**RESULTS**

We performed genome-wide association (GWA) for eQTL detection across all genes expressed in the *B. cinerea* - *A. thaliana* pathosystem, including transcripts from the host and from the pathogen. Previous analysis showed that the vast majority of transcripts are affected by variation in the *B. cinerea* genome. Of the 9,284 differentially expressed *B. cinerea* genes, 74% (5,244) showed significant variation in their expression due to *B. cinerea* genetic variation, with an average broad-sense heritability of 0.152 attributed to *B. cinerea* isolate (Zhang, Corwin et al. 2018). Of the 23,898 differentially expressed *A. thaliana* genes, 85% (20,328) showed significant variation in their expression due to *B. cinerea* genetics, with an average broad-sense heritability of 0.108 due to *B. cinerea* isolate (Zhang, Corwin et al. 2017). As such, we conducted genome-wide association (GWA) to look for evidence of loci in *B. cinerea* that may be modulating this heritable expression variation.

**eQTL indicate polygenic transcriptome modulation**

To look for expression quantitative trait loci (eQTL) for both host and pathogen, we associated the expression profiles of these 9,267 *B. cinerea* genes and 23,947 *A. thaliana* genes to genome-wide SNP variation in the *B. cinerea* genome. We performed Genome-wide Efficient Mixed Model Association (GEMMA) (Zhou and Stephens 2012) to rapidly estimate the significance of all markers for each expression profile. We first controlled for the effects of population structure within our *B. cinerea* isolates by calculating and including a relatedness matrix in the downstream analysis. GEMMA estimates the significance of effects of each SNP on the focal trait as a p-value. For individual expression traits, we find 0 to XX loci with significant p-values under XX for *B. cinerea*, and 0 to XX loci for *A. thaliana* transcripts (Figure N1). In total, genetic variation in *B. cinerea* appeared to significantly affect the expression of X *A. thaliana* genes and X *B. cinerea* genes. Overall, we find a highly polygenic basis of loci modulating transcriptome variation.

**Randomization to define hotspots**

To validate SNPs as significantly associated with transcript variation, we performed a comparative analysis of randomized phenotypes. Taking each transcriptional profile, we randomized the assignment of phenotypes across the 96-isolate collection. This analysis includes 9,267 randomized *B. cinerea* phenotypes and 23,947 randomized *A. thaliana* phenotypes, one from each measured expression profile. We repeated this randomization in a 5x permutation. We ran GEMMA on each of these permutations, and plotted SNP p-value vs. position (Figure N5). Permutation approaches are often more effective than p-value thresholding for determining significance across GWA studies with many phenotypes {CITE}.

**Lack of evidence for cis-effect patterns transcriptome-wide**

Previous eQTL mapping studies show evidence for large-effect *cis-*eQTL {CITE}. To search for a pattern of transcriptome-wide *cis* effects in this study, we calculated the distance between the center of each transcript and the top associated SNP. If *cis*- acting loci contribute the bulk 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 observe that distances between transcript center and top SNP as far as 2 Mb are common (Figure N3). These distances are similar to those from the association of random transcript profiles to top SNPs (Figure SX1). As such, we do not see evidence for *cis*-effect loci overrepresented in the top candidates for control of expression variation. Rather, most control of gene expression variation in *B. cinerea* on *A. thaliana* appears to be *trans*-acting.

In these transcript-to-SNP associations, we find that *cis*- control of gene expression is largely drowned out by patterns of *trans*-acting variation. We first focused on the single top SNP hit per transcript, with the highest probability (lowest p-value) of significant effect on expression in the gene of interest. If control of gene expression is localized to the gene itself or to *cis*-acting loci, we would expect a strong linear (*cis*-diagonal) association between the center of each gene and the genomic location of its top SNP hit. However, we find that few genes have a top SNP hit within the same chromosome, and even fewer within 1Mb (Figure N2). We again conclude that most of this genetic variation is *trans*-acting, as we do not see a strong *cis*-diagonal signal when comparing transcript center to top SNP hit. This pattern holds whether we examine the top 1 SNP per transcript (Figure N2a) or the top 10 SNPs per transcript (FigureN2b).

**Search for cis effects through focus on gene networks with presence-absence polymorphism**

We did not see enrichment for local SNPs as the top loci controlling expression variation; rather, the top SNPs are distributed across the genome. Thus, we are mostly detecting *trans*-eQTL. As we did not detect a transcriptome-wide pattern of *cis*-effect variation controlling expression patterns of *B. cinerea* on *A. thaliana*, we narrowed our focus to a subset of localized biosynthetic gene networks with known presence-absence polymorphism to identify the role of *cis*-eQTL. These focal networks were among those most highly conserved across *B. cinerea* isolates (Zhang, Corwin et al. 2018). These include the botcynic acid biosynthetic pathway (13 genes, 55.8 kb), botrydial biosynthesis (7 genes, 26 kb), and Network 5 (Zhang, Corwin et al. 2018), which contains XX (10 genes, 46.5 kb).

We focused analysis on the botcynic acid biosynthesis network, which has a known presence-absence polymorphism. Upon visualizing SNP-level variation surrounding genes in this pathway, we found evidence of a deletion common to 12 of the 96 isolates (Figure N4c). The major deletion extends 53.5 kb and includes SNP 4kb from the 5’ end of the chromosome, indicating a teleomeric loss on chromosome 1. We selected a focal region encompassing the deletion endpoints (1.4029, 1.82614) and an additional 2 genes beyond the deletion boundaries (Bcin01g00170, Bcin01g00190) (Figure N4c). We removed 10 SNPs that were likely miscalled (SNP state ~ inverse compared to surrounding region) and called all SNPs within the deletion region as missing. We find evidence of many haplotypes within the pathway. Hierarchical clustering based on these loci assigned the 96 isolates into three major clusters, and one small two-isolate cluster (B05.10, Fd1) (Figure N4a). The largest cluster (cluster 4) includes a sub-cluster containing all of the isolates in the deletion group (cluster 3). We removed isolates with the whole-region deletion, then performed ANOVA across the 3 major clusters (1, 4, 5) to see whether SNP variation within the botcynic acid biosynthetic network predicts expression level of genes within the region. We found no significant effect of cluster membership on expression profile (F(1,74)=0.36, p=0.55). This suggests that the major *cis*-effect loci controlling expression variation in this network are not captured by SNP-level variation. Rather, our GWA analysis misses the major cis-eQTL signal of the network deletion.

To look for patterns of expression variation that may be controlled by undetected *cis*-acting loci, we examined mean botcynic acid pathway expression across all isolates, grouped by cluster membership. Expression levels across the network are reduced in the deletion group (cluster 3) but we also see independent outliers of low-expression isolates with additional loss-of-function polymorphisms (Noble Rot, 01.04.03, Apple 517, 02.04.09) (Figure N4b). These isolates also contain deletions within the botcynic acid biosynthetic network (Figure N4c), likely of independent origin. As such, the major *cis­*-effect variation controlling expression in the botcynic acid region appears to be explained by deletion events, which are not captured by our SNP analysis. If insertion and deletion events account for the majority of localized control of expression variation, our GWA analysis will not detect these *cis*-effect loci.

The botrydial biosynthetic network, and network 5, exhibit similar cis-effect patterns to the botcynic acid biosynthetic network. Hierarchical clustering within each of these networks by genic SNP variation divides the isolate population into two groups, and mean pathway expression across all isolates is not differentiated by this clustering for either network (Figure S1, Figure S2). We find that SNP state does not detect the major cis-effects polymorphisms… We conclude that large cis effect loci exist for transcriptional regulation in this pathosystem, but that GWA across SNPs misses these patterns. This missing cis-effect likely amplifies the apparent magnitude of trans-acting loci.

**Hotspots indicate mostly trans-eQTL transcriptome-wide**

To search for hotspots of eQTL, we plotted the p-value and location of the top SNP with the strongest evidence (lowest p-value) of association by GEMMA per transcript. We defined hotspot permutation thresholds from the maximum hotspot size across the 5x permutations and applied these thresholds to all data to define significant hotspots. The maximum permuted hotspot size for *B. cinerea* was XX, and XX for *A. thaliana*. We conservatively set our permutation thresholds to 20 linked transcripts for *B. cinerea* and 150 transcripts for *A. thaliana*. As such, hotspots are defined as peaks of significant SNPs above the hotspot permutation threshold. The significant regulatory hotspots we identified are spread throughout the genome, present on chromosomes 2-6, 8-10, 12-15 (Figure N6). These hotspots of controlling variation may be due to a mix of *cis*- and *trans*-eQTL. Genome-wide, we identified 22 hotspots, 11 linked to *B. cinerea* gene expression and 11 linked to *A. thaliana* gene expression, from X to X per chromosome (Figure N6, Figure N7). Hotspots were defined to the gene level, with 1 to X significant SNPs. To determine whether few hotspots were identified due to our SNP selection approach, we repeated the full analysis by selecting the top 10 SNPs per transcript. We found that few major hotspots are identified by this expanded approach (Table SX1; Figure SX3).

To look for *B. cinerea* loci that are *trans*-acting to control expression variation in the affected host, we also examined patterns of association between *A. thaliana* expression variation and *B. cinerea* genomic variation. We identified hotspots putatively controlling *A. thaliana* expression variation for all *B. cinerea* chromosomes except 17 and 18 (Figure X4b). These *B. cinerea* loci may contribute to regulation of gene expression in the host. We looked for evidence of eQTL hotspots that are common across both *B. cinerea* and *A. thaliana* (Figure XX). We find that X% of hotspots are shared among both transcriptomes, X% are unique to *B. cinerea*, and X% are unique to *A. thaliana.*

**Gene functions of expression modulating hotspots**

We mapped genes to selected hotspots to identify possible functions associated with our eQTL. At hotspot XX shared between *B. cinerea* and *A. thaliana*…

**Detection and annotation of *trans*-eQTL hotspots**

Our 11 significant *trans­-*eQTL hotspots for *B. cinerea* range from 22 to 129 linked genes, with minimal overlap to expression modulation in *A. thaliana* (a maximum of 56 genes). These hotspots are dispersed across the genome, at least 0.1 Mb apart and across 9 chromosomes (Figure N7). The 11 significant cross-species *trans*-eQTL hotspots for *A. thaliana* are also dispersed across the genome of *B. cinerea*, covering 8 chromosomes with at least 0.1 Mb between hotspots (Figure N7). These range from 114 to 634 linked *A. thaliana* transcripts, with very low overlap with *B. cinerea* transcripts (a maximum of 3 genes).

**Annotation of eQTL hotspots**

We annotated the genes at these eQTL hotspots with functional information, including links to coexpression networks from previous RNAseq analysis of these transcripts (Zhang, Corwin et al. 2017, Zhang, Corwin et al. 2018) (Table X1). Among the 11 *A. thaliana* hotspot trans-eQTL, two of these genes are correlated to *B. cinerea* lesion size across all tested *A. thaliana* genotypes, including immune pathway mutants. An additional gene is correlated with lesion size variation on Col-0 *A. thaliana*, and one on *coi1-1* *A. thaliana.* Among the 11 *B. cinerea* hotspot *trans-*eQTL, three of these genes are linked to lesion size variation across all tested *A. thaliana* genotypes. A fourth gene is linked to lesion size variation on Col-0 *A. thaliana*. This gene is also linked to the major vesicle/ virulence network of *B. cinerea* coexpression on Col-0 *A. thaliana* (Zhang, Corwin et al. 2018). If these eQTL are modulating expression of many genes, and affecting lesion size, they may be major *B. cinerea* control points in the plant-pathogen interaction.

**eQTL correlation to co-expression networks**

Previous analysis of these RNAseq data clustered transcripts in *B. cinerea* and *A. thaliana* into coexpression networks (Zhang, Corwin et al. 2017, Zhang, Corwin et al. 2018). eQTL hotspots linked to these coexpression networks could indicate regulatory points for these modules of expression variation. We looked for gene overlap between transcripts linked to eQTL hotspots and coexpression networks. Nine of the *B. cinerea* eQTL hotspots were also linked to genes in one or more of four major *B. cinerea* coexpression networks on *A. thaliana* (Figure N8). In particular, two of these were host-specific networks functionally associated with virulence; in total, 7 of the 11 *B. cinerea* eQTL hotspots was associated with one of these virulence coexpression networks. As such, we hypothesize that these major points of *B. cinerea* gene expression modulation may also exhibit regulation of virulence strategies on *A. thaliana*.

Nine of the *A. thaliana* eQTL hotspots were also linked to genes in one or more of six major genotype-dependent *A. thaliana* coexpression networks when infected with *B. cinerea* (Figure 9). These networks are host genotype-dependent, and either contain genes pointing to network function in jasmonate and salicylic acid signaling processes and camalexin biosynthesis (Network I), or photosynthesis in the host (Network IV). We observe particularly strong links between 4 of the 11 eQTL hotspots of *B. cinerea* modulation of *A. thaliana* gene expression and immune networks in *A. thaliana* (Figure N8). These frequent links suggest that our eQTL hotspots may exhibit regulatory control over coexpressed modules of genes active in virulence interactions between *B. cinerea* and its host.

**METHODS**

**Experimental design**

We used a previously described collection of *B. cinerea* genotypes that were isolated as single spores from natural infections of fruit and vegetable tissues collected in California and internationally (Atwell, Corwin et al. 2015, Zhang, Corwin et al. 2017, Fordyce, Soltis et al. 2018, Zhang, Corwin et al. 2018). We focused analysis on the *A. thaliana* accession Columbia-0 (Col-0), and all plants were grown as described in a previous study, with 4-fold replication of the full randomized complete block experimental design across two independent experiments (Zhang, Corwin et al. 2017, Zhang, Corwin et al. 2018). Leaves were harvested 5 weeks after sowing, and inoculated in a detached leaf assay with spores of each of 96 *B. cinerea* isolates (Zhang, Corwin et al. 2017, Zhang, Corwin et al. 2018).

**Expression analysis**

RNASeq libraries were prepared as previously described (Kumar, Ichihashi et al. 2012, Zhang, Corwin et al. 2017, Zhang, Corwin et al. 2018). Briefly, we prepared mRNA from leaves frozen at 16 hours post inoculation, and pooled amplified, size-selected libraries into four replicate groups of 96 barcoded libraries. Sequencing was completed on a single Illumina HiSeq 2500 (San Diego, CA) lane as single 50bp reads at the U.C. Davis Genome Center- DNA Technologies Core (Davis, CA). Individual libraries were then separated by adapter index from fastq files, evaluated for read quality and overrepresentation (FastQC Version 0.11.3, www.bioinformatics.babraham.ac.uk/projects/), and trimmed (fastx, http://hannonlab.cshl.edu/fastx\_toolkit/commandline.html). Reads were aligned to the *A. thaliana* TAIR10.25 cDNA reference genome, followed by the *B. cinerea* B05.10 cDNA reference genome, and we pulled gene counts (Langmead, Trapnell et al. 2009, Li, Handsaker et al. 2009, Van Kan, Stassen et al. 2017, Zhang, Corwin et al. 2017, Zhang, Corwin et al. 2018). We summed counts across gene models, and normalized gene counts as previously described (Zhang, Corwin et al. 2017, Zhang, Corwin et al. 2018).

We used as input the model-adjusted means per transcript from previously published studies in the *A. thaliana* transcriptome and *B. cinerea* transcriptome (Zhang, Corwin et al. 2017, Zhang, Corwin et al. 2018). *A. thaliana* and *B. cinerea* transcript phenotypes were from least square means of normalized gene counts in a negative binomial generalized linear model (nbGLM) (Zhang, Corwin et al. 2017, Zhang, Corwin et al. 2018). We calculated linear models from the transcript data including the effects of isolate and host genotype. We z-scaled all transcript profiles prior to GWA.

**Genome wide association**

For GEMMA mapping, we used 95 isolates with a total of 237,878 SNPs against the *B. cinerea* B05.10 genome (Atwell, Corwin et al. 2018). We used haploid binary SNP calls with MAF > 0.20 and <20% missingness. We ran GEMMA once per phenotype, across 9,267 *B cinerea* gene expression profiles and 23,947 *A. thaliana* gene expression profiles.

**Permutation and defining hotspots**

We plotted the number of transcripts linked to each SNP, summed across all 5 permutations, to calculate permuted hotspot size. For any SNPs that linked to permuted hotspots of over 5 transcripts in *B. cinerea* or 20 transcripts in *A. thaliana*, we removed these SNPs from further analysis as likely false positives. We then conservatively defined actual hotspots as SNP peaks exceeding 20 transcripts in *B. cinerea* and 150 transcripts in *A. thaliana*. We then collapsed hotspots into genes, such that all SNPs were annotated to the nearest gene within 2kb.

**Pathway focus**

We focused further *cis*-effects analysis on three networks which were highly conserved across *B. cinerea* isolates (Zhang, Corwin et al. 2018). We clustered isolates by SNP data within focal networks. Hierarchical clustering was computed using the R package pvclust based on mean linkage (UPGMA), with correlation distance and 1000 bootstrap replications (Suzuki and Shimodaira 2015). 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%.

**FIGURE LEGENDS**

**Figure N1. Manhattan plot examples for 1 transcript per species.** Panel a is an example plot of p-values for all *B. cinerea* SNP associations to a single *B. cinerea* transcript, from Bcin01g00170. Panel b is an example plot of p-values for all *B. cinerea* SNP associations to a single *A. thaliana* transcript.

**Figure N2. *cis*-diagonal plot comparing *B. cinerea* gene center to position of top associated SNP, for all 9,284 transcripts.** We retained only the SNPs with highest probability (lowest p-value) of significant effect on expression for each transcript. Panel a depicts the single top SNP per transcript. Panel b depicts the top 10 SNPs per transcript. Chromosomes are delimited by red bars along the x-axis.

**Figure N3. Distance between transcript center and top SNP location for all *B. cinerea* expression profiles on Col-0 *A. thaliana*.** Data include the top 1 SNP identified by GEMMA association with each transcript expression profile (lowest p-value for association). Distances are in Mb, including only top SNPs on the same chromosome as the focal gene.

**Figure N4. Cis-effect analysis of the botcynic acid biosynthetic gene network.** Panel a is Hierarchical clustering of *B. cinerea* isolates from SNPs within the botcynic acid biosynthetic gene network. Clustering was based on mean linkage (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 botcynic acid network-level expression within *B. cinerea* clusters. Isolates are clustered based membership in groups defined by hierarchical clustering of the SNPs within the botcynic acid biosynthesis network (Figure X5). Panel c is the gene models of the biosynthetic gene network, with the cluster 3 deletion indicated as a triangle.

**Figure N5. Manhattan-type plot of GEMMA results of transcriptome-wide *B. cinerea* expression phenotypes.** 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 N6. All eQTL hotspots across the *B. cinerea* and *A. thaliana* transcriptomes.** We counted the number of genes (transcripts) associated with each SNP. Panel a is for all *B. cinerea* transcripts, panel b is all *A. thaliana* transcripts.

**Figure N7. Interspecific 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 N8. Genes linked to eQTL hotspots are in virulence and defense coexpression networks.** Circles along the *B. cinerea* genome map are hotspots, centered at the gene containing the eQTL and with radius proportional to the number of transcripts linked to this hotspot. eQTL for *B. cinerea* transcripts are drawn in blue, eQTL for *A. thaliana* transcripts are drawn in green. Links between hotspots and coexpression networks are drawn according to the number of genes shared between them, with variable widths; 2 pixels for 1 to 5 genes shared, 5 pixels for 6 to 10 genes, 10 pixels for 11 to 19 genes, 15 pixels for 20 to 114 genes.

REFERENCES

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

Atwell, S., J. Corwin, N. Soltis, A. Subedy, K. Denby and D. J. Kliebenstein (2015). "Whole genome resequencing of Botrytis cinerea isolates identifies high levels of standing diversity." Frontiers in microbiology **6**: 996.

Fordyce, R., N. Soltis, C. Caseys, G. Gwinner, J. Corwin, S. Atwell, D. Copeland, J. Feusier, A. Subedy, R. Eshbaugh and D. Kliebenstein (2018). "Combining Digital Imaging and GWA Mapping to Dissect Visual Traits in Plant/Pathogen Interactions." Plant Physiology.

Kumar, R., Y. Ichihashi, S. Kimura, D. H. Chitwood, L. R. Headland, J. Peng, J. N. Maloof and N. R. Sinha (2012). "A high-throughput method for Illumina RNA-Seq library preparation." Frontiers in plant science **3**.

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**(3): R25.

Li, H., B. Handsaker, A. Wysoker, T. Fennell, J. Ruan, N. Homer, G. Marth, G. Abecasis and R. Durbin (2009). "The sequence alignment/map format and SAMtools." Bioinformatics **25**(16): 2078-2079.

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

Van Kan, J. A., J. H. Stassen, A. Mosbach, T. A. Van Der Lee, L. Faino, A. D. Farmer, D. G. Papasotiriou, S. Zhou, M. F. Seidl and E. Cottam (2017). "A gapless genome sequence of the fungus Botrytis cinerea." Molecular plant pathology **18**(1): 75-89.

Zhang, W., J. A. Corwin, D. Copeland, J. Feusier, R. Eshbaugh, F. Chen, S. Atwell and D. J. Kliebenstein (2017). "Plastic transcriptomes stabilize immunity to pathogen diversity: the jasmonic acid and salicylic acid networks within the Arabidopsis/Botrytis pathosystem." The Plant Cell: tpc. 00348.02017.

Zhang, W., J. A. Corwin, D. Copeland, J. Feusier, R. Eshbaugh, D. E. Cook, S. Atwell and D. J. Kliebenstein (2018). "Network connections across kingdoms illuminate a potential metabolic battlefield." bioRxiv.

Zhou, X. and M. Stephens (2012). "Genome-wide efficient mixed-model analysis for association studies." Nature genetics **44**(7): 821.