days of exposure. Sterile deionized water was then used to remove as much contaminated medium from the root system as possible. After the root was rinsed, a sterile scalpel was used to remove 3 cm of root tissue from the tip. The excised tissue was then transferred to Qiagen deepwell strip-tubes, preloaded with a 3 mm metallic bead, and stored on dry ice while in the greenhouse. Each seedling was then transferred to a 66 cm^3, 2.5 cm x 16cm deep Ray Leach Cone-tainer™ (Stuewe & Sons, Inc., Corvallis, OR, USA) filled with sterile potting medium. Each transferred seedling was then monitored for 16 weeks to observe the phenotype of Phytophthora root rot resistance level assessed via percent shoot necrosis.

3.2.3 RNA Library preparation

RNA was extracted from the root tip samples following the method of Matallana et al. (2019). Briefly, the frozen tissue samples were ground to a powder between two 3 mm, metallic beads using a Qiagen mixer mill. The powdered samples were then processed using buffers adapted from an insect DNA extraction protocol (Ivanova et al. 2006). The RNA samples were then quantified using a Nano-drop microspectrophotometer. A selection of the RNA extractions were then pooled into 10 groups of 10 individuals based on the experimental phenotype of mortality due to Phytophthora root rot. In total 4 pools of resistant, 4 pools of susceptible, and 2 pools of control (non-inoculated) seedlings were made; the individuals selected, and their quality scores are listed in table 3.1.

The pooled samples were then sent to the sequencing center at DHMRI (Kannapolis, NC) for ribosomal depletion and RNA-seq library preparation. Ribosomal depletion was performed following the protocol of Adiconis et al. (2013) on the extracted root RNA, using

custom oligonucleotides complementary to Fraser fir ribosomal RNA sequences. Appendix A lists the oligo baits complementary to Fraser fir mitochondrial, chloroplast, and nuclear ribosomal RNAs used for depletion; these sequences were derived from rRNA sequences detected in a Fraser fir transcriptome assembly using RNAmmer (Lagesen et al. 2017) (Whetten unpublished). No further sample shearing was used following ribosomal depletion prior to reverse transcription and library preparation. The rRNA-depleted RNAseq libraries were then sequenced on an Illumina Hi-Seq 2500 flowcell with 100 bp single-end reads.

3.2.4 Bioinformatic analysis

FastQC was used to check sequence results for basic quality and screen out low-quality sequences. The sequencing reads were then normalized using the normalize-by-median plugin of khmer suite (Crusoe et al. 2015) combining the k=20 and k=31. The normalized reads were then assembled into contigs using BridgeR (Chang et al. 2015), with CAP3 (Huang and Madan 1999) used to annotate the resulting contigs. Bbduk from the BBMap suite (Bushell 2014) was used to screen and filter out remnant rRNA (Appendix A) and *Phytophthora cinnamomi* contamination using the published *P. cinnamomi* genome (Longmuir et al. 2018) respectively. Mash (Ondov et al. 2016) was used to identify microbially derived transcripts within the differentially-expressed transcripts (Chapter 4) of each phenotypic group: control, susceptible, and resistant roots.

Kraken2 (Wood et al. 2019) with a custom database comprised of the complete bacterial and fungal genomes from RefSeq was used to screen those transcripts and assign them to microbial genera. Krona (Ondov et al. 2011) was then used to visualize the resulting taxonomic assignments of transcripts from each phenotypic group of RNA samples. Finally,

Pavian (Breitwieser and Salzberg, 2016) was used to further refine the taxonomic assignments and identify taxons significantly enriched between phenotypic groups.

3.3 Results

3.3.1 Sequencing results

Table 3.2 describes the results of sequencing and screening for rRNA and *Phytophthora* contaminants within the 10 libraries. A total of 244,783,105 reads were sequenced, encompassing 18.67 Gb of sequence. Ribosomal depletion reduced rRNA in the libraries to 22,726,155 reads or about 9.28% of the overall sequencer reads. Filtering the dataset against the *Phytophthora cinnamomi* genome (Longmuir et al. 2018) identified a further 7,357,988 contaminant reads. The remaining 217,443,743 reads assembled into 96,524 contigs ranging from 201 bp to 20.3 kb. The reduction of rRNA using ribosomal depletion was roughly equivalent (Table 3.3) to the results from a subsequent RNAseq experiment where standard poly-A selection was utilized (Chapter 4).

3.3.2 Metagenomic analysis

The Mash software sketches of the total filtered RNA assembly identified 24% of RNA reads in PRR resistant roots, 37% of reads in PRR Susceptible, and 31% of control roots as microbially derived. Kraken2 analysis of the *Phytophthora*-filtered transcripts identified a total of 25.7% microbially derived transcripts when compared to the fungal and bacterial complete genome refseq collections. Krona generated hierarchal cladograms of bacterial and fungal genera identified by Kraken2 for each phenotypic group, as shown in Figure 3.1 and 3.2 respectively. Taxon assignment by Kraken2 for each combined

phenotypic group is shown in Figure 3.3. The most represented microbial genera and species and their proportion of the microbial derived RNA are listed in tables 3.4 and 3.5 respectively.

3.4 Discussion

3.4.1 rRNA depletion effective for use in conifer tissue

To date, rRNA depletion use for conifer RNAseq library preparation has not been published. In this study ribosomal RNA depletion achieved an overall reduction of rRNA comparable to that found in traditional oligo-dT-based methods of other conifer species; with less than ~10% of the total reads identified as rRNA contaminants for both methods (table 3.3). However, the rRNA depletion disproportionately reduced 18S rRNA over 28S relative to oligo(dT) with an average 28S/18S of 69.5 and 2.33 respectively (table 3.3).

The rRNA oligo bait developed from Fraser fir (Appendix A) was effective for removing the majority of the cellular rRNA. However, subsequent rRNA depletion library preparations may be improved by utilizing 28S/18S oligos developed from remnant Trojan fir rRNAs identified using RNAmmer (Lagesen et al. 2017) in this study.

3.4.2 Successful sequencing of Root RNA from Necrotic root tissue

The results of the ribosomal depletion technique are encouraging. First off, the use of ribosomal depletion facilitated the preparation and sequencing of the necrotized, low-quality RNA samples. Most sequencing centers advise against or even refuse to prepare libraries from samples with RIN scores as low as in this case (Table 3.2). Low RIN scores are problematic for the standard oligo-dT preparation technique as much of the highly

fragmented mRNA would be screened out with the rRNA and fail to make it to the sequencing flowcell. However, the yield and quality of RNAseq samples following ribosomal depletion (table 3.2) were more than sufficient for analysis of differential expression of the Trojan fir roots response to *Phytophthora* inoculation (Chapter 4).

Often sample availability or budgetary constraints limit the collection and preparation of sequencing libraries to a single attempt, where all the available tissue is used for RNA extraction. Each low-quality sample can represent weeks to months of time and energy lost, especially with slow growing organisms such as conifers. Thus, this study shows that ribosomal depletion-based library preparation is an option for salvaging just such otherwise lost samples.

3.4.2 Preliminary metagenomic assay

The depletion of rRNA using complementary rRNA-specific oligonucleotides has an additional consequence or utility, depending on perspective. Prokaryotic RNAs, normally screened due to oligo-dT selection, are still present in the rRNA depleted sample. Soil represents one of the most diverse microbial communities with upwards of 4×10^7 prokaryotic cells contained within a single gram of forest soil (Richter and Markewitz 1995). The relationships between microbial members are extremely diverse and range from pathogenic pests, *Pseudomonas syringae* and *Ralstonia solanacearum*, to beneficial symbionts and everything in between (Daniel 2005, Nesme et al. 2016). These microbial communities are important to understanding the complex response to root pathogens. There are examples of root symbionts that provide resistance to other potentially pathogenic microbes through either the excretion of antimicrobial factors or by out-competing other

microbes for physical space around the roots (Richter and Markewitz 1995, Toju and Sato 2018). *Firmicutes* and *Pseudomonads* have already been used for the pro-biotic effects in modern agriculture and there are many commercially available potting media additives containing beneficial mycorrhizae purported to improve disease resistance and plant growth. Although the potting medium that was colonized with *Phytophthora* was sterilized prior to the seedling inoculation, the experimental conditions permitted a reintroduction of microbes to the medium. A combination of the open-air exposure, microbes in the watering supply, and movement of insects around the experiment site all likely introduced the microbes detected in this study.

Mash's k-mer matching program was useful in mapping the subset of transcripts likely to be contributed to microbial contaminants of the root preparation and required less intensive compute power than Kraken2, but Mash lacked many classification tools available in Kraken2. Kraken2 has a high fidelity for classifying microbial community members within a mixed RNA sample (Wood et al. 2019). Figure 3.7.1 displays the distribution of bacterial communities associated with highly necrotic susceptible roots versus the healthy resistant roots. Of note, the root systems associated with resistant seedlings had a higher proportion Proteobacteria (79%) as compared to the susceptible seedlings (60%) (Figure 3.1). This increase in Proteobacteria in the resistant roots is in proportion to a reduction in Actinobacteria; 31% in susceptible roots and 11% in resistant roots respectively (figure 3.1). Many bacteria with the proteobacteria phylum, have been studied for their root associations and interactions. The *Pseudomonas* genus alone has wide variety of interactions with the rhizosphere, with the opportunistic pathogens such as *P. syringiae* to the commensal relationship formed between roots and *P. fluorescens* (Sitaraman 2015).

3.5 Conclusions and Future Work

This project validates ribosomal depletion as an effective alternative to oligo-dT selection for conifer RNAseq projects. This tool will be invaluable to future projects with equally problematic tissues samples; whether they be from diseased tissues as is described here or from samples that degrade during collections in non-laboratory environments such as field plots.

Although this study was not originally designed to delve into the metagenomics of the rhizosphere, the capture and retention of microbial RNA yielded by this technique could be refined for future exploration of the complex populations found around root systems. The presence of certain microbes within the rhizosphere may be important for the interaction between *Phytophthora* and fir roots as some research has shown (Sukhada et al., 2011).

Additionally, the use of ribosomal-depleted prepared RNAseq libraires would allow for the study of non-coding RNA in conjunction with other expression profiling of tissues. Non-coding RNAs include long noncoding RNA (lncRNA), microRNA (miRNA), and small interfering RNA (siRNA). SiRNA are important for post-transcriptional gene silencing in plants (Vaucheret et al. 2001). Known functions of noncoding RNAs are expanding as deeper sequencing of single-cell RNA samples identifies these new low-abundance transcripts and represent an additional level of transcription regulation important to many physiological traits.

3.6 Tables

Table 3.1: Quality and yield of RNA samples

The RNA yield and RNA integrity number (RIN) for sequencing library. Each sample is a pool of ten individual root tips pooled By Phytophthora root rot phenotype and non-inoculated controls.

Sample	Group	RIN	Conc. [ng/µl]
S01	Resistant	4.0	49.8
S02	Resistant	5.9	34.4
S03	Resistant	2.3	42.5
S04	Resistant	3.6	45.3
S05	Control	4.2	54.3
S06	Control	2.6	42.0
S07	Susceptible	4.0	48.1
S08	Susceptible	3.6	44.5
S09	Susceptible	5.8	50.3
S10	Susceptible	5.7	54.4

Table 3.2: Sequencing results and filtering summary

The total amount of sequencing reads per Kaz24 root RNA library with the number of ribosomal RNA and *Phytophthora* RNA contaminants removed during filtering.

Sample	**Total reads**	**rRNA contaminants**	**Filtered**	***Phytophthora* contaminant (% of filtered)**	**Final Read Count**
S01	26,712,035	2,721,904 (10.19%)	23,474,643	251,520 (1.07%)	23,223,123
S02	27,307,492	503,255 (1.84%)	26,518,325	229,887 (0.87%)	26,288,438
S03	25,866,957	2,561,489 (9.90%)	25,866,957	2,718,138 (10.51%)	23,148,819
S04	24,061,240	1,445,168 (6.01%)	24,061,240	1,631,936 (6.78%)	22,429,304
S05	20,286,281	1,587,932 (7.83%)	18,441,030	235,966 (1.28%)	18,205,064
S06	24,453,117	878,827 (3.59%)	23,219,294	333,046 (1.43%)	22,886,248
S07	26,246,034	1,384,760 (5.28%)	26,246,034	1,579,462 (6.02%)	24,666,572
S08	22,737,058	3,838,565 (16.88%)	18,587,349	223,530 (1.20%)	18,363,819
S09	23,769,205	5,586,922 (23.50%)	17,837,376	113,233 (0.64%)	17,724,143
S10	23,343,686	2,217,333 (9.50%)	20,549,483	41,270 (0.20%)	20,508,213
Total	244,783,105	22,726,155 (9.28%)	224,801,731	7,357,988 (3.27%)	217,443,743

Table 3.3: rRNA Reduction in Sequenced RNAseq Libraries

Comparisons between the efficacy of ribosomal RNA reduction between Oligo(dT) and rRNA depleted RNA library preparations.

Library Preparation	N	Mean Total % of rRNA	Median Total % of rRNA	Standard Deviation	Mean 28S/18S Ratio	Median 28S/18S Ratio
Oligo(dT)	13	8.82	10.49	5.86	2.33	2.57
rRNA Depletion	10	9.45	8.66	6.48	69.50	62.13

Table 3.4: Top 25 microbial genera by percent total read identified in RNAseq

A list of the most abundant microbial genera, their taxonomic ID, lineage, and percentage of total microbially-derived RNA assigned to the species within each PRR phenotypic group and non-inoculated controls

Name	TID	Max (%)	Combined Resistant	Combined Susceptible	Combined Control	Lineage
Fusarium	5506	53.68	41.88%	53.68%	38.55%	Fungi>Dikarya>Ascomycota>saccharomyceta>Pezizomycotina>leotiomyceta>sordariomyceta>Sordariomycetes>Hypocreomycetidae>Hypocreales>Nectriaceae
Pochonia	243023	10.93	4.39%	4.98%	10.93%	Fungi>Dikarya>Ascomycota>saccharomyceta>Pezizomycotina>leotiomyceta>sordariomyceta>Sordariomycetes>Hypocreomycetidae>Hypocreales>Clavicipitaceae
Thermothielavioides	2609811	10.74	5.60%	3.09%	10.74%	Fungi>Dikarya>Ascomycota>saccharomyceta>Pezizomycotina>leotiomyceta>sordariomyceta>Sordariomycetes>Sordariomycetidae>Sordariales>Chaetomiaceae
Colletotrichum	5455	5.597	2.63%	1.61%	5.60%	Fungi>Dikarya>Ascomycota>saccharomyceta>Pezizomycotina>leotiomyceta>sordariomyceta>Sordariomycetes>Hypocreomycetidae>Glomerellales>Glomerellaceae
Aspergillus	5052	5.111	5.11%	2.94%	3.91%	Fungi>Dikarya>Ascomycota>saccharomyceta>Pezizomycotina>leotiomyceta>Eurotiomycetes>Eurotiomycetidae>Eurotiales>Aspergillaceae
Paraburkholderia	1822464	5.041	4.60%	5.04%	1.06%	Bacteria>Proteobacteria>Betaproteobacteria>Burkholderiales>Burkholderiaceae
Nocardioides	1839	3.221	2.05%	3.22%	1.07%	Bacteria>Terrabacteria group>Actinobacteria>Actinomycetia>Propionibacteriales>Nocardioidaceae
Talaromyces	5094	3.016	2.50%	2.12%	3.02%	Fungi>Dikarya>Ascomycota>saccharomyceta>Pezizomycotina>leotiomyceta>Eurotiomycetes>Eurotiomycetidae>Eurotiales>Trichocomaceae
Alicyclobacillus	29330	2.5	2.50%	2.13%	1.04%	cellular organisms>Bacteria>Terrabacteria group>Firmicutes>Bacilli>Bacillales>Alicyclobacillaceae
Neurospora	5140	2.066	1.30%	0.81%	2.07%	cellular Fungi>Dikarya>Ascomycota>saccharomyceta>Pezizomycotina>leotiomyceta>sordariomyceta>Sordariomycetes>Sordariomycetidae>Sordariales>Sordariaceae
Streptomyces	1883	1.896	0.87%	0.88%	1.90%	Bacteria>Terrabacteria group>Actinobacteria>Actinomycetia>Streptomycetales>Streptomycetaceae

Table 3.4 (continued).

Paenibacillus	**44249**	**1.645**	**1.65%**	**1.40%**	**0.51%**	**Bacteria>Terrabacteria group>Firmicutes>Bacilli>Bacillales>Paenibacillaceae**
Dyella	231454	1.394	0.57%	0.83%	1.39%	Bacteria>Proteobacteria>Gammaproteobacteria>Xanthomonadales>Rhodanobacteraceae
Thermothelomyces	1920207	1.33	1.33%	0.41%	0.82%	Fungi>Dikarya>Ascomycota>saccharomyceta>Pezizomycotina>leotiomyceta>sordariomyceta>Sordariomycetes>Sordariomycetidae>Sordariales>Chaetomiaceae
Pantoea	53335	1.024	0.35%	1.02%	0.18%	Bacteria>Proteobacteria>Gammaproteobacteria>Enterobacterales>Erwiniaceae
Halobacteriovorax	1652133	1.007	1.01%	0.43%	0.30%	Bacteria>Proteobacteria>Oligoflexia>Bacteriovoracales>Halobacteriovoraceae
Rhodanobacter	75309	0.9842	0.36%	0.45%	0.98%	Bacteria>Proteobacteria>Gammaproteobacteria>Xanthomonadales>Rhodanobacteraceae
Curtobacterium	2034	0.9764	0.43%	0.98%	0.27%	Bacteria>Terrabacteria group>Actinobacteria>Actinomycetia>Micrococcales>Microbacteriaceae
Botrytis	33196	0.8906	0.63%	0.51%	0.89%	Fungi>Dikarya>Ascomycota>saccharomyceta>Pezizomycotina>leotiomyceta>sordariomyceta>Leotiomycetes>Helotiales>Sclerotiniaceae
Zymoseptoria	1047167	0.8073	0.64%	0.64%	0.81%	Fungi>Dikarya>Ascomycota>saccharomyceta>Pezizomycotina>leotiomyceta>dothideomyceta>Dothideomycetes>Dothideomycetidae>Mycosphaerellales>Mycosphaerellaceae
Burkholderia	32008	0.6883	0.66%	0.69%	0.36%	Bacteria>Proteobacteria>Betaproteobacteria>Burkholderiales>Burkholderiaceae
Pseudomonas	286	0.6718	0.67%	0.35%	0.36%	Bacteria>Proteobacteria>Gammaproteobacteria>Pseudomonadales>Pseudomonadaceae
Azoarcus	12960	0.6216	0.62%	0.20%	0.17%	Bacteria>Proteobacteria>Betaproteobacteria>Rhodocyclales>Zoogloeaceae
Brettanomyces	13366	0.6141	0.61%	0.23%	0.06%	cellular organisms>Eukaryota>Opisthokonta>Fungi>Dikarya>Ascomycota>saccharomyceta>Saccharomycotina>Saccharomycetes>Saccharomycetales>Pichiaceae
Methylobacterium	407	0.5913	0.59%	0.16%	0.22%	cellular organisms>Bacteria>Proteobacteria>Alphaproteobacteria>Rhizobiales>Methylobacteriaceae

Table 3.5: Top 25 Microbial Species by percent total read identified in RNAseq

A list of the most abundant microbial species, their taxonomic ID, and percentage of total microbially-derived RNA assigned to the species within each PRR phenotypic group and non-inoculated controls

Name	**TID**	**Max%**	**Combined Resistant**	**Combined Susceptible**	**Combined Control**
Fusarium oxysporum	5507	22.93	13.58%	22.93%	11.13%
Pochonia chlamydosporia	280754	16.83	7.53%	9.80%	16.83%
Thermothielavioides terrestris	2587410	16.54	9.62%	6.07%	16.54%
Colletotrichum higginsianum	80884	8.621	4.51%	3.17%	8.62%
Fusarium fujikuroi	5127	7.462	7.46%	6.08%	6.49%
Talaromyces rugulosus	121627	4.646	4.30%	4.17%	4.65%
Fusarium verticillioides	117187	3.665	1.92%	3.67%	1.42%
Neurospora crassa	5141	3.182	2.23%	1.59%	3.18%
Aspergillus fumigatus	746128	2.361	2.36%	1.04%	1.51%
Thermothelomyces thermophilus	78579	2.285	2.29%	0.81%	1.26%
Aspergillus oryzae	5062	1.896	1.90%	1.83%	1.42%
Paraburkholderia tropica	92647	1.881	1.02%	1.88%	0.02%
Halobacteriovorax marinus	97084	1.727	1.73%	0.85%	0.45%
Alicyclobacillus sp. SO9	2665646	1.468	1.47%	1.29%	0.54%
Botrytis cinerea	40559	1.372	1.08%	1.00%	1.37%
Zymoseptoria tritici	1047171	1.255	1.09%	1.26%	1.24%
Nocardioides baekrokdamisoli	1804624	1.231	0.73%	1.23%	0.39%
Pantoea dispersa	59814	1.066	0.06%	1.07%	0.04%
Brettanomyces nanus	13502	1.055	1.06%	0.46%	0.09%
Azoarcus olearius	418699	1.055	1.06%	0.39%	0.25%
Candidatus Chazhemtobacterium aquaticus	2715735	0.9736	0.97%	0.36%	0.22%

Table 3.5 (continued).

Rhodanobacter glycinis	**582702**	**0.9704**	**0.34%**	**0.47%**	**0.97%**
Nocardioides sp. zg-1228	2763008	0.9374	0.48%	0.94%	0.05%
Nocardioides cynanchi	2558918	0.9153	0.30%	0.92%	0.11%
Alicyclobacillus acidocaldarius	405212	0.9037	0.90%	0.52%	0.27%

3.7 Figures

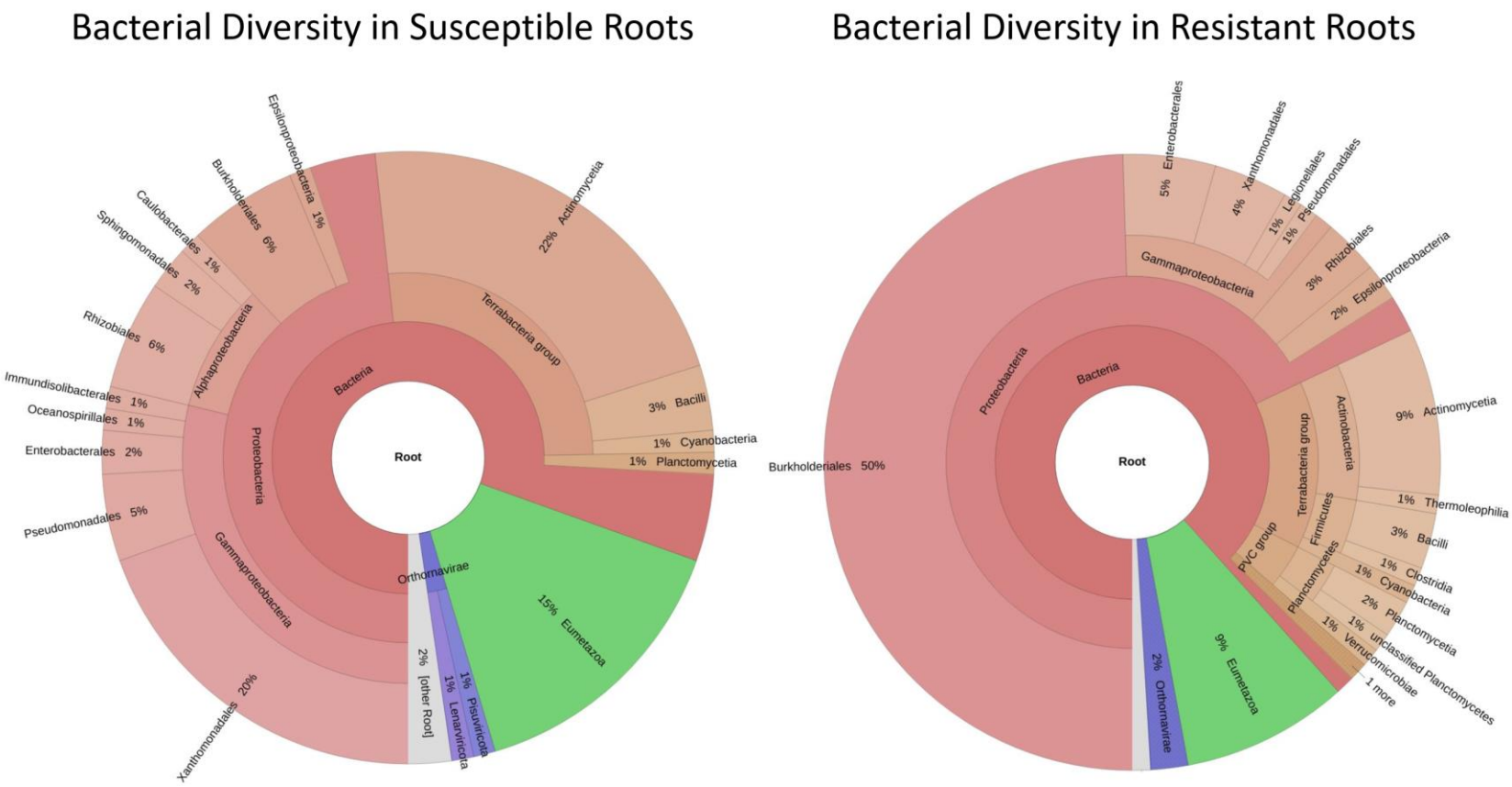


Figure 3.1: Differences in bacterial communities associated with PRR disease response

Krona-generated cladogram of the percentage bacteria-assigned transcripts. Bacterial RNA comprised 11% of the total RNA in PRR resistant root libraries and 16% of the total RNA in the PRR susceptible root libraries.

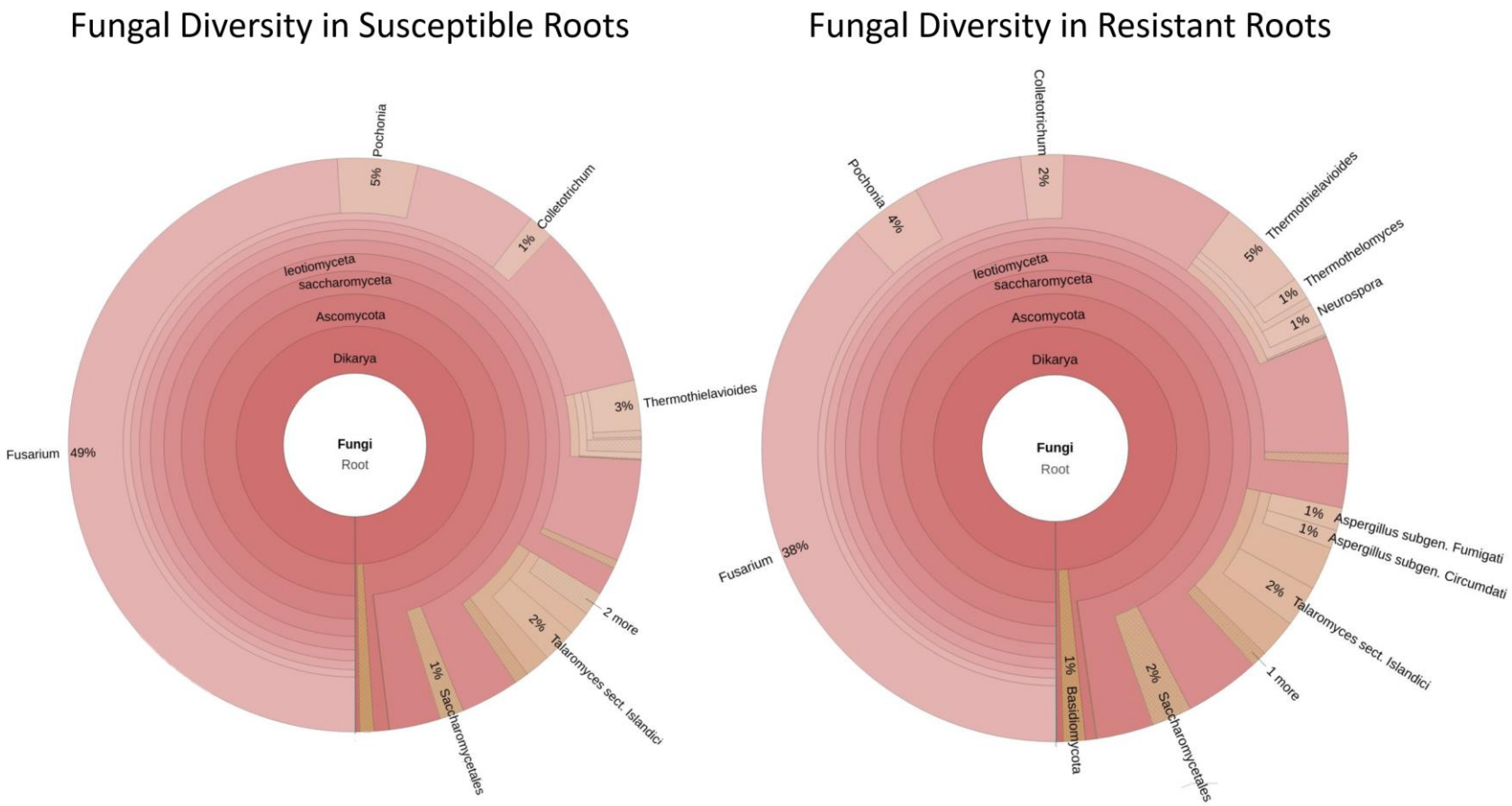


Figure 3.2: Differences in fungal communities associated with PRR disease response

Krona-generated cladogram of the percentage fungal-assigned transcripts. Fungal RNA comprised 13% of the total RNA in PRR resistant root libraries and 21% of the total RNA in the PRR susceptible root libraries.

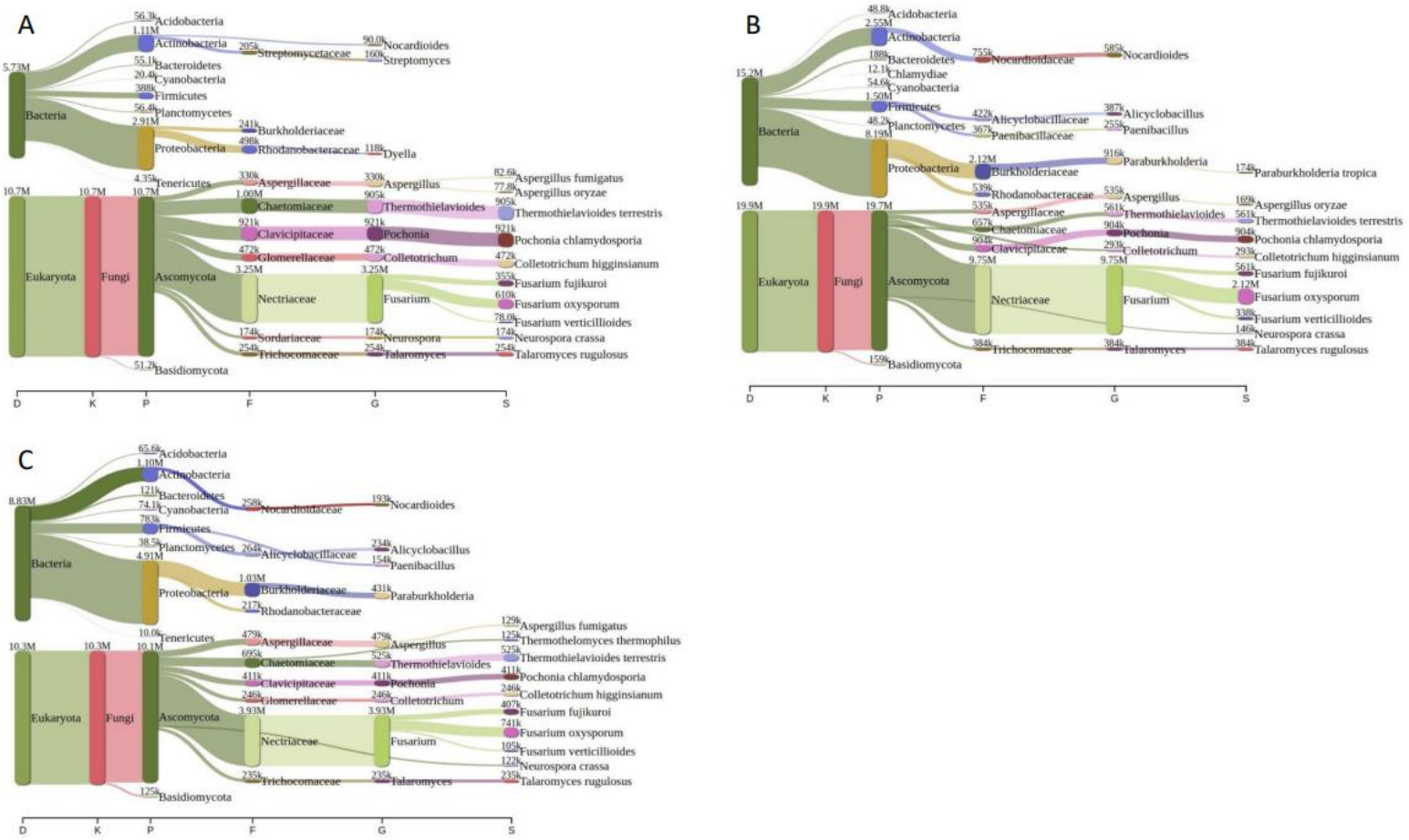


Figure 3.3: Microbial Taxonomic Assignment of Reads by PRR Phenotypic Group

Distribution of RNAseq reads assigned to taxonomic group associated with PRR phenotype.
A: Combined non-inoculated Control (N=2), B: Combined PRR Susceptible (N=4), C: Combined PRR Resistant (N=4).

3.8 Literature cited

Adiconis, Xian, et al. “Comparative Analysis of RNA Sequencing Methods for Degraded or Low-Input Samples.” Nature Methods, vol. 10, no. 7, July 2013, pp. 623–29.

Aviv, H., and P. Leder. “Purification of Biologically Active Globin Messenger RNA by Chromatography on Oligothymidylic Acid-Cellulose.” Proceedings of the National Academy of Sciences, vol. 69, no. 6, June 1972, pp. 1408–12.

“BBMap.” SourceForge, https://sourceforge.net/projects/bbmap/. Accessed May 2019.

Breitwieser, Florian P., and Steven L. Salzberg. Pavian: Interactive Analysis of Metagenomics Data for Microbiomics and Pathogen Identification. preprint, Bioinformatics, 31 Oct. 2016. DOI.org (Crossref), https://doi.org/10.1101/084715.

Chang, Zheng, et al. “Bridger: A New Framework for de Novo Transcriptome Assembly Using RNA-Seq Data.” Genome Biology, vol. 16, no. 1, 2015, p. 30. Crossref, https://doi.org/10.1186/s13059-015-0596-2.

Crusoe, Michael R., et al. “The Khmer Software Package: Enabling Efficient Nucleotide Sequence Analysis.” F1000Research, vol. 4, Sept. 2015. PubMed Central, https://doi.org/10.12688/f1000research.6924.1.

Daniel, Rolf. “The Metagenomics of Soil.” Nature Reviews Microbiology, vol. 3, no. 6, June 2005, pp. 470–78. DOI.org (Crossref), https://doi.org/10.1038/nrmicro1160.

Duponnois, R., Garbaye, J. “Effect of dual inoculation of Douglas fir with the ectomycorrhizal fungus Laccaria laccata and mycorrhization helper bacteria (MHB) in two bare-root forest nurseries.” Plant Soil, vol 138, 1991, pp. 169–176

Frampton, J., Isik, F. & Benson, D.M. “Genetic variation in resistance to Phytophthora cinnamomi in seedlings of two Turkish Abies species.” Tree Genetics & Genomes, vol 9, 2013, pp. 53–63.

Gallego Romero, Irene, et al. “RNA-Seq: Impact of RNA Degradation on Transcript Quantification.” BMC Biology, vol. 12, no. 1, Dec. 2014, p. 42. DOI.org (Crossref), https://doi.org/10.1186/1741-7007-12-42.

Hardham, Adrienne R. “Phytophthora Cinnamomi.” Molecular Plant Pathology, vol. 6, no. 6, Nov. 2005, pp. 589–604. Crossref, https://doi.org/10.1111/j.1364-3703.2005.00308.x.

Huang, Xiaoqiu, and Anup Madan. “CAP3: A DNA Sequence Assembly Program.” Genome Research, vol. 9, no. 9, Sept. 1999, pp. 868–77. genome.cshlp.org, https://doi.org/10.1101/gr.9.9.868.

Inglis, Peter W., et al. “Fast and Inexpensive Protocols for Consistent Extraction of High Quality DNA and RNA from Challenging Plant and Fungal Samples for High-Throughput SNP Genotyping and Sequencing Applications.” PLOS ONE, edited by Ruslan Kalendar, vol. 13, no. 10, Oct. 2018, p. e0206085. DOI.org (Crossref), https://doi.org/10.1371/journal.pone.0206085.

Kohlway, W. H., et al. “Response of Turkish and Trojan Fir to Phytophthora Cinnamomi and P. Cryptogea.” Scandinavian Journal of Forest Research, vol. 32, no. 5, July 2017, pp. 406–11. Crossref, https://doi.org/10.1080/02827581.2017.1280076.

Kunin, Victor, et al. “A Bioinformatician’s Guide to Metagenomics.” Microbiology and Molecular Biology Reviews, vol. 72, no. 4, Dec. 2008, pp. 557–78. DOI.org (Crossref), https://doi.org/10.1128/MMBR.00009-08.

Lagesen, Karin, et al. “RNAmmer: Consistent and Rapid Annotation of Ribosomal RNA Genes.” Nucleic Acids Research, vol. 35, no. 9, May 2007, pp. 3100–08. DOI.org (Crossref), https://doi.org/10.1093/nar/gkm160.

Longmuir, Amy L., et al. “Draft Genomes of Two Australian Strains of the Plant Pathogen, Phytophthora Cinnamomi.” F1000Research, vol. 6, Feb. 2018. PubMed Central, https://doi.org/10.12688/f1000research.12867.2.

Morlan, John D., et al. “Selective Depletion of RRNA Enables Whole Transcriptome Profiling of Archival Fixed Tissue.” PLoS ONE, edited by Soheil S. Dadras, vol. 7, no. 8, Aug. 2012, p. e42882. DOI.org (Crossref), https://doi.org/10.1371/journal.pone.0042882.

Nesme, Joseph, et al. “Back to the Future of Soil Metagenomics.” Frontiers in Microbiology, vol. 7, Feb. 2016. DOI.org (Crossref), https://doi.org/10.3389/fmicb.2016.00073.

Ondov, Brian D, et al. “Interactive Metagenomic Visualization in a Web Browser.” BMC Bioinformatics, vol. 12, no. 1, Dec. 2011, p. 385. DOI.org (Crossref), https://doi.org/10.1186/1471-2105-12-385.

Ondov, Brian D., et al. “Mash: Fast Genome and Metagenome Distance Estimation Using MinHash.” Genome Biology, vol. 17, no. 1, Dec. 2016. Crossref, https://doi.org/10.1186/s13059-016-0997-x.

Richter, Daniel D., and Daniel Markewitz. How Deep Is Soil? no. 9, 1995, p. 11.

Sangwan, Seema, and Radha Prasanna. “Mycorrhizae Helper Bacteria: Unlocking Their Potential as Bioenhancers of Plant–Arbuscular Mycorrhizal Fungal Associations.” Microbial Ecology, vol. 84, no. 1, July 2022, pp. 1–10. DOI.org (Crossref), https://doi.org/10.1007/s00248-021-01831-7.

Schroeder, Andreas, et al. “The RIN: An RNA Integrity Number for Assigning Integrity Values to RNA Measurements.” BMC Molecular Biology, vol. 7, no. 1, 2006, p. 3. DOI.org (Crossref), https://doi.org/10.1186/1471-2199-7-3.

Sitaraman, Ramakrishnan. “Pseudomonas Spp. as Models for Plant-Microbe Interactions.” Frontiers in Plant Science, vol. 6, Sept. 2015. PubMed Central, https://doi.org/10.3389/fpls.2015.00787.

Sukhada, M., Manjula, R., & Rawal, R. D. “Evaluation of arbuscular mycorrhiza and other biocontrol agents against Phytophthora parasitica var. Nicotianae infecting papaya (Carica papaya cv. Surya) and enumeration of pathogen population using immunotechniques. Biological Control, 58(1),2011, p. 22–29. https://doi.org/10.1016/j.biocontrol.2011.03.013

Toju, Hirokazu, and Hirotoshi Sato. “Root-Associated Fungi Shared Between Arbuscular Mycorrhizal and Ectomycorrhizal Conifers in a Temperate Forest.” Frontiers in Microbiology, vol. 9, Mar. 2018, p. 433. DOI.org (Crossref), https://doi.org/10.3389/fmicb.2018.00433.

Vaucheret, Hervé, et al. “Post-Transcriptional Gene Silencing in Plants.” Journal of Cell Science, vol. 114, no. 17, 2001, p. 9.

Wood, Derrick E., et al. Improved Metagenomic Analysis with Kraken 2. preprint, Bioinformatics, 7 Sept. 2019. DOI.org (Crossref), https://doi.org/10.1101/762302.

CHAPTER 4

4.1 Introduction

RNA-Seq has become a quick, cost-effective method for obtaining a wealth of genetic data in non-model organisms that lack a sequenced genome (Wang et al. 2009). Rather than sequence genomic DNA, RNA-seq sequences actively transcribed RNA within the isolated tissue. With RNA-seq, sequence variants within actively transcribed genes can also be identified. However, to identify SNPs and other sequence variants in the RNA-seq samples, deep coverage is necessary, as different transcriptional isoforms derived from alternate splicing of genes can confound analyses. An added benefit of RNA-seq is that the abundance of each transcript is measured for each sample sequenced. With proper sample normalization, relative transcription levels can be compared between the two phenotypic groups. Transcriptional profiles are variable between different tissue types; for instance photosynthetic-foliage has different expression profile than the roots of the same plant. This means that the selection of what tissue is collected is important for detection of genes critical to a trait, such as host defense and great care must be taken in the experimental design (Auer and Doerge, 2010).

Phytophthora cinnamomi Rands is a soilborne pathogen that is primarily spread via water movement, but as the motile zoospore detects root hormones, it uses its flagella to move to actively growing root tip where it encysts and infects the root (Hardham 2005). The little-known-about conifer and other woody root systems molecular response to *Phytophthora* has been reviewed earlier in Chapter 1. In previous studies the Trojan fir family Kaz24 displayed Phytophthora root rot (PRR) resistance rate of 30-50% (Frampton et

al. 2013, Kohlway et al. 2017) to the aggressive isolate of *Phytophthora cinnamomi*, 23ss04, employed by NCSU (Kohlway et al. 2015).

The goal of this chapter is two-fold; first, to establish a reliable and effective Phytophthora root rot resistance screen and second, to compare the transcriptional response of Trojan fir roots to *Phytophthora cinnamomi* between resistant and susceptible individuals to identify PRR resistance or susceptibility genes. To those ends, two different *Phytophthora cinnamomi* inoculation trials using a novel root inoculation and sampling method were performed. In the first, the Trojan fir root transcriptional response to long-term exposure to *P. cinnamomi* was measured and in the second trial, two shorter term exposures were used.

4.2 Materials and Methods

4.2.1 Long-term Trojan fir root exposure to Phytophthora

The full protocol for the long-term exposure group is described in Chapter 3, but the relevant points will be reiterated here. 300 Kaz24 Trojan fir seed were cold-stratified for one month and germinated on vermiculite. Two-week-old, germinated seedlings were then transferred to a greenhouse and into a box of Phytophthora-colonized potting medium.

Each seedling was gently removed from the Phytophthora-ridden media after 8 days of exposure. The roots were rinsed, and 3 cm of root tissue was removed from the root tip. The excised tissue was then transferred to Qiagen deepwell strip-tubes, preloaded with a 3 mm metallic bead, and stored on dry ice while in the greenhouse. The seedling was then transferred to a Cone-tainer tube filled with sterile potting medium. The transferred seedling was then monitored for 16 weeks to phenotype Phytophthora root rot resistance level via percent shoot necrosis or time to mortality.

4.2.2 Short-term Trojan fir root exposure to Phytophthora

The preparation of the Trojan fir seedlings for the short-term exposure study followed much of the same process as the long-term exposure study, albeit within a sterile laboratory setting rather than a greenhouse environment. Two sets of 200 Kaz24 Trojan fir seed were cold-stratified for a month, separated three weeks apart. The cold-stratified seed were sown in vermiculite to germinate over two weeks. In parallel, a 50.8x35.56x15.24 cm plastic bin was filled with 0.014158 m^3 of potting media mixed with *Phytophthora cinnamomi* covered rice grains (Holmes and Benson, 1994) and kept moist to promote full *Phytophthora* growth over the two week incubation. After two weeks a total of 206 seedlings germinated and had grown to a length of 15.24 cm with an unbranched root system. 20 seedlings, five for each replication and exposure length were transferred from vermiculite to sterile potting medium filled Cone-tainers for use as non-exposed controls. The remaining 186 seedlings was transferred to the two bins containing the *Phytophthora* colonized potting media to begin exposure of the roots to the pathogen. 48 hours post exposure, the first batch of seedlings were removed from the media, washed, and 3 cm of the root tip were excised before transferring the seedling to sterile potting medium. The remaining seedlings were incubated for an additional 48 hours in the *Phytophthora* inoculation media for a total exposure time of 96 hours before the root tip harvesting procedure was repeated. PRR disease progression was monitored based on percent foliage and stem necrosis weekly for a total of 16 weeks. A reinoculation of the transferred seedlings with fresh *P. cinnamomi* rice grain inoculum was carried out 8 weeks into the study to ensure proper exposure of each sample to *Phytophthora*. Following the end of the screening a random sample of 50 seedlings were tested with Agdia ImmunoStrips^{tm}.

4.2.2 RNA library preparation and sequencing

The RNA library preparation of the 10 long-term exposure libraries is described in chapter 3.2.3. Briefly, 10 libraries each containing 10 pooled individuals from the same Kaz24 family; 4 resistant, 4 susceptible, and 2 control libraries respectively, were prepared using ribosomal depletion to recover the low-quality RNA (Adiconis et al. 2013). The 10 libraries were then sequenced on a single lane of Hiseq 2500.

The short-term exposure RNA libraries were prepared using oligo-dT to select poly-adenylated mRNA for RNAseq library preparation. A total of 28 RNAseq libraries were prepared by the Genome Sequencing Lab (GSL) at NCSU; Table 4.1 displays the overall library pool guide and average Rin score. Each RNAseq library was a pool of 5 individuals each from the same Kaz24 family, grouped by week 16 disease resistance phenotype, with 400 ng of RNA per individual in the pool. The 28 RNAseq libraries were then sequenced on 7 lanes of Hiseq 2500 with single-end reads. 12 libraries were combined for each of the 7 lanes with 3 technical replicates for each library across the different lanes to control for lane effects.

4.2.3 Trojan fir root transcriptome assembly

Three separate software pipelines were used to assemble the Trojan fir root transcriptomes. The first was described in Chapter 3 using Khmer, BridgeR, and Cap3. The second was with the Trans-ABySS transcriptome assembler (Robertson et al. 2010). The third and final assembly was made using the CLC Genomic Workbench v.12 DeNovo assembler using the default DeNovo pipeline settings. EnTAP rnaQUAST (Bushmanova et

al. 2016) was used to score each assembly. Transrate (Smith-Unna et al. 2016) was used to merge each assembly to one final transcriptome.

4.2.4 Differential Expression Analysis

Differential expression between phenotypic groups was performed using the R package DESeq2 (Love et al. 2014). Contrasts between the phenotypic groups, PRR resistant and susceptible, were made against the controls and each other. Comparisons between each of the two time points (48 and 96 hours) were made, first treating each time point separately, and secondly with the susceptible and resistant samples pooled across both time points. Counts of the number of transcripts that were two- and four-fold up/down-regulated at an adjusted p-value of 0.1 (corrected for multiple testing) were collected for each comparison. Finally, contigs found differently expressed were further characterized for Gene Ontology (GO) terms by screening against the EggNOG (Huerta-Cepas et al. 2016) database using EnTAP (Hart and Wegrzyn 2017). A GO term enrichment network was generated using REVIGO (Supek et al. 2011). In parallel, a non-model organism RNAseq analysis pipeline, Seq2Fun, was used matching against the plant taxonomic group (Liu et al. 2021).

4.3 Results

4.3.1 Phytophthora root rot resistance screen

The results of two Phytophthora root rot resistance screens are displayed in Table 4.2. The 8-day exposure had the highest mortality with an overall mortality of 81% which is higher than the 50-71% mortality rate previously observed of the Kaz24 Trojan fir family (Frampton et al. 2013, Kohlway 2017). The controls all survived, barring a single sample that

was accidentally destroyed during handling. The 2- and 4-day incubation periods had average mortalities of 22.25% and 51.85% respectively. *Phytophthora cinnamomi* was successfully reisolated from a sample of the media from both a resistant and susceptible seedling after 16 weeks. The controls tested negative for *Phytophthora* with Agdia ImmunoStrips^{tm}.

4.3.2 Trojan fir transcriptome assemblies

The summary of each transcriptome assembly is shown in table 4.3. The Trojan fir root transcriptome generated using ribosomal rRNA depleted library preparation has a total of 59,853 contigs for a total of 49,673,231 bp in length with a BUSCO score of 54.1%. The transcriptome from the second RNAseq experiment has 105,296 contigs covering 52,282,480 bp with a BUSCO score of 35.3%. The Transrate merged, final transcriptome has 96,526 contigs for a total length of 61,059,892 bp and a BUSCO score of 70.6%.

4.3.3 Differential expression analysis of Phytophthora infested roots

The results of the differential expression analysis are summarized in table 4.4. In short, a total of 710 differentially expressed (DE) genes were identified between resistant and susceptible groups after 8-day exposure, 37,082 DE genes after 4-day exposure, and 38,154 DE genes after 2-day exposure. When comparing PRR resistant to PRR susceptible phenotypes a total of 212 DE genes after 8-day exposure, 508 DE genes after 4-day exposure, 755 DE genes after 2-day exposure, and 1192 DE genes when the time points were combined by phenotype were identified. The GO term distribution of the differential expressed transcripts is displayed table 4.5 and a GO term enrichment network is displayed in figure 4.1. DE genes between resistant and susceptible phenotypes with EnTAP-assigned

known function are displayed in 2-day (table 4.6), 4-day (table 4.7), and combined time points (table 4.8),

4.4 Discussion

4.4.1 Novel Phytophthora root resistance screen method

Overall, the fir seedling root inoculation method was very successful. The resilience of the seedlings to the excision of the ends of their roots and the ability of the plants to continue growing after sample collection allowed both early-time-point sampling for RNA-seq experiments and 16-week survival analysis of the effects of Phytophthora infestation. The improvement of the experimental conditions in the second, short-term seedling inoculation resulted in higher quality, less fragmented RNA from the roots harvested. The high-quality RNA in the second experiment allowed for the use of standard oligo-d(T) RNAseq library preparation, preventing any bacterial-derived RNA from occupying sequencing space as they did in the rRNA-depleted libraries and improving the yield of informative, Trojan fir-derived RNA.

The 4-day incubation period mortality averages of around 42% were closest to the 50% resistance originally observed for the Kaz24 family (Frampton and Isik 2012). The advanced state of necrosis in the roots of the 8-day long exposed samples, and the poor quality of the RNA isolated from those necrotic roots (Table 3.2), suggests that the choice of incubation length is critical to the quality of sample recovery, and also highlights the need to consider other environmental variables. For instance, with the 8-day screen, the seedlings were exposed to higher mean temperatures in the greenhouse and a heavier water flow. Excessive watering expedites the spread of *Phytophthora* zoospores and hyphae throughout

the potting media which could overwhelm partially PRR resistant or tolerant individuals. Although the shorter two- and four-day exposures may allow for unchallenged individuals to escape, the re-inoculation of the seedlings post-transfer will confirm PRR phenotype.

The reduction of time required for waiting for the seedlings to mature and the associated handling costs is a major advantage of this technique. The most recent survey of Christmas tree farms identified two new *Phytophthora* species from diseased trees (Martin et al. 2017). With the spread of new species and isolates, a rapid screening technique would allow for the screening of numerous genotypes and fir species at an accelerated rate. With additional screening new PRR resistant fir materials could be identified and brought into breeding programs or used as root stock for grafts. The previously relied upon PRR tolerance of Canaan fir for use as root stock has been proven to be unsustainable as the warmer winters allow for *Phytophthora* to spread longer into the season and accumulate to levels that overwhelm the fast-growing root system of Canaan fir (*Abies balsamea* var. *phanerolepis*). The best choice for future plantings would be to use fully *Phytophthora* resistant exotic fir species. As new varieties of Turkish and Trojan fir are evaluated for use as Christmas trees (Kurt et al. 2016) it will be paramount for their success to screen the aesthetically-marketable families for PRR resistance.

This new method is similar in sample generation time to the stem inoculation technique (Barret et al. 2003). Stem inoculation is another PRR resistance screen method which involves wounding the plant host's above soil stem and place *Phytophthora* inoculum directly onto the wound. The progression of the *Phytophthora* caused lesion is then measured a few days post-inoculation and the stem tissue can be collected for expression analysis. However, stem wounding still requires older plant material as the inoculation site needs to be

large enough to support the mechanical damage. The location of the inoculation is also potentially problematic. In nature, *Phytophthora cinnamomi* attacks the actively growing regions of the fir root system (Hardham 2005). As much of the fir and *Phytophthora* host-pathogen interaction is unknown, there may be different mechanisms of resistance between the different tissue types.

4.3.2 Trojan fir root transcriptome assemblies

The separate transcriptomes assembled from each of the individual studies merged into a combined final assembly helped to refine the Kaz24 root transcriptome to a quality level that is in line and even surpasses other conifer transcriptomes. The final Kaz24 transcriptome with a total length of 61Gb, 96,526 contigs and a BUSCO score of 70.6% represents a major improvement over the two separate assemblies that scored 54.1% and 35.3%. BUSCO scores are useful for gauging the completeness of a transcriptome assembly (Simão et al. 2015). The suite of evolutionarily conserved single-copy gene orthologs in the plant database used for BUSCO analysis works well for non-model organisms like fir, where little previous genomic data is available.

In a recent study of two conifers, *Juniperus flaccida* and *Pinus cembroides,* novel transcriptomes were generated using from a single tissue type (Webster er al. 2022). The BUSCO scores were lower with 36.3% and 30.1% for *J. flaccida* and *P. cembroides* respectively with a similar range of total transcripts (60,000-159,000 contigs). The transcriptome of Norway spruce (*Picea abies*) has also been extensively researched (Reza et al. 2018; Bag et al. 2021). The currently available transcriptome contains 70,736 genes and

was developed from multiple tissue types. These results are similar to the transcriptomes generated here prior to the refinement gained by merging the assemblies.

4.3.3 Differential expression after long-term exposure

A total of 6298 differentially expressed (DE) genes were identified between resistant and susceptible groups. Within the differential expressed genes SAUR72 (small auxin upregulated RNA), and HSP90-1 (heat-shock protein) were significantly upregulated in susceptible trees, and five OMP16 (outer membrane lipoprotein) homologs (ranging from 4.7 to 9.8-fold) greatly upregulated in resistant trees. A total of 1545 genes were differentially expressed when comparing resistant seedlings to the uninoculated control groups (1219 upregulated in control, 326 in resistant) and the comparison between susceptible and control groups yielded 4837 DE genes (115 upregulated in control, 4722 in susceptible). Overall, there was an observed increased expression of xanthine and alcohol dehydrogenases in resistant trees relative to the control group. A total of 38 annotated genes were downregulated in susceptible trees and upregulated in resistant trees relative to control expression levels; these included YSl5 (YELLOW STRIPE like 5), cspA (cold-shock protein A) and several ribosomal proteins. Findings also included significant upregulation of apoptosis related genes in resistant plants across comparisons, including XCP1, a cysteine peptidase which is involved in tracheary element autolysis.

As noted before, the long-term exposure turned out to be much too long, as most of the samples had already been necrotized by the time the RNA was extracted. Therefore, it is not surprising that there was an abundance of cell death related expression pathways in the PRR susceptible groups.

4.3.4 Differential expression after short-medium term exposure

Whereas the 8-day incubation period captured the transcriptional profiles of roots in the aftermath of *Phytophthora* infection the PCA shown in figure 4.2 PC1 separates the treated samples from the untreated controls, and PC2 separates the two incubation periods. From this analysis, the impact of incubation length is extremely important during the root response to *P. cinnamomi*. Within each treatment group, there is a division between resistant and susceptible phenotypes. The transplantation of seedlings into *Phytophthora* colonized medium versus sterile medium had a much larger impact on gene expression than PRR response as the differential expression profiles reflect. On average there are 19,899 differential expressed transcripts in the comparison to controls while treated samples had an average of 818 differential expressed transcripts. Due to this, it is likely better to look between the resistant versus susceptible treatment groups for genes responsible for survival against *Phytophthora*.

The Trojan fir roots exposed at the 2-day *Phytophthora* exposure time point had larger magnitude of differential expression with a max of 3-to 7-fold-change relative to the much smaller magnitude of around 2-3 fold change in the 4-day exposure. Additionally, there were more total differentially expressed transcripts when comparing PRR resistant versus PRR susceptible samples at 2-days (755) than at 4-days (508).

Comparing the PRR resistant roots versus PRR susceptible roots after the short, 2-day exposure genes, such as Very-long-chain 3-oxoacyl-CoA reductase 1 (LFC 7.47), Methylglutaconyl-CoA hydratase (LFC 7.27), and Dual specificity protein phosphatase 1B (LFC 6.19) are notably upregulated. Very-long-chain 3-oxoacyl-CoA reductase 1 and Methylglutaconyl-CoA hydratase are of note due to their connection with the lignin

biosynthesis pathway (Xie et al. 2018). Dual specificity protein phosphatase 1B (LFC 6.19) is also of note as it regulates Mitogen-activated protein kinase (MAPK) which regulates pathogenesis related (PR) genes (Vileal et al. 2010). Serine/threonine-protein kinase AFC2 (LFC 5.52) is also linked to MAPK regulation in response to pathogen infection (Cross et al. 2000).

Of note in the differentially expressed genes in the 4-day PRR resistant versus susceptible comparison was an upregulation of laccase (LFC 0.584). In conifers and other higher plants laccases are involved in lignification of xylem tissues (LaFayette et al. 1999). Naringenin,2-oxoglutarate 3-dioxygenase (LFC 1.65) is also an interesting upregulated gene as it is part of the flavonoid synthesis pathway (Besseau et al. 2007). Cytosolic Sulfotransferase 12 (LFC 1.27) is also linked to the flavonoid synthesis pathway (Hirschmann et al. 2014).

An interesting transcript downregulated in resistant phenotypes is SWEET1a (-4.8 log fold), a member of SWEET (Sugars Will Eventually be Exported Transporters) sugar transporters. Increased SWEET1a and other SWEET sugar transporter expression has been linked with pathogen susceptibility (Chen 2016). Plant pathogens can induce the expression of SWEET sugar transporters to increase the amount of sugar expelled from the phloem. Both biotrophic bacteria and fungi have been proven to elicit SWEET sugar transporter expression (Chen *et al.* 2010) and *Phytophthora parasitica* induces the expression of similar sugar transporters in tomatoes (García-rodríguez et al. 2005).

4.3.5 Trends in Trojan fir root response to Phytophthora cinnamomi

The lignin biosynthesis pathway had upregulated transcripts differentially expressed in the PRR resistant seedlings. The lignin biosynthesis pathway contains many different effects including lignification of cell walls and the initiation of pathogen response (Xie et al. 2018). In a recent in histopathological study of *Phytophthora cinnamomi* infection of *Castanea* species (Fernandes et al. 2021), the growth of *Phytophthora* hyphae was inhibited within the root cortex by a thicker, more densely-lignified cell wall in the resistant *C. crenata* relative to the susceptible *C. sativa*.

The flavonoid biosynthesis pathway also had a few upgregulated genes in PRR resistant roots. Flavonoids are plant secondary metabolites that have a large range of biological roles including protection against plant pathogens (Falcone et al. 2012). The flavonoids Coumestrol, biochanin A, genistein, naringenin and isorhamnetin have even been reported to inhibit *Phytophthora sojae* growth (Rivera-Vargas et al. 1993).

4.5 Conclusions and Future Work

The results of this study are very promising. First, the confirmation of a new *Phytophthora* inoculation method will be useful for future fir PRR screening projects. The reduction in time and cost for the plant preparation will enable the rapid screening of future cohorts. Previous fir PRR resistance screenings used 4- to 5-year-old seedlings, which is a significant investment of time and budget to generate and maintain.

The robust Trojan fir root transcriptome represents a great tool for future genomic analyses. The transcriptome can be used to make probes for exome capture, reducing the cost of genotyping the large, complex conifer genome (Rellstab et al. 2018). RNAseq of different

Trojan fir tissues during *Phytophthora* exposure would be an interesting future project as well, however many of the transcripts have no known function which would make parsing the genomic information difficult.

The differentially expressed genes identified in this study represent the first major steps into understanding the complex interaction of Trojan fir and *Phytophthora cinnamomi*. The PRR resistant upregulated genes in the lignin and flavonoid biosynthesis pathways along with the upregulation of PR-regulatory genes like Dual specificity protein phosphatase 1B and Serine/threonine-protein kinase AFC2 point to a combinatorial response of physical barrier to *Phytophthora cinnamomi* penetration, and the activation of PR genes specific to *P. cinnamomi*. Those PR genes may be found in the non-annotated differentially expressed transcripts that made up the majority of the differentially expressed genes. The next step would be to use quantitative PCR to confirm these expressional trends in other resistant lines of Turkish and Trojan fir families. A further step would be to do functional validation of the of the biological role of these differentially expressed gene.

Additionally, seed from Kaz24 have been transformed into somatic embryogenic cultures (unpublished, Matallana). Somatic embryogenic (SE) lines represent an essential tool for genetic manipulation using different techniques including CRISPR/Cas9 which has found success in the woody poplar trees (An et al. 2020). One potential target could be the SWEET1a transcript, a sugar transporter that is greatly upregulated in the PRR susceptible individuals and may be linked with susceptibility to *Phytophthora*. SWEET1a knockdown may indeed stymie the growth of *Phytophthora* long enough for other plant defense system to respond.

Additionally, once SE lines have been fully established, the clonal propagation of materials will be useful for reducing the variance of the gene expression that pooling half-sibs puts into the study. With the clonal refinement, the genes responsible for PRR resistance in Trojan fir would be easier to identify from the background gene expression.

4.6 Tables

Table 4.1: Short-term RNA Library Pools and Average Quality

The number and average RNA integrity number (RIN) of pooled RNAseq libraries per PRR phenotypic group and non-inoculated controls. Each pool contained five root samples.

Library	# of Pools	Average Rin
48 hr Deceased	4	7.48
48 hr Alive	10	7.93
96 hr Deceased	6	6.05
96 hr Alive	4	7.15
48 hr Control	2	6.80
96 hr Control	2	7.65
Total	28	7.24

Table 4.2: Final mortality counts after 16 weeks post exposure for each study

Summary of the results for each of Phytophthora root rot resistance screens with final count of living and dead Trojan fir seedlings per *Phytophthora* exposure length.

Inoculation	Exposure Time (days)	Survived	Died	Mortality	Total
Study 1	8	35	145	0.806	180
Study 2 Rep 1	2	52	18	0.257	70
	4	22	31	0.585	53
	Total	**74**	**49**	**0.39**	**123**
Study 2 Rep 2	2	26	6	0.188	32
	4	17	14	0.452	31
	Total	**43**	**20**	**0.318**	**63**

Table 4.3: Trojan fir root transcriptome assembly

Summary of the Trojan fir root transcriptome for each of the independent RNAseq studies and the final merged transcriptome.

Trojan fir Root Transcriptome Assembly	1st Study Transcriptome	2nd Study Transcriptome	Final Transcriptome
Average Length	398bp	496bp	633 bp
Max Length	20,258bp	11,965	20,258 bp
BUSCO Score	54.1%	35.3%	70.6 %
Total Transcript Count	59,854 contigs	105,298 contigs	96,526 contigs
Total Length	49,673,231 bp	52,282,480bp	61,059,892 bp

Table 4.4: Results of Differential Expression Analysis

Summary of the results of the DEseq2 differential expression analysis by pairwise comparison of PRR phenotypic category. SeD: Dead after 48hr *Phytophthora* exposure, SeS: Survived after 48hr exposure, MeD: Dead after 96hr exposure, MeS: Survived after 96hr exposure, LD: Dead after 8-day exposure LS: Survived after 8-day exposure. Con: non-inoculated control group.

Comparison	4-fold Down	2-fold Down	2-fold Up	4-fold Up	Total
SeD vs Con	240	666	2087	12854	18497
SeS vs Con	188	830	2458	11990	18902
SeD vs SeS	24	6	188	465	755
MeD vs Con	181	433	2368	15647	21706
MeS vs Con	179	846	1855	9656	14868
MeD vs MeS	6	11	136	33	508
CD vs Con	68	404	2580	18878	25319
CS vs Con	42	684	2483	13574	20106
CD vs CS	6	8	374	122	1192
LD vs Con	63	34	20	124	246
LS vs Con	158	72	4	15	252
LD vs LS	38	19	49	103	212

Table 4.5: GO Term Enrichment

Summary of the counts of differentially expressed transcripts assigned to Gene Ontology terms by combined PRR phenotype.

Gene Ontology Term	**Comb_Sus vs Con (%total)**		**Comb_Res vs Con (%total)**		**Comb_Res vs Comb_Sus (%total)**	
GO:0090304-nucleic acid metabolic process(L=4)	3318	(13.10)	2590	(12.88)	172	(14.43)
GO:0006796-phosphate-containing compound metabolic process(L=4)	3095	(12.22)	2361	(11.74)	157	(13.17)
GO:0009059-macromolecule biosynthetic process(L=4)	2975	(11.75)	2288	(11.38)	133	(11.15)
GO:0010467-gene expression(L=4)	2962	(11.70)	2286	(11.37)	128	(10.74)
GO:0044267-cellular protein metabolic process(L=4)	2936	(11.60)	2275	(11.32)	126	(10.57)
GO:0034645-cellular macromolecule biosynthetic process(L=4)	2902	(11.46)	2264	(11.26)	126	(10.57)
GO:0031323-regulation of cellular metabolic process(L=4)	2868	(11.33)	2154	(10.71)	125	(10.49)
GO:0080090-regulation of primary metabolic process(L=4)	2854	(11.27)	2146	(10.67)	124	(10.40)
GO:0060255-regulation of macromolecule metabolic process(L=4)	2620	(10.35)	1966	(9.78)	109	(9.14)
GO:0016070-RNA metabolic process(L=4)	2541	(10.04)	1953	(9.71)	105	(8.81)

Table 4.6: Resistant vs susceptible after two-day exposure differentially expressed transcripts with known function

List of differentially expressed transcripts with EnTAP assigned function between PRR resistant and PRR susceptible samples following two-day *Phytophthora* exposure. Transcripts are sorted by descending logarithmic fold change (LFC) with their associated p-value corrected for multiple testing.

Query Sequence	Percent Identical	E Value	Description	Species	LFC	P-adj
Contig75915	30.5	$3.40E^{-20}$	sp|Q8L9C4|KCR1_ARATH Very-long-chain 3-oxoacyl-CoA reductase 1	*Arabidopsis thaliana*	7.47	0.0247
Contig77993	51.3	$2.70E^{-159}$	sp|Q29504|UBA1_RABIT Ubiquitin-like modifier-activating enzyme 1	*Oryctolagus cuniculus*	7.37	0.00267
Contig61979	66.7	$4.60E^{-83}$	sp|Q54HG7|AUHM_DICDI Methylglutaconyl-CoA hydratase, mitochondrial	*Dictyostelium discoideum*	7.27	0.0385
Contig61891	36.3	$1.20E^{-52}$	sp|Q75WF2|DNA2_ACAPL Plancitoxin-1	*Acanthaster planci*	7.22	0.0843
Contig64991	30	$1.60E^{-19}$	sp|P59941|SIR6_MOUSE NAD-dependent protein deacetylase sirtuin-6	*Mus musculus*	7.11	0.0593
Contig52862	35.2	$9.40E^{-75}$	sp|Q54C16|SGMB_DICDI Sphingomyelin phosphodiesterase B	*Dictyostelium discoideum*	6.99	0.0112
Contig76632	45	$2.60E^{-74}$	sp|Q53FA7|QORX_HUMAN Quinone oxidoreductase PIG3	*Homo sapiens*	6.97	0.0277
Contig5685	34.8	$5.50E^{-22}$	sp|Q9M8K7|DUS1B_ARATH Dual specificity protein phosphatase 1B	*Arabidopsis thaliana*	6.94	3.22E-06
Contig79363	26.7	$2.90E^{-14}$	sp|Q58588|Y1187_METJA Uncharacterized protein MJ1187	*Methanocaldococcus jannaschi*	6.8	0.0259
Contig18665	58	$1.30E^{-87}$	sp|Q54KB7|DHE3_DICDI Glutamate dehydrogenase, mitochondrial	*Dictyostelium discoideum*	6.64	0.00788
Contig52785	31.2	$5.70E^{-41}$	sp|Q54JH4|P2XE_DICDI P2X receptor E	*dictyostelium discoideum*	6.6	0.00267
Contig34149	48.4	$3.20E^{-79}$	sp|Q8T293|DIAC2_DICDI Probable di-N-acetylchitobiase 2	*Dictyostelium discoideum*	6.56	0.0277
Contig60403	60.4	$3.30E^{-133}$	sp|Q9DC48|PRP17_MOUSE Pre-mRNA-processing factor 17	*Mus musculus*	6.56	0.0478
Contig75499	36.2	$1.60E^{-167}$	sp|P98200|AT8A2_MOUSE Phospholipid-transporting ATPase IB	*Mus musculus*	6.36	0.0277
Contig75802	58.5	$3.50E^{-112}$	sp|Q54PQ4|GEFA_DICDI Ras guanine nucleotide exchange factor A	*Dictyostelium discoideum*	6.36	0.0518

Table 4.6 (continued).

Contig54903	23.8	$3.90E^{-13}$	sp|Q8BWD2|IP6K3_MOUSE Inositol hexakisphosphate kinase 3	*Mus musculus*	6.27	0.0826
Contig77499	31.1	$2.30E^{-44}$	sp|Q9M884|MPI1_ARATH Mannose-6-phosphate isomerase 1	*Arabidopsis thaliana*	6.22	0.0095
Contig29539	68	$3.30E^{-89}$	sp|Q54DW2|ERGI3_DICDI Probable endoplasmic reticulum-Golgi intermediate compartment protein 3	*Dictyostelium discoideum*	6.19	0.00553
Contig64873	32.6	$2.30E^{-69}$	sp|A2BFP5|S12A9_DANRE Solute carrier family 12 member 9	*Danio rerio*	6.14	0.052
Contig75625	61.3	$1.90E^{-130}$	sp|Q54MZ4|MCFB_DICDI Mitochondrial substrate carrier family protein B	*Dictyostelium discoideum*	6.04	0.0173
Contig75990	32.5	$1.60E^{-56}$	sp|P42892|ECE1_HUMAN Endothelin-converting enzyme 1	*Homo sapiens*	6.04	0.0234
Contig75504	61.2	$3.20E^{-95}$	sp|Q550A8|ATG3_DICDI Autophagy-related protein 3	*Dictyostelium discoideum*	6.03	0.00197
Contig69559	54.9	$1.50E^{-75}$	sp|Q9LTV6|DECR2_ARATH Peroxisomal 2,4-dienoyl-CoA reductase	*Arabidopsis thaliana*	5.96	0.0311
Contig77969	31.6	$9.40E^{-43}$	sp|Q9SKI4|RFA1A_ARATH Replication protein A 70 kDa DNA-binding subunit A	*Arabidopsis thaliana*	5.83	0.0424
Contig62814	70.5	$2.20E^{-63}$	sp|Q17QG5|UB2G2_BOVIN Ubiquitin-conjugating enzyme E2 G2	*Bos taurus*	5.81	0.0747
Contig54904	27	$2.40E^{-22}$	sp|Q96PC2|IP6K3_HUMAN Inositol hexakisphosphate kinase 3	*Homo sapiens*	5.79	0.0169
Contig76557	59.8	$2.60E^{-100}$	sp|Q5V9F0|AMPL_DICDI Cytosol aminopeptidase	*Dictyostelium discoideum*	5.73	0.0553
Contig75619	29.7	$1.40E^{-18}$	sp|P28562|DUS1_HUMAN Dual specificity protein phosphatase 1	*Homo sapiens*	5.54	0.0503
Contig75958	54.6	$2.00E^{-98}$	sp|P51567|AFC2_ARATH Serine/threonine-protein kinase AFC2	*Arabidopsis thaliana*	5.52	0.0878
Contig75380	52.9	$1.40E^{-68}$	sp|P26368|U2AF2_HUMAN Splicing factor U2AF 65 kDa subunit	*Homo sapiens*	5.51	0.0181
Contig75419	22.7	$7.50E^{-33}$	sp|Q9Y2G0|EFR3B_HUMAN Protein EFR3 homolog B	*Homo sapiens*	5.44	0.0423
Contig75232	49.3	$6.50E^{-152}$	sp|Q54BM8|U652_DICDI UPF0652 protein	*Dictyostelium discoideum*	5.42	0.0252
Contig53129	73.3	$3.30E^{-121}$	sp|Q86AD7|MYLKB_DICDI Probable myosin light chain kinase DDB_G0271550	*Dictyostelium discoideum*	5.31	0.00588

Table 4.6 (continued).

Contig59820	81.6	$2.70E^{-41}$	sp|Q94694|RAP1_PHYPO Ras-related protein Rap-1	*Physarum polycephalum*	5.25	0.00489
Contig102383	41.7	$1.40E^{-15}$	sp|Q54P23|SDF2_DICDI Stromal cell-derived factor 2-like protein	*Dictyostelium discoideum*	5.23	0.000145
Contig58497	38.7	$2.90E^{-24}$	sp|Q54IX6|STX8B_DICDI Probable syntaxin-8B	*Dictyostelium discoideum*	5.23	0.0777
Contig59069	49.8	$1.60E^{-75}$	sp|Q55GV3|PAKC_DICDI Serine/threonine-protein kinase pakC	*Dictyostelium discoideum*	5.21	0.0268
Contig55600	47	$2.60E^{-76}$	sp|Q9AV81|PRP19_ORYSJ Pre-mRNA-processing factor 19	*Oryza sativa subsp. japonica*	5.1	0.0252
Contig56434	61.8	$1.60E^{-96}$	sp|Q54MB4|ASNS_DICDI Probable asparagine synthetase [glutamine-hydrolyzing]	*Dictyostelium discoideum*	5.03	0.0843
Contig65950	43	$5.90E^{-66}$	sp|Q94AQ6|SIR4_ARATH NAD-dependent protein deacylase SRT2	*Arabidopsis thaliana*	5.02	0.00434
Contig77856	34.3	$2.30E^{-27}$	sp|B5BLW5|ARE_SACSO Arylesterase	*Saccharolobus solfataricus*	4.99	0.0554
Contig84718	28.6	$4.20E^{-18}$	sp|Q5SRT8|CCNJL_MOUSE Cyclin-J-like protein	*Mus musculus*	4.92	0.000647
Contig75194	66.4	$9.50E^{-179}$	sp|Q86A85|PUR1_DICDI Amidophosphoribosyltransferase	*Dictyostelium discoideum*	4.92	0.0837
Contig7137	37.6	$1.10E^{-21}$	sp|O17433|1CPX_DIRIM 1-Cys peroxiredoxin	*Dirofilaria immitis*	4.89	0.0925
Contig75656	70	$5.00E^{-174}$	sp|Q9VKD3|NFS1_DROME Cysteine desulfurase, mitochondrial	*Drosophila melanogaster*	4.7	0.0572
Contig57675	58.7	$3.80E^{-60}$	sp|P08239|GNAO_BOVIN Guanine nucleotide-binding protein G(o) subunit alpha	*Bos taurus*	4.65	0.0361
Contig1461	45.7	$2.60E^{-207}$	sp|Q869Q3|NAPA_DICDI Nck-associated protein 1 homolog	*Dictyostelium discoideum*	4.63	0.0165
Contig1932	47.7	$2.50E^{-19}$	sp|Q9D0P0|EBPL_MOUSE Emopamil-binding protein-like	*Mus musculus*	4.41	0.0985
Contig99746	71.9	$4.40E^{-41}$	sp|P05714|RAB4A_RAT Ras-related protein Rab-4A	*Rattus norvegicus*	4.4	0.0341
Contig76094	43.2	$1.80E^{-37}$	sp|Q05975|RAB2_LYMST Ras-related protein Rab-2	*Lymnaea stagnalis*	4.4	0.0417
Contig73111	46	$1.80E^{-87}$	sp|Q8H116|MNS2_ARATH Mannosyl-oligosaccharide 1,2-alpha-mannosidase MNS2	*Arabidopsis thaliana*	4.39	0.00193
Contig58197	55.4	$2.40E^{-50}$	sp|P42322|CANB1_NAEGR Calcineurin subunit B	*Naegleria gruberi*	4.34	0.0237

Table 4.6 (continued).

Contig66216	37.1	$1.40E^{-51}$	sp|Q54KZ8|EIF3M_DICDI Eukaryotic translation initiation factor 3 subunit M	*Dictyostelium discoideum*	4.27	0.00491
Contig57549	25.9	$6.00E^{-12}$	sp|Q7REH6|GST_PLAYO Glutathione S-transferase	*Plasmodium yoelii yoelii*	4.22	0.0967
Contig53498	47.2	$5.10E^{-44}$	sp|Q9C6B3|GCA2_ARATH Gamma carbonic anhydrase 2, mitochondrial	*Arabidopsis thaliana*	4.16	0.00588
Contig91429	43.8	$1.00E^{-21}$	sp|Q9SHU5|ARF4_ARATH Probable ADP-ribosylation factor At2g15310	*Arabidopsis thaliana*	3.93	0.0339
Contig46531	65.4	$2.30E^{-55}$	sp|P34149|RACC_DICDI Rho-related protein racC	*Dictyostelium discoideum*	3.91	2.90E-05
Contig42958	43	$1.40E^{-89}$	sp|Q19013|GLS2_CAEEL Putative glutaminase 2	*Caenorhabditis elegans*	3.78	0.0169
Contig61816	46.9	$4.30E^{-19}$	sp|Q94518|NACA_DROME Nascent polypeptide-associated complex subunit alpha	*Drosophila melanogaster*	3.68	0.00588
Contig67017	66	$5.60E^{-64}$	sp|F4KGQ0|ALFC4_ARATH Fructose-bisphosphate aldolase 4, cytosolic	*Arabidopsis thaliana*	3.64	0.00473
Contig57526	38.5	$6.90E^{-42}$	sp|Q4V9P6|ADAL_DANRE Adenosine deaminase-like protein	*Danio rerio*	3.47	0.00842
Contig79283	23.8	$2.20E^{-19}$	sp|Q84JB8|RUS3_ARATH Protein root UVB sensitive 3	*Arabidopsis thaliana*	3.39	0.0748
Contig75171	61.8	$1.60E^{-151}$	sp|P11141|HSP7F_CAEEL Heat shock 70 kDa protein F, mitochondrial	*Caenorhabditis elegans*	3.36	0.0475
Contig60092	73.1	$1.40E^{-155}$	sp|Q869Z4|PCKGM_DICDI Phosphoenolpyruvate carboxykinase [GTP], mitochondrial	*Dictyostelium discoideum*	3.24	0.0189
Contig78485	71.2	$1.80E^{-125}$	sp|Q641W4|RFC2_RAT Replication factor C subunit 2	*Rattus norvegicus*	3.24	0.0172
Contig58440	39	$4.80E^{-42}$	sp|Q86HD3|MRT4_DICDI Ribosome assembly factor mrt4	*Dictyostelium discoideum*	3.2	0.0169
Contig103736	65.8	$2.90E^{-124}$	sp|P11141|HSP7F_CAEEL Heat shock 70 kDa protein F, mitochondrial	*Caenorhabditis elegans*	3.14	0.00238
Contig82107	74.4	$6.60E^{-66}$	sp|Q54X16|GID8_DICDI Glucose-induced degradation protein 8 homolog	*Dictyostelium discoideum*	3.13	0.04
Contig63371	50	$7.10E^{-66}$	sp|Q2KJJ0|PRP6_BOVIN Pre-mRNA-processing factor 6	*Bos taurus*	3.13	0.0321
Contig56093	56.9	$3.70E^{-87}$	sp|Q86L14|PUR9_DICDI Bifunctional purine biosynthesis protein purH	*Dictyostelium discoideum*	1.73	0.038

Table 4.6 (continued).

Contig49886	31.5	$4.50E^{-43}$	sp|Q9DCN7|RNFT1_MOUSE E3 ubiquitin-protein ligase RNFT1	*Mus musculus*	1.58	0.0615
Contig94826	62.2	$1.20E^{-154}$	sp|Q54JP5|OAT_DICDI Probable ornithine aminotransferase	*Dictyostelium discoideum*	1.52	0.0368
Contig68478	49	$1.70E^{-51}$	sp|Q54Z01|TBCB_DICDI Tubulin-specific chaperone B	*Dictyostelium discoideum*	1.21	0.0978
Contig82800	52.6	$8.10E^{-129}$	sp|Q8W471|AAE15_ARATH Long-chain-fatty-acid--[acyl-carrier-protein] ligase AEE15, chloroplastic	*Arabidopsis thaliana*	0.567	0.0809
Contig70611	54	$1.70E^{-135}$	sp|Q84WB7|GINT1_ARATH Glucosamine inositolphosphorylceramide transferase 1	*Arabidopsis thaliana*	0.549	0.0753
Contig61935	32.6	$8.40E^{-66}$	sp|P93479|RETO_PAPSO Reticuline oxidase	*Papaver somniferum*	-1.1	0.0942
Contig146	66.6	$3.10e^{-312}$	sp|Q6ESI7|TPPII_ORYSJ Tripeptidyl-peptidase 2	*Oryza sativa subsp. japonica*	-1.2	0.0985
Contig74658	31.8	$4.10E^{-47}$	sp|P21836|ACES_MOUSE Acetylcholinesterase	*Mus musculus*	-5.91	0.0212

Table 4.7: Resistant vs Susceptible after Four-day exposure differentially expressed transcripts with known function

List of differentially expressed transcripts with EnTAP assigned function between PRR resistant and PRR susceptible samples following four-day *Phytophthora* exposure. Transcripts are sorted by descending logarithmic fold change (LFC) with their associated p-value corrected for multiple testing.

Query Sequence	Percent Identical	E Value	Description	Species	LFC	Padj
Contig65760	96.4	$1.00E^{-69}$	sp|P26354|TBP_ACACA TATA-box-binding protein	*Acanthamoeba castellanii*	2.09	0.00946
Contig34047	50.4	$1.10E^{-108}$	sp|Q5XGS8|GTPB1_XENLA GTP-binding protein 1	*Xenopus laevis*	2.04	0.0634
Contig95600	67.2	$6.40E^{-118}$	sp|P80317|TCPZ_MOUSE T-complex protein 1 subunit zeta	*Mus musculus*	2.01	0.00472
Contig74988	62.9	$6.90E^{-64}$	sp|Q9FE06|EXOL2_ARATH Protein EXORDIUM-like 2	*Arabidopsis thaliana*	1.68	0.0136
Contig68907	46.3	$7.50E^{-73}$	sp|Q05964|FL3H_DIACA Naringenin,2-oxoglutarate 3-dioxygenase	*Dianthus caryophyllus*	1.65	0.077
Contig62088	66.2	$6.20E^{-93}$	sp|Q9LPU2|URGT2_ARATH UDP-rhamnose/UDP-galactose transporter 2	*Arabidopsis thaliana*	1.5	0.0194
Contig91671	27.6	$2.50E^{-19}$	sp|Q58087|Y674_METJA Uncharacterized protein MJ0674	*Methanocaldococcus jannaschii*	1.5	0.0702
Contig74361	27.2	$7.10E^{-16}$	sp|Q9M202|ZAT9_ARATH Zinc finger protein ZAT9	*Arabidopsis thaliana*	1.43	0.00194
Contig55342	35.4	$2.70E^{-47}$	sp|Q9AR73|HQGT_RAUSE Hydroquinone glucosyltransferase	*Rauvolfia serpentina*	1.35	0.0236
Contig95290	40.3	$4.90E^{-55}$	sp|P52839|SOT12_ARATH Cytosolic sulfotransferase 12	*Arabidopsis thaliana*	1.27	0.0375
Contig41722	63.8	$4.80E^{-51}$	sp|P19036|HSP17_ARATH 17.4 kDa class I heat shock protein	*Arabidopsis thaliana*	1.26	0.0191
Contig74545	90.2	$1.20E^{-206}$	sp|A3C4S4|GME1_ORYSJ GDP-mannose 3,5-epimerase 1	*Oryza sativa subsp. japonica*	1.1	0.0138
Contig91531	42.1	$5.80E^{-35}$	sp|O24407|IAA16_ARATH Auxin-responsive protein IAA16	*Arabidopsis thaliana*	1.04	0.00256
Contig36224	67	$3.20E^{-92}$	sp|Q9LQI7|CIA30_ARATH Probable complex I intermediate-associated protein 30	*Arabidopsis thaliana*	0.924	0.0762

Table 4.7 (continued).

Contig102550	45	$6.10E^{-29}$	sp|P35135|UBC4_SOLLC Ubiquitin-conjugating enzyme E2-17 kDa	*Solanum lycopersicum*	0.922	0.02
Contig95516	62.6	$2.40E^{-109}$	sp|Q9ZVF7|ESMD1_ARATH Protein ESMERALDA	*Arabidopsis thaliana*	0.908	0.0138
Contig62470	77.2	$3.70E^{-44}$	sp|P67884|RL30_OPHHA 60S ribosomal protein L30 O	*Ophiophagus hannah*	0.773	0.0469
Contig8	41.8	$2.30E^{-109}$	sp|Q2TA14|PCP_BOVIN Lysosomal Pro-X carboxypeptidase	*Bos taurus*	0.725	0.0666
Contig42331	57.4	$1.90E^{-197}$	sp|Q941X2|LAC3_ORYSJ Laccase-3	*Oryza sativa subsp. japonica*	0.584	0.0724
Contig53799	34.7	$3.50E^{-82}$	sp|Q9SI53|PP147_ARATH Pentatricopeptide repeat-containing protein At2g03880	*Arabidopsis thaliana*	0.512	0.0748
Contig73926	74	$0.00E^{+00}$	sp|Q10NY2|TPR3_ORYSJ Protein TPR3	*Aryza sativa subsp. japonica*	-0.854	0.0139
Contig92760	49.3	$7.50E^{-53}$	sp|Q9LT10|CXE18_ARATH Probable carboxylesterase 18	*Arabidopsis thaliana*	-0.873	0.0732
Contig68839	29.2	$1.40E^{-71}$	sp|Q8C0L8|COG5_MOUSE Conserved oligomeric Golgi complex subunit 5	*Mus musculus*	-0.885	0.011
Contig55431	83.7	$1.20E^{-92}$	sp|A8IU92|CFA20_CHLRE Cilia- and flagella-associated protein 20	*Chlamydomonas reinhardtii*	-0.976	0.00187
Contig55344	39.9	$1.10E^{-40}$	sp|Q9AR73|HQGT_RAUSE Hydroquinone glucosyltransferase	*Rauvolfia serpentina*	-1.95	0.0146
Contig105220	45.5	$3.40E^{-119}$	sp|O65351|SBT17_ARATH Subtilisin-like protease SBT1.7	*Arabidopsis thaliana*	-3.88	3.02E-05

Table 4.8: Combined Resistant vs Susceptible differentially expressed transcripts with known function

List of differentially expressed transcripts with EnTAP assigned function between combined PRR resistant and PRR susceptible samples. Transcripts are sorted by descending logarithmic fold change (LFC) with their associated p-value corrected for multiple testing.

Query Sequence	Percent Identical	E Value	Description	Species	LFC	Padj
61891	36.3	$1.20E^{-52}$	sp|Q75WF2|DNA2_ACAPL Plancitoxin-1	*Acanthaster planci*	5.82	0.0352
66260	47.2	$7.50E^{-23}$	sp|Q54WR8|GNA1_DICDI Glucosamine 6-phosphate N-acetyltransferase 1	*Dictyostelium discoideum*	5.76	0.00952
77525	22.7	$5.90E^{-13}$	sp|P94026|RBCMT_TOBAC Ribulose-1,5 bisphosphate carboxylase/oxygenase large subunit N-methyltransferase, chloroplastic	*Nicotiana tabacum*	5.19	0.062
52785	31.2	$5.70E^{-41}$	sp|Q54JH4|P2XE_DICDI P2X receptor E	*Dictyostelium discoideum*	4.76	0.0137
60403	60.4	$3.30E^{-133}$	sp|Q9DC48|PRP17_MOUSE Pre-mRNA-processing factor 17	*Mus musculus*	4.52	0.0683
80917	49.7	$5.70E^{-179}$	sp|P51061|CAPP2_SOYBN Phosphoenolpyruvate carboxylase	*Glycine max*	4.25	0.0606
77993	51.3	$2.70E^{-159}$	sp|Q29504|UBA1_RABIT Ubiquitin-like modifier-activating enzyme 1	*Oryctolagus cuniculus*	4.25	0.0456
56434	61.8	$1.60E^{-96}$	sp|Q54MB4|ASNS_DICDI Probable asparagine synthetase [glutamine-hydrolyzing]	*Dictyostelium discoideum*	4.14	0.0381
29539	68	$3.70E^{-89}$	sp|Q54DW2|ERGI3_DICDI Probable endoplasmic reticulum-Golgi intermediate compartment protein 3	*Dictyostelium discoideum*	3.96	0.0328
75625	61.3	$1.90E^{-130}$	sp|Q54MZ4|MCFB_DICDI Mitochondrial substrate carrier family protein B	*Dictyostelium discoideum*	3.92	0.0524
76557	59.8	$2.60E^{-100}$	sp|Q5V9F0|AMPL_DICDI Cytosol aminopeptidase	*Dictyostelium discoideum*	3.88	0.0683
75958	54.6	$3.60E^{-108}$	sp|P51567|AFC2_ARATH Serine/threonine-protein kinase AFC2	*Arabidopsis thaliana*	3.83	0.0945
69521	66.9	$4.60E^{-92}$	sp|Q94BX4|PIGA_ARATH Phosphatidylinositol N-acetylglucosaminyltransferase subunit A	*Arabidopsis thaliana*	3.79	0.0762
75194	66.4	$9.50E^{-179}$	sp|Q86A85|PUR1_DICDI Amidophosphoribosyltransferase	*Dictyostelium discoideum*	3.64	0.0685
69559	54.9	$1.50E^{-75}$	sp|Q9LTV6|DECR2_ARATH Peroxisomal 2,4-dienoyl-CoA reductase	*Arabidopsis thaliana*	3.63	0.0843

Table 4.8(continued).

55600	47	$2.60E^{-76}$	sp|Q9AV81|PRP19_ORYSJ Pre-mRNA-processing factor 19	*Oryza sativa subsp. japonica*	3.53	0.0667
5685	34.8	$5.70E^{-22}$	sp|Q9M8K7|DUS1B_ARATH Dual specificity protein phosphatase 1B	*Arabidopsis thaliana*	3.35	0.0146
53129	73.3	$3.30E^{-121}$	sp|Q86AD7|MYLKB_DICDI Probable myosin light chain kinase DDB_G0271550	*Dictyostelium discoideum*	3.3	0.0477
59621	44.9	$1.00E^{-29}$	sp|A5JYX5|DHS3_CAEEL Protein dhs-3	*Caenorhabditis elegans*	3.28	0.0685
75232	49.3	$6.40E^{-152}$	sp|Q54BM8|U652_DICDI UPF0652 protein	*Dictyostelium discoideum*	3.2	0.0911
75504	61.2	$3.20E^{-95}$	sp|Q550A8|ATG3_DICDI Autophagy-related protein 3	*Dictyostelium discoideum*	3.18	0.0782
75380	52.9	$1.40E^{-68}$	sp|P26368|U2AF2_HUMAN Splicing factor U2AF 65 kDa subunit	*Homo sapiens*	3.16	0.0897
57675	58.7	$3.80E^{-60}$	sp|P08239|GNAO_BOVIN Guanine nucleotide-binding protein G(o) subunit alpha	*Bos taurus*	3.11	0.0585
73111	46	$1.80E^{-87}$	sp|Q8H116|MNS2_ARATH Mannosyl-oligosaccharide 1,2-alpha-mannosidase MNS2	*Arabidopsis thaliana*	3.05	0.0224
84718	28.6	$4.20E^{-18}$	sp|Q5SRT8|CCNJL_MOUSE Cyclin-J-like protein	*Mus musculus*	2.87	0.0307
58197	55.4	$2.40E^{-50}$	sp|P42322|CANB1_NAEGR Calcineurin subunit B	*Naegleria gruberi*	2.71	0.0705
66216	37.1	$1.40E^{-51}$	sp|Q54KZ8|EIF3M_DICDI Eukaryotic translation initiation factor 3 subunit M	*Dictyostelium discoideum*	2.66	0.0383
57526	38.5	$6.90E^{-42}$	sp|Q4V9P6|ADAL_DANRE Adenosine deaminase-like protein	*Danio rerio*	2.64	0.0155
79283	24	$8.10E^{-19}$	sp|Q84JB8|RUS3_ARATH Protein root UVB sensitive 3	*Arabidopsis thaliana*	2.47	0.0575
52299	41.3	$1.60E^{-63}$	sp|Q54IP0|DNJC7_DICDI DnaJ homolog subfamily C member 7 homolog	*Dictyostelium discoideum*	2.38	0.0153
60092	73.1	$1.40E^{-155}$	sp|Q869Z4|PCKGM_DICDI Phosphoenolpyruvate carboxykinase [GTP], mitochondrial	*Dictyostelium discoideum*	2.34	0.0327
91749	62.2	$1.20E^{-81}$	sp|P0CD61|GPT_DICDI UDP-N-acetylglucosamine--dolichyl-phosphate N-acetylglucosaminephosphotransferase	*Dictyostelium discoideum*	2.32	0.0781
46531	65.4	$2.30E^{-55}$	sp|P34149|RACC_DICDI Rho-related protein racC	*Dictyostelium discoideum*	2.3	0.0164
58440	39	$4.80E^{-42}$	sp|Q86HD3|MRT4_DICDI Ribosome assembly factor mrt4	*Dictyostelium discoideum*	2.25	0.0328

Table 4.8(continued).

53174	58.3	$7.30E^{-106}$	sp|Q17M80|MMSA_AEDAE Probable methylmalonate-semialdehyde dehydrogenase [acylating], mitochondrial	*Aedes aegypti*	2.18	0.0707
75171	61.8	$1.60E^{-151}$	sp|P11141|HSP7F_CAEEL Heat shock 70 kDa protein F, mitochondrial	*Caenorhabditis elegans*	2.06	0.097
82107	74.4	$6.60E^{-66}$	sp|Q54X16|GID8_DICDI Glucose-induced degradation protein 8 homolog	*Dictyostelium discoideum*	1.98	0.0843
63371	50.2	$2.60E^{-74}$	sp|Q91YR7|PRP6_MOUSE Pre-mRNA-processing factor 6	*Mus musculus*	1.98	0.0945
61816	46.9	$4.30E^{-19}$	sp|Q94518|NACA_DROME Nascent polypeptide-associated complex subunit alpha	*Drosophila melanogaster*	1.97	0.0941
103736	65.8	$2.90E^{-124}$	sp|P11141|HSP7F_CAEEL Heat shock 70 kDa protein F, mitochondrial	*Caenorhabditis elegans*	1.93	0.0348
78485	71.2	$1.80E^{-125}$	sp|Q641W4|RFC2_RAT Replication factor C subunit 2	*Rattus norvegicus*	1.86	0.0947
56850	60	$6.40E^{-56}$	sp|O76767|ERD2_DROME ER lumen protein-retaining receptor	*Drosophila melanogaster*	1.85	0.0192
86756	54.4	$3.20E^{-54}$	sp|Q8GXS3|CFIS2_ARATH Pre-mRNA cleavage factor Im 25 kDa subunit 2	*Arabidopsis thaliana*	1.57	0.0636
89396	31.8	$2.60E^{-39}$	sp|Q29RL2|CHRD1_BOVIN Cysteine and histidine-rich domain-containing protein 1	*Bos taurus*	1.2	0.0869
53707	74.2	$1.80E^{-101}$	sp|P51893|SAHHA_XENLA Adenosylhomocysteinase A	*Xenopus laevis*	1.17	0.0785
52934	75	$5.00E^{-72}$	sp|Q9LT08|PSDE_ARATH 26S proteasome non-ATPase regulatory subunit 14 homolog	*Arabidopsis thaliana*	1.13	0.0464
49886	31.5	$4.50E^{-43}$	sp|Q9DCN7|RNFT1_MOUSE E3 ubiquitin-protein ligase RNFT1	*Mus musculus*	1.09	0.0761
94826	62.2	$1.20E^{-154}$	sp|Q54JP5|OAT_DICDI Probable ornithine aminotransferase	*Dictyostelium discoideum*	1.03	0.0672
16227	32.9	$1.60E^{-37}$	sp|O14734|ACOT8_HUMAN Acyl-coenzyme A thioesterase 8	*Homo sapiens*	1.03	0.0856
99238	95.8	$1.90E^{-141}$	sp|P53455|ACT_AJECG Actin	*Ajellomyces capsulatus*	0.88	0.0585
102550	45	$6.10E^{-29}$	sp|P35135|UBC4_SOLLC Ubiquitin-conjugating enzyme E2-17 kDa	*Solanum lycopersicum*	0.86	0.00947
4374	31.7	$2.00E^{-07}$	sp|Q8VY97|Y4213_ARATH ER membrane protein complex subunit 7 homolog	*Arabidopsis thaliana*	0.82	0.0913

Table 4.8(continued).

14983	96.9	$1.40E^{-148}$	sp|Q9SGD6|AROD6_ARATH Arogenate dehydratase/prephenate dehydratase 6, chloroplastic	*Arabidopsis thaliana*	-0.7	0.0478
74263	68.6	$1.10E^{-62}$	sp|Q6NQ66|GGLO2_ARATH L-gulonolactone oxidase 2	*Arabidopsis thaliana*	-0.7	0.0911
51003	47.4	$7.20E^{-38}$	sp|Q9M1K9|ATA1_ARATH Short-chain dehydrogenase reductase ATA1	*Arabidopsis thaliana*	-0.8	0.0793
52248	67	$3.60E^{-162}$	sp|Q9S7Y7|XYL1_ARATH Alpha-xylosidase 1	*Arabidopsis thaliana*	-0.8	0.0431
17	67	$3.90E^{-128}$	sp|Q8VYZ3|PME53_ARATH Probable pectinesterase 53	*Arabidopsis thaliana*	-0.9	0.0362
55343	24.6	$2.70E^{-47}$	sp|Q9AR73|HQGT_RAUSE Hydroquinone glucosyltransferase	*Rauvolfia serpentina*	-1.3	0.0393
55344	43.4	$1.10E^{-40}$	sp|Q9AR73|HQGT_RAUSE Hydroquinone glucosyltransferase	*Rauvolfia serpentina*	-2	0.0292
77413	74.2	$5.50E^{-42}$	sp|P0DKJ3|SWT1A_SORBI Bidirectional sugar transporter SWEET1a	*Sorghum bicolor*	-4.8	0.0337

4.7 Figures

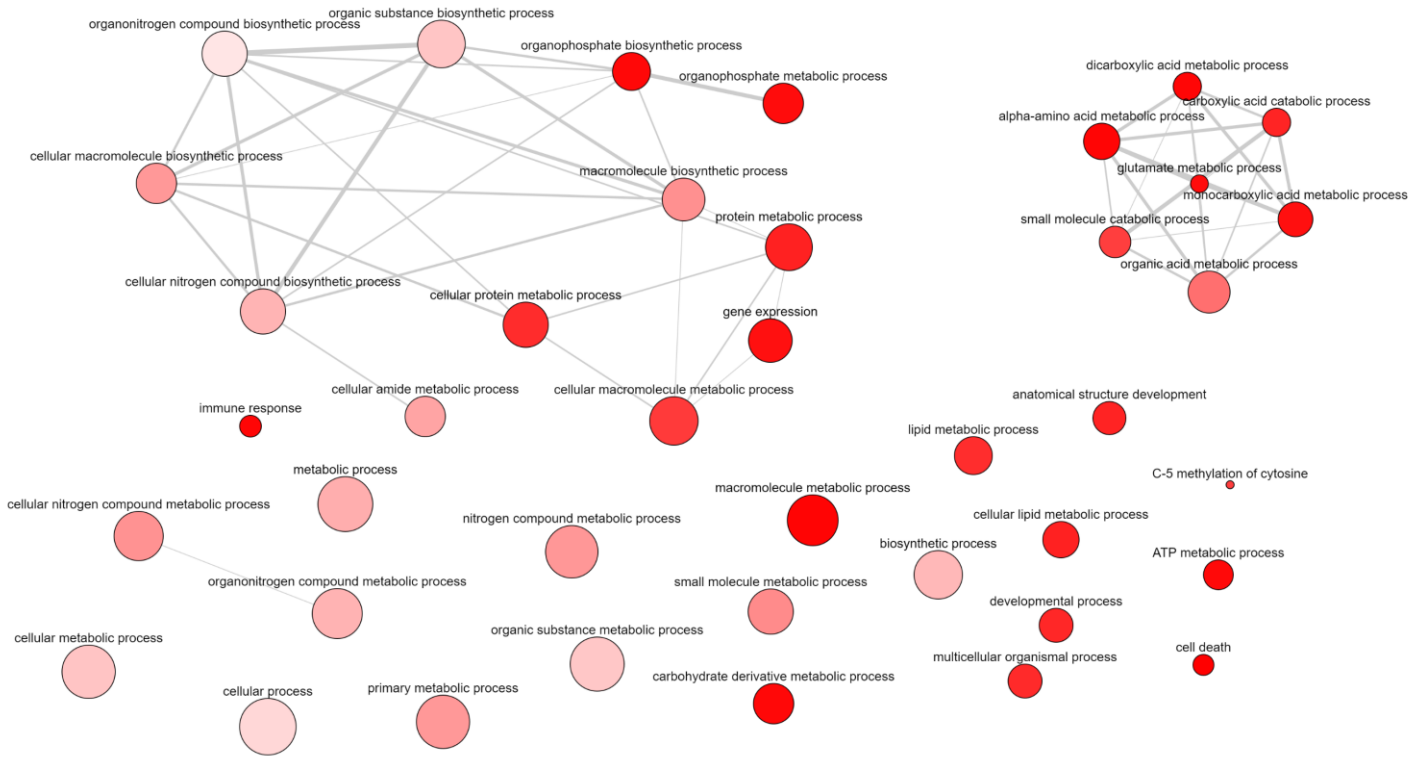


Figure 4.1: GO term enrichment network for combined PRR resistant vs susceptible DE genes

Revigo generated GO term network of transcripts differentially expressed between the combined PRR resistant and PRR susceptible samples. The color of the bubble corresponds to the GO term enrichment p-value. The size of the bubble corresponds to the LogSize value for the GO Term.

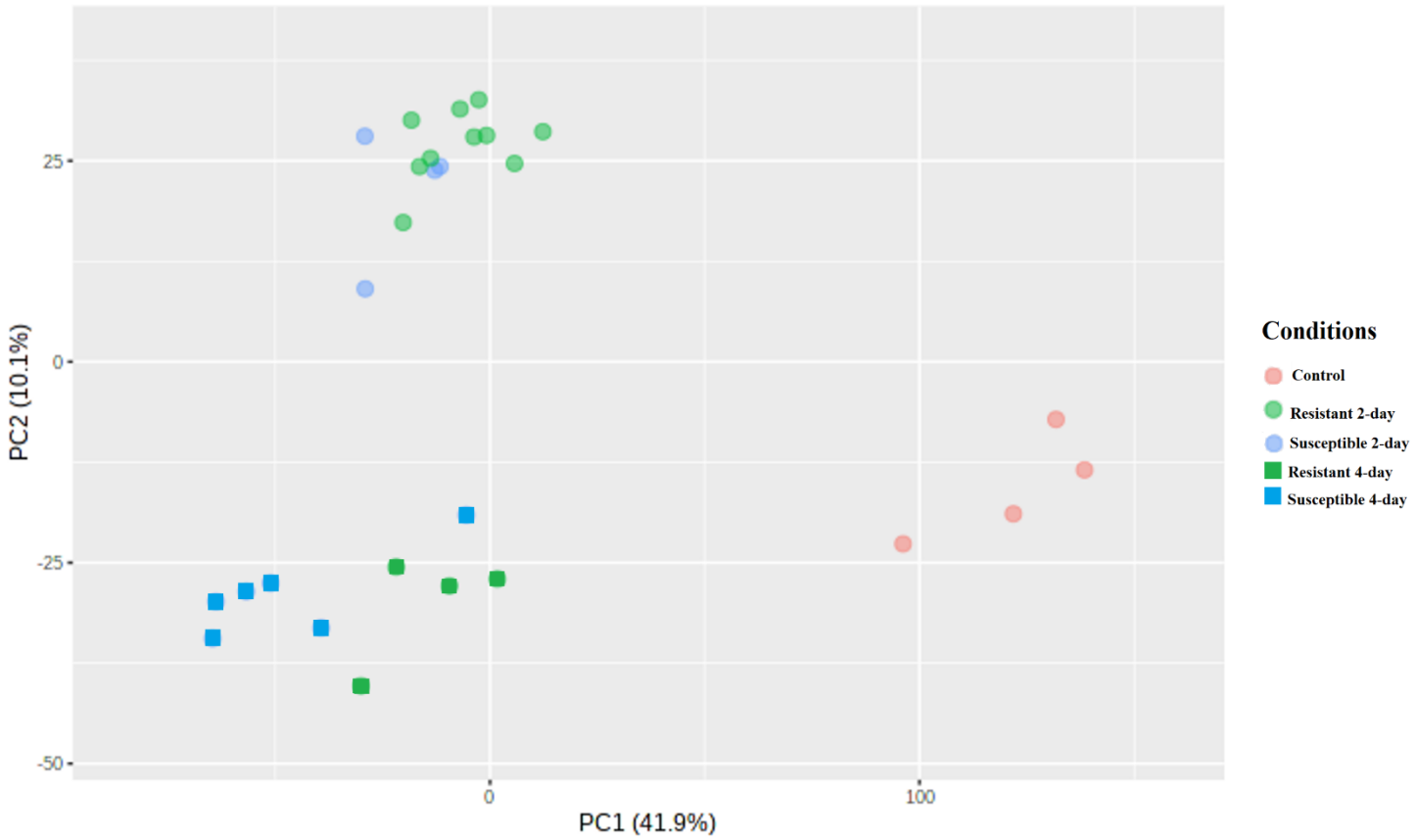


Figure 4.2: PCA of normalized RNAseq libraries from second study

Principle component analysis of normalized RNAseq read libraries. The inoculated samples and non-inoculated controls are separated by PC1 and the *Phytophthora* exposure length is separated by PC2.

4.8 Literature Cited

Auer, P. L., & Doerge, R. W. (2010). Statistical Design and Analysis of RNA Sequencing Data. Genetics, 185(2), 405–416. https://doi.org/10.1534/genetics.110.114983

Besseau S, Hoffmann L, Geoffroy P, Lapierre C, Pollet B, Legrand M. 2007). Flavonoid accumulation in Arabidopsis repressed in lignin synthesis affects auxin transport and plant growth. Plant Cell,19(1):148-62.

Chen, L. (2014). SWEET sugar transporters for phloem transport and pathogen nutrition. New Phytologist, 201(4), 1150–1155. https://doi.org/10.1111/nph.12445

Chen, L.-Q., Hou, B.-H., Lalonde, S., Takanaga, H., Hartung, M. L., Qu, X.-Q., Guo, W.-J., Kim, J.-G., Underwood, W., Chaudhuri, B., Chermak, D., Antony, G., White, F. F., Somerville, S. C., Mudgett, M. B., & Frommer, W. B. (2010). Sugar transporters for intercellular exchange and nutrition of pathogens. Nature, 468(7323), 527–532. https://doi.org/10.1038/nature09606

Conesa, A., Madrigal, P., Tarazona, S., Gomez-Cabrero, D., Cervera, A., McPherson, A., Szcześniak, M. W., Gaffney, D. J., Elo, L. L., Zhang, X., & Mortazavi, A. (2016). A survey of best practices for RNA-seq data analysis. Genome Biology, 17(1). https://doi.org/10.1186/s13059-016-0881-8

Cross, T. G., Scheel-Toellner, D., Henriquez, N. V., Deacon, E., Salmon, M., & Lord, J. M. (2000). Serine/threonine protein kinases and apoptosis. Experimental cell research, 256(1), 34–41. https://doi.org/10.1006/excr.2000.4836

Feng, B., & Li, P. (2012). Genome-wide identification of laccase gene family in three Phytophthora species. Genetica, 140(10–12), 477–484. https://doi.org/10.1007/s10709-012-9696-z

Fernandes, P., Machado, H., Silva, M. do C., & Costa, R. L. (2021). A Histopathological Study Reveals New Insights Into Responses of Chestnut (Castanea spp.) to Root Infection by Phytophthora cinnamomi. Phytopathology®, 111(2), 345–355. https://doi.org/10.1094/PHYTO-04-20-0115-R

García-Rodríguez, S., Pozo, M. J., Azcón-Aguilar, C., & Ferrol, N. (2005). Expression of a tomato sugar transporter is increased in leaves of mycorrhizal or Phytophthora parasitica-infected plants. Mycorrhiza, 15(7), 489–496. https://doi.org/10.1007/s00572-005-0354-5

Grabherr, M. G., Haas, B. J., Yassour, M., Levin, J. Z., Thompson, D. A., Amit, I., Adiconis, X., Fan, L., Raychowdhury, R., Zeng, Q., Chen, Z., Mauceli, E., Hacohen, N., Gnirke, A., Rhind, N., di Palma, F., Birren, B. W., Nusbaum, C., Lindblad-Toh, K., … Regev, A. (2011). Full-length transcriptome assembly from RNA-Seq data without a reference genome. Nature Biotechnology, 29(7), 644–652. https://doi.org/10.1038/nbt.1883

Hirschmann, F., Krause, F., & Papenbrock, J. (2014). The multi-protein family of sulfotransferases in plants: composition, occurrence, substrate specificity, and functions. Frontiers in plant science, 5, 556

Huang, D. W., Sherman, B. T., & Lempicki, R. A. (2009). Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. Nature Protocols, 4(1), 44–57. https://doi.org/10.1038/nprot.2008.211

Islam, M. T., Hussain, H. I., Rookes, J. E., & Cahill, D. M. (2018). Transcriptome analysis, using RNA-Seq of Lomandra longifolia roots infected with Phytophthora cinnamomi reveals the complexity of the resistance response. Plant Biology, 20(1), 130–142.

Jiang H, Wang S, Dang L, Wang S, Chen H, Wu Y, Jiang X, Wu P. A novel short-root gene encodes a glucosamine-6-phosphate acetyltransferase required for maintaining normal root cell shape in rice. Plant Physiol. 2005 May;138(1):232-42. doi: 10.1104/pp.104.058248

Kohlway, W. H., et al. "Response of Turkish and Trojan Fir to Phytophthora Cinnamomi and P. Cryptogea." Scandinavian Journal of Forest Research, vol. 32, no. 5, July 2017, pp. 406–11. Crossref, https://doi.org/10.1080/02827581.2017.1280076.

Kurt, Y., Frampton, J., Isik, F., Landgren, C., & Chastagner, G. (2016). Variation in needle and cone characteristics and seed germination ability of Abies bornmuelleriana and Abies equi-trojani populations from Turkey. TURKISH JOURNAL OF AGRICULTURE AND FORESTRY, 40, 169–176. https://doi.org/10.3906/tar-1502-101

LaFayette, P.R., Eriksson, KE.L. & Dean, J.F. (1999). Characterization and heterologous expression of laccase cDNAs from xylem tissues of yellow-poplar (Liriodendron tulipifera). Plant Mol Biol 40, 23–35. https://doi.org/10.1023/A:1026437406859

Love, M. I., Huber, W., & Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biology, 15(12). https://doi.org/10.1186/s13059-014-0550-8

Pettersson, M., Frampton, J., Rönnberg, J., Shew, H. D., Benson, D. M., Kohlway, W. H., Cubeta, M. A. (2017). Increased diversity of Phytophthora species in Fraser fir Christmas tree plantations in the Southern Appalachians. Scandinavian Journal of Forest Research, 32(5), 412–420. https://doi.org/10.1080/02827581.2016.1265144

Rellstab, C., Dauphin, B., Zoller, S., Brodbeck, S., & Gugerli, F. (2019). Using transcriptome sequencing and pooled exome capture to study local adaptation in the giga-genome of Pinus cembra. Molecular ecology resources, 19(2), 536–551. https://doi.org/10.1111/1755-0998.12986

Rivera-Vargas, L. I., Schmitthenner, A. F., & Graham, T. L. (1993). Soybean flavonoid effects on and metabolism by Phytophthora sojae. The International Journal of Plant Biochemistry, 32(4), 851–857. https://doi.org/10.1016/0031-9422(93)85219-H

Robertson, G., Schein, J., Chiu, R., Corbett, R., Field, M., Jackman, S. D., Mungall, K., Lee, S., Okada, H. M., Qian, J. Q., Griffith, M., Raymond, A., Thiessen, N., Cezard, T., Butterfield, Y. S., Newsome, R., Chan, S. K., She, R., Varhol, R., … Birol, I. (2010). De novo assembly and analysis of RNA-seq data. Nature Methods, 7(11), 909–912. https://doi.org/10.1038/nmeth.1517

Simão, F. A., Waterhouse, R. M., Ioannidis, P., Kriventseva, E. V., & Zdobnov, E. M. (2015). BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs. Bioinformatics, 31(19), 3210–3212. https://doi.org/10.1093/bioinformatics/btv351

Smith-Unna, R., Boursnell, C., Patro, R., Hibberd, J. M., & Kelly, S. (2016). TransRate: Reference-free quality assessment of de novo transcriptome assemblies. Genome Research, 26(8), 1134–1144. https://doi.org/10.1101/gr.196469.115

Supek, F., Bošnjak, M., Škunca, N., & Šmuc, T. (2011). REVIGO Summarizes and Visualizes Long Lists of Gene Ontology Terms. PLoS ONE, 6(7), e21800. https://doi.org/10.1371/journal.pone.0021800

Vilela, B., Pagès, M., & Lumbreras, V. (2010). Regulation of MAPK signaling and cell death by MAPK phosphatase MKP2. Plant signaling & behavior, 5(11), 1497–1500.

Wang, Z., Gerstein, M., & Snyder, M. (2009). RNA-Seq: A revolutionary tool for transcriptomics. Nature Reviews Genetics, 10(1), 57–63. https://doi.org/10.1038/nrg2484

Webster, C., Figueroa-Corona, L., Méndez-González, I. D., Álvarez-Soto, L., Neale, D. B., Jaramillo-Correa, J. P., Wegrzyn, J. L., & Vázquez-Lobo, A. (2022). Comparative analysis of differential gene expression indicates divergence in ontogenetic strategies of leaves in two conifer genera. Ecology and Evolution, 12(2), e8611. https://doi.org/10.1002/ece3.8611

Xie, M., Zhang, J., Tschaplinski, T. J., Tuskan, G. A., Chen, J.-G., & Muchero, W. (2018). Regulation of Lignin Biosynthesis and Its Role in Growth-Defense Tradeoffs. Frontiers in Plant Science, 9. https://www.frontiersin.org/articles/10.3389/fpls.2018.01427

An, Y., Geng, Y., Yao, J., Fu, C., Lu, M., Wang, C., & Du, J. (2020). Efficient Genome Editing in Populus Using CRISPR/Cas12a. Frontiers in plant science, 11, 593938. https://doi.org/10.3389/fpls.2020.593938

Zhao, Q.-Y., Wang, Y., Kong, Y.-M., Luo, D., Li, X., & Hao, P. (2011). Optimizing de novo transcriptome assembly from short-read RNA-Seq data: A comparative study. BMC Bioinformatics, 12(Suppl 14), S2. https://doi.org/10.1186/1471-2105-12-S14-S2

CHAPTER 5

5.1 Summary of total work

The main goal of this project was to generate new genomic resources for Trojan fir to explore the host-pathogen interaction of fir with *Phytophthora cinnamomi*. That goal was achieved, albeit not to the extent that was planned at the onset of the project. The production of the novel Trojan fir root transcriptome and megagametophyte genetic markers will be invaluable for future genomic studies. Additionally, this project confirmed a novel Phytophthora root rot resistance fir seedling screening technique as an effective alternative to past methods. This study also pioneered the use of rRNA depletion with conifer tissue as an alternative to standard oligo(dT) RNAseq library preparation to rescue degraded RNA samples. RNA degradation is a constant concern when collecting samples from the field, often hours away from a lab. Finally, the analysis of Trojan fir root response to *Phytophthora cinnamomi* revealed new insights into the complex interaction of *Phytophthora* with fir roots.

5.2 Goals not met and lessoned learned

Of the original goals of this project, the major target missed was a chromosome-scale genetic map and the mapping PRR resistance makers to it. As described in chapter 2, there were a few unexpected problems with the collected seed, but the experimental design may have been too naïve to begin with. With as large and complex of a genome as Trojan fir, 63 GBS mapping individuals was not enough to capture enough recombination to construct a robust linkage map. Indeed, the large amount of missing data within the GBS libraries confounded the marker screening process and overcame the benefits the haploid megagametophytic tissue provided. The lack of a known parental generation also further

complicated the analysis, as without a known maternal genotype, the created artificial Kaz24 parent likely introduced additional error. If this experiment was repeated, a much larger mapping population would need to be selected and if possible, they should be progeny from parents that could be genotyped themselves. The larger mapping population could then be screened for PRR resistance to check for genetic linkage to PRR in Trojan fir.

5.3 Future work

Trojan fir and other exotic fir species will continue to be of great interest to Christmas tree growers, as Phytophthora continues to spread. A test plot of Trojan fir survivors from a previous PRR resistance screen (Frampton et al. 2005) have matured enough to begin to produce cones (unpublished). Access to those trees will allow for controlled crosses with fully genotyped parents and will vastly improve the quality of the mapping population(Lauer and Isik 2021).

There remains much unknown about the *Phytophthora*-fir interaction, but there is still a lot of information buried within the Trojan fir root transcriptome. As fir somatic embryogenic cultures become available, there will be the opportunity to perturb the expression of some of the differentially expressed genes with unknown function identified in this study. Fraser fir has no innate resistance to Phytophthora (Frampton et al. 2005) but remains the most valuable Christmas tree to NC growers. Therefore, the development of *Phytophthora* resistance in Fraser fir may rely on artificial genetic modification of key genes like those in the lignin biosynthesis pathway using tools like CRISPR that have already found success in woody plants (An et al. 2020). Finally, this study focused on the Trojan fir root

tissue, but future sequencing of the *Phytophthora cinnamomi* isolate used in this study will help to identify pathogenicity factors and how they interact with the Trojan fir roots.

5.4 Literature cited

An, Y., Geng, Y., Yao, J., Fu, C., Lu, M., Wang, C., & Du, J. (2020). Efficient Genome Editing in Populus Using CRISPR/Cas12a. Frontiers in plant science, 11, 593938. https://doi.org/10.3389/fpls.2020.593938

Frampton J, Fikret I, Benson DM (2013). Genetic variation in resistance to Phytophthora cinnamomi in seedlings of two Turkish Abies species. Tree Genetics & Genomes. 9:53-63

Lauer E, Isik F. (2021). Major QTL confer race-nonspecific resistance in the co-evolved Cronartium quercuum f. sp. fusiforme–Pinus taeda pathosystem. Heredity. 1–12.

APPENDIX

Appendix A

Oligo Name	Sequence
18srRNA.R01	TTAATCATGGCCTCAGTTCCGAAAACCAACAAAATAGAACCGCGGTCCTATTCC ATTATTCCTAGCTGCGGTATCCAGGC
18srRNA.R02	GCTTTCGCTCTGGTCCGTCTTGCGCCGGTCCAAGAATTTCACCTCTAGCGGCGC AATACGAATGCCCCCGGCCGTCCCTC
18srRNA.R03	TGGTCGGAACTACGACGGTATCTGATCGTCTTCGAACCTCCGACTTTCGTTCTT GATTAATGAAAACATTCTTGGCAAAT
18srRNA.R04	TCCTCGTTCATGGGGAATAATTGCAATCCCCGATCCCCATCACGAATGGGGTTC AACGGGTTACCCGCGCCTGCCGGCGT
18srRNA.R05	ATCGGTAGTAGCGACGGGCGGTGTGTACAAAGGGCAGGGACTTAATCAACGCA AGCTTATGACCCGCACTTACTGGGAAT
18srRNA.R06	TTCGACCGTCTTCTCAGCGCTCCGCCAGGCCGTGGGCCGACCCCGGCGGGGCC GATCCGAGGGCCTCACTAAACCATCCA
18srRNA.R07	GGGGGCTCGAGGACGGGCCCGGCGCCCCGCAAGCGAGGAGGACGACGGACGG ACGG
18srRNA.R08	TTTGAGACAAGCATATGCTACTGGCAGGATCAACCAGGTAGGTAGGTAGAGCG CGGCGAGGCCCCGACGCGGCCGGACGG
18srRNA.R09	AAGGAACCATAACTGATTTAATGAGCCATTCGCAGTTTCACTGTACCGGCCGTG CGTACTTAGACATGCATGGCTTAATC
18srRNA.R10	CGAAGGGGGTCAGCGCCCGTCGGCATGTATTAGCTCTAGAATTACCACAGTTA TCCAAGTAGGAGAGGAGCGAGCGACCA
18srRNA.R11	CCGCGGCCCGCCCCCCGGCCGGGGCCGGAGAGGGGCTGACCGGGTTGGTTTTG ATCTGATAAATGCACGCATCCCCCCCG
18srRNA.R12	TAGGGCAGACGTTCGAATGGGTCGTCGCCGCCACGGGGGGCGTGCGATCGGCC CGAGGTTATCTAGAGTCACCAAAGCCG
18srRNA.R13	CTCCCTCTCCGGAATCGAACCCTGATTCCCCGTCACCCGTGGTCACCATGGTAG GCACGGCGACTACCATCGAAAGTTGA
18srRNA.R14	TTCGTCACTACCTCCCCGGGTCGGGAGTGGGTAATTTGCGCGCCTGCTGCCTTC CTTGGATGTGGTAGCCGTTTCTCAGG
18srRNA.R15	CTCCAATGGATCCTCGTTAAAGGATTTAAAGTGGACTCATTCCAATTACAGGGC CTCGAAAGAGTCCTGTATTGTTATTT
18srRNA.R16	CTACGAGCTTTTTAACTGCAGCAACTTTAATATACGCTATTGGAGCTGGAATTA CCGCGGCTGCTGGCACCAGACTTGCC
18srRNA.R17	AGGGGGCGCCGAGAGGCAAGGGGCGGGGACGGGCGGTGGCTCGCCTCGCGGC GGACCGCCCGCCCGCTCCCAAGATCCAA
18srRNA.R18	GGCTCGGGCCTGCTTTGAACACTCTAATTTTTTCAAAGTAAACGCTTCGGGCCC CGCGGGACACTCAGCTAAGAGCATCG
18srRNA.R19	ACCCAAAGACTTTGGTTTCCCGGAAGCTGCCCGGCGGGTCATGGGAATAACGC CGCCGCATCGCCGGTCGGCATCGTTTA
18srRNA.R20	TAAGCCGCAGGCTCCACTCCTGGTGGTGCCCTTCCGTCAATTCCTTTAAGTTTC AGCTTTGCAACCATACTCCCCCCGGA
18srRNA.R21	CACGGAATCGAGAAAGAGCTATCAATCTGTCAATCCTGTCCGTGTCCGGGCCG GGTGAGGTTTCCCGTGTTGAGTCAAAT
18srRNA.R22	AGCATGCCAGAGTCTCGTTCGTTATCGGAATTAACCAGACAAATCGCTCCACCA ACTAAGAACGGCCATGCACCACCACC
18srRNA.R23	TCAATCTCGGGTGGCTGAACGCCACTTGTCCCTCTAAGAAGTTGGGGGACGCC GACCGCTCGGGGGTCGCGTAACTAGTT
18srRNA.R24	AGGGTAGGCACACGCTGAGCCAGTCAGTGTAGCGCGCGTGCAGCCCCGGACAT CTAAGGGCATCACAGACCTGTTATTGC
18srRNA.R25	TAATGATCCTTCCGCAGGTTCACCTACGGAAACCTTGTTACGACTTTTACTTCCT CTAGATAGTCAAG

28srRNA.R01	GGAATCCTGGTTAGTTTCTTTTCCTCCGCTGACTAATATGCTTAAATTCAGCGG GTCGCCACGTCTGATCTGAGGTCGCG
28srRNA.R02	TTCCGTACGCCACATGTCCCGCGCCCCGCGGGGCGGGGATTCGGCGCTGGGCT CTTCCCTGTTCACTCGCCGTTACTGAG
28srRNA.R03	TACCGGCCTCACACCGTCCACGGGCTGGGCCTCGATCAGAAGGACTTGGGCCC CCCACGAGCGGCGCCGGGGAGCGGGTC
28srRNA.R04	TTAGATGGAGTTTACCACCCGCTTTGGGCTGCATTCCCAAGCAACCCGACTCCG GGAAGACCCGGGCGCGCGCCGGCCGC
28srRNA.R05	CTTGAACTCTCTCTTCAAAGTTCTTTTCAACTTTCCCTTACGGTACTTGTTGACT ATCGGTCTCGTGCCGGTATTTAGCC
28srRNA.R06	GCCGGACCCGCCGCCGGGTTGAATCCTCCGGGCGGACTGCGCGGACCCCACCC GTTTACCTCTTAACGGTTTCACGCCCT
28srRNA.R07	GGGGCGGCGGGGGAAGGGAGGGCGGGTGGAGGGGTCGGGAGGAACGGGGGG CGGGAAAGATCCGCCGGGCCGCCGACACG
28srRNA.R08	TCCCCCGCCGACCCCACCCCCGGCCCCGCCCGCCCACCCCCGCACCCGCCGGA GCCCGCCCCCTCCGGGGAGGAGGAGGA
28srRNA.R09	TCCCAGCCGTCCCGGAGCCGGTCGCGGCGCACCGCCTGGAAATGCGCCCGGCG GCGGCCGGTCGCCGGTCCGGGGGACGG
28srRNA.R10	GGGGCCGGGGGGCGGAGACGGGGGAGGAGGAGGACGGACGGACGGACGGGG CCCCCCGAGCCACCTTCCCCGCCGGGCCT
28srRNA.R11	CGGTCCCGCCGCCCCCGCCGCCGCCGCCACCGCCGCCGCCGCCGCCCCGA CCCGCGCGCCCTCCCGAGGGAGGACGC
28srRNA.R12	CGGCGACGGGTCTCGCTCCCTCGGCCCCGGGATTCGGCGAGTGCTGCTGCCGG GGGGGCTGTAACACTCGGGGGGGGTTT
28srRNA.R13	AGGAGACGCCGGCGCCGCGCCGGGGGAGACCCCCCTCGCGGGGATTCCCGCGG GGGTGGGCGCCGGGAGGGGGGAGAGCG
28srRNA.R14	CCCGTCGCCGGGGCGGGGGCGCGGGGAGGAGGGGTGGGAGAGCGGTCGCGCC GTGGGAGGGGTGGCCCGGCCCCCCCACG
28srRNA.R15	CCCCGGGCCCGACGGCGCGACCCGCCCGGGGCGCACTGGGGACAGTCCGCCCC GCCCCCCGACCCGCGCGCGGCACCCCC
28srRNA.R16	CCGACGTCGCCGCCGACCCCGTGCGCTCGCTCCGCCGTCCCCCTCTTCGGGGGA CGCGCGCGTGGCCCCGAGAGAACCTC
28srRNA.R17	CGGCGGCTTTCGTGCGAGCCCCCGACTCGCGCACGTGTTAGACTCCTTGGTCCG TGTTTCAAGACGGGTCGGGTGGGTAG
28srRNA.R18	GGTGCCCCTCGGCGGACTGGAGAGGCCTCGGGATCCCACCTCGGCCGGCGAGC GCGCCGGCCTTCACCTTCATTGCGCCA
28srRNA.R19	GGCATAGTTCACCATCTTTCGGGTCCTAACACGTGCGCTCGTGCTCCACCTCCC CGGCGCGGCGGGCGAGACGGGCCGGT
28srRNA.R20	TACCCAGGTCGGACGACCGATTTGCACGTCAGGACCGCTACGGACCTCCACCA GAGTTTCCTCTGGCTTCGCCCTGCCCA
28srRNA.R21	TCTGCGAGAGCGCCAGCTATCCTGAGGGAAACTTCGGAGGGAACCAGCTACTA GATGGTTCGATTAGTCTTTCGCCCCTA
28srRNA.R22	ATAGGTTGAGATCGTTTCGGCCCCAAGACCTCTAATCATTCGCTTTACCGGATA AAACTGCGTGGCGGGGGTGCGTCGGG
28srRNA.R23	AAAGTGGCCCACTAGGCACTCGCATTCCACCCCGGCTCCACGCCAGCGAGCCG GGCTTCTTACCCATTTAAAGTTTGAGA
28srRNA.R24	TTCTGGGGTCTGATGAGCGTCGGCATCGGGCGCCTTAACCCGGCGTTCGGTTCA TCCCGCAGCGCCAGTTCTGCTTACCA
28srRNA.R25	GCAGGTGAGTTGTTACACACTCCTTAGCGGATTCCGACTTCCATGGCCACCGTC CTGCTGTCTATATCAACCAACACCTT
28srRNA.R26	CGTCCACTCTCGACTGCCGGCGACGGCCGGGTATGGGCCCGACGCTCCAGCGC CATCCATTTTCAGGGCTAGTTGATTCG
28srRNA.R27	CCCACACCCCCGCCGCCGCCGCCGCCGCCCTCCGACGCACACCACACGCG CGCGCGCGCCGCCGCCCCCGCCGCTCC

28srRNA.R28	TCCGCGGGGCTCCGGGGGCGGGGAGCGGGGCGTGGGCGGGAGGAGGGGAGGA GGCGTGGGGGGGGGGGCGGGGGAAGGAC
28srRNA.R29	CCTGCGGCGGCCTCCACCCGGGCCCGCGCCCTAGGCTTCAAGGCTCACCGCAG CGGCCCTCCTACTCGTCGCGGCGTAGC
28srRNA.R30	TGCTGTTCACATGGAACCCTTCTCCACTTCGGCCTTCAAAGTTCTCGTTTGAATA TTTGCTACTACCACCAAGATCTGCA
28srRNA.R31	CCGAGGGCAACGGAGGCCATCGCCCGTCCCTTCGGAACGGCGCTCGCCCATCT CTCAGGACCGACTGACCCATGTTCAAC
28srRNA.R32	GTTACCGCACTGGACGCCTCGCGGCGCCCATCTCCGCCACTCCGGATTCGGGGA TCTGAACCCGACTCCCTTTCGATCGG
28srRNA.R33	CCATTCCAGGGCGCCCTGCCCTTCACAAAGAAAAGAGAACTCTCCCCGGGGCT CCCGCCGGCTTCTCCGGGATCGGTCGC
28srRNA.R34	TCAAGGGCCAGCGAGAGCTCACCGGACGCCGCCGGAACCGCGACGCTTTCCAA GGCACGGGCCCCTCTCTCGGGGCGAAC
28srRNA.R35	CCAGAGGCTGTTCACCTTGGAGACCTGCTGCGGATATGGGTACGGCCCGGCGC GAGATTTACACCCTCTCCCCCGGATTT
28srRNA.R36	CCCAGCCCTTAGAGCCAATCCTTATCCCGAAGTTACGGATCCGGCTTGCCGACT TCCCTTACCTACATTGTTCCAACATG
28srRNA.R37	TGCCCCGGGCGTGGGGGGGGCGCGCGCCTCGTCCAGCCGCGGCGCGCGCCCAG CCCCGCTTCGCGCCCCAGCCCGACCGA
28srRNA.R38	AGAGAGAGAGAGAGAGGGCGCGGGGTGGGGAGGGAGCGAGCGGCGCGCGCG GGTGGGGCGGGGGAGGGCCGCGAGGGGGG
28srRNA.R39	CCTGCCGCCCCGACCCTTCTCCCCCCGCCGCGCCCCCACGCGGCGCTCCCCCGG GGAGGGGGGAGGACGGGGAGCGGGGG
28srRNA.R40	GCCGCCGGCCCCCCGGGTCCCCGGGGCCCCCCTCGCGGGGACCTGCCCCCGCC GGCCGCCCCGGCGGCCGCCGCGCGGCC
28srRNA.R41	CTCCCCGGGGGCGGCCGCGACGCCCGCCGCAGCTGGGGCGATCCACGGGAAGG GCCCGGCTCGCGTCCAGAGTCCGCGCC
28srRNA.R42	CGACCGCTCCCGCCCCCAGCGGACGCGCGCGCGACCGAGACGTGGGGTGGGGG TGGGGGGCGCGCCGCGCCGCCGCCGGG
28srRNA.R43	GAACGGGGGGCGGACGGGGCCGGGGGGGTAGGGCGGGGGGACGAACCGCCCC GCCCCGCCGCCCGCCGACCGCCGCCGCC
28srRNA.R44	GGCGGACCCGGCGGGGGGGGACCGGCCCGCGGCCCCTCCGCCGCCTGCCGCCGC CGCCGCCGCGCGCCGAGGAGGAGGGGG
28srRNA.R45	ATTCCCCTGGTCCGCACCAGTTCTAAGTCGGCTGCTAGGCGCCGGCCGAGGCG AGGCGCGCGCGGAACCGCGGCCCCGGG
28srRNA.R46	TCAGAGCACTGGGCAGAAATCACATCGCGTCAACACCCGCCGCGGGCCTTCGC GATGCTTTGTTTTAATTAAACAGTCGG
28srRNA.R47	GAGGCATTTGGCTACCTTAAGAGAGTCATAGTTACTCCCGCCGTTTACCCGCGC TTCATTGAATTTCTTCACTTTGACAT
28srRNA.R48	TGGCTGTGGTTTCGCTGGATAGTAGGTAGGGACAGTGGGAATCTCGTTCATCCA TTCATGCGCGTCACTAATTAGATGAC
28srRNA.R49	CATGTCTCTTCACCGTGCCAGACTAGAGTCAAGCTCAACAGGGTCTTCTTTCCC CGCTGATTCCGCCAAGCCCGTTCCCT
28srRNA.R50	CAGGGCCGCGGACCCCGCCCCGGGCCCCTCGCGGGGACACCGGGGGGGCGCC GGGGGCCTCCCACTTATTCTACACCTCT
28srRNA.R51	AGAGCCCCTCGGGCTCGCCCCCCCGCCTCACCGGGTCAGTGAAAAAACGATCA GAGTAGTGGTATTTCACCGGCGGCCCG
28srRNA.R52	CGCCCCAGTCAAACTCCCCACCTGGCACTGTCCCCGGAGCGGGTCGCGCCCGG CCGGGCGGGCGCTTGGCGCCAGAAGCG
28srRNA.R53	TTGCCCTTCTGCTCCACGGGAGGTTTCTGTCCTCCCTGAGCTCGCCTTAGGACA CCTGCGTTACCGTTTGACAGGTGTAC
28srRNA.R54	AAACCCAAAAGGTCAGAAGGATCGTGAGGCCCCGCTTTCACGGTCTGTATTCG TACTGAAAATCAAGATCAAGCGAGCTT

28srRNA.R55	CAAAAAGCGACGTCGCTATGAACGCTTGGCCGCCACAAGCCAGTTATCCCTGTGGTAACTTTTCTGACACCTCCTGCTTA
28srRNA.R56	CGTTCCCTATTAGTGGGTGAACAATCCAACGCTTGGCGAATTCTGCTTCACAATGATAGGAAGAGCCGACATCGAAGGAT
28srRNA.R57	CTGAGCAGGATTACCATGGCAACAACACATCATCAGTAGGGTAAAACTAACCTGTCTCACGACGGTCTAAACCCAGCTCA
28srRNA.R58	CCCACAGATGGTAGCTTCGCCCCATTGGCTCCTCAGCCAAGCACATACACCAAATGTCTGAACCTGCGGTTCCTCTCGTA
28srRNA.R59	CCGAGGCCAACCGAGGCTCCGCGGCGCTGCCGTATCGTTCGCCTGGGCGGGATTCTGACTTAGAGGCGTTCAGTCATAAT
28srRNA.R60	GCGGGGCACGCGCCCTCCCGCGGCGGGGCGCGTGGAGGGGGGGGCGGCCCGCCGGCGGGGACAGGCGGGGGACCGGCTAT
28srRNA.R61	AGGGGGCGGCCGCCTTTCCGGCCGCGCCCCGTTTCCCAGGACGAAGGGCACTCCGCACCGGACCCCGGTCCCGGCGCGCG
28srRNA.R62	CGAAACCCCGACCCAGAAGCAGGTCGTCTACGAATGGTTTAGCGCCAGGTTCCCCACGAACGTGCGGTGCGTGACGGGCG
28srRNA.R63	GACAAACCCTTGTGTCGAGGGCTGACTTTCAATAGATCGCAGCGAGGGAGCTGCTCTGCTACGTA