Host plant genetic control of associated fungal and insect species in a Populus hybrid

Sandra Simon[1], Timothy Tschaplinski[2], Jared Leboldus[3], Ken Keefover-Ring[4], Muhammad Azeem[4], Jin-Gui Chen[2], David Macaya-Sanz[1], William MacDonald[1], Wellington Muchero[2], Gerald Tuskan[2], and Stephen DiFazio[1]

[1]West Virginia University
[2]Oak Ridge National Laboratory
[3]Oregon State University
[4]University of Wisconsin Madison

May 5, 2020

Abstract

Plants employ a diverse set of defense mechanisms to mediate interactions with insects and fungi. These relationships can leave lasting impacts on host plant genome structure such as rapid expansion of gene families through tandem duplication. These genomic signatures provide important clues about the complexities of plant/biotic stress interactions and evolution. We used a pseudo-backcross hybrid family to identify Quantitative Trait Loci (QTL) controlling associations between Populus trees and several common Populus diseases and insects. Using whole genome sequences from each parent, we identified candidate genes that may mediate these interactions. Candidates were partially validated using mass spectrometry to identify corresponding QTL for defensive compounds. We detected significant QTL for two interacting fungal pathogens and three insects. The QTL intervals contained candidate genes potentially involved in physical and chemical mechanisms of host-plant resistance and susceptibility. In particular, we identified overlapping QTLs for a phenolic glycoside and Phyllocolpa sawfly abundance. There was also significant enrichment of recent tandem duplications in the genomic intervals of the native parent, but not the exotic parent. Tandem gene duplication may be an important mechanism for rapid response to biotic stressors, enabling trees with long juvenile periods to reach maturity despite many coevolving biotic stressors.

Keywords

Community genetics – Comparative genomics – Plant-biotic interactions – *Populus* – Tandem duplication

1 INTRODUCTION

In natural ecosystems, the dynamics of plant interactions with other living organisms are complex. This is especially true for organisms that rely on plants for shelter, nutrition, and reproduction, such as fungi and insects (Chisholm, Coaker, Day, & Staskawicz, 2006; Panda & Khush, 1995). Although fungi and insects can provide some of the same useful services in return, those that do not can be extremely harmful to plant productivity. To mitigate the effects of biotic stress, plants employ a diverse set of defense mechanisms including chemical, protein-derived molecules, and physical barriers (Panda & Khush, 1995). Insects and fungi must develop strategies in parallel to overcome these obstacles to survive (Chisholm et al., 2006; Mello & Silva-Filho, 2002). The theory of gene-for-gene coevolution has frequently been used to describe this host plant genetic relationship to its arthropod and fungal communities (Chisholm et al., 2006; Ehrlich & Raven, 1964; Mello & Silva-Filho, 2002; Thompson, 1988).

The gene-for-gene theory suggests a very simple dynamic for the genetic interactions that occur between two species. A gene in the host plant that is important in biotic relationships has a corresponding, coevolving gene from a pathogen/insect which can lead to resistance or susceptibility depending on the life history of the pathogen/insect (Flor, 1971; Friesen, Meinhardt, & Faris, 2007). Much of the evidence for these interactions has been found in crop systems where plant species often have dominant, single-gene resistance to feeding (Thompson, 1988). For example, there are over twenty different genes in wheat (*Triticum aestivum* L.) that each confer resistance to the Hessian fly, *Mayetiola destructor* (Thompson & Burdon, 1992). Exposure of Hessian fly populations to these resistant varieties of wheat created selection pressure that led to increased virulent gene combinations in the pest (Gallun, 1977; Panda & Khush, 1995). Similarly, in plant-fungal systems breeding for dominant resistance in cereal crops resulted in new selective forces that increased virulent gene frequencies in *Puccinia* spp. cereal rusts (X. M. Chen, 2005; Pretorius, Singh, Wagoire, & Payne, 2000). This in turn can lead to an evolutionary arms race between plants, insects, and fungi with the continual development of mechanisms to overcome both genetic defenses and virulent attacks (Bergelson, Kreitman, Stahl, & Tian, 2001; Thompson & Burdon, 1992).

The relationships of host plant genetics and biotic association can also be more complex than these crop breeding systems suggest, and they can leave a lasting impact on genome structure (Lefebvre & Chèvre, 1995). Host plant and biotic associations can lead to the expansion of gene families responsible for the host plant response to biotic stress. For example, the Kunitz protease inhibitors (KPIs) in *Populus* are important in defense responses against insects by inhibition of herbivore digestion (Haruta, Major, Christopher, Patton, & Constabel, 2001; Major & Constabel, 2008). The KPI gene family has greatly expanded in response to insect attack through tandem duplication events (Philippe, Ralph, Külheim, Jancsik, & Bohlmann, 2009). Similarly, plant resistance (*R*) genes, which are important in the defense response of plants to fungal pathogen attack, have also expanded through tandem and segmental duplication events due to biotic pressures (Hulbert, Webb, Smith, & Sun, 2001; Leister, 2004). Analyzing how the genome is structured in the host plant when it associates with fungi and insects is important for studying these relationships and understanding the complexities of their genetic interactions.

Given their rapid growth and vegetative reproduction, *Populus* species have become a focus for research into biofuel production making them a valuable commercial crop (Meilan et al., 2002; Stanton, B. Neale, & Li, 2010; Taylor, 2002). *Populus* has also become an important genetic model for research into a wide variety of ecological and adaptive traits (McKown et al., 2014), including interactions with the biotic community (Crutsinger et al., 2014; Whitham et al., 2006). In particular, interspecific hybrids of *P. trichocarpa* x *P. deltoides* segregate for a wide variety of traits including resistance to insect and fungal attack (G. Newcombe, 1998; George Newcombe & Ostry, 2001b). Such hybrid family crosses can be used to identify regions of the genome that are important in mediating biotic stress.

In this study, we investigated the genome composition of loci associated with insect and fungal species in an interspecific *Populus* family. We surveyed insects and fungal pathogens in a *P. deltoides x P. trichocarpa* pseudo-backcross family and used quantitative trait locus (QTL) analysis and comparative genomics to address three main questions: (1) Is there heritable, host genetic control of fungal and insect species? (2) What protein domains and gene ontology terms are enriched in the QTL intervals in the genomes of each *Populus* species? (3) What candidate genes in the intervals are unique to each species when comparing the *P. trichocarpa* and *P. deltoides* genomes?

2 METHODS

2.1 QTL mapping pedigree

The 52-124 family was developed by crossing a male *Populus deltoides*, ILL-101 from southern Illinois, with a female *Populus trichocarpa* clone, 93-968 from western Washington State. The F_1 clone, 52-225, was then crossed with a male *Populus deltoides* clone, D124 from Minnesota, to generate the final 749 progeny in the pseudo-backcross population.

2.2 Field Site Descriptions

Clonal cuttings of the 52-124 progeny were obtained from the University of Minnesota and the University of Florida. The individuals were propagated in the West Virginia University (WVU) greenhouses in March 2006. A hay production field was tilled and disked at the WVU Agronomy farm (39°39'32"N 79deg54'19"W) in Morgantown, West Virginia for planting. In July 2007 the rooted cuttings were planted with two clonal replicates for each of 749 genotypes at 2 m x 2 m spacing in an interlocking block design. The plantation was thinned in December of 2008 by removing 50% of the trees in a diamond fashion resulting in 2.83 m x 4 m spacing of the remaining trees. A plantation of the same family cross was established at the Westport Research Station in the Columbia River floodplain (46deg07'58.8"N 123deg22'05.0"W) in April, 2016. The trial was planted with a total of 339 progeny replicated in a randomized complete block design with three blocks.

2.3 SNP genotyping and genetic map construction

SNP loci were selected from whole genome resequencing data generated for the parent trees of the pedigree, focusing on loci that were fixed for different alleles in *P. deltoides* and *P. trichocarpa* . This yielded loci that were maternally informative for the SNP derived from *P. trichocarpa* . These were incorporated into an Illumina Infinium Bead Array which was used to genotype 3,568 SNP loci in 692 of the progeny. This data was then used to create the genetic map which was composed of 19 linkage groups corresponding to 19 *Populus* chromosomes. Genotyping and map construction are described in more detail elsewhere (Muchero et al., 2015).

2.4 Family 52-124 Parent and Progeny Phenotyping

In order to identify regions of the *Populus* genome associated with biotic stress, a variety of fungal pathogens and insect herbivore species were surveyed in the WVU Morgantown plantation site. *Melampsora* sp. leaf rust (Figure 1a) and *Sphaerulina musiva* fungal leaf spot symptoms (Figure 1b) were scored on a 0-3 scale of disease severity, with 0 indicating absence of symptoms and 3 indicating high degree of pathogen leaf damage, in the 2008 growing season for all 1353 tree canopies. In the fall of 2014, stem canker symptoms caused by the same fungus *Sphaerulina musiva* (Figure 1c) were scored on a 0-5 disease severity scale for a subset 498 unique genotypes and a total of 580 trees. Upon further examination of field conditions for *S. musiva* disease severity it was determined that none of the progeny displayed complete resistance. The original 0-5 scale was binarized with scores from 2.5-5 scaled to 1 and 0-2 scaled to 0. The new scale indicated the progression of infection with 1 specifying severe canker development and 0 indicating less aggressive canker symptoms.

Two abundant species of galling aphid were also surveyed during the 2016 growing season. To equalize the biomass surveyed for the aphid insect observations on each tree, branches of equal diameter were selected for insect counts and the remaining branches in the canopy were not surveyed. Tree canopy presence or absence for the petiole galling aphid *Pemphigus populitransversus* (Figure 1d) and the branch galling aphid *Mordwilkoja vagabunda* (Figure 1e) was recorded for 201 unique genotypes and a total of 218 trees. For all of the damage scores in every survey year, 0 indicated a complete absence of fungal or insect presence in tree canopy.

In early August 2017 at the Westport site full canopies were scored by counting galls of the leaf folding sawfly *Phyllocolpa* sp. (Figure 1f), on 534 unique genotypes and a total of 1020 trees. To estimate productivity of individuals and confirm that availability of resources did not drive insect attraction or feeding, main stem diameter in millimeters was recorded for all trees to be used as a covariate in the analysis.

2.5 Leaf Metabolites

The metabolomic profiles for 211 genotypes of Family 52-124 were determined for leaves (leaf plastochron index 7 +- 2) sampled on September 14, 2006 at a USDA-FS field site near Grand Rapids, MI from plants in their first growing season in the field. Leaves were quickly collected and flash-frozen on dry ice, prior to shipping back to Oak Ridge National Laboratory, where ~200 mg fresh weight per sample were extracted in 80% ethanol (aqueous), as described in Zhao *et al.* , (2010) and analyzed for the major aromatic metabolites following trimethylsilylation and analysis by gas chromatography-mass spectrometry (GC-MS) using

3

electron impact ionization (70 eV). The analytical and data extraction protocols were similar to that reported previously in Weston *et al.* , (2015). The major phenolic metabolites quantified included salicortin, α-salicyloylsalicin, salicyl-salicylic acid-2-O-glucoside, salirepin (gentisyl alcohol 2-O-glucoside), gentisic acid 5-O-glucoside, 1,2,4-benzenetriol, and gentisyl alcohol 5-O-glucoside.

Statistical Analysis

2.6 Within-Family Broad-Sense Heritability (H^2) Calculation

Surveys of the *Melampsora* sp. leaf rust, *S. musiva* leaf spot and counts of the *Phyllocolpa* sp. galls had within-garden microsite variation removed using thin-plate spline regression (*fields* R package). Residuals of the spatial correction were added to the mean of each survey dataset for each tree observation to spatially correct and rescale the values. These corrected values were then used to calculate the proportion of variance in the fungal and insect distributions that was due to genotype using a linear-mixed model (*lmer* function of the *lme4* R package) with insect counts and fungal scores as the response, tree genotype as the predictor, and stem diameter biomass estimates as a covariate where applicable. Broad-sense heritability was calculated as $H^2 = \delta_g^2/(\delta_g^2 + \delta_e^2)$, where δ_g^2 is the genetic variance due to genotype and δ_e^2 is the residual variance. Rapid, simulation-based exact likelihood ratio tests were used to evaluate the significance of variation due to genotype for each linear model (*exactRLRT* function of the *RLRsim* R package).

SAS software version 9.4 (2013) was used to test for the normality of all data sets and for all subsequent statistical analyses, transformations were conducted when necessary. Data from all chemical assays were analyzed by single-factor ANOVAs (PROC GLM) to test for differences between low and high *Phyllocolpa* sp. abundance.

2.7 Quantitative Trait Loci (QTL) Analysis

The R software package *R/qtl* (Broman, Wu, Sen, & Churchill, 2003) was used for all QTL analyses described below. Composite interval mapping (CIM) was used to associate *Melampsora* sp. leaf rust,*S. musiva* leaf spot, and *Phyllocolpa* sp. to QTL positions (*cim* function). Single QTL mapping was used to associate the binary scores for the *S. musiva* canker, *M. vagabunda* and*P. populitransversus* to QTL positions (*scanone* function). The method used for both mapping approaches was the expectation-maximization (EM) algorithm. Estimation of QTL interval significance was completed by performing 1000 permutations. Intervals with logarithm of odds (LOD) scores that were above the p-value threshold (alpha = 0.05), as determined from the permutation tests, were selected for further analysis. The percent variance explained by significant markers for fungal and insect surveys, that were mapped using CIM, were calculated by extracting significant marker positions and creating a fit QTL model (*fitQTL* function). The positive allele contributing to an increase in susceptibility to fungi and insects was found by generating effect plots for each phenotype and its significant marker position (*effectplot* function).

2.8 Physical Genome Intervals

Physical genome intervals in the *P. trichocarpa* genome (v3.0) were examined for each significant QTL for biotic associations. The intervals were defined as 1 Mb regions centered on the marker with the highest LOD score. Fixed physical genome sizes were used rather than intervals defined based on LOD scores due to the large variation in magnitude of LOD observed for the significant QTL. For example, intervals of 1 LOD centered on the QTL ranged in size from 169 to 4620 kb. Much of this variation was likely due to variation in marker density and local recombination rates, in addition to phenotyping and genotyping error. We believe that a fixed 1 Mb interval is a more consistent and conservative approach. On average, this represents approximately 6.34 cM, based on a total map size of 2617 cM and a total assembled genome length of 420 Mb.

Orthologous intervals were identified in the *P. deltoides* clone WV94 reference genome (v2.1) obtained from Phytozome (Goodstein et al., 2012). Orthology was determined using a combination of protein sequence conservation and synteny using MCScanX (Wang et al., 2012). Briefly, all proteins were compared in all-vs-all searches using blastp both within genomes and between genomes. These were then chained into collinear

segments using the MCScanX algorithm. Orthologous segments were identified based on the presence of large numbers of gene pairs in collinear order with high sequence identity (median blastp E score <1e-180) (Figure 2). Synonymous (Ks) and nonsynonymous Ka) nucleotide substitution rates were calculated using the Bioperl DNAstatistics module (Stajich, 2002) (Table S1), domain composition (Table S2), and Gene Ontology (GO) terms (Table S3) were obtained for each genome from Phytozome (v12.1). Intervals were customized for the grandparents of the pseudo backcross progeny (clones 93-968 for *P. trichocarpa* and D124 for *P. deltoides*) using ˜150X of 2x250 paired end Illumina sequences. These were aligned to the respective reference genome for each species using bwa mem with default parameters. SNPs and small indels were identified using samtools mpileup (Li, 2011; Li et al., 2009), and sequence depth was extracted using vcftools (Danecek et al., 2011). Sequences were converted using the vcftools utility vcf-consensus. Genes with no coverage in the alignments were excluded from the intervals for each species.

2.9 Tandem Duplications

Tandemly duplicated genes were identified using all-vs-all blastp searches within each genome for biotic stress associated intervals (Table S4). Genes with blastp E scores < 1e-180 that were located within 500 kb of one another were considered to be recent tandem duplications. The window size was determined by testing a range of values and choosing a window size at which the number of newly-discovered tandem duplicates began to decline (Figure 3). The stringent E score cutoff was intended to focus the analysis on genes that are recently duplicated and therefore potentially differentially duplicated between the species. The QTL intervals were tested for significant enrichment of tandem duplicates by using a Monte Carlo simulation. Sets of contiguous genes equal in number to those contained in each QTL interval were randomly selected from the whole genome, and the number of sampled tandem duplications was counted for each iteration. This was repeated 10,000 times, and the observed number of tandem duplicates was compared to the simulated distribution to derive an empirical P-value.

3 RESULTS

3.1 Heritability of Fungal and Insect Associations

Clonal repeatability (or within-family broad-sense heritability) was estimated for each categorical survey trait (Table 1). There was a significant host plant genetic contribution of moderate effect controlling the association of *S. musiva* leaf spot disease severity ($H^2 = 0.227$, p-value < 0.0001). There was a strong effect of host plant genetics on associations of *Melampsora* sp. disease severity ($H^2 = 0.631$, p-value < 0.0001) and the continuous count of leaf-folding galls of *Phyllocolpa* sp. ($H^2 = 0.391$, p-value < 0.0001).

3.2 QTL Mapping of Fungal and Insect Surveys

Three QTL intervals containing a total of 26 markers were significantly associated with the abundance of fungal pathogens (Table 2, Figure 4a,b). An overlapping QTL on Chr04 (Figure 4a) was associated with both *Melampsora* sp. (marker position = 3.50687, p-value = 0.001) and *S. musiva* leaf spot disease severity (marker position = 15.4912824, p-value = 0.026). The marker with the highest LOD score explained 54.1% of the variance in the disease severity of *Melampsora* sp. within the Chr04 QTL interval, whereas the top marker within the overlapping interval only explained 3.31% of the variance in the disease severity of *S. musiva* . The positive allele contributing to an increase in fungal infection for both species was derived from *P. trichocarpa* . Binary presence of the canker symptoms caused by *S. musiva* (Figure 4b) was found to be associated with a QTL located on Chr16 (marker position = 60.81429, p-value = 0.022). The positive allele contributing to increase in presence of *S. musiva* canker was from *P. trichocarpa* .

Three QTL intervals containing a total of 40 markers were significantly associated with insect abundance (Table 2, Figure 4c,d). A QTL on Chr05 (marker position = 21.83874, p-value = 0.001) was associated with binary presence of *M. vagabunda* branch galls (Figure 4c). The positive allele contributing to an increase in presence of *M. vagabunda* was from *P. trichocarpa* . Two QTLs, one located on Chr10 (marker position = 57.69914343, p-value = 0.015) and one located on Chr13 (marker position = 84.43857329, p-value = 0.045) were associated with the number of leaf-folding galls from *Phyllocolpa* sp. (Figure 4d). The marker with

the highest LOD score explained 9.65% of the variance in oviposition gall count while 8.82% of the variance was explained by the top marker located on Chr13. For both markers, the positive allele contributing to an increase in the number of female oviposition galls was from *P. deltoides* . There were no QTL intervals that passed the permutation threshold for the binary presence of *P. populitransversus* petiole galls (Figure 4c).

3.3 QTL Mapping of Leaf Metabolites

Only one tested compound, gentisyl alcohol 5-O-glucoside levels, yielded significant QTL. Three QTL intervals containing a total of 56 significant markers were significantly associated with gentisyl alcohol 5-O-glucoside levels (Table 2, Figure 4d). This included markers on chromosome 17 (marker position = 21.54739208, p-value < 0.0001) which explained 25.8% of variation, Chr10 (marker position = 109.8469079, p-value < 0.0001) which explained 23.1% of variation, and chromosome 14 (marker position = 91.51828521, p-value = 0.0100) which explained 7.30% of variation. The positive allele that contributed to an increase in gentisyl alcohol 5-O-glucoside levels for all markers was derived from *P. trichocarpa* .

3.4 Interaction between *Melampsora* sp. and *S. musiva* Leaf Infection

Melampsora sp. leaf rust infection severity was dependent upon the disease severity of *S. musiva* leaf spot symptoms (F = 35.2, p-value = 0.0001) (Table 3). The severity of *Melampsora* sp. infection for individuals with no *S. musiva* leaf spot was significantly higher than for individuals with *S. musiva* leaf spot symptoms (Figure 5). Similarly, *S. musiva* leaf spot disease severity was dependent upon the levels of *Melampsora* sp. leaf rust (F = 31.5, p-value = 0.0001). *S. musiva* leaf spot infection severity was significantly lower for individuals with a*Melampsora* sp. leaf spot score of 3 compared to individuals with less severe *Melampsora* sp. leaf spot symptoms (Figure 5).

3.5 PFAM, GO Term and Tandem Duplication Enrichment

The total number of genes present in the QTL intervals of each parental species was 174 on Chr04, 81 on Chr16, 161 on Chr10, 445 on Chr13, and 156 on Chr05 (Table 4). There were 40 PFAM domains enriched relative to the rest of the *P. trichocarpa* genome across all QTL intervals (alpha = 0.05; p-value threshold < 0.00014), and 25 PFAM domains enriched in the *P. deltoides* QTL intervals (alpha = 0.05, p-value < 0.000059) (Table S2). Additionally, there were 21 GO terms enriched in the *P. trichocarpa* intervals (alpha = 0.05, p-value < 0.000226) and 46 GO terms enriched in the *P. deltoides* intervals (alpha = 0.05, p-value < 0.000214) (Table S3). Finally seven of the genes in these intervals showed evidence of positive selection based on the ratio of nonsynonymous to synonymous nucleotide substitutions between the *P. deltoides* and *P. trichocarpa* orthologs (Table 5).

In total, there were 107 recent tandem duplicates in the *P. trichocarpa* intervals and 195 in the *P. deltoides* intervals for all biotic QTLs (Figure 6, Table 6). This includes 23 and 41 recent tandem duplicates for the *Melampsora sp.* intervals; 6 and 10 for the *S. musiva* intervals; 32 and 96 for the *Phyllocolpa* sp. intervals; and 47 and 46 for the *M. vagabunda* intervals, for*P. trichocarpa* and *P. deltoides* , respectively. The total number of recent tandem duplicates was significantly enriched relative to tandem counts for random intervals of the same size as the QTL intervals for the *P. deltoides* grandparent (p-value = 0.0118) but not for the *P. trichocarpa* grandparent (p-value = 0.1191).

4 DISCUSSION

The goal of our research was to utilize QTL analysis as a tool to identify regions of the *Populus* genome that were important in mediating biotic interactions. Upon identification of these regions we were able to directly compare the parental genomes of the hybrid cross to look for similarity in content and potential gene-for-gene interactions reflected in recent tandem duplication expansion. We found that the host plant genotype had a significant effect on fungi and insects in our study. The progeny segregated for varying resistance to fungal and insect pathogens and pests that were inherited from the grandparents and parents of the 52-124 family cross. Novel alleles (i.e. those from the grandparent that was not native to the region where the trials occurred) were important for resistance to *Melampsora* sp. fungus, *M. vagabunda* galls, and *Phyllocolpa* sp. oviposition. In contrast, the interval associated with *S. musiva*canker symptoms was the

only case in which susceptibility to the fungus was dominant and inherited from the non-coevolved host (*P. trichocarpa*), as has been previously observed (Muchero et al., 2018; George Newcombe & Ostry, 2001a).

We detected a major QTL for *Melampsora sp.* resistance on Chr04. This region is known to contain the *MXC3* locus which confers resistance to infection of many species of the *Melampsora* leaf rusts (George Newcombe, Stirling, & Bradshaw, 2001; Yin et al., 2004). Based on mean parental infection scores and the allelic effects at the QTL, progeny in family 52-124 inherited this resistance from the *P. trichocarpa* grandmother 93-968. Furthermore, the QTL interval contained two tandem repeats of stigma-specific proteins (Stig1) in both the *P. trichocarpa* genome and the orthologous interval in the *P. deltoides* genome. Several of these showed evidence of positive selection based on Ka/Ks ratios (Table 5). Stigma-specific proteins, specifically Stig1, have been found to be associated with female sterility in tobacco (*Nicotiana tabacum*) and petunia (*Petunia hybrida*) (Goldman, Goldberg, & Mariani, 1994; Verhoeven et al., 2005). Stig1 is known to mediate secretion of exudate lipids in the intercellular spaces and high expression of the protein inhibits pollen grains from penetrating style tissue preventing fertilization (Verhoeven et al., 2005). Most lipid transfer proteins, such as Stig1, are important in plant cell-wall loosening and their expression can prevent penetration of plant tissues (Nieuwland et al., 2005). Diversification of the protein family containing Stig1 may play an important role in lowering fungal infection in *Populus* by providing a physical barrier to resist the *Melampsora* sp. hyphae. Another protein domain that was found to be enriched in the *P. trichocarpa* Chr04 interval was the malonyl-CoA decarboxylase C-terminal domain. Similarly, the gene ontology function for malonyl-CoA decarboxylase activity was also enriched in *P. trichocarpa* in the same interval. Genes which are capable of transforming malonyl-CoA could be beneficial in resistance to *Melampsora* sp. as it is an important precursor in the production of several pathogen defensive compounds, such as isoprenoids in the Mevalonate pathway (H. Chen, Kim, Weng, & Browse, 2011; Dixon, 2001).

Interestingly, an overlapping interval on Chr04 was found to be associated with the activity of the leaf spot symptoms of the *S. musiva* fungus. Upon further investigation we found that trees that were not infected by the *S. musiva* leaf spot had more severe symptoms of the *Melampsora* sp. leaf rust. This suggests that a competitive interaction occurred between the two pathogens in the field during the year of survey, and was reflected in a false association on Chr04 with the *S. musiva* leaf spot score. Competition among fungal pathogens is not uncommon in field conditions with most examples focused on different genotypes within the same fungal species (Abdullah et al., 2017). The outcome of within host tissue colonization of multiple strains or species often relies upon the genetic similarity of the pathogens (Abdullah et al., 2017; Koskella, Giraud, & Hood, 2006). In our case, we had two very different pathogens that were utilizing the leaf tissue in vastly dissimilar ways. *S. musiva* is a necrotrophic fungus, which requires dead tissue to reproduce, and *Melampsora* sp. is biotrophic, requiring living tissue to reproduce. The presence of *S. musiva* on the trees had a much larger effect on the occurrence of *Melampsora* sp. infection. This may indicate that the *S. musiva* fungus had a competitive advantage over *Melampsora* sp. in utilization of the *Populus* leaf tissue at our site. Further supporting this finding was the complete loss of *Melampsora* sp. symptoms at the plantation site over the course of ten years and the continual presence of the *S. musiva* leaf spot. A similar interaction has been reported in wheat between the biotroph *Blumeria graminis* f.sp. *tritici* (*Bgt*), the powdery mildew pathogen, and necrotroph *Zymoseptoria tritici* , the cause of *Septoria tritici* blotch (Orton & Brown, 2016). Similarly, the outcome of the interaction of the two pathogens in wheat was competitive. However, the necrotroph was found to actually be capable of reducing the reproductive capability of the biotroph which indicates pathogen-pathogen interactions can be more direct rather than relying solely on the host plant genetics (Orton & Brown, 2016).

The *S. musiva* stem canker symptoms were found to be associated with a QTL interval on Chr16. In this study, susceptibility to the necrotroph appeared to be originating from the presence of *P. trichocarpa* alleles in the progeny. Previous work on a similar *Populus* hybrid cross supports this with susceptibility to *S. musiva* stem canker originating primarily from dominant alleles derived from *P. trichocarpa* (G. Newcombe, 1998; George Newcombe & Ostry, 2001a). Interestingly, the loci conferring susceptibility in family 52-124 did not overlap with the four loci that were uncovered in a previous genome-wide association study of *S. musiva* susceptibility in *P. trichocarpa* (Muchero et al., 2018), suggesting that different mechanisms may

7

be involved in hybrid interactions with this pathogen. However, the QTL did contain a tandem repeat of a G-type lectin receptor-like protein kinase that was expanded in *P. deltoides* . This protein could play a similar role to a receptor-like kinase from the same family (Yang et al., 2016) that was associated with susceptibility to *S. musiva* in the *P. trichocarpa* study (Muchero et al., 2018).

M. vagabunda has been recorded completing its life cycle on several species of *Populus* including *P. deltoides* and *P. tremuloides* (Floate, 2010). Although the aphid's life history has been documented, little is known about the influence of host plant genetics on gall formation or resistance to feeding (Floate, 2010; Ignoffo & Granovsky, 1961). We detected a QTL on Chr05 in which alleles inherited from *P. deltoides* were positively associated with gall occurrence. Genes conferring lipoxygenase and oxidoreductase activity were enriched in both the *P. deltoides* as well as the *P. trichocarpa* QTL intervals. Lipoxygenase genes are known to be associated with the *Populus* response to both abiotic and biotic stressors (Cheng et al., 2006; Ralph et al., 2006). They are often upregulated in the presence of mechanical damage, fungal pathogen invasion, and exposure to simulated insect feeding (F. Chen, Liu, Tschaplinski, & Zhao, 2009; Cheng et al., 2006). The lipoxygenases are important in the formation of jasmonic acid, the signaling molecule that upregulates plant defenses against herbivore feeding (F. Chen et al., 2009).

The *M. vagabunda* QTL interval on Chr05 also contained a tandem array of resistance genes (R-genes) that encoded disease resistance proteins (TIR-NBS-LRR class) that was greatly expanded in *P. trichocarpa* compared to *P. deltoides* , as well as repeats of the leucine-rich repeat protein kinase family proteins in both species. These protein families are well known for their roles in the recognition and upregulation of host plant defenses against bacterial and fungal infection (Bergelson et al., 2001; Martin, Bogdanove, & Sessa, 2003). Diversification of their function through tandem duplication due to gene-for-gene coevolution has also been demonstrated in many plant-fungal pathosystems (Leister, 2004). Although R-genes have not been recognized as frequently in plant-insect interactions, those that have been discovered typically mediate interactions with piercing-sucking insects such as aphids and whiteflies (Kaloshian, 2004).

In addition to the R-genes, there were a series of genes encoding cytochrome P450 family proteins in both species as well as a unique tandem set that was only present in the *P. trichocarpa* genome. P450 enzymes are important in the production of many classes of secondary metabolites such as furanocoumarins and terpenoids which are highly toxic to insects (Keeling & Bohlmann, 2006; Schuler, 2011). Alternatively, cytochrome P450's may also be implicated in the susceptibility of the host plant to galling aphids. They are important in the synthesis of fatty acids and production of suberin in plant tissues (Höfer et al., 2008; Pinot & Beisson, 2011). Typically, suberin is important in separation of different tissues as well as in the establishment of apoplastic barriers that restrict nutrient/water loss as well as pathogen invasion (Höfer et al., 2008; Qin & LeBoldus, 2014). Several insects are known to produce suberized spherical galls on leaves including hymenopteran pests of Rosacea species and dipteran pests of Fabaceae (Krishnan & Franceschi, 1988; Oliveira et al., 2016). If aphids induce the suberization mechanism of the plant genome it may lead to the increased toughening of aphid galls, much like the woody structures *M. vagabunda* leaves behind on branches once they have finished feeding.

We detected two QTL for *Phyllocolpa* sp. with *P. trichocarpa* alleles being positively associated with leaf fold occurrence in both cases. The *P. trichocarpa* genomic interval corresponding to the Chr10 QTL was enriched for several domains and GO terms that could be involved in gall development. For example, the interval was enriched for sugar transporters as well as GO terms for sucrose transporter activity. Often in the case of galling insects plant tissue is modified in such a way as to act as a sugar sink, thereby enhancing its nutritional value to larvae (Larson & Whitham, 1991; Nyman, Widmer, & Roininen, 2000; Wool, 2004) . The presence of these combinations of sugar transporter genes may be mediating a similar interaction between the *Phyllocolpa* sp. female sawflies and their chosen *Populus* hosts.

Phyllocolpa sp. galls are formed early in the season when a female sawfly selects a leaf and injects the longitudinal fold with small amounts of fluid on the underside of young leaves (Fritz & Price, 1988; Kopelke, 2007). The adult sawfly will proceed to oviposit near the base of the leaf and after 1-2 days the leaf fold forms and the newly hatched larvae feeds on the inside of the gall (Smith & Fritz, 1996). The *Phyllocolpa* sp. galls

8

were a unique biotic phenotype to this study as they were an estimate of female sawfly ovipositional choice rather than feeding success. Host selection for oviposition is initially driven by visual cues and reinforced by females assessing the nutrition and chemical cues of foliage (Boeckler, Gershenzon, & Unsicker, 2011; Panda & Khush, 1995). Previous research in *Salix* (closely related to *Populus*) has shown that common phenolic glycosides in leaf tissue are important in host choice in both the free-feeding *Nematus oligospilus* and galling *Euura amerinae* specialist sawflies (Fernández et al., 2019; Kolehmainen, Roininen, Julkunen-Tiitto, & Tahvanainen, 1994). We therefore pursued QTL mapping of metabolites to determine if phenolic compounds might be important determinants of the potential ovipositional relationship.

We detected three significant QTL for gentisyl alcohol 5-O-glucoside levels. One of these overlapped with the *Phyllocolpa* sp. leaf fold QTL on Chr10. The QTL on Chr14 also overlapped with a suggestive QTL peak for *Phyllocolpa sp* . (LOD=3.96). This is the first case of an association of gentisyl alcohol 5-O -glucoside with a specialist arthropod in *Populus* , and is, in fact, the first report of this metabolite in *Populus* . Salirepin (gentisyl alcohol 2-O -glucoside), a closely related metabolite, is a well-known constituent in *Populus* sp, (Busov et al., 2006; Tschaplinski et al., 2019; Veach et al., 2018) and leaves of *P. deltoides* and *P. trichocarpa x deltoides* have a lower abundance, later eluting metabolite with a nearly identical fragmentation pattern to salirepin that we tentatively identify as gentisyl alcohol 2-O -glucoside. The QTL contains three putative candidate genes, including an aldehyde dehydrogenase 5F1 (Podel.10G175800) that may be involved in the reduction of gentisyl aldehyde to the alcohol, and two UDP-glycosyltransferases (Podel.10G184800, Podel.10G185000) that may be involved in the gentisyl alcohol conjugation to glucose. Specific substrates have yet to be determined for these genes, but a previous report suggests aldehyde dehydrogenase 5F1 genes are likely involved in the basic metabolism

of *Populus* (Tian et al., 2015).

Phyllocolpa sp. sawflies are considered a keystone species as the abandoned or unused leaf folds are often used as a habitat for many other species such as aphids and spiders (Bailey & Whitham, 2007). The presence of folds in aspen forests is associated with a two-fold increase in arthropod species richness and around a four-fold increase in arthropod abundance relative to forests where the insect is absent (Bailey & Whitham, 2003). This in turn makes the host plant and sawfly relationship important in examining how shifts in the genes of a population ultimately structure whole communities, effectively linking ecology and evolutionary biology. Further investigation of this potential relationship could be key to connecting *Populus* genetics to the assemblage of the surrounding communities of organisms.

A striking finding in this study was enrichment of recent tandem duplications in the *P. deltoides* genome but not the *P. trichocarpa* genome for the biotic QTL intervals. Out of the six chromosomes that yielded significant QTL results, four were associated with phenotypes that were fungi and insects native to the distribution of *P. deltoides* , but not *P. trichocarpa* . Given that *P. deltoides* has been co-evolving with the majority of the surveyed fungi and insects, it was not unexpected that there were more recent tandem duplicates in biotic intervals in its genome as there is more selective pressure on the native species to overcome biotic stress (Constabel & Lindroth, 2010; George Newcombe, Martin, & Kohler, 2010). However, given the high amount of novel resistance occurring in the progeny, recent tandem duplication may also be important in naïve host resistance.

Our study has demonstrated that many recent tandem duplications, found across biotic stress QTL intervals, have functional annotations that are involved in physical, chemical, and tolerance mechanisms of host plant resistance as well as a few that may be implicated in host plant susceptibility. The enrichment of recent tandem duplications may be a signature of gene-for-gene interactions, and a mechanism that protects long-lived plants such as trees, enabling them to reach maturity despite many coevolving biotic stressors. In addition to the direct effect of the host plant genetics on associated organisms, we have also demonstrated an indirect community effect that was mediated by the *Populus* progeny (Whitham et al., 2006). The fungal pathogen succession highlights the complexities of how host plant-pathogen coevolution can affect multiple species interactions in more subtle ways than expected in a natural environment.

ACKNOWLEDGEMENTS

CONFLICT OF INTEREST

None declared.

AUTHORS' CONTRIBUTIONS

S.S., S.D., and D.M.-S. identified, collected, and analyzed insect surveys. W.M., J.L., J.C., and W.M. identified and collected fungal species datasets. K.K.-R., M.A., and T.T. collected/analyzed metabolite data. S.S. and S.D. analyzed the data. W.M., G.T. and S.D. designed and supervised the study. S.S., S.D., J.L., T.T., and K.K.-R. wrote the manuscript.

DATA ACCESSIBILITY

DNA sequences are deposited at NCBI Sequence Read Archive under the following accession numbers: SRP112750, SRP109803 and SRP190316. The annotated *P. deltoides* clone WV94 genome assembly is available through Phytozome (https://phytozome.jgi.doe.gov/pz/portal.html).

Table 1. Linear-mixed model output for biotic surveys. Broad-sense heritability (H^2) denotes the contribution of all host plant genetic factors to total variance in the biotic phenotype. R package RLRsim exactRLRT function was used to test significance of effects in mixed-model.

Fungus/Insect	H^2	Genetic variance	Error variance	RLRatio	p-value
Melampsora sp. leaf rust	0.622	0.631	0.382	294	< 0.0001
S. musiva leaf spot	0.227	0.099	0.337	32.0	< 0.0001
Phyllocolpa sp.	0.391	11.9	18.5	129	< 0.0001

Table 2. Summary of QTL permutation test output. Percent variance in surveys for nsects and fungi explained by significant marker indicated for composite interval mapping models. Positive (+) allele specifies genotype at significant interval that results in an increase in susceptibility. D indicates progeny are homozygous for *P. deltoides* alleles and T indicates progeny are heterozygous for *P. deltoides* and *P. trichocarpa* alleles.

Model	Chrom.	Position	LOD score	p-value	% variance	+ allele	Garden
Melampsora sp. leaf rust	Chr04	3.50	15.8	0.001	54.1	D	Morgantown
S. musiva leaf spot	Chr04	15.49	4.75	0.026	3.31	D	Morgantown
S. musiva canker	Chr16	60.81	4.34	0.022	NA	T	Morgantown
M. vagabunda	Chr05	21.83	5.39	0.001	NA	D	Morgantown

Model	Chrom.	Position	LOD score	p-value	% variance	+ allele	Garden
P. populi-transversus	Chr03	83.06	2.70	>0.05	NA	NA	Morgantown
Phyllocolpa sp.	Chr10	57.69	6.60	0.015	9.65	T	Westport
Phyllocolpa sp.	Chr13	84.43	4.99	0.045	8.82	T	Westport
Gentisyl alcohol 5-O-glucoside	Chr17	21.54	12.2	<0.0001	25.760	T	Grand Rapids
Gentisyl alcohol 5-O-glucoside	Chr10	109.84	10.3	<0.0001	23.104	T	Grand Rapids
Gentisyl alcohol 5-O-glucoside	Chr14	91.51	5.96	0.0100	7.296	T	Grand Rapids

Table 3. One-way ANOVA, with Genotype as a covariate, analyzing the effect of infection severity of the genus of one fungus on the infection severity of the competing leaf fungi in 2008.

Factor	DF	Sum of squares	Mean square	F-ratio	p-value
Dependent Variable- *Melampsora* sp. leaf rust infection	Dependent Variable- *Melampsora* sp. leaf rust infection	Dependent Variable- *Melampsora* sp. leaf rust infection	Dependent Variable- *Melampsora* sp. leaf rust infection	Dependent Variable- *Melampsora* sp. leaf rust infection	Dependent Variable- *Melampsora* sp. leaf rust infection
Genotype	732	1120	1.53	4.69	<0.0001
S. musiva score	3	34.4	11.5	35.2	<0.0001
Residuals	602	197	0.326		
Dependent Variable- *S. musiva* leaf spot infection	Dependent Variable- *S. musiva* leaf spot infection	Dependent Variable- *S. musiva* leaf spot infection	Dependent Variable- *S. musiva* leaf spot infection	Dependent Variable- *S. musiva* leaf spot infection	Dependent Variable- *S. musiva* leaf spot infection
Genotype	732	379	0.518	1.77	<0.0001
Melampsora sp. score	3	27.6	9.19	31.5	<0.0001
Residuals	602	176	0.292		

Table 4. Number of genes in QTL intervals in parental genomes for biotic associations.

Fungus/Insect	QTL Interval	# Genes in *P. trichocarpa*	# Genes in *P. deltoides*
Melampsora sp.	Chr04	159	147
S. musiva canker	Chr16	71	68
Phyllocolpa sp.	Chr10	146	154
Phyllocolpa sp.	Chr13	252	425
M. vagabunda	Chr05	129	147

11

Table 5. Candidate genes under positive selection (Ka/Ks>1) in genetic intervals associated with fungi and insects.

QTL	Arabidopsis function	*P. trichocarpa* gene	*P. deltoides* gene	Ka/Ks
Chr04	Chitin elicitor receptor kinase 1	Potri.004G005800	Podel.04G004900.1.p	1.35
Chr04	Cysteine-rich RLK (RECEPTOR-like protein kinase)	Potri.004G012600	Podel.04G010900.1.p	2.23
Chr04	Stigma-specific Stig1 family protein	Potri.004G006800	Podel.04G005700.1.p	1.31
Chr04	Stigma-specific Stig1 family protein	Potri.004G007200.1	Podel.04G006100.1.p	1.00
Chr04	Stigma-specific Stig1 family protein	Potri.004G007200	Podel.04G006100.1.p	1.00
Chr05	Pentatricopeptide repeat (PPR) superfamily protein	Potri.005G038500.1	Podel.05G042200.1.p	1.52
Chr10	Copper amine oxidase family protein	Potri.010G088700	Podel.10G084500.1.p	1.19

Table 6. Tandem duplication profiles for genetic intervals. Number of copies next to species gene name indicates the size of tandem expansion for the gene.

Arabidopsis function	Arabidopsis function
ERD (early- responsive to dehydration stress) family protein	ERD (early- responsive to dehydration stress) family p
Pentatricopeptide repeat (PPR) superfamily protein	Pentatricopeptide repeat (PPR) superfamily protein
Peroxidase superfamily protein	Peroxidase superfamily protein
Disease resistance protein (TIR-NBS- LRR class)	Disease resistance protein (TIR-NBS- LRR class)
Shikimate O- Hydroxycinnamoyl- transferase	Shikimate O- Hydroxycinnamoyl- transferase
Protein tyrosine kinase (Pkinase_Tyr) Leucine rich repeat (LRR_8)	Protein tyrosine kinase (Pkinase_Tyr) Leucine rich rep
Lipoxygenase	Lipoxygenase
Cytochrome P450, family 721, subfamily A, polypeptide 1	Cytochrome P450, family 721, subfamily A, polypepti
Cytochrome P450, family 76, subfamily G, polypeptide 1	Cytochrome P450, family 76, subfamily G, polypeptid
Cytochrome P450, family 93, subfamily D, polypeptide 1	Potri.005G037100
Cytochrome P450, family 76, subfamily G, polypeptide 1	NA
Disease resistance protein (TIR-NBS class), putative	Potri.005G032000
Receptor-like kinase in flowers	Potri.005G040200
Pentatricopeptide repeat (PPR) superfamily protein	Potri.010G083800
Copper amine oxidase family protein	Potri.010G088800
Ankyrin repeat family protein	Potri.013G133400
O-methyltransferase family protein	Potri.013G143700
O-methyltransferase family protein	NA

12

Arabidopsis function	Arabidopsis function
G-type lectin receptor-like protein kinase	Potri.016G102500

REFERENCES

Abdullah, A. S., Moffat, C. S., Lopez-Ruiz, F. J., Gibberd, M. R., Hamblin, J., & Zerihun, A. (2017). Host–multi-pathogen warfare: pathogen interactions in co-infected plants. *Frontiers in Plant Science* , *8* , 1806. https://doi.org/10.3389/fpls.2017.01806

Bailey, J. K., & Whitham, T. G. (2003). Interactions among elk, aspen, galling sawflies and insectivorous birds. *Oikos* , *101* (1), 127–134. https://doi.org/10.1034/j.1600-0706.2003.12185.x

Bailey, J. K., & Whitham, T. G. (2007). Biodiversity is related to indirect interactions among species of large effect. In *Ecological Communities: Plant Mediation in Indirect Interaction Webs* (pp. 306–328). Cambridge: Cambridge University Press.

Bergelson, J., Kreitman, M., Stahl, E. A., & Tian, D. (2001). Evolutionary dynamics of plant R-Genes. *Science* ,*292* (5525), 2281–2285. https://doi.org/10.1126/science.1061337

Boeckler, G. A., Gershenzon, J., & Unsicker, S. B. (2011). Phenolic glycosides of the Salicaceae and their role as anti-herbivore defenses.*Phytochemistry* , *72* (13), 1497–1509. https://doi.org/10.1016/j.phytochem.2011.01.038

Broman, K. W., Wu, H., Sen, S., & Churchill, G. A. (2003). R/qtl: QTL mapping in experimental crosses. *Bioinformatics* , *19* (7), 889–890. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/12724300

Busov, V., Meilan, R., Pearce, D. W., Rood, S. B., Ma, C., Tschaplinski, T. J., & Strauss, S. H. (2006). Transgenic modification of gai or rgl1 causes dwarfing and alters gibberellins, root growth, and metabolite profiles in *Populus* . *Planta* , *224* (2), 288–299. https://doi.org/10.1007/s00425-005-0213-9

Chen, F., Liu, C.-J., Tschaplinski, T. J., & Zhao, N. (2009). Genomics of secondary metabolism in *Populus* : interactions with biotic and abiotic environments. *Critical Reviews in Plant Sciences* ,*28* (5), 375–392. https://doi.org/10.1080/07352680903241279

Chen, H., Kim, H. U., Weng, H., & Browse, J. (2011). Malonyl-CoA synthetase, encoded by ACYL AC-TIVATING ENZYME13, is essential for growth and development of Arabidopsis. *The Plant Cell* , *23* (6), 2247–2262. https://doi.org/10.1105/tpc.111.086140

Chen, X. M. (2005). Epidemiology and control of stripe rust [*Puccinia striiformis* f. sp. *tritici*] on wheat.*Canadian Journal of Plant Pathology* , *27* (3), 314–337. https://doi.org/10.1080/07060660509507230

Cheng, Q., Zhang, B., Zhuge, Q., Zeng, Y., Wang, M., & Huang, M. (2006). Expression profiles of two novel lipoxygenase genes in*Populus deltoides* . *Plant Science* , *170* (6), 1027–1035. https://doi.org/10.1016/J.PLANTSCI.2005.12.012

Chisholm, S. T., Coaker, G., Day, B., & Staskawicz, B. J. (2006). Host-microbe interactions: shaping the evolution of the plant immune response. *Cell* , *124* (4), 803–814. https://doi.org/10.1016/J.CELL.2006.02.008

Constabel, C. P., & Lindroth, R. L. (2010). The impact of genomics on advances in herbivore defense and secondary metabolism in*Populu* s. In S. Jansson, R. Bhalerao, & A. Groover (Eds.),*Genetics and Genomics of* Populus (pp. 279–305). NY: Springer New York.

Crutsinger, G. M., Rudman, S. M., Rodriguez-Cabal, M. A., McKown, A. D., Sato, T., MacDonald, A. M., . . . El-Sabaawi, R. W. (2014). Testing a 'genes-to-ecosystems' approach to understanding aquatic-terrestrial linkages. *Molecular Ecology* , *23* (23), 5888–5903. https://doi.org/10.1111/mec.12931

Danecek, P., Auton, A., Abecasis, G., Albers, C. A., Banks, E., DePristo, M. A., . . . Durbin, R. (2011). The variant call format and VCFtools. *Bioinformatics* , *27* (15), 2156–2158. https://doi.org/10.1093/bioinformatics/btr330

Dixon, R. A. (2001). Natural products and plant disease resistance.*Nature* , *411* (6839), 843–847. https://doi.org/10.1038/35081178

Ehrlich, P. R., & Raven, P. H. (1964). Butterflies and plants: a study in coevolution. *Evolution* , *18* (4), 586–608. https://doi.org/10.2307/2406212

Fernández, P. C., Braccini, C. L., Dávila, C., Barrozo, R. B., Aráoz, M. V. C., Cerrillo, T., . . . Zavala, J. A. (2019). The use of leaf surface contact cues during oviposition explains field Preferences in the willow sawfly *Nematus oligospilus* . *Scientific Reports* ,*9* (1). https://doi.org/10.1038/s41598-019-41318-7

Floate, K. D. (2010). Gall-inducing aphids and mites associated with the hybrid complex of cottonwoods, *Populus* spp. (Salicaceae) on Canada's grasslands. *Arthropods of Canadian Grasslands* , *1* , 281–300. https://doi.org/10.3752/9780968932148.ch13

Flor, H. H. (1971). Current status of the gene-for-gene concept.*Annual Review of Phytopathology* , *9* (1), 275–296. Retrieved from www.annualreviews.org

Friesen, T. L., Meinhardt, S. W., & Faris, J. D. (2007). The*Stagonospora nodorum* -wheat pathosystem involves multiple proteinaceous host-selective toxins and corresponding host sensitivity genes that interact in an inverse gene-for-gene manner. *Plant Journal* , *51* (4), 681–692. https://doi.org/10.1111/j.1365-313X.2007.03166.x

Fritz, R. S., & Price, P. W. (1988). Genetic variation among plants and insect community structure: willows and sawflies. *Ecology* ,*69* (3), 845–856.

Gallun, R. L. (1977). Genetic basis of Hessian fly epidemics.*Annals of the New York Academy of Sciences* , *287* (1), 223–229. https://doi.org/10.1111/j.1749-6632.1977.tb34241.x

Goldman, M. H., Goldberg, R. B., & Mariani, C. (1994). Female sterile tobacco plants are produced by stigma-specific cell ablation. *The EMBO Journal* , *13* (13), 2976–2984. https://doi.org/10.1002/J.1460-2075.1994.TB06596.X

Goodstein, D. M., Shu, S., Howson, R., Neupane, R., Hayes, R. D., Fazo, J., . . . Rokhsar, D. S. (2012). Phytozome: a comparative platform for green plant genomics. *Nucleic Acids Research* , *40* (D1), D1178–D1186. https://doi.org/10.1093/nar/gkr944

Haruta, M., Major, I. T., Christopher, M. E., Patton, J. J., & Constabel, C. P. (2001). A Kunitz trypsin inhibitor gene family from trembling aspen (*Populus tremuloides* Michx.): cloning, functional expression, and induction by wounding and herbivory. *Plant Molecular Biology* , *46* (3), 347–359. https://doi.org/10.1023/A:1010654711619

Hofer, R., Briesen, I., Beck, M., Pinot, F., Schreiber, L., & Franke, R. (2008). The *Arabidopsis* cytochrome P450 CYP86A1 encodes a fatty acid ω-hydroxylase involved in suberin monomer biosynthesis.*Journal of Experimental Botany* , *59* (9), 2347–2360. Retrieved from https://academic.oup.com/jxb/article-abstract/59/9/2347/461748

Hulbert, S. H., Webb, C. A., Smith, S. M., & Sun, Q. (2001). Resistance gene complexes: evolution and utilization. *Annual Review of Phytopathology* , *39* (1), 285–312. https://doi.org/10.1146/annurev.phyto.39.1.285

Ignoffo, C. M., & Granovsky, A. A. (1961). Life history and gall development of Mordwilkoja vagabunda (Homoptera: Aphidae) on*Populus deltoides* . *Annals of the Entomological Society of America* , *54* (4), 486–499. https://doi.org/10.1093/aesa/54.4.486

Kaloshian, I. (2004). Gene-for-gene disease resistance: bridging insect pest and pathogen defense. *Journal of Chemical Ecology* , *30* (12), 2419–2438. https://doi.org/10.1007/s10886-004-7943-1

Keeling, C. I., & Bohlmann, J. (2006). Genes, enzymes and chemicals of terpenoid diversity in the constitutive and induced defence of conifers against insects and pathogens. *New Phytologist* , *170* (4), 657–675. https://doi.org/10.1111/j.1469-8137.2006.01716.x

Kolehmainen, J., Roininen, H., Julkunen-Tiitto, R., & Tahvanainen, J. (1994). Importance of phenolic glucosides in host selection of shoot galling sawfly, *Euura amerinae* , on *Salix pentandra* .*Journal of Chemical Ecology* , *20* (9), 2455–2466. https://doi.org/10.1007/BF02033213

Kopelke, J.-P. (2007). The European species of the genus*Phyllocolpa* , Part I: The leucosticta-group (Insecta, Hymenoptera, Tenthredinidae, Nematinae). *Senckenbergiana Biologica* ,*87* (1), 75–109.

Koskella, B., Giraud, T., & Hood, M. E. (2006). Pathogen relatedness affects the prevalence of within-host competition. *The American Naturalist* , *168* (1), 121–126. Retrieved from http://www.journals.uchicago.edu/t-and-c

Krishnan, H. B., & Franceschi, V. R. (1988). Anatomy of some leaf galls of *Rosa woodsii* (Rosaceae). *American Journal of Botany* ,*75* (3), 396–376. Retrieved from https://bsapubs.onlinelibrary.wiley.com/doi/abs/10.1002/j.1537-2197.1988.tb13451.x

Larson, K. C., & Whitham, T. G. (1991). Manipulation of food resources by a gall-forming aphid: the physiology of sink-source interactions.*Oecologia* , *88* (1), 15–21. https://doi.org/10.1007/BF00328398

Lefebvre, V., & Chèvre, A. (1995). Tools for marking plant disease and pest resistance genes: a review. *Agronomie* , *15* (1), 3–19. Retrieved from https://hal.archives-ouvertes.fr/hal-00885666

Leister, D. (2004). Tandem and segmental gene duplication and recombination in the evolution of plant disease resistance genes.*Trends in Genetics* , *20* (3), 116–122. https://doi.org/10.1016/J.TIG.2004.01.007

Li, H. (2011). A statistical framework for SNP calling, mutation discovery, association mapping and population genetical parameter estimation from sequencing data. *Bioinformatics* , *27* (21), 2987–2993. https://doi.org/10.1093/bioinformatics/btr509

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., . . . Durbin, R. (2009). The Sequence Alignment/Map format and SAMtools. *Bioinformatics* , *25* (16), 2078–2079. https://doi.org/10.1093/bioinformatics/btp352

Major, I. T., & Constabel, C. P. (2008). Functional analysis of the Kunitz trypsin inhibitor family in poplar reveals biochemical diversity and multiplicity in defense against herbivores. *Plant Physiology* ,*146* (3), 888–903. https://doi.org/10.1104/pp.107.106229

Martin, G. B., Bogdanove, A. J., & Sessa, G. (2003). Understanding the functions of plant disease resistance proteins. *Annual Review of Plant Biology* , *54* (1), 23–61. https://doi.org/10.1146/annurev.arplant.54.031902.135035

McKown, A. D., Klapste, J., Guy, R. D., Geraldes, A., Porth, I., Hannemann, J., . . . Douglas, C. J. (2014). Genome-wide association implicates numerous genes underlying ecological trait variation in natural populations of *Populus trichocarpa* . *New Phytologist* , *203* (2), 535–553. https://doi.org/10.1111/nph.12815

Meilan, R., Han, K. H., Ma, C., DiFazio, S. P., Eaton, J. A., Hoien, E. A., . . . Strauss, S. H. (2002). The CP4 transgene provides high levels of tolerance to Roundup® herbicide in field-grown hybrid poplars.*Canadian Journal of Forest Research* , *32* (6), 967–976. https://doi.org/10.1139/x02-015

Mello, M. O., & Silva-Filho, M. C. (2002). Plant-insect interactions: an evolutionary arms race between two distinct defense mechanisms.*Brazilian Journal of Plant Physiology* , *14* (2), 71–81. https://doi.org/10.1590/S1677-04202002000200001

15

Muchero, W., Guo, J., DiFazio, S. P., Chen, J.-G., Ranjan, P., Slavov, G. T., . . . Tuskan, G. A. (2015). High-resolution genetic mapping of allelic variants associated with cell wall chemistry in *Populus* . *BMC Genomics* , *16* (1), 24. https://doi.org/10.1186/s12864-015-1215-z

Muchero, W., Sondreli, K. L., Chen, J. G., Urbanowicz, B. R., Zhang, J., Singan, V., . . . LeBoldus, J. M. (2018). Association mapping, transcriptomics, and transient expression identify candidate genes mediating plant-pathogen interactions in a tree. *Proceedings of the National Academy of Sciences of the United States of America* ,*115* (45), 11573–11578. https://doi.org/10.1073/pnas.1804428115

Newcombe, G. (1998). A review of exapted resistance to diseases of*Populus* . *Forest Pathology* , *28* (4), 209–216. https://doi.org/10.1111/j.1439-0329.1998.tb01175.x

Newcombe, George, Martin, F., & Kohler, A. (2010). Defense and nutrient mutualisms in *Populus* . In S. Jansson, R. Bhalerao, & A. Groover (Eds.), *Genetics and Genomics of Populus* (pp. 247–277). NY: Springer New York.

Newcombe, George, & Ostry, M. (2001a). Recessive resistance to Septoria stem canker of hybrid poplar. *Phytopathology* , *91* (11), 1081–1084. https://doi.org/10.1094/PHYTO.2001.91.11.1081

Newcombe, George, & Ostry, M. (2001b). Recessive Resistance to Septoria Stem Canker of Hybrid Poplar. *Phytopathology* , *91* (11), 1081–1084. https://doi.org/10.1094/PHYTO.2001.91.11.1081

Newcombe, George, Stirling, B., & Bradshaw, H. D. (2001). Abundant pathogenic variation in the new hybrid rust *Melampsora* ×*columbiana* on hybrid Poplar. *Phytopathology* ,*91* (10), 981–985. https://doi.org/10.1094/PHYTO.2001.91.10.981

Nieuwland, J., Feron, R., Huisman, B. A. H., Fasolino, A., Hilbers, C. W., Derksen, J., & Mariani, C. (2005). Lipid transfer proteins enhance cell wall extension in tobacco. *The Plant Cell Online* ,*17* (7), 2009–2019. https://doi.org/10.1105/tpc.105.032094

Nyman, T., Widmer, A., & Roininen, H. (2000). Evolution of gall morphology and host-plant relationships in willow-feeding sawflies (Hymenoptera: Tenthradinidae). *Evolution* , *54* (2), 526–533. https://doi.org/10.1111/j.0014-3820.2000.tb00055.x

Oliveira, D. C., Isaias, R. M., Fernandes, G. W., Ferreira, B. G., Carneiro, R. G., & Fuzaro, L. (2016). Manipulation of host plant cells and tissues by gall-inducing insects and adaptive strategies used by different feeding guilds. *Journal of Insect Physiology* ,*1* (84), 103–113. https://doi.org/10.1016/j.jinsphys.2015.11.012

Orton, E. S., & Brown, J. K. M. (2016). Reduction of growth and reproduction of the biotrophic fungus *Blumeria graminis* in the presence of a necrotrophic pathogen. *Frontiers in Plant Science* ,*7* (MAY2016), 742. https://doi.org/10.3389/fpls.2016.00742

Panda, N., & Khush, G. S. (1995). *Host Plant Resistance to Insects* . Wallingford: CAB International.

Philippe, R. N., Ralph, S. G., Külheim, C., Jancsik, S. I., & Bohlmann, J. (2009). Poplar defense against insects: genome analysis, full-length cDNA cloning, and transcriptome and protein analysis of the poplar Kunitz-type protease inhibitor family. *New Phytologist* ,*184* (4), 865–884. https://doi.org/10.1111/j.1469-8137.2009.03028.x

Pinot, F., & Beisson, F. (2011). Cytochrome P450 metabolizing fatty acids in plants: characterization and physiological roles. *The FEBS Journal* , *278* (2), 195–205. https://doi.org/10.1111/j.1742-4658.2010.07948.x

Pretorius, Z. A., Singh, R. P., Wagoire, W. W., & Payne, T. S. (2000). Detection of Virulence to Wheat Stem Rust Resistance Gene *Sr31* in*Puccinia graminis* . f. sp. *tritici* in Uganda. *Plant Disease* , *84* (2), 203–203. https://doi.org/10.1094/PDIS.2000.84.2.203B

Qin, R., & LeBoldus, J. M. (2014). The infection biology of*Sphaerulina musiva* : Clues to understanding a forest pathogen.*PLoS ONE* , *9* (7). https://doi.org/10.1371/journal.pone.0103477

Ralph, S., Oddy, C., Cooper, D., Yueh, H., Jancsik, S., Kolosova, N., ... Bohlmann, J. (2006). Genomics of hybrid poplar (*Populus trichocarpa × deltoides*) interacting with forest tent caterpillars (*Malacosoma disstria*): normalized and full-length cDNA libraries, expressed sequence tags, and a cDNA microarray for the study of insect-induced defences. *Molecular Ecology* ,*15* (5), 1275–1297. https://doi.org/10.1111/j.1365-294X.2006.02824.x

Schuler, M. A. (2011). P450s in plant–insect interactions.*Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics* ,*1814* (1), 36–45. https://doi.org/10.1016/J.BBAPAP.2010.09.012

Smith, D. R., & Fritz, R. S. (1996). Review of the eastern United States species of the leaf-folding sawflies of the genus*Phyllocolpa* Benson (Hymenoptera: Tenthredinidae).*Proceedings- Entomological Society of Washington* , *98* , 695–707. Retrieved from https://www.researchgate.net/publication/260798186

Stajich, J. E. (2002). The Bioperl Toolkit: Perl Modules for the Life Sciences. *Genome Research* , *12* (10), 1611–1618. https://doi.org/10.1101/gr.361602

Stanton, B. J., B. Neale, D., & Li, S. (2010). Populus breeding: from the classical to the genomic approach. In S. Jansson, R. Bhalerao, & A. Groover (Eds.), *Genetics and Genomics of Populus* (pp. 309–348). https://doi.org/10.1007/978-1-4419-1541-2_14

Taylor, G. (2002). *Populus* : Arabidopsis for forestry. Do we need a model tree? *Annals of Botany* , *90* (6), 681–689. https://doi.org/10.1093/aob/mcf255

Thompson, J. N. (1988). Coevolution and alternative hypotheses on insect/plant interactions. *Ecology* , *69* (4), 893–895. https://doi.org/10.2307/1941238

Thompson, J. N., & Burdon, J. J. (1992). Gene-for-gene coevolution between plants and parasites. *Nature* , *360* (6400), 121–125. https://doi.org/10.1038/360121a0

Tian, F.-X., Zang, J.-L., Wang, T., Xie, Y.-L., Zhang, J., & Hu, J.-J. (2015). Aldehyde Dehydrogenase Gene Superfamily in *Populus* : Organization and Expression Divergence between Paralogous Gene Pairs.*PLOS ONE* , *10* (4), e0124669. https://doi.org/10.1371/journal.pone.0124669

Tschaplinski, T. J., Abraham, P. E., Jawdy, S. S., Gunter, L. E., Martin, M. Z., Engle, N. L., ... Tuskan, G. A. (2019). The nature of the progression of drought stress drives differential metabolomic responses in *Populus deltoides* . *Annals of Botany* ,*124* (4), 617–626. https://doi.org/10.1093/aob/mcz002

Veach, A. M., Yip, D., Engle, N. L., Yang, Z. K., Bible, A., Morrell-Falvey, J., ... Schadt, C. W. (2018). Modification of plant cell wall chemistry impacts metabolome and microbiome composition in*Populus* PdKOR1 RNAi plants. *Plant and Soil* ,*429* (1–2), 349–361. https://doi.org/10.1007/s11104-018-3692-8

Verhoeven, T., Feron, R., Wolters-Arts, M., Edqvist, J., Gerats, T., Derksen, J., & Mariani, C. (2005). STIG1 Controls Exudate Secretion in the Pistil of Petunia and Tobacco. *Plant Physiology* ,*138* (1), 153–160. https://doi.org/10.1104/pp.104.054809

Wang, Y., Tang, H., Debarry, J. D., Tan, X., Li, J., Wang, X., ... Paterson, A. H. (2012). MCScanX: A toolkit for detection and evolutionary analysis of gene synteny and collinearity. *Nucleic Acids Research* , *40* (7), 1–14. https://doi.org/10.1093/nar/gkr1293

Weston, D. J., Rogers, A., Tschaplinski, T. J., Gunter, L. E., Jawdy, S. A., Engle, N. L., ... Wullschleger, S. D. (2015). Scaling nitrogen and carbon interactions: what are the consequences of biological buffering? *Ecology and Evolution* , *5* (14), 2839–2850.

Whitham, T. G., Bailey, J. K., Schweitzer, J. A., Shuster, S. M., Bangert, R. K., LeRoy, C. J., ... Wooley, S. C. (2006). A framework for community and ecosystem genetics: from genes to ecosystems.*Nature Reviews Genetics* , *7* (7), 510–523. https://doi.org/10.1038/nrg1877

Wool, D. (2004). Galling aphids: specialization, biological complexity, and variation. *Annual Review of Entomology* , *49* (1), 175–192. https://doi.org/10.1146/annurev.ento.49.061802.123236

Yang, Y., Labbé, J., Muchero, W., Yang, X., Jawdy, S. S., Kennedy, M., . . . Chen, J. G. (2016). Genome-wide analysis of lectin receptor-like kinases in *Populus* . *BMC Genomics* ,*17* (1), 699. https://doi.org/10.1186/s12864-016-3026-2

Yin, T.-M., DiFazio, S. P., Gunter, L. E., Jawdy, S. S., Boerjan, W., & Tuskan, G. A. (2004). Genetic and physical mapping of Melampsora rust resistance genes in Populus and characterization of linkage disequilibrium and flanking genomic sequence. *New Phytologist* ,*164* (1), 95–105. https://doi.org/10.1111/j.1469-8137.2004.01161.x

Zhao, N., Guan, J., Ferrer, J. L., Engle, N., Chern, M., Ronald, P., . . . Chen, F. (2010). Biosynthesis and emission of insect-induced methyl salicylate and methyl benzoate from rice. *Plant Physiology and Biochemistry* , *48* (4), 279–287. https://doi.org/10.1016/j.plaphy.2010.01.023
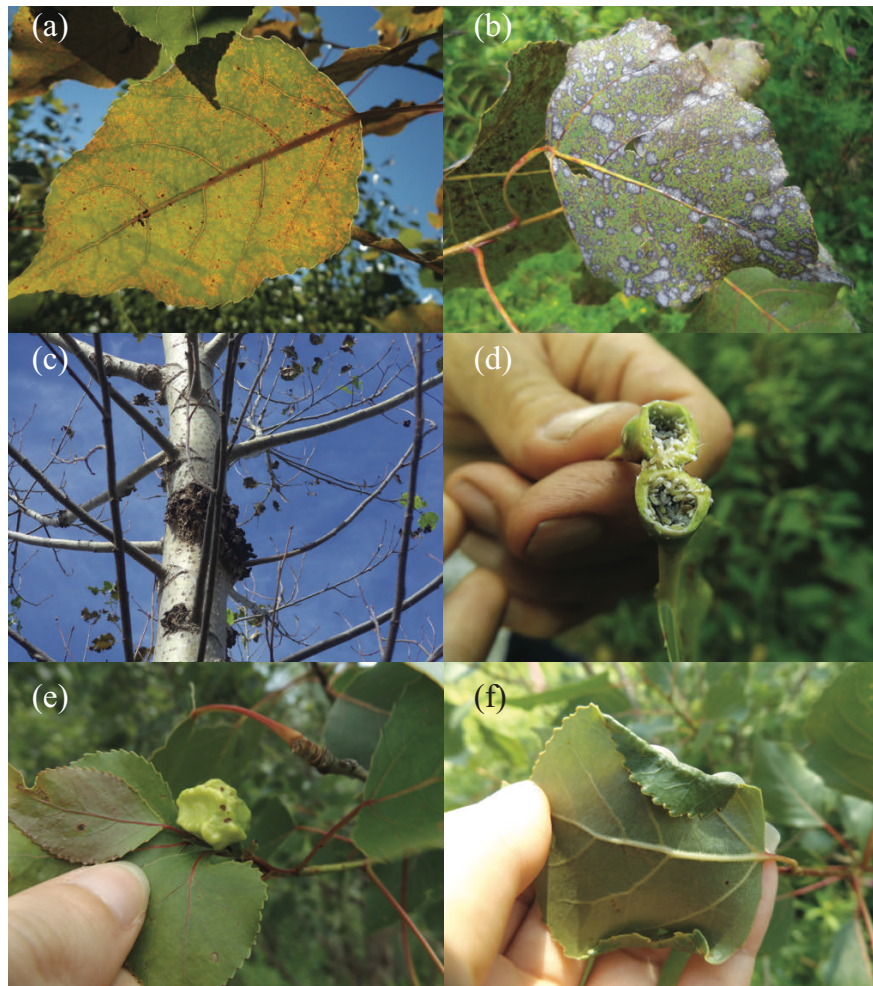
figures/Figure2/Figure2-eps-converted-to.pdf

figures/Figure3/Figure3-eps-converted-to.pdf

figures/Figure4/Figure4-eps-converted-to.pdf

figures/Figure5/Figure5-eps-converted-to.pdf

figures/Figure6/Figure6-eps-converted-to.pdf