



Infection by *Pseudocercospora musae* leads to an early reprogramming of the *Musa paradisiaca* defense transcriptome

Sudarshana Borah^{1,2} · Dipsikha Bora² · Priyadarshini Bhorali¹

Received: 25 November 2021 / Accepted: 27 June 2022 / Published online: 16 July 2022
© King Abdulaziz City for Science and Technology 2022

Abstract

Deep sequencing technologies such as RNA sequencing can help unravel mechanisms governing defense or resistance responses in plant–pathogen interactions. Several studies have been carried out to investigate the transcriptomic changes in *Musa* germplasm against Yellow Sigatoka disease, but the defense response of *Musa paradisiaca* has not been investigated so far. We carried out transcriptome sequencing of *M. paradisiaca* var. Kachkal infected with the pathogen *Pseudocercospora musae* and found that a vast set of genes were upregulated while many genes were downregulated in the resistant cultivar as a result of infection. After transcriptome assembly and differential gene expression analysis, 429 upregulated and 156 downregulated genes were filtered out (considering fold change ± 2 , $p < 0.01$). Functional annotation of the differentially expressed genes (DEGs) enriched the upregulated genes into 49 gene ontology (GO) classes of biological processes (BP), 20 classes of molecular function (MF) and 9 classes of cellular component (CC). Similarly, the downregulated genes were classified into 35 GO classes of BP, 28 classes of MF and 6 classes of CC. The KEGG enrichment analysis revealed that the upregulated genes were most highly represented in ‘metabolic’ and ‘biosynthesis of secondary metabolites’ pathways. Additionally, ‘plant hormone signal transduction’, ‘plant–pathogen interaction’ and ‘phenylpropanoid biosynthesis’ pathways were also significantly enriched indicating their involvement in resistance responses against the pathogen. The RNA-seq analysis also depicts that a range of important defense-related genes are modulated as a result of infection, all of which are responsible for either mediating or activating resistance responses in the host. We studied and validated the expression profiles of ten important defense-related genes potentially involved in conferring resistance to the pathogen through qRT-PCR. Almost all the selected defense-related genes were found to be highly and significantly upregulated within 24 h post inoculation (hpi) and for some genes, the expression remained consistently high till the later time point of 72 hpi. These results, thus, indicate that the infection by *P. musae* leads to a rapid reprogramming of the defense transcriptome of the resistant banana cultivar. The defense-related genes identified to be modulated in response to infection are important not only for pathogen recognition and perception but also for activation and persistence of defense in the host.

Keywords *Musa paradisiaca* · Sigatoka · Host–pathogen interaction · Transcriptome · Differentially expressed genes

Introduction

Bananas including plantains are one of the most important fruits crops all over the world. From the production context, after rice, wheat and maize, bananas are the fourth most important crop and are grown in all the tropical and

subtropical regions including coastal, deltaic, plain inland to hills, up to an altitude of 1800 m (Mustaffa and Sathiamoorthy 2004). According to the Food and Agriculture Organization of the United Nations, the world’s banana production was 116 million tons during 2017–2019. It is a highly nutritious fruit crop and is a major source of carbohydrates and has high contents of potassium, vitamin B and C. Most of the modern cultivated bananas that produce edible fruits are hybrids of intra- or interspecies crosses between two diploid wild species, *Musa acuminata* (A genome) and *M. balbisiana* (B genome) both belonging to Eumusa section of the Musaceae family (Simmonds and Sheppard 1955; Heslop-Harrison and Schwarzacher 2007). *M. paradisiaca* (ABB) is

✉ Priyadarshini Bhorali
priyadarshini.bhorali@aau.ac.in

¹ Department of Agricultural Biotechnology, Assam Agricultural University, Jorhat, Assam 785013, India

² Department of Life Sciences, Dibrugarh University, Dibrugarh, Assam 786004, India

a hybrid between *M. acuminata* and *M. balbisiana*, and the vernacular name of the cultivar in Assam is ‘Kachkal’. It is a culinary banana found in the entire North-Eastern region of India and has been reported to be relatively resistant to common banana pests and pathogens (Mishra 2000).

The global production of bananas is threatened by several biotic stresses caused by pests and pathogens. Among the fungal diseases, the Sigatoka leaf spot disease complex is one of the most devastating foliar diseases of bananas. The causal agent of Sigatoka disease comprises of three species of *Mycosphaerella*. Black Sigatoka is caused by *Mycosphaerella fijiensis* (anamorph *Pseudocercospora fijiensis*), Yellow Sigatoka is caused by *M. musicola* (anamorph *P. musae*) and Eumusae leaf spot is caused by *M. eumusae* (anamorph *P. eumusae*) (Carlier et al. 2000; Jones 2000; Arzanlou et al. 2007). Black Sigatoka disease is the most virulent among the three leaf spot diseases and is distributed worldwide affecting a wider range of banana genotypes (Carlier et al. 2000). But in India, Yellow Sigatoka and Eumusae leaf spot are more prevalent and severe. In the North-Eastern regions, Sigatoka disease results in significant yield loss predominantly in the cultivars of the Cavendish subgroup. Although studies are being carried out worldwide concerning host resistance in banana against Black Sigatoka disease, molecular studies for an incompatible interaction between *Musa* spp. and the Yellow Sigatoka pathogen is still lagging behind, particularly in India. Management of the Sigatoka disease in commercial bananas is highly dependent on the use of fungicides and cultural practices. But the use of chemicals brings collateral problems such as environmental pollution, human/animal health hazards and pathogen resistance.

In recent years, high-throughput next-generation sequencing (NGS) platforms have evolved as revolutionary tools for gaining complete insight into genomes through large scale genome and transcriptome sequencing. RNA sequencing (RNA-seq) is an affordable, accurate and high-resolution deep sequencing technology that sequences the entire transcriptome of an organism within a short period. It is being widely used for transcriptome sequencing in many plant species to understand the mechanisms of plant defense responses to biotic stresses, and to identify defense or resistance-related genes implicated in host–pathogen interactions. Such defense-related or resistance-related candidate genes could further be utilized for crop improvement through plant genetic modification and marker-assisted breeding technologies. A number of studies have been carried out in several genotypes of *Musa* spp. to unravel host responses to various biotic and abiotic stresses including responses to the Sigatoka pathogen, by investigating transcriptomic changes and identifying differentially expressed genes (DEGs). In a study by Passos et al. (2013) where *M. acuminata* was challenged by *M. musicola*, several defense-related genes

and transcription factors such WRKY transcription factors, MAP kinase, pathogenesis-related protein1, putative chitinase, endochitinase, serine/threonine-protein kinase, NBS-LRR protein, etc. were found to be upregulated. In another study by Sun et al. (2019), where banana was infected with *Fusarium oxysporum* f. sp. *cubense* (*Foc*) TR4, 9612 DEGs were identified, of which 72.5% were assigned in biological process, 42.2% assigned in cellular component and 82.2% in molecular functioning. Similarly, a comparative transcriptome analysis was carried out in resistant banana ‘Pahang’ and susceptible cultivar ‘Brazilian’ infected by *Foc* TR4 (Zhang et al. 2019) in which the resistant genotype was found to exhibit constitutive defense responses even before pathogen inoculation and inducible defense responses ahead of the susceptible genotype during initial stages of *Foc* infection. In the present study, several transcription factors particularly the NAC, MAPK and WRKY transcription factors have also been identified to be associated with stress responses in banana. Recently, Yan et al. (2021) have reported that the transcriptional activator *NAC42* regulates fruit ripening in banana under oxidative stress. Previous studies have reported *MaNAC42* and *MusaNAC68* to be stress-related transcription factors in banana, as their over-expression in transgenic lines improved tolerance to salinity and drought stresses (Negi et al. 2016; Tak et al. 2017). Similarly, the *MpSNAC67* transcription factor was also demonstrated to be a positive regulator of senescence in banana that regulates the expression of chlorophyll catabolic genes (Tak et al. 2018). Furthermore, *MusaSNAC1*, which modulates stomatal closure and H_2O_2 content in response to drought tolerance in banana, was found to have a potential role in cross talk between shoot multiplication and drought responses (Negi et al. 2018, 2021). Mitogen-activated protein kinases (MAPKs) are proteins involved in many functions such as signaling, stress response and plant development (Hamel et al. 2006). The MAPK gene *MusaMPK5* has been reported to be involved in the regulation of cold tolerance in banana and is found to be associated with phosphorylation of *NAC42* and *SNAC67* (Tak et al. 2020). The WRKY transcription factors have several established biological roles in banana plants such as those associated with stress responses, fruit ripening, magnesium deficiency, etc. (Yang et al. 2021; Jia et al. 2022). The *MusaWRKY18* was found to be strongly induced in transgenic tobacco lines in response to abiotic stresses (cold, salinity and drought) and applications of several plant growth regulators, indicating its role under multiple stress conditions (Tak et al. 2021). To devise a complete and effective management strategy against diseases, understanding the host–pathogen interactions at the molecular level is a prerequisite. Investigating such interactions can unravel the host defense mechanisms and provide information on genes involved in resistance responses. Further, elucidation of the expression patterns of such genes

would help in the identification of candidate genes involved in conferring resistance. With this background, the present study was designed to undertake a comprehensive study on the banana defense transcriptome during the '*M. paradisiaca*–*P. musae*' interaction, to identify specific molecular events as well as candidate genes associated with defense or resistance against the pathogen.

Materials and methods

Plant preparation, pathogen inoculation and sample collection

The banana cultivar Kachkal (*M. paradisiaca*, Genome ABB), which is resistant to Yellow Sigatoka disease, was chosen for the present study. We collected young and healthy banana suckers with 1–2 opened leaves from the Horticulture Orchard of Assam Agricultural University, Jorhat, and planted in pots (size 40 cm height × 35 cm width) containing mixture of compost and soil (in the ratio 1:2). The plants were grown inside a green house till complete emergence of 4–5 leaves (Supplementary file 1). Leaf samples showing typical Yellow Sigatoka leaf spot symptoms were collected from a susceptible banana cultivar *M. accuminata* var. Jahaji (AAA), cut into 2 cm pieces and surface sterilized using 0.1% mercuric chloride. The pathogen *P. musae* was isolated from infected leaves of banana identified by their typical symptoms following methodology described by Selvarajan et al. (2001). Leaves with typical lesions of Sigatoka were collected, cut into small sections of 2 cm² sizes and dipped in sterile distilled water for 15 min. The sections were then stapled to sterile filter paper and placed on the lid of a petriplate facing 3% water agar. The plates were incubated at room temperature and the release of spores was monitored regularly under a microscope. The spores identified based on their morphology were then transferred to V8 juice agar medium and incubated at 28 °C. For microscopic detection of the fungus, a small amount of the culture was stained with Lactophenol blue and observed under a light microscope (Supplementary file 1). For molecular detection, genomic DNA of the fungus was isolated by CTAB method followed by PCR amplification using specific primers MMus137/R635 described by Johanson and Jeger (1993) (Supplementary file 1). Young plants of cv. Kachkal with 3–4 opened leaves were artificially inoculated using *P. musae* spore suspension prepared at a concentration of 10⁶ spores mL⁻¹. The suspension was inoculated into the plants by spraying on both the upper and lower leaf surfaces. The plants meant to be used as control were mock inoculated by spraying with sterile distilled water. Leaf samples were collected from the pathogen challenged and mock inoculated plants in triplicates at three time points, i.e., 24 h post inoculation (hpi),

48 hpi and 72 hpi. Samples were collected in liquid nitrogen and immediately stored at – 80 °C. The plants were subsequently monitored for development of symptoms of infection.

Total RNA isolation and RNA-seq

Total RNA was isolated from the leaf samples collected at different time points using the RNeasy Plant Mini Kit (QIAGEN). After checking the quality and quantity of the RNA samples using a Nanodrop 1000 spectrophotometer (Thermo Scientific), we pooled the total RNA extracted from three biological replicates for each treatment (control and *P. musae*-infected tissues at 24, 48 and 72 hpi) for RNA-seq. The samples were also checked by electrophoresing in a 1% denaturing agarose gel (Supplementary file 1). The samples collected at 48 hpi were then precipitated using 0.1 volume of 3 M Sodium Acetate (NaOAc, pH 5.5) and 100% ethanol and shipped in dry ice to Bencos Pvt. Ltd., Mumbai, for sequencing works. RNA-seq was carried out using the Illumina HiSeq 2500 platform following 2 × 101 bp chemistry. Paired-end sequencing libraries were prepared from the total RNA using the Illumina TruSeq RNA sample preparation kit (Cat. No. RS-122-2101). The sequenced raw data was processed to obtain high-quality clean reads using Sequencing Control Software HCS v2.2. The BCL (base calls) binary was converted into FASTQ utilizing the Illumina package bcl2fastq (v1.8.4). The low-quality bases and adapters were trimmed away and the high-quality filtered reads were aligned to the reference *M. acuminata* genome using the software Hisat2. The draft genome and its GTF files were downloaded from the NCBI database to assist the reference-guided assembly. The aligned reads were further processed using the Cufflink software package. The assembled reads were merged using Cuffmerge (a component of Cufflink). The GTF reference file was used to merge new isoforms and known isoforms and maximize overall assembly quality. The Cuffdiff program further estimated the significant changes in the expression of the transcripts. The transcripts were then classified as upregulated and downregulated based on their Generalized Fold change (GFold) value. We considered the transcripts with a difference of at least two-fold change and p-value ≤ 0.01.

Differential gene expression analysis and functional annotation of transcripts regulated during *P. musae* infection in *M. paradisiaca*

The differentially expressed transcripts of both control and treated samples were with the BLASTx program (NCBI) for homology searches against publicly available protein database Swissprot (<http://www.ebi.ac.uk/uniprot/>), annotated and classified on Gene Ontology (<http://www.geneontology.org/>) and

the KEGG pathway (<http://www.genome.jp/kegg/>) with a threshold E value of 10^{-5} for functional annotation and classification. The GO terms were further classified into biological process, molecular function, and cellular components using the WEGO software. Pathview was used to generate KEGG pathways and Mapman was used to view the DEGs in specific biological processes. Venn diagrams were constructed using the online tool Venny (version 2.1.0). The biological pathways for DEGs were assigned by mapping with KEGG database by inbuilt tool of Blast2GO and KAAS annotator.

Validation of RNA-seq data by quantitative real-time PCR (qRT-PCR)

To validate the gene expression patterns obtained through RNA-seq experiments, we carried out qRT-PCR based expression analysis for all the three time points (24, 48, and 72 hpi). We synthesized first-strand cDNA using the High-Capacity cDNA Reverse Transcription kit (Applied Biosystems, USA) and diluted (1:10) with nuclease-free water. The gene-specific primers were designed using the NCBI Primer-Blast tool (<http://www.ncbi.nlm.nih.gov/tools/primer-blast/>) for amplification of selected DEGs based on the functional annotation analysis, considering FC value and p-value. A SYBR green assay was performed using a QuantStudio 5 Real-Time PCR system (Applied Biosystems, USA) using three biological replicates in three technical replicates. β -Tubulin gene was taken as the endogenous control gene. A reaction volume of 10 μ L was set up for each primer set with 5 μ M of forward and reverse primers, 1X SYBR Green PCR Master Mix and nuclease-free water. Cycling conditions consisted of initial template denaturation at 95 °C for 5 min followed by 30 cycles of denaturation at 94 °C for 20 s, annealing at 50–60 °C (primer specific) for 20 s, extension at 72 °C for 25 s and, final extension at 72 °C for 7 min. Finally, the expression level of each gene was relatively quantified based on calculated normalized ratio using advanced relative quantification using QuantStudio5 software version v1.5.x. The mean delta delta Ct value ($2^{-\Delta\Delta Ct}$ value) or RQ (relative quantification) value of technical replicates was taken and presented as a histogram to illustrate the expression patterns. One-way ANOVA was tested between the RQ values obtained from qRT-PCR. Dunnett's multiple comparisons test was done and significance was tested at p-value < 0.005 to test the statistical significance of the expression parameters at different time points for individual genes.

Results

RNA sequencing and transcriptome assembly of DEGs

Total RNA extraction was done from *P. musae* infected *M. paradisiaca* var. Kachkal samples and water-treated controls (mock treated) at different time points 24, 48 and 72 hpi. Paired-end RNA-seq of leaf samples of Kachkal at 48 hpi by *P. musae* (KIc) along with water-treated control (KC) were done in Illumina HiSeq 2500. The transcriptome sequencing generated a total of 43,506,340 reads from mock-treated control sample and 45,459,572 reads from the *P. musae*-infected sample. The ratio of reads having Phred quality score of over 20, i.e., Q20 (%) of control and *P. musae*-infected Kachkal samples were found to be 97.79% and 97.866%, respectively (base call accuracy 99%). The ratio of reads that have Phred quality score of over 30, i.e., Q30 (%) of control and *P. musae*-infected Kachkal samples were found to be 96.207% and 96.353%, respectively (base call accuracy 99.9%). The GC content of the mock-treated control and *P. musae*-infected samples were found to be 49.502% and 50.339%, respectively. After reference-based transcriptome assembly and differential gene expression analysis, it was observed that out of a total of 45,857 genes, 9851 genes were upregulated and 1303 genes were downregulated in the cultivar Kachkal infected with *P. musae* (Supplementary file 2). After applying a threshold value of fold change ± 2 and p-value > 0.01, 429 upregulated genes and 156 downregulated genes were filtered out. The results show that 17 DEGs were exclusively expressed in the control and 52 DEGs were exclusively expressed in *P. musae*-infected Kachkal plant. The heat map of the DEGs in *M. paradisiaca* upon *P. musae* infection is shown in Fig. 1 which depicts the vast array of genes modulated in the resistant banana cultivar upon infection at 48 hpi. The sequences were submitted to nucleotide sequence database of NCBI under Sequence Read Archive (SRA) (Accession: PRJNA493890 ID: 493890).

Functional annotation through gene ontology and KEGG analysis

To understand the molecular mechanism of the interaction between *M. paradisiaca* var. Kachkal and the Sigatoka pathogen *P. musae*, GO (gene ontology) enrichment analyses of the DEGs in Kachkal compared with mock inoculated control sample was carried out using the WEGO software. The upregulated genes were classified into 49 GO classes of biological process (BP), 20 classes of molecular

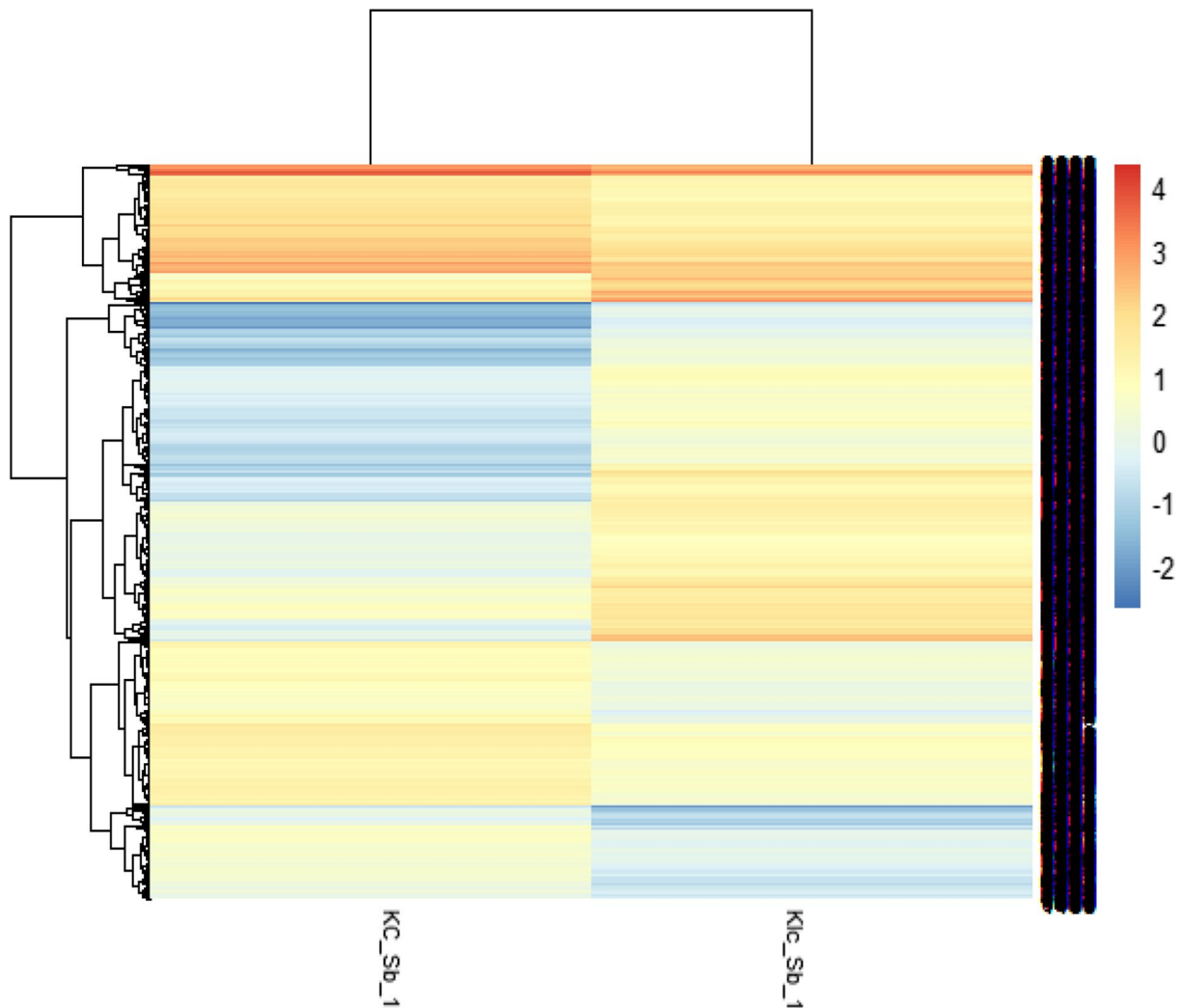


Fig. 1 Heat map of differentially expressed genes (DEGs) in *M. paradisiaca* upon *P. musae* infection

function (MF) and 9 classes of cellular component (CC) (Supplementary file 3). Similarly, the downregulated genes were classified into 35 GO classes of BP, 28 classes of MF and 6 classes of CC (Supplementary file 4). The graphical representations of the top 20 GO categories under BP, MF and CC in which the upregulated and downregulated DEGs associated with the *M. paradisiaca*–*P. musae* interaction were enriched, are depicted in Figs. 2 and 3, respectively. The Venn diagram (Fig. 4) of the enrichment data depicts that among the upregulated genes, 51 were common between BP and MF categories, 32 were common among BP and CC, 36 are common among MF and CC, while 15 genes were common in all the 3 GO categories. In case of downregulated genes, BP and MF share 29 common genes, BP and CC share 80 common genes, CC and MF share 45 common genes, while 16 genes are commonly represented in all the categories (Fig. 4).

Some important defense-related genes significantly upregulated in *M. paradisiaca* var. Kachkal in response to *P. musae* infection at 48 hpi, include LRR receptor-like serine/threonine-protein kinase, non-specific lipid-transfer protein, ethylene-responsive factor, cytochrome P450, peroxidase, phenylalanine ammonia-lyase, plant intracellular Ras-group-related LRR, PR1, germin-like protein, GDSL esterase/lipase, UDP-glucose 6-dehydrogenase, PR4, chitinase, mannose-specific lectin, polyphenol oxidase, WRKY, MAPK, Bowman–Birk-type proteinase inhibitor, zinc finger protein, O-methyltransferase and F-box/LRR-repeat protein. The G-fold change in expression at 48 hpi and specific annotations of the genes as obtained from the RNA-seq analysis are presented in Table 1. Apart from genes, several transcription factors known to be involved in biotic stress responses have been identified to be upregulated as a result of infection. The study shows that 34 transcription factors including

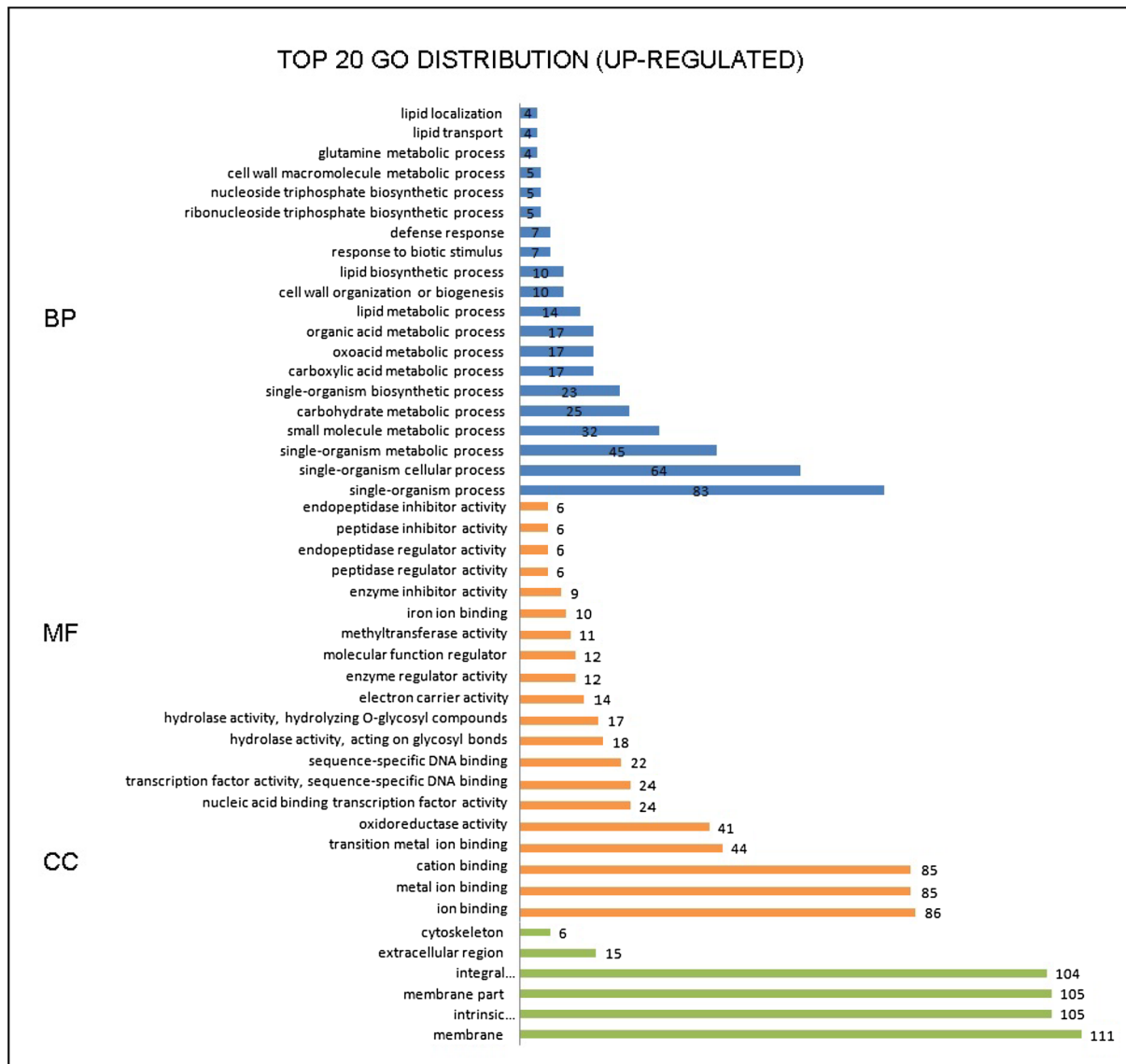


Fig. 2 Graphical representation of the top 20 GO classes, under which the genes upregulated during the *M. paradisiaca*–*P. musae* interaction were enriched

WRKY transcription factors, ethylene-responsive transcription factors, bHLH transcription factors, MYB family transcription factors, heat stress transcription factors, GATA transcription factors, MYC transcription factors, etc. are involved in the interaction, of which 27 were overexpressed while 7 were underexpressed (Table 2).

To elucidate the functional involvement of transcripts in metabolic pathways, the DEGs were annotated against the KEGG database. The results of the KEGG analysis depict that, the DEGs which were upregulated in response to infection were significantly enriched in 20 KEGG pathways (Supplementary file 3). The pathways in which a large number of genes were represented include metabolic pathways (286 genes), biosynthesis of secondary metabolites (173 genes),

plant hormone signal transduction (47 genes), and phenylpropanoid biosynthesis (27 genes). The downregulated genes were annotated into 17 pathways (Supplementary file 4) and the highly enriched ones include metabolic pathways (286 genes), biosynthesis of secondary metabolites (173 genes), plant hormone signal transduction (62 genes), amino sugar and nucleotide sugar metabolism (46 genes). Furthermore, to visualize the functional involvement of DEGs in biological pathways administrating plant defense in *M. paradisiaca* in response to Sigatoka infection at 48 hpi, GO results assigned were mapped using Mapman for plant–pathogen interaction and plant hormone signal transduction pathways. The plant–pathogen interaction pathway generated from the analysis indicates several upregulated DEGs (Fig. 5) such

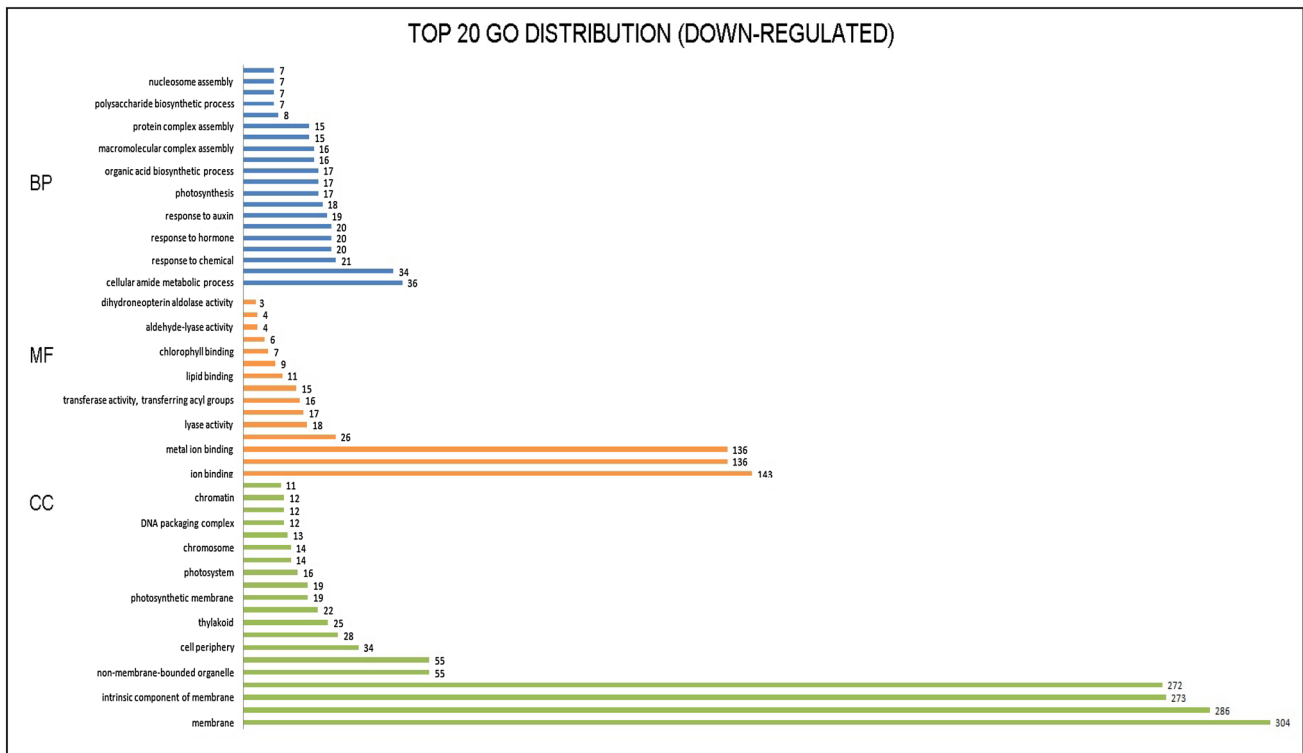


Fig. 3 Graphical representation of the top 20 GO classes, under which the genes downregulated during the *M. paradisiaca*–*P. musae* interaction were enriched

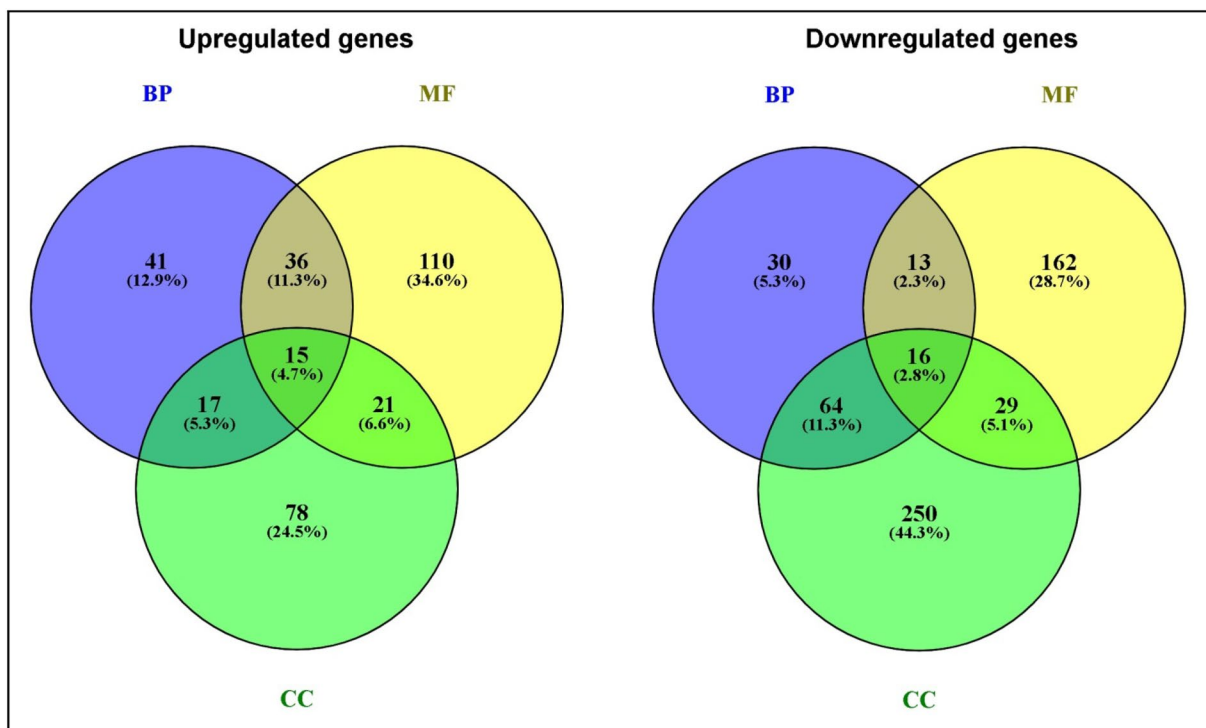


Fig. 4 Venn diagram depicting GO distribution showing numbers of overlapping genes (upregulated and downregulated genes) among the classes of biological process, molecular function and cellular component

Table 1 Defense-response genes highly upregulated during *M. paradisiaca*–*P. musae* interaction at 48 hpi

Sl. no.	Gene	G-Fold	Gene ID	GO class
1	Non-specific lipid-transfer protein-like protein At2g13820	2.94874	Ma02_g19210	Metabolic pathways
2	Phenylalanine ammonia-lyase	3.71392	Ma05_g03720	Biosynthesis of secondary metabolites
3	Ethylene responsive	2.53381	Ma11_g20400	Small molecule metabolic process
4	Cytochrome P450 86B1	3.35509	Ma04_g21730	Oxidoreductase activity
5	Peroxidase 4 (P7 like)	3.08634	Ma10_g07300	Oxidoreductase activity
6	Plant intracellular Rasgroup-related LRR protein 1	2.6941	Ma03_g13420	Intracellular membrane bounded organelle
7	Pathogenesis-related protein 1-like	3.2666	Ma03_g08160	Response to biotic stimulus
8	Germin-like protein 3–8	2.93965	Ma04_g17850	Extra cellular space
9	GDSL esterase/lipase At5g55050-like	2.80955	Ma09_g24530	Hydrolase activity
10	UDP-glucose 6-dehydrogenase 5-like	3.62154	Ma06_g12550	Metabolic pathways
11	Pathogen-related protein-like	2.0226	Ma06_g18530	Response to biotic stimulus
12	Pathogenesis-related protein PR-4-like	3.63418	Ma06_g03120	Response to biotic stimulus
13	Chitinase 1-like	2.48003	Ma03_g28030	Catalytic activity
14	Mannose-specific lectin	3.28988	Ma09_g10450	Mannose binding
15	Polyphenol oxidase	2.119	Ma08_g09160	Biosynthesis of secondary metabolites
16	WRKY transcription factor	2.78011	Ma04_g24790	Sequence-specific DNA binding
17	Mitogen-activated protein kinase 5-like	3.15461	Ma09_g04230	MAPK cascade
18	Bowman–Birk-type proteinase inhibitor-like	3.4837	Ma04_g24980	Peptidase inhibitor activity
19	Zincfinger protein WIP3-like	2.70682	Ma11_g17420	Metal ion binding
20	Putative O-methyltransferase	4.0605	Ma05_g16730	O-methyltransferase activity
21	Pathogenesis-related protein 1-like	3.02678	Ma03_g08140	Regulation of biological process
22	Peroxidase 5	2.57759	Ma05_g15710	Oxidoreductase
23	F-box/LRR-repeat protein 3	2.56387	Ma11_g16610	Protein kinase activity
24	Probable LRR receptor-like serine/threonine-protein kinase	13.255	Ma10_g09270	Protein kinase activity
25	GDSL esterase/lipase	2.69	Ma04_g29060	Hydrolase activity
26	Germin-like protein 2	2.769	Ma03_g13630	Extra cellular space

as CDPK, Rboh, Cam/CML, Pti1, Pti6, PR1, etc. which are important factors involved in PAMP-triggered immunity. Other factors such as CERK1, PBS1, and WRKY associated with effector-triggered immunity were also found to be upregulated. All these genes are important components of the plant–pathogen interaction pathway that are associated with hypersensitive response (HR), programmed cell death (PCD) and defense-related gene induction. A map of the plant hormone signal transduction pathway depicts several genes/factors associated with signal transduction mechanisms leading to cell enlargement (AUX/IAA, GH3, SAUR), cell division (AHP, ARR), stem growth (GID, TF), stomatal closure (PYR/PYL, PP2C, SnRK2), senescence/stress response (JAR1, JAZ, MYC2), disease resistance (NPR1, TGA, PR1) (Fig. 6).

Validation of RNA-seq data through expression profiling of defense response genes

To confirm and validate the differential expression profiles obtained from the transcriptome analysis data, we randomly selected a total of 10 defense-related DEGs for qRT-PCR

analysis. The genes exhibiting defense response-related traits (based on their functions related to defense as established by previous studies) which were significantly upregulated in the resistant host upon infection were chosen for qRT-PCR validation *viz.* pathogenesis-related protein PR1-like (PR1), PR4, pathogen-related protein-like (PR), ethylene-responsive transcription factor (ERF), plant intracellular Ras-group-related LRR protein 1-like (Ras-LRR), WRKY transcription factor (WRKY), LRR receptor-like serine/threonine-protein kinase (LRR-STK), putative O-methyltransferase (OMT), mitogen-activated protein kinase (MAPK) and germin-like protein (GLP) (Table 3). The qRT-PCR-based expression profiling revealed the differential expression of the selected genes at three time points *viz.* 24, 48 and 72 hpi (Fig. 7) and also validated the transcriptome expression data. The RNA-seq study was carried out using the samples (mock inoculated control and infected) from 48 hpi and the fold change in expression was found to be analogous to the expression patterns of the genes in qRT-PCR analysis.

The qRT-PCR study revealed that the PR1 gene was significantly upregulated by 2.064-fold within 24 hpi and the expression further increased at 48 hpi (2.867-fold).

Table 2 Transcription factors differentially expressed in the *M. paradisiaca*–*P. musae* interaction

Sl no	ID	Transcription factor	GFold (0.1)
1	Ma10_g06760	Transcription factor SPEECHLESS-like	4.77132
2	Ma10_g25820	Transcription initiation factor TFIID subunit 9-like	4.45832
3	Ma08_g17860	Transcription factor MYB108-like	4.28183
4	Ma10_g11100	Transcriptional activator Myb	3.17716
5	Ma04_g02810	Probable WRKY transcription factor 71	2.98827
6	Ma04_g26400	Heat stress transcription factor B-2c-like	2.9187
7	Ma07_g11370	Transcription factor ILI5-like	2.76168
8	Ma01_g16650	Transcription factor LUX-like	2.69033
9	Ma01_g03440	Transcription factor MYC3-like	2.63309
10	Ma04_g03780	AP2/ERF and B3 domain-containing transcription factor RAV1-like	2.58357
11	Ma11_g20400	Putative Ethylene-responsive transcription factor 1B	2.45023
12	Ma07_g03600	Probable transcription factor KAN4	2.43873
13	Ma10_g12120	ANAC083 transcription factor	2.43226
14	Ma08_g17710	AP2/ERF and B3 domain-containing transcription repressor RAV2like	2.35534
15	Ma06_g24790	Ethylene-responsive transcription factor RAP2-1-like	2.32036
16	Ma04_g24790	WRKY transcription factor 22%2C transcript variant X1	2.28186
17	Ma11_g21710	transcription factor MYC3-like	2.27829
18	Ma04_g24790	WRKY transcription factor 22%2C transcript variant X1	2.27617
19	Ma04_g35590	Heat stress transcription factor B-2b-like	2.24199
20	Ma06_g28830	Probable WRKY transcription factor 71	2.20857
21	Ma10_g16990	ANAC083 transcription factor	2.17946
22	Ma04_g37350	Putative GATA transcription factor 22	2.15238
23	Ma03_g22010	AP2/ERF and B3 domain-containing transcription repressor RAV2like	2.15097
24	Ma06_g11380	Probable transcription factor GLK2	2.12882
25	Ma06_g22210	ANAC083 transcription factor	2.12378
26	Ma09_g28400	Transcription factor bhlh130-like%2C transcript variant X2	2.07411
27	Ma07_g22320	Ethylene-responsive transcription factor RAP2-9-like	2.05575
28	Ma02_g20270	MYB family transcription factor%2C putative%2C expressed	– 2.98808
29	Ma06_g07330	Probable WRKY transcription factor 41	– 2.8705
30	Ma01_g17260	Putative Transcription factor MYB86	– 2.71337
31	Ma08_g14720	MYB family transcription factor%2C putative%2C expressed	– 2.45973
32	Ma08_g27190	Transcription factor bhlh18-like	– 2.36151
33	Ma01_g23360	Probable WRKY transcription factor 53	– 2.13854
34	Ma01_g12390	Transcription factor bhlh76-like%2C transcript variant X2	– 2.02894

However, its expression rapidly declined at 72 hpi. In case of PR4, there was no significant change in expression at 24 hpi but the gene was significantly highly upregulated at 48 hpi by 3.949-fold. Interestingly, its expression also declined quite rapidly at 72 hpi. Similarly, a pathogen-related protein showed a slight induction after 48 hpi by 2.326-fold but was rapidly and highly upregulated by 6.314-fold at 72 hpi as compared to control (Fig. 7).

In case of ERF (Ethylene-responsive transcription factor ERF037-like), there was an increase in expression by 1.903-fold at 24 hpi, and 1.799-fold at 48 hpi, but the gene was highly upregulated by 2.911-fold at 72 hpi. On the other hand, the Ras-group-related LRR protein 1-like gene (Ras-LRR) did not show significant upregulation at 24 hpi but

its expression increased at 48 hpi significantly by 2.152-fold. Later at 72 hpi, the gene was highly upregulated by 5.018-fold. A similar expression pattern was also seen in case of WRKY. The WRKY transcription factor showed no significant change in expression at the initial time point but later, at 48 hpi, it was highly upregulated by 2.986-fold. At the last time point (72 hpi), WRKY showed a sudden spike in expression by 7.533-fold (Fig. 7).

Further, in case of LRR-STK (LRR receptor-like serine/threonine-protein kinase) significantly high upregulation by 21.95-fold was observed at the initial time point. In the later time points, at 48 and 72 hpi, although its expression was observed to be gradually declining, it remained significantly higher as compared to control (12.324- and 10.39-fold

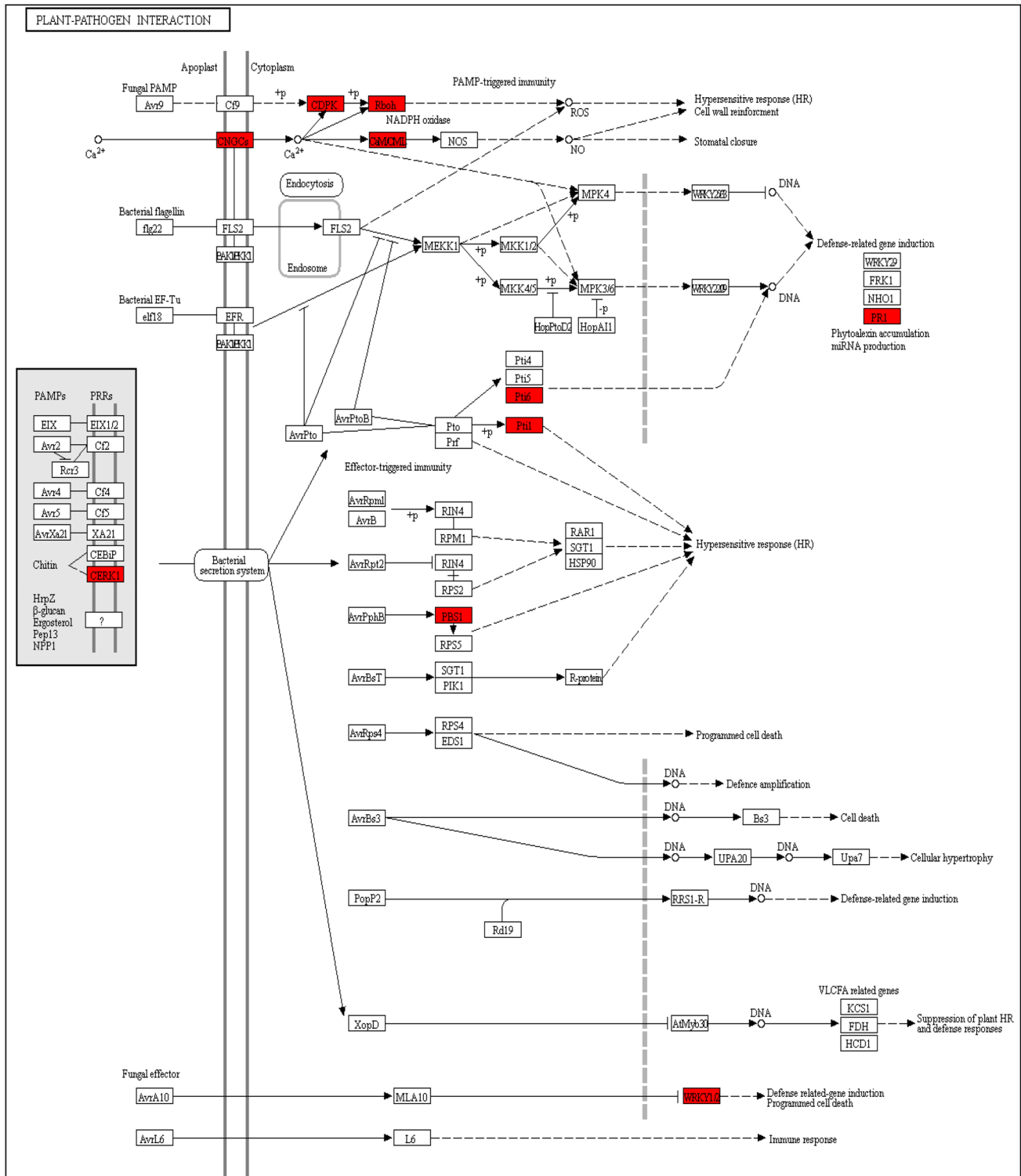


Fig. 5 Visualization of DEGs involved in the plant–pathogen interaction pathway in *M. paradisiaca* upon *P. musae* infection at 48 hpi. White box or no color fill means no DEG was assigned to that ontology term and red represents upregulation

during 48 hpi and 72 hpi, respectively). In contrast, although no significant change in expression was observed in OMT (O-methyltransferase ZRP4) at 24 hpi, the expression

gradually increased at the later time points. OMT showed significantly high upregulation at 48 hpi (4.592-fold) and 72 hpi (7.917-fold). Similarly, MAPK showed no upregulation

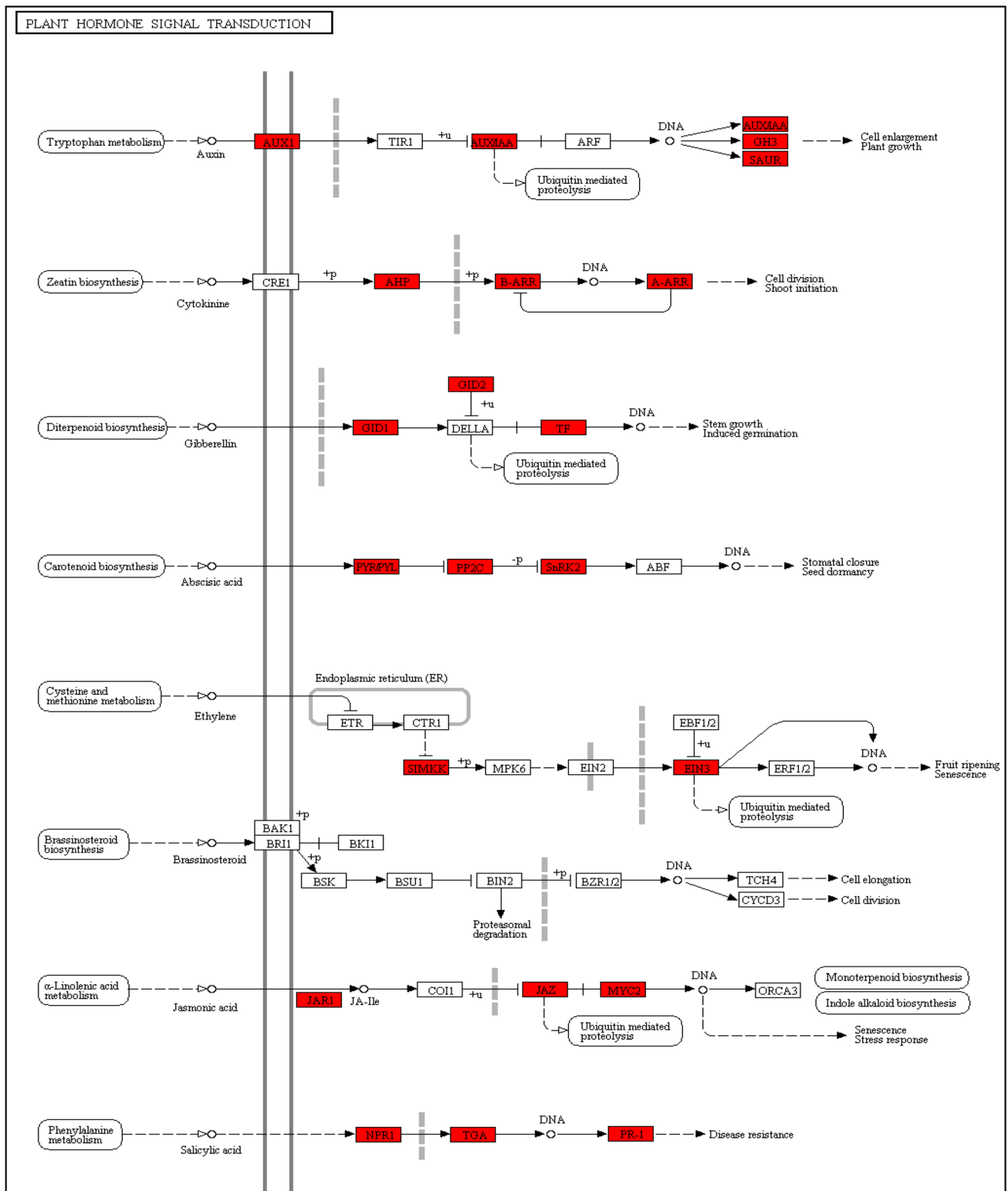


Fig. 6 Visualization of DEGs involved in the plant hormone signal transduction pathway in *M. paradisiaca* upon *P. musae* infection at 48 hpi. White box or no color fill means no DEG was assigned to that ontology term and red represents upregulation

Table 3 Defense-related genes associated with *M. paradisiaca* defense against *P. musae* infection, selected for qRT-PCR showing fold changes in expression (at 48 hpi) as observed from the RNA-seq and qRT-PCR analyses

Sl. no.	Attribute	Description	Gene ID	Function (and GO category)	Gfold at 48 hpi (RNA-seq)	Fold change in expression at 48 hpi (qRT-PCR)
1	Ma03_g08160	Pathogenesis-related protein PR1-like	PR1	Response to biotic stimulus/ defense response (BP)	2.817	2.867
2	Ma06_g03120	Pathogenesis-related protein PR4-like	PR4	Response to biotic stimulus/ defense response/ defense against fungus (BP)	2.926	3.949
3	Ma09_g05410	Pathogen-related protein-like	PR-like	Response to biotic stimulus/ defense response (BP)	3.572	2.326
4	Ma11_g20400	Ethylene-responsive transcription factor ERF037-like	ERF	Molecular transcription (MF)	2.973	1.799
5	Ma03_g13420	Plant intracellular Ras-group-related LRR protein 1-like	Ras -LRR	Protein kinase activity (MF)	2.163	2.152
6	Ma04_g02810	WRKY transcription factor WRKY71	WRKY	DNA-binding transcription factor activity (MF)	2.988	2.986
7	Ma10_g09270	LRR receptor-like serine/threonine-protein kinase	LRR-STK	Protein kinase activity (MF)	13.255	12.324
8	Ma05_g16730	Putative O-methyltransferase ZRP4	OMT	Methyltransferase activity (MF)	4.024	4.592
9	Ma09_g04230	Mitogen-activated protein kinase 5-like	MAPK	MAP kinase activity (MF)	2.274	3.398
10	Ma04_g17850	Germin-like protein 3–8	GLP	Defense response to fungus (MF)	2.939	2.445

at 24 hpi but at 48 hpi, the gene was significantly upregulated by 3.398-fold and was highly upregulated by 9.024-fold at 72 hpi. Finally, in case of Germin-like protein (GLP), although no change in expression was seen initially, the gene was significantly upregulated by 2.445- and 4.384-fold at 48 hpi and 72 hpi, respectively (Fig. 7).

The investigation, thus, established that a resistant host can quickly deploy defense strategies within 24–48 h of infection. Majority of the defense-related genes selected for the validation study were found to be highly and significantly upregulated within these time points. Moreover, for most of the genes, the expression remained consistently high till the later time point of 72 hpi. These results, thus, indicate that the early defense-related genes identified from the study are important not only for pathogen recognition and perception but also for activation of defense-related genes and persistence of defense in *M. paradisiaca* against *P. musae* invasion.

Discussion

High throughput next generation sequencing technologies such as RNA-seq can help understand mechanisms underlying resistance responses in plant–pathogen interactions. Several studies have been carried out to study the transcriptomic changes in *Musa* germplasms against Sigatoka leaf

spot disease (Portal et al. 2011; Passos et al. 2013; Rodriguez et al. 2016), however, the defense transcriptome of *M. paradisiaca* in response to Sigatoka disease has not been investigated so far. In the current investigation, the transcriptome sequencing of *M. paradisiaca* var. Kachkal, during Sigatoka infection caused by *P. musae* was carried out using the Illumina platform. The differential gene expression analysis revealed that a vast set of genes were upregulated while many genes were downregulated in Kachkal at 48 hpi as a result of *P. musae* infection. The RNA-seq analysis also depicted that a wide range of important defense-related genes are upregulated in ‘Kachkal’ in response to *P. musae* infection with functions related to oxidoreductase activity, response to biotic stimulus, protein kinase activity, hydrolase activity, metabolic pathways, catalytic activity, mannose binding, sequence-specific DNA binding, MAPK cascade, etc., all of which are responsible for either mediating or activating resistance responses in the host. The KEGG enrichment analysis shows that the upregulated genes were most highly represented in ‘metabolic’ and ‘biosynthesis of secondary metabolites’ pathways. Additionally, ‘plant hormone signal transduction’, ‘plant–pathogen interaction’ and ‘phenylpropanoid biosynthesis’ pathways were also highly significantly enriched indicating their involvement in imparting resistance to the pathogen. The analysis also indicated the involvement of several classes of transcription factors in the interaction such as WRKY, ERF, bHLH, MYB, most of

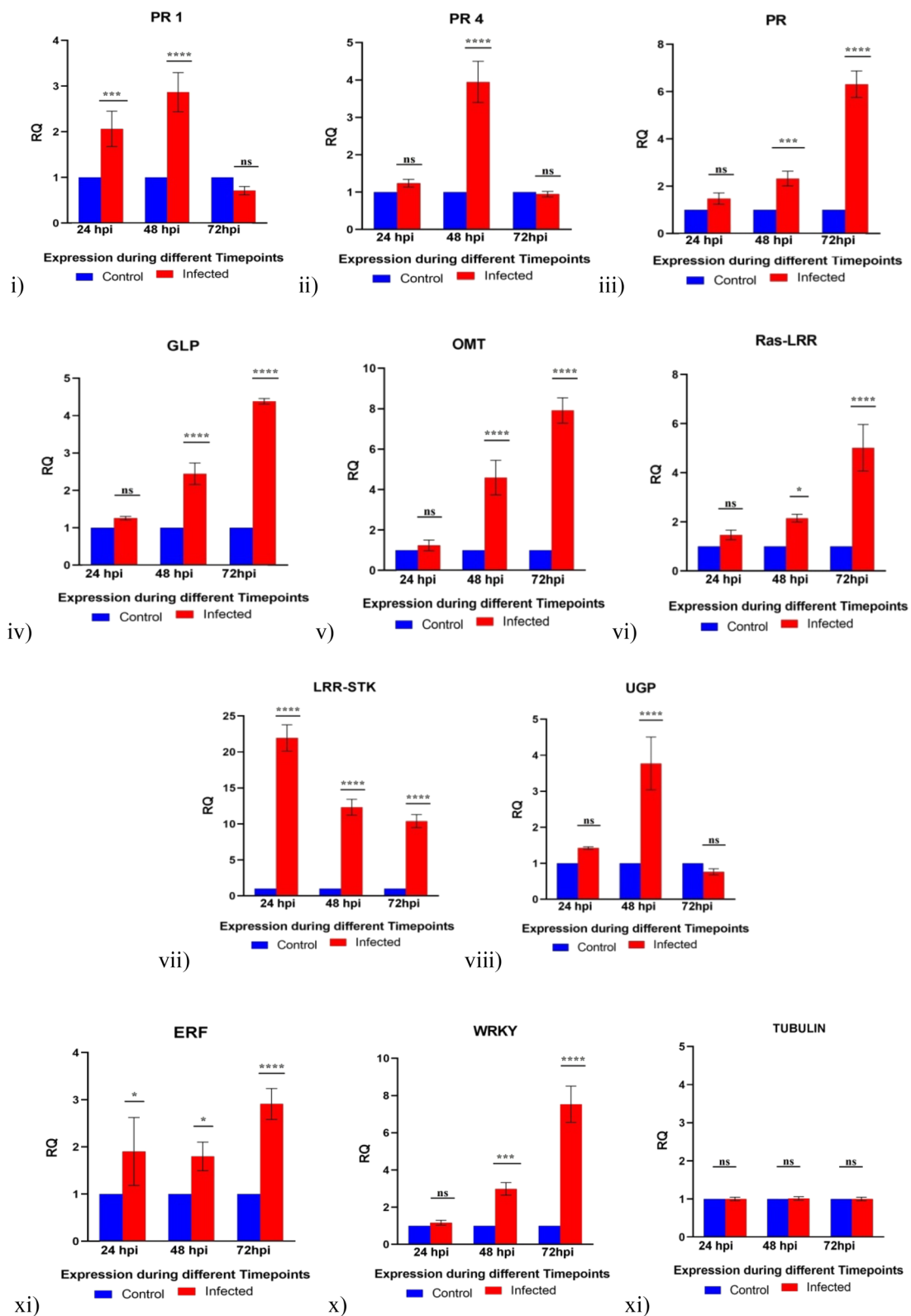


Fig. 7 qRT-PCR-based differential expression analysis of selected defense-related genes upregulated in *M. paradisiaca* in response to infection by *P. musae*: (i) PR1 (ii) PR4, (iii) PR-like, (iv) GLP (v) OMT, (vi) Ras-LRR, (vii) LRR-STK, (viii) UGP, (ix) ERF, (x)

WRKY, across 24, 48 and 72 hpi using (xi) Tubulin (as internal control) (Y axis –RQ obtained from RT-PCR analysis of the genes; X axis samples at different conditions *viz.* control and infected leaf tissue at 24 hpi, 48 hpi and 72 hpi)

which are known to be involved in regulating downstream defense gene activation.

The qRT-PCR study evaluated the expression patterns of the selected defense response genes in the host banana plant Kachkal in response to infection by *P. musae*. The study also corroborated the RNA-seq results, as the fold change in expression patterns of the selected genes at 48 hpi, obtained from qRT-PCR analysis, was found to be comparable to those obtained from the transcriptome data. The genes selected for the validation study have defense associated functions. The PR proteins, induced during pathogen attack as well as other stresses, belong to multi-gene families. Several PR proteins have been identified to exhibit plant defenses against a range of plant pathogens including fungi (van Loon and van Strien 1999), and are a vital part of plant innate immunity. PR1 is one of the most abundant groups of PR proteins mostly induced by pathogen attack or salicylic acid and is used as a marker for pathogen-induced systemic acquired resistance (SAR) (van Loon et al. 2006). Another PR protein, PR4, is reported to have antifungal activity and is induced by pathogen infection or SA, JA and ethylene (Van Loon and Van Strien 1999; Elvira et al. 2008). PR4 can disrupt cell polarity by binding to chitin present in the fungal cell wall (Portal et al. 2011). In the present study, a rapid but slight induction in expression of PR1 was observed in the resistant cultivar immediately after infection at 24 hpi which further enhanced at 48 hpi as revealed by qRT-PCR. However, its expression declined rapidly at 72 hpi. In case of PR4, although no significant change in expression was seen at 24 hpi but the gene was highly upregulated at 48 hpi. Interestingly, just like PR1, the expression of PR4 also rapidly declined at 72 hpi. The pathogen-related protein also did not show any significant upregulation at 24 hpi but was upregulated at 48 hpi. At 72 hpi, it showed a quite rapid and high upregulation as compared to control. These results are in agreement with Rodriguez et al. (2016) who reported that a resistant variety exhibits an earlier and stronger gene expression as compared to a susceptible one. They had conducted microarray analysis in resistant *M. acuminata* ssp. *Burmannicoides* ‘Calcutta 4’ after inoculation with *M. fijiensis* and detected high levels of relative gene expression in PR1 and PR4 at 24 hours after inoculation. In contrast, these genes were downregulated in the susceptible genotype ‘Williams’ (Rodriguez et al. 2016). Zhang et al. (2019) carried out comparative transcriptome analysis in forms of resistant banana ‘Pahang’ and susceptible Cavendish cultivar ‘Brazilian’ infected by *Foc* TR4. They observed that the resistant genotype exhibited constitutive defense responses before pathogen inoculation and inducible defense responses before the susceptible genotype at the initial stages of *Foc* infection. They further demonstrated that several defense response DEGs including PR1 were upregulated rapidly within one day post inoculation. Earlier, Bai et al. (2013)

had also reported that the resistant cultivar ‘Yueyoukang 1’ could trigger a much faster defense response against *Foc* TR4 infection compared to ‘Brazilian’.

The Ethylene Response Factor (ERF) superfamily has been reported to play important roles in response to several stresses including pathogen infection (Mizoi et al. 2012). The ethylene regulated expression of secondary transcription factors like ERF can eventually regulate the expression of downstream defense-related genes in abiotic as well as biotic stresses (Berrocal-Lobo et al. 2002; Zhang et al. 2004; Zhang et al. 2007). In the present investigation, it was found that the expression of ERF increased immediately at 24 hpi as shown by qRT-PCR analysis. At 48 hpi, its expression was still significantly high as compared to the control and was further upregulated at the late time point. This gradual increase in the expression of ERF protein over time in the resistant host upon infection by *P. musae* indicates its possible role in conferring defense. McGrath et al. (2005) have observed that ERFs show altered transcript levels following infection with the incompatible pathogen *Alternaria brassicicola* as early as at 6 hpi. They have further identified that 10 members of the APETALA2/ethylene response factor family were induced by *F. oxysporum* as well as methyl jasmonate treatments. In tomato, the AP2/ERF transcription factors were found to induce defense-related genes, which when overexpressed in transgenic *Arabidopsis*, led to resistance against *Erysiphe orontii* and *P. syringae* (Gu et al. 2002). Again in transgenic wheat, overexpression of ERF has been found to confer enhanced resistance to fungal pathogen *Rhizoctonia cerealis* (Zhu et al. 2014).

The WRKY transcription factors, along with other signaling proteins, regulate the defense-related signal transduction pathways as well as the overall expression of defense-response genes (Mao et al. 2011). Several studies have reported and established the functions of WRKY in plant resistance responses particularly in response to fungal pathogens (AbuQamar et al. 2006; Xu et al. 2006; Zheng et al. 2006). Our study indicates strong involvement of WRKY in mediating defense-related signaling in response to pathogen infection. Although no significant change in expression of WRKY was noticed in Kachkal at 24 hpi upon infection by *P. musae*, it upregulated at 48 hpi, and its expression increased further at 72 hpi. This is in equivalence with the study by Aamir et al. (2018) which reported that *F. oxysporum* f. sp. *lycopersici* challenged tomato shows increased expression of two WRKY genes (*SolyWRKY33* and *SolyWRKY37*) at 48 h. Moreover, the expression of these genes was lesser at the earlier phase of infection (0–24 h) as compared to that in 48 h, which also corroborates our results. In *Brassica rapa*, the relative expression of WRKY genes was recorded to be highest on the 6th day post inoculation by fungus *F. oxysporum* sp. *conglutinans* and on the 7th day post inoculation by bacterium *Pectobacterium carotovorum*

sub. sp. *carotovorum* (Kayum et al. 2014). In banana, Niu et al. (2018) have reported the upregulation of *MaWRKY50* against *F. oxysporum* infection in a resistant cultivar, both before and after pathogen infection. In contrast to these reports, Tripathi et al. (2019) reported that the expression of nine WRKYs transcription factors was suppressed in response to infection by the bacterial pathogen *Xanthomonas campestris* cv. *musacearum* in the Banana *Xanthomonas* Wilt-resistant genotype indicating their involvement as negative regulators of defense.

Plant innate immunity is composed of pathogen-associated molecular pattern (PAMP)-triggered immunity (PTI) and effector-triggered immunity (ETI). PTI is initiated by the recognition of the molecular structure of pathogens and it often activates the downstream mitogen-activated protein kinase (MAPK) signaling cascades leading to expression of a range of defense genes. MAPKs can activate and regulate the expression of many defense-related or resistance-related genes through the phosphorylation of proteins and transcription factors (Zhang and Klessig 2001). Wang et al. (2017) have demonstrated that the MAPK signaling cascade participates in multiple stress signaling in bananas. They identified 10 MAPKK and 77 MAPKKK genes in the banana genome involved in tissue development, fruit development and ripening, and response to abiotic stress. In the present study, it was observed that MAPK was significantly upregulated in Kachkal in response to *P. musae* at 48 hpi as revealed by qRT-PCR. At 72 hpi, MAPK expression increased to a great extent. These results indicate the importance of the MAPK cascade in mediating resistance to the pathogen. A transcriptome analysis involving two contrasting barley genotypes in response to *Blumeria graminis* f. sp. *hordei* infection have reported that several DEGs involved in the MAPK signaling pathway are upregulated at 24 and 48 hpi in the resistant genotype Feng 7 (Li et al. 2019). Similarly, Sun et al. (2019) have also reported the upregulation of a MAPK pathway gene in response to Fusarium wilt infection in the resistant banana cultivar ‘Guijiao 9’. Studies in sugarcane and soybean have also reported a peak in expression of MAPK genes in response to infection by fungal pathogens at 72 hpi, particularly in resistant cultivars (Wu et al. 2013; Lanubile et al. 2015).

Plants have pattern recognition receptors (PRRs), belonging to classes of receptor-like kinases (RLKs) which can recognize conserved molecular signatures of invading pathogens and initiate PTI. The leucine-rich repeat (LRR)-RLKs are the largest group of plasma membrane proteins that are known to be involved in pathogen perception and recognition and mediating the downstream responses. The Ras-group are a class of intracellular LRR proteins that interact directly with Ras GTPases, and are involved in defense signal transduction in animals, lower eukaryotes and plants (Claudianos and Campbell 1995). In the present investigation, the

Ras-group-related LRR protein 1-like gene did not show significant upregulation at 24 hpi but its expression was significantly induced at 48 hpi and later, was highly upregulated at 72 hpi. A receptor-like serine/threonine-protein kinase (LRR-STK) was found to be one of the most highly upregulated gene in the study, thus indicating its possible involvement in inducing defense responses against the pathogen. The gene was highly upregulated at all the time points, however, the expression showed a decreasing trend. Several studies in bananas have reported differential upregulation of LRR receptor-like serine/threonine-protein kinases with higher expression levels in the resistant cultivar compared to the susceptible in response to infection by *Foc* (Niu et al. 2018), *M. musicola* (Passos et al. 2013) and *X. campestris* cv. *musacearum* (Tripathi et al. 2019) establishing the role of these LRR kinases in mediating pathogen resistance.

The plant O-methyltransferases (OMTs) enzymes, involved in the biosynthesis of several compounds essential for plant defense and growth, methylate the oxygen atom (O-methylation reaction) of plant defense-related secondary metabolites such as alkaloids, flavonoids, phenylpropanoids etc., thus participating in lignification, stress tolerance and disease resistance (Lam et al. 2007). In pea, alfalfa, medicago and chickpea, the O-methylation of isoflavonoids leads to the biosynthesis of defense chemicals called ‘phytoalexins’, which are essential for disease resistance (Liu et al. 2006). Yang et al. (2017) have reported that a gene coding for a caffeoyl-CoA O-methyltransferase (*ZmCCoAOMT2*) is associated with the phenylpropanoid pathway and lignin production confers quantitative resistance to southern leaf blight and gray leaf spot disease in maize. The present investigation also establishes a possible role of an OMT in conferring resistance to *P. musae* infection in the cultivar Kachkal, since it was found to be highly and significantly upregulated at 48 hpi as well as 72 hpi. Our findings are consistent with several other studies that have demonstrated that OMTs are differentially expressed in response to infection by fungal pathogens such as *Foc* TR4 (in banana) (Zhang et al. 2019) and *Sclerotinia sclerotiorum* (in oilseed rape) (Jian et al. 2018) within 48–72 hpi.

Germin-like proteins (GLPs) have a crucial role in basal resistance and are induced due to pathogen infection or application of disease resistance-associated chemicals such as H_2O_2 , SA, and ethylene (Zimmermann et al. 2006). These proteins function as cofactors of plant cell wall reinforcement through facilitating cross-linking of the cell wall. This leads to the generation of H_2O_2 which in turn activates the SA/JA signaling cascades leading to the expression of defense genes and induction of SAR. Several studies have established the involvement of GLPs in enhancing defense responses and broad spectrum disease resistance to many pathogens (Breen and Bellgard 2010; Liu et al. 2016). The expression study revealed that in Kachkal, a GLP has

a potential role in conferring resistance to *P. musae* as it was significantly upregulated during 48–72 hpi. Similar expression patterns of induction of GLP at 72 hpi were also reported by Lanubile et al. (2015) in a soybean cultivar resistant to *F. oxysporum*. However, a study by Tripathi et al. (2019) on disease resistance of *M. balbisiana* against banana *Xanthomonas* wilt disease, found that GLP was activated in both BXW-susceptible and BXW-resistant genotypes at 12 hpi and 48 hpi, respectively.

The present study, thus, reveals that a number of genes and transcription factors are induced in *M. paradisiaca* immediately after infection by *P. musae* which are important for pathogen perception and activation of defense responses in the host. From the qRT-PCR results, the genes such as PR-like, GLP, OMT and LRR-STK and the transcription factors ERF and WRKY were observed to be highly induced within 24–48 hpi. Moreover, their expression was found to remain consistently high till the last time point under study, thus indicating their roles in defending the fungus attack. Therefore, these genes can be further studied to be used as potential candidates for engineering *P. musae* resistant banana varieties using biotechnological approaches.

Conclusion

The study aimed to investigate the defense-related transcriptomic responses triggered in the resistant cultivar *M. paradisiaca* var. Kachkal in response to infection by the Sigatoka pathogen *P. musae*, for the first time. We further envisaged identifying a set of defense-related genes modulated as a result of the infection with the hypothesis that their expression should be relatively high during the initial stages of infection (24 and 48 hpi) to provide the host plant a long-lasting and broad spectrum of resistance. The transcriptomic as well qRT-PCR analysis established the fact that a resistant host tries to initiate and establish defense response by deploying an array of genes and transcription factors involved in a wide range of biological processes and pathways. The results also indicate that the early defense-related genes identified from the study are important not only for pest/pathogen recognition and perception but also for activation and persistence of defense in *M. paradisiaca*. These genes would serve as important potential candidates for future banana genetic improvement programs for the development of disease-resistant varieties.

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s13205-022-03245-9>.

Acknowledgements Science & Engineering Research Board (SERB), Department of Science & Technology, Government of India, for funding the work; Department of Agricultural Biotechnology and DBT-NECAB, Assam Agricultural University, Jorhat, for providing the

infrastructural facilities; Dr. Pranab Dutta, Department of Plant Pathology, AAU, Jorhat, for assistance in works related to pathogen isolation, identification and inoculation.

Declarations

Conflict of interest The authors declare that they have no conflict of interest.

References

- Aamir M, Singh VK, Dubey MK, Kashyap SP, Zehra A, Upadhyay RS, Singh S (2018) Structural and functional dissection of differentially expressed tomato WRKY transcripts in host defense response against the vascular wilt pathogen (*Fusarium oxysporum* f. sp. *lycopersici*). PLoS One. <https://doi.org/10.1371/journal.pone.0193922>
- AbuQamar S, Chen X, Dahwan R, Bluhm B, Salmeron J, Lam S, Dietrich RA, Mengiste T (2006) Expression profiling and mutant analysis reveals complex regulatory networks involved in *Arabidopsis* response to *Botrytis* infection. Plant J 48:28–44
- Arzanlou M, Abeln E, Kema G, Waalwijk C, Carlier J, de Vries I, Guzmán M, Crous P (2007) Molecular diagnostics for the sigatoka disease complex of banana. Phytopathol 97:1112–1118
- Bai TT, Xie WB, Zhou PP, Wu ZL, Xiao WC, Zhou L, Sun J, Ruan XL, Li HP (2013) Transcriptome and expression profile analysis of highly resistant and susceptible banana roots challenged with *Fusarium oxysporum* f. sp. *cubense* tropical race 4. PLoS One 8(9):73945
- Berrocal-Lobo M, Molina A, Solano R (2002) Constitutive expression of ethylene-response-factor1 in *Arabidopsis* confers resistance to several necrotrophic fungi. Plant J 29:23–32
- Breen J, Bellgard M (2010) Germin-like proteins (GLPs) in cereal genomes: gene clustering and dynamic roles in plant defence. Funct Integr Genomics 10(4):463–476
- Carlier J, Zapater M, Lapeyre F, Jones D, Mourichon X (2000) Septoria leaf spot of banana: a newly discovered disease caused by *Mycosphaerella eumusae* (Anamorph *Septoria eumusae*). Phytopathol 90:884–890
- Claudianos C, Campbell HD (1995) The novel flightless-I gene brings together two gene families, actin-binding proteins related to gelsolin and leucine-rich-repeat proteins involved in Ras signal transduction. Mol Biol Evol 12:405–414
- Elvira MI, Molina Galdeano M, Gilardi P, García-Luque I, Serra MT (2008) Proteomic analysis of pathogenesis-related proteins (PRs) induced by compatible and incompatible interactions of pepper mild mottle virus (PMMoV) in *Capsicum chinense* L3 plants. J Exp Bot 59:1253–1265
- Gu YQ, Wildermuth MC, Chakravarthy S, Loh YT, Yang C, He X, Han Y, Martin GB (2002) Tomato transcription factors Pti4, Pti5, and Pti6 activate defense responses when expressed in *Arabidopsis*. Plant Cell 14(1):817–831
- Hamel LP, Nicole MC, Sritubtim S, Morency MJ, Ellis M, Ehrling J, Beaudoin N, Barbazuk B, Klessig D, Lee J, Martin G, Mundy J, Ohashi Y, Scheel D, Sheen J, Xing T, Zhang S, Seguin A, Ellis BE (2006) Ancient signals: comparative genomics of plant MAPK and MAPKK gene families. Trends Plant Sci 11(4):192–198
- Heslop-Harrison P, Schwarzacher T (2007) Domestication, genomics and the future for banana. Ann Bot 100:1073–1084
- Jia C, Wang Z, Wang J, Miao H, Zhang J, Xu B, Liu J, Jin Z, Liu J (2022) Genome-wide analysis of the banana WRKY transcription factor gene family closely related to fruit ripening and stress. Plants (Basel) 11(5):662

- Jian H, Ma J, Wei L, Liu P, Zhang A, Yang B, Li J, Xu X, Liu L (2018) Integrated mRNA, sRNA, and degradome sequencing reveal oilseed rape complex responses to *Sclerotinia sclerotiorum* (Lib.) infection. *Sci Rep* 8:10987
- Johanson A, Jeger MJ (1993) Use of PCR for detection of *Mycosphaerella fijiensis* and *M. musicola*, the causal agents of Sigatoka leaf spots in banana and plantain. *Mycol Res* 96:670–674
- Jones DR (ed) (2000) Diseases of banana, Abaca and Enset. CAB International, CABI Publications, Wallingford
- Kayum MA, Jung HJ, Park JI, Ahmed NU, Saha G, Yang TJ, Nou IS (2014) Identification and expression analysis of WRKY family genes under biotic and abiotic stresses in *Brassica rapa*. *Mol Genet Genomics* 290(1):79–95
- Lam KC, Ibrahim RK, Behdad B, Dayanandan S (2007) Structure, function, and evolution of plant O-methyltransferases. *Genome* 50(11):1001–1013
- Lanubile A, Muppirala UK, Severin AJ, Marocco A, Munkvold GP (2015) Transcriptome profiling of soybean (*Glycine max*) roots challenged with pathogenic and non-pathogenic isolates of *Fusarium oxysporum*. *BMC Genomics* 16:1089
- Li Y, Guo G, Zhou L, Chen Y, Zong Y, Huang J, Lu R, Liu C (2019) Transcriptome analysis identifies candidate genes and functional pathways controlling the response of two contrasting barley varieties to powdery mildew infection. *Int J Mol Sci* 21(1):151
- Liu CJ, Deavours BE, Richard SB, Ferrer JL, Blount JW, Huhman D, Dixon RA, Noel JP (2006) Structural basis for dual functionality of isoflavonoid O-methyltransferases in the evolution of plant defense responses. *Plant Cell* 18(12):3656–3669
- Liu Q, Yang J, Yan S, Zhang S, Zhao J, Wang W, Yang T, Wang X, Mao X, Dong J, Zhu X, Liu B (2016) The germin-like protein OsGLP2-1 enhances resistance to fungal blast and bacterial blight in rice. *Plant Mol Biol* 92:411–423
- Mao G, Meng X, Liu Y, Zheng Z, Chen Z, Zhang S (2011) Phosphorylation of a WRKY transcription factor by two pathogen-responsive MAPKs drives phytoalexin biosynthesis in *Arabidopsis*. *Plant Cell* 23(4):1639–1653
- McGrath KC, Dombrecht B, Manners JM, Schenk PM, Edgar CI, Maclean DJ, Scheible WR, Udvardi MK, Kazan K (2005) Repressor- and activator type ethylene response factors functioning in jasmonate signaling and disease resistance identified via a genome-wide screen of *Arabidopsis* transcription factor gene expression. *Plant Physiol* 139(2):949–959
- Mishra AK (2000) Epidemiology and management of Sigatoka disease of banana. Thesis, Assam Agricultural University, Jorhat
- Mizoi J, Shinozaki K, Yamaguchi-Shinozaki K (2012) AP2/ERF family transcription factors in plant abiotic stress responses. *Biochim Biophys Acta* 1819:86–96
- Mustaffa MM, Sathiamoorthy S (2004) Current status of banana research in India. In: Molina AB, Eusebio JE, Roa VN, Van Den Bergh I, Maghuyop MAG, Borromeo KH (eds) Advancing banana and plantain R and D in Asia and the Pacific, vol 12. International Network for the Improvement of Banana and Plantain, Philippines, pp 65–80
- Negi S, Tak H, Ganapathi TR (2016) Expression analysis of MusaNAC68 transcription factor and its functional analysis by overexpression in transgenic banana plants. *Plant Cell Tiss Organ Cult (PCTOC)* 125:59–70
- Negi S, Tak H, Ganapathi TR (2018) A banana NAC transcription factor (MusaSNAC1) impart drought tolerance by modulating stomatal closure and H₂O₂ content. *Plant Mol Biol* 96(4–5):457–471
- Negi S, Tak H, Ganapathi TR (2021) Overexpression of MusaSNAC1 improves shoot proliferation in transgenic banana lines. *3 Biotech* 11(4):188
- Niu Y, Hu B, Li X, Chen H, Takac T, Samaj J, Xu C (2018) Comparative digital gene expression analysis of tissue-cultured plantlets of highly resistant and susceptible banana cultivars in response to *Fusarium oxysporum*. *Int J Mol Sci* 19(2):350
- Passos MAN, de Cruz VO, Emediato FL, de Teixeira CC, Azevedo VCR, Brasileiro ACM, Amorim EP, Ferreira CF, Martins NF, Togawa RC, Pappas Junior GJ, de Silva Jr OB, Miller RNG (2013) Analysis of the leaf transcriptome of *Musa acuminata* during interaction with *Mycosphaerella musicola*: gene assembly, annotation and marker development. *BMC Genomics* 14:78
- Portal O, Izquierdo Y, de Vleeschauwer D, Sanchez-Rodriguez A, Mendoza-Rodriguez M, Acosta Suarez M, Ocana B, Jimenez E, Hofte M (2011) Analysis of expressed sequence tags derived from a compatible *Mycosphaerella fijiensis*–banana interaction. *Plant Cell Rep* 30:913–928
- Rodriguez HA, Rodriguez-Arango E, Morales JG, Kema G, Arango RE (2016) Defense gene expression associated with biotrophic phase of *Mycosphaerella fijiensis* *M. Morelet* infection in banana. *Plant Dis* 100:1170–1175
- Selvarajan R, Uma S, Sathiamoorthy S (2001) Etiology and survey of banana leaf spot diseases in India. In Advancing banana and plantain R & D in Asia and the Pacific (eds. Molina AB, Roa VN, Maghuyop MAG), Proceedings of the 10th INIBAP-ASP-NET Regional Advisory Committee meeting held at Bangkok 10:94–102
- Simmonds NW, Shepherd K (1955) The taxonomy and origins of the cultivated bananas. *J Linn Soc* 55:302–312
- Sun J, Zhang J, Fang H, Peng L, Wei S, Li C, Zheng S, Lu J (2019) Comparative transcriptome analysis reveals resistance-related genes and pathways in *Musa acuminata* banana “Guijiao 9” in response to Fusarium wilt. *Plant Physiol Biochem* 141:83–94
- Tak H, Negi S, Ganapathi TR (2017) Banana NAC transcription factor MusaNAC042 is positively associated with drought and salinity tolerance. *Protoplasma* 254:803–816
- Tak H, Negi S, Gupta A, Ganapathi TR (2018) A stress associated NAC transcription factor MpSNAC67 from banana (*Musa x paradisiaca*) is involved in regulation of chlorophyll catabolic pathway. *Plant Physiol Biochem* 132:61–71
- Tak H, Negi S, Rajpurohit YS, Misra HS, Ganapathi TR (2020) MusaMPK5, a mitogen activated protein kinase is involved in regulation of cold tolerance in banana. *Plant Physiol Biochem* 146:112–123
- Tak H, Negi S, Ganapathi TR (2021) The 5'-upstream region of WRKY18 transcription factor from banana is a stress inducible promoter with strong expression in guard cells. *Physiol Plant* 173(4):1335–1350
- Tripathi L, Tripathi JN, Shah T, Muiruri KS, Katari M (2019) Molecular basis of disease resistance in banana progenitor *Musa balbisiana* against *Xanthomonas campestris* pv. *musacearum*. *Sci Rep* 9(1):7007
- van Loon LC, van Strien EA (1999) The families of pathogenesis related proteins, their activities, and comparative analysis of PR-1 type proteins. *Physiol Mol Plant Pathol* 55(2):85–97
- van Loon LC, Rep M, Pieterse CM (2006) Significance of inducible defense-related proteins in infected plants. *Annu Rev Phytopathol* 44:135–162
- Wang L, Hu W, Tie W, Ding Z, Ding X, Liu Y, Yan Y, Wu C, Peng M, Xu B, Jin Z (2017) The MAPKKK and MAPKK gene families in banana: identification, phylogeny and expression during development, ripening and abiotic stress. *Sci Rep* 7:1159
- Wu QB, Xu LP, Guo JL, Su YC, Que YX (2013) Transcriptome profile analysis of sugarcane responses to *Sporisorium scitaminea* infection using Solexa sequencing technology. *Biomed Res Int*. <https://doi.org/10.1155/2013/298920>
- Xu X, Chen C, Fan B, Chen Z (2006) Physical and functional interactions between pathogen-induced *Arabidopsis* WRKY18, WRKY40 and WRKY60 transcription factors. *Plant Cell* 18:1310–1326

- Yan H, Jiang G, Wu F, Li Z, Xiao L, Jiang Y, Duan X (2021) Sulfoxidation regulation of transcription factor NAC42 influences its functions in relation to stress-induced fruit ripening in banana. *J Exp Bot* 72(2):682–699
- Yang Q, He Y, Kabahuma M, Chaya T, Kelly A, Borrego E, Bian Y, El Kasmi F, Yang L, Teixeira P, Kolkman J, Nelson R, Kolomiets ML, Dangl J, Wissner R, Caplan J, Li X, Lauter N, Balint-Kurti P (2017) A gene encoding maize caffeoyl-CoA O-methyltransferase confers quantitative resistance to multiple pathogens. *Nat Genet* 49:1364–1372
- Yang Y, Li X, Kan B, He H, Li T, Ding Y, Du P, Lai W, Hu H, Huang J (2021) Transcriptome analysis reveals MYB and WRKY transcription factors involved in banana (*Musa paradisiaca* AA) magnesium deficiency. *Planta* 254(6):115
- Zhang S, Klessig DF (2001) MAPK cascades in plant defense signaling. *Trends Plant Sci* 6(11):520–527
- Zhang F, Zhu L, He G (2004) Differential gene expression in response to brown planthopper feeding in rice. *J Plant Physiol* 161:53–62
- Zhang J, Broeckling CD, Sumner LW, Wang ZY (2007) Heterologous expression of two *Medicago truncatula* putative ERF transcription factor genes, WXP1 and WXP2, in Arabidopsis led to increased leaf wax accumulation and improved drought tolerance, but differential response in freezing tolerance. *Plant Mol Biol* 64:265–278
- Zhang L, Cenci A, Rouard M, Zhang D, Wang Y, Tang W, Zheng SJ (2019) Transcriptomic analysis of resistant and susceptible banana corms in response to infection by *Fusarium oxysporum* f. sp. *cubense* tropical race 4. *Sci Rep* 9:8199
- Zheng Z, Qamar SA, Chen Z, Mengiste T (2006) Arabidopsis WRKY33 transcription factor is required for resistance to necrotrophic fungal pathogens. *Plant J* 48(4):592–605
- Zhu XL, Qi L, Liu X, Cai SB, Xu HJ, Huang RF, Li JR, Wei XN, Zhang ZY (2014) The wheat ethylene response factor transcription factor PATHOGEN-INDUCED ERF1 mediates host responses to both the necrotrophic pathogen *Rhizoctonia cerealis* and freezing stresses. *Plant Physiol* 164:1499–1514
- Zimmermann G, Baumlein H, Mock H-P, Himmelbach A, Schweizer P (2006) The multigene family encoding germin-like proteins of barley. Regulation and function in basal host resistance. *Plant Physiol* 142(1):181–92