Bioinformatics and Computational Plant Pathology

Identification of Novel Basil Downy Mildew Resistance Genes Using De Novo Comparative Transcriptomics

Kelly S. Allen,[1] Gregory A. Delulio,[1] Robert Pyne,[2] Jacob Maman,[1] Li Guo,[3] Rebecca Lyon,[4] Eric T. Johnson,[4] Robert L. Wick,[5] James E. Simon,[2] Anne Gershenson,[1] and Li-Jun Ma[1,†]

[1] Department of Biochemistry and Molecular Biology, University of Massachusetts Amherst, Amherst, MA, U.S.A.
[2] Department of Plant Biology, Rutgers University, New Brunswick, NJ, U.S.A.
[3] Institute of Advanced Agricultural Sciences, Peking University, Weifang, Shandong, China
[4] Crop Bioprotection Research Unit, National Center for Agricultural Utilization Research, U.S. Department of Agriculture-Agricultural Research Service, Peoria, IL, U.S.A.
[5] Stockbridge School of Agriculture, University of Massachusetts Amherst, Amherst, MA, U.S.A.

Abstract

Sweet basil (*Ocimum basilicum*) production is threatened by the oomycete pathogen *Peronospora belbahrii*, causing basil downy mildew (BDM). BDM-resistant cultivar 'Mrihani' (MRI) was identified in a germplasm screen and bred with BDM-susceptible 'Newton' (SB22) to produce resistant cultivars, but the molecular mechanisms conferring resistance in MRI and the progeny remained unknown. A comparative transcriptomic approach was used to identify candidate resistance genes and potential mechanisms for BDM resistance. To differentiate the host–pathogen interactions in resistant and susceptible plants, RNA samples from BDM-infected MRI and SB22 plants were harvested at four time points during the first 3 days of infection, with mock-inoculated controls for both genotypes. Three categories of genes uniquely transcribed in the resistant MRI upon pathogen challenge were identified: nucleotide-binding leucine-rich repeat proteins (NLRs), multifunctional receptor-like kinases (RLKs), and secondary metabolic enzymes. Validation of the top resistance candidate NLR gene confirmed its unique presence in MRI and two of four resistant MRI × SB22 F_2 progeny. In MRI, pathogen challenge also induced differential regulation in members of the salicylic acid synthesis pathway, suggesting its role in BDM resistance. Overall, our study demonstrates the utility of de novo comparative transcriptomics to identify resistance genes and mechanisms in non-model crops.

Keywords: bioinformatics, computational biology, disease resistance, molecular, oomycetes, plant immune responses

Basil (genus *Ocimum*) is an agronomically and culturally significant herb crop with diverse species and cultivars possessing distinct phenotypes in plant size, leaf shape, aroma, and flavor

(Vieira and Simon 2006). Sweet basil (*Ocimum basilicum* L.) is the most popular basil and is cultivated for culinary use and essential oil production for applications including medicine, health care products, and food additives. The estimated market value of sweet basil was reported to be $57 million in 2020 and is projected to reach $63 million by 2027 (Brindisi et al. 2024).

Basil downy mildew (BDM), caused by the biotrophic oomycete *Peronospora belbahrii*, has become the most critical disease posing a severe threat to global basil production since its introduction and movement in Europe in 2001 and the United States in 2007 (Belbahri et al. 2005; Wyenandt et al. 2015). The pathogen enters plant tissue through open stomata or direct penetration of the upper cuticle and colonizes leaf tissue intercellularly, causing characteristic symptoms of interveinal chlorosis several days after infection (Cohen et al. 2017; Wyenandt et al. 2015). Following sustained periods of high relative humidity (>85%), the pathogen produces sporangiophores bearing infective sporangia, which emerge through stomata and create a brown to dark-gray discoloration corresponding to interveinal chlorosis (Cohen et al. 2017; Wyenandt et al. 2015). The sporangia are aerially dispersed and cause polycyclic infections throughout large production areas (Cohen et al. 2017; Wyenandt et al. 2015). After prolonged infection, the leaves desiccate and are abscised from the plant (Cohen et al. 2017; Wyenandt et al. 2015). Symptoms and signs of BDM disease render plants unfit for commercial sale during the early stages of infection and rapidly cause increasing yield losses as the disease progresses (Wyenandt et al. 2015).

First reported in Uganda in 1930, BDM reemerged and began to attract attention in 2001 when disease instances were increasingly reported from the Americas, Asia, and Europe (Choi et al. 2016; Garibaldi et al. 2004, 2005; Kanetis et al. 2014; Khateri et al. 2007;

Martínez de la Parte et al. 2010; McLeod et al. 2006; Nagy and Horváth 2011; Roberts et al. 2009; Ronco et al. 2009; Šafránková and Holková 2014; Smith and Urrea 2018; Yadav et al. 2022). BDM was first reported in the United States in 2007 in Florida and then in the Northeast the following year (Roberts et al. 2009; Wyenandt et al. 2015). As of 2021, BDM has been reported in 44 U.S. states, including Hawaii and the District of Columbia (McGrath 2021, 2023a; Wyenandt et al. 2015), threatening productivity in every growing region.

Chemical and cultural control methods to prevent infection and reduce disease spread include conventional and organic fungicides (Homa et al. 2014; McGrath 2020, 2023b), nocturnal fanning (Cohen and Ben-Naim 2016), nocturnal illumination (Cohen et al. 2013), and daytime solar heating (Cohen and Rubin 2015). However, these measures can be prohibitively time-consuming and costly, necessitating the development of improved basil cultivars with BDM resistance. Initial screening of basil germplasm identified the *O. basilicum* 'Mrihani' (MRI) (Horizon Seed Co., Williams, OR) with significant resistance to BDM (Pyne et al. 2014, 2015). This cultivar has a unique anise/fennel aroma and flavor profile and a distinctive phenotype compared with other *O. basilicum*

cultivars selected for culinary use. For example, the MRI leaves are smaller than other basil cultivars and have serrated leaves, compared with the large downward cupped smooth leaves of commercial sweet basil (Fig. 1A). Most importantly, the taste and smell of MRI differ considerably from eugenol-enriched sweet basil due to its unique methyl chavicol chemotype (Pyne et al. 2015). This cultivar was used in a parental cross with BDM-susceptible and *Fusarium*-resistant 'Newton' (also referred to as Rutgers breeding line SB22) and successfully produced fertile offspring. From the six-generation breeding design, four BDM-resistant cultivars ('Devotion', 'Obsession', 'Passion', and 'Thunderstruck') were selected from the backcrossed population progeny with improved downy mildew resistance and desirable phenotypes and chemotypes (Simon et al. 2018).

Screening of the full-sibling family offspring of the MRI × SB22 cross revealed additive and dominant gene effects of the MRI-conferred resistance, leading to the hypothesis that dominant alleles are involved in resistance (Pyne et al. 2015). A quantitative trait locus (QTL) analysis of the F_2 mapping population from the cross between MRI and SB22 identified a single locus (*dm11.1*) that accounts for 20 to 28% of the variance observed among the F_2
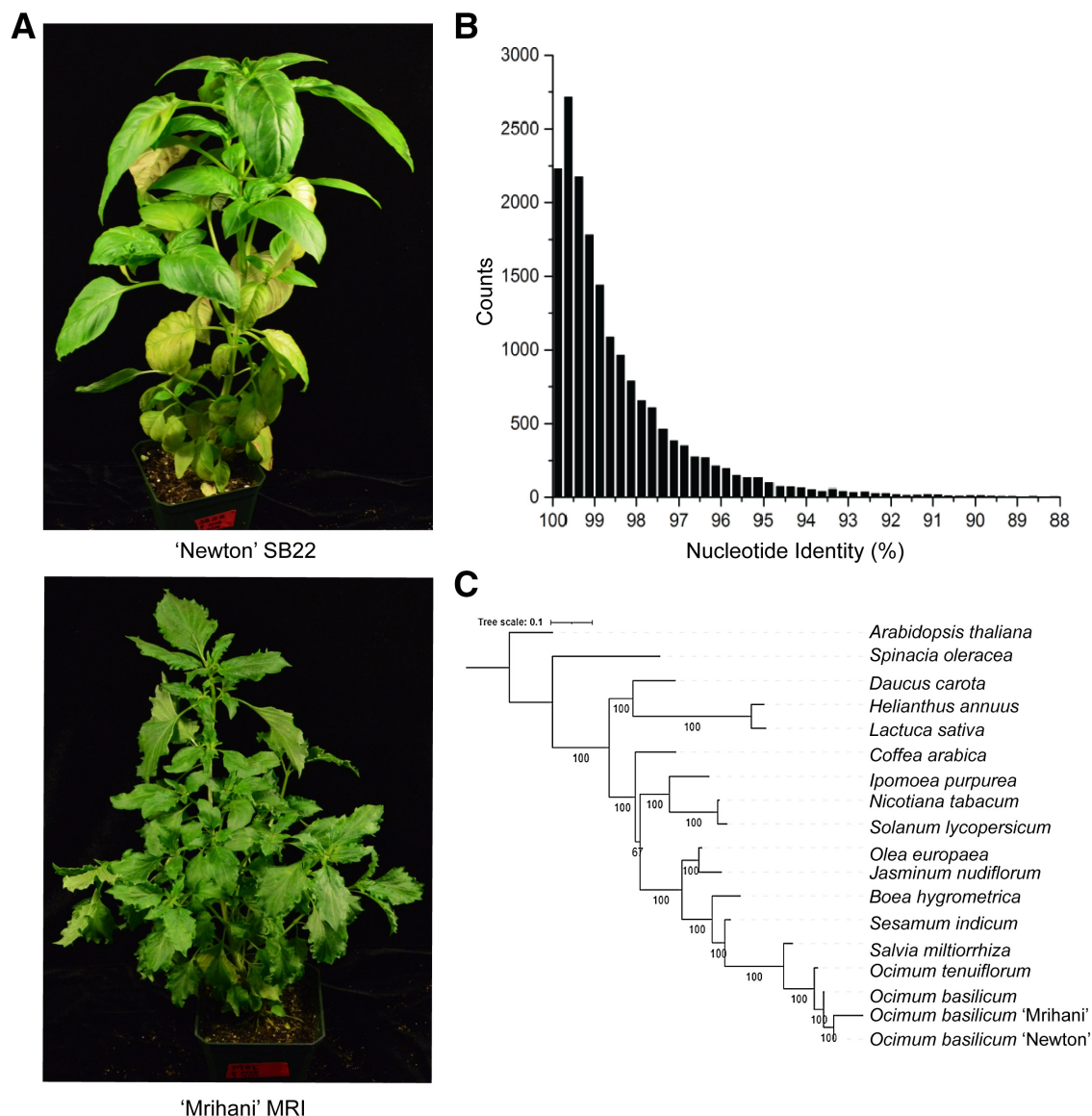
Fig. 1. Basil cultivar phenotypic and genetic diversity. **A,** Basil downy mildew-susceptible 'Newton' (SB22) and -resistant 'Mrihani' (MRI) cultivars inoculated with *Peronospora belbahrii*. **B,** Bidirectional blast hit among 20,943 genes shared between the SB22 and MRI water references. The nucleotide identity for each top BLAST hit is graphed here with a bin size of 0.25%. **C,** Multigene phylogeny of chloroplast genome orthologs across members of the asterid clade.

population (Pyne et al. 2017). In addition, this study also identified two minor loci (*dm9.1* and *dm14.1*) that respectively contributed 5 to 16% and 4 to 18% of the F$_2$ population's phenotypic variation. Resistance (*R*) genes involved in quantitative disease resistance map to QTLs (Nelson et al. 2018), and the linkage mapping results performed in the MRI × SB22 F$_2$ mapping population suggest that these loci may contain genes conferring quantitative disease resistance in MRI. Without genome assemblies of the basil cultivars, the underlying causative genes remain unknown.

Two genetic loci linked to BDM resistance were identified in a different resistant lineage of basil (Ben-Naim and Weitman 2022), and transcriptomics was also used to study susceptible BDM interactions during later stages of infection (Johnson et al. 2022). However, these studies do not address our unique needs of identifying sources of quantitative resistance. Therefore, we designed a comparative transcriptomics and identified MRI-unique transcripts, including nucleotide-binding leucine-rich repeat proteins (NLRs) and multifunctional receptor-like kinases (RLKs). We further confirmed the unique presence of the top NLR gene, demonstrating the utility of de novo comparative transcriptomics in identifying resistance genes in non-model crops.

Materials and Methods

Sample preparation

Inbred *O. basilicum* genotypes SB22 (*P. belbahrii* susceptible) and MRI (*P. belbahrii* resistant) plants were grown from seed as described previously (Pyne et al. 2014). Previously infected sweet basil leaves with fresh *P. belbahrii* sporulation were harvested and agitated for 5 min in sterile distilled water (diH$_2$0). The inoculum mixture was filtered with 40-μm nylon mesh. A 1-ml subsample from the filtered inoculum was pipetted into an Eppendorf tube and frozen at −80°C to serve as a pathogen control. The remaining inoculum was centrifuged at $1,000 \times g$ for 1 min and diH$_2$0 decanted. The resulting sporangia pellet was resuspended in diH$_2$O, and the inoculum concentration was adjusted to 1×10^5 sporangia/ml. Four- to six-week-old MRI and SB22 plants were spray-inoculated at the 6-leaf (3 true leaf set) growth stage with approximately 1 ml/leaf, and plants were incubated at 100% relative humidity for 24 h. A set of MRI and SB22 plants were sprayed with diH$_2$O in triplicate to serve as the mock-inoculated control.

Four disks per true leaf were sampled from both genotypes at 12, 24, 48, and 72 h postinoculation (hpi) and immediately flash frozen in liquid nitrogen (Supplementary Fig. S1). The water control leaves were harvested at 12 hpi only. Total RNA was extracted from freshly ground tissue using the Spectrum Plant Total RNA Kit (Sigma-Aldrich). RNA samples were used to generate sequence libraries using a library prep kit from New England Biolabs (NEB #E7530). Paired-end sequence reads of 75 bp were generated at the TUFTs genomic center at the Tufts University School of Medicine using the Hi-Seq Illumina platform. Higher coverage analyses were specifically designed using increased sequencing depth for inoculated samples that contain both the pathogen and the host due to the increased complexity of these samples, allowing for further study of pathogen expression.

Generating transcript assemblies and fragments per kilobase exon per million mapped reads (FPKM) expression

FASTQC version 0.11.5 (Andrews 2010) was used to assess average read quality. Paired-end reads (fastq files) were provided to Trinity version 2.4.0 and assembled using default parameters (Grabherr et al. 2011). Datasets were assembled including single sets using either all MRI datasets and the sporangia control (MRI combined assembly) or all SB22 datasets including the sporangia control (SB22 combined assembly). Separately assembled control data provided organism-specific databases of genes and transcripts. All sequence data are available at NCBI under GEO NUMBER: GSE111387.

The resulting output files served as the references for expression quantification. RSEM version 1.2.29 (Li and Dewey 2011) and bowtie version 1.0.0 (Langmead et al. 2009) were used to calculate FPKM values for assembled contigs while tracking replicate information. RNA sequencing datasets from both MRI and SB22 were mapped to the infected MRI Combined Assembly to standardize the reference, which allowed us to use previously generated gene annotations from our existing pipeline (Guo et al. 2016) and to cluster genes from both cultivars together. In all cases, the standard settings were used for de novo assembly in Trinity and transcript quantification (Grabherr et al. 2011). Additionally, edgeR (Robinson et al. 2010) was used to calculate differential gene expression. Expression data from both MRI and SB22 data mapped to the MRI Combined Assembly using Trinity and edgeR was used to assess differential expression between all time points within a single cultivar. Trinity differentially expressed gene output data were filtered for genes with a *P* value less than 0.05 and false discovery rate less than 0.01. An additional alignment to the reference genome was generated for each assembly. The raw reads were trimmed using Trimmomatic version 0.38 (Bolger et al. 2014). The trimmed reads were then aligned to the *P. belbahrii* reference genome (Thines et al. 2020) using STAR version 2.7.9a using 2-pass mapping (Dobin et al. 2013).

Reciprocal BLAST hits and reference gene phylogeny

To explore the overall sequence conservation between MRI and SB22, we performed a reciprocal BLAST using all sequences within the MRI and SB22 water control assemblies. Briefly, all MRI sequences were compared with the SB22 transcriptome, and all SB22 sequences were compared with the MRI transcriptome. The hit with the highest BLAST score for each gene was chosen. Results were compared, and pairs of top-scoring genes were considered reciprocal BLAST hits (i.e., MRI gene X BLASTs to SB22 gene Z, and SB22 gene Z BLASTs to MRI gene X).

Sequence conservation between MRI and SB22 was further assessed by performing a phylogenetic analysis using nine protein-coding chloroplast genome genes (*rpoC2*, *rpoB*, *psaA*, *ndhD*, *atpB*, *petB*, *rps16*, *psbA*, and *rps4*) based on a prior analysis (Rastogi et al. 2015). The MRI and SB22 chloroplast genes were identified using BLAST against the MRI and SB22 water control assemblies, and the top hits with the highest bit scores were chosen and translated into coding sequences. Sequence alignment of the MRI and SB22 coding sequences was performed against sequences from 14 asterid lineage plants downloaded from NCBI Organelle Genome Resources database, with *Spinacia oleracea* L. and *Arabidopsis thaliana* L. set as outgroups (Rastogi et al. 2015). The gene sequences were aligned using MAFFT (Madeira et al. 2019), and the tree was generated using IQ-TREE (Minh et al. 2020).

Sequence translation and annotation

We generated a database of sequence annotations for MRI genes. All MRI genes with an expression of FPKM > 1 in at least one time point were chosen, and the longest transcript associated with that gene was compared with the NCBI non-redundant database using cloud BLAST through Blast2GO. Annotations were saved as a searchable database in text format. Genes were filtered by taxonomic hit to verify their species of origin as needed.

To facilitate easier searches for gene families of interest, we translated the longest nucleotide sequence associated with each gene into six-frame translated protein sequences using EMBOSS (Madeira et al. 2019), searching for only those translated sequences between START and STOP codons longer than 30 amino acids. In many cases, to verify protein domain structure, the nucleotide or protein sequence was analyzed using either the NCBI conserved domain finder (Marchler-Bauer et al. 2015), PFAM (Finn et al. 2016), or InterProScan (Jones et al. 2014). Sequences were aligned using MEGA 6 for visual inspection (Tamura et al. 2013).

Expression clustering toward candidate resistance gene identification

Genes were clustered using Trinity version 2.2.0 based on read counts following the steps outlined in the Trinity manual (Grabherr et al. 2011). Expression data generated by mapping all datasets to the MRI Combined Assembly were used for clustering. A matrix of gene expression at all time points and replicates was used to define clusters with the edgeR function associated with Trinity (Robinson et al. 2010), using $P = 50$ and $P = 20$ (a grouping parameter for cluster creation, with higher numbers forming larger and broader clusters). The resulting clusters, available in pdf format, were visually examined for clusters that displayed the target expression profile.

BLAST search for secondary metabolite enzyme and resistance genes

To analyze MRI-unique gene families and defense hormone signaling genes, we performed a BLASTp search against MRI and SB22 translated nucleotide sequences using *A. thaliana*, or in some cases sweet basil, protein sequences retrieved from NCBI. Generally, the hit with the highest bit score was chosen as the top hit for each sequence. In cases of short alignment length or low sequence identity, the recovered MRI or SB22 hit was compared with the green plant database on NCBI. BLAST version 2.2.22 was used in all cases to compare protein sequences (Altschul et al. 1990) at the Massachusetts Green High Performance Computing Center.

PCR screen of parent and cultivar genomic DNA for unique genes

Genomic DNA was prepared from approximately 80 mg of newly emerging leaf tissue of MRI, SB22, Devotion, Obsession, Passion, and Thunderstruck cultivars using the E.Z.N.A. SP Plant DNA Kit (Omega BioTek, Norcross, GA) (Pyne et al. 2017). Primers amplifying transcript sequences were designed for MRI and SB22 shared genes as well as MRI-unique genes, including the comp160460c0 transcript, and were ordered from IDT (Coralville, IA). The primers were either external primers designed to amplify the whole gene (MRI_134-F and MRI_2869-R) or internal primers for specific regions of interest (Supplementary Table S2). Primers amplifying a 198-bp fragment of the *O. basilicum* Actin gene (ObActin_2-F, ObActin_2-R) were used in positive control reactions for all cultivars. PCR was performed with Q5 High-Fidelity DNA Polymerase (New England Biolabs, Ipswich, MA) using manufacturer-recommended cycling conditions with a 30-s denaturation cycle to ensure full denaturation of genomic DNA, a 30-s extension time to amplify the 198-bp ObActin region, and a 120-s extension time to amplify the approximately 3.5-kb comp160460-encoded gene on a Mastercycler Pro S (Eppendorf, Hamburg, Germany). Water was used as a negative control template in all reaction sets. Amplicons were visualized on 1.5% agarose gels stained with SYBR Safe DNA Gel Stain (Invitrogen, Waltham, MA) and imaged under UV light.

NLR allele analysis

Successfully amplified comp160460c0 products from MRI, Devotion, and Obsession were purified using a QIAquick PCR purification kit (QIAGEN, Hilden, Germany), and single amplicon copies were ligated into the pMiniT 2.0 vector using the NEB PCR Cloning Kit (New England Biolabs). Individual clones were selected and confirmed via colony PCR using Q5 High-Fidelity DNA Polymerase (New England Biolabs) with primers 1BF and 8R, which were designed to amplify the coiled-coil (CC) and NB-ARC domain coding sequences of MRI-R1. Plasmid DNA was prepared from confirmed clones using the Zyppy Plasmid Miniprep Kit (Zymo Research, Irvine, CA).

Individual comp160460c0 clones from MRI, Devotion, and Obsession were sequenced using NEB PCR Cloning Kit Cloning Analysis Forward and Reverse primers to flank the insert, as well as internal primers (Supplementary Table S2). Twelve clones were sequenced from MRI, and six clones were sequenced from both Devotion and Obsession. Consensus sequences for each clone were assembled and annotated to identify intron regions by alignment to transcript sequences; coding regions were confirmed and annotated using InterProScan (Jones et al. 2014). Assembled sequences were identified as individual alleles, nucleotide and predicted protein sequences were aligned using the EMBL-EBI Clustal Omega multiple sequence alignment tool (Madeira et al. 2022), and alignments were visualized using Jalview 2.11.1.4 (Waterhouse et al. 2009).

Protein structures were predicted using RoseTTAFold (Baek et al. 2021), and the predicted secondary structure was added to the Clustal Omega allele sequence alignment using ESPript 3.0 (Robert and Gouet 2014). Allele structures were used as a query search against the Protein Data Bank in DALI, and structure pairwise comparison was performed (Holm 2020). Allele structures were further analyzed and aligned using UCSF Chimera (Pettersen et al. 2004). Allele expression was analyzed by mapping the variable coding regions to the RNA sequencing data using the Burrows-Wheeler Aligner software package (BWA-MEM) (Li and Durbin 2009), and results were visually examined using the Integrated Genomics Viewer (Robinson et al. 2011).

Results

Sequencing data reflect the phylogenetic relatedness of MRI and SB22 and increasing pathogen load in infected samples

Inbred *O. basilicum* genotypes SB22 (*P. belbahrii* susceptible) and MRI (*P. belbahrii* resistant) leaves were spray-inoculated with fresh *P. belbahrii* sporangia. RNA samples from MRI and SB22 inoculated with the pathogen *P. belbahrii* were extracted at 12, 24, 48, and 72 hpi, representing roughly the germination, penetration, and intercellular growth stages (Supplementary Fig. S1). We generated 12.8 million (MRI), 14.3 million (SB22), and 9.9 million (sporangia) high-quality paired-end reads per replicate for three controls. Considering the increased complexity of infected samples, we doubled the sequence coverage and generated an average 24.8 million and 27.6 million high-quality Illumina paired-end reads per replicate per infection sample for MRI and SB22, respectively (Table 1; Supplementary Fig. S2). All sequence data were deposited at NCBI under GEO NUMBER: GSE111387.

In total, 240.2 million paired-end reads were used to generate the MRI Combined Assembly, containing 341,633 unique transcripts corresponding to 133,441 genes called by Trinity (Grabherr et al. 2011). The SB22 Combined Assembly was generated using 263.9 million paired-end reads and contained 118,296 genes and a total of 322,696 unique transcripts. The MRI and SB22 plant-only control assemblies contained 66,930 and 52,844 genes, respectively, and the sporangia control contained 10,144 assembled genes. As expected, more genes were assembled in infected samples, representing both host and pathogen transcripts, genes expressed only during infection.

Though fewer reads were sequenced for both water and sporangia control samples, the sporangia control produced the highest sequence coverage ($>90\times$) and longest average gene length (1,603 bp) for assembled genes, likely due to the smaller genome/transcriptome size of the (inactive) pathogen. The average assembled gene length of *O. basilicum* transcripts from a previously published transcriptome was 1,363 bp (Rastogi et al. 2014), larger than our average assembly size. The average gene size of the oomycete *Phytophthora infestans* was 1,523 bp (Haas et al. 2009), roughly equivalent to the sporangia control assembly.

To assess the genetic diversity between MRI and SB22, we performed a BLAST search between MRI and SB22 water control assemblies and identified 20,943 reciprocal hits, likely representing orthologs between these two plants. These orthologs are highly similar, with an average pairwise sequence identity of 98.66% (Fig. 1B). We further assessed the genetic similarity between the cultivars by conducting a phylogenetic analysis of nine protein-coding chloroplast genome orthologs across members of the asterid

clade, to which *O. basilicum* belongs, and with *A. thaliana* and *S. oleracea* set as outgroups (Rastogi et al. 2015). The analysis showed high sequence conservation among the *Ocimum* spp., with MRI and SB22 grouped together (Fig. 1C). We anticipate that unique or differentially regulated genes will likely contribute to the MRI and SB22 phenotypic variations.

Using expression profiles in the individual assemblies, MRI transcripts can be divided into 36,414 predicted plant transcripts (present in the MRI water control), 9,988 predicted pathogen transcripts (sequences from MRI infected samples that are present in the sporangia control), and 29,502 infection unique transcripts (absent in both plant and pathogen controls) (Table 2). Similarly, the SB22 transcripts include 31,702 transcripts with a plant origin, 9,426 transcripts of pathogen origin (sequences from SB22-infected samples that are present in the sporangia control), and 26,486 transcripts uniquely present in the infection samples. Consistent with SB22 susceptibility, we saw a significant increase in the number of expressed pathogen genes in the inoculated SB22 samples, observing an almost 20-fold increase from 12 to 72 hpi, compared with only a twofold increase for MRI (Fig. 2; Supplementary Fig. S3).

A recent high-quality reference genome for *P. belbahrii* (Thines et al. 2020) allowed us to benchmark our de novo analyses by aligning the control and infection assembly reads. The sporangia control reads had a high percentage mapped to the *P. belbahrii* reference genome using STAR alignment (85.65%), consistent with de novo mapping to the control assembly using Trinity (84.1%). The *P. belbahrii* genome mapped to a low percentage of reads in the MRI assemblies (Supplementary Table S1), consistent with the low percentage of sporangia control reads mapped to the MRI assemblies in the de novo analysis (Fig. 2; Supplementary Fig. S2). The SB22 assemblies increased in mapping percentage to the *P. belbahrii* reference over the time course (Supplementary Table S1), consistent with the de novo analysis (Fig. 2). As the *P. belbahrii* reference represents a European isolate, and mapping against the reference did not appear to significantly change the alignment results compared with using the sporangia control in the de novo analyses, we are reporting the Trinity alignment for gene expression and clustering analyses.

Gene Ontology (GO) analysis

The MRI combined assembly was annotated for top blast hit and GO terms using Blast2GO, using the longest sequence per component (i.e., the longest "alternative splice variant") as a representative of that component (Conesa et al. 2005). Of the 76,018 unique MRI sequences provided to Blast2GO (Table 2), 35,294 returned no tblastx hit. The remaining 40,725 sequences were separated based on the genus name of the top BLAST hit. As our experiments were conducted in a non-sterile greenhouse where only basil plants were grown, we removed matches to non-oomycete microbes and insects.

These removed sequences matched to the Basidiomycete yeast genera, *Tilletiaria* (308 sequences) and *Pseudozyma* (91 sequences), and common greenhouse pests, *Tetranychus urticae* (4,588 sequences), the red or two-spotted spider mite, and *Bemisia tabaci* (1,716 sequences), the silverleaf whitefly, among others. We kept 25,045 sequences that matched to the plant kingdom as potential *O. basilicum* sequences and 4,430 matches to the genus *Phytophthora* as *P. belbahrii* genes.

Clustering analysis highlights transcripts with potential functions involved in host–pathogen interactions

Based on overall ortholog identity, we were confident that SB22 reads from orthologous genes would align to the MRI assembly. Using the MRI Combined Assembly as a reference, we mapped all the SB22 and MRI reads using RSEM with default parameters (Li and Dewey 2011). Initial coarse clustering using all genes from MRI and SB22 mapped data resulted in 12 clusters, with an average of 2,810 genes per cluster (Supplementary Table S3). Three clusters that lacked consistency among biological replicates were not included in further analyses. Three clusters (A-1, A-2, and A-3) characterized transcripts primarily belonging to the pathogen (Fig. 3A), as all transcripts showed significant expression in sporangia pathogen control samples (Fig. 3, gray bar in the middle) but no expression in both water-inoculated plant control samples (Fig. 3, two blue bars representing MRI-water only and SB22-water only).

Six clusters containing primarily plant transcripts were identified. All transcripts showing significant expression in plant and infected plant samples absent from sporangia pathogen control

TABLE 2. Expressed genes across time points[a]

Dataset	Total genes	Number of genes FPKM > 1		
		Plant control	Sporangia control	Infection unique
MRI water	36,528	36,414	0	0
MRI 12 h	39,349	29,247	472	9,534
MRI 24 h	40,171	29,074	389	10,606
MRI 48 h	40,973	28,852	179	11,840
MRI 72 h	39,279	28,365	825	9,987
Sporangia	10,102	0	9,988	0
Total unique	76,018	36,414	9,988	29,502
SB22 water	31,794	31,702	0	0
SB22 12 h	35,561	25,770	274	9,431
SB22 24 h	36,732	25,584	699	10,364
SB22 48 h	40,216	25,746	1,788	12,598
SB22 72 h	40,838	25,414	5,375	9,962
Sporangia	9,518	0	9,426	0
Total unique	67,706	31,702	9,426	26,486

[a] The gene counts of expressed transcripts at each time point reflect the mapping to either 'Mrihani' (MRI)-combined or 'Newton' (SB22)-combined assemblies. FPKM, fragments per kilobase exon per million mapped reads.

TABLE 1. Summary of assembled transcripts by time point[a]

Dataset (organism)	Read pairs (millions)	Mapping percentage	Average base quality	Number of genes	Average gene length	Average coverage
Sporangia (*Peronospora belbahrii*)	29.7	84.1	37.28	10,144	1,603	230.4
MRI water (*Ocimum basilicum* MRI)	42.7	62.0	37.20	66,930	893	66.4
MRI 12 hpi (*O. basilicum* MRI, *P. belbahrii*)	75.9	63.3	37.21	81,873	897	98.1
MRI 24 hpi (*O. basilicum* MRI, *P. belbahrii*)	75.3	63.3	37.22	89,302	864	92.6
MRI 48 hpi (*O. basilicum* MRI, *P. belbahrii*)	77.0	63.1	37.23	85,832	860	98.7
MRI 72 hpi (*O. basilicum* MRI, *P. belbahrii*)	69.8	63.4	37.22	78,061	909	93.5
SB22 water (*O. basilicum* SB22)	38.3	62.9	37.18	52,844	977	69.9
SB22 12 hpi (*O. basilicum* SB22, *P. belbahrii*)	85.9	63.9	37.18	74,918	934	117.6
SB22 24 hpi (*O. basilicum* SB22, *P. belbahrii*)	83.2	63.3	37.19	72,699	949	114.5
SB22 48 hpi (*O. basilicum* SB22, *P. belbahrii*)	75.9	65.2	37.15	84,478	870	100.9
SB22 72 hpi (*O. basilicum* SB22, *P. belbahrii*)	86.6	66.6	37.20	80,675	903	117.7
MRI-combined assembly	370.4	64.8	37.22	133,441	765	352.6
SB22-combined assembly	399.6	66.0	37.19	118,296	692	477.7

[a] MRI, 'Mrihani'; SB22, 'Newton'; and hpi, hours postinoculation.

samples were filtered (Fig. 3B and C). The remaining genes from the six plant gene clusters represent transcripts uniquely expressed during pathogen challenge. Transcripts within three clusters, B-1, B-2, and B-3, had comparable expression profiles between MRI and SB22, indicating conserved functions between the two different plant hosts. Both B-1 and B-2 clusters show a pattern consistent with a 12-h shift, but these clusters respond in opposite directions. Cluster B-1 was upregulated at 24, 48, and 72 hpi and was enriched for metabolism, oxidation reduction, and photosynthesis functions. No GO terms were significantly enriched in cluster B-2, which showed downregulation at 24, 48, and 72 hpi; however, of those GO terms annotated by Blast2GO, metal ion functions were predominant. The largest cluster, B-3, roughly represents stably expressed genes. Cluster B-3 is enriched for many categories, including various metabolic processes, protein modification, and protein localization.

Transcripts in three clusters, C-1, C-2, and C-3, displayed differential regulation responses between MRI and SB22 during pathogen challenge (Fig. 3B). In Cluster C-1, MRI expression was high but almost completely absent from the SB22 transcriptome. In Cluster C-2, MRI expression was higher than SB22 and was enriched for genes related to defense response, response to stress, response to stimulus, and DNA integration. In Cluster C-3, MRI gene expression was lower than SB22. No GO terms were enriched in Cluster C-3.

MRI-unique expressed genes include NLRs, RLKs, and secondary metabolic enzymes

To understand potential mechanisms underlying BDM resistance, we repeated the clustering with higher stringency ($P = 20$, see methods for details), resulting in 188 clusters. Eight clusters were chosen, as they were expressed in MRI, minimally expressed in SB22, and showed no expression in the sporangia control. A comprehensive filtering process of the eight clusters to remove genes with mapped reads to SB22 or bidirectional BLAST hits between MRI and SB22 resulted in 376 MRI-unique candidate genes. Of these 376 MRI-unique candidates, 95 were identified as MRI-unique

candidates after removing hypothetical protein and transposable elements. These MRI-unique candidate genes contained secondary metabolic enzymes (22 genes), immunity-related genes, including 18 NLR genes and 12 RLKs or receptor-like proteins, and others (Supplementary Table S5).

Detecting secondary metabolic enzymes as MRI-unique genes is expected, as these two basil genotypes produce distinct secondary compounds. A previous investigation revealed that SB22 accumulates a significant amount of eugenol, whereas MRI predominantly accumulates methyl chavicol (R. Pyne, *unpublished data*). Two enzymes involved in the methyl chavicol biosynthesis pathway, cinnamate p-coumarate carboxyl methyltransferase (comp153726; involved in methyl cinnamate biosynthesis) and chavicol/eugenol O-methyltransferase (comp171322; catalyzing the conversion of chavicol to methyl chavicol), were uniquely detected in MRI (Supplementary Table S5).

Unique expression of RLKs and NLRs, both immunity-related protein families, in MRI is of particular interest. Basic plant immunity consists of pattern-triggered immunity and effector-triggered immunity (Cui et al. 2015; Jones and Dangl 2006). Plant pattern-triggered immunity uses receptor-like proteins/RLKs, a large gene family (Chern et al. 2016; Mendy et al. 2017) that has roles as sensors of microbe-associated molecular patterns and induce downstream defense reactions (Bigeard et al. 2015; Dodds and Rathjen 2010; Jones and Dangl 2006; Zhou and Zhang 2020). Plant effector-triggered immunity employs intracellular NLRs that play roles in sensing effector proteins secreted by pathogens and regulating downstream defense signaling (Cesari 2018; Cui et al. 2015; Jones and Dangl 2006; Monteiro and Nishimura 2018; Wang and Chai 2020).

To investigate basil pattern-triggered immunity, we further characterized the RLK BLAST hits. The best RLK candidate, selected based on the unique presence and upregulation in resistant MRI, is transcript mri_comp170662 encoding a full-length malectin-like RLK (Fig. 4A). Several predicted RLKs were truncated fragments, including three transcripts containing only the leucine-rich repeat (LRR) domain, two containing only protein kinase domains, and
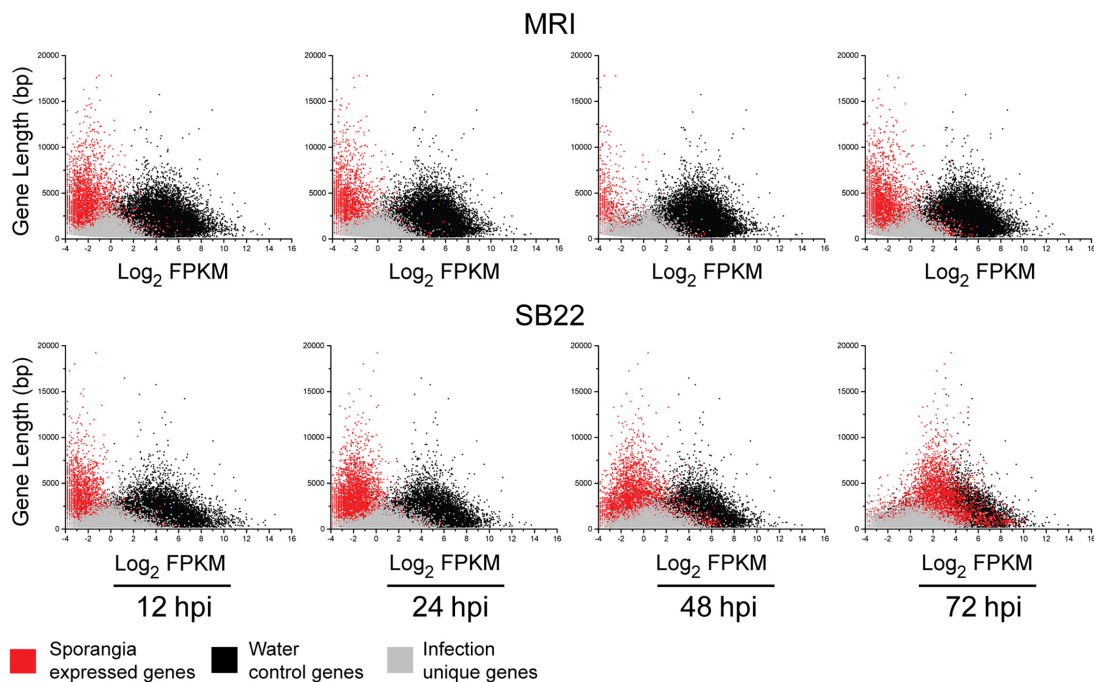

Fig. 2. Mapping of transcripts at each time point to the sporangia-only and 'Mrihani' (MRI) mock-infected water control-only assemblies (top) and the sporangia-only and the 'Newton' (SB22) mock-infected water control assemblies (bottom). Dataset and replicate are listed on the bottom *x* axis. Transcripts only mapped to the sporangia assembly are highlighted in red, and transcripts mapped to the mock-infected water controls are in black for MRI (top) and SB22 (bottom). All other transcripts are defined as infection unique and highlighted in gray. FPKM, fragments per kilobase exon per million mapped reads; and hpi, hours postinoculation.

two that were too short to have domain affiliation. Two transcripts, comp109241 and comp171939, both contained an RLK-like domain and shared a homologous sequence with OsRMC, a negative regulator of salt stress response, and the antifungal protein Ginkbilobin-2, respectively (Sawano et al. 2007; Zhang et al. 2009). Using the full-length top candidate RLK transcript mri_comp170662, we identified 12 and 10 homologs in the MRI and SB22 assemblies, respectively. Seven SB22/MRI orthologous pairs can be readily identified between these two sister cultivars based on a phylogeny (three sequences lacking a malectin domain were excluded from the tree) (Fig. 4A). Among four MRI orphan RLK transcripts, only this top RLK candidate, mri_comp170662 (boxed in Fig. 4A), showed increased expression at 12 and 24 hpi compared with the water control (Fig. 4B) and contained LRR, malectin, and kinase domains (Fig. 4C).

Differential regulation of NLRs in MRI during infection

We identified two putative NLR resistance genes, comp_160460_c0 and comp_178221_c0, encoding a late blight resistance protein homolog R1A (gi:848916018 and gi:848932751) from spotted monkey flower (*Erythranthe guttata*), belonging to the order Lamiales along with basil. We focused on our top candidate comp160460_c0, named MRI Resistance gene 1 (MRI-R1), as it encodes a full-length CC-NLR protein with three canonical functional domains and was induced upon pathogen challenge. Specifically, MRI-R1 was twofold upregulated between 12 and 24 hpi, and upregulated expression was maintained at 48 and 72 hpi (Supplementary Table S5).

To determine whether MRI-R1 activity could be attributed to a presence/absence polymorphism between MRI and SB22 or the differential regulation at the transcriptional level, we designed primers to amplify the predicted coding region of the full-length MRI-R1 transcript using PCR. Forward primer 1BF flanks the 5′ end of the coding sequence, and reverse primer 14R targets the 3′ end of the gene. This primer pair produced an amplicon of the MRI-R1 gene from MR1, but not from SB22 or the water template negative control (Fig. 5A). Internal primers were designed to map and sequence MRI-R1-like genes in MRI and SB22 further (Supplementary Table S2). PCR using internal forward primer 9F, which flanks the NB-ARC domain, paired with 3′ end primer 14R, resulted in gene amplification from MRI and potentially off-target or homologous amplification from SB22 (Fig. 5A).
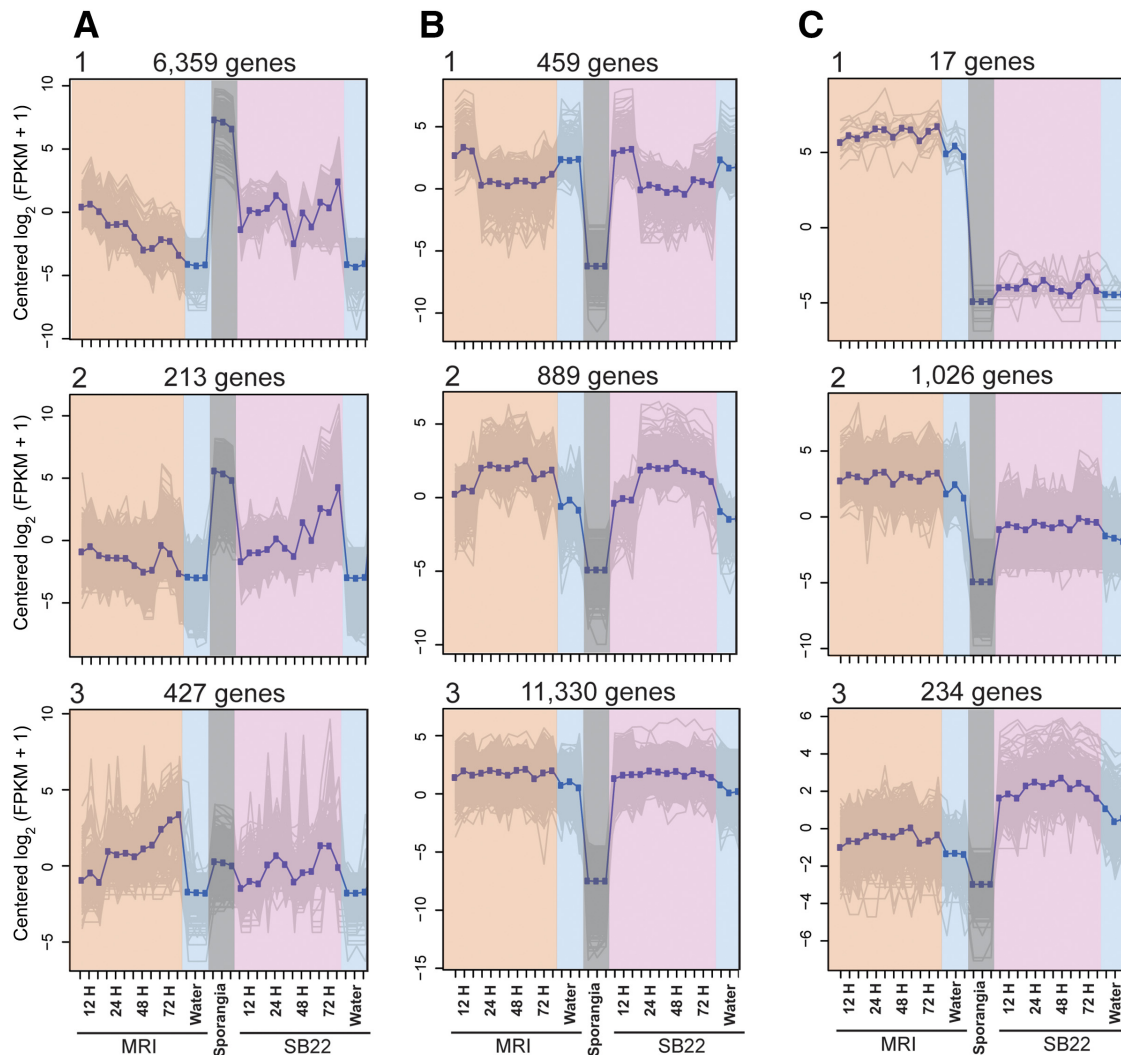

Fig. 3. Clusters with expressed plant genes. Nine gene clusters produced by coarse clustering are displayed here. Expression value, *y* axis, is on a $\log_2$ scale. Three replicates for each dataset are represented by tick marks on the bottom *x* axis, with mean data plotted as dark blue points and gray shading representing the standard deviation. Datasets from 'Mrihani' (MRI) are highlighted in orange with three replicates from 12 to 72 h postinoculation (hpi), SB22 datasets are colored pink with three replicates from 12 to 72 hpi, both water controls are labeled blue (MRI left, SB22 right), and the pathogen sporangia control data are colored gray. **A,** Pathogen gene clusters, as all transcripts showed significant expression in sporangia pathogen control samples but were absent from plant and infected plant samples. **B and C,** Gene clusters contain primarily plant transcripts, as well as differentially expressed pathogen transcripts, with significant expression in plant and infected plant samples, and no expression in sporangia pathogen control samples. B, Differentially expressed plant and pathogen transcripts with similar expression profiles in both cultivars. C, Plant and pathogen transcripts with different expression profiles between cultivars. FPKM, fragments per kilobase exon per million mapped reads.

The MRI-R1 amplicon detected in the MRI genomic DNA appeared to be approximately 800 bp larger than predicted from the RNA transcripts, indicating the potential presence of a non-coding intronic region (Fig. 5A). We also observed sequence polymorphisms among MRI-R1 transcripts with six different isoforms predicted for the same gene (two with full coding sequences, comp160460_c0_seq2 and comp160460_c0_seq5). To investigate transcript polymorphisms, we isolated and sequenced individual clones from the PCR products. A total of 12 full-length amplicons were cloned from three MRI plants. Due to the length of the transcript and to help validate the sequences, 12 to 25 overlapping sequences were generated from each clone. Assembly and annotation of these MRI-R1 sequences revealed four separate alleles of the gene, supporting the allelic polymorphisms observed in the transcriptomic data (Fig. 5B). The pairwise nucleotide sequence identity among the four MRI-R1 alleles range from 80 to 96.45%. Alleles 1 and 2 share an identical 827-nucleotide intron sequence, and alleles 3 and 4 share an identical 376-nucleotide intron sequence, which accounts for the size discrepancy in the agarose gel electrophoresis result (Fig. 5A).

Analysis of the protein sequences using InterProScan (Jones et al. 2014) showed that all four alleles contain the CC, NB-ARC, and LRR domains with canonical functional motifs (Supplementary Fig. S4). The N-terminus of all four alleles begins with an identical MADA motif, a functional motif conserved in approximately 20% of CC-NLR immune receptors across distantly related plant species (Adachi et al. 2019; Bentham et al. 2018). The MADA motif is necessary for *Nicotiana benthamiana* NRC4 cell death like the ZAR1 resistosome and suggests the likely role of MRI-R1 as a helper NLR (Adachi et al. 2019; Contreras et al. 2023). Similarly, the EDVID motif was found in all four allele coding sequences; this motif is involved in self-association, direct interactions with cofactors, and, in some cases, cell death signaling resulting in a hypersensitive response (Bentham et al. 2018). The NB-ARC domain of all four
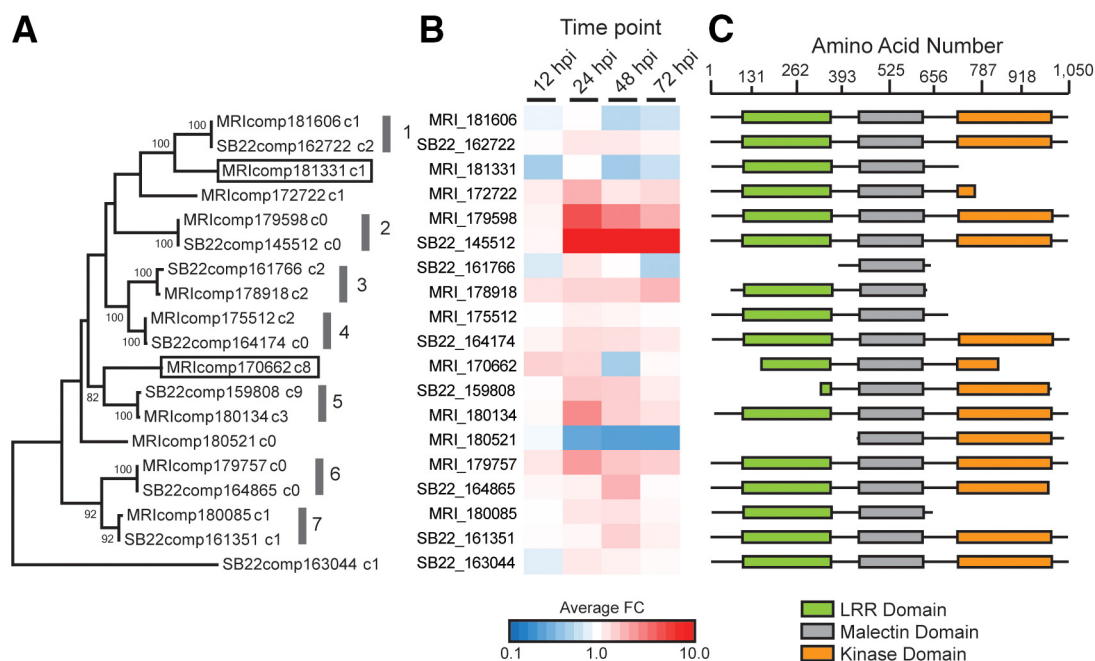

Fig. 4. Malectin-like receptor-like kinase (RLK) proteins in 'Mrihani' (MRI) and 'Newton' (SB22). **A,** Alignment of 19 of 22 RLK proteins by their conserved malectin domain. Gray bars with numbers represent orthologous pairs, and boxes indicate sequences absent from the SB22 cultivar. **B,** Fold change (FC) compared with water for MRI and SB22 RLKs across infected plant samples. **C,** Protein domain structure of 22 RLK hits generated from translated nucleotide sequences. Transcript IDs are those of the adjacent FC row. LRR, leucine-rich repeat.
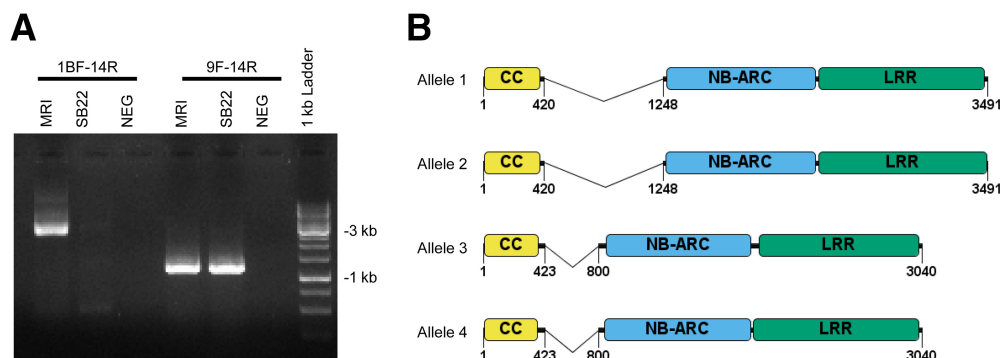

Fig. 5. Unique presence of 'Mrihani' (MRI) resistance gene 1 (MRI-R1) and models of four alleles in MRI. **A,** Unique amplification of MRI-R1 from MRI (3,673 bp) using the external 1BF forward and 14R reverse primers. The internal primer, 9F, paired with 14R produces amplicons of expected size 1,206 bp in both MRI and SB22. The negative control (NEG) is water in place of genomic DNA template. **B,** Gene models of the four MRI-R1 alleles with domains and subdomains predicted by InterProScan colored as follows: coiled-coil (CC, gold), NB-ARC (blue), and leucine-rich repeat (LRR, green). Introns between the CC and NB-ARC domains are represented as a single line.

alleles also contains the Walker A (P-loop) motif (GMFGLGKT) (Ramakrishnan et al. 2002), which is critical for nucleotide binding (Steele et al. 2019). Also present in the four alleles is the MHD-like motif IHD, which is involved in the inhibition of autoactivation of R proteins in the absence of a pathogen (van Ooijen et al. 2008). The MHD motif is proposed to act as a molecular switch for R protein activation, and the histidine and aspartate residues are the most highly conserved across R proteins, with the histidine occupying a critical position in an ADP-binding pocket (van Ooijen et al. 2008). The most variable regions among these four alleles are in and between the CC and NB-ARC domains (Supplementary Fig. S4).

Protein structural models of all four alleles were generated in the Robetta server utilizing RoseTTAFold, a top-ranked deep-learning-based protein structure prediction method (Baek et al. 2021; Du et al. 2021). The resulting allele protein structures were queried using a distance matrix alignment in the DALI server to search for the closest structural homologs (Holm 2020).

The resulting top hits for MRI-R1 included plant proteins such as the NB-ARC domain from the tomato immune receptor NRC1 (6S2P), LRR receptor-like serine/threonine-protein kinase FLS2 (4MN8), and LRR receptor-like serine/threonine-protein kinase GSO1 (6S6Q). The top structural homologs also included human and animal LRR proteins, such as LRR transmembrane neuronal protein 2 (5Z8X), dimeric bovine tissue-extracted decorin (1XCD), and human osteomodulin (5YQ5). The top 10 structural homologs ranged in Z scores from 18 to 22.2 (with Z scores above 20 indicating definite homology and above 8 indicating probable homology), and average deviation in distance between aligned Cα atoms in 3D superimpositions, indicated by root mean-squared deviation, ranged from 2.2 to 8.5 angstroms (Supplementary Table S4) (Holm 2020).

Structural modeling and homology comparison of each allele were used to refine the initial sequence-based boundary predictions of the CC, NB-ARC, and LRR domains (Supplementary Fig. S5A). All four allele structural models were aligned, revealing the conservation of predicted functional domain structures and active sites. As expected, based on high sequence identity, all four alleles shared high structural similarity, with slight differences in domain boundaries predicted by the models (Supplementary Fig. S5B). Allele expression analysis showed a clear bias for expression of MRI-R1 allele 1 by MRI in response to BDM infection.

The presence of MRI-R1 in backcross progenies Devotion and Obsession

To assess the contribution of MRI-R1 in the four new downy mildew-resistant (DMR) basil cultivars that integrated MRI resistance genes, we tested each for the presence of the MRI-R1 gene.

A partial sequence of ObActin, the ubiquitous positive control, was successfully amplified from BDM-susceptible SB22, BDM-resistant MRI, and all four DMR cultivars, whereas no amplification was detected from the negative water template control (Fig. 6). Interestingly, MRI-R1 was detected in only two, Devotion and Obsession, out of the four DMR cultivars using primers designed to amplify the full coding region of the gene (Fig. 6). The two amplicons are estimated to be 3,064 and 3,515 bp, corresponding to allele 4 and allele 1 for Devotion and Obsession, respectively. These amplicon sizes include 24 bp of the 5′ untranslated region (beginning with primer 1BF) and the 827-bp intron (allele 1) or the 376-bp intron (allele 4). The six individual clones selected and sequenced from Devotion and Obsession were uniform, indicating that only one allele of MRI-R1 was passed from MRI to these offspring.

These four new DMR cultivars were selected from genetic breeding efforts. Briefly, F_1 progeny were generated through the MRI (female) × SB22 (male) cross, exhibiting dominant gene action (Pyne et al. 2015). An F_2 family was generated after F_1 self-pollination, and a single resistant individual, RUMS469-11, was selected for hybridization with elite sweet basil inbred line 'SB13', an Italian-style sweet basil with observed disease tolerance to downy mildew and the Fusarium wilt pathogen *Fusarium oxysporum* f. sp. *basilici*. Twenty individuals achieving the highest category of reduced disease severity from the RUMS469-11 (female) × SB13 (male) cross were self-pollinated to generate full sibling families evaluated for response to downy mildew. The cultivars Devotion, Obsession, Passion, and Thunderstruck were selected from these inbred lines. These four selected DMR cultivars have inherited genes involved in quantitative resistance from both MRI and SB13.

Differential upregulation of salicylic acid (SA) biosynthesis pathways in MRI

Both NLR and RLK genes interact with hormone signaling pathways to activate host defense (McHale et al. 2006). To understand the involvement of plant hormone signaling pathways involved in susceptibility and resistance responses, we examined the expression of genes involved in ethylene, jasmonic acid, abscisic acid, indole-3-acetic acid (auxin), gibberellic acid, and SA based on *A. thaliana* annotation (1e-20, sequence similarity >40%). No significant differences were observed between the MRI and SB22 expression pattern in ethylene, jasmonic acid, abscisic acid, or gibberellic acid pathway genes (Supplementary Fig. S6). We saw a difference in expression profile for two of five auxin genes; YUC1 and TAA1 were upregulated at early time points in MRI, but they were not statistically significantly differentially expressed ($P > 0.05$).

The most striking difference was observed among genes required for the synthesis of SA (Fig. 7). Plants use two biosynthesis

Fig. 6. 'Mrihani' (MRI) resistance gene 1 (MRI-R1) detection in downy mildew-resistant (DMR) offspring. Gel electrophoresis showing amplification of ObActin fragment positive control from all basil cultivars tested on the left, MRI-R1 amplification from MRI, 'Devotion', and 'Obsession' on the right using the external 1BF forward and 14R reverse primers. The negative control (NEG) is water in place of genomic DNA template.
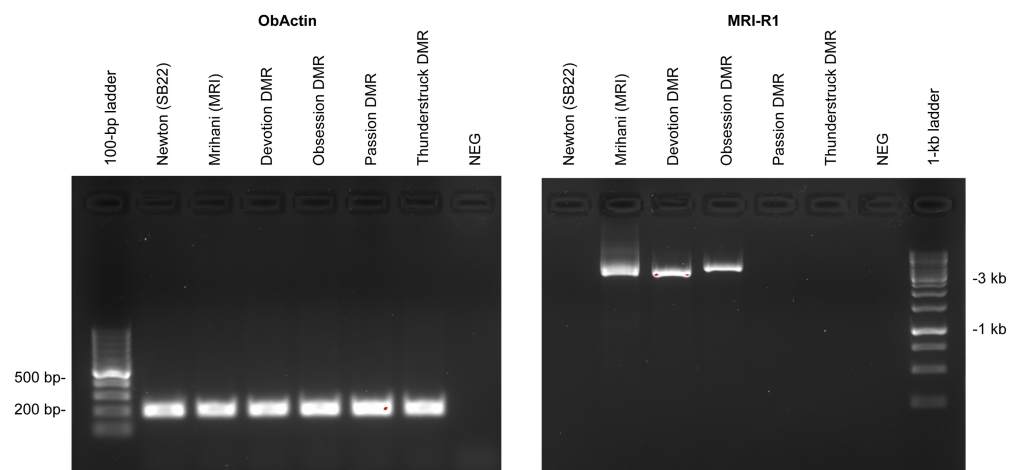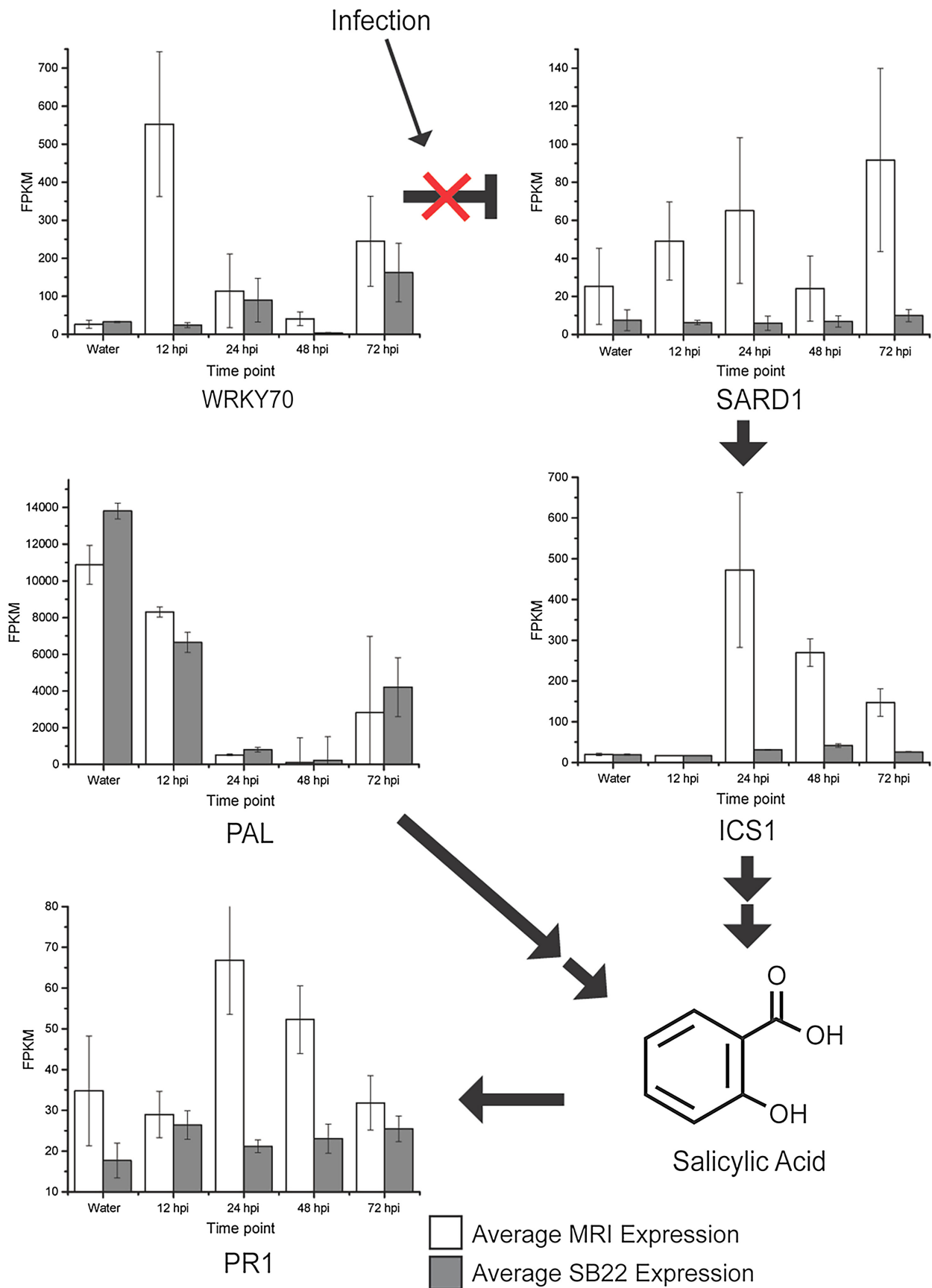
Fig. 7. 'Mrihani' (MRI) and 'Newton' (SB22) diverge transcriptionally at salicylic acid (SA) synthesis. Bar graphs are the average and standard deviation of three replicates. White bars indicate MRI data, and gray bars indicate SB22 for each time point. Infection releases the repression WRKY70 has on SARD1 expression. Double arrows leading from ICS1 and PAL indicate more than one step to the SA molecule. FPKM, fragments per kilobase exon per million mapped reads; hpi, hours postinoculation.

pathways to synthesize SA, both starting from chorismate, but subsequent steps involve either isochorismate synthase (ICS) (Lefevere et al. 2020; Wildermuth et al. 2001) or phenylalanine ammonia-lyase (PAL) (Olsen et al. 2008). There was no significant difference in PAL expression between the two cultivars, but we did observe a drastic induction of ICS1 at 24 hpi in MRI. For the water control and 12 hpi samples, there was no significant difference in ICS1 expression between MRI and SB22. Compared with SB22, ICS1 expression in the MRI cultivar is 15-, 6-, and 5-fold higher at 24, 48, and 72 hpi, respectively.

SARD1 is upstream of ICS1 in the SA synthesis pathway, and the average SARD1 expression across all time points is sevenfold higher in MRI compared with SB22 and has roughly 2-, 2.5-, 0-, and 3.6-fold increases at 12, 24, 48, and 72 hpi compared with the water control. Similarly, PR1, a commonly used marker gene downstream of SA synthesis, is significantly induced in MRI upon pathogen challenge starting at 24 hpi, confirming the unique increase of SA in MRI upon pathogen challenge.

Our data suggest that MRI-derived BDM resistance involves the induction of SA biosynthesis and signaling through the ICS pathway. The most significant change occurred at 24 hpi, 1 day after the plant was challenged with the pathogen. WRKY70 is known to repress SARD1 expression in the absence of pathogens and is required to activate some defense genes (Li et al. 2004; Zhou et al. 2018). In MRI, we saw a strong induction in the transcription of WRKY70 at 12 hpi, increasing expression of SARD1 from water to 24 hpi, and a 23-fold rise in the relative normalized expression of ICS1 at 24 hpi relative to 12 hpi. In SB22, we observed only a slight increase in WRKY70 expression, which was delayed until 24 hpi, and ICS1 expression rose only 2.5-fold by 24 hpi and remained effectively stable at later time points. Similarly, there was a roughly twofold induction in the expression of PR1 in MRI between 12 and 24 hpi, consistent with the upregulation of ICS1, and contrastingly, there was no increase in PR1 expression in SB22 between 12 and 72 hpi.

Discussion

Here, we report the results of transcriptomic sequencing of basil cultivars infected with *Peronospora belbahrii* with the goals of identifying genes likely involved in resistance and the associated signaling pathways. We sequenced 33 datasets to a reasonable coverage, likely capturing most transcripts, although the average predicted gene length was shorter than a previous study in sweet basil (Rastogi et al. 2014), and Trinity's estimate of gene count in the plant samples was greater than expected. Nonetheless, we identified roughly 21,000 orthologous genes between cultivars with overall high sequence conservation consistent with closely related individuals. Phylogenetic analysis supports the hypothesis that SB22 and MRI are closely related, intraspecific genotypes, which was previously shown using simple sequence repeats to confirm conservation of the orthologs in the two breeding parents (Pyne et al. 2018). These results underscore the close relationship between cultivars and their sexual compatibility, likely providing us with a pool of true genes shared in MRI and SB22. At the time of this study, no sweet basil genome was available for use in a reference-based assembly, though a draft genome for the Genovese-type cultivar 'Perrie' has since been published (Gonda et al. 2020; Zhang et al. 2025). Furthermore, several unique MRI gene candidate sequences were used to search the *O. basilicum* draft genome, and no significant BLAST hits were retrieved.

Gene expression clustering identified patterns consistent with MRI-unique infection-expressed genes. Annotated MRI-unique candidates were found to be enriched for NLR, RLK, and secondary metabolite proteins. Analysis of the NLR and malectin-like RLK genes in both MRI and SB22 identified orthologous pairs and cultivar-unique genes. Members of the NLR and RLK families are known to act alongside or upstream of disease signaling pathways

and serve as good candidate resistance genes. Secondary metabolite genes identified in the two cultivars correlate well with known chemotypic characteristics differentiating MRI and SB22.

The top candidate NLR prediction is supported by PCR screening revealing its presence in resistant MRI and absence in susceptible SB22, though amplification of a fragment in SB22 suggests partial sequence conservation in a polymorphic gene of unknown functional status in SB22. Amplification of the full sequence was further supported by the cloning and sequencing-based confirmation of MRI-R1 alleles, as well as predicted amino acid sequence analysis and protein structure modeling, supporting the hypothesis that this is a unique NLR likely involved in immune responses to pathogen infection. The conserved MADA motif further indicates that MRI-R1 is likely a helper NLR, functioning to promote sensor NLR activation and mediating signal transduction pathways (Adachi et al. 2019).

MRI-R1 was detected in only two of the four resistant cultivars, Devotion and Obsession. However, recent QTL analysis detected at least two major genomic regions (LOD > 4.0) that control downy mildew resistance in the MRI × SB22 F2 mapping population (Pyne et al. 2017), suggesting that the predicted involvement of MRI-R1 in quantitative disease resistance may be redundant or have shared function(s) with other helper NLRs, consistent with the overlapping NRC network functions (Contreras et al. 2022, 2023). We cannot conclusively determine the functionality of the alleles simply based on the bias for MRI-R1 allele expression in MRI. The presence of allele 1 in Obsession and allele 4 in Devotion indicates that there may be functional redundancy within this allelic series, and further understanding of the conserved motifs in these alleles suggests that there are likely interacting partners that have an impact on the activity of the encoded proteins. Nevertheless, we are confident that our transcriptomic analyses were robust in detecting unique genes involved in the MRI background involved in resistance.

In addition to the prediction of specific genes likely conferring resistance, this comparative transcriptomic approach was also valuable in revealing physiological mechanisms involved in BDM resistance. SA signaling is an integral part of plant defense responses and has been demonstrated to be involved in defense against downy mildews and other biotrophic pathogens (Delaney et al. 1994; Mohr et al. 2010). Pathogen virulence strategies have developed to overcome and inhibit SA defenses, thus enhancing susceptibility to biotrophic pathogens such as downy mildew organisms (Caillaud et al. 2013, 2016). This suggests a likely role of SA signaling in MRI BDM resistance, though the specific mechanisms of signaling induction remain unknown. We hypothesize that multiple NLRs and RLKs are active and have interacting or redundant roles in mediating the SA signaling pathway in MRI. If this hypothesis is correct, utilizing multiple targets in a marker-assisted selective breeding program will be more effective and robust to pass resistance from parents to progeny to slow the evolution of new pathogen races.

This study utilized mRNA sequencing over an infection time course to provide strong evidence linking susceptibility to known mechanisms that control defense responses and predicted genes that regulate the resistant phenotype. Transcriptomics without a reference genome is a powerful tool for comparative analyses given the availability of data annotation and pattern identification methods. The strong resistance phenotype of MRI compared with SB22 likely led to the strength of the visible signal between the cultivars. NLRs that confer resistance to biotrophic pathogens have increased genetic diversity, and here, we observed that, despite high genetic conservation between orthologs (Fig. 1), there are distinct genetic differences in the NLR repertoire (Van de Weyer et al. 2019).

Generating cultivars with quantitative resistance has been shown to produce durable, non-race-specific resistance, typically through the activity of multiple minor-effect genes (Brown 2015; Niks et al. 2015). Other BDM-resistant cultivars have been produced through interspecific hybridization of *O. basilicum* with *O. americanum* var. *pilosum*, leading to dominant resistance against two races of

P. belbahrii (Ben-Naim and Weitman 2022). We have observed that the DMR basil cultivars succumb to infection in production systems in different regions, consistent with the report of emerging pathogen races (Ben-Naim and Weitman 2022). Therefore, understanding the physiological and molecular bases of host–pathogen interactions is critical to rapidly develop improved cultivars, monitoring strategies, and management practices. Identification of suitable molecular markers conferring multiple sources of resistance will improve breeding for BDM resistance and advance the understanding of molecular mechanisms of resistance and pathogenicity. These developments in approaches and knowledge can be broadly utilized by plant breeders and pathologists working with downy mildew pathogens on many different crops. After further validation, the genes predicted here can likely serve as molecular markers for the selection of DMR sweet basils, and investigation and validation of physiological responses to infection may aid in developing more robust phenotyping assays. This comparative transcriptomics approach not only revealed candidate resistance genes but also offered us new insights into differential infection responses in resistant and susceptible cultivars, which will open new avenues of investigation to further combat BDM.

This method of analysis is broadly applicable to two-organism biological systems where the identification of genes involved in any specific interaction is desired. Although genomes for holy basil (*O. tenuiflorum*) had been published prior to this study and sweet basil (*O. basilicum*) and *P. belbahrii* genomes have been published since, the methods described here worked exclusively from RNA sequencing data and did not require the whole-genome sequence. Any system utilizing two organisms from different kingdoms could be examined using these methods, as DNA sequences are differentiable down to reasonable taxonomic levels.

Literature Cited

Adachi, H., Contreras, M. P., Harant, A., Wu, C.-h., Derevnina, L., Sakai, T., Duggan, C., Moratto, E., Bozkurt, T. O., Maqbool, A., Win, J., and Kamoun, S. 2019. An N-terminal motif in NLR immune receptors is functionally conserved across distantly related plant species. eLife 8:e49956.

Altschul, S. F., Gish, W., Miller, W., Myers, E. W., and Lipman, D. J. 1990. Basic local alignment search tool. J. Mol. Biol. 215:403-410.

Andrews, S. 2010. FastQC. A quality control tool for high throughput sequence data. Babraham Bioinformatics. https://www.bioinformatics.babraham.ac.uk/projects/fastqc/

Baek, M., DiMaio, F., Anishchenko, I., Dauparas, J., Ovchinnikov, S., Lee, G. R., Wang, J., Cong, Q., Kinch, L. N., Schaeffer, R. D., Millán, C., Park, H., Adams, C., Glassman, C. R., DeGiovanni, A., Pereira, J. H., Rodrigues, A. V., van Dijk, A. A., Ebrecht, A. C., Opperman, D. J., Sagmeister, T., Buhlheller, C., Pavkov-Keller, T., Rathinaswamy, M. K., Dalwadi, U., Yip, C. K., Burke, J. E., Garcia, K. C., Grishin, N. V., Adams, P. D., Read, R. J., and Baker, D. 2021. Accurate prediction of protein structures and interactions using a three-track neural network. Science 373:871-876.

Belbahri, L., Calmin, G., Pawlowski, J., and Lefort, F. 2005. Phylogenetic analysis and real time PCR detection of a presumably undescribed *Peronospora* species on sweet basil and sage. Mycol. Res. 109:1276-1287.

Ben-Naim, Y., and Weitman, M. 2022. Joint action of *Pb1* and *Pb2* provides dominant complementary resistance against new races of *Peronospora belbahrii* (basil downy mildew). Phytopathology 112:595-607.

Bentham, A. R., Zdrzałek, R., De la Concepcion, J. C., and Banfield, M. J. 2018. Uncoiling CNLs: Structure/function approaches to understanding CC domain function in plant NLRs. Plant Cell Physiol. 59:2398-2408.

Bigeard, J., Colcombet, J., and Hirt, H. 2015. Signaling mechanisms in pattern-triggered immunity (PTI). Mol. Plant 8:521-539.

Bolger, A. M., Lohse, M., and Usadel, B. 2014. Trimmomatic: A flexible trimmer for Illumina sequence data. Bioinformatics 30:2114-2120.

Brindisi, L. J., Mattera, R., III, Mudiyala, S., Honig, J., and Simon, J. E. 2024. Genetic linkage mapping and quantitative trait locus (QTL) analysis of sweet basil (*Ocimum basilicum* L.) to identify genomic regions associated with cold tolerance and major volatiles. PLoS One 19:e0299825.

Brown, J. K. M. 2015. Durable resistance of crops to disease: A Darwinian perspective. Annu. Rev. Phytopathol. 53:513-539.

Caillaud, M.-C., Asai, S., Rallapalli, G., Piquerez, S., Fabro, G., and Jones, J. D. G. 2013. A downy mildew effector attenuates salicylic acid-triggered immunity in Arabidopsis by interacting with the host mediator complex. PLoS Biol. 11:e1001732.

Caillaud, M.-C., Asai, S., Rallapalli, G., Piquerez, S., Fabro, G., and Jones, J. D. G. 2016. Correction: A downy mildew effector attenuates salicylic acid-triggered immunity in Arabidopsis by interacting with the host mediator complex. PLoS Biol. 14:e1002408.

Cesari, S. 2018. Multiple strategies for pathogen perception by plant immune receptors. New Phytol. 219:17-24.

Chern, M., Xu, Q., Bart, R. S., Bai, W., Ruan, D., Sze-To, W. H., Canlas, P. E., Jain, R., Chen, X., and Ronald, P. C. 2016. A genetic screen identifies a requirement for cysteine-rich–receptor-like kinases in rice NH1 (OsNPR1)-mediated immunity. PLoS Genet. 12:e1006049.

Choi, Y. J., Choi, I. Y., Lee, K. J., and Shin, H. D. 2016. First report of downy mildew caused by *Peronospora belbahrii* on sweet basil (*Ocimum basilicum*) in Korea. Plant Dis. 100:2335.

Cohen, Y., and Ben-Naim, Y. 2016. Nocturnal fanning suppresses downy mildew epidemics in sweet basil. PLoS One 11:e0155330.

Cohen, Y., Ben Naim, Y., Falach, L., and Rubin, A. E. 2017. Epidemiology of basil downy mildew. Phytopathology 107:1149-1160.

Cohen, Y., and Rubin, A. E. 2015. Daytime solar heating controls downy mildew *Peronospora belbahrii* in sweet basil. PLoS One 10:e0126103.

Cohen, Y., Vaknin, M., Ben-Naim, Y., and Rubin, A. E. 2013. Light suppresses sporulation and epidemics of *Peronospora belbahrii*. PLoS One 8:e81282.

Conesa, A., Götz, S., García-Gómez, J. M., Terol, J., Talón, M., and Robles, M. 2005. Blast2GO: A universal tool for annotation, visualization and analysis in functional genomics research. Bioinformatics 21:3674-3676.

Contreras, M. P., Lüdke, D., Pai, H., Toghani, A., and Kamoun, S. 2023. NLR receptors in plant immunity: Making sense of the alphabet soup. EMBO Rep. 24:e57495.

Contreras, M. P., Pai, H., Tumtas, Y., Duggan, C., Him Yuen, E. L., Cruces, A. V., Kourelis, J., Ahn, H.-K., Lee, K.-T., Wu, C.-H., Bozkurt, T. O., Derevnina, L., and Kamoun, S. 2022. Sensor NLR immune proteins activate oligomerization of their NRC helpers in response to plant pathogens. EMBO J. 42:e111519.

Cui, H., Tsuda, K., and Parker, J. E. 2015. Effector-triggered immunity: From pathogen perception to robust defense. Annu. Rev. Plant Biol. 66:487-511.

Delaney, T. P., Uknes, S., Vernooij, B., Friedrich, L., Weymann, K., Negrotto, D., Gaffney, T., Gut-Rella, M., Kessmann, H., Ward, E., and Ryals, J. 1994. A central role of salicylic acid in plant disease resistance. Science 266:1247-1250.

Dobin, A., Davis, C. A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M., and Gingeras, T. R. 2013. STAR: Ultrafast universal RNA-seq aligner. Bioinformatics 29:15-21.

Dodds, P. N., and Rathjen, J. P. 2010. Plant immunity: Towards an integrated view of plant–pathogen interactions. Nat. Rev. Genet. 11:539-548.

Du, Z., Su, H., Wang, W., Ye, L., Wei, H., Peng, Z., Anishchenko, I., Baker, D., and Yang, J. 2021. The trRosetta server for fast and accurate protein structure prediction. Nat. Protoc. 16:5634-5651.

Finn, R. D., Coggill, P., Eberhardt, R. Y., Eddy, S. R., Mistry, J., Mitchell, A. L., Potter, S. C., Punta, M., Qureshi, M., Sangrador-Vegas, A., Salazar, G. A., Tate, J., and Bateman, A. 2016. The Pfam protein families database: Towards a more sustainable future. Nucleic Acids Res. 44:D279-D285.

Garibaldi, A., Minuto, A., and Gullino, M. L. 2005. First report of downy mildew caused by *Peronospora* sp. on basil (*Ocimum basilicum*) in France. Plant Dis. 89:683.

Garibaldi, A., Minuto, A., Minuto, G., and Gullino, M. L. 2004. First report of downy mildew on basil (*Ocimum basilicum*) in Italy. Plant Dis. 88:312.

Gonda, I., Faigenboim, A., Adler, C., Milavski, R., Karp, M.-J., Shachter, A., Ronen, G., Baruch, K., Chaimovitsh, D., and Dudai, N. 2020. The genome sequence of tetraploid sweet basil, *Ocimum basilicum* L., provides tools for advanced genome editing and molecular breeding. DNA Res. 27:dsaa027.

Grabherr, M. G., Haas, B. J., Yassour, M., Levin, J. Z., Thompson, D. A., Amit, I., Adiconis, X., Fan, L., Raychowdhury, R., Zeng, Q., Chen, Z., Mauceli, E., Hacohen, N., Gnirke, A., Rhind, N., di Palma, F., Birren, B. W., Nusbaum, C., Lindblad-Toh, K., Friedman, N., and Regev, A. 2011. Full-length transcriptome assembly from RNA-Seq data without a reference genome. Nat. Biotechnol. 29:644-652.

Guo, L., Allen, K. S., Deiulio, G., Zhang, Y., Madeiras, A. M., Wick, R. L., and Ma, L.-J. 2016. A *de novo*-assembly based data analysis pipeline for plant obligate parasite metatranscriptomic studies. Front. Plant Sci. 7:925.

Haas, B. J., Kamoun, S., Zody, M. C., Jiang, R. H. Y., Handsaker, R. E., Cano, L. M., Grabherr, M., Kodira, C. D., Raffaele, S., Torto-Alalibo, T., Bozkurt, T. O., Ah-Fong, A. M. V., Alvarado, L., Anderson, V. L., Armstrong, M. R., Avrova, A., Baxter, L., Beynon, J., Boevink, P. C., Bollmann, S. R., Bos, J. I. B., Bulone, V., Cai, G., Cakir, C., Carrington, J. C., Chawner, M., Conti, L., Costanzo, S., Ewan, R., Fahlgren, N., Fischbach, M. A., Fugelstad, J., Gilroy, E. M., Gnerre, S., Green, P. J., Grenville-Briggs, L. J., Griffith, J., Grünwald, N. J., Horn, K., Horner, N. R., Hu, C.-H., Huitema, E., Jeong, D.-H., Jones, A. M. E., Jones, J. D. G., Jones, R. W., Karlsson, E. K., Kunjeti, S. G., Lamour, K., Liu, Z., Ma, L., MacLean, D., Chibucos, M. C., McDonald, H., McWalters, J., Meijer, H. J. G., Morgan, W., Morris, P. F., Munro, C. A., O'Neill, K., Ospina-Giraldo, M., Pinzón, A., Pritchard, L., Ramsahoye, B., Ren, Q., Restrepo, S., Roy, S., Sadanandom, A., Savidor, A., Schornack, S., Schwartz, D. C., Schumann, U. D., Schwessinger, B., Seyer, L., Sharpe, T., Silvar, C., Song, J., Studholme, D. J., Sykes, S., Thines, M., van de Vondervoort, P. J. I., Phuntumart, V., Wawra, S., Weide, R., Win, J., Young, C., Zhou, S., Fry, W., Meyers, B. C., van West, P., Ristaino, J., Govers, F., Birch, P. R. J., Whisson, S. C., Judelson, H. S., and Nusbaum, C. 2009. Genome sequence and analysis of the Irish potato famine pathogen *Phytophthora infestans*. Nature 461:393-398.

Holm, L. 2020. Using Dali for protein structure comparison. Pages 29-42 in: Structural Bioinformatics: Methods and Protocols. Methods in Molecular Biology. Z. Gáspári, ed. Humana Press, New York.

Homa, K., Barney, W. P., Ward, D. L., Wyenandt, C. A., and Simon, J. E. 2014. Evaluation of fungicides for the control of *Peronospora belbahrii* on sweet basil in New Jersey. Plant Dis. 98:1561-1566.

Johnson, E. T., Kim, H.-S., Tian, M., Dudai, N., Tal, O., and Gonda, I. 2022. Dual transcriptional analysis of *Ocimum basilicum* and *Peronospora belbahrii* in susceptible interactions. Plant Gene 29:100350.

Jones, J. D. G., and Dangl, J. L. 2006. The plant immune system. Nature 444:323-329.

Jones, P., Binns, D., Chang, H.-Y., Fraser, M., Li, W., McAnulla, C., McWilliam, H., Maslen, J., Mitchell, A., Nuka, G., Pesseat, S., Quinn, A. F., Sangrador-Vegas, A., Scheremetjew, M., Yong, S.-Y., Lopez, R., and Hunter, S. 2014. InterProScan 5: Genome-scale protein function classification. Bioinformatics 30:1236-1240.

Kanetis, L., Vasiliou, A., Neophytou, G., Samouel, S., and Tsaltas, D. 2014. First report of downy mildew caused by *Peronospora belbahrii* on sweet basil (*Ocimum basilicum*) in Cyprus. Plant Dis. 98:283.

Khateri, H., Calmin, G., Moarrefzadeh, N., Belbahri, L., and Lefort, F. 2007. First report of downy mildew caused by *Peronospora* sp. on basil in Northern Iran. J. Plant Pathol. 89:S70.

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

Lefevere, H., Bauters, L., and Gheysen, G. 2020. Salicylic acid biosynthesis in plants. Front. Plant Sci. 11:338.

Li, B., and Dewey, C. N. 2011. RSEM: Accurate transcript quantification from RNA-Seq data with or without a reference genome. BMC Bioinform. 12:323.

Li, H., and Durbin, R. 2009. Fast and accurate short read alignment with Burrows–Wheeler transform. Bioinformatics 25:1754-1760.

Li, J., Brader, G., and Palva, E. T. 2004. The WRKY70 transcription factor: A node of convergence for jasmonate-mediated and salicylate-mediated signals in plant defense. Plant Cell 16:319-331.

Madeira, F., Park, Y. m., Lee, J., Buso, N., Gur, T., Madhusoodanan, N., Basutkar, P., Tivey, A. R. N., Potter, S. C., Finn, R. D., and Lopez, R. 2019. The EMBL-EBI search and sequence analysis tools APIs in 2019. Nucleic Acids Res. 47:W636-W641.

Madeira, F., Pearce, M., Tivey, A. R. N., Basutkar, P., Lee, J., Edbali, O., Madhusoodanan, N., Kolesnikov, A., and Lopez, R. 2022. Search and sequence analysis tools services from EMBL-EBI in 2022. Nucleic Acids Res. 50:W276-W279.

Marchler-Bauer, A., Derbyshire, M. K., Gonzales, N. R., Lu, S., Chitsaz, F., Geer, L. Y., Geer, R. C., He, J., Gwadz, M., Hurwitz, D. I., Lanczycki, C. J., Lu, F., Marchler, G. H., Song, J. S., Thanki, N., Wang, Z., Yamashita, R. A., Zhang, D., Zheng, C., and Bryant, S. H. 2015. CDD: NCBI's conserved domain database. Nucleic Acids Res. 43:D222-D226.

Martínez de la Parte, E., Pérez-Vicente, L., Bernal, B., and García, D. 2010. First report of *Peronospora* sp. on sweet basil (*Ocimum basilicum*) in Cuba. Plant Pathol. 59:800.

McGrath, M. T. 2020. Efficacy of conventional fungicides for downy mildew in field-grown sweet basil in the United States. Plant Dis. 104:2967-2972.

McGrath, M. T. 2021. Where in the USA is Basil Downy Mildew? Cornell University. https://blogs.cornell.edu/livegpath/extension/basil-downy-mildew/where-in-the-usa-is-basil-downy-mildew/

McGrath, M. T. 2023a. Basil Downy Mildew. Cornell Vegetables. Cornell University. https://www.vegetables.cornell.edu/pest-management/disease-factsheets/basil-downy-mildew/

McGrath, M. T. 2023b. Efficacy of organic fungicides for downy mildew in field-grown sweet basil. Plant Dis. 107:2467-2473.

McHale, L., Tan, X., Koehl, P., and Michelmore, R. W. 2006. Plant NBS-LRR proteins: Adaptable guards. Genome Biol. 7:212.

McLeod, A., Coertze, S., and Mostert, L. 2006. First report of a *Peronospora* species on sweet basil in South Africa. Plant Dis. 90:1115.

Mendy, B., Wang'ombe, M. W., Radakovic, Z. S., Holbein, J., Ilyas, M., Chopra, D., Holton, N., Zipfel, C., Grundler, F. M. W., and Siddique, S. 2017. Arabidopsis leucine-rich repeat receptor–like kinase NILR1 is required for induction of innate immunity to parasitic nematodes. PLoS Pathog. 13:e1006284.

Minh, B. Q., Schmidt, H. A., Chernomor, O., Schrempf, D., Woodhams, M. D., von Haeseler, A., and Lanfear, R. 2020. IQ-TREE 2: New models and efficient methods for phylogenetic inference in the genomic era. Mol. Biol. Evol. 37:1530-1534.

Mohr, T. J., Mammarella, N. D., Hoff, T., Woffenden, B. J., Jelesko, J. G., and McDowell, J. M. 2010. The *Arabidopsis* downy mildew resistance gene *RPP8* is induced by pathogens and salicylic acid and is regulated by W box *cis* elements. Mol. Plant-Microbe Interact. 23:1303-1315.

Monteiro, F., and Nishimura, M. T. 2018. Structural, functional, and genomic diversity of plant NLR proteins: An evolved resource for rational engineering of plant immunity. Annu. Rev. Phytopathol. 56:243-267.

Nagy, G., and Horváth, A. 2011. Occurrence of downy mildew caused by *Peronospora belbahrii* on sweet basil in Hungary. Plant Dis. 95:1034.

Nelson, R., Wiesner-Hanks, T., Wisser, R., and Balint-Kurti, P. 2018. Navigating complexity to breed disease-resistant crops. Nat. Rev. Genet. 19:21-33.

Niks, R. E., Qi, X., and Marcel, T. C. 2015. Quantitative resistance to biotrophic filamentous plant pathogens: Concepts, misconceptions, and mechanisms. Annu. Rev. Phytopathol. 53:445-470.

Olsen, K. M., Lea, U. S., Slimestad, R., Verheul, M., and Lillo, C. 2008. Differential expression of four *Arabidopsis PAL* genes; *PAL1* and *PAL2* have functional specialization in abiotic environmental-triggered flavonoid synthesis. J. Plant Physiol. 165:1491-1499.

Pettersen, E. F., Goddard, T. D., Huang, C. C., Couch, G. S., Greenblatt, D. M., Meng, E. C., and Ferrin, T. E. 2004. UCSF Chimera—A visualization system for exploratory research and analysis. J. Comput. Chem. 25:1605-1612.

Pyne, R., Honig, J., Vaiciunas, J., Koroch, A., Wyenandt, C., Bonos, S., and Simon, J. 2017. A first linkage map and downy mildew resistance QTL discovery for sweet basil (*Ocimum basilicum*) facilitated by double digestion restriction site associated DNA sequencing (ddRADseq). PLoS One 12:e0184319.

Pyne, R. M., Honig, J. A., Vaiciunas, J., Wyenandt, C. A., and Simon, J. E. 2018. Population structure, genetic diversity and downy mildew resistance among *Ocimum* species germplasm. BMC Plant Biol. 18:69.

Pyne, R. M., Koroch, A. R., Wyenandt, C. A., and Simon, J. E. 2014. A Rapid screening approach to identify resistance to basil downy mildew (*Peronospora belbahrii*). HortScience 49:1041-1045.

Pyne, R. M., Koroch, A. R., Wyenandt, C. A., and Simon, J. E. 2015. Inheritance of resistance to downy mildew in sweet basil. J. Am. Soc. Hortic. Sci. 140:396-403.

Ramakrishnan, C., Dani, V. S., and Ramasarma, T. 2002. A conformational analysis of walker motif A [GXXXXGKT (S)] in nucleotide-binding and other proteins. Protein Eng. Des. Sel. 15:783-798.

Rastogi, S., Kalra, A., Gupta, V., Khan, F., Lal, R. K., Tripathi, A. K., Parameswaran, S., Gopalakrishnan, C., Ramaswamy, G., and Shasany, A. K. 2015. Unravelling the genome of holy basil: An "incomparable" "elixir of life" of traditional Indian medicine. BMC Genomics 16:413.

Rastogi, S., Meena, S., Bhattacharya, A., Ghosh, S., Shukla, R. K., Sangwan, N. S., Lal, R. K., Gupta, M. M., Lavania, U. C., Gupta, V., Nagegowda, D. A., and Shasany, A. K. 2014. *De novo* sequencing and comparative analysis of holy and sweet basil transcriptomes. BMC Genomics 15:588.

Robert, X., and Gouet, P. 2014. Deciphering key features in protein structures with the new ENDscript server. Nucleic Acids Res. 42:W320-W324.

Roberts, P. D., Raid, R. N., Harmon, P. F., Jordan, S. A., and Palmateer, A. J. 2009. First report of downy mildew caused by a *Peronospora* sp. on basil in Florida and the United States. Plant Dis. 93:199.

Robinson, J. T., Thorvaldsdóttir, H., Winckler, W., Guttman, M., Lander, E. S., Getz, G., and Mesirov, J. P. 2011. Integrative genomics viewer. Nat. Biotechnol. 29:24-26.

Robinson, M. D., McCarthy, D. J., and Smyth, G. K. 2010. edgeR: A Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics 26:139-140.

Ronco, L., Rollán, C., Choi, Y. J., and Shin, H. D. 2009. Downy mildew of sweet basil (*Ocimum basilicum*) caused by *Peronospora* sp. in Argentina. Plant Pathol. 58:395.

Šafránková, I., and Holková, L. 2014. The first report of downy mildew caused by *Peronospora belbahrii* on sweet basil in greenhouses in the Czech Republic. Plant Dis. 98:1579.

Sawano, Y., Miyakawa, T., Yamazaki, H., Tanokura, M., and Hatano, K.-i. 2007. Purification, characterization, and molecular gene cloning of an antifungal protein from *Ginkgo biloba* seeds. Biol. Chem. 388:273-280.

Simon, J. E., Wyenandt, C. A., and Pyne, R. M. 2018. Downy Mildew Resistant/Tolerant Sweet Basil Varieties. U.S. Patent No. US 10,159,212 B1. Date of Patent issue: December 25, 2018.

Smith, S., and Urrea, K. 2018. First report of downy mildew caused by *Peronospora belbahrii* on sweet basil (*Ocimum basilicum*) in Arkansas. Plant Dis. 102:686.

Steele, J. F. C., Hughes, R. K., and Banfield, M. J. 2019. Structural and biochemical studies of an NB-ARC domain from a plant NLR immune receptor. PLoS One 14:e0221226.

Tamura, K., Stecher, G., Peterson, D., Filipski, A., and Kumar, S. 2013. MEGA6: Molecular evolutionary genetics analysis version 6.0. Mol. Biol. Evol. 30:2725-2729.

Thines, M., Sharma, R., Rodenburg, S. Y. A., Gogleva, A., Judelson, H. S., Xia, X., van den Hoogen, J., Kitner, M., Klein, J., Neilen, M., de Ridder, D., Seidl, M. F., van den Ackerveken, G., Govers, F., Schornack, S., and Studholme, D. J. 2020. The genome of *Peronospora belbahrii* reveals high heterozygosity, a low number of canonical effectors, and TC-rich promoters. Mol. Plant-Microbe Interact. 33:742-753.

Van de Weyer, A.-L., Monteiro, F., Furzer, O. J., Nishimura, M. T., Cevik, V., Witek, K., Jones, J. D. G., Dangl, J. L., Weigel, D., and Bemm, F. 2019. A species-wide inventory of NLR genes and alleles in *Arabidopsis thaliana*. Cell 178:1260-1272.e14.

van Ooijen, G., Mayr, G., Kasiem, M. M. A., Albrecht, M., Cornelissen, B. J. C., and Takken, F. L. W. 2008. Structure–function analysis of the NB-ARC domain of plant disease resistance proteins. J. Exp. Bot. 59:1383-1397.

Vieira, R. F., and Simon, J. E. 2006. Chemical characterization of basil (*Ocimum* spp.) based on volatile oils. Flavour Fragr. J. 21:214-221.

Wang, J., and Chai, J. 2020. Structural insights into the plant immune receptors PRRs and NLRs. Plant Physiol. 182:1566-1581.

Waterhouse, A. M., Procter, J. B., Martin, D. M. A., Clamp, M., and Barton, G. J. 2009. Jalview Version 2—A multiple sequence alignment editor and analysis workbench. Bioinformatics 25:1189-1191.

Wildermuth, M. C., Dewdney, J., Wu, G., and Ausubel, F. M. 2001. Isochorismate synthase is required to synthesize salicylic acid for plant defence. Nature 414:562-565.

Wyenandt, C. A., Simon, J. E., Pyne, R. M., Homa, K., McGrath, M. T., Zhang, S., Raid, R. N., Ma, L.-J., Wick, R., Guo, L., and Madeiras, A. 2015. Basil downy mildew (*Peronospora belbahrii*): Discoveries and challenges relative to its control. Phytopathology 105:885-894.

Yadav, S. S., Suryavanshi, P., Nishad, I., and Sinha, S. 2022. First report of downy mildew on sweet basil caused by *Peronospora belbahrii* in India. Plant Dis. 106:318.

Zhang, J., Abu-Abied, M., Milavski, R., Adler, C., Shachter, A., Kahane-Achinoam, T., Melnik-Ben-Gera, H., Davidovich-Rikanati, R., Powell, A. F., Chaimovitsh, D., Carmi, G., Dudai, N., Strickler, S. R., and Gonda, I. 2025. Chromosome-level assembly of basil genome unveils the genetic variation driving Genovese and Thai aroma types. Plant J. 121: e17224.

Zhang, L., Tian, L.-H., Zhao, J.-F., Song, Y., Zhang, C.-J., and Guo, Y. 2009. Identification of an apoplastic protein involved in the initial phase of salt stress response in rice root by two-dimensional electrophoresis. Plant Physiol. 149:916-928.

Zhou, J.-M., and Zhang, Y. 2020. Plant immunity: Danger perception and signaling. Cell 181:978-989.

Zhou, M., Lu, Y., Bethke, G., Harrison, B. T., Hatsugai, N., Katagiri, F., and Glazebrook, J. 2018. WRKY70 prevents axenic activation of plant immunity by direct repression of *SARD1*. New Phytol. 217:700-712.