Running title: Botrytis-Arabidopsis transcriptome eGWA

**Pathogen genetic control of transcriptome variation in the *Arabidopsis thaliana* – *Botrytis cinerea* pathosystem**

**Nicole E. Soltis1,2, Celine Caseys1, Wei Zhang3, Jason A. Corwin4, Susanna Atwell2, and Daniel J. Kliebenstein1,2,5\***

1Department of Plant Sciences, University of California, Davis, One Shields Avenue, Davis, CA, 95616, USA

2Plant Biology Graduate Group, University of California, Davis, One Shields Avenue, Davis, CA, 95616, USA

3Department of Plant Pathology, Kansas State University, 1712 Clafflin Road, Throckmorton Hall, Manhattan, KS, 66506, USA

4Department of Ecology and Evolution Biology, University of Colorado, 1900 Pleasant Street, 334 UCB, Boulder, CO, 80309-0334, USA

5DynaMo Center of Excellence, University of Copenhagen, Thorvaldsensvej 40, DK-1871, Frederiksberg C, Denmark

**\*Correspondence:** Daniel J. Kliebenstein, Department of Plant Sciences, University of California, Davis, One Shields Ave, Davis, CA, 95616, USA.

Kliebenstein@ucdavis.edu

**Keywords: Host-Pathogen Interaction, Pathosystem, Arabidopsis, *B. cinerea*, RNA-Seq, Genome Wide Association, Dual transcriptome**

**ABSTRACT**

Disease symptoms arise from the interaction of the host and pathogen genomes. However, little is known about how genetic variation in the interaction modulates both organisms’ transcriptomes, especially in complex interactions like those between generalist pathogens and multiple plant hosts. In this study, we used the *Botrytis cinerea* - *Arabidopsis thaliana* pathosystem to examine how genomic variations in the pathogen influences both host and pathogen transcriptomes. Using whole-genome sequencing derived single nucleotide polymorphisms (SNPs) in a collection of 96 *B. cinerea* isolates, we performed genome-wide association (GWA) for each of 23,947 measurable transcripts in Arabidopsis (host), and 9,267 measurable transcripts in *B.cinerea* (pathogen). Unlike other eGWA studies, we detected a relative absence of *cis*-eQTL partly caused by structural variants and allelic heterogeneity hindering their identification. The identified *trans*-eQTL could be linked to eQTL hotspots dispersed across Botrytis genome that altered only Botrytis transcripts, only Arabidopsis transcripts, or transcripts from both species. Gene membership in the *trans-*eQTL hotspots suggests links to several known and many novel virulence mechanisms in this pathosystem. Genes annotated to these hotspots provide potential targets for blocking manipulation of the host response by this ubiquitous generalist necrotrophic pathogen. This shows that genetic control over the co-transcriptome is polygenic, similar to the virulence outcome in the interaction of *Botrytis cinerea* on *Arabidopsis thaliana*.

100 Word Summary:

Disease arises from the interaction of host and pathogen genomes. We mapped transcriptome variation in the *Botrytis cinerea – Arabidopsis thaliana* pathosystem to genetic variation in the pathogen. We used data from a previously published co-transcriptome analysis across 96 diverse *B. cinerea* isolates infected on the Arabidopsis wildtype, Col-0 to perform genome-wide association (GWA) for thousands of measurable transcripts across the host or pathogen. Identified *trans*-eQTL in the pathogen linked to eQTL hotspots dispersed across the pathogen genome, influencing known and novel virulence mechanisms. The polygenic genetic control of the co-transcriptome mirrors the polygenic virulence of *Botrytis cinerea* on *Arabidopsis thaliana*.

**INTRODUCTION**

Infectious disease results from an interaction between host and pathogen driven by the genetics of both organisms. The mechanisms of plant-pathogen interactions are often divided into qualitative, in which a few genetic variants of large effect shape binary disease outcomes, or quantitative, in which a spectrum of outcomes arise from the interaction of polygenic variation in the host and pathogen. The past decades have witnessed the unraveling of the molecular basis of large-effect loci on both the host and the pathogen sides that control qualitative interactions (Giraldo and Valent 2013; Marone *et al.* 2013; Meng and Zhang 2013; Cui *et al.* 2015; Lo Presti *et al.* 2015). In the qualitative model, alternative alleles of the genes create sweeping differences in the transcriptome and phenotypic responses to infection in both the host and pathogen via differential recognition events surrounding their proteins. However, plant-pathogen interactions cover a full range of genetic architectures, from few genes of large effect to many genes of small effect (Poland *et al.* 2009; Kou and Wang 2010; Lannou 2012). In contrast to qualitative systems, quantitative plant-pathogen interactions exhibit a lack of virulence/ resistance genes that explain large proportions of the variance in the disease outcome in the population (Poland *et al.* 2009; Kou and Wang 2010; St. Clair 2010; Roux *et al.* 2014). These interactions are highly polygenic with genetic variation influencing diverse molecular mechanisms, extending beyond host-pathogen perception and large effect arms-race loci (Glazebrook 2005; Nomura *et al.* 2005; Goss and Bergelson 2006; Rowe and Kliebenstein 2008; Barrett *et al.* 2009; Corwin *et al.* 2016a; Bartoli and Roux 2017; Wu *et al.* 2017; Atwell *et al.* 2018b; Fordyce *et al.* 2018; Soltis *et al.* 2019). It is, however, unclear how these polygenic molecular interactions alter higher-order phenotypes such as virulence or the transcriptome of both species. There is conflicting evidence on the balance of the system. Some studies and traits indicate that genetic variation in the pathogen dominates the interaction (Bartha *et al.* 2017; Wang *et al.* 2018a). Other studies find a balanced contribution of plant and pathogen genetics (Corwin *et al.* 2016a; Soltis *et al.* 2019). Thus, there is a need to develop genomic approaches to understand how polygenic variations affect the genomic response of both organisms.

The polygenic variation in the pathogen should influence numerous genes that consequently shift the pathogen’s transcriptome and cause differential expression of various virulence mechanisms. This variation in virulence mechanism will then impact the host and lead to shifts in the host’s resistance-associated transcriptome. Thus, by measuring the transcriptome in both the pathogen and the host, it should be possible to map how genetic variation in the pathogen is conveyed through the pathogen’s transcriptome and concurrently how the host’s transcriptome responds. Recent work has shown that it is possible to measure the pathogen’s transcriptome in planta in the *A. thaliana* - *Pseudomonas syringae* pathosystem leading to new hypotheses about virulence (Nobori *et al.* 2018). In the *A. thaliana* - *B. cinerea* pathosystem, the genetic interactions are dominated by complex small-effect loci that display a high degree of interaction between the host and pathogen (Denby *et al.* 2004b; Finkers *et al.* 2007; Finkers *et al.* 2008; Rowe and Kliebenstein 2008; Anuradha *et al.* 2011; Fu *et al.* 2017; Fordyce *et al.* 2018). In this pathosystem, a co-transcriptome study with simultaneous analysis of the host and pathogen’s transcripts was recently done through single-sample RNA-Seq (Zhang *et al.* 2017; Zhang *et al.* 2019). This co-transcriptome approach allowed the description of key virulence networks in the pathogen and resistance responses within the host (Zhang *et al.* 2017; Zhang *et al.* 2019). Further, this study revealed a single network of transcripts from pathogen and host species. However, these studies did not assess the genetic architecture behind these co-transcriptome interactions.

GWA that identifies expression quantitative trait loci (eQTL, SNPs correlated with variations in transcript expression) can reveal the genetic architecture behind these co-transcriptome interactions. Previous eQTL studies revealed that the SNPs that cause differential transcript accumulation can be parsed into *cis* or *trans* effects. Locally acting (*cis*) eQTL indicate regulatory variation within or near the expressed gene itself. *trans*-eQTL reveal SNPs that are acting at a distance to affect regulatory processes influencing the expression of the transcript. A *trans*-eQTL that affects many transcripts is classified as a hotspot. Such trans-acting hotspot SNPs may influence regulatory processes that in turn influence numerous transcripts. eQTL analysis has been utilized to study host-pathogen interactions, albeit with a focus either on host or pathogen. Frequently, these studies focus on the host’s response, such as mapping how host loci control host gene expression over time using either traditional QTL mapping or GWA analysis (Chen *et al.* 2010; Hsu and Smith 2012; Zou *et al.* 2012; Allen *et al.* 2016; Christie *et al.* 2017). Additional studies have begun to invert this scheme by looking at how genetic variation in the pathogen influences the host transcriptome to identify pathogen loci modulating host expression levels, and thus to identify candidate loci for interspecific signals (Saeij *et al.* 2007; Wu *et al.* 2015; Guo *et al.* 2017). These studies attest to the potential to identify pathogen loci that influence host gene expression. However, previous studies have thus far addressed pathogen populations with limited genetic variation, and thus identify the few polymorphic loci between strains with strongest effects on transcriptomic variation (Wu *et al.* 2015; Guo *et al.* 2017).

Expanding the scope of these studies, we performed a co-transcriptome analysis wherein both the wild type *A. thaliana* Col-0 host and *B. cinerea* pathogen transcriptomes are measured using a diverse *B. cinerea* population (Zhang *et al.* 2017; Zhang *et al.* 2019). We conducted a GWA analysis of both host and pathogen transcriptomes to identify loci in the *B. cinerea* genome that may affect the transcriptomes of either or both organisms (Zhang *et al.* 2017; Zhang *et al.* 2019). The genomes of both the host and the pathogen harbor extensive genetic diversity that has been successfully used for genetic mapping to identify loci controlling virulence in combination with transcriptomics and genomics (Atwell *et al.* 2018b; Soltis *et al.* 2019) (Denby *et al.* 2004a; Rowe and Kliebenstein 2008; Dalmais *et al.* 2011; Schumacher *et al.* 2012; Zhang *et al.* 2017). The loci tagged by these SNPs have an explicit directionality of effect, as genetic causality must arise within the pathogen and then extend to the host. Our analysis found mostly small-effect polymorphisms dispersed throughout the *B. cinerea* genome, with several *trans*-eQTL hotspots. These hotspot loci are associated with specific host or pathogen transcript co-expression modules and variation in lesion size. There was no identifiable overlap in the hotspots that influenced the host’s or the pathogen’s transcriptome, suggesting a surprisingly independent basis of transcriptional regulation of host and pathogen by the *B. cinerea* genome. All of the hotspot loci tagged genes with no previous association to plant-pathogen virulence interactions. This generates a set of *B.cinerea* loci that have regulatory potential in controlling the interaction via modulation of gene expression to influence the lesions outcome.

**RESULTS**

**eQTL indicate polygenic transcriptome modulation**

To understand how natural genetic variation in the pathogen influences both the host and pathogen transcriptomes, we performed expression GWA across all genes expressed in each species within the *B. cinerea* - *A. thaliana* pathosystem. This analysis incorporated the expression profiles of 9,267 *B. cinerea* genes and 23,947 Col-0 *A. thaliana* genes,mappedeach as individual traits across 96 diverse *B. cinerea* isolates. For each trait, we used a Genome-wide Efficient Mixed Model Association (GEMMA) univariate linear mixed model (Zhou and Stephens 2012) with a previously described *B. cinerea* genome-wide SNP dataset of 237,878 SNPs with a conservative minimum minor allele frequency of 0.20 (Atwell *et al.* 2018a; Soltis *et al.* 2019). GEMMA estimates the significance of effects of each SNP on the focal trait as a *P*-value after accounting for potential effects of population structure within the *B. cinerea* isolates. Given the scale of this dataset at 33,214 transcript phenotypes, it was not computationally viable to estimate empirical significance thresholds for every transcript using at least 1,000 permutations. However, we conducted five permutation tests in which we randomized the phenotype-genotype associations and repeated GEMMA across all the traits (individual expression profiles in *B. cinerea* and *A. thaliana*) to query for potential patterns that may exist randomly. Using these permutations, we compiled the top 5% of random p-values from each gene across all the permutations and utilized the median of these values as a conservative genome wide-permutation threshold, 1.96e-5 for the *B. cinerea* transcriptome and 2.9e-5 for the *A. thaliana* transcriptome. In *B. cinerea*, this threshold identified an average of 4756 SNPs associated with the expression of 461 transcripts (range 3843 to 5584 SNPs and 410 to 499 transcripts) across the five random permutations. In *A. thaliana*, this threshold identified an average of 16,446 SNPs associated with the expression of 1,201 transcripts (range 13,040 to 22,359 SNPs and 1129 to 1260 transcripts) across the five random permutations.

Using GEMMA with these conservative genome wide-permutation thresholds and the real transcriptome datasets identified 461,723 *B. cinerea* SNPs associated with transcriptional variation in 1,616 *B. cinerea* genes and 978,693 SNPs associated with 5,213 *A. thaliana* genes (Figure 1). In comparison to the random permutations, this suggests an alpha error rate for SNPs of approximately 1% using the 1.96e-5 threshold for *B. cinerea* (4,756 permutation/461,723 empirical) and 1.6% using the 2.9e-5 threshold for *A. thaliana* (16,446 permutation /978,693 empirical).

The range of significant SNPs per transcript is wide with 0- 16,814 SNPs for *B. cinerea* and 0-24,622 SNPs for *A. thaliana*. The observation of transcript variation associated with a large number of significant SNPs has been seen in other eQTL analyses (Figure S1) (West *et al.* 2007; Wang *et al.* 2018b). Given the wide-range of significant SNPs per transcript and how it could shape the analysis, we assessed the definition provided by analysis of the top SNP per gene (SNP with lowest *P*-value) to survey genomic patterns. We thus tested how conservative our threshold was by comparing the permuted minimum to the empirical minimum *P*-value across all SNPs per individual transcript. This showed that the empirical minimum SNP per trait for 69% of the *B. cinerea* transcripts and for 58% of the *A. thaliana* transcripts is lower than transcripts minimum permuted SNP across all five permutations. This indicates that using the most significant SNP per transcript will have a sufficient signal to noise ratio to identify over-arching patterns.

Therefore, we focused on the top significant SNP per transcript for the remaining analysis. This restrictive analysis on SNPs that are most likely to be associated with the gene expression accounts for the fact that nearly every transcript has a highly significant heritability ascribed to the *B. cinerea* genome (Zhang *et al.* 2017; Zhang *et al.* 2019). To account for the potentially low resolution in genomic signal encompassed by the single top SNP, we also assessed any general pattern using the top 10 SNPs.

**Absence of observed transcriptome cis-effect**

A hallmark of within species eQTL mapping studies using either GWA or structured mapping populations in a wide range of species is the occurrence of cis-diagonals (Brem *et al.* 2002; Schadt *et al.* 2003; Monks *et al.* 2004; Keurentjes *et al.* 2007; West *et al.* 2007; Zou *et al.* 2012). This occurs because polymorphisms proximal to the gene, *cis-*eQTL or *cis*-SNPs, frequently have large-effects on the transcripts accumulation measured as either fraction of population variance or allele substitution effects. To test whether our analysis identifies a similar *cis*-diagonal in the *B. cinerea* transcriptome, we plotted the position of the transcript’s genomic position against the top SNP for all the *B. cinerea* transcripts. As cis-diagonals are generally enriched in SNPs with the largest effects on transcript abundance, the top GWA SNP should identify this diagonal assuming there are no intervening technical or biological issues. However, this study found no evidence supporting a *cis*-diagonal (Figure 2). This pattern held whether we examined the top SNP per transcript (Figure 2a) or the top 10 SNPs per transcript (Figure 2b). In contrast, there is evidence for *trans*-eQTL hotspots; distant loci which modulate expression variation across many of the pathogen genes (Figure 2).

To test whether the absence of *cis*-effect resulted from a genomic distance bias, we calculated the genomic distance between the center of each transcript and the top associated SNP. If *cis*- acting loci contributed a majority of genetic control of expression variation, we would expect to see a high frequency of short-distance associations, and a rapid decline to a plateau moving away from the gene of interest. However, we observed that distances between transcript center and top SNP as far as 2 Mb are common (Figure S3). These distances are similar to what would happen if the causal SNPs are *trans*-effects scattered across the genome instead of local *cis*-effects (Figure S3).

Because any significance threshold may complicate the identification of a *cis*-signal, we converted all the P-values assigned to SNPs per transcript into their quantile rank. This conversion was performed by sorting all p-values in ascending order, assigning ranks values to each p-value and converting these values to percentage of the total number of SNPs. We then identified all the SNPs within each *B. cinerea* transcript or within 1kb of the start or end of the transcript and extracted the quantile ranks for these SNPs. The number of local SNPs ranged from 1-352 SNPs, with an average of 16 local SNPs per transcripts. The distribution of ranks across all the local SNPs was flat and similar to a random sampling of *P*-values further indicating no general cis-enrichment within this dataset (Figure S4). We next limited this quantile rank analysis to the local top SNP for each transcript because each cis-region may have only one or a few true associations masked by non-causal SNPs. As a random comparison, we simulated a dataset of quantile rank by random draw of 16 SNPs out of all SNPs defined cis-defined boundaries repeated 9267 times (the number of *B. cinerea* transcripts). For each simulated transcript, we extracted the top SNP and repeated the whole simulation five times. This showed that that cis-SNPs are not enriched for higher significance than the genome and might even be less frequent than expected by random chance (Figure S4). As such, we do not detect evidence for over-representation of *cis*-effect loci for the control of gene expression variation.

**Allelic heterogeneity and structural variation potentially masking *cis*-effects**

Haplotype heterogeneity or allele frequency may complicate the accurate identification of *cis*-polymorphisms and explain the absence of detection of a *cis*- pattern (Chan *et al.* 2010; Rivas *et al.* 2011; Visscher *et al.* 2017). To assess these possibilities we used three clustered biosynthetic pathways ; the botcinic acid biosynthetic pathway (13 genes, 55.8 kb, Chromosome 1), botrydial biosynthetic pathway (7 genes, 26 kb, Chromosome 12), and a putative cyclic peptide pathway (10 genes, 46.5 kb, Chromosome 1) (Deighton *et al.* 2001; Colmenares *et al.* 2002; Porquier *et al.* 2016; Zhang *et al.* 2019). These pathways have known presence-absence polymorphisms segregating in the studied *B.cinerea* population and were expected to hold *cis-*patterns, yet none were detected by our GWA analysis (Siewers *et al.* 2005; Pinedo *et al.* 2008). Additionally, genes of the botcinic acid and botrydial biosynthetic pathways are variable across the *Botrytis* genus (Valero-Jiménez *et al.* 2019). Critically, the transcription profiles of the genes within these pathways are highly correlated across isolates, suggesting pathway-specific control of expression variation (Zhang *et al.* 2019). Thus, the absence of *cis*-SNPs in these loci may represent *cis*-pattern false-negatives. To test if these pathways have undetected *cis*-SNPs, we extracted all the SNPs within each biosynthetic cluster and investigated haplotype diversity across the 96 *B. cinerea* isolates.

We first investigated the botcinic acid cluster that contains multiple distinct haplotypes (Figure 3a). We tested the haplotype specific effects on transcript expression within that pathway. To estimate the pathway expression in each isolate, we z-scored and averaged the expression profile of each individual gene (Kliebenstein *et al.* 2006). These pathway expression profiles identified a group of 12 isolates (Cluster 4 in Figure 3b) with a distinctly lower level of expression than the other groups, with all four groups significantly distinct at >95% bootstrap. The Investigation of the short-reads sequences and SNP calls showed that these 12 isolates share a 78.6 kb deletion that removes the entire botcinic acid biosynthetic cluster, from Bcboa1 to Bcboa17 (Figure 3). After accounting for this major deletion, we found no remaining significant effect of group membership in the remaining 3 major groups (Clusters 2, 5 and 6) on expression profile (F(1,74)=0.36, p=0.55). However, within each of these groups, independent isolates have low pathway expression, suggesting loss-of-expression polymorphisms (e.g. isolates Noble Rot and Apple517 in cluster 5, and isolates 01.04.03 and 1.03.20 in cluster 6) (Figure 3b). While each of these isolates contain independent smaller deletions the origins of the loss of expression of the genes of botcinic acid pathway remain unknown. These patterns may suggest undetected *cis*-effect polymorphisms in addition to the large common deletion and independent additional events. Functional analysis of the botcinic acid biosynthetic cluster has thus far identified one transcription factor (Bcboa13) that controls the expression of the cluster (Porquier *et al.* 2019). However, none of the SNPs within or near Bcboa13 were significantly associated with variation in expression of the botcinic acid biosynthesis genes.

We then investigated the botrydial and putative cyclic peptide biosynthetic pathways for additional evidence of un-detected *cis*-acting genetic variation. These two pathways exhibit a lack of *cis*-effect SNP patterns similar to the botcinic acid pathway. Hierarchical clustering within each of these pathways by genic SNP variation divided the isolate population into two groups that were not associated with mean pathway expression (Figure S5, Figure S6, both bipartite divisions were supported by >95% of bootstraps). While there was no obvious structural variation in the botrydial pathway, the cyclic peptide pathway contained small deletions within the intergenic regions that correlate with low expression. Furthermore, two isolates with partial deletions within the genes early in the pathway exhibited very low pathway expression (1.05.16, 1.05.22) (Figure S6).

Overall, we identified *cis*-acting deletions in two of three of the clustered biosynthetic pathways. This highlights the potential of structural variants that often fall below the minor allele cutoffs and compromise the detection of *cis*-effects by GWA. Testing whether insertion and deletion events account for the majority of localized control of expression variation would require both long-read sequencing to accurately identify these structural variants and computational approaches that can blend SNP and insertion-deletion information.

**Detection of *trans*-eQTL hotspots**

The GWA analysis identified a strong signature of *trans*-eQTL hotspots (Figure 2), causal polymorphisms that may influence the regulation of numerous genes in *trans*. These hotspots were detected using both the *B. cinerea* and *A. thaliana* transcriptome with solely the top SNP per transcript (Figure 4, Figure 5). To fix the thresholds that define trans-eQTL hotspots, we calculated the maximum number of transcripts associated to a single SNP across the five random permutations. These maximum eQTL hotspot sizes were 11 *B. cinerea* and 80 *A. thaliana* transcripts. We thus fixed conservative thresholds for eQTL hotspots to 20 transcripts for *B. cinerea* and 150 transcripts for *A. thaliana*. These thresholds identified 13 SNPs as potential *trans*-eQTL hotspots for the *B. cinerea* transcriptome and 12 SNPs as potential cross-specieseQTL influencing *the A. thaliana* transcriptome (Figure 5, Figure 6). These eQTL hotspots affecting both species transcriptomes are distributed throughout *B. cinerea* genome (Figure 6, Table 1). The hotspots SNPs had an average minor allele frequency (MAF) of 27% (1% S.E.) with no correlation between the MAF and the number of transcripts associated to a particular SNP within either the Arabidopsis or Botrytis transcripts. Further, the hotspot SNPs were not co-segregating across the isolates.

The use of a co-transcriptome approach should theoretically allow the identification of *trans*-eQTL hotspot affecting *B. cinerea* transcripts that by extension create an eQTL hotspot affecting the host’s transcriptome. However, the GWA analysis detected no significant cross-species connected eQTL hotspots across the two transcriptomes. This result was consistent across analysis of the single top SNP per transcript and the top 10 SNPs per transcript (Figure S7). The absence of cross-species overlap in eQTL hotspots suggests that the pathogen’s influence on the host’s transcriptome is not solely limited to major interactions between *tran*s-eQTL hotspots but can involve molecularly constrained changes in the pathogen that are magnified in the host’s response. One possible scenario that fit this model is a SNP that alters the expression of a single effector gene or mechanism in the pathogen that doesn’t impact the pathogen but impacts the host. For example, altering expression of the botcinic acid biosynthetic cluster would alter the accumulation of that phytotoxic metabolite and cause large responses in the host. Similarly, mutations that do not affect transcript abundance like missense polymorphisms that alter protein function could equally lead to a lack of overlap in hotspots. In addition, it is possible that the high level of genetic variation in *B. cinerea* may decrease our power to detect the cross-species trans-eQTL hotspots. Deeper analysis into the transcriptome and downstream responses could elucidate how restricted responses in the pathogen transcriptome translate to sweeping responses in the host. Future studies using these eQTL hotspots as *a priori* candidates for control of transcript variation in both host and pathogen may need to increase power to detect more modulation overlap across the two transcriptomes.

**eQTL hotspot modules**

To better understand the gene sets under the influence of the trans-eQTL hotspots, we examined the genes influenced by each hotspot. We first collected the gene ontology (GO) annotations within each species to address whether hotspots were enriched for some specific functionality. The *B. cinerea* GO annotations showed a preponderance of enzyme, signal peptides for secretion, and transcription factor annotations but no specific molecular insights arose, largely because the majority of genes had no annotation (Table 1, Table S1, File S1, File S4). In contrast, GO analysis of the *A. thaliana* transcripts showed that three of the hotspots have an overrepresentation of photosynthesis-related functions within their targeted genes (Table 1, File S2, File S3). Downregulation of photosynthesis gene expression is a hallmark of plant immune processes (Bilgin *et al.* 2010; Jiang *et al.* 2017). Two of the hotspots primarily affect *A. thaliana* genes associated with abiotic stress responses. Only two of the hotspots influence expected plant defense loci, including chitin response and microbe defenses. This suggests that the *B. cinerea* genes underlying these hotspots have specific effects on defined networks within the host and are not causing nonspecific responses.

In previous work, we had defined key transcript modules within both the host and pathogen transcriptomes that connected to virulence (Zhang *et al.* 2017; Zhang *et al.* 2019). Thus we tested for overlap between the *trans*-eQTL hotspot defined modules and previously defined transcript modules. We compared the gene lists of module and hotspot membership using a hypergeometric enrichment test (Zhang *et al.* 2017; Atwell *et al.* 2018b; Zhang *et al.* 2019). Nine of the 11 *B. cinerea* eQTL hotspots were enriched for transcripts present in one or more of four major *B. cinerea* modules on *A. thaliana* (Figure 6). An additional six *B. cinerea* modules did not share any gene membership with detected eQTL hotspots. Similarly, nine of the *A. thaliana* eQTL hotspots were enriched for transcripts from two of the major *A. thaliana* modules when infected with *B. cinerea* (Figure 6). These two *A. thaliana* modules contain genes that function in jasmonate and salicylic acid signaling processes and camalexin biosynthesis (Network I), or photosynthesis (Network IV). Interestingly, these links are not limited to a single hotspot, but have strong connections across several different eQTL hotspots suggesting that these *A. thaliana* modules are influenced by the pathogens polygenic architecture(Figure 6).

**eQTL hotspot candidate genes**

To generate working hypotheses on the possible causal basis of the eQTL hotspots, we investigated the candidate genes underlying the associated SNPs. The 12 *B. cinerea* hotspots that influence *A. thaliana* transcripts are located within 11 genes, including 4 enzymes and 2 genes associated with isolate compatibility (Table 1). The 13 *B. cinerea* hotspots that influence *B. cinerea* expression profiles were associated to 11 genes, including 4 enzymes (Table 1). However, only one of these 22 genes was previously linked to virulence functions in *B. cinerea* or other fungi. Bccwh41 (Bcin16g01950), a glycoside hydrolase whose homolog shows increased expression in virulent strains of *Ustilago maydis* on *A. thaliana* (Martínez-Soto *et al.* 2013). To test if any of these 22 eQTL hotspot candidate genes may influence virulence in *B. cinerea*, we used the existing co-expression and virulence data to compare the level of expression of these 22 genes to existing virulence measurements (Zhang *et al.* 2017; Atwell *et al.* 2018b; Zhang *et al.* 2019). Transcript accumulation for three *B. cinerea* hotspot genes and two of the *A. thaliana* hotspot genes are strongly positively correlated to lesion size variation while none are negatively correlated with lesion size (Table 1) (Zhang *et al.* 2019). Further, we utilized a previous GWA analysis of virulence of these same isolates on *A. thaliana* to test for any overlap. This showed that one of the *B. cinerea* hotspot genes (Bcin16g00010, SsuA/THI5-like) is a top GWA hit controlling lesion size across host genotypes and association methods (Table 1) (Atwell *et al.* 2018b). Together, this suggests that these genes are likely candidates for controlling transcriptome responses in both the host and pathogen.

**DISCUSSION**

(Wang *et al.* 2018a)**Dispersed interactions across host and pathogen genomes**

Using co-transcriptome GWA, we identified 25 *trans-*eQTL hotspots dispersed across the *B. cinerea* genome that modulate either the host or pathogen transcriptomes. This contrasts with previous cross-species eQTL studies, which identified one or only a few cross-species eQTL hotspots (Wu *et al.* 2015; Guo *et al.* 2017). Further, most of the genetic variation detected in our study is distant from the affected transcripts, *i.e*. located in *trans*. These *trans-*eQTL hotspots influence expression variation for five major *B. cinerea* modules containing genes dispersed across the genome (Zhang *et al.* 2019). In particular, the eQTL hotspots influenced the expression of many genes from the previously identified *B. cinerea trans*-co-expression networks (vesicle/virulence, translation/growth, exocytosis regulation, peptidase). Interestingly, the candidate polymorphisms are spread throughout the genome and the detected eQTL hotspots are not in regions of the genome with significantly elevated genetic variation. Further, the genetic targets of these eQTL are dispersed across the plant and pathogen genomes (Zhang *et al.* 2019). As such, *B. cinerea* does not fit the model of what might be expected in filamentous fungi showing multiple-speed genome evolution due to varying selective pressures influencing the genome. In such specialist fungi with closer co-evolution to their host species than Botrytis, diverse fungal virulence effectors can be enriched in regions of the genome containing enhanced rates of mutation and polymorphism while the rest of the genome shows slower evolutionary rates (Dong *et al.* 2015). If this scenario were true for Botrytis genome, it would predict clustering of the great majority of GWA hits to a few locations, rather than distribution of eQTL across the genome as was found. This is consistent with previous findings of high genome-wide diversity in *B. cinerea*, and virulence and host specificity mapping to large swaths of the pathogen genome including 16 of 18 chromosomes (Atwell *et al.* 2018b; Caseys *et al.* 2018). Similar eQTL analyses in the multi-speed genome filamentous fungi are required to test whether eQTL in pathogens with a multi-speed genome truly cluster within the highly polymorphic regions. Together, these findings provide evidence for polygenic *trans*-regulation of gene expression in *B. cinerea* interactions that then coalesces around specific transcriptional modules to influence virulence.

**Polygenic modules and pleiotropy in cross-species eQTL**

Previous pathogen eQTL studies typically identified qualitative patterns whereby each host expression profile was explained by only a single major-effect pathogen locus (Guo *et al.* 2017) or each pathogen eQTL connected to a specific host network (Wu *et al.* 2015). In contrast, co-transcriptome GWA with *B. cinerea* identified a more complex picture with numerous *trans­-*eQTL hotspots altering multiple transcriptome modules in either the host or the pathogen. This suggests that the polygenic architecture of the pathogen may at least in part function by influencing these defined modules rather than functioning as thousands of individual genes each separately targeting the host. This gives us an overarching pattern of polygenic and pleiotropic genetic regulation, as the disease outcome is based on extensive genetic variation in both the host and pathogen. In effect, we see polygenicity of host expression regulation by the pathogen at the gene level, and at the network level. It remains to be ascertained if this system functions to create robustness in these connections in the face of changes to the pathogen or host genetics or if, alternatively, this is an indication of existence of discrete set of interaction mechanisms between the host and the pathogen.

**Candidate causal loci encode diverse mechanisms**

Querying the putative function of the candidate loci underlying the different *trans*-eQTL hotspots identified an array of potential molecular mechanisms. While one might assume that transcription factors are the most likely genes for containing genetic variation that would lead to *trans*-eQTL hotspots, we found instead enrichment for enzyme-encoding genes among these loci. Eight enzyme-encoding genes contained the SNPs for four of the *B. cinerea* *trans*-eQTL and four of the *A. thaliana* *trans*-eQTL hotspots. Interestingly, these enzymes have potential activities in sugar release from the plant cell wall, or reactions involving sugar-phosphates (Table 1). In addition to enzymatic activity, four of the Botrytis *trans*-eQTL candidate genes encoded transcriptional regulators. *Bcin10g05900*, a putative winged helix TF is predicted to have pathway specific effects while the other three are putative general transcriptional regulators including *Bcin12g00330*, a putative Topoisomerase II-associated protein PAT1, and *Bcin09g06590*, a putative helicase (Table 1). Interestingly, the putative winged helix TF trans-eQTL hotspot alters the expression of *A. thaliana* genes associated with water deprivation responses. This suggests that this putative winged helix TF may influence a specific virulence factor that influences this *A. thaliana* network. Interestingly, while the candidate genes connect to processes that likely influence virulence, none of them have been explicitly shown to influence virulence in *B. cinerea*. Future work is necessary to test these loci and discover the mechanisms by which they influence virulence and the host/pathogen co-transcriptome.

**Complications in detection of *cis*-acting loci**

The vast majority of eQTL studies identify a strong signature of *cis*-acting loci. However, in *B. cinerea*, the main detected pattern was *trans*-eQTL with few identified *cis*-eQTL. A deeper investigation suggested that this may be due to genetic factors that complicate the ability to identify *cis*-acting SNPs. *B. cinerea* has high haplotype diversity, and in three gene clusters investigated, potential rare *cis*-acting variants fell below the minor allele cutoff for GWA potentially leading to false negative detection errors (Tabangin *et al.* 2009). Further, a number of deletions were found in this locus, which also complicates the detection of a *cis*-eQTL signature by introducing non-SNP variation unused by current GWA algorithms. Additional *cis*-acting variants may also be hidden by undetected transposon variation (Porquier *et al.* 2019).

A full understanding of the pattern of potential *cis*-acting loci in *B. cinerea* would require a detailed investigation of structural variation by incorporating long-read sequencing in a larger population size. Additionally, the GWA algorithms would need to be written to allow for simultaneous use of both SNP and presence/ absence polymorphism data; one option is to code deletions as an additional state for each genotyped variant (Wang *et al.* 2018a). It should be noted that these same difficulties would also create undetected *trans*-eQTL. This does suggest that there is likely a significant fraction of undetected *cis*-eQTLs within *B. cinerea*, caused by the high polymorphism rate within this species.

**Conclusion**

Previous work in the *B. cinerea* – *A. thaliana* pathosystem established connections between host polymorphisms and lesion growth, between gene expression and lesion size, and between transcriptomes of the host and pathogen (Corwin *et al.* 2016b; Zhang *et al.* 2017; Zhang *et al.* 2019). This study establishes a foundation to study how genetic variation in the pathogen can manipulate the host transcriptome to influence disease progression. In the Arabidopsis/Botrytis pathosystem, this connection had a preponderance of *trans*-acting polymorphisms with mainly moderate to small effects, suggesting that a polygenic architecture underlies the transcriptome variation, similar to the polygenic architecture observed for virulence. Using previously defined transcriptome modules showed that there may be a modular structure to these effects, with specific pathogen SNPs linking to specific modules in either the host or the pathogen. However, future validation work will be required to test the directionality and mechanism of this crosstalk. Similar work in other systems will help to build our functional knowledge of cross-kingdom communication between host and pathogen.

**METHODS**

**Previously published transcriptomics**

For this study, we utilized the exact previously published transcriptome data involving 1164 RNA-seq libraries available under the NCBI accession number SRP149815 (Zhang *et al.* 2017; Zhang *et al.* 2019)A brief recitation of the materials and methods used to generate the transcriptome data is provided below.(Zhang *et al.* 2017; Zhang *et al.* 2019) This large number of libraries measure the interaction of three *A. thaliana* genotypes (Col-0, coi1-1, npr1-1) with a collection of 96 *B. cinerea* genotypes that were isolated as single spores from natural infections of fruit and vegetable tissues collected in California and internationally (Zhang *et al.* 2017; Atwell *et al.* 2018b; Caseys *et al.* 2018; Zhang *et al.* 2019). We focused the analysis for this eGWA study on the *A. thaliana* wild-type accession Columbia-0 (Col-0). The transcriptome data was generated using 4-fold replication of the full-randomized complete block experimental design across two independent experiments for all interactions. Leaves from Col-0 were harvested five weeks after sowing, and individually inoculated in a detached leaf assay with spores of each of 96 *B. cinerea* isolates (Zhang *et al.* 2017; Zhang *et al.* 2019). Whole leaves were sampled at 16 hours post inoculation, prior to visible lesion formation, and flash-frozen for RNA isolation while lesion area was measured at 72 hours post inoculation.RNASeq libraries were prepared following (Kumar *et al.* 2012; Zhang *et al.* 2017; Zhang *et al.* 2019). Reads were aligned to both the *A. thaliana* TAIR10.25 and the *B. cinerea* B05.10 ASM83294v1 cDNA reference genomes, and gene counts were pulled, summed across gene models, and normalized (Langmead *et al.* 2009; Li *et al.* 2009; Van Kan *et al.* 2017).

**Genome wide association of gene expression profiles**

As phenotype for the GWA, we used the z-scaled model-adjusted least square means of normalized gene counts obtained from a negative binomial generalized linear model (nbGLM) for both the *A. thaliana* and *B. cinerea* transcriptomes (Zhang *et al.* 2017; Zhang *et al.* 2019). GWA was implemented using a univariate linear mixed model in GEMMA (Zhou and Stephens 2012). GWA was performed using PLINK binary ped format files with a centered relatedness matrix calculated in Gemma. The relatedness matrix accounts for population structure among *B. cinerea* isolates. Individual SNP significances were extracted as P-values from the score test. We used haploid binary SNP calls with less than 10% missing values and the default 1% MAF was increased to 20%. The use of MAF> 0.20 should help to limit the false positive error given the population size (Tabangin *et al.* 2009). The mapping was based 96 isolates with a total of 237,878 SNPs mapped to the *B. cinerea* B05.10 genome (Atwell *et al.* 2018a). We ran GEMMA once per phenotype, across 9,267 *B. cinerea* gene expression profiles and 23,947 *A. thaliana* gene expression profiles.(Atwell *et al.* 2018b)

**Significance threshold by GWA of permuted transcripts**

Given the large number of traits used for GWA, 33,214 traits, full permutation tests assessing the significance of all associations is unfeasible. To query for potential patterns of association that may exist randomly and fix global thresholds of likely non-random association of SNPs with transcript variation, we performed a comparative GWA analysis of randomized assignment of each transcriptional profile across the 96-isolate collection. This analysis included five independent permutations of 9,267 randomized *B. cinerea* transcripts and 23,947 randomized *A. thaliana* transcripts. To fix the P-value significance threshold, we compiled the smallest 5% *P*-values for each transcript in each permutation and then pooled those values across the five permutations. A conservative threshold was fixed as the median P-value, *p* < 1.96e-05 for *B. cinerea* and *p* < 2.90e-05 for *A. thaliana*. Permutation approaches are often more effective than a priori *P*-value thresholds to determine significance thresholds of GWA studies with large number of phenotypes (Evans and Cardon 2006).

**eQTL hotspot identification, significance and annotation**

An eQTL hotspot was defined as a SNP having the most significant effect (top SNP) on the abundance of multiple transcripts. To assess the significance of these hotspots, we used the five random permutations GWA analysis to query and measure the number of transcripts randomly associated with each SNP. The SNP with the maximum random hotspot size across any of the 5 permutations was linked to variation in 11 transcripts in *B. cinerea* and 80 transcripts in *A. thaliana*. Using this, we conservatively defined significant hotspots as SNPs associated to more than 20 transcripts in *B. cinerea* and 150 transcripts in *A. thaliana*. Furthermore, SNPs associated with more than 5 *B. cinerea* or 10 A. thaliana transcripts in the random permutation analysis were removed from eQTL hotspot analysis as SNPS likely to be false positives hotspots.

To document the putative functionality of eQTL hotspots, we annotated SNPs to genes using the identity of the nearest gene within a 2kb window. This window was chosen because the average linkage disequilibrium (LD) decay in the *B. cinerea* genome is < 1kb (Atwell *et al.* 2018). This loss of allelic association over short distances allows tagging SNPs to particular genes with some confidence. While the majority of annotated genes had single SNP hotspots, three genes had two independent SNP hotspots. (Atwell *et al.* 2018b)Two genes on chromosome 12 denoting *A. thaliana* gene expression hotspots appear closely linked; but are separated by ~80kb on the *B. cinerea* genome.

**Annotation of gene ontology and modules**

We annotated functions to *B. cinerea* genes using the BotPortal resource (<http://dx.doi.org/10.15454/IHYJCX>) and looked for patterns indicating signal peptides for secretion using the SignalP-5.0 Server (<http://www.cbs.dtu.dk/services/SignalP-5.0/>). We looked for functional overrepresentation among the genes targeted by each *A. thaliana* eQTL hotspot using the PANTHER overrepresentation test implemented by plant GO term enrichment from TAIR (Lamesch *et al.* 2011; Mi *et al.* 2013). Over-enrichment of genes found in the previous *B. cinerea* and *A. thaliana* transcriptome modules and the eQTL hotspots were tested by the hypergeometric test (Subramanian *et al.* 2005; Zhang *et al.* 2017; Zhang *et al.* 2019).

**Biosynthetic Pathway focused cis-eQTL analysis**

All of the SNPs within the three clustered biosynthetic pathways were collected for analysis (Zhang *et al.* 2019). Within each pathway, the SNPs were used to cluster the isolates by hierarchical clustering using the R package pvclust. Distances were estimated using the unweighted pair group method with arithmetic mean (UPGMA) distance, with correlation distance and 1000 bootstrap replications (Suzuki and Shimodaira 2015). The 95% bootstrap probability (BP) values are show in green while the approximately unbiased (AU) *P*-values are reported in red. Clustering is drawn according to those edges with strong support under both estimations.

**Data Availability Statement**

Strains are available upon request. Supplemental files are available at FigShare. File S1 contains information on functional annotation of the *B. cinerea* genes targeted by the *B. cinerea trans*-eQTL hotspots. File S2 contains information on functional annotation of the *A. thaliana* genes targeted by the *B. cinerea trans*-eQTL hotspots. File S3 contains gene ontology summary analysis of the *A. thaliana* genes targeted within each *B. cinerea trans*-eQTL hotspot. File S4 contains summary information on the top SNP hits from GWA of each *B. cinerea* expression trait (transcript). Code used for data analysis can be found at <https://github.com/nicolise/BcAt_RNAGWAS>.

**FIGURE LEGENDS**

**Figure 1. Manhattan plot resulting from GWAS in Botrytis genome.** Panel a: Log-scaled P-values of SNPs associations to *B. cinerea* transcript from Bcin06g05580. Dashed line indicate significance at p=1.96e-05. Panel B: Log-scaled P-values of SNPs associations to *A. thaliana* transcript from AT2G02780. Dashed lines indicate significance at p=2.9e-5.

**Figure 2. Distance to top associated SNP shows the absence of cis-diagonal and presence of trans-eQTL hotspots.**  Panel a depicts the single top SNP per transcript. Panel b depicts the top 10 SNPs per transcript. The 18 *B. cinerea* chromosomes are delimited by red bars along the x-axis, and positions indicate individual SNPs. The y-axis depicts the same chromosome alignment, but positions are the center of each mapped transcript. Vertical striping of SNP positions indicates genomic locations of putative *trans*-eQTL hotspots.

**Figure 3. Analysis of the botcinic acid gene cluster on *B. cinerea* chromosome 1.** Panel a is a SNP based hierarchical clustering of *B. cinerea* isolates. The hierarchical clustering was based on mean distance (UPGMA), with 1000 bootstrap replications. Red boxes indicate clustering with > 95% confidence. Panel b is Violin plots of botcinic acid network-level expression within *B. cinerea* groups. The mean network expression is obtained by converting expression of each gene across the isolates into its corresponding z-score and then averaging across the z-scores. Below is the gene model of the gene cluster, with each rainbow colored box delimiting single genes (including Bcboa1 to Bcboa13 and 5 additional genes) and with the cluster 4 major deletion indicated as a triangle.

**Figure 4. Frequency and positions of trans-eQTL hotspots in the *B. cinerea* genome.** Panel a is the number of Botrytis transcripts associated to trans-eQTL SNPs aligned along the 18 Botrytis chromosomes (x-axis). Panel b is the number of Arabidopsis transcripts associated to cross-species eQTL SNPs aligned along the 18 Botrytis chromosomes (x-axis).

**Figure 5. Cross-species hotspot comparison on the *B. cinerea* genome.** For each SNP that is a top hit for one or more transcripts, the number of associated transcripts is counted, across both the *B. cinerea* transcriptome and the *A. thaliana* transcriptome.

**Figure 6. Genes associated to eQTL hotspots are in virulence and defense co-expression networks.** Circles along the *B. cinerea* 18 chromosomesare eQTL hotspots, centered at the gene containing the eQTL and with radius proportional to the number of transcripts associated to this hotspot. The gene center is marked with a white dot. Hotspots for *B. cinerea* transcripts are drawn in blue, hotspots for *A. thaliana* transcripts are drawn in green. The network names are based on biological functions from gene ontology analysis of network members, from Figure 4 of Zhang *et al*. 2019 and Figure 6 of Zhang *et al*. 2017. The *A. thaliana* networks depicted are the most inclusive of the host-dependent networks, from *npr1-1*. Links between hotspots and co-expression networks are drawn according to the number of genes shared between them. Variable line weight represents the percent of hotspot target genes shared with the co-expression network; 1-25% is dashed, 25-50% is dotted, 50-75% is solid, 75-100% is heavy solid.

**TABLES  
Table 1. Annotation of the genes identified from *B. cinerea* and *A. thaliana* eQTL hotspots.** Each row identifies a significant eQTL hotspot SNP associated with transcript variations in *B. cinerea* or in *A. thaliana*. Gene functions are from BotPortal, Arabidopsis GO overrepresentation are from PANTHER. The table also presents whether the gene’s transcript is correlated with virulence or is associated with virulence via GWA from previous analysis (Zhang *et al.* 2017; Atwell *et al.* 2018b; Zhang *et al.* 2019).

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **hotspot gene** | **hotspot SNP** | **Botrytis Transcripts** | **Arabidopsis**  **Transcripts** | **Name** | **gene function** | **Virulence Correlation** | **Virulence GWA** | **Arabidopsis GO overrepresentation** |
| Bcin01g01610 | 629157 | 0 | 219 |  | Glucose/ribitol dehydrogenase | NA |  | amylopectin, glycogen, chlorophyll, chloroplast, photosynthesis |
| Bcin02g02480 | 910858 | 31 | 6 |  |  | p=0.008 |  | NA |
| Bcin02g02850 | 1032242 | 32 | 0 |  | Fructosamine-3-kinase |  |  | NA |
| Bcin03g00960 | 336467 | 1 | 122 |  | GTP cyclohydrolase I | NA |  | carotenoid biosynthesis, chloroplast organization |
| Bcin03g05020 | 1695103 | 24 | 15 |  |  | NA |  | NA |
| Bcin04g00830 | 310509 | 0 | 144 |  | NACHT nucleoside triphosphatase |  |  | nucleic acid metabolism |
| Bcin04g04700 | 1633072 | 0 | 634 |  | Heterokaryon incompatibility | NA |  | photosynthesis, light, translation |
| Bcin04g05160 | 1791561 | 3 | 114 |  |  | NA |  | chitin response |
| Bcin05g02780 | 1015184 | 22 | 2 |  |  | NA |  | NA |
| Bcin06g05680 | 1952544 | 0 | 157 |  | WLM | p < 0.001 |  |  |
| Bcin06g05790 | 1988179 | 25 | 0 |  |  | NA |  | NA |
| Bcin08g05340 | 2014228 | 27 | 0 |  | F-box domain | NA |  | NA |
| Bcin09g03390 | 1235841 | 73 | 6 |  | 6-phosphogluconate dehydrogenase | NA |  | NA |
| Bcin09g06590 | 2330312 | 106 | 56 |  | SNF2-related; Helicase, superfamily 1/2 | NA |  | metabolism |
| Bcin09g06590 | 2334368 | 23 | 0 |  | SNF2-related; Helicase, superfamily 1/2 |  |  | NA |
| Bcin10g00940 | 383007 | 23 | 1 |  |  | p < 0.001 |  | NA |
| Bcin10g05900 | 2268522 | 0 | 117 |  | Winged helix-turn-helix Transcription factor |  |  | water stress |
| Bcin12g00330 | 115491 | 37 | 1 | Bcpat1 | Topoisomerase II-associated protein PAT1 | p < 0.001 |  | NA |
| Bcin12g00330 | 115511 | 64 | 3 | Bcpat1 | Topoisomerase II-associated protein PAT1 | p < 0.001 |  | NA |
| Bcin12g02130 | 758420 | 0 | 110 |  |  | NA |  | response to stimulus |
| Bcin12g02130 | 760499 | 2 | 265 |  |  |  |  | JA, fungal response, microbe defense, biotic stress |
| Bcin12g02340 | 842369 | 1 | 449 | Bccds1 | Phosphatidate cytidylyltransferase | p < 0.001 |  | primary metabolism, amino acid biosynthesis, salt stress, biotic response |
| Bcin13g02930 | 1026752 | 0 | 123 |  | SET domain | NA |  | transport |
| Bcin16g00010 | 76596 | 23 | 2 |  | SsuA/THI5-like | NA | Yes | NA |
| Bcin16g01950 | 787512 | 0 | 240 | Bccwh41 | Glycoside hydrolase, family 63 | NA |  | photosynthesis |

**SUPPLEMENTAL FIGURE AND TABLE LEGENDS**

**Table S1. Annotation of the *B. cinerea* genetic targets of *B. cinerea* hotspots.** Columns include the hotspot SNP and the nearest gene, and all other columns pertain to the target gene modulated by the eQTL. Additional information about the target gene includes gene name, gene function, annotation as an enzyme, and target transcript. Gene functions are IPR numbers from the InterPro database.

**File S1. Functional summary of the *B. cinerea* genetic targets of *B. cinerea* hotspots.** These count occurrences of major functional categories among the hotspot target genes.

**File S2. Annotation of the *A. thaliana* genetic targets of *B. cinerea* hotspots.** Columns include the *B. cinerea* hotspot SNP and the nearest *B. cinerea* gene, and the *A. thaliana* target gene.

**File S3. Gene ontology analysis of the *A. thaliana* genetic targets of *B. cinerea* hotspots.** These include all PANTHER overrepresentation test outputs for target gene sets within each eQTL hotspot. Hotspots are labeled by SNP and nearest *B. cinerea* gene. All calculations are from Bonferroni-corrected Fisher’s exact tests, and only significant GO categories are presented.

**File S4. Distance from B. cinerea transcripts to top SNP hits.** Dataset includes the lowest 50 *P*-values of SNP-transcript associations and the top 50 effect size estimates (beta) of SNP-transcript associations. Distance column indicates whether SNP and transcript are on separate chromosomes (*trans*), or if on the same chromosome (*cis*), the distance between the SNP and the transcript center.

**Figure S1. Distribution of number of associations and *P*-values for SNP-transcript associations.** *B. cinerea* transcripts with significant SNP associations (1,616 total) are a and c, *A. thaliana* transcripts with significant SNP associations (5,213 total) are b and d. In a and b, histograms depict the distribution of number of significant SNP associations per each transcript. In c and d, histograms encompass the top 1 SNP associated with each transcript. Box edges delimit the first and third quartile, the thick center line delimits the median. Whiskers extend to 1.5 times the interquartile range and additional points indicate outliers.

**Figure S2. Manhattan-type plot of GEMMA results of transcriptome-wide *B. cinerea* expression phenotypes.** Each point represents a single transcript-SNP *P*-value of association. Panel a is a Manhattan-type plot of the top 1 SNP hit per *B. cinerea* transcript on Col-0 *A. thaliana*. Panel b is a Manhattan-type plot of the top 1 SNP hit per *A. thaliana* transcript when infected by *B. cinerea*.

**Figure S3. Distance between transcript center and top SNP location for all *B. cinerea* expression profiles on Col-0 *A. thaliana*.** Distances are in Mb, including only top SNPs on the same chromosome as the focal gene. Panel a data include the top 1 SNP identified by GEMMA association with each transcript expression profile (lowest *P*-value for association). Panel b describes the length of individual chromosomes. Panel c data include the shortest distance between transcript genomic location and top 1 SNP identified by GEMMA association with each transcript expression profile (lowest *P*-value for association) out of 5 permutations.

**Figure S4. Distribution of the rank percentile p-values for local SNPs.** Local SNPs are defined as within the transcript or within 1kb of the beginning or end of the transcript. Panel a shows the rank percentiles of local top SNPs for 9,267 *B. cinerea* transcripts. The median was 16 SNPs per transcript. Panel b shows the distribution of the local top SNP for each transcript, in black. In grey are shown five simulations of the distribution of top SNPs from 9267 random sets of 16 p-values.

**Figure S5. *cis*-effect analysis of the botrydial biosynthetic gene network.** Panel a is hierarchical clustering of *B. cinerea* isolates from SNPs within the botrydial biosynthetic gene network. Clustering was based on mean distance (UPGMA), with correlation distance and 1000 bootstrap replications. AU *P*-values are reported in red, BP values in green. Edges with high AU values are considered strongly supported by the data, and clustering is drawn according to these edges with AU > 95%. Panel b is Violin plots of botrydial network-level expression within *B. cinerea* clusters. Isolates are clustered based membership in groups defined by hierarchical clustering of the SNPs within the botrydial biosynthesis network.

**Figure S6. *cis*-effect analysis of the cyclic peptide biosynthetic gene network.** Panel a is hierarchical clustering of *B. cinerea* isolates from SNPs within the cyclic peptide biosynthetic gene network. Clustering was based on mean distance (UPGMA), with correlation distance and 1000 bootstrap replications. AU *P*-values are reported in red, BP values in green. Edges with high AU values are considered strongly supported by the data, and clustering is drawn according to these edges with AU > 95%. Panel b is Violin plots of cyclic peptide network-level expression within *B. cinerea* clusters. Isolates are clustered based membership in groups defined by hierarchical clustering of the SNPs within the botrydial biosynthesis network.

**Figure S7. Interspecific hotspot comparison on the *B. cinerea* genome with the top 10 genes per SNP.** For each SNP that is a top hit for one or more transcripts, the number of associated transcripts is counted, across both the *B. cinerea* transcriptome and the *A. thaliana* transcriptome.

**REFERENCES**

Allen, M., M. M. Carrasquillo, C. Funk, B. D. Heavner, F. Zou *et al.*, 2016 Human whole genome genotype and transcriptome data for Alzheimer’s and other neurodegenerative diseases. Scientific data 3**:** 160089.

Anuradha, C., P. M. Gaur, S. Pande, K. K. Gali, M. Ganesh *et al.*, 2011 Mapping QTL for resistance to botrytis grey mould in chickpea. Euphytica 182**:** 1-9.

Atwell, S., J. Corwin, N. Soltis and D. Kliebenstein, 2018a Resequencing and association mapping of the generalist pathogen Botrytis cinerea. bioRxiv.

Atwell, S., J. Corwin, N. Soltis, W. Zhang, D. Copeland *et al.*, 2018b Resequencing and association mapping of the generalist pathogen Botrytis cinerea. BioRxiv**:** 489799.

Barrett, L. G., J. M. Kniskern, N. Bodenhausen, W. Zhang and J. Bergelson, 2009 Continua of specificity and virulence in plant host–pathogen interactions: causes and consequences. New Phytologist 183**:** 513-529.

Bartha, I., P. J. McLaren, C. Brumme, R. Harrigan, A. Telenti *et al.*, 2017 Estimating the respective contributions of human and viral genetic variation to HIV control. PLoS computational biology 13**:** e1005339.

Bartoli, C., and F. Roux, 2017 Genome-Wide Association Studies In Plant Pathosystems: Toward an Ecological Genomics Approach. Frontiers in plant science 8.

Bilgin, D. D., J. A. Zavala, J. Zhu, S. J. Clough, D. R. Ort *et al.*, 2010 Biotic stress globally downregulates photosynthesis genes. Plant, cell & environment 33**:** 1597-1613.

Brem, R. B., G. Yvert, R. Clinton and L. Kruglyak, 2002 Genetic dissection of transcriptional regulation in budding yeast. Science 296**:** 752-755.

Caseys, C., G. Shi, N. Soltis, R. Gwinner, J. Corwin *et al.*, 2018 A generalist pathogen view of plant evolution. bioRxiv**:** 507491.

Chan, E. K., H. C. Rowe, B. G. Hansen and D. J. Kliebenstein, 2010 The complex genetic architecture of the metabolome. PLoS Genet 6**:** e1001198.

Chen, X., C. A. Hackett, R. E. Niks, P. E. Hedley, C. Booth *et al.*, 2010 An eQTL analysis of partial resistance to Puccinia hordei in barley. PLoS One 5**:** e8598.

Christie, N., A. A. Myburg, F. Joubert, S. L. Murray, M. Carstens *et al.*, 2017 Systems genetics reveals a transcriptional network associated with susceptibility in the maize–grey leaf spot pathosystem. The Plant Journal 89**:** 746-763.

Colmenares, A. J., J. Aleu, R. Duran-Patron, I. G. Collado and R. Hernandez-Galan, 2002 The putative role of botrydial and related metabolites in the infection mechanism of *Botrytis cinerea*. Journal of Chemical Ecology 28**:** 997-1005.

Corwin, J. A., D. Copeland, J. Feusier, A. Subedy, R. Eshbaugh *et al.*, 2016a The quantitative basis of the Arabidopsis innate immune system to endemic pathogens depends on pathogen genetics. PLoS Genet 12**:** e1005789.

Corwin, J. A., A. Subedy, R. Eshbaugh and D. J. Kliebenstein, 2016b Expansive phenotypic landscape of Botrytis cinerea shows differential contribution of genetic diversity and plasticity. Molecular Plant-Microbe Interactions 29**:** 287-298.

Cui, H., K. Tsuda and J. E. Parker, 2015 Effector-triggered immunity: from pathogen perception to robust defense. Annual review of plant biology 66**:** 487-511.

Dalmais, B., J. Schumacher, J. Moraga, P. Le Pecheur, B. Tudzynski *et al.*, 2011 The Botrytis cinerea phytotoxin botcinic acid requires two polyketide synthases for production and has a redundant role in virulence with botrydial. Molecular Plant Pathology 12**:** 564-579.

Deighton, N., I. Muckenschnabel, A. J. Colmenares, I. G. Collado and B. Williamson, 2001 Botrydial is produced in plant tissues infected by Botrytis cinerea. Phytochemistry 57**:** 689-692.

Denby, K. J., P. Kumar and D. J. Kliebenstein, 2004a Identification of Botrytis cinerea susceptibility loci in Arabidopsis thaliana. The Plant Journal 38**:** 473-486.

Denby, K. J., P. Kumar and D. J. Kliebenstein, 2004b Identification of *Botrytis cinerea* susceptibility loci in *Arabidopsis thaliana*. Plant J 38**:** 473-486.

Dong, S., S. Raffaele and S. Kamoun, 2015 The two-speed genomes of filamentous pathogens: waltz with plants. Current opinion in genetics & development 35**:** 57-65.

Evans, D. M., and L. R. Cardon, 2006 Genome-wide association: a promising start to a long race. Trends in Genetics 22**:** 350-354.

Finkers, R., Y. L. Bai, P. van den Berg, R. van Berloo, F. Meijer-Dekens *et al.*, 2008 Quantitative resistance to Botrytis cinerea from Solanum neorickii. Euphytica 159**:** 83-92.

Finkers, R., P. van den Berg, R. van Berloo, A. ten Have, A. W. van Heusden *et al.*, 2007 Three QTLs for *Botrytis cinerea* resistance in tomato. Theoretical and Applied Genetics 114**:** 585-593.

Fordyce, R. F., N. E. Soltis, C. Caseys, R. Gwinner, J. A. Corwin *et al.*, 2018 Digital imaging combined with genome-wide association mapping links loci to Plant-Pathogen interaction traits. Plant Physiol 178**:** 1406-1422.

Fu, Y. Q., A. van Silfhout, A. Shahin, R. Egberts, M. Beers *et al.*, 2017 Genetic mapping and QTL analysis of Botrytis resistance in Gerbera hybrida. Molecular Breeding 37.

Giraldo, M. C., and B. Valent, 2013 Filamentous plant pathogen effectors in action. Nature Reviews Microbiology 11**:** 800.

Glazebrook, J., 2005 Contrasting mechanisms of defense against biotrophic and necrotrophic pathogens. Annu. Rev. Phytopathol. 43**:** 205-227.

Goss, E. M., and J. Bergelson, 2006 Variation in resistance and virulence in the interaction between Arabidopsis thaliana and a bacterial pathogen. Evolution 60**:** 1562-1573.

Guo, Y., S. Fudali, J. Gimeno, P. DiGennaro, S. Chang *et al.*, 2017 Networks underpinning symbiosis revealed through cross-species eQTL mapping. Genetics**:** genetics. 117.202531.

Hsu, J., and J. D. Smith, 2012 Genome wide studies of gene expression relevant to coronary artery disease. Current opinion in cardiology 27**:** 210.

Jiang, Z., F. He and Z. Zhang, 2017 Large-scale transcriptome analysis reveals arabidopsis metabolic pathways are frequently influenced by different pathogens. Plant molecular biology 94**:** 453-467.

Keurentjes, J. J., J. Fu, I. R. Terpstra, J. M. Garcia, G. van den Ackerveken *et al.*, 2007 Regulatory network construction in Arabidopsis by using genome-wide gene expression quantitative trait loci. Proceedings of the National Academy of Sciences 104**:** 1708-1713.

Kliebenstein, D. J., M. A. West, H. van Leeuwen, O. Loudet, R. W. Doerge *et al.*, 2006 Identification of QTLs controlling gene expression networks defined a priori. Bmc Bioinformatics 7**:** 308.

Kou, Y., and S. Wang, 2010 Broad-spectrum and durability: understanding of quantitative disease resistance. Current opinion in plant biology 13**:** 181-185.

Kumar, R., Y. Ichihashi, S. Kimura, D. H. Chitwood, L. R. Headland *et al.*, 2012 A high-throughput method for Illumina RNA-Seq library preparation. Frontiers in plant science 3.

Lamesch, P., T. Z. Berardini, D. Li, D. Swarbreck, C. Wilks *et al.*, 2011 The Arabidopsis Information Resource (TAIR): improved gene annotation and new tools. Nucleic acids research 40**:** D1202-D1210.

Langmead, B., C. Trapnell, M. Pop and S. L. Salzberg, 2009 Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. Genome biology 10**:** R25.

Lannou, C., 2012 Variation and selection of quantitative traits in plant pathogens. Annual Review of Phytopathology 50**:** 319-338.

Li, H., B. Handsaker, A. Wysoker, T. Fennell, J. Ruan *et al.*, 2009 The sequence alignment/map format and SAMtools. Bioinformatics 25**:** 2078-2079.

Lo Presti, L., D. Lanver, G. Schweizer, S. Tanaka, L. Liang *et al.*, 2015 Fungal effectors and plant susceptibility. Annual review of plant biology 66**:** 513-545.

Marone, D., M. Russo, G. Laidò, A. De Leonardis and A. Mastrangelo, 2013 Plant nucleotide binding site–leucine-rich repeat (NBS-LRR) genes: active guardians in host defense responses. International journal of molecular sciences 14**:** 7302-7326.

Martínez-Soto, D., A. M. Robledo-Briones, A. A. Estrada-Luna and J. Ruiz-Herrera, 2013 Transcriptomic analysis of U stilago maydis infecting Arabidopsis reveals important aspects of the fungus pathogenic mechanisms. Plant signaling & behavior 8**:** e25059.

Meng, X., and S. Zhang, 2013 MAPK cascades in plant disease resistance signaling. Annual review of phytopathology 51**:** 245-266.

Mi, H., A. Muruganujan, J. T. Casagrande and P. D. Thomas, 2013 Large-scale gene function analysis with the PANTHER classification system. Nature protocols 8**:** 1551.

Monks, S., A. Leonardson, H. Zhu, P. Cundiff, P. Pietrusiak *et al.*, 2004 Genetic inheritance of gene expression in human cell lines. The American Journal of Human Genetics 75**:** 1094-1105.

Nobori, T., A. C. Velásquez, J. Wu, B. H. Kvitko, J. M. Kremer *et al.*, 2018 Transcriptome landscape of a bacterial pathogen under plant immunity. Proceedings of the National Academy of Sciences 115**:** E3055-E3064.

Nomura, K., M. Melotto and S.-Y. He, 2005 Suppression of host defense in compatible plant–Pseudomonas syringae interactions. Current opinion in plant biology 8**:** 361-368.

Pinedo, C., C.-M. Wang, J.-M. Pradier, B. Dalmais, M. Choquer *et al.*, 2008 Sesquiterpene synthase from the botrydial biosynthetic gene cluster of the phytopathogen Botrytis cinerea. ACS chemical biology 3**:** 791-801.

Poland, J. A., P. J. Balint-Kurti, R. J. Wisser, R. C. Pratt and R. J. Nelson, 2009 Shades of gray: the world of quantitative disease resistance. Trends in plant science 14**:** 21-29.

Porquier, A., J. Moraga, G. Morgant, B. Dalmais, A. Simon *et al.*, 2019 Botcinic acid biosynthesis in Botrytis cinerea relies on a subtelomeric gene cluster surrounded by relics of transposons and is regulated by the Zn 2 Cys 6 transcription factor BcBoa13. Current genetics**:** 1-16.

Porquier, A., G. Morgant, J. Moraga, B. Dalmais, I. Luyten *et al.*, 2016 The botrydial biosynthetic gene cluster of Botrytis cinerea displays a bipartite genomic structure and is positively regulated by the putative Zn (II) 2Cys6 transcription factor BcBot6. Fungal genetics and biology 96**:** 33-46.

Rivas, M. A., M. Beaudoin, A. Gardet, C. Stevens, Y. Sharma *et al.*, 2011 Deep resequencing of GWAS loci identifies independent rare variants associated with inflammatory bowel disease. Nature genetics 43**:** 1066.

Roux, F., D. Voisin, T. Badet, C. Balagué, X. Barlet *et al.*, 2014 Resistance to phytopathogens e tutti quanti: placing plant quantitative disease resistance on the map. Molecular plant pathology 15**:** 427-432.

Rowe, H. C., and D. J. Kliebenstein, 2008 Complex genetics control natural variation in *Arabidopsis thaliana* resistance to *Botrytis cinerea*. Genetics 180**:** 2237-2250.

Saeij, J. P. J., S. Coller, J. P. Boyle, M. E. Jerome, M. W. White *et al.*, 2007 Toxoplasma co-opts host gene expression by injection of a polymorphic kinase homologue. Nature 445**:** 324-327.

Schadt, E. E., S. A. Monks, T. A. Drake, A. J. Lusis, N. Che *et al.*, 2003 Genetics of gene expression surveyed in maize, mouse and man. Nature 422**:** 297.

Schumacher, J., J. M. Pradier, A. Simon, S. Traeger, J. Moraga *et al.*, 2012 Natural Variation in the VELVET Gene bcvel1 Affects Virulence and Light-Dependent Differentiation in Botrytis cinerea. Plos One 7.

Siewers, V., M. Viaud, D. Jimenez-Teja, I. G. Collado, C. S. Gronover *et al.*, 2005 Functional analysis of the cytochrome P450 monooxygenase gene bcbot1 of Botrytis cinerea indicates that botrydial is a strain-specific virulence factor. Molecular plant-microbe interactions 18**:** 602-612.

Soltis, N. E., S. Atwell, G. Shi, R. Fordyce, R. Gwinner *et al.*, 2019 Crop domestication and pathogen virulence: Interactions of tomato and *Botrytis* genetic diversity. The Plant Cell 31**:** 502-519.

St. Clair, D. A., 2010 Quantitative disease resistance and quantitative resistance loci in breeding. Annual review of phytopathology 48**:** 247-268.

Subramanian, A., P. Tamayo, V. K. Mootha, S. Mukherjee, B. L. Ebert *et al.*, 2005 Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles. Proceedings Of The National Academy Of Sciences Of The United States Of America 102**:** 15545-15550.

Suzuki, R., and H. Shimodaira, 2015 pvclust: Hierarchical Clustering with P-Values via Multiscale Bootstrap Resampling. . R package version 2.0-0.

Tabangin, M. E., J. G. Woo and L. J. Martin, 2009 The effect of minor allele frequency on the likelihood of obtaining false positives. BMC proceedings 3 Suppl 7**:** S41-S41.

Valero-Jiménez, C. A., J. Veloso, M. Staats and J. A. van Kan, 2019 Comparative genomics of plant pathogenic Botrytis species with distinct host specificity. BMC Genomics 20**:** 203.

Van Kan, J. A., J. H. Stassen, A. Mosbach, T. A. Van Der Lee, L. Faino *et al.*, 2017 A gapless genome sequence of the fungus Botrytis cinerea. Molecular plant pathology 18**:** 75-89.

Visscher, P. M., N. R. Wray, Q. Zhang, P. Sklar, M. I. McCarthy *et al.*, 2017 10 years of GWAS discovery: biology, function, and translation. The American Journal of Human Genetics 101**:** 5-22.

Wang, M., F. Roux, C. Bartoli, C. Huard-Chauveau, C. Meyer *et al.*, 2018a Two-way mixed-effects methods for joint association analysis using both host and pathogen genomes. Proceedings of the National Academy of Sciences 115**:** E5440-E5449.

Wang, X. F., Q. Y. Chen, Y. Y. Wu, Z. H. Lemmon, G. H. Xu *et al.*, 2018b Genome-wide Analysis of Transcriptional Variability in a Large Maize-Teosinte Population. Molecular Plant 11**:** 443-459.

West, M. A. L., K. Kim, D. J. Kliebenstein, H. van Leeuwen, R. W. Michelmore *et al.*, 2007 Global eQTL mapping reveals the complex genetic architecture of transcript level variation in Arabidopsis. Genetics 175**:** 1441-1450.

Wu, J., B. Cai, W. Sun, R. Huang, X. Liu *et al.*, 2015 Genome-wide analysis of host-Plasmodium yoelii interactions reveals regulators of the type I interferon response. Cell reports 12**:** 661-672.

Wu, J. Q., S. Sakthikumar, C. Dong, P. Zhang, C. A. Cuomo *et al.*, 2017 Comparative genomics integrated with association analysis identifies candidate effector genes corresponding to Lr20 in phenotype-paired Puccinia triticina isolates from Australia. Frontiers in plant science 8.

Zhang, W., J. A. Corwin, D. Copeland, J. Feusier, R. Eshbaugh *et al.*, 2017 Plastic transcriptomes stabilize immunity to pathogen diversity: The jasmonic acid and salicylic acid networks within the Arabidopsis/Botrytis pathosystem. Plant Cell 29**:** 2727-2752.

Zhang, W., J. A. Corwin, D. Copeland, J. Feusier, R. Eshbaugh *et al.*, 2019 Plant-Necrotroph Co-transcriptome Networks Illuminate a Metabolic Battlefield. eLife 8**:** e44279.

Zhou, X., and M. Stephens, 2012 Genome-wide efficient mixed-model analysis for association studies. Nature genetics 44**:** 821.

Zou, F., H. S. Chai, C. S. Younkin, M. Allen, J. Crook *et al.*, 2012 Brain expression genome-wide association study (eGWAS) identifies human disease-associated variants. PLoS genetics 8**:** e1002707.