Plant Pathology

WILEY

Plant Pathology

ORIGINAL ARTICLE OPEN ACCESS

The Transcriptome Response of African and South American Cassava (*Manihot esculenta*) to Infection by East African Cassava Mosaic Virus-Uganda

Erica Ngwensang | Sophie Bouvaine | Midatharahally N. Maruthi

Agriculture, Health and Environment Department, Natural Resources Institute, University of Greenwich, Chatham, UK

Correspondence: Midatharahally N. Maruthi (m.n.maruthi@greenwich.ac.uk)

Received: 24 October 2024 | **Revised:** 29 January 2025 | **Accepted:** 4 February 2025

Funding: This work was supported by the African Union Commission, AURG/2/141 and AURG II-1-060-, 2016.

Keywords: cassava | differentially expressed genes | EACMV-Ug | resistance | RNA sequencing

ABSTRACT

Cassava mosaic begomoviruses (CMBs) cause the economically important cassava mosaic disease (CMD) in cassava. In this research, we investigated potential new sources of resistance to CMD. Fourteen cassava varieties were tested for resistance to the major CMB, East African cassava mosaic virus-Uganda (EACMV-Ug). Six African cassava varieties (72-TME 14, TME 204, TZ 130, Nase 1, Nase 3 and Nase 14) expressed no or mild disease symptoms with low viral load and therefore have been classified as resistant. South American varieties showed severe leaf symptoms and supported high viral load. Nase 3 (resistant), TZ 130 (resistant) and Ebwanateraka (susceptible) were subjected to RNA-sequencing (RNA-Seq) to identify putative CMD resistance genes and mechanisms. Transcriptome analysis of three cassava varieties at 2, 4, 7, and 28 days after grafting (DAG) revealed the largest number of differentially expressed genes (DEGs) upon virus infection in Nase 3 (4228) with 1725 genes uniquely over-expressed. Additionally, early induction of heat shock proteins and transcription factors such as ethylene-responsive transcription factor (ERF), teosinte-like, cycloidea and PCF1 (TCPs), heat stress transcription factor and basic leucine zipper were observed in the resistant varieties. These results suggest that resistant varieties maintain a low titre of EACMV-Ug by activating specific stress-response genes early. The role of these genes in plants remains to be investigated but offers insights into the molecular mechanism of resistance to CMD.

1 | Introduction

Cassava (*Manihot esculenta*; Euphorbiaceae) is an important food security crop for over 450 million people in Africa, and its production is threatened by cassava mosaic disease (CMD) (Legg et al. 2015). At least 11 distinct cassava mosaic begomoviruses (CMBs) cause CMD, which can only be effectively controlled by growing resistant varieties (Legg et al. 2011). One of the strains of CMB is East African cassava mosaic virus-Uganda (EACMV-Ug), which is a member of the genus *Begomovirus*, family *Geminiviridae*, based on the guidelines of the International Committee on Taxonomy of Viruses (ICTV) in 2022. EACMV-Ug is a recombinant strain that has been associated with the CMD pandemic in eastern Africa since the late 1980s (Legg et al. 2006). Regardless of the virus species, symptoms range from green to yellow mosaic together with leaf deformation. Early infection of susceptible varieties causes severe stunting and a reduction in yield (Alabi et al. 2011). The most sustainable way to control CMD is by growing resistant varieties, and several resistant varieties have been identified (Legg et al. 2011; Nicaise 2014). However, the mechanisms underlying the resistance have been poorly understood.

Transcriptome sequencing by RNA-sequencing (RNA-Seq) is one of the ways of understanding the molecular processes

underlying disease resistance by identifying changes in gene expression during viral infections (Fan et al. 2015; Nicaise 2014; Stobbe and Roossinck 2014). By comparing susceptible and resistant varieties, RNA-Seq identifies gene expression changes during the defence response and reveals complex resistance mechanisms. Few RNA-Seq studies have been performed to evaluate the response of cassava varieties challenged with cassava viruses. Transcriptome analysis has been conducted in cassava following inoculation with a different strain of CMB, South African cassava mosaic virus (SACMV), a member of the genus *Begomovirus*. Allie et al. (2014) compared transcripts of cassava genotypes and showed that 4181 and 1008 transcripts in total were differentially expressed in response to SACMV in susceptible (T200) and tolerant (TME3) varieties, respectively. GOslim functional group demonstrated that in T200 and TME3, differentially expressed genes were over-represented in the cellular component category of plasma membrane and nucleus. Alterations in the expression of transcription factors (TFs), resistance (R) genes and histone/DNA methylation associated genes were observed (Allie et al. 2014). Transcriptome studies have also been conducted following inoculation with two RNA virus species, cassava brown streak virus (CBSV) and Ugandan cassava brown streak virus (UCBSV) (Patil et al. 2015; Tomlinson et al. 2018), together known as cassava brown streak ipomoviruses (CBSIs) belonging to family *Potyviridae*, genus *Ipomovirus,* which cause cassava brown streak disease (CBSD) (Maruthi et al. 2017). RNA-Seq data identified the over-expression of phenylalanine ammonia-lyase 1 and 2 (PAL1, PAL2), cinnamic acid and two chalcone synthase genes in a resistant variety Kaleso, and silencing of PAL1 by RNA interference led to increased susceptibility to CBSD (Kavil et al. 2021). Other defence-related genes such as NAM, ATAF and CUC (NAC) transcription factors and elongation factor eIF(iso)4E were over-expressed after CBSV infection (Maruthi et al. 2014). In Kaleso, several defence response genes, such as genes encoding leucine-rich repeats (LRR), nucleotide binding domain (NBARC), pathogenesis-related, late embryogenesis abundant, selected transcription factors, chaperones and heat shock proteins, were highly over-expressed upon UCBSV infection (Amuge et al. 2017). Defence-related GO terms of translational elongation, translation factor activity, ribosomal subunit and phosphorelay signal transduction were also over-represented (Amuge et al. 2017). The goal of this study was to identify genes that respond early to infection as well as genes modulated during virus replication and movement. Here, we characterised the transcriptomic response of EACMV-Ug-resistant, tolerant and susceptible cassava varieties between 2 and 24 days after infection to characterise their defence response.

2 | Materials and Methods

2.1 | Planting Materials and Graft Inoculation

Cassava varieties were obtained from Latin America (supplied by the International Center for Tropical Agriculture, CIAT-Colombia) and from the National Agricultural Research Systems (NARS) of Kenya and Uganda (Table 1) (Maruthi et al. 2019). For each variety, 15 virus-free cuttings of 10 cm length were planted and grown in glasshouse conditions of 28°C ± 5°C and 50%–60% relative humidity (RH). The resistant variety Nase 3 and the susceptible variety Col 2246 were used as controls.

Virus inoculum was obtained from variety Ebwanateraka exhibiting severe symptoms, and end-point PCR using ACMV-CP/R3 and UV-AL1/F1 primers (Zhou et al. 1997) confirmed the presence of EACMV-Ug (product 1.6 kb). At 12 weeks after planting, four plants per variety were virus-inoculated by side grafting

TABLE 1 | East African and Latin American cassava varieties screened for resistance to cassava mosaic disease (CMD).

Geographical origin	Variety	Resistance to CMD	References
Kenya	Kibandameno	Susceptible	Maruthi et al. (2019), Tumwegamire et al. (2018)
Uganda	72-TME 14	Resistant	Tumwegamire et al. (2018)
	Ebwanateraka	Susceptible	Otim-Nape et al. (1997)
	Nase 1	Resistant	Maruthi et al. (2019), Tumwegamire et al. (2018)
	Nase 3	Tolerant	Maruthi et al. (2019), Tumwegamire et al. (2018)
	Nase 14	Resistant	Maruthi et al. (2019), Tumwegamire et al. (2018)
	TME 204	Resistant	Maruthi et al. (2019), Tumwegamire et al. (2018)
	TZ 130	Resistant	Maruthi et al. (2019), Tumwegamire et al. (2018)
South America	Col 2246	Susceptible	
	PER 317	Unknown	
	PER 335	Unknown	
	PER 368	Unknown	
	PER 415	Unknown	
	PER 608	Unknown	

using EACMV-Ug scions or healthy scions (mock-inoculated) as control (Mohammed et al. 2012).

2.2 | Symptom Scoring and Sample Collection

Leaf symptom severity was scored weekly on all plants using the five-point scale (Beyene et al. 2016), where 0 represents no disease symptoms and 5 represents very severe mosaic on leaves affecting 50%–80% of the leaf area. For enhancing symptom expression, the plants were cut back to 20 cm above soil level at 16 weeks after grafting (WAG) and all leaves were removed. Symptom development was assessed on newly formed leaves after 4 weeks using the same scale.

In this time series experiment, samples were collected at 1, 2, 4 and 7 days after grafting (DAG) and 2, 4, 8, 12, 16 and 24 WAG from EACMV-Ug- and mock-inoculated plants (10 time points × 14 varieties × 3 plants × 2 treatments). For each sample, three leaf discs were collected from the top, middle and bottom of each plant, but avoiding the scions. Samples were frozen immediately in liquid nitrogen before storage at −80°C. For each time point, three biological replicates were processed independently for each treatment and mock-inoculated plants.

2.3 | Total Nucleic Acid Extraction and Real-Time PCR

Total nucleic acid was extracted from samples using the cetyltrimethylammonium bromide (CTAB) method (Abarshi et al. 2010; Otti et al. 2016). Pellets were suspended in 100 μL water, total nucleic acid concentration was determined using a NanoDrop 2000 spectrometer (Thermo Fisher Scientific) and samples were stored at −20°C.

EACMV-Ug titre was measured in each sample by real-time quantitative PCR (qPCR) in duplicate using primers and probes (Otti et al. 2016) with a minor modification of the EACMV probe, which contained no secondary quencher. Co-amplification of EACMV-Ug with the endogenous target PP2A (Moreno et al. 2011) in a single reaction was done to measure relative virus titres. To assess the effect of multiplexing PP2A and EACMV-Ug, C_q values from both multiplex and uniplex qPCR assays were compared and the Wilcoxon test was used to test the significant change in rising C_q values after multiplexing. Samples with C_q values > 35 were considered negative and relative quantities of positive samples were calculated using the $2^{-\Delta\Delta Cq}$ method with *PP2A* as an internal housekeeping gene (Abarshi et al. 2010).

2.4 | Data Analysis

Varieties were classified into three categories: resistant, tolerant or susceptible, based on foliar severity and relative virus quantity over time (Fargette et al. 1996; Ogbe et al. 2003). Resistant plants showed little or no foliar symptoms with low virus titre, tolerant plants expressed foliar symptoms with low virus titre and susceptible plants showed severe foliar symptoms with high virus concentrations (Kuria et al. 2017). Analysis of variance was also carried out on means of all replicates for CMD leaf symptom severity at 2, 4, 8, 12, 16 and 24 WAG, and their average score was compared to the resistant control Nase 3, which is currently one of the best sources of resistance to CMD (Thresh et al. 1994). An agglomerative hierarchical clustering tree was generated using average disease severity and virus quantities to classify the varieties (Houngue et al. 2019).

2.5 | RNA-Seq and Transcriptome Analysis

Leaf samples collected at 2, 4, 7 and 28 DAG from each EACMV-Ug- and mock-inoculated plant were subjected to RNA-Seq. Total RNA was isolated using the CTAB method (Maruthi et al. 2002) and lithium chloride with slight modifications. Flash-frozen leaf tissues (18 leaf discs) were ground in liquid nitrogen before being transferred into 1 mL of pre-heated CTAB buffer and mixed. Samples were heated at 60°C for 10 min and approximately 800 μL of sample mix was transferred to an equal volume of phenol:chloroform:isoamyl alcohol (25:24:1) before being centrifuged at 21,129 *g* for 10 min at 4°C. The top aqueous phase was transferred to an equal volume of chloroform, mixed by inversion and centrifuged at 21,129 *g* for 20 min at 4°C. The RNA was then precipitated from the supernatant by adding 0.3 volumes of 8 M LiCl and incubated at 4°C overnight. Samples were then centrifuged at 21,129 *g* for 10 min at 4°C to pellet the RNA. The pellet was washed with 500 μL of 70% ethanol, followed by centrifuging at 18,218 *g* for 5 min at 4°C before being dried for 10 min at low temperature in a spin vac and resuspended in 30 μL of RNase-free water. DNA was removed using the DNase Max kit QuickStart protocol (Qiagen). RNA quantities were measured using a NanoDrop spectrophotometer (Thermo Fisher Scientific). The RNA integrity was then confirmed using the Agilent 2100 Bioanalyzer instrument according to the manufacturer's protocol (Agilent Technologies). cDNA libraries and RNA-Seq were performed by Macrogen (Europe) for generating 100 bp paired-end reads using the Illumina NovaSeq next-generation sequencing system.

Paired-end reads from samples were aligned to the *M. esculenta* v6.1 reference genome downloaded from Phytozome 12 (https://phytozome.jgi.doe.gov/pz/portal.html) with HISAT2 (https://www.ccb.jhu.edu/software/hisat/manual.shtml) (Prochnik et al. 2012). Samtools (v. 1.9) was used to convert and sort HISAT2 output into sorted bam files, which were then assembled into genes using Cufflinks (Trapnell et al. 2012). Assemblies were merged into one file using the Cuffmerge utility program, with the Mesculenta_305_v6.1.gene.gff3.gz file serving as a reference annotation.

Expression values were normalised as fragments per kilobase per million reads (FPKM) values. Cuffdiff was used to identify differentially expressed genes (DEGs) in pairwise comparisons between control and virus-infected samples for each variety and time points. To reduce artefacts, genes with FPKM values ≤ 1 in all samples were discarded from the analysis. Genes having $\log_2$ fold change ≥ 1 and $p \leq 0.05$ were considered up-regulated, whereas those with fold change ≤ −1 and $p \leq 0.05$ were deemed to be down-regulated upon EACMV-Ug infection.

DEGs were queried against the *M. esculenta* genome database at g:Profiler to determine functional annotation. Functional enrichment analysis was performed using g:-Profiler (version e98_eg45_p14_ce5b097) with g:SCS multiple testing correction methods applying a significance threshold of 0.05 (Raudvere et al. 2019).

3 | Results

3.1 | Varieties Response to Virus Infection

About 95% of all cassava plants grafted established a successful graft union between 3 and 4 WAG. Early mild symptoms (score 2) were identified on Col 2246 at 8, 12 and 16 WAG, followed by PER 608 at 16 WAG and all other varieties remained symptomless. There was a significant difference between the CMD severity scores of the 14 varieties at 12 WAG ($df = 13$, $p < 0.0001$).

At 24 WAG, five varieties inoculated with EACMV-Ug remained symptomless (72-TME 14, Nase 1, Nase 3, Nase 14 and TZ 130) (Figures 1 and 2), TME 204 had a severity score of 0–2 and the other eight varieties had a severity score of 2–5 (Col 2246, Ebwanateraka, Kibandameno, PER 317, PER 335, PER 368, PER 415 and PER 608) (Figures 1 and 2). PER 415, PER 335, PER 317, PER 368, PER 335, Ebwanateraka and Kibandameno each had a score of 3 and greater at 24 WAG after pruning (Figures 1 and 2).

Three clusters were identified, with Group 1 having six varieties (72-TME 14, Nase 1, Nase 3, Nase 14, TZ 130 and TME 204), and Groups 2 and 3 having eight varieties (Col 2246, Ebwanateraka, Kibandameno, PER 317, PER 335, PER 368, PER 415 and PER 608). Group 1 contained varieties with a severity score of 0–2 and was considered resistant/tolerant (R/T). Groups 2 and 3 varieties had a severity score of 3–5 and were considered susceptible (S) (Figure 2).

3.2 | Quantification of EACMV-Ug Using qPCR

The Wilcoxon signed rank test showed that multiplexing did not significantly change the ΔC_q value at a threshold of $p = 0.05$; the multiplex qPCR test was then used for the rest of the study. The virus was detected in all varieties and virus titre fold change was calculated using Col 2246 at 2 WAG as a sample calibrator. The virus load was relatively low at 8 WAG in Ebwanateraka (0.1 fold change) and PER 368 at 12 WAG (4.3 fold change). It increased to 1390 and 6894 fold change in Ebwanateraka and PER 368, respectively, at 24 WAG. In Col 2246, the virus quantity increased to the highest level at 8 WAG and decreased at 16 WAG (1865 to 584 fold change) before rising to 1409 fold change at 24 WAG. Varieties with low relative virus load included 72-TME 14 (1.1 fold change), Nase 1 (0.9 fold change), Nase 3 (1.5 fold change), Nase 14 (0.1 fold change), TME 204 (16.1 fold change) and TZ 130 (3.7 fold change). The virus concentration in PER 317 and PER 368 was much higher at 24 than at 16 WAG, while it was similar at 16 and 24 WAG in Col 2246 and PER 608 (Figure 3). Spearman's rank correlation rho analysis showed a positive correlation for Col 2246 (0.59), Ebwanateraka (0.77), Kibandameno (0.66), PER 317 (0.81), PER 335 (0.76), PER 368 (0.99), PER 415 (0.99), PER 608 (0.71) and TME 204 (0.99) between virus titre and mean foliar symptom score, while 72-TME 14, Nase 1, Nase 3, Nase 14 and TZ 130 had relatively low virus loads with no foliar severity at 24 WAG (Figure 3).

3.3 | Transcriptome Analysis and GO Term Enrichment of DEGs

Transcriptome analyses of Ebwanateraka (susceptible), Nase 3 and TZ 130 (resistant) produced 700 million reads generated from 67 samples. FPKM values were generated for each gene in the cassava reference genome. Approximately 70% of the reads were uniquely mapped to the genome, which represent 17,000 genes expressed with FPKM value ≥1 in each sample (Tables S1 and S2). DEGs were quantified between virus- and

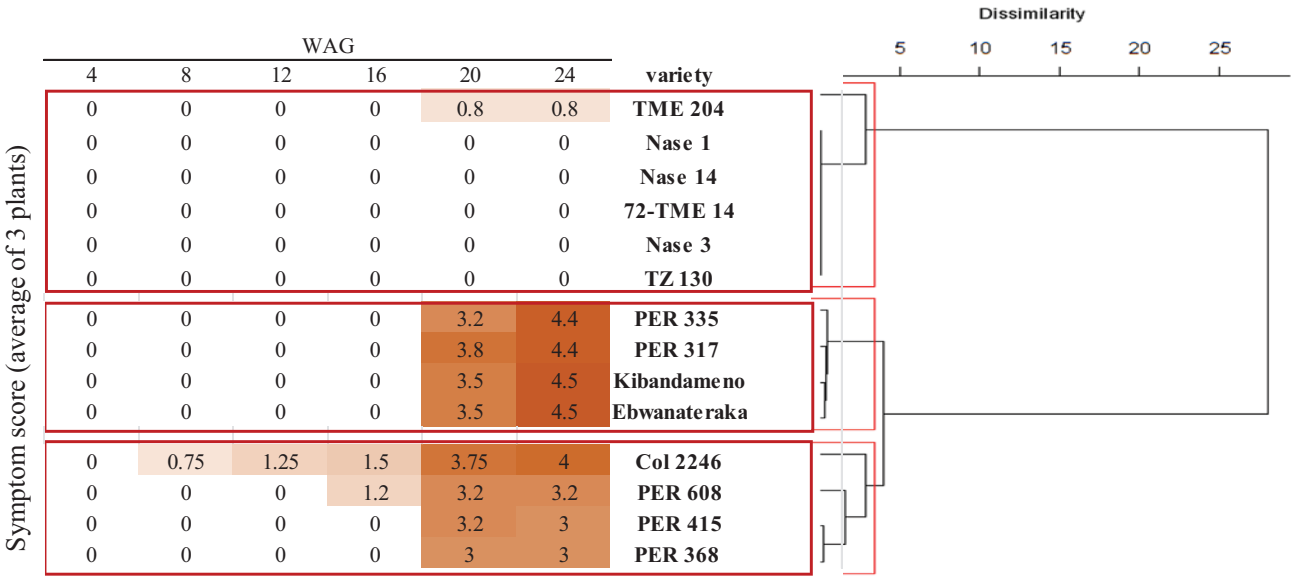

FIGURE 1 | Symptom recording of cassava varieties from 4 to 24 weeks after grafting (WAG) with EACMV-Ug (left) and hierarchical clustering (right) of symptom values.

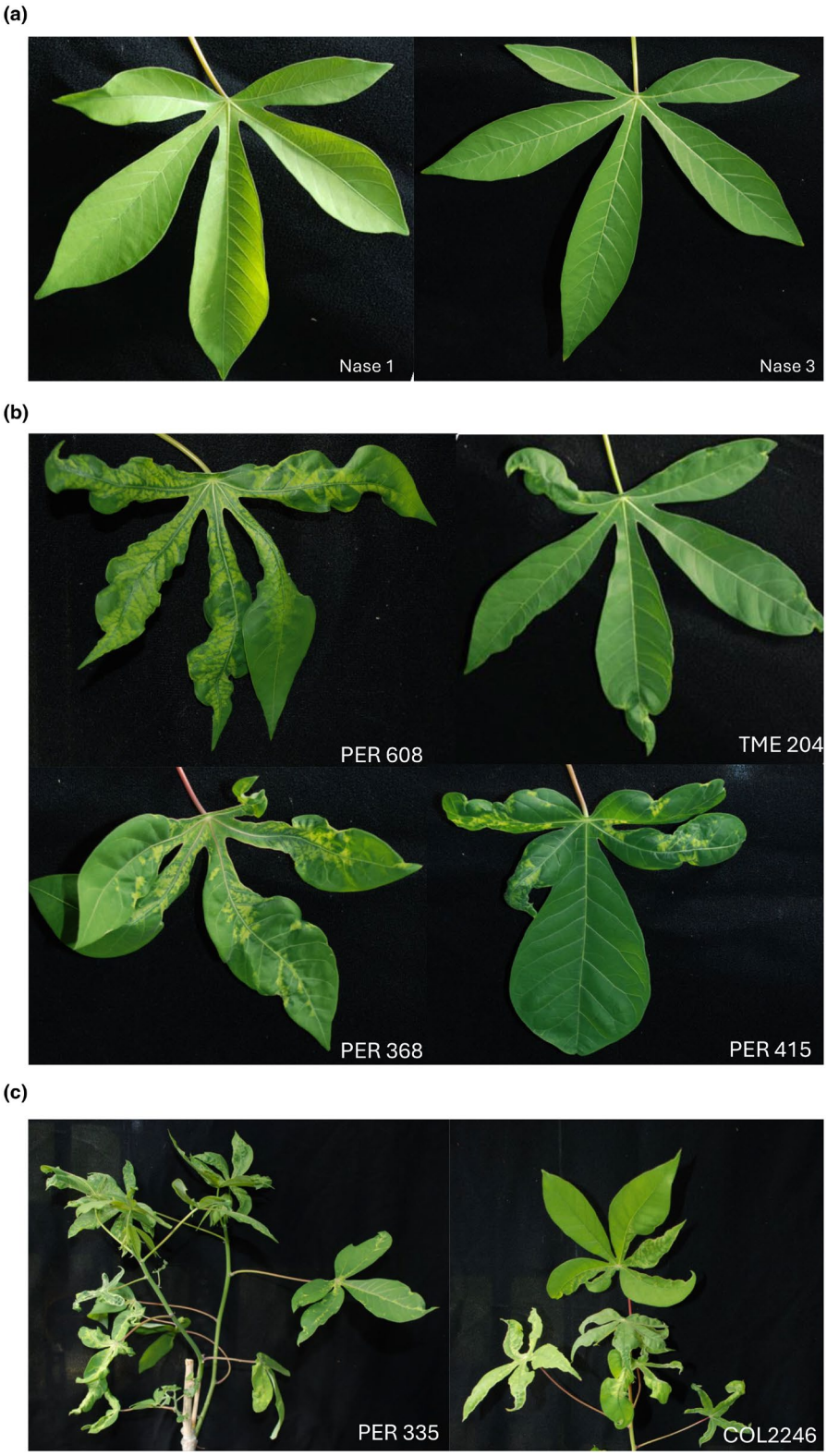

FIGURE 2 | Cassava mosaic disease symptom expression on cassava varieties at 24 weeks after grafting.

mock-inoculated samples, except for Nase 3 at 2 DAG, where no DEGs were identified.

The total transcripts after the quality filtering of the data (fold change cut-off, $p \leq 0.05$) were used for downstream analysis. Nase 3 had the highest number of DEGs (1940 genes up-regulated; 2288 genes down-regulated), followed by TZ 130 (983 genes up-regulated; 1567 genes down-regulated) when compared to Ebwanateraka (910 genes up-regulated; 333 genes down-regulated) over the 28 DAG period (Table S3). Notwithstanding

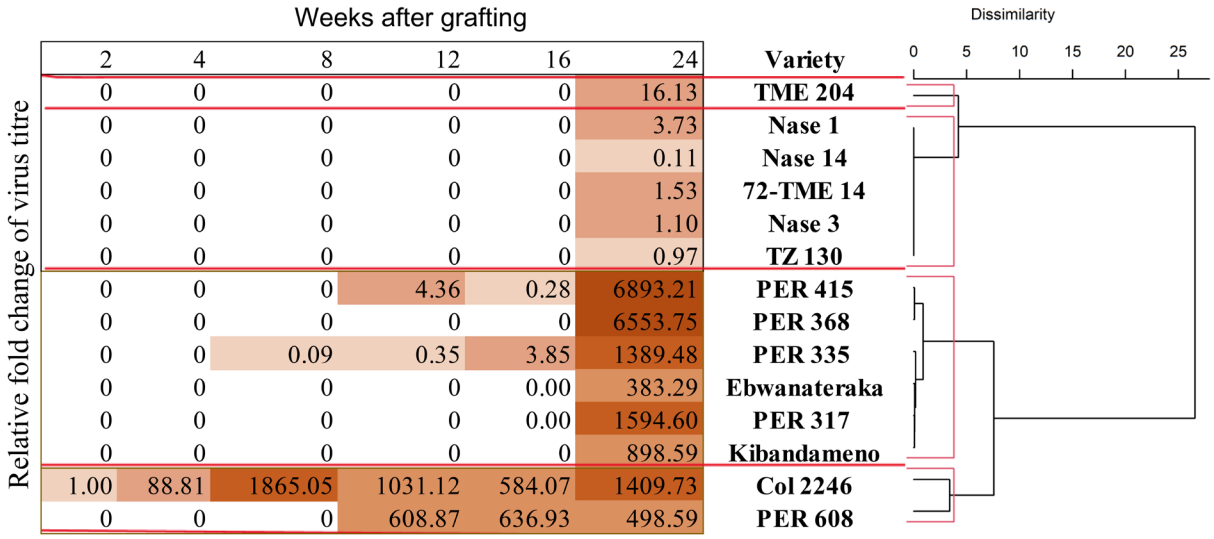

FIGURE 3 | Time course quantification of EACMV-Ug titre from 4 to 24 weeks after grafting with EACMV-Ug (left) and hierarchical clustering (right) of relative fold change of virus titre.

the genetic background of these varieties, the highest number of DEGs was seen among TZ 130 and Nase 3 at 4 and 7 DAG compared with other combinations of varieties (Nase 3 and Ebwanateraka, TZ 130 and Ebwanateraka) (Figure 4).

All DEGs were further assigned into three groups based on their functional annotation and gene ontology (GO) terms: molecular functions (MF), cellular components (CC) and biological processes (BP) (Table S4 and Figures S1–S3). Among all the varieties, MF and BP associated with transcription were enriched. Specifically, response to abiotic stimulus BP was enriched in Nase 3 at 4 DAG for up-regulated genes but not at any later time points after virus inoculation. BP was also absent in Ebwanateraka and TZ 130 DEGs at all time points (Figure S3).

3.4 | Putative CMD Defence Genes

Several defence genes were observed to be differentially expressed across all time points in all varieties. The observed transcripts included heat shock proteins (HSPs), lactoylglutathione lyase glyoxalase I (GLO1), late embryogenesis abundant protein (LEA), histone-related genes, pathogenesis-related (PR) proteins and TFs that are involved in stress tolerance and defence response (Tables S5–S8).

Three MeHSP20-like chaperone superfamily proteins (Manes.01G119000, Manes.02G124600 and Manes.02G124700) and Manes.16G083600 were up-regulated at 4 DAG (fold change = 1.0–5.8) in the resistant/tolerant Nase 3 and TZ 130. Manes.16G083600 was up-regulated in Nase 3 at 28 DAG. These HSPs were not differentially expressed at any time point in the susceptible Ebwanateraka (Table S5).

Two GLO1 genes (Manes.02G101800 and Manes.01G143500) were down-regulated at 4 DAG in Nase 3 and TZ 130. These were not differentially expressed in Ebwanateraka at any time point (Table S5).

MeLEA_3 (Manes.03G197100) was up-regulated in Nase 3 at 4 and 7 DAG and in TZ 130 at 4 DAG. This was not differentially expressed at 2 DAG in Ebwanateraka (Table S5).

Five histone-related genes, Manes.01G202800, histone H2A (Manes.05G070300) and histone H3 (Manes.14G013600, Manes.13G097500, Manes.12G129100) were all down-regulated in Nase 3 and TZ 130, while they were not differentially expressed at any time points in Ebwanateraka (Table S5).

Among the four MePR genes analysed, PR-10f5 (Manes.15G008100), PR-5c (Manes.11G095900) and PR-5h2 (Manes.06G007400 and Manes.14G170800), Manes.15G008100 was down-regulated at 4 DAG in TZ 130 and Nase 3, and at 28 DAG in Ebwanateraka (Table S5).

3.4.1 | Ethylene-Responsive Transcription Factor (ERF)

Manes.13G093300 and Manes.18G069000 were up-regulated at 4 DAG in Nase 3 and down-regulated in TZ 130 at 7 DAG, while Manes.12G087500 was up-regulated in Nase 3 and Ebwanateraka at 4 and 7 DAG. Manes.16G111000 was down-regulated in Ebwanateraka and Manes.15G066800 was up-regulated in Nase 3 at 4 DAG. At the same time point, Manes.07G114500 was induced in Nase 3 and TZ 130. ERF-5 (Manes.01G085200) was induced at 4 DAG in Nase 3 and TZ 130 and at 7 DAG in Ebwanateraka. The same gene was down-regulated in TZ 130 at 7 DAG.

3.4.2 | Teosinte-Like, Cycloidea and PCF1 (TCP)

Four of the five TCP genes were up-regulated only in Nase 3 at different time points: MeTCP11d (Manes.04G016700), MeTCP7 (TCP21-RELATED; Manes.11G108500), MeTCP19 (Manes.06G141800) at 4 DAG and MeTCP15d (TCP family transcription factor; Manes.14G077200) at 28 DAG. MeTCP20a (Manes.04G088500) was up-regulated in Ebwanateraka, but down-regulated in Nase 3 at 4 DAG. Finally, MeTCP9a

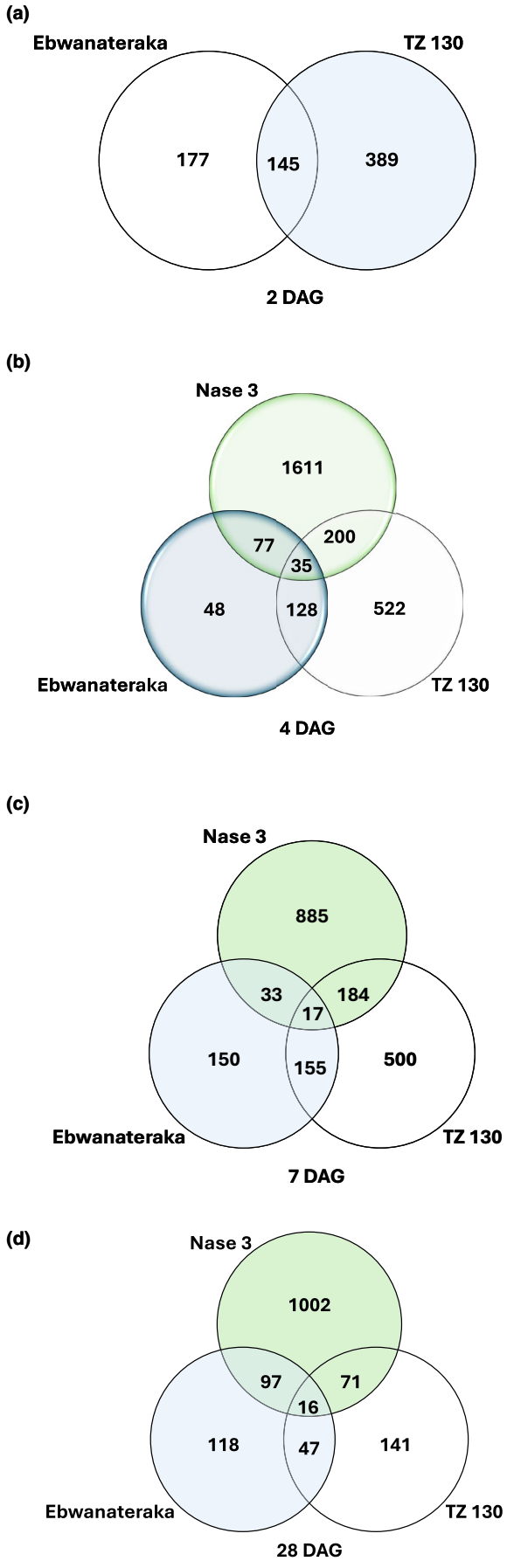

FIGURE 4 | Comparisons of differentially expressed genes between Ebwanateraka, Nase 3. DAG, days after grafting.

(Manes.01G263300) was down-regulated in both Nase 3 and TZ 130 at 7 DAG and up-regulated in Nase 3 at 4 and 28 DAG.

3.4.3 | Heat Stress Transcription Factor (HSF)

Only six HSFs were observed in this study. Four of them were Nase 3-specific DEGs: MeHSFB-2B (Manes.01G065200), MeHSFB-1 (Manes.12G101000), MeHSFA-3 (Manes.08G020700) and MeHSFC-1 (Manes.16G116200), which were up-regulated at 4 DAG, while MeHSFB-1 (Manes.13G124500) was down-regulated at 28 DAG in TZ 130. HSFB-2A (Manes.14G027700) was up-regulated at 4 DAG in both Nase 3 and TZ 130.

3.4.4 | Basic Leucine Zipper (bZIP)

All three bZIP genes identified in this study were Nase 3 specific, while Manes.18G001000 and Manes.18G005000 encoded MebZIP60. Manes.18G001000 was up-regulated at 4 DAG and down-regulated at 28 DAG, while Manes.18G005000 was up-regulated only at 4 DAG. Manes.01G227900 was up-regulated at 4 and 28 DAG.

3.4.5 | Nuclear Factor Y (NF-Y)

MeNF-YA-10-RELATED (Manes.09G044200) was up-regulated in Ebwanateraka at 2 DAG, while MeNF-YA-10-RELATED, A-3-RELATED (Manes.04G142600) and ALPHA-RELATED (Manes.16G097900) were up-regulated at 4 and 28 DAG. A-1-RELATED (Manes.14G003100) and ALPHA-RELATED (Manes.09G025200) were up-regulated at 4 DAG, and ALPHA-RELATED (Manes.10G141400) at 7 DAG in Nase 3.

3.4.6 | Other Transcription Factors

3.4.6.1 | NAC. NAC DOMAIN-CONTAINING PROTEIN 72 (MeNAC72; Manes.15G084800) was up-regulated in Ebwanateraka at 4 and 28 DAG but was not differentially expressed in Nase 3 and TZ 130.

3.4.6.2 | MYB. Manes.12G085600 was up-regulated in TZ 130 and Ebwanateraka at different time points (2 and 4 DAG respectively). At 7 DAG, it was down-regulated in TZ 130, but up-regulated in Ebwanateraka.

3.4.6.3 | bHLH. MebHLH57 (Manes.05G015800) was shown to be down-regulated in TZ 130 and Nase 3 (at 2 and 7 DAG) and up-regulated at 4 DAG in Ebwanateraka.

3.4.6.4 | WRKY. WRKY DNA-binding domain (WRKY 70; Manes.08G170600) was up-regulated in Nase 3 and not differentially expressed in TZ 130 and Ebwanateraka.

4 | Discussion

In this study, we measured symptom expression and EACMV-Ug relative loads over time in 14 cassava varieties and used the

RNA-Seq method to identify DEGs between EACMV-Ug-inoculated and mock-inoculated plants across time with three cassava varieties: Nase 3 (resistant), TZ 130 (resistant) and Ebwanateraka (susceptible). Col 2246 was the first to show symptoms at 8 WAG; this is consistent with previous observations that susceptible varieties displayed early symptoms (Houngue et al. 2019; Kuria et al. 2017; Ogbe et al. 2003). Ebwanateraka, PER 368 and PER 608 supported high virus concentration as much as Col 2246 but not the other varieties (72-TME 14, Kibandameno, Nase 1, Nase 3, Nase 14, TZ 130, TME 204, PER 317, PER 335 and PER 415) before 16 WAG. This could be attributed to their ability to restrict virus multiplication and limit symptom development. Pruning had a significant effect on symptom severity and virus quantity ($df = 13$; $p < 0.0001$). Col 2246, Ebwanateraka, Kibandameno, PER 317, PER 335, PER 368, PER 415 and PER 608 expressed severe symptoms with high virus quantities and were considered susceptible. TME 204 expressed mild symptoms with lower virus quantities compared to Col 2246; hence, it was considered tolerant (Fargette et al. 1996; Thresh et al. 1994). The definitions of resistance, tolerance and susceptibility adopted in this work are those established earlier (Ogbe et al. 2003), where resistance corresponds to no or very low accumulation of the virus with no or mild foliar symptoms. Tolerance is an interaction in which viruses accumulate but with mild foliar symptom severity, while susceptibility is high virus concentration with severe disease symptoms. Based on these definitions, 72-TME 14, TZ 130, Nase 1, Nase 3 and Nase 14 were considered CMD resistant. These varieties were also classified as resistant based on field screening (Maruthi et al. 2019; Tumwegamire et al. 2018). Col 2246, Ebwanateraka, Kibandameno, PER 317, PER 335, PER 368, PER 415 and PER 608 accumulated high viral titres and displayed severe symptoms, which grouped them as susceptible, while only TME 204 grouped as tolerant with mild symptoms and low viral titres. 72-TME 14, Nase 1, Nase 3 and Nase 14 are improved cassava varieties obtained from Uganda (Otim-Nape et al. 1997) and TZ 130 is an improved variety released by both Uganda and Tanzania (Sserubombwe et al. 2001). These varieties can thus provide relief from the impacts of CMD in these countries.

Transcriptomes (RNA-Seq) of Nase 3 and TZ 130 (resistant), together with the susceptible Ebwanateraka, were analysed to understand the mechanisms of resistance and to identify putative EACMV-Ug resistance genes. About 80% of the sequences from all three varieties mapped to the cassava reference genome (except two Nase 3 samples that mapped poorly [51%]). The non-alignment of the remaining 20% sequences could be due to their diverse origin as the sequences used in this study were of African origin, while the reference genome was from a partial inbred line called AM560-2 from Latin America (Otim-Nape et al. 1997; Prochnik et al. 2012; Sserubombwe et al. 2001). Relatively fewer DEGs were observed in the susceptible Ebwanateraka (1243) over the 28 DAG period compared to the resistant Nase 3 (4228) and TZ 130 (2550). The total number of genes repressed by the virus (Nase 3, 2067 and TZ 130, 1176) was greater than the number of induced genes in the resistant varieties (Nase 3, 1725 and TZ 130, 897) and vice versa for the susceptible variety Ebwanateraka. A similar pattern was observed in the transcriptome response of cassava to SACMV (Allie et al. 2014).

Many GO terms were enriched in Nase 3 followed by TZ 130 at 4, 7 and 28 DAG. This was particularly acute at 7 DAG, where more defence genes were expressed in Nase 3 and TZ 130 compared to Ebwanateraka, suggesting a rapid response to EACMV-Ug, which can be a characteristic of resistant varieties. Five major groups of TFs (WRKY, HSF, bZIP, TCP and ERF) were up-regulated in Nase 3, and HSP was up-regulated in TZ 130. These TFs were previously reported to play crucial roles in immune responses against pathogens (Buscaill and Rivas 2014; Noman et al. 2017; Phukan et al. 2016). WRKY are the largest families of DNA-binding TFs involved in plant development as well as responses to biotic and abiotic stress (Birkenbihl et al. 2018; Lai et al. 2008). WRKY DNA-binding domain (WRKY 70; Manes.08G170600) was up-regulated in Nase 3, which is involved in defence responses (Li et al. 2006, 2013). In *Arabidopsis*, AtWRKY70 is involved in the cross-talk between salicylic acid- and jasmonic acid-mediated signalling pathways (Li et al. 2006). We speculate that MeWRKY70 identified in this study may also be involved in a similar cross-talk in cassava. Another class of transcription factor, HSF, plays a role in plant response to several biotic and abiotic stressors (Fragkostefanakis et al. 2015). Four MeHSFs genes were up-regulated at 4 DAG in Nase 3 and TZ 130. MeHSFB-1 (Manes.13G124500) and HSFB-2A in particular were both up-regulated at 4 DAG, possibly indicating the early response by the resistant varieties. In *Arabidopsis*, HsfB1 and HsfB2b suppress the general heat shock response in the absence of extreme heat, but support acquired thermotolerance (Ikeda et al. 2011). Three bZIP genes were up-regulated in Nase 3 at most time points, which are also known to respond to biotic and abiotic stress in plants (Alves et al. 2013; Wei et al. 2012).

For BP, response to abiotic stimulus was highly expressed in Nase 3. The majority of the selected DEGs were shown to be mostly PR genes, HSPs, LEA and histone-related genes involved in basal immune and other metabolic processes. Data from all three varieties were scrutinised to identify differences between genes of resistant and susceptible varieties. PR proteins are an important part of plant immunity. In this study, PR-10f5 was down-regulated in Nase 3 and TZ 130 at 4 DAG and Ebwanateraka at a much later point (28 DAG), while the PR-5c and PR-5h2 were induced in Ebwanateraka and suppressed in TZ 130 at 2 DAG and suppressed in Nase 3 and TZ 130 at 7 DAG. Our results are similar to those of an earlier study where transcript levels of PR-5 (thaumatin superfamily protein) increased after SACMV infection in the susceptible cassava T200 (Allie et al. 2014). Results upon UCBSV infection (another cassava virus) showed that 012383 was up-regulated at 1 and 45 DAG in Albert (susceptible) (Amuge et al. 2017). PR-5 in cassava and *Arabidopsis* is also induced upon whitefly infestation (Irigoyen et al. 2020; Zarate et al. 2007). Results from the different studies show that despite the induction of PR genes in susceptible varieties, they were incapable of preventing viral replication and led to a progressive increase in symptom severity and virus titre. The HSPs are involved in regulatory, signalling and defence pathways. In this study, MeHSP20-like chaperone superfamily protein (Manes.02G124600, Manes.02G124600 and Manes.02G124700) was up-regulated in Nase 3 and TZ 130 at 4 DAG after virus infection, suggesting an early defence response. Similarly, 14 MeHSPs were induced in the resistant variety Namikonga, while the susceptible Albert had only two genes induced after UCBSV infection (Amuge et al. 2017). MeHSP20-like chaperone superfamily proteins are also induced 24 h after infestation by mealybug in the resistant variety AR23.1 and repressed in the

susceptible variety (P40/1) (Rauwane et al. 2018). LEA proteins also play an important role in abiotic stress adaptation in plants, and they were expressed in both TZ 130 and Nase 3. Similarly, two LEA genes were induced in resistant cassava varieties, but remained unchanged in a susceptible variety upon infection by UCBSV (Amuge et al. 2017). Histones are well known to play an important role in DNA replication and transcription regulation (Liu et al. 2010; Saze et al. 2012), and were down-regulated (H3 and H2A) in Nase 3 (at 7 DAG) and TZ 130 (at 2, 4 and 7 DAG), but not in Ebwanateraka. Histone H3 plays a role in geminivirus replication and movement (Kong and Hanley-Bowdoin 2002; Zhou et al. 2011). The H2A was over-expressed in the susceptible cassava variety T200 and *Nicotiana benthamiana*, but not in the tolerant variety TME3 (Allie et al. 2014). These consensual data on the down-regulation of H3 and H2A in resistant/tolerant plants perhaps suggest their role in the inhibition of geminivirus DNA replication in resistant/tolerant plants. The next steps in understanding the mechanisms of resistance to EACMV-Ug should include validating the role of these putative genes by gene up-regulation and knockdown experiments.

In conclusion, we unambiguously confirmed that 72-TME 14, TME 204, TZ 130, Nase 1, Nase 3 and Nase 14 from Africa were resistant to the most severe strain of the CMBs, EACMV-Ug, by molecular approaches. These will make a reliable source of resistance to CMD in Africa and thus can be safely cultivated and promoted for mitigating the impact of the devastating CMD pandemic in affected countries. All the South American varieties were susceptible to EACMV-Ug, further confirming their limited utility as a source of resistance to the begomoviruses and highlighting the risks of accidentally introducing the African CMBs to South America. The early defence response of the resistant varieties was key for their resistance to virus infection, while the response of the susceptible variety was delayed. Also, fewer unique DEGs were expressed by the susceptible variety compared to the resistant varieties. Genes encoding heat shock proteins and transcription factors such as ethylene-responsive transcription factor (ERFs), teosinte-like, cycloidea and PCF1, heat stress transcription factor and basic leucine zipper were observed to be up-regulated early in resistant varieties. The down-regulation of PR genes in Nase 3 and TZ 130 was surprising and requires further investigations.

Acknowledgements

We are grateful to Drs. Luis Augusto Bacerra Lopez-Lavalle and Adriana Bohoroquez from the International Center for Tropical Agriculture (CIAT), Colombia, for providing the South American cassava varieties in the African cassava whitefly project and the many African partners for providing the African cassava varieties in the 5CP project that were used in this study. We thank Natalie Morley for helping to maintain cassava plants. This work was supported by the African Union Commission (grants AURG/2/141 and AURG II-1-060-, 2016).

Conflicts of Interest

The authors declare no conflicts of interest.

Data Availability Statement

The data that supports the findings of this study are available in the Supporting Information of this article.

References

Abarshi, M. M., I. U. Mohammed, P. Wasswa, et al. 2010. "Optimization of Diagnostic RT-PCR Protocols and Sampling Procedures for the Reliable and Cost-Effective Detection of Cassava Brown Streak Virus." *Journal of Virological Methods* 163: 353–359.

Alabi, O., P. Kumar, and R. Naidu. 2011. "Cassava Mosaic Disease: A Curse to Food Security in SubSaharan Africa." *APSnet Feature*. https://10.1094/APSnetFeature-2011-0701.

Allie, F., E. J. Pierce, M. J. Okoniewski, and C. Rey. 2014. "Transcriptional Analysis of South African Cassava Mosaic Virus-Infected Susceptible and Tolerant Landraces of Cassava Highlights Differences in Resistance, Basal Defense and Cell Wall Associated Genes During Infection." *BMC Genomics* 15: 1006.

Alves, M. S., S. P. Dadalto, A. B. Gonçalves, G. B. De Souza, V. A. Barros, and L. G. Fietto. 2013. "Plant bZIP Transcription Factors Responsive to Pathogens: A Review." *International Journal of Molecular Sciences* 14: 7815–7828.

Amuge, T., D. K. Berger, M. S. Katari, A. A. Myburg, S. L. Goldman, and M. E. Ferguson. 2017. "A Time Series Transcriptome Analysis of Cassava (*Manihot esculenta* Crantz) Varieties Challenged With Ugandan Cassava Brown Streak Virus." *Scientific Reports* 7: 9749.

Beyene, G., R. D. Chauhan, H. Wagaba, et al. 2016. "Loss of CMD2-Mediated Resistance to Cassava Mosaic Disease in Plants Regenerated Through Somatic Embryogenesis." *Molecular Plant Pathology* 17: 1095–1110.

Birkenbihl, R. P., B. Kracher, A. Ross, K. Kramer, I. Finkmeier, and I. E. Somssich. 2018. "Principles and Characteristics of the *Arabidopsis* WRKY Regulatory Network During Early MAMP-Triggered Immunity." *Plant Journal* 96: 487–502.

Buscaill, P., and S. Rivas. 2014. "Transcriptional Control of Plant Defence Responses." *Current Opinion in Plant Biology* 20: 35–46.

Fan, H., Y. Zhang, H. Sun, et al. 2015. "Transcriptome Analysis of *Beta macrocarpa* and Identification of Differentially Expressed Transcripts in Response to Beet Necrotic Yellow Vein Virus Infection." *PLoS One* 10: 0132277.

Fargette, D., L. T. Colon, R. Bouveau, and C. Fauquet. 1996. "Components of Resistance of Cassava to African Cassava Mosaic Virus." *European Journal of Plant Pathology* 102: 645–654.

Fragkostefanakis, S., S. Röth, E. Schleiff, and K. D. Scharf. 2015. "Prospects of Engineering Thermotolerance in Crops Through Modulation of Heat Stress Transcription Factor and Heat Shock Protein Networks." *Plant, Cell & Environment* 38: 1881–1895.

Houngue, J. A., M. Zandjanakou-Tachin, H. B. Ngalle, et al. 2019. "Evaluation of Resistance to Cassava Mosaic Disease in Selected African Cassava Cultivars Using Combined Molecular and Greenhouse Grafting Tools." *Physiological and Molecular Plant Pathology* 105: 47–53.

Ikeda, M., N. Mitsuda, and M. Ohme-Takagi. 2011. "*Arabidopsis* HsfB1 and HsfB2b Act as Repressors of the Expression of Heat-Inducible *Hsf*s but Positively Regulate the Acquired Thermotolerance." *Plant Physiology* 157: 1243–1254.

Irigoyen, M. L., D. C. Garceau, A. Bohorquez-Chaux, et al. 2020. "Genome-Wide Analyses of Cassava Pathogenesis-Related (PR) Gene Families Reveal Core Transcriptome Responses to Whitefly Infestation, Salicylic Acid and Jasmonic Acid." *BMC Genomics* 21: 93.

Kavil, S., G. Otti, S. Bouvaine, A. Armitage, and M. N. Maruthi. 2021. "*PAL1* Gene of the Phenylpropanoid Pathway Increases Resistance to the *Cassava brown streak virus* in Cassava." *Virology Journal* 18: 184.

Kong, L. J., and L. Hanley-Bowdoin. 2002. "A Geminivirus Replication Protein Interacts With a Protein Kinase and a Motor Protein That Display Different Expression Patterns During Plant Development and Infection." *Plant Cell* 14: 1817–1832.

Kuria, P., M. Ilyas, E. Ateka, et al. 2017. "Differential Response of Cassava Genotypes to Infection by Cassava Mosaic Geminiviruses." *Virus Research* 227: 69–81.

Lai, Z., K. M. Vinod, Z. Zheng, B. Fan, and Z. Chen. 2008. "Roles of *Arabidopsis* WRKY3 and WRKY4 Transcription Factors in Plant Responses to Pathogens." *BMC Plant Biology* 8: 68.

Legg, J. P., S. C. Jeremiah, H. M. Obiero, et al. 2011. "Comparing the Regional Epidemiology of the Cassava Mosaic and Cassava Brown Streak Virus Pandemics in Africa." *Virus Research* 159: 161–170.

Legg, J. P., P. L. Kumar, T. Makeshkumar, et al. 2015. "Cassava Virus Diseases: Biology, Epidemiology, and Management." In *Advances in Virus Research*, vol. 91, 85–142. Elsevier.

Legg, J. P., B. Owor, P. Sseruwagi, and J. Ndunguru. 2006. "Cassava Mosaic Virus Disease in East and Central Africa: Epidemiology and Management of a Regional Pandemic." *Advances in Virus Research* 67: 355–418.

Li, J., S. Besseau, P. Törönen, et al. 2013. "Defense-Related Transcription Factors WRKY 70 and WRKY 54 Modulate Osmotic Stress Tolerance by Regulating Stomatal Aperture in *Arabidopsis*." *New Phytologist* 200: 457–472.

Li, J., G. Brader, T. Kariola, and E. T. Palva. 2006. "WRKY70 Modulates the Selection of Signaling Pathways in Plant Defense." *Plant Journal* 46: 477–491.

Liu, C., F. Lu, X. Cui, and X. Cao. 2010. "Histone Methylation in Higher Plants." *Annual Review of Plant Biology* 61: 395–420.

Maruthi, M. N., S. Bouvaine, H. A. Tufan, I. U. Mohammed, and R. J. Hillocks. 2014. "Transcriptional Response of Virus-Infected Cassava and Identification of Putative Sources of Resistance for Cassava Brown Streak Disease." *PLoS One* 9: 0096642.

Maruthi, M. N., J. Colvin, S. Seal, G. Gibson, and J. Cooper. 2002. "Co-Adaptation Between Cassava Mosaic Geminiviruses and Their Local Vector Populations." *Virus Research* 86: 71–85.

Maruthi, M. N., S. C. Jeremiah, I. U. Mohammed, and J. P. Legg. 2017. "The Role of the Whitefly, *Bemisia tabaci* (Gennadius), and Farmer Practices in the Spread of Cassava Brown Streak Ipomoviruses." *Journal of Phytopathology* 165: 707–717.

Maruthi, M. N., E. C. Whitfield, G. Otti, et al. 2019. "A Method for Generating Virus-Free Cassava Plants to Combat Viral Disease Epidemics in Africa." *Physiological and Molecular Plant Pathology* 105: 77–87.

Mohammed, I. U., M. M. Abarshi, B. Muli, R. J. Hillocks, and M. N. Maruthi. 2012. "The Symptom and Genetic Diversity of Cassava Brown Streak Viruses Infecting Cassava in East Africa." *Advances in Virology* 2012: 795697.

Moreno, I., W. Gruissem, and H. Vanderschuren. 2011. "Reference Genes for Reliable Potyvirus Quantitation in Cassava and Analysis of Cassava Brown Streak Virus Load in Host Varieties." *Journal of Virological Methods* 177: 49–54.

Nicaise, V. 2014. "Crop Immunity Against Viruses: Outcomes and Future Challenges." *Frontiers in Plant Science* 5: 660.

Noman, A., Z. Liu, M. Aqeel, et al. 2017. "Basic Leucine Zipper Domain Transcription Factors: The Vanguards in Plant Immunity." *Biotechnology Letters* 39: 1779–1791.

Ogbe, F. O., G. Thottappilly, A. G. O. Dixon, G. I. Atiri, and H. D. Mignouna. 2003. "Variants of *East African cassava mosaic virus* and Its Distribution in Double Infections With *African cassava mosaic virus* in Nigeria." *Plant Disease* 87: 229–232.

Otim-Nape, G. W., J. M. Thresh, and M. W. Shaw. 1997. "The Effects of Cassava Mosaic Virus Disease on Yield and Compensation in Mixed Stands of Healthy and Infected Cassava." *Annals of Applied Biology* 130: 503–521.

Otti, G., S. Bouvaine, B. Kimata, et al. 2016. "High-Throughput Multiplex Real-Time PCR Assay for the Simultaneous Quantification of DNA and RNA Viruses Infecting Cassava Plants." *Journal of Applied Microbiology* 120: 1346–1356.

Patil, B. L., J. P. Legg, E. Kanju, and C. M. Fauquet. 2015. "Cassava Brown Streak Disease: A Threat to Food Security in Africa." *Journal of General Virology* 96: 956–968.

Phukan, U. J., G. S. Jeena, and R. K. Shukla. 2016. "WRKY Transcription Factors: Molecular Regulation and Stress Responses in Plants." *Frontiers in Plant Science* 7: 760.

Prochnik, S., P. R. Marri, B. Desany, et al. 2012. "The Cassava Genome: Current Progress, Future Directions." *Tropical Plant Biology* 5: 88–94.

Raudvere, U., L. Kolberg, I. Kuzmin, et al. 2019. "G:Profiler: A Web Server for Functional Enrichment Analysis and Conversions of Gene Lists (2019 Update)." *Nucleic Acids Research* 47, no. W1: W191–W198.

Rauwane, M. E., D. A. Odeny, I. Millar, C. Rey, and J. Rees. 2018. "The Early Transcriptome Response of Cassava (*Manihot esculenta* Crantz) to Mealybug (*Phenacoccus manihoti*) Feeding." *PLoS One* 13: 202541.

Saze, H., K. Tsugane, T. Kanno, and T. Nishimura. 2012. "DNA Methylation in Plants: Relationship to Small RNAs and Histone Modifications, and Functions in Transposon Inactivation." *Plant and Cell Physiology* 53: 766–784.

Sserubombwe, W. S., J. M. Thresh, G. W. Otim-Nape, and D. O. S. Osiru. 2001. "Progress of Cassava Mosaic Virus Disease and Whitefly Vector Populations in Single and Mixed Stands of Four Cassava Varieties Grown Under Epidemic Conditions in Uganda." *Annals of Applied Biology* 138: 161–170.

Stobbe, A. H., and M. J. Roossinck. 2014. "Plant Virus Metagenomics: What We Know and Why We Need to Know More." *Frontiers in Plant Science* 5: 150.

Thresh, J. M., D. Fargette, and G. W. Otim-Nape. 1994. "Effects of African Cassava Mosaic Geminivirus on the Yield of Cassava." *Tropical Science* 34: 26–42.

Tomlinson, K. R., A. M. Bailey, T. Alicia, S. Seal, and G. D. Foster. 2018. "Cassava Brown Streak Disease: Historical Timeline, Current Knowledge and Future Prospects." *Molecular Plant Pathology* 19: 1282–1294.

Trapnell, C., A. Roberts, L. Goff, et al. 2012. "Differential Gene and Transcript Expression Analysis of RNA-Seq Experiments With TopHat and Cufflinks." *Nature Protocols* 7: 562–578.

Tumwegamire, S., E. Kanju, J. Legg, et al. 2018. "Exchanging and Managing In-Vitro Elite Germplasm to Combat Cassava Brown Streak Disease (CBSD) and Cassava Mosaic Disease (CMD) in Eastern and Southern Africa." *Food Security* 10: 351–368.

Wei, K., J. Chen, Y. Wang, et al. 2012. "Genome-Wide Analysis of bZIP-Encoding Genes in Maize." *DNA Research* 19: 463–476.

Zarate, S. I., L. A. Kempema, and L. L. Walling. 2007. "Silverleaf Whitefly Induces Salicylic Acid Defenses and Suppresses Effectual Jasmonic Acid Defenses." *Plant Physiology* 143: 866–875.

Zhou, X., Y. Liu, L. Calvert, et al. 1997. "Evidence That DNA-A of a Geminivirus Associated With Severe Cassava Mosaic Disease in Uganda Has Arisen by Interspecific Recombination." *Journal of General Virology* 78: 2101–2111.

Zhou, Y., M. R. Rojas, M. R. Park, Y. S. Seo, W. J. Lucas, and R. L. Gilbertson. 2011. "Histone H3 Interacts and Colocalizes With the Nuclear Shuttle Protein and the Movement Protein of a Geminivirus." *Journal of Virology* 85: 11821–11832.

Supporting Information

Additional supporting information can be found online in the Supporting Information section.