

Article

Network Analysis of Publicly Available RNA-seq Provides Insights into the Molecular Mechanisms of Plant Defense against Multiple Fungal Pathogens in *Arabidopsis thaliana*

Cynthia Soto-Cardinault ^{1,†}, Kevin L. Childs ² and Elsa Góngora-Castillo ^{3,*}

¹ Unidad de Biotecnología, Centro de Investigación Científica de Yucatán, Mérida 97205, Mexico; cyntsc10@gmail.com or ccardin2@jh.edu

² Plant Biology Department, Michigan State University, East Lansing, MI 48824, USA; kchilds@msu.edu

³ CONAHCYT-Unidad de Biotecnología, Centro de Investigación Científica de Yucatán, Mérida 97205, Mexico

* Correspondence: elsa.gongora@cicy.mx

† Current address: Liber Institute for Brain Development, Baltimore, MD 21205, USA.

Abstract: Fungal pathogens can have devastating effects on global crop production, leading to annual economic losses ranging from 10% to 23%. In light of climate change-related challenges, researchers anticipate an increase in fungal infections as a result of shifting environmental conditions. However, plants have developed intricate molecular mechanisms for effective defense against fungal attacks. Understanding these mechanisms is essential to the development of new strategies for protecting crops from multiple fungi threats. Public omics databases provide valuable resources for research on plant–pathogen interactions; however, integrating data from different studies can be challenging due to experimental variation. In this study, we aimed to identify the core genes that defend against the pathogenic fungi *Colletotrichum higginsianum* and *Botrytis cinerea* in *Arabidopsis thaliana*. Using a custom framework to control batch effects and construct Gene Co-expression Networks in publicly available RNA-seq dataset from infected *A. thaliana* plants, we successfully identified a gene module that was responsive to both pathogens. We also performed gene annotation to reveal the roles of previously unknown protein-coding genes in plant defenses against fungal infections. This research demonstrates the potential of publicly available RNA-seq data for identifying the core genes involved in defending against multiple fungal pathogens.

Keywords: public RNA-seq datasets; weighted gene co-expression network analysis; gene module; arabidopsis; fungal pathogens; plant defense



Citation: Soto-Cardinault, C.; Childs, K.L.; Góngora-Castillo, E. Network Analysis of Publicly Available RNA-seq Provides Insights into the Molecular Mechanisms of Plant Defense against Multiple Fungal Pathogens in *Arabidopsis thaliana*. *Genes* **2023**, *14*, 2223. <https://doi.org/10.3390/genes14122223>

Academic Editors: Paola Vittorioso and Zhiqiang Wu

Received: 10 November 2023

Revised: 6 December 2023

Accepted: 12 December 2023

Published: 16 December 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Pathogenic fungi are a significant factor in global crop production. Pathogen-related crop damage leads to annual losses ranging from 10% to 23%—equivalent to over USD 200 billion in economic damages [1–3]. Climate change has led to unprecedented challenges in food production, including the growing resistance to antifungal agents. In a warmer world, fungal infections are expected to become more widespread among crops, as the life cycle of fungi is directly influenced by temperature and humidity [4–6]. According to prediction models, the prevalence of pathogenic fungal infections in commercial crops will rise by 5% to 100% by 2050 as a result of climate change [7,8].

Pathogenic fungi can be classified into three groups: biotrophs, which feed on living tissues; necrotrophs, which kill the host cell and obtain nutrients from dead tissues; and hemibiotrophs, which exhibit an intermediate lifestyle, combining traits of both biotrophs and necrotrophs [9–12]. One highly important pathogen is *Colletotrichum higginsianum*, an ascomycete fungus with a hemibiotrophic nature. This fungus causes anthracnose, a disease that affects numerous monocotyledonous and dicotyledonous plants worldwide [9,13,14]. Another important pathogen, *Botrytis cinerea*, is a necrotrophic fungus with a broad range of

hosts that has devastating effects on various plant tissues, including foliage, stems, flowers, and fruits [13,15–17].

Plants have developed complex molecular mechanisms to protect themselves from invading pathogens, including the ability to identify pathogens via pathogen-associated molecular patterns (PAMPs) with the help of pattern recognition receptors (PRRs). This leads to a process known as PAMP-triggered immunity (PTI), which limits the progression of disease. This pattern recognition is typically linked with calcium influx, callose deposition, bursts of reactive oxygen species (ROS), activation of miRNA pathways, activation of MAPK cascades, and increased expression of various defense-related genes, such as disease-related proteins (PR) [18–20]. In some cases, effectors produced by pathogens are recognized by plant resistance proteins, promoting a response known as effector-triggered immunity (ETI). One of the roles of ETI is to stimulate the hypersensitive response (HR), which results in rapid cell death at the site of pathogen invasion, limiting the spread of the pathogen. In addition, various signal transduction events, such as nitrous oxide, lipids, and various phytohormones that identify pathogen invasions, initiate plant immune defense responses [21–24]. The PTI and ETI immune responses are interconnected, as the activation of ETI improves the PTI signaling pathway. Both responses share functions related to the recognition of danger signals and the activation of defense mechanisms [18,25–28].

In the era of next-generation sequencing and machine learning, and considering the need for proactive responses to climate change, it is crucial to implement more comprehensive strategies in modern agriculture. One useful resource is the NCBI SRA, which contains more than 17 petabytes of public sequence data [29]. This underused data repository can be exploited using novel computational methods. A large portion of the sequence data is relevant to the agricultural sector, and opportunities are emerging to implement machine learning strategies that incorporate this immense pool of genomic data.

There is a growing interest in using public RNA-seq datasets to investigate intricate biological processes such as plant defense [30–32]. Publicly available RNA-seq expression data are a valuable resource for improving crop quality, particularly when conducting experiments under various environmental stressors or diseases is not feasible. However, the reutilization of public RNA-seq data is limited by the challenge of integrating batch datasets generated from different studies due to experimental variation, which can introduce bias; this must be carefully considered during analysis to ensure accurate and reliable results [30].

Traditionally, transcriptome studies with RNA-seq have been widely used for the identification of specific genes expressed during the fungal infection process, with a focus on specific characteristics of the plant or fungus [9,33–37]. In this study, our objective is to identify the core genes involved in defending against *B. cinerea* and *C. higginsianum* pathogens. We define core defense genes as those that confer a broad-spectrum defense mechanism and are not restricted to specific pathogens. We used publicly available RNA-seq expression information from the *C. higginsianum*–*A. thaliana* and *B. cinerea*–*A. thaliana* pathosystems. Both of these valuable models have been extensively studied with regard to specific responses during plant–pathogen interaction [14–17,38–41]. Additionally, only a limited number of studies have been conducted to understand common defense mechanisms against biotic stress, and our understanding of this topic is still at an early stage [42,43]. To date, the common reactions that these fungal pathogens can elicit in the host remain unclear.

One strategy, Weighted Gene Co-expression Network Analysis (WGCNA) [44], has contributed to crop improvement by revealing complex gene–gene interactions and helping to identify key genes that regulate important traits [45]. For instance, in soybeans, a gene module related to seed development was recognized and was found to include several genes involved in fatty acid biosynthesis [46]. In wheat, a set of genes associated with resistance to Fusarium head blight, a devastating fungal disease, was identified [47]. Furthermore, gene co-expression networks have been constructed using publicly available expression datasets. One of these studies used gene expression data to functionally annotate the rice proteome [48]. Moreover, other studies used gene co-expression networks to identify core genes in *A. thaliana* that respond to biotic stress using microarray data [42]

and those in maize that respond to fungal infection using RNA-seq data [43]. However, neither of these studies considered batch-effect correction.

To identify core genes expressed during *C. higginsianum* and *B. cinerea* defense in *A. thaliana*, publicly available RNA-seq data were preprocessed for batch-effect correction using an in-house method, allowing us to obtain integrated gene expression data that could be used with the WGCNA [44]. Our preprocessing framework includes data integration, normalization and standardization based on variability identification while emphasizing a minimal loss of data, and it is based on statistics such as Gaussian distributions, kernel density estimation (KDE), and quantiles, which assist with the generation of evaluation metrics [49]. A module containing core genes that respond to infection with *B. cinerea* and *C. higginsianum* was identified. Using the DAVID agglomeration method [50], 33 genes were annotated by function across six categories, providing new insights into the functions of currently unknown protein-coding genes. Our findings demonstrate that publicly available RNA-seq data are a valuable resource for identifying broad-spectrum genes related to the defense against multiple fungal pathogens.

2. Materials and Methods

2.1. RNA-Seq Data Selection

Twenty-five RNA-seq datasets of fungus-infected leaves of *A. thaliana* (Col-0) were obtained from the National Centre for Biotechnology Information's Sequence Read Archive (NCBI SRA) database. The *C. higginsianum*–*A. thaliana* dataset was composed of 8 samples from BioProject accession number PRJNA148307 [9,51]; the *B. cinerea*–*A. thaliana* dataset comprised 6 samples from BioProjects with accession numbers PRJNA315516 and PRJNA593073; the *Sclerotinia sclerotiorum*–*A. thaliana* dataset consisted of 3 samples from BioProject accession number PRJNA418121 [52], and this dataset was used as an outlier control. The control dataset comprised 8 samples of healthy plants from BioProjects PRJNA315516 and PRJNA418121. The sample collection was performed from 12 to 30 h for control samples, and from 12 to 40 h for fungus-infected leaf samples (Supplementary Table S1). The transcriptomes were sequenced on Illumina platforms. The number of reads ranged from 10 to 30 million, with read lengths ranging from 93 to 150 bp (Table 1).

Table 1. Description of public RNA-Seq datasets downloaded from the SRA.

BioProject	Sample ID	HPI ¹	Run	Layout	Reads (M ²)	% Clean Reads	% Aligned Reads
PRJNA148307	Ch22	22	SRR364389	SE	12.6	92.12	87.17
	Ch22.1		SRR364390	SE	12.4	92.03	77.45
	Ch22.2		SRR364391	SE	12.4	92.18	77.4
	Ch22.3		SRR364392	SE	12.2	92.08	87.29
	Ch40	40	SRR364400	SE	11.9	91.40	82.08
	Ch40.1		SRR364401	SE	11.9	91.40	82.25
	Ch40.2		SRR364398	SE	13.2	92.55	80.25
	Ch40.3		SRR364399	SE	13.2	92.73	80.25
PRJNA315516 PRJNA593073	Bc12	12	SRR3383696	SE	12.1	100.00	98.2
	Bc12.1		SRR3383697	SE	15	100.00	98.14
	Bc18	18	SRR3383779	SE	10.3	97.37	93.55
	Bc18.1		SRR3383780	SE	13.6	97.26	95.2
	Bc24	24	SRR10586397	PE	22.2	95.53	89.07
	Bc24.1		SRR10586399	PE	22	95.70	88.69
PRJNA418121	Ss30	30	SRR6283146	SE	20.8	95.01	36.48
	Ss30.1		SRR6283147	SE	20.9	96.09	32.35
	Ss30.2		SRR6283148	SE	21.2	92.90	29.5

Table 1. Cont.

BioProject	Sample ID	HPI ¹	Run	Layout	Reads (M ²)	% Clean Reads	% Aligned Reads
PRJNA315516 PRJNA418121	healthy12	12	SRR3383640	SE	10.9	97.12	98.31
	healthy12.1		SRR3383641	SE	22.6	97.50	98.34
	healthy18	18	SRR3383782	SE	29.8	97.62	98.13
	healthy18.1		SRR3383783	SE	14.2	97.69	98.34
	healthy24	24	SRR3383821	SE	15	97.18	98.09
	healthy24.1		SRR3383822	SE	10.2	95.98	97.87
	healthy30	30	SRR6283144	SE	22.1	95.31	97.51
	healthy30.1		SRR6283145	SE	19.7	95.72	97.99

¹ Hours Post Inoculation; ² Million reads.

2.2. Sequence Quality Filtering and Expression Estimation

Reads from each dataset were processed for quality and filtered using FastQC v0.11.5 [53] and Trimmomatic v0.38 [54] tools, using a Phred Score ≥ 20 and a minimum sequence length ≥ 40 bp. Filtered sequences were aligned to the *A. thaliana* genome TAIR10 [55] and the latest Araport 11 annotation (GenBank accessions CP002684–CP002688), using STAR alignment tool v2.7.5 [56], with parameters sjdbOverhang = 92 and genomeSAindexNbases = 7 to adjust the tool to the genome size and alignIntronMin = 8 and alignIntronMax = 1999 to adjust the minimum and maximum intron lengths [57]. The alignment quality reported in SAM/BAM files [58] were evaluated with the HTSeq-qa tool v0.11.1 and the expression abundance was estimated using the HTSeq-count tool v0.11.1 [59].

2.3. Data Preprocessing and Batch-Effect Correction

Central tendency and statistical dispersion measures were calculated, including the mean (μ), standard deviation (σ), and range (R), and histograms with KDE of distributions and percentiles were created using Seaborn [49] (Figure 1).

2.3.1. Integration of Raw Counts and Data Normalization (Figure 1a,b)

A single expression matrix was generated using the raw count files for each dataset. Protein-coding genes (henceforth referred to as genes) with expression values equal to zero were eliminated. Transcripts per million (TPM) [60] were used to perform data normalization with the TPM normalization tool v0.9.1 using Gene_length_extraction_from_GTF.ipynb [61]. For TPM normalization, the gene length was obtained from the *A. thaliana* genome GTF/GFF file [55]. The script 1_Step1_integrating_raw_counts.ipynb and script 2_Step2_TPM_normalization.ipynb were used to organize the gene expression data and perform data normalization, respectively (Table S2).

2.3.2. Data Standardization and Outlier Identification (Figure 1c,d)

The gene expression values were divided into percentiles, and top and bottom extreme tails were calculated. Genes with TPM values between the 1st and 99th percentile were retained, while genes beyond this range were discarded. After filtering, TPM values were transformed to $\log_2(\text{TPM} + 1)$. To evaluate the efficiency of our methods in detecting samples with atypical distributions, we included 3 samples with known atypical distribution as negative controls. The 3_Step3_TPM_standardization.ipynb and 4_Step4_Log2_scale.ipynb scripts were used (Table S2).

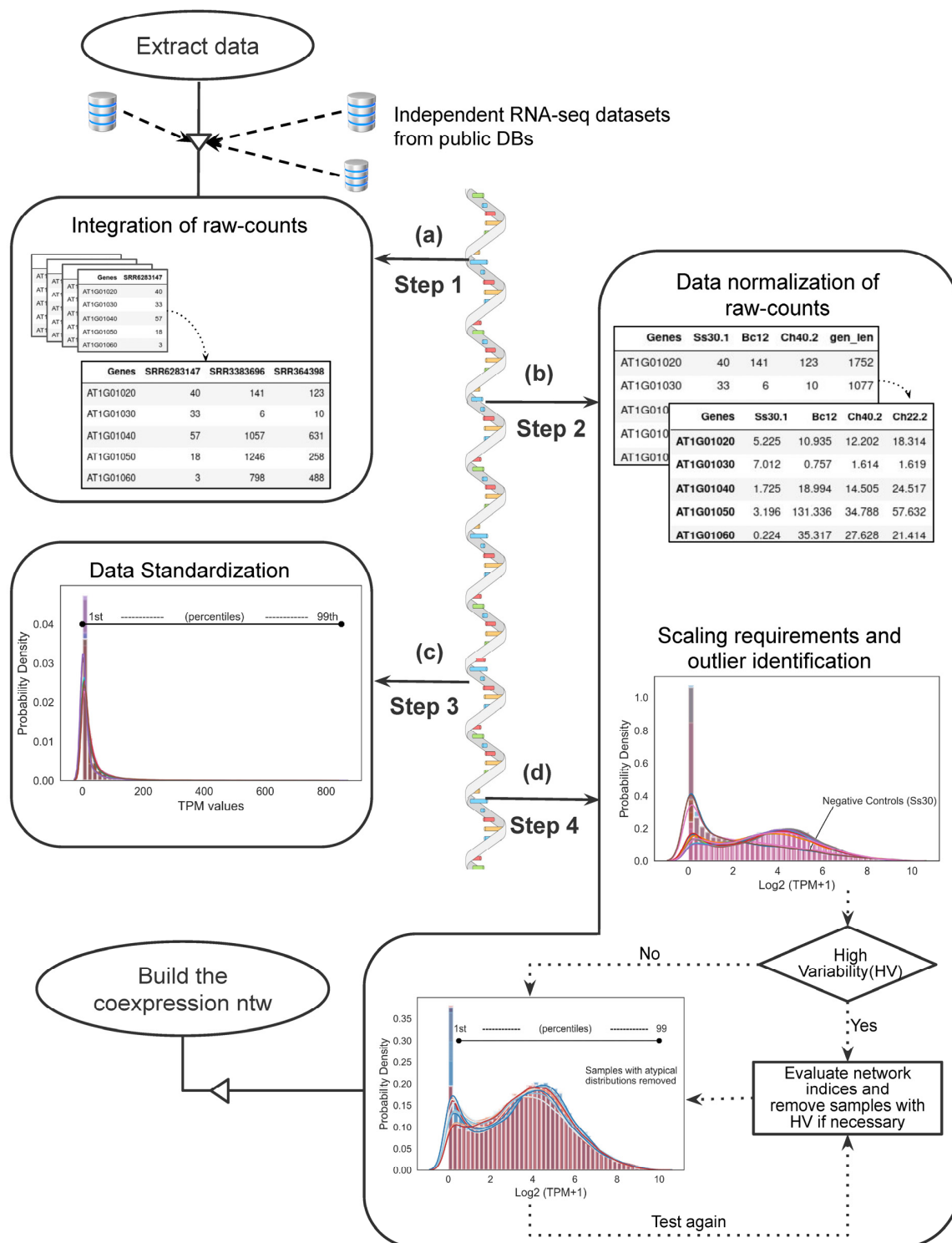


Figure 1. Framework for integrating publicly available RNA-seq data. (a) Step 1. Raw counts are integrated in a single expression matrix and evaluation metrics (standard deviation (σ), mean (μ) and range (R)) are calculated. Genes with zero expression values are removed. (b) Step 2. Data arrays are normalized by TPM and evaluation metrics are recalculated. (c) Step 3. Percentiles and distribution are calculated to identify tail values within the distribution. TPM values below the 1st and above the 99th percentile are removed, and evaluation metrics are recalculated. (d) Step 4. Normalized and filtered TPM values are transformed to $\log_2(\text{TPM} + 1)$ to identify atypical distributions, and evaluation metrics are recalculated.

2.4. Weighted Gene Co-Expression Networks and Identification of Core Genes

Weighted gene co-expression network analysis (WGCNA v1.69-81) was performed using R package [44]. The *pickSoftThreshold* function according to the Scale Free Topology model fit was used to define the appropriate soft-thresholding power. The adjacency matrix was generated using a signed topological overlap matrix (TOM), along with dissimilarity values (1-TOM). Gene clusters were detected using a dynamic cutting algorithm with a minimum module size of 20 and a default cutoff height (0.99). Merged networks were derived at 0.1 distances. Pearson's correlation analysis ($r^2 > 0.75$) was used to identify modules linked to phenotype. The *moduleEigens*, *cor*, *corPvalueStudent* functions were used to define module membership (MM) and gene significance (GS) with a cutoff of $r^2 > 0.75$, $r^2 < 0.75$ and p -value < 0.05 .

To identify the core genes, genetic modules from healthy-control and infected-plant gene networks were compared. Modules in the infected-plant gene network with a difference $>75\%$ in genes (unique genes) were selected as exclusive gene modules for the infected-plant gene co-expression network. The correlation cutoff was set at $R^2 > 0.50$. Gene modules that fit these criteria were selected as core genes.

2.5. Functional Annotation of the Gene Co-Expression Networks and Core Genes

A Gene Ontology (GO) overrepresentation test was performed using PANTHER v17.0 [62,63]. Gene modules with coefficients of $R^2 > 0.75$ were selected, and a binomial test with Bonferroni correction was used to identify GO-Slim terms for biological process and molecular function GO modules [64,65] (<http://geneontology.org/>, accessed on 11 July 2023).

The core genes were analyzed using The Database for Annotation, Visualization and Integrated Discovery (DAVID, <https://david.ncifcrf.gov/home.jsp>, accessed on 15 July 2023). Nine of the fourteen DAVID annotation categories were used in this study. To identify closely related genes within the module, the DAVID Gene Functional Annotation tool applies a *kappa* score distribution and an agglomeration method based on heuristic fuzzy multiple-linkage partitioning. These methods are implemented through the tool's web-based interface, and allow the user to select different levels of stringency, ranging from low to very high [66] (in this study, medium and high levels were selected). The selected gene groups were those repeated across the selection criteria. To identify the functional gene groups most relevant to the study (e.g., those of plants infected with pathogenic fungi), the DAVID tool uses the EASE score, which is a modified Fisher's exact test, to assign an enrichment score to each annotated gene group. The cutoff was >0.80 [50,66].

The Arabidopsis Information Portal (Araport) through ThaleMine webtool (<https://bar.utoronto.ca/thalemine/>, accessed on 4 September 2023) was used to identify each *A. thaliana* gene [67,68].

2.6. Implementation

Data processing was implemented using Jupyter Notebooks [69] for Python v3.8.10 and WGCNA v1.69-81 [44] for R v4.1.0. The R scripts and code are freely available via GitHub at <https://github.com/cyntsc/RNA-Seq-raw-integration> (accessed on 28 September 2023) (DOI 10.5281/zenodo.7076416) (Table S2).

3. Results

To improve our understanding of the core genes involved in defending against multiple fungal infections, we conducted a comparative study using publicly accessible transcriptome data from *C. higginsianum*– and *B. cinerea*–*A. thaliana* pathosystems (Tables 1 and S1). First, we controlled the batch effects of public data to reduce data variability using an in-house framework (Figure 1). Next, we performed a weighted gene co-expression network analysis (WGCNA) [44] of the preprocessed data to identify the core genes. These core genes were then subjected to further agglomerative analysis using DAVID [50] to identify functions in closely related gene groups.

3.1. Controlling the Batch Effect in Public RNA-seq Data to Construct Gene Co-Expression Networks (GCN)

RNA-seq sequences were filtered for quality, recovering more than 92% of the reads. These were aligned to the *A. thaliana* TAIR10 genome [55], achieving an alignment coverage rate of 97.5% to 98.3% for the healthy samples and 77.4% to 98.2% for the infected samples. *S. sclerotiorum* samples (Ss30, Ss30.1 and Ss30.2), with alignment coverages ranging from 29.5% to 36.5%, were included as an outlier control to track data changes (Table 1, Figure 2).

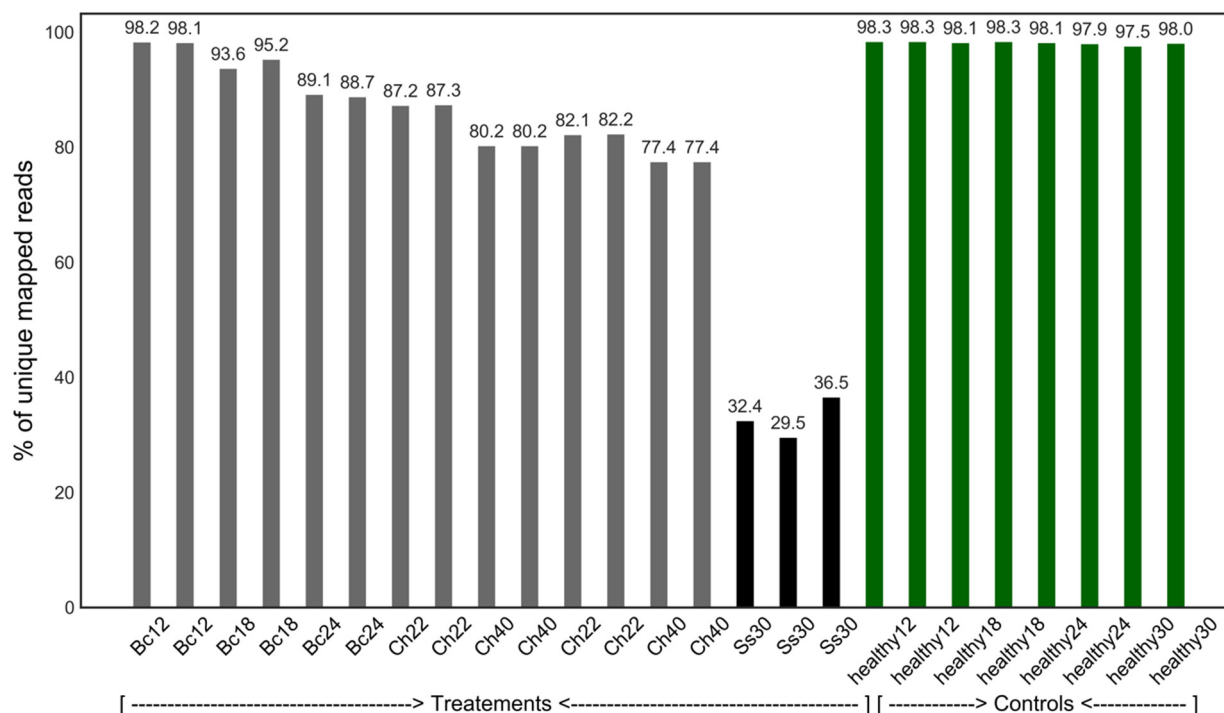


Figure 2. Rates of read alignments against the *Arabidopsis thaliana* genome. Percentage of unique reads mapped to the *A. thaliana* genome (TAIR11) from healthy-control and infected-plant samples.

After read mapping, two gene expression matrices were obtained, one containing more than 221 thousand expression values (27,655 genes $\times$ 8 samples) derived from healthy-control samples and another containing more than 470 thousand expression values (27,655 genes $\times$ 17 samples) from infected-plant samples. Genes with an expression value equal to zero were eliminated (Figure S1), resulting in a total of 22,426 and 24,239 genes in the healthy-control and infected expression matrices, respectively.

The expression values were normalized to TPM to set the means of the distributions to the same point, and the evaluation metrics showed a $\mu = 44.6$, $\sigma = 333 \pm 6$ and $R = [0 \text{ to } 36,543]$ for the healthy-control gene expression matrix and a $\mu = 41.2$, $\sigma = 284.5 \pm 84.5$ and $R = [0 \text{ to } 30,000]$ for the infected-plant gene expression matrix (Figure 3a). Extreme values in both expression matrices were identified and removed to reduce data variability. In the healthy-control expression matrix, 2262 genes were below the 1st percentile, in which TPM values were smaller than 0.1, and 372 genes were above the 99th percentile, with TPM values larger than 840. A total of 2634 genes were removed. In the infected-plant expression matrix, a total of 3965 genes were filtered. Of these, 3495 genes were below the 1st percentile and 470 genes were above the 99th percentile, where the TPM values were smaller than 0.1 and larger than 845, respectively. In total, 19,792 and 20,274 were obtained for the healthy-control and infected-plant expression matrices, respectively (Figure 3b). Data were transformed to $\log_2(\text{TPM} + 1)$ to reduce the scale, and the metrics for the healthy-control samples were $\sigma = 2.1 \pm 0.1$, $\mu = 3.4 \pm 0.2$ and $R = [0 \text{ to } 9.28]$ while for the infected-plant samples, the metrics were $\sigma = 2.1 \pm 0.1$, $\mu = 2.9 \pm 0.8$ and $R = [0 \text{ to } 9.72]$ (Figure 3c).

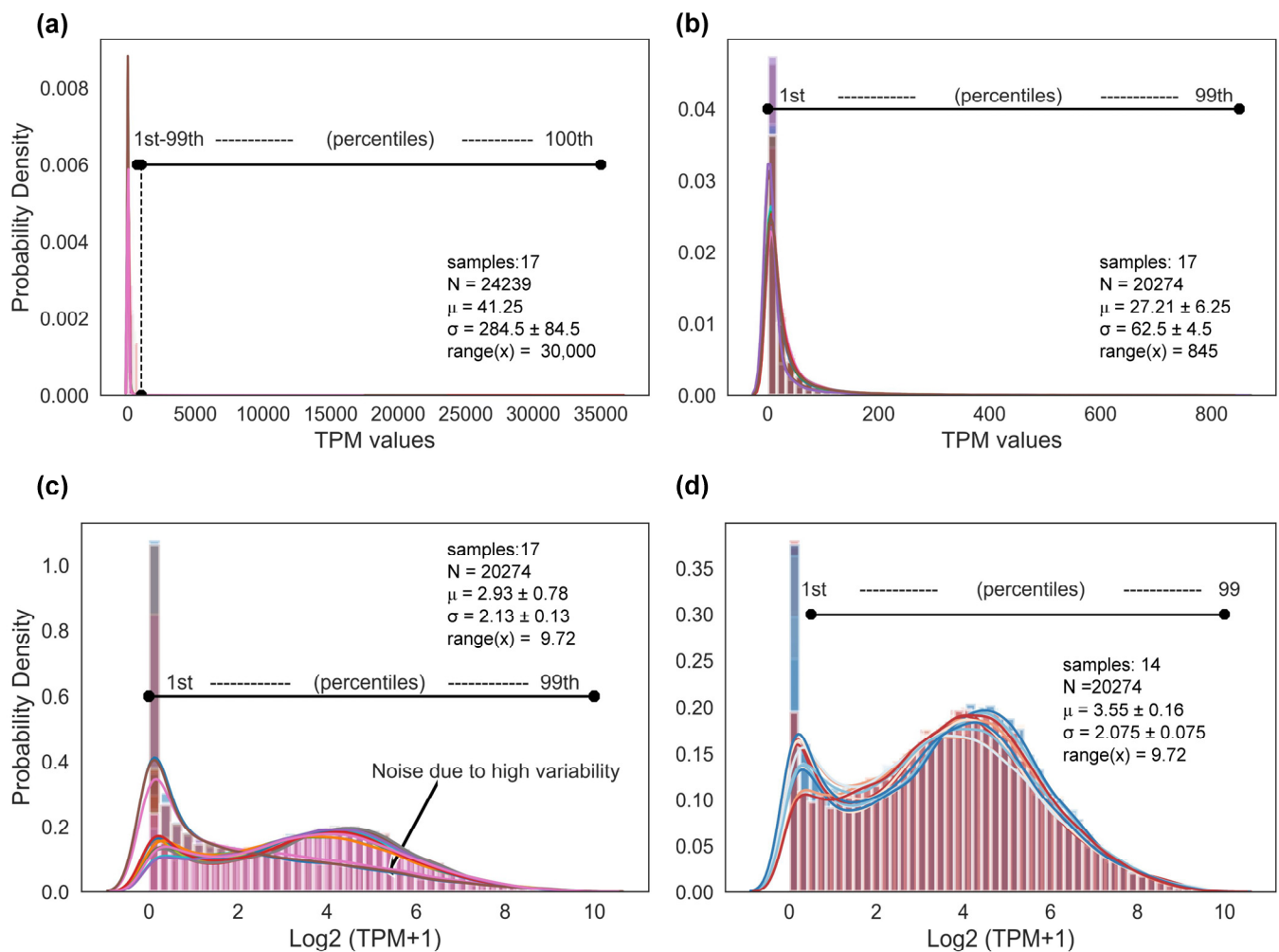


Figure 3. Integration of RNA-seq datasets. (a) The expression matrix was normalized to TPM and evaluation metrics were calculated, $\mu = 41.2$, $\sigma = 284.5 \pm 84.5$, and $R = [0 \text{ to } 30,000]$. (b) Percentiles and distribution were calculated to identify tail values within the distribution. TPM values below the 1st percentile (TPM < 0.1) and above the 99th percentile (TPM > 845) were removed. Evaluation metrics were calculated ($\mu = 27.21 \pm 6.25$, $\sigma = 62.5 \pm 4.5$ and $R = [0 \text{ to } 845]$). (c) TPM values were transformed to $\log_2(\text{TPM} + 1)$ to reduce the scale and to identify atypical distributions. Evaluation metrics were calculated ($\sigma = 2.93 \pm 0.78$, $\mu = 2.13 \pm 0.13$, and $R = [0 \text{ to } 9.72]$). (d) Fungal-infected samples with atypical distributions were removed (Ss30, Ss30.1 and Ss30.2) and evaluation metrics were calculated ($\sigma = 2.075 \pm 0.075$, $\mu = 3.55 \pm 0.16$, and $R = [0 \text{ to } 9.72]$). The colored lines represent the RNA-seq data used in this study; colored bars represent percentiles.

To understand and demonstrate the effect of highly variable gene expression, Ss30, Ss30.1 and Ss30.2 samples were added to introduce variability in the data collection, with the aim of monitoring the efficacy of the workflow corrections. The histograms with KDE reveal a clear contrast between the preponderant and the outlier distributions of the Ss30, Ss30.1 and Ss30.2 samples (Figure 3c). The violin plots with KDE show individual distributions and reveal a pronounced difference in the distributions of the Ss30, Ss30.1 and Ss30.2 samples compared to other samples (Figure S2). Once the Ss30, Ss30.1 and Ss30.2 samples were eliminated, a more homogeneous dataset was obtained for the co-expression analysis (Figures 3d and S2).

3.2. Gene Co-Expression Network Analyses and Gene Ontology Overrepresentation Test for Healthy and Infected Plants with Fungal Pathogens

Two gene co-expression networks were built. In the healthy-control gene network, a coefficient $R^2 = 0.80$ was reached, and 237 clusters were identified with a Node-Connectivity Mean (NCM) of 374 genes per cluster ($\beta = 28$; $R^2 = 0.78$). The merged network provided 23 clusters (Figure 4a, Table 2). In the infected-plant gene network, a coefficient $R^2 = 0.83$ was reached, and 100 clusters with a NCM of 270 genes per cluster were identified ($\beta = 27$; $R^2 = 0.79$). The merged network contained 36 clusters (Figure 4b, Table 2). A correlation of module eigengenes to disease trait data was performed, with a significant cutoff value of $|GS| > 0.75$ and p -value < 0.05 (Datas S1 and S2).

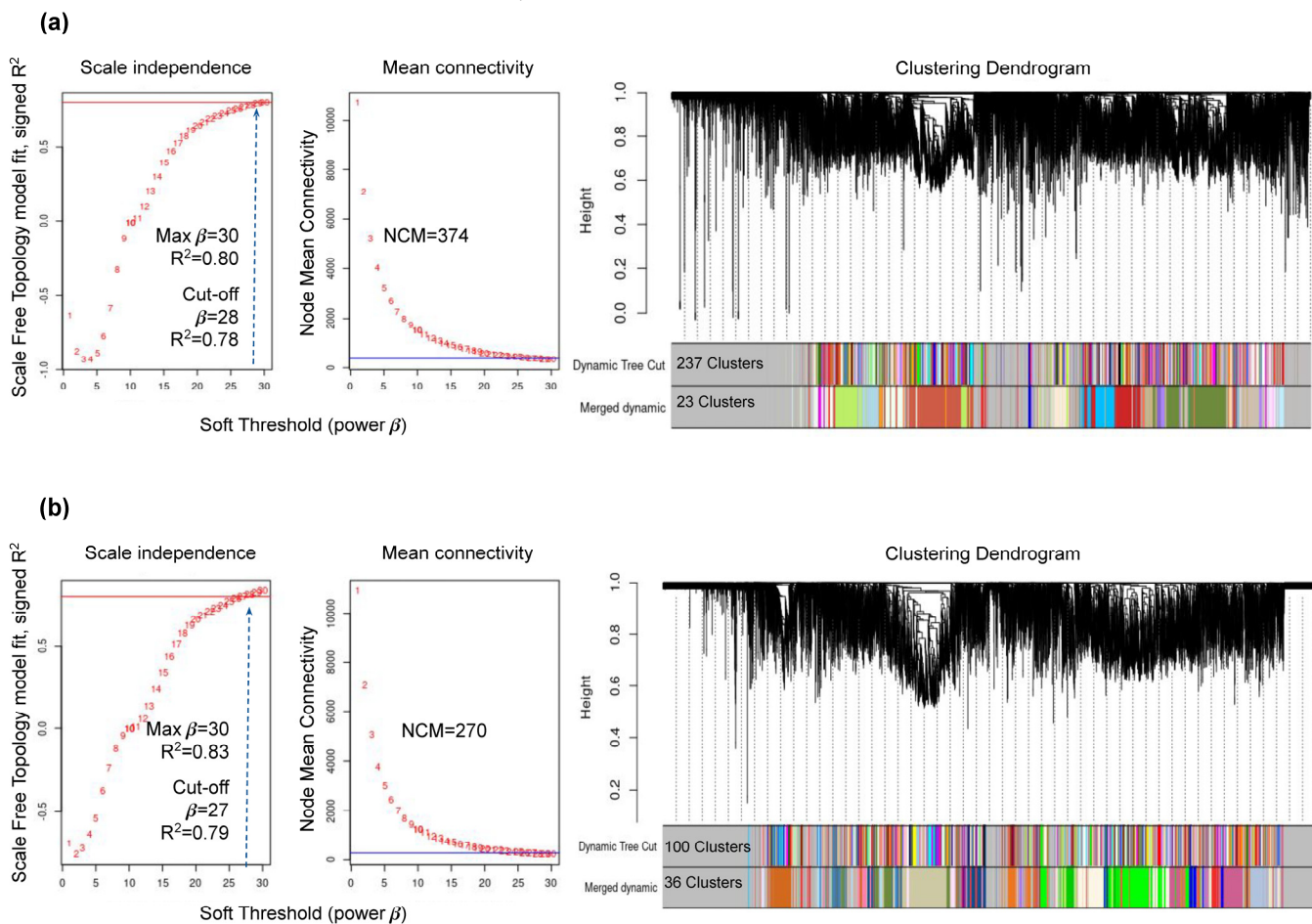


Figure 4. Co-expression network and clustering. The Scale-Free-Topology model was used to assess healthy-control samples and infected-plant gene networks. A signed network was built using Pearson's method. **(a)** Control -sample gene network reached a maximum of $R^2 = 0.80$. The soft threshold cutoff was set at $\beta = 28$ for network discovery, which resulted in $R^2 = 0.78$ and a Node Mean Connectivity (NMC) of 374 genes. The merged network comprised 23 gene clusters. **(b)** Infected-plant gene network reached a maximum of $R^2 = 0.83$. The soft threshold cutoff was $\beta = 27$, which resulted in a model fit of $R^2 = 0.79$ with an NCM of 270 genes. The merged network contained 36 gene clusters.

Table 2. TOM network results and the Node-Connectivity Mean (NCM).

Condition	R^2	# Clusters of Standard Network	NCM	# Clusters of Merged Network
Healthy-control samples	$\beta = 28$; 0.78	237	374	23
Infected-plant samples	$\beta = 27$; 0.79	100	270	36

To confirm the biological relevance of the identified networks, biological and molecular functions for each gene network were identified via a Gene Ontology (GO) overrepresentation gene test on the genetic modules with coefficients $R^2 > 0.75$. Three and four genetic modules were analyzed to identify molecular function and biological processes from the healthy-control and infected-plant gene networks, respectively (Table S3). The number of genes in each module of the healthy-control sample gene network ranged from 79 to 2024, and 34–96% of these were assigned to ontology classes. In contrast, the infected-plant gene network modules contained 72 to 818 genes, 25–30% of which were classified into ontology classes (Table S4).

The results for the healthy-control gene network showed overrepresentation of 16 molecular functions mainly related to binding (GO:0005488) and translation (GO:0045182) in the “Coral3” module, which contains 1991 genes. The “Navajowhite3” module has 489 genes, and two biological processes related to vesicle-mediated transport (GO:0016192) and Golgi vesicle transport (GO:0048193) were overrepresented. The “Blue3” module contains 79 genes, and response to light (GO:0009416) and radiation (GO:0009314) are the main functions associated with this module (Figure 5a,b). In contrast the infected-plant gene network modules showed overrepresentation of 19 molecular functions in the 805-gene “Chocolate” module, in which protein binding (GO: 0005515) and kinase activity (GO: 0004672) are the main molecular functions. The “Chocolate2” and “Green3” modules each had only one molecular function class. The molecular function for the “Chocolate2” module is related to RNA binding (GO:0003723), while the “Green3” module is associated with unfolded protein binding (GO:0051082) (Figure 5c, Data S3).

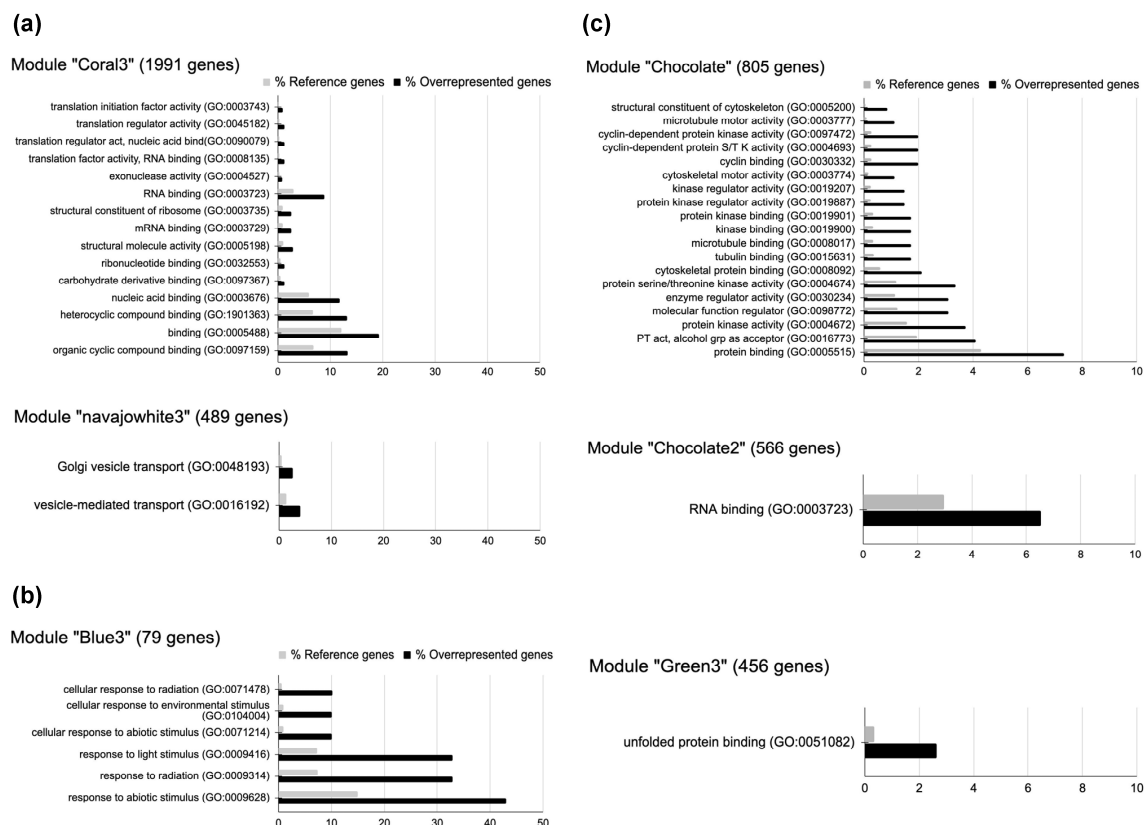


Figure 5. Gene Ontology (GO) overrepresentation test. (a) Bar plots of overrepresented molecular function GO-slim classes in Coral3 gene module and overrepresented biological process GO-slim classes in Navajowhite3 gene module obtained for control sample gene network. (b) Bar plot of overrepresented biological process GO-slim classes in Blue3 gene module obtained for control sample gene network. (c) Bar plots of overrepresented molecular function GO-slim classes in Chocolate, Chocolate2 and Green3 gene modules obtained for infected-plant gene expression network.

3.3. Identification and Functional Annotation of Core Defense Genes in *A. thaliana*

To identify defense core genes related to multi-fungal infections, healthy-control and infected-plant networks modules were compared, and modules in the infected-plant gene network with at least 75% unique or different genes were identified (Figure S3). Finally, modules that showed a positive correlation ($R^2 > 0.50$) were selected. It was determined that the “Darkmagenta” module fit our criteria, with a correlation coefficient of $R^2 = 0.75$ for *B. cinerea* at 24 hpi and a $R^2 = 0.59$ for *C. higginsianum* at 22 hpi (Figure 6). It contains 113 genes, and of these, 103 (77.4%) were exclusively in this module (Figure S3). The functional annotation of this module was performed using the Database for Annotation, Visualization and Integrated Discovery (DAVID, <https://david.ncicrf.gov/home.jsp>, accessed on 28 September 2023), which helps to identify closely related gene groups [50,66].

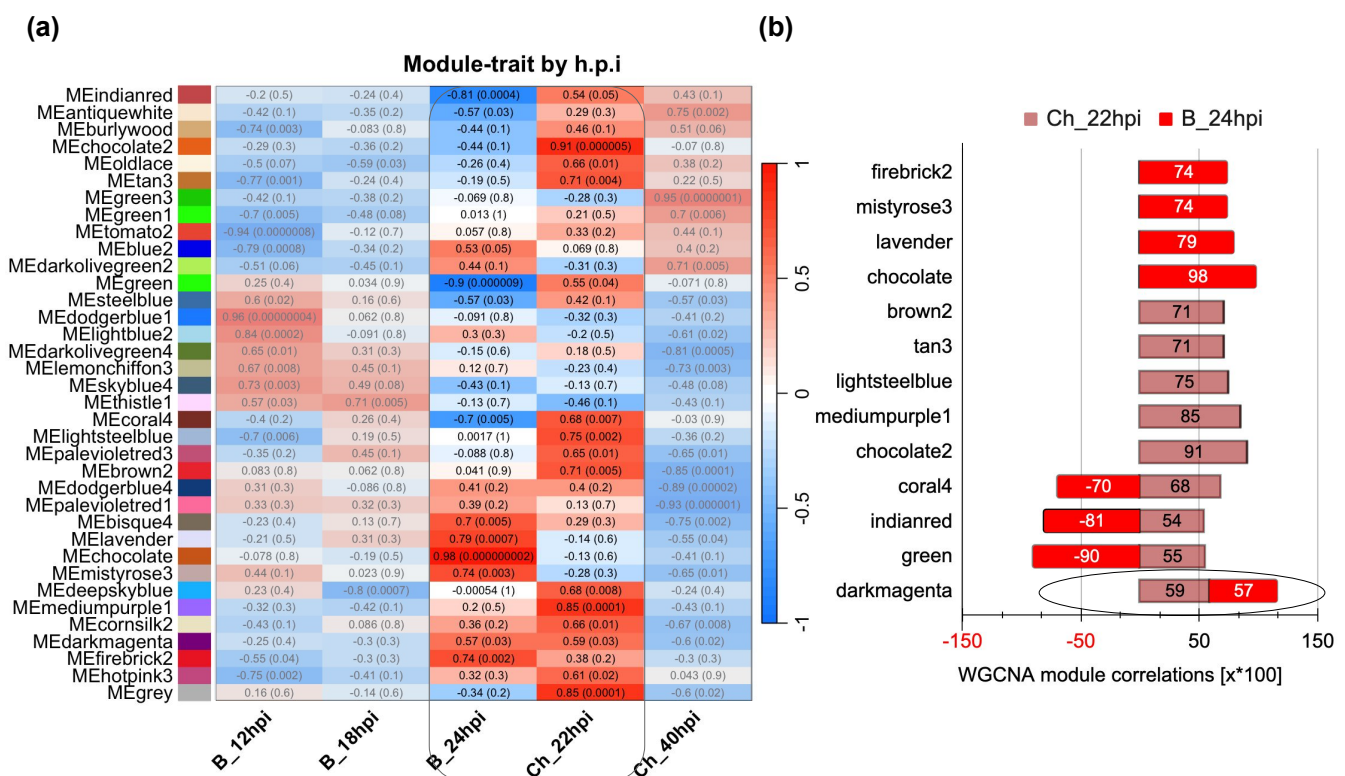


Figure 6. Core gene module. (a) The heatmap shows the different treatments analyzed in the study. The rectangle highlights the treatments with significant correlation. B_24 hpi represents the sample with *B. cinerea* at 24 h pos infection. Ch_22 hpi corresponds to the sample with *C. higginsianum* at 22 h postinfection. (b) Selected modules are displayed. The ellipses represent the consensus module “Darkmagenta”, with a $R^2 > 0.59$ for *C. higginsianum* and a $R^2 > 0.57$ for *B. cinerea*.

The 113 genes in the “Darkmagenta” module were mapped to nine DAVID functional categories. To categorize highly related genes into functional groups relevant to the pathogenic fungal infection, the DAVID gene functional annotation tool [66] employs a $kappa$ distribution score, an agglomeration method and an EASE enrichment score. The web tool implementation enables users to select different stringency levels. We identified gene groups that remained consistent at medium and high stringency. Thus, 14 enriched functional groups were obtained and then filtered using an EASE score > 0.80 and p -value < 0.05 , resulting in six enriched functional groups (Tables 3 and S5). Of these six groups, seven genes were classified in the “WD40/YVTN repeat-like-containing” group, three genes in the “sterol metabolism” group, ten genes in the “glycosyltransferase” group, seven genes in the “intracellular protein transport” group, four genes in “Zinc finger, FYVE/PHD-type”, and six genes in “Methyltransferase” (Table 3). Four of these thirty-seven genes have multiple

functions (AT5G06050, AT5G13960, AT5G40870 and AAT5G45660) and are associated with at least two functional groups (Table 4). Finally, 33 defense-related genes and their roles in *B. cinerea* and *C. higginsianum* infection in *A. thaliana* at 22 and 24 hpi were identified through DAVID functional annotation (Table 4).

Table 3. Enriched functional groups from “Darkmagenta” gene module.

Category	Term (Group #)	p Value	P Benjamin	#Genes	Gene	E. Score
IPR015943	WD40/YVTN repeat-like-containing (group 1)	0.0029	0.5307	7	AT5G23430, AT3G13340, AT3G18060, AT5G51980, VPS11, AT1G79990, VCS	1.5527
KW-1207 KW-0752 GO:0016126	Sterol metabolism (group 2)	0.0125 0.0275 0.0277	0.4613 0.7459 0.9995	3	SMT1, SMO1-1, 3BETAHSD/D1	1.3511
GO:0016757 KW-0328	Glycosyltransferase (group 3)	0.0113 0.0346	0.8358 0.7467	10	ALG3, FUT13, CALS1, UK/UPRT1, AT4G38040, GUT2, PARVUS, SETH2, AT5G45660, AT1G34270	1.3489
KW-0968 GO:0006886	Intracellular protein transport (group 4)	0.0066 0.0428	0.1592 0.9999	7	PLA2- α , PAT2, AT1G60070, AT4G13730, AT1G14910, VPS11, AT1G79990	1.1112
IPR011011 IPR019787	Zinc finger, FYVE/PHD-type (group 5)	0.0190 0.0477	0.9928 0.9999	4	EMB1135, ATX2, AT5G12350, AT1G50620	1.0173
KW-0489	Methyltransferase (group 6)	0.0403	0.7991	6	AT3G15530, SUVH4, AT2G34300, SMT1, ATX2, AT5G06050	0.8795

Table 4. Description of core genes involved in multiple fungal infections.

Group	TAIR ID	Gene	Chr	Description
1	AT5G23430	AT5G23430	Chr5	Transducin/WD40 repeat-like superfamily protein.
1	AT3G13340	AT3G13340	Chr3	Transducin/WD40 repeat-like superfamily protein
1	AT3G18060	AT3G18060	Chr3	Transducin/WD-40 repeat family protein
1	AT5G51980	AT5G51980	Chr5	Transducin/WD40 repeat-like superfamily protein
1–4	AT1G79990	AT1G79990	Chr1	Coatmer subunit β -2. WD repeat COPB2 family
1–4	AT2G05170	VPS11	Chr2	Vacuolar protein sorting 11
1	AT3G13300	VCS	Chr3	VARICOSE. Transducin/WD40 repeat-like superfamily protein
2–6	AT5G13710	SMT1	Chr5	Sterol methyltransferase 1
2	AT4G12110	SMO1-1	Chr4	Sterol-4 α -methyl oxidase 1-1
2	AT1G47290	3BETAHSD/D1	Chr1	3 β -hydroxysteroid-dehydrogenase/decarboxylase isoform 1
3	AT2G47760	ALG3	Chr2	Asparagine-linked glycosylation 3
3	AT1G71990	FUT13	Chr1	Fucosyltransferase 13
3	AT1G05570	CALS1	Chr1	Callose synthase 1
3	AT5G40870	UK/UPRT1	Chr5	Uridine kinase/uracil phosphoribosyltransferase 1
3	AT1G27440	GUT2	Chr1	Exostosin family protein
3	AT5G12350	AT5G12350	Chr5	Exostosin family protein(AT4G38040)
3	AT1G05570	CALS1	Chr1	Uridine kinase/uracil phosphoribosyltransferase 1(UK/UPRT1)
3	AT1G19300	PARVUS	Chr1	Nucleotide-diphospho-sugar transferases superfamily
3	AT3G45100	SETH2	Chr3	UDP-Glycosyltransferase superfamily protein
3	AT4G38040	AT4G38040	Chr4	Exostosin family protein. Glycosyltransferase 47 family
3	AT5G45660	AT5G45660	Chr5	Adenine phosphoribosyltransferase
3	AT1G34270	AT1G34270	Chr1	Exostosin family protein

Table 4. Cont.

Group	TAIR ID	Gene	Chr	Description
4	AT2G06925	PLA2- α	Chr2	Phospholipase A2 family protein
4	AT3G55480	PAT2	Chr3	Protein affected trafficking 2
4	AT1G60070	AT1G60070	Chr1	Adaptor protein complex AP-1
4	AT4G13730	AT4G13730	Chr4	Ypt/Rab-GAP domain of gyp1p superfamily protein
4	AT1G14910	AT1G14910	Chr1	ENTH/ANTH/VHS superfamily protein
5	AT1G79350	EMB1135	Chr1	RING/FYVE/PHD zinc finger superfamily protein
5–6	AT1G05830	ATX2	Chr1	Trithorax-like protein 2
5	AT5G12350	AT5G12350	Chr5	RCC1 family with FYVE zinc finger domain-containing protein
5	AT1G50620	AT1G50620	Chr1	RING/FYVE/PHD zinc finger superfamily protein
6	AT5G13960	SUVH4	Chr5	Histone-lysine N-methyltransferase, H3 lysine-9
6	AT3G15530	AT3G15530	Chr3	S-adenosyl-L-methionine-dependent methyltransferases superfamily
6	AT2G34300	AT2G34300	Chr2	S-adenosyl-L-methionine-dependent methyltransferases superfamily
6	AT5G06050	AT5G06050	Chr5	Putative methyltransferase family protein

4. Discussion

Pathogenic fungi-induced diseases are among the main causes of losses in commercial crops. In light of major challenges such as climate change and fungicide resistance, it is imperative to develop global strategies to tackle these losses. Conducting experiments under a wide range of environmental stresses can be both challenging and time consuming. However, the advent of NGS technologies has allowed the collection of large amounts of data that could help to address the challenges associated with modern agriculture. The utilization of publicly available RNA-seq expression data holds great potential for further study of complex plant–pathogen interactions, leading to the identification of candidate genes and molecular markers and optimization of breeding strategies by targeting specific genes. However, utilizing publicly available RNA-seq data remains challenging. To tackle this, we based our study on computational approaches using public data from RNA-seq transcriptome studies of *A. thaliana*–*B. cinerea* and *A. thaliana*–*C. higginsianum* pathosystems to identify core genes that respond to multiple types of fungal infection using GCN. Reanalysis of public RNA-seq expression data requires the minimization of technical biases, known as the batch effect, to obtain high-quality co-expression networks [30]. Although several tools may be used for batch-effect correction (such as ComBat-Seq [70], ComBat [71], svaseq [72]), further efforts may be required, particularly when sample sizes are limited (<30) [30]. Therefore, we developed a straightforward workflow to preprocess data in order to correct the batch effect that allowed us to recover over 80% of genes for WGCNA (Figure 1).

Previous studies revealed that gene modules generated by WGCNA from condition-dependent gene expression experiments are more informative than gene modules identified by combining the entire dataset regardless of condition [48]. Thus, two gene co-expression networks were constructed using filtered data from healthy (control) and infected *A. thaliana* plants with *B. cinerea* and *C. higginsianum*. The network for healthy-control plants included 23 gene modules with an NCM of 374, while for the network for infected plants, 36 gene modules with an NCM of 270 were obtained (Figure 4, Table 2). Additionally, the quality of each network was confirmed through GO overrepresentation analysis, resulting in biological processes and functions that are overrepresented in each network and linked to each phenotype. Modules with an R^2 values greater than 0.75 were selected for both networks, and it was determined that within the healthy-control plant network, genetic modules coral3, blue3 and navajowhite3, modules with positive correlation (R^2 of 0.98, 0.96, and 0.79, respectively) were related to plant development and growth, exhibiting functions such as translation and protein binding (GO:0005488), vesicle-mediated protein transport (GO:0016192), as well as response to light and radioactivity (GO:0016192) (Figure 5, Table S2). The gene modules from the plant-infected network, chocolate, dodgerblue1,

green and chocolate2 with positive correlations (R^2 of 0.98, 0.96, 0.95, 0.91, respectively), displayed overrepresented molecular functions of genes associated with regulating gene expression and signal transduction, which are the primary stress response mechanisms. These functions included protein binding (GO:0005515), kinase activity (GO:0004672), RNA binding (GO:000373) and unfolded protein binding (GO:0051082) (Figure 5, Table S2) [73,74].

To identify core genes that respond to *B. cinerea* and *C. higginsianum* infection in *A. thaliana*, the genetic modules in the network of infected plants should differ by at least 75% compared to the network of healthy-control plants with a positive correlation ($R^2 > 0.5$) for the infection conditions. Therefore, we selected the “Darkmagenta” module due to its differentiation in 77.4% of genes compared to the control. This module displayed an $R^2 = 0.57$ with the *B. cinerea* treatment at 24 hpi and an $R^2 = 0.59$ with the *C. higginsianum* treatment at 22 hpi, as shown in the heatmap (Figure 6). This module contained 113 genes that were functionally annotated using the DAVID tool [66], revealing 33 genes distributed across six functional categories (Tables 3, 4 and S5).

The first group of genes in the “Darkmagenta” module consisted of seven genes (AT5G23430, AT3G13340, AT3G18060, AT5G51980, AT1G79990, VPS and VSC) that contained the WD40/YVTN repeat-like domain (IPR015943). In *A. thaliana*, 237 proteins have been reported to contain four or more copies of the WD40 domain and participate in diverse biological processes, including defense against pathogens [75,76]. The most extensively studied of the WDR proteins are the G β proteins, which are associated with type-1 membrane receptors in the plant innate immunity signaling pathway. Interactions between G β and type-1 membrane receptors convert extracellular signals into intracellular chemical defense responses, including reactive oxygen species (ROS) production, callose deposition, MAPK activation, defense gene activation and programmed cell death [77]. In this group, we identified at least 3 G β -type genes (AT5G23430, AT5G51980, AT1G79990) (Table 4).

In the second group, we identified three CPG for key enzymes in sterol biosynthesis, SMT1, SMO1-1, and 3BETAHSD/D1 (Table 4). Sterols are a complex mixture of organic compounds that are synthesized in plants and function as structural components of cell membranes. They play a vital role in processes such as channel regulation, protein trafficking, signal transduction and plant–pathogen interactions [78]. The SMT1, 4-methyl sterol oxidase (SMO) and 3-BETAHSD enzymes participate in the biosynthesis of c24-alkyl sterol and cholesterol in plants [79]. The role of sterols in plant defense has been studied more comprehensively in relation to abiotic stressors, such as drought, salinity, and cold, than against biotic stress. However, certain sterols such as cholesterol activate biotic stress response genes [78], while pathogenic bacteria and ROS stimulate biosynthesis of stigmasterol. Moreover, PATHOGENESIS-RELATED PROTEIN 1 (PR-1) can bind various sterols, including stigmasterol, to inhibit pathogen growth by sequestering sterols from pathogens [80]. Although there is evidence suggesting that sterols might function as a crucial element in plant defense, further studies are needed to demonstrate their efficacy.

The glycosyltransferase genes in group three consist of 10 CPG (ALG3, FUT13, CALS1, UK/UPRT1, AT4G38040, GUT2, PARVUS, SETH2, AT5G45660, AT1G34270). Glycosylation is one of the major posttranslational modifications in plants, playing an important role in defense against pathogens by inactivating toxic microbial compounds and strengthening the cell wall. Glycosylation also regulates secondary metabolite levels by increasing their water solubility to facilitate metabolite transport throughout the plant, increasing the metabolites’ stability to avoid degradation and even altering their biological activity to increase toxicity to pathogens or decrease bioactivity for plants. An example of this is the glycosylation of salicylic acid (SA), a phytohormone that functions as a signal molecule during plant defense against pathogens, activating systemic acquired resistance (SAR). This glycosylation is believed to facilitate detoxification of SA, allowing it to be safely stored and moved within the plant, which helps to avoid excessive activation of defense responses that could harm the plant’s growth and development [81–83]. Interestingly, within this group of genes, we have identified CALS1 (AT1G05570), which responds to SA (Table 4).

Four genes were identified as Zinc FYVE/PHD group (EMB1135, ATX2, AT5G12350, and AT1G5062). The zinc finger protein (ZFP) transcription factors (TFs) constitute a vast family in the plant kingdom and play a crucial role in stress tolerance. The ZFP TFs mainly regulate genes involved in antioxidation activity, which reduce ROS accumulation. Controlling the ROS level is critical, as it helps to avoid damage in molecules, including DNA, proteins and lipids, and to enhance stress tolerance. AT1G50620 and EMB1135 (alternatively known as the FGT 1 gene, AT1G79350), containing the RING/FYVE/PHD-type (RFP) domain (IPR011011) [84,85], have a zinc ion binding function, and recent studies indicate that this function may help with ROS scavenging [85,86]. Additionally, rapid accumulation of ROS during the development of fungal appressoria has been observed as a defense mechanism to prevent penetration of the pathogenic fungi [87,88]. This is consistent with the infection period of *C. higginsianum*, since at 22 h post inoculation, the fungus is in the prepenetration stage and is engaged in forming the appressorium [9]. Additionally, FGT 1 (EMB1135) encodes the FORGETTER 1 protein that binds directly to a specific class of heat-inducible genes through nucleosome remodeling [89]. These types of ZFP TFs regulate cellular responsiveness by binding to promoter regions of actively expressed genes in a heat-dependent fashion, such as heat shock proteins (HSP). The HSPs are induced by a variety of stresses, including oxidative stress; however, the crosstalk between abiotic and biotic stress responses in the ROS network remains poorly understood [90].

A group of six methyltransferases were identified (AT3G15530, SUVH4, AT2G34300, SMT1, ATX2, AT5G06050); all of these genes are involved in methylation. However, to date, there have been few studies on the role of methylation in plant–pathogen interactions, particularly with regard to infections caused by pathogenic fungi. The available evidence suggests that epigenetic control of gene expression contributes to the fine-tuning of plant defenses in response to pathogens [91,92]. Crespo-Salvador et al. (2018) demonstrated that differentially expressed genes responding to *B. cinerea* infection were induced 24 h after infection and displayed similar histone modification patterns, suggesting that epigenetics marks may impact the transcriptional regulation in this particular pathosystem [93]. *A. thaliana* hypomethylated mutants showed resistance against the biotrophic pathogen *Hyaloperonospora arabidopsidis*, whereas hypermethylated mutants were more susceptible to this pathogen [92,94]. Additionally, research on geminivirus infections has shown that the SUVH4 protein is a virus-silencing target that aims to evade host defenses [95]. Interestingly, our results highlighted the presence of SUVH4 (AT5G13960) and ATX2 (AT1G05830), which are involved in the maintenance of epigenetic control [96].

5. Conclusions

Core gene clusters that respond to multiple fungal infections were identified. However, further studies are needed to elucidate the role of several important genes in plant immune responses. Identifying coexpressed gene associations may provide insights into the functions of currently unknown coding genes. In the context of climate change, genomic data repositories are important resources for developing new tools for crop improvement and disease prevention, positively influencing productivity, sustainability and food security.

Supplementary Materials: The following supporting information can be downloaded at <https://www.mdpi.com/article/10.3390/genes14122223/s1>: Figure S1: Venn diagrams for nonexpressed genes; Figure S2: Distribution of atypical samples in infected-plant gene networks; Figure S3: Unique genes identified in infected-plant gene modules; Table S1: Sample information obtained from NCBI SRA; Table S2: Databases, Bioinformatic tools and source code; Table S3: Significant gene modules identified in healthy-control and fungal-infected samples; Table S4: Gene Ontology overrepresentation test summary; Table S5: Genes functional groups related to multiple fungal infections identified by DAVID; Data S1: Normalized expression data for control samples; Data S2: Normalized expression for infected-plant samples; and Data S3: Gene Ontology overrepresentation analysis results.

Author Contributions: Conceptualization, E.G.-C.; methodology, software, validation, data curation, and visualization, C.S.-C.; formal analysis, C.S.-C. and E.G.-C.; investigation, C.S.-C. and E.G.-C.; resources, E.G.-C. and K.L.C.; writing—original draft preparation, E.G.-C. and C.S.-C.; writing—review and editing, E.G.-C. and K.L.C.; supervision, E.G.-C.; project administration, E.G.-C.; funding acquisition, E.G.-C. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by Consejo Nacional de Humanidades, Ciencia y Tecnología (CONAHCYT), Mexico. Grant number INFR201601-269833, and scholarship 552999.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The code developed for this study is freely available via GitHub at <https://github.com/cyntsc/RNA-Seq-raw-integration> (accessed on 28 September 2023) (DOI 10.5281/zenodo.7076416).

Acknowledgments: The authors thank Adriana Quiroz-Moreno for providing technical advice and facilities throughout the project.

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

References

1. Avery, S.V.; Singleton, I.; Magan, N.; Goldman, G.H. The Fungal Threat to Global Food Security. *Fungal Biol.* **2019**, *123*, 555–557. [CrossRef] [PubMed]
2. Kettles, G.J.; Luna, E. Food Security in 2044: How Do We Control the Fungal Threat? *Fungal Biol.* **2019**, *123*, 558–564. [CrossRef] [PubMed]
3. Steinberg, G.; Gurr, S.J. Fungi, Fungicide Discovery and Global Food Security. *Fungal Genet. Biol.* **2020**, *144*, 103476. [CrossRef] [PubMed]
4. Miedaner, T.; Juroszek, P. Climate Change Will Influence Disease Resistance Breeding in Wheat in Northwestern Europe. *Theor. Appl. Genet.* **2021**, *134*, 1771–1785. [CrossRef] [PubMed]
5. Fisher, M.C.; Hawkins, N.J.; Sanglard, D.; Gurr, S.J. Worldwide Emergence of Resistance to Antifungal Drugs Challenges Human Health and Food Security. *Science* **2018**, *360*, 739–742. [CrossRef] [PubMed]
6. Almeida, F.; Rodrigues, M.L.; Coelho, C. The Still Underestimated Problem of Fungal Diseases Worldwide. *Front. Microbiol.* **2019**, *10*, 214. [CrossRef] [PubMed]
7. Chaloner, T.M.; Gurr, S.J.; Bebber, D.P. Plant Pathogen Infection Risk Tracks Global Crop Yields under Climate Change. *Nat. Clim. Change* **2021**, *11*, 710–715. [CrossRef]
8. Bregaglio, S.; Donatelli, M.; Confalonieri, R. Fungal Infections of Rice, Wheat, and Grape in Europe in 2030–2050. *Agron. Sustain. Dev.* **2013**, *33*, 767–776. [CrossRef]
9. O’Connell, R.J.; Thon, M.R.; Hacquard, S.; Amyotte, S.G.; Kleemann, J.; Torres, M.F.; Damm, U.; Buiate, E.A.; Epstein, L.; Alkan, N.; et al. Lifestyle Transitions in Plant Pathogenic *Colletotrichum* Fungi Deciphered by Genome and Transcriptome Analyses. *Nat. Genet.* **2012**, *44*, 1060–1065. [CrossRef]
10. Mendgen, K.; Hahn, M. Plant Infection and the Establishment of Fungal Biotrophy. *Trends Plant Sci.* **2002**, *7*, 352–356. [CrossRef]
11. Laluk, K.; Mengiste, T. Necrotroph Attacks on Plants: Wanton Destruction or Covert Extortion? *Arab. Book* **2010**, *8*, e0136. [CrossRef] [PubMed]
12. Fatima, U.; Senthil-Kumar, M. Plant and Pathogen Nutrient Acquisition Strategies. *Front. Plant Sci.* **2015**, *6*, 750. [CrossRef] [PubMed]
13. Dean, R.; Van Kan, J.A.L.; Pretorius, Z.A.; Hammond-Kosack, K.E.; Di Pietro, A.; Spanu, P.D.; Rudd, J.J.; Dickman, M.; Kahmann, R.; Ellis, J.; et al. The Top 10 Fungal Pathogens in Molecular Plant Pathology. *Mol. Plant Pathol.* **2012**, *13*, 414–430. [CrossRef] [PubMed]
14. Yan, Y.; Yuan, Q.; Tang, J.; Huang, J.; Hsiang, T.; Wei, Y.; Zheng, L. *Colletotrichum higginsianum* as a Model for Understanding Host–Pathogen Interactions: A Review. *Int. J. Mol. Sci.* **2018**, *19*, 2142. [CrossRef] [PubMed]
15. Muckenschnabel, I.; Goodman, B.A.; Williamson, B.; Lyon, G.D.; Deighton, N. Infection of Leaves of *Arabidopsis thaliana* by *Botrytis cinerea*: Changes in Ascorbic Acid, Free Radicals and Lipid Peroxidation Products. *J. Exp. Bot.* **2002**, *53*, 207–214. [CrossRef] [PubMed]
16. Windram, O.; Madhou, P.; McHattie, S.; Hill, C.; Hickman, R.; Cooke, E.; Jenkins, D.J.; Penfold, C.A.; Baxter, L.; Breeze, E.; et al. *Arabidopsis* Defense against *Botrytis cinerea*: Chronology and Regulation Deciphered by High-Resolution Temporal Transcriptomic Analysis. *Plant Cell* **2012**, *24*, 3530–3557. [CrossRef]

17. Veillet, F.; Gaillard, C.; Lemonnier, P.; Coutos-Thévenot, P.; La Camera, S. The Molecular Dialogue between *Arabidopsis thaliana* and the Necrotrophic Fungus *Botrytis cinerea* Leads to Major Changes in Host Carbon Metabolism. *Sci. Rep.* **2017**, *7*, 17121. [CrossRef]
18. Jones, J.D.G.; Dangl, J.L. The Plant Immune System. *Nature* **2006**, *444*, 323–329. [CrossRef]
19. Zipfel, C.; Robatzek, S.; Navarro, L.; Oakeley, E.J.; Jones, J.D.G.; Felix, G.; Boller, T. Bacterial Disease Resistance in *Arabidopsis* through Flagellin Perception. *Nature* **2004**, *428*, 764–767. [CrossRef]
20. Navarro, L.; Jay, F.; Nomura, K.; He, S.Y.; Voinnet, O. Suppression of the MicroRNA Pathway by Bacterial Effector Proteins. *Science* **2008**, *321*, 964–967. [CrossRef]
21. Vleeshouwers, V.G.A.A.; Oliver, R.P. Effectors as Tools in Disease Resistance Breeding Against Biotrophic, Hemibiotrophic, and Necrotrophic Plant Pathogens. *MPMI* **2014**, *27*, 196–206. [CrossRef] [PubMed]
22. Landete, J.M. Effector Molecules and Regulatory Proteins: Applications. *Trends Biotechnol.* **2016**, *34*, 777–780. [CrossRef] [PubMed]
23. Khan, M.; Seto, D.; Subramaniam, R.; Desveaux, D. Oh, the Places They'll Go! A Survey of Phytopathogen Effectors and Their Host Targets. *Plant J.* **2018**, *93*, 651–663. [CrossRef] [PubMed]
24. Koeck, M.; Hardham, A.R.; Dodds, P.N. The Role of Effectors of Biotrophic and Hemibiotrophic Fungi in Infection: Effectors of Biotrophic Fungi. *Cell. Microbiol.* **2011**, *13*, 1849–1857. [CrossRef] [PubMed]
25. Bigeard, J.; Colcombet, J.; Hirt, H. Signaling Mechanisms in Pattern-Triggered Immunity (PTI). *Mol. Plant* **2015**, *8*, 521–539. [CrossRef]
26. Thomma, B.P.H.J.; Nürnberger, T.; Joosten, M.H.A.J. Of PAMPs and Effectors: The Blurred PTI-ETI Dichotomy. *Plant Cell* **2011**, *23*, 4–15. [CrossRef]
27. Tsuda, K.; Katagiri, F. Comparing Signaling Mechanisms Engaged in Pattern-Triggered and Effector-Triggered Immunity. *Curr. Opin. Plant Biol.* **2010**, *13*, 459–465. [CrossRef]
28. Dodds, P.N.; Rathjen, J.P. Plant Immunity: Towards an Integrated View of Plant–Pathogen Interactions. *Nat. Rev. Genet.* **2010**, *11*, 539–548. [CrossRef]
29. GenBank and WGS Statistics. Available online: <https://www.ncbi.nlm.nih.gov/genbank/statistics/> (accessed on 30 August 2022).
30. Vandenbon, A. Evaluation of Critical Data Processing Steps for Reliable Prediction of Gene Co-Expression from Large Collections of RNA-Seq Data. *PLoS ONE* **2022**, *17*, e0263344. [CrossRef]
31. Crandall, S.G.; Gold, K.M.; del Mar Jiménez-Gasco, M.; Filgueiras, C.C.; Willett, D.S. A Multi-Omics Approach to Solving Problems in Plant Disease Ecology. *PLoS ONE* **2020**, *15*, e0237975. [CrossRef]
32. Subramanian, I.; Verma, S.; Kumar, S.; Jere, A.; Anamika, K. Multi-Omics Data Integration, Interpretation, and Its Application. *Bioinform. Biol. Insights* **2020**, *14*, 1177932219899051. [CrossRef] [PubMed]
33. Bhadauria, V.; Vijayan, P.; Wei, Y.; Banniza, S. Transcriptome Analysis Reveals a Complex Interplay between Resistance and Effector Genes during the Compatible Lentil-Colletotrichum Lentis Interaction. *Sci. Rep.* **2017**, *7*, 42338. [CrossRef] [PubMed]
34. Zhang, L.; Huang, X.; He, C.; Zhang, Q.-Y.; Zou, X.; Duan, K.; Gao, Q. Novel Fungal Pathogenicity and Leaf Defense Strategies Are Revealed by Simultaneous Transcriptome Analysis of Colletotrichum Fructicola and Strawberry Infected by This Fungus. *Front. Plant Sci.* **2018**, *9*, 434. [CrossRef] [PubMed]
35. Pérez-Torres, C.-A.; Ibarra-Laclette, E.; Hernández-Domínguez, E.-E.; Rodríguez-Haas, B.; Pérez-Lira, A.-J.; Villafán, E.; Alonso-Sánchez, A.; de Jesus García-Ávila, C.; Ramírez-Pool, J.-A.; Sánchez-Rangel, D. Molecular Evidence of the Avocado Defense Response to *Fusarium Kuroshium* Infection: A Deep Transcriptome Analysis Using RNA-Seq. *PeerJ* **2021**, *9*, e11215. [CrossRef] [PubMed]
36. Xu, X.-H.; Wang, C.; Li, S.-X.; Su, Z.-Z.; Zhou, H.-N.; Mao, L.-J.; Feng, X.-X.; Liu, P.-P.; Chen, X.; Hugh Snyder, J.; et al. Friend or Foe: Differential Responses of Rice to Invasion by Mutualistic or Pathogenic Fungi Revealed by RNAseq and Metabolite Profiling. *Sci. Rep.* **2015**, *5*, 13624. [CrossRef] [PubMed]
37. Takahara, H.; Dolf, A.; Endl, E.; O'Connell, R. Flow Cytometric Purification of *Colletotrichum higginsianum* Biotrophic Hyphae from *Arabidopsis* Leaves for Stage-specific Transcriptome Analysis. *Plant J.* **2009**, *59*, 672–683. [CrossRef]
38. Mulema, J.M.K.; Denby, K.J. Spatial and Temporal Transcriptomic Analysis of the *Arabidopsis thaliana*–*Botrytis cinerea* Interaction. *Mol. Biol. Rep.* **2012**, *39*, 4039–4049. [CrossRef]
39. La Camera, S.; L'Haridon, F.; Astier, J.; Zander, M.; Abou-Mansour, E.; Page, G.; Thurow, C.; Wendehenne, D.; Gatz, C.; Métraux, J.-P.; et al. The Glutaredoxin ATGRXS13 Is Required to Facilitate *Botrytis cinerea* Infection of *Arabidopsis thaliana* Plants: Role of ATGRXS13 during *B. cinerea* Infection. *Plant J.* **2011**, *68*, 507–519. [CrossRef]
40. Narusaka, Y.; Narusaka, M.; Park, P.; Kubo, Y.; Hirayama, T.; Seki, M.; Shiraishi, T.; Ishida, J.; Nakashima, M.; Enju, A.; et al. *RCH1*, a Locus in *Arabidopsis* That Confers Resistance to the Hemibiotrophic Fungal Pathogen *Colletotrichum higginsianum*. *MPMI* **2004**, *17*, 749–762. [CrossRef]
41. Takahara, H.; Hacquard, S.; Kombrink, A.; Hughes, H.B.; Halder, V.; Robin, G.P.; Hiruma, K.; Neumann, U.; Shinya, T.; Kombrink, E.; et al. *Colletotrichum higginsianum* Extracellular LysM Proteins Play Dual Roles in Appressorial Function and Suppression of Chitin-triggered Plant Immunity. *New Phytol.* **2016**, *211*, 1323–1337. [CrossRef]

42. Amrine, K.C.H.; Blanco-Ulate, B.; Cantu, D. Discovery of Core Biotic Stress Responsive Genes in Arabidopsis by Weighted Gene Co-Expression Network Analysis. *PLoS ONE* **2015**, *10*, e0118731. [[CrossRef](#)] [[PubMed](#)]
43. Kumar, A.; Kanak, K.R.; Arunachalam, A.; Dass, R.S.; Lakshmi, P.T.V. Comparative Transcriptome Profiling and Weighted Gene Co-Expression Network Analysis to Identify Core Genes in Maize (*Zea mays* L.) Silks Infected by Multiple Fungi. *Front. Plant Sci.* **2022**, *13*, 985396. [[CrossRef](#)] [[PubMed](#)]
44. Langfelder, P.; Horvath, S. WGCNA: An R Package for Weighted Correlation Network Analysis. *BMC Bioinform.* **2008**, *9*, 559. [[CrossRef](#)] [[PubMed](#)]
45. Serin, E.A.R.; Nijveen, H.; Hilhorst, H.W.M.; Ligterink, W. Learning from Co-Expression Networks: Possibilities and Challenges. *Front. Plant Sci.* **2016**, *7*, 444. [[CrossRef](#)] [[PubMed](#)]
46. Han, X.; Zhang, Y.-W.; Liu, J.-Y.; Zuo, J.-F.; Zhang, Z.-C.; Guo, L.; Zhang, Y.-M. 4D Genetic Networks Reveal the Genetic Basis of Metabolites and Seed Oil-Related Traits in 398 Soybean RILs. *Biotechnol. Biofuels* **2022**, *15*, 92. [[CrossRef](#)] [[PubMed](#)]
47. Sari, E.; Cabral, A.L.; Polley, B.; Tan, Y.; Hsueh, E.; Konkin, D.J.; Knox, R.E.; Ruan, Y.; Fobert, P.R. Weighted Gene Co-Expression Network Analysis Unveils Gene Networks Associated with the Fusarium Head Blight Resistance in Tetraploid Wheat. *BMC Genom.* **2019**, *20*, 925. [[CrossRef](#)] [[PubMed](#)]
48. Childs, K.L.; Davidson, R.M.; Buell, C.R. Gene Coexpression Network Analysis as a Source of Functional Annotation for Rice Genes. *PLoS ONE* **2011**, *6*, e22196. [[CrossRef](#)]
49. Waskom, M. Seaborn: Statistical Data Visualization. *JOSS* **2021**, *6*, 3021. [[CrossRef](#)]
50. Sherman, B.T.; Hao, M.; Qiu, J.; Jiao, X.; Baseler, M.W.; Lane, H.C.; Imamichi, T.; Chang, W. DAVID: A Web Server for Functional Enrichment Analysis and Functional Annotation of Gene Lists (2021 Update). *Nucleic Acids Res.* **2022**, *50*, gkac194. [[CrossRef](#)]
51. Robin, G.P.; Kleemann, J.; Neumann, U.; Cabre, L.; Dallery, J.-F.; Lapalu, N.; O'Connell, R.J. Subcellular Localization Screening of *Colletotrichum higginsianum* Effector Candidates Identifies Fungal Proteins Targeted to Plant Peroxisomes, Golgi Bodies, and Microtubules. *Front. Plant Sci.* **2018**, *9*, 562. [[CrossRef](#)]
52. Badet, T.; Voisin, D.; Mbengue, M.; Barascud, M.; Sucher, J.; Sadon, P.; Balagué, C.; Roby, D.; Raffaele, S. Parallel Evolution of the PQR Prollyl Oligo Peptidase Gene Conferring Plant Quantitative Disease Resistance. *PLOS Genet.* **2017**, *13*, e1007143. [[CrossRef](#)] [[PubMed](#)]
53. Andrews, S. FastQC. A Quality Control Tool for High Throughput Sequence Data. 2010. Available online: <https://www.bioinformatics.babraham.ac.uk/projects/fastqc/> (accessed on 8 February 2021).
54. Bolger, A.M.; Lohse, M.; Usadel, B. Trimmomatic: A Flexible Trimmer for Illumina Sequence Data. *Bioinformatics* **2014**, *30*, 2114–2120. [[CrossRef](#)] [[PubMed](#)]
55. Lamesch, P.; Berardini, T.Z.; Li, D.; Swarbreck, D.; Wilks, C.; Sasidharan, R.; Muller, R.; Dreher, K.; Alexander, D.L.; Garcia-Hernandez, M.; et al. The Arabidopsis Information Resource (TAIR): Improved Gene Annotation and New Tools. *Nucleic Acids Res.* **2012**, *40*, D1202–D1210. [[CrossRef](#)] [[PubMed](#)]
56. Dobin, A.; Davis, C.A.; Schlesinger, F.; Drenkow, J.; Zaleski, C.; Jha, S.; Batut, P.; Chaisson, M.; Gingeras, T.R. STAR: Ultrafast Universal RNA-Seq Aligner. *Bioinformatics* **2013**, *29*, 15–21. [[CrossRef](#)] [[PubMed](#)]
57. Chang, N.; Sun, Q.; Hu, J.; An, C.; Gao, A.H. Large Introns of 5 to 10 Kilo Base Pairs Can Be Spliced out in Arabidopsis. *Genes* **2017**, *8*, 200. [[CrossRef](#)]
58. Li, H.; Handsaker, B.; Wysoker, A.; Fennell, T.; Ruan, J.; Homer, N.; Marth, G.; Abecasis, G.; Durbin, R. 1000 Genome Project Data Processing Subgroup The Sequence Alignment/Map Format and SAMtools. *Bioinformatics* **2009**, *25*, 2078–2079. [[CrossRef](#)] [[PubMed](#)]
59. Anders, S.; Pyl, P.T.; Huber, W. HTSeq—A Python Framework to Work with High-Throughput Sequencing Data. *Bioinformatics* **2015**, *31*, 166–169. [[CrossRef](#)]
60. Zhao, Y.; Li, M.-C.; Konaté, M.M.; Chen, L.; Das, B.; Karlovich, C.; Williams, P.M.; Evrard, Y.A.; Doroshov, J.H.; McShane, L.M. TPM, FPKM, or Normalized Counts? A Comparative Study of Quantification Measures for the Analysis of RNA-Seq Data from the NCI Patient-Derived Models Repository. *J. Transl. Med.* **2021**, *19*, 269. [[CrossRef](#)]
61. Renesh, B. Reneshbedre/Bioinfokit: Bioinformatics Data Analysis and Visualization Toolkit | Zenodo. Available online: <https://zenodo.org/record/3965241> (accessed on 21 January 2022).
62. Mi, H.; Thomas, P. PANTHER Pathway: An Ontology-Based Pathway Database Coupled with Data Analysis Tools. *Methods Mol. Biol.* **2009**, *563*, 123–140. [[CrossRef](#)]
63. Mi, H.; Muruganujan, A.; Ebert, D.; Huang, X.; Thomas, P.D. PANTHER Version 14: More Genomes, a New PANTHER GO-Slim and Improvements in Enrichment Analysis Tools. *Nucleic Acids Res.* **2019**, *47*, D419–D426. [[CrossRef](#)]
64. The Gene Ontology Consortium. Expansion of the Gene Ontology Knowledgebase and Resources. *Nucleic Acids Res.* **2017**, *45*, D331–D338. [[CrossRef](#)] [[PubMed](#)]
65. The Gene Ontology Consortium. The Gene Ontology Resource: Enriching a Gold Mine. *Nucleic Acids Res.* **2021**, *49*, D325–D334. [[CrossRef](#)] [[PubMed](#)]
66. Huang, D.W.; Sherman, B.T.; Tan, Q.; Collins, J.R.; Alvord, W.G.; Roayaei, J.; Stephens, R.; Baseler, M.W.; Lane, H.C.; Lempicki, R.A. The DAVID Gene Functional Classification Tool: A Novel Biological Module-Centric Algorithm to Functionally Analyze Large Gene Lists. *Genome Biol.* **2007**, *8*, R183. [[CrossRef](#)] [[PubMed](#)]

67. Pasha, A.; Subramaniam, S.; Cleary, A.; Chen, X.; Berardini, T.; Farmer, A.; Town, C.; Provart, N. Araport Lives: An Updated Framework for Arabidopsis Bioinformatics. *Plant Cell* **2020**, *32*, 2683–2686. [\[CrossRef\]](#) [\[PubMed\]](#)
68. Krishnakumar, V.; Hanlon, M.R.; Contrino, S.; Ferlanti, E.S.; Karamycheva, S.; Kim, M.; Rosen, B.D.; Cheng, C.-Y.; Moreira, W.; Mock, S.A.; et al. Araport: The Arabidopsis Information Portal. *Nucleic Acids Res.* **2015**, *43*, D1003–D1009. [\[CrossRef\]](#) [\[PubMed\]](#)
69. Perkel, J.M. Why Jupyter Is Data Scientists' Computational Notebook of Choice. *Nature* **2018**, *563*, 145–146. [\[CrossRef\]](#) [\[PubMed\]](#)
70. Zhang, Y.; Parmigiani, G.; Johnson, W.E. ComBat-Seq: Batch Effect Adjustment for RNA-Seq Count Data. *NAR Genom. Bioinform.* **2020**, *2*, lqaa078. [\[CrossRef\]](#)
71. Leek, J.T.; Johnson, W.E.; Parker, H.S.; Jaffe, A.E.; Storey, J.D. The Sva Package for Removing Batch Effects and Other Unwanted Variation in High-Throughput Experiments. *Bioinformatics* **2012**, *28*, 882–883. [\[CrossRef\]](#)
72. Leek, J.T. Svaseq: Removing Batch Effects and Other Unwanted Noise from Sequencing Data. *Nucleic Acids Res.* **2014**, *42*, e161. [\[CrossRef\]](#)
73. Backlund, M.; Stein, F.; Rettel, M.; Schwarzl, T.; Perez-Perri, J.I.; Brosig, A.; Zhou, Y.; Neu-Yilik, G.; Hentze, M.W.; Kulozik, A.E. Plasticity of Nuclear and Cytoplasmic Stress Responses of RNA-Binding Proteins. *Nucleic Acids Res.* **2020**, *48*, 4725–4740. [\[CrossRef\]](#)
74. Liang, X.; Zhang, J. Regulation of Plant Responses to Biotic and Abiotic Stress by Receptor-like Cytoplasmic Kinases. *Stress Biol.* **2022**, *2*, 25. [\[CrossRef\]](#) [\[PubMed\]](#)
75. van Nocker, S.; Ludwig, P. The WD-Repeat Protein Superfamily in Arabidopsis: Conservation and Divergence in Structure and Function. *BMC Genom.* **2003**, *4*, 50. [\[CrossRef\]](#) [\[PubMed\]](#)
76. Li, Q.; Zhao, P.; Li, J.; Zhang, C.; Wang, L.; Ren, Z. Genome-Wide Analysis of the WD-Repeat Protein Family in Cucumber and Arabidopsis. *Mol. Genet. Genom.* **2014**, *289*, 103–124. [\[CrossRef\]](#) [\[PubMed\]](#)
77. Miller, J.C.; Chezem, W.R.; Clay, N.K. Ternary WD40 Repeat-Containing Protein Complexes: Evolution, Composition and Roles in Plant Immunity. *Front. Plant Sci.* **2016**, *6*, 1108. [\[CrossRef\]](#) [\[PubMed\]](#)
78. Rogowska, A.; Szakiel, A. The Role of Sterols in Plant Response to Abiotic Stress. *Phytochem. Rev.* **2020**, *19*, 1525–1538. [\[CrossRef\]](#)
79. Du, Y.; Fu, X.; Chu, Y.; Wu, P.; Liu, Y.; Ma, L.; Tian, H.; Zhu, B. Biosynthesis and the Roles of Plant Sterols in Development and Stress Responses. *Int. J. Mol. Sci.* **2022**, *23*, 2332. [\[CrossRef\]](#) [\[PubMed\]](#)
80. Aboobucker, S.I.; Suza, W.P. Why Do Plants Convert Sitosterol to Stigmasterol? *Front. Plant Sci.* **2019**, *10*, 354. [\[CrossRef\]](#)
81. Zhang, W.; Wang, S.; Yang, J.; Kang, C.; Huang, L.; Guo, L. Glycosylation of Plant Secondary Metabolites: Regulating from Chaos to Harmony. *Environ. Exp. Bot.* **2022**, *194*, 104703. [\[CrossRef\]](#)
82. Lim, E.-K.; Doucet, C.J.; Li, Y.; Elias, L.; Worrall, D.; Spencer, S.P.; Ross, J.; Bowles, D.J. The Activity of Arabidopsis Glycosyltransferases toward Salicylic Acid, 4-Hydroxybenzoic Acid, and Other Benzoates. *J. Biol. Chem.* **2002**, *277*, 586–592. [\[CrossRef\]](#)
83. Huang, X.; Zhu, G.; Liu, Q.; Chen, L.; Li, Y.; Hou, B. Modulation of Plant Salicylic Acid-Associated Immune Responses via Glycosylation of Dihydroxybenzoic Acids. *Plant Physiol.* **2018**, *176*, 3103–3119. [\[CrossRef\]](#)
84. Noman, A.; Aqeel, M.; Khalid, N.; Islam, W.; Sanaullah, T.; Anwar, M.; Khan, S.; Ye, W.; Lou, Y. Zinc Finger Protein Transcription Factors: Integrated Line of Action for Plant Antimicrobial Activity. *Microb. Pathog.* **2019**, *132*, 141–149. [\[CrossRef\]](#) [\[PubMed\]](#)
85. Zang, D.; Li, H.; Xu, H.; Zhang, W.; Zhang, Y.; Shi, X.; Wang, Y. An Arabidopsis Zinc Finger Protein Increases Abiotic Stress Tolerance by Regulating Sodium and Potassium Homeostasis, Reactive Oxygen Species Scavenging and Osmotic Potential. *Front. Plant Sci.* **2016**, *7*, 1272. [\[CrossRef\]](#) [\[PubMed\]](#)
86. Davletova, S.; Schlauch, K.; Coutu, J.; Mittler, R. The Zinc-Finger Protein Zat12 Plays a Central Role in Reactive Oxygen and Abiotic Stress Signaling in Arabidopsis. *Plant Physiol.* **2005**, *139*, 847–856. [\[CrossRef\]](#) [\[PubMed\]](#)
87. Brown, S.H.; Yarden, O.; Gollop, N.; Chen, S.; Zveibil, A.; Belasov, E.; Freeman, S. Differential Protein Expression in *Colletotrichum Acutatum*: Changes Associated with Reactive Oxygen Species and Nitrogen Starvation Implicated in Pathogenicity on Strawberry. *Mol. Plant Pathol.* **2008**, *9*, 171–190. [\[CrossRef\]](#) [\[PubMed\]](#)
88. Foley, R.C.; Kidd, B.N.; Hane, J.K.; Anderson, J.P.; Singh, K.B. Reactive Oxygen Species Play a Role in the Infection of the Necrotrophic Fungi, *Rhizoctonia Solani* in Wheat. *PLoS ONE* **2016**, *11*, e0152548. [\[CrossRef\]](#)
89. Brzezinka, K.; Altmann, S.; Czesnick, H.; Nicolas, P.; Gorka, M.; Benke, E.; Kabelitz, T.; Jähne, F.; Graf, A.; Kappel, C.; et al. Arabidopsis FORGETTER1 Mediates Stress-Induced Chromatin Memory through Nucleosome Remodeling. *eLife* **2016**, *5*, e17061. [\[CrossRef\]](#)
90. Driedonks, N.; Xu, J.; Peters, J.L.; Park, S.; Rieu, I. Multi-Level Interactions Between Heat Shock Factors, Heat Shock Proteins, and the Redox System Regulate Acclimation to Heat. *Front. Plant Sci.* **2015**, *6*, 999. [\[CrossRef\]](#)
91. Ding, B.; Wang, G.-L. Chromatin versus Pathogens: The Function of Epigenetics in Plant Immunity. *Front. Plant Sci.* **2015**, *6*, 675. [\[CrossRef\]](#)
92. Alonso, C.; Ramos-Cruz, D.; Becker, C. The Role of Plant Epigenetics in Biotic Interactions. *New Phytol.* **2019**, *221*, 731–737. [\[CrossRef\]](#)
93. Crespo-Salvador, Ó.; Escamilla-Aguilar, M.; López-Cruz, J.; López-Rodas, G.; González-Bosch, C. Determination of Histone Epigenetic Marks in Arabidopsis and Tomato Genes in the Early Response to *Botrytis cinerea*. *Plant Cell Rep.* **2018**, *37*, 153–166. [\[CrossRef\]](#)
94. López Sánchez, A.; Stassen, J.H.M.; Furci, L.; Smith, L.M.; Ton, J. The Role of DNA (de)Methylation in Immune Responsiveness of Arabidopsis. *Plant J.* **2016**, *88*, 361–374. [\[CrossRef\]](#) [\[PubMed\]](#)

-
95. Castillo-González, C.; Liu, X.; Huang, C.; Zhao, C.; Ma, Z.; Hu, T.; Sun, F.; Zhou, Y.; Zhou, X.; Wang, X.-J.; et al. Geminivirus-Encoded TrAP Suppressor Inhibits the Histone Methyltransferase SUVH4/KYP to Counter Host Defense. *eLife* **2015**, *4*, e06671. [[CrossRef](#)] [[PubMed](#)]
 96. Malagnac, F. An Arabidopsis SET Domain Protein Required for Maintenance but Not Establishment of DNA Methylation. *EMBO J.* **2002**, *21*, 6842–6852. [[CrossRef](#)] [[PubMed](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.