Bioenerg. Res. (2017) 10:129–145
DOI 10.1007/s12155-016-9783-6

Comparative NGS Transcriptomics Unravels Molecular Components Associated with Mosaic Virus Infection in a Bioenergy Plant Species, *Jatropha curcas* L.

Archit Sood[1] · **Rajinder Singh Chauhan**[1]

Published online: 2 August 2016

Abstract *Jatropha curcas*, a promising bioenergy crop is becoming vulnerable to various biotic stresses due to large-scale cultivation of selected genotypes, thereby affecting its yield potential. Virus causing mosaic disease is prevalent in Jatropha plantations and causing significant reduction in seed yield and quality. To better understand the molecular mechanisms associated with virus infection response, we performed RNA-seq-based comprehensive transcriptome sequencing of symptomatic virus-infected (JV) and healthy (JH) leaf tissues of *J. curcas*. Through reference genome-based mapping approach, 55,755 genes expressed in both samples were identified. Differential expression analysis identified genes linked to various pathways, upregulated and downregulated during mosaic virus infection. Upon KEGG-based functional annotation, it was observed that various metabolism-associated processes along with oxidative phosphorylation, endocytosis, terpenoid biosynthesis, and hormone signal transduction were upregulated whereas photosynthesis, anthocyanin biosynthesis, plant-pathogen interaction, and calcium signaling were downregulated in response to virus infection. Significantly, genes associated with hormone signal transduction were upregulated as physiological symptoms induced upon mosaic virus infection is due to the interplay of various phytohormones regulating general growth and development of plant. Also, many genes regulating photosynthesis which were downregulated during virus infection showed repressed rate of photosynthesis and also reduction in seed yield and oil content upon mosaic virus infection in *J. curcas*. RT-qPCR-based experimental validation approach was supplemented to confirm the computational identification. The study provides a repertoire of molecular components which have been affected in response to virus infection, and their precise role can be further functionally validated.

Keywords Jatropha · Healthy · Virus infected · Transcriptome · Pathways

Abbreviations

ABA	Abscisic acid
CDS	Coding sequence
CML	Calcium-binding protein
CMV	Cucumber mosaic virus
CNGCs	Cyclic nucleotide-gated ion channels
FPKM	Fragments per kilobase of transcripts per million mapped reads
GO	Gene ontology
JA	Jasmonic acid
*Jc*MD	*Jatropha curcas* mosaic disease
JH	Healthy
JV	Virus infected
KEGG	Kyoto Encyclopedia of Genes and Genomes
NGS	Next-generation sequencing
RNA-seq	Ribose nucleic acid sequencing
RPM1	Resistance to *Pseudomonas syringae* PV Maculicola 1
RPS2	Resistance to *Pseudomonas syringae* 2
RT-qPCR	Reverse transcription-quantitative polymerase chain reaction
SA	Salicylic acid

Electronic supplementary material The online version of this article (doi:10.1007/s12155-016-9783-6) contains supplementary material, which is available to authorized users.

✉ Rajinder Singh Chauhan
rajinder.chauhan@juit.ac.in

[1] Department of Biotechnology and Bioinformatics, Jaypee University of Information Technology, Waknaghat, Solan, H P 173 215, India

Introduction

The overall increase in energy consumption necessitates searching for alternative sources of fuel other than fossil fuels. Biodiesel, majorly produced from plants, is an alternate to fossil fuel as it is nontoxic, biodegradable, and emits lower amount of carbon monoxides and hydrocarbons than petro-diesel. The crop plants being used for the production of biodiesel are edible and causing scarcity for the overall food supply. For effective and profitable biodiesel production, the whole emphasis should shift from edible to nonedible crop plants so that there is no competition with food security [1]. There is a hurdle in the effective management of these energy crops as they have been plagued with diseases that thwart production and causes a global yearly average yield loss of 16 % [2].

Jatropha curcas L. (Jatropha) is a nonedible plant species of family euphorbiaceae, considered as promising nonedible crop plant for biodiesel production due to the presence of unsaturated fatty acids in a large ratio. Large-scale cultivation of Jatropha is becoming vulnerable to biotic stresses like diseases and pests which are directly affecting the yield potential of this bioenergy crop [3]. Mosaic disease caused by *J. curcas* mosaic virus (*Jc*MV) has been reducing overall seed yield and quality in *J. curcas* as well as other species. It is characterized by leaf curling resulting in reduction of fruit size and distortion. Viruses like begomovirus and *Cucumber mosaic virus* (CMV) has been identified in large plantations [4, 5]. The outbreak of viruses in Jatropha has been reported in other parts of the world such as Ramkat et al. [6] reported the incidence of Jatropha infection with cassava mosaic viruses in Europe. Begomovirus associated with mosaic disease of *J. curcas* has been identified in Nigeria [7]. However, as of today, no information exists on molecular components associated with viral disease in Jatropha.

In recent years, next-generation sequencing (NGS) has provided novel and advanced ways to fasten the identification of genes in many plant species, mainly those under biotic and abiotic stresses [8, 9]. Transcriptome analysis technique has been done to gain molecular understanding regarding virus infection response in many plants. Lu et al. [10] performed deep sequencing of healthy and mosaic virus-infected plants of tobacco and found that biological processes, such as pigment metabolism, photosynthesis, and plant-pathogen interaction, were linked with virus symptom development. In another study, the transcriptome sequencing of African cassava mosaic virus-infected cassava leaves provided novel insights into the upregulation of genes associated with degradation of chlorophyll and thus photosynthesis [11]. Recently, Choi et al. [12] carried out transcriptomic investigation of chrysanthemum in response to tomato spotted wilt virus, cucumber mosaic virus, and potato virus X. Genes involved in stress response such as ethylene-mediated signaling pathway and chitin response were upregulated. Tomato spotted wilt virus infection downregulated genes related to DNA metabolic process such as DNA replication, cytokinesis, histone modification, and chromatin organization. Genes related to photosynthesis and flowering were downregulated whereas genes linked to metabolic pathways, transcription, and stress responses were found to be upregulated in response to stripe virus infection in rice as revealed by transcriptome profiling of infected seedlings [13]. Similarly, transcriptome sequencing approach has been employed to understand the molecular basis of disease response in plants like alfalfa, beet, capsicum, tomato, and white pine [8, 14–17]. For Jatropha also, many transcriptome-based sequencing methodologies have been employed to elucidate molecular responses to important biological processes such as oil biosynthesis, flower formation, abiotic stress, waterlogging, etc. [18–22]. However, the molecular insights regarding mosaic virus infection in Jatropha are not understood. Only few reports regarding identification and characterization of mosaic virus in Jatropha exist [4, 5, 23]. Although there are reports regarding the overall understanding of molecular mechanism behind disease resistance in *J. curcas* [24, 25], no studies on investigating molecular components regulating disease response in *J. curcas* exist till date.

Therefore, to uncover insights into the molecular mechanisms in response to virus infection in Jatropha, we employed NextSeq 500 platform (Illumina) of RNA-seq technology in symptomatic virus-infected (JV) and healthy (JH) leaves of *J. curcas*, identified a total of 12,226,847 and 10,548,434 high-quality reads, respectively, and 55,755 genes expressed in both samples through reference genome-based mapping approach. Upon Kyoto Encyclopedia of Genes and Genomes (KEGG)-based functional annotation, it was observed that various metabolism-associated processes like oxidative phosphorylation (energy metabolism), arginine and proline metabolism, ascorbate metabolism, lipid metabolism, fatty acid metabolism, and sugar metabolism were upregulated in response to virus infection. Other major processes like endocytosis, terpenoid biosynthesis, and hormone signal transduction were also upregulated whereas photosynthesis, anthocyanin biosynthesis, and pathways related to defense responses, i.e., plant-pathogen interaction and calcium signaling, were suppressed during virus infection. Gene ontology (GO)-based functional annotation revealed correlation with KEGG-based pathway analysis. RT-qPCR-based experimental validation approach was supplemented to confirm the computational identification. Coexpression network analysis indicated presence of other genes associated with the identified upregulated and downregulated genes. To our knowledge, this is the first comprehensive transcriptomic analysis for understanding the molecular mechanisms underlying disease response in Jatropha. The present study could lead to genetic interventions for broadening disease control methodologies, specifically virus control, in Jatropha to maximize its potential to be an ideal bioenergy crop.

Materials and Methods

Plant Material

Two to three healthy and symptomatic virus-infected younger apical leaves were collected from different mature plants of *J. curcas* genotype IC561235 from experimental farm of Himalayan Forest Research Institute at Jwalaji (508-m altitude, 31° 51′ N, 76° 18′ E), Himachal Pradesh, India (Supplementary Fig. S1). The selected plants were of same developmental stage (8 years old). Leaf tissues were frozen immediately in liquid nitrogen and stored at −80 °C until use.

DNA Isolation and Detection of Virus

Total DNA was isolated from healthy and symptomatic virus-infected leaves (each in two biological replicates) using CTAB extraction method with minor modifications. Primers were designed for Jatropha mosaic virus coat protein gene (NCBI accession no. 9247600). PCR amplification was performed on Thermocycler (Applied Biosystems, USA) in 25-μl reaction volume. The PCR was performed on 30 ng of genomic DNA with primer pairs, Mg2+, dNTPs (Intron Technologies, USA), and Taq DNA polymerase (Intron Technologies, USA). Amplification programs included 94 °C for 5 min, 30 cycles of 94 °C for 45 s, annealing temperature of 54.6 °C for 45 s, 72 °C for 2 min, and a final extension of 7 min at 72 °C. Amplified products on 1 % agarose gel were analyzed using the gel documentation system AlphaImager EP (Alpha Innotech Corp., USA).

Isolation of Total RNA

Total RNA was isolated from virus-infected (JV) and healthy (JH) leaves (each in two biological replicates) using RaFlex Total RNA isolation Kit (Genei, India) as per manufacturer's instruction. The quality of total RNA was checked on 1 % denatured Agarose gel (1 μg) for the presence of 28S and 18S bands. The gel was run at 100 V for 30 min. Further, total RNA was quantified using Qubit RNA BR kit (fluorometer).

Illumina NextSeq 2 × 150 PE Library Preparation

Paired-end complementary DNA (cDNA) sequencing libraries were prepared using Illumina TruSeq stranded Total RNA Library Preparation Kit and as per protocol described by manufacturer (Illumina). Briefly, ribosomal RNA (rRNA) was depleted from total RNA followed by fragmentation. The fragmented rRNA depleted RNA was converted into first-strand cDNA, followed by second-strand cDNA generation, A-tailing, adapter ligation, and finally ended by index PCR amplification of adaptor-ligated library. Library quantification and validation was performed using Qubit dsDNA HS kit and High Sensitivity Assay Kit, respectively. RNA-Seq libraries were prepared using Illumina TrueSeq stranded total RNA HT (with Ribo-Zero plant) kit using 1 μg of total RNA. The mean size of the fragment distribution ranged from 550 to 700 bp.

Data Generation and Mapping of Reads to Reference Genome

The raw data was generated on NextSeq. The raw reads were filtered using Trimmomatic (v 0.30) with quality value QV > 20 and other contaminants such as adapters were also trimmed. The reference genome of *J. curcas* was downloaded from Jatropha Genome Database (http://www.kazusa.or.jp/jatropha/). The Illumina NextSeq transcriptome data for both samples were separately mapped to the Jatropha reference genome using BWA version 0.7.5a (http://bio-bwa.sourceforge.net/) with default settings. The software package SAMtools (http://samtools.sourceforge.net/) was used to convert sequence alignment/map (SAM) file to sorted binary alignment/map (BAM) file. Mapped reads ratio (MRR) to the reference in each dataset was calculated by applying flagstat command of SAMtools software to the BAM file.

Differential Gene Expression Analysis

The CDS from Jatropha GFF file were used to study the gene expression analysis. The expression analysis of these genes was carried out with R package DESeq. The expression of genes was calculated in terms of fragment per kilobase per million mapped reads (FPKM). The FPKM values for each gene were calculated for healthy (JH) and virus-infected (JV) samples with DESeq package. These FPKM values were further used to calculate the log fold change [$\log_2$ (FPKM_JV/FPKM_JH)]. The analysis was carried out to identify commonly expressed genes between JH and JV samples, respectively. These genes were divided on the basis of their statistical significance (depending on whether p value is less than 0.05 for their significant expression). These genes were further categorized as upregulated and downregulated in JV as compared to JH. Genes that exhibited p value <0.05 and estimated absolute log2 fold were determined to be significantly expressed genes.

Heat Map Analysis

A complete linkage hierarchical cluster analysis was performed on top 100 differentially expressed genes (50 upregulated and 50 downregulated in JH or JV) obtained from DESeq using Multiple experiment viewer (MEV v4.8.1) (Fig. 1). Heat map was constructed using log-transformed and normalized value of genes based on Pearson's uncentered correlation distance as well as based on complete linkage method.

Gene Ontology Analysis

For functional annotation of the predicted CDS in both JV and JH, BLAST2GO program was used with default parameters to retrieve GO annotation which distinguishes on the basis of molecular function, biological process, and cellular component ontologies [26]. Gene ontology analysis specifies all the annotated nodes comprising of GO functional groups such as cellular component, biological process, and molecular function. Main GO categories were determined after the genes were further analyzed for BLAST, gene mapping, and annotation.

Pathway Analysis

CDSs of JV and JH samples were functionally annotated by KEGG Automatic Annotation Server (KAAS) with BLAST comparisons against KEGG GENES database. KEGG GENES database has an advantage over other databases, as it is a single resource for cross-species depiction by assigning KEGG orthology to all existing genomes. The bi-directional best hit (BBH) option was used to assign KO terms. For pathway mapping, KEGG Orthology database (http://www.genome.jp/kegg/ko.html) was used.

Coexpression Network Analysis

A gene coexpression network (GCN) is an undirected graph, where each node corresponds to a gene, and a pair of nodes is connected with an edge if there is a significant coexpression relationship between them [27]. BLAST analysis was performed for transcriptome sequences of both JH and JV samples with NCBI nonredundant database. Annotated Jatropha scaffold ID were mapped with sequences and genes were extracted from annotated list on the basis of fold change (threshold twofold) using in house perl scripts. Constructing gene coexpression network involved combination of Pearson's correlation coefficient (PCC), gene ontology score, abundance score, and codon score that was calculated as the geometric mean of the correlation rank of gene X to gene Y and of gene Y to gene X. Cytoscape network construction and Network analyzer plugin were used for coexpression network analysis.

RT-qPCR Based Experimental Validation

For experimental validation, RT-qPCR approach was used to confirm the transcriptome data. First-strand cDNA synthesis was done using Verso cDNA synthesis kit (Thermo Scientific, USA) from independent RNA samples isolated from the same tissues in two biological replicates, which were used for RNA isolation in transcriptome sequencing, as per manufacturer's instructions. The expression profiles were investigated by RT-qPCR. Using gene-specific primers, RT-qPCR was performed on CFX96 system (Bio-Rad, USA) with the iScript one step RT PCR kit (Bio-Rad, USA). The PCR protocol was as follows: denaturation for 5 min at 94 °C, followed by 40 cycles each of denaturation for 20 s at 94 °C, annealing for 30 s at 50–55 °C, followed by one elongation step for 20 s at 72 °C. All quantitative PCR experiments were repeated in triplicate. For calculating transcript abundances, 26S rRNA and GAPDH were used as internal controls. Standard deviation with percentage error was used to statistically evaluate significant differences between treatments.

Results

Identification of Virus

The leaves of virus-infected plants (JV) were showing symptomatic conditions like mosaic, blistering, and mottling in the experimental farm. Leaves were characterized by reduced size, chlorotic spots, and rolling [5, 7] (Supplementary Fig. S1). We also observed overall reduction in fruit number and size in plants infected with mosaic virus as compared to healthy plants. Upon PCR amplification with mosaic virus coat protein gene, an amplicon of 700 bp was clearly observed in JV sample as compared to JH which confirmed the presence of virus in the infected leaf sample.

Reduction in Fruits Size, Seed Yield, and Oil Content in Response to Virus Infection

The data on parameters like fruit size, seed yield, and oil content was recorded consecutively for 2 years for healthy and virus-infected plants (100 plants in each category) of high oil content genotype IC 561235 at experimental farm of Himalayan Forest Research Institute, Jwalaji (Himachal Pradesh). It was observed that the plants infected with mosaic virus (JV) had reduced fruit size characterized with dark brown colored spots as compared to healthy plants (JH). The mature fruit showed the presence of one to two seeds in plants with virus infection as compared to healthy plants having three seeds on average. Further, the overall number and weight of seeds was less in virus-infected plants compared to healthy plants (JH) (Table 1). We also observed reduction in oil content on comparing virus-infected (JV) and healthy (JH) plants; however, it was not significant. The oil content in JH and JV was 42 and 37 %, respectively (Table 1).

Transcriptome Sequencing, Data Generation, and Differential Gene Expression Analysis

The libraries were sequenced on NextSeq 500 platform of Illumina using 2 × 150 PE Chemistry. The raw reads were filtered using Trimmomatic v0.30 to filter out the adaptor

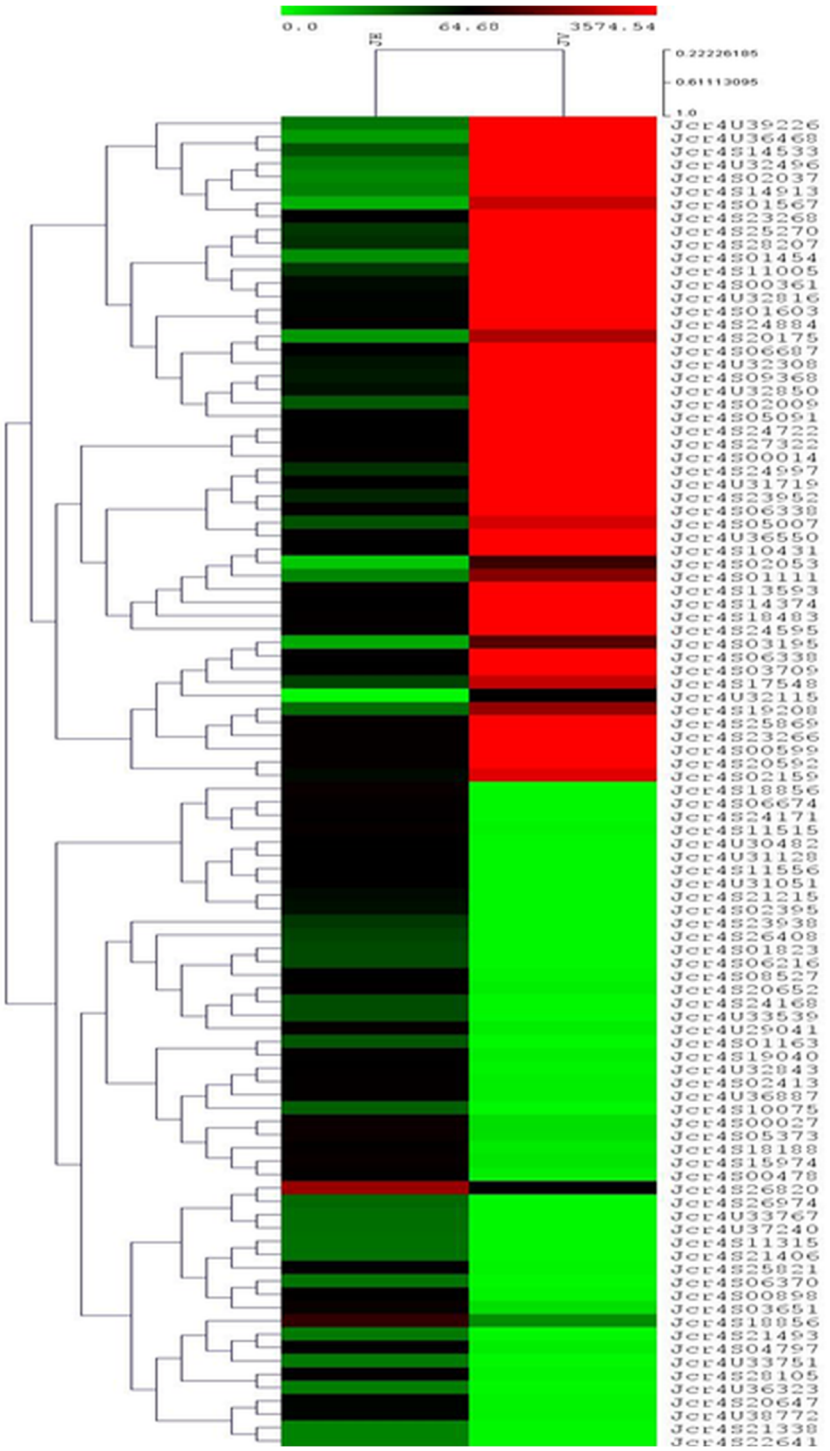

Fig. 1 Heat map representing top 100 differentially expressed genes (50 upregulated and 50 downregulated in JH or JV)

Table 1 Effect of mosaic virus infection on Jatropha yield and oil content (average data of total 100 plants for each category, i.e., virus-infected plants (JV) and healthy plants (JH) for two consecutive years; ± values denote SEM)

Yield parameters	Healthy plant (JH)	Virus-infected plant (JV)
Seeds per fruit	3 ± 0.5	1 ± 0.5
Number of seeds per plant	620 ± 5	280 ± 2.5
Weight of seeds per plant (g)	400 ± 4	190 ± 2.5
100-seed weight (g)	60 ± 5	39.98 ± 1
Oil content (%)	42 ± 0.45	37 ± 0.25

SEM standard error of the mean

contamination and low-quality (reads with QV < 20) reads. The high-quality reads (QV > 20) were used for the RNA-Seq analysis. A total of 10,548,434 and 12,226,847 high-quality reads were obtained in JH and JV, respectively (Table 2). High-quality reads were mapped on to the reference genome using mapping software BWA with optimized parameters. Further, the expression analysis of genes was carried out with R package DESeq and was calculated in terms of FPKM. Upon annotation, it was observed that a total of 55,755 transcripts associated to multiple pathways were expressed in both the samples, i.e., JH and JV on the basis of FPKM values. Six hundred nineteen and 330 transcripts were expressed uniquely in JH and JV, respectively (Fig. 2a), whereas 685 and 2132 transcripts were upregulated and downregulated, respectively in JV (Fig. 2b).

Gene Ontology, KEGG Analysis Based Functional Classification, and Identification of Pathways Upregulated in Response to Viral Infection

Gene ontology (GO) is a standardized gene functional cataloging system whose terms are derived from ontologies that can be used to describe the function of genes and their products in any organism. Three ontologies, i.e., molecular function, cellular component, and biological process, are linked to GO database which are the mainstay of any GO annotation. The predicted CDS in response to virus infection (JV) and healthy leaf (JH) were annotated by BLAST2GO to know about functional classification. The analysis revealed that majority of assignments were in category "biological process" followed by "molecular function" and "cellular component" categories in both JV and JH (Figs. 3 and 4). Further, KAAS was used for functional annotation in both the samples with BLAST comparisons against KEGG GENES database. KEGG Orthology database was used for pathway mapping. Upon annotation, maximum number of transcripts, i.e., 773, corresponded to the category "Signal transduction" followed by 704 in "Carbohydrate metabolism" and 591 in "Translation" (Supplementary Fig. S2). The other associated pathways are provided in the supplementary file 2. Under the pathway category "Environmental adaptation," there is a subcategory "Plant-pathogen interaction" which had 127 transcripts. Upon

Table 2 Statistics of generated reads

Sample	Number of reads	Number of high-quality reads (PE)	Total data (Gb)
JH	20,950,364	10,548,434	3
JV	19,958,750	12,226,847	3.50

PE paired end, *Gb* gigabytes

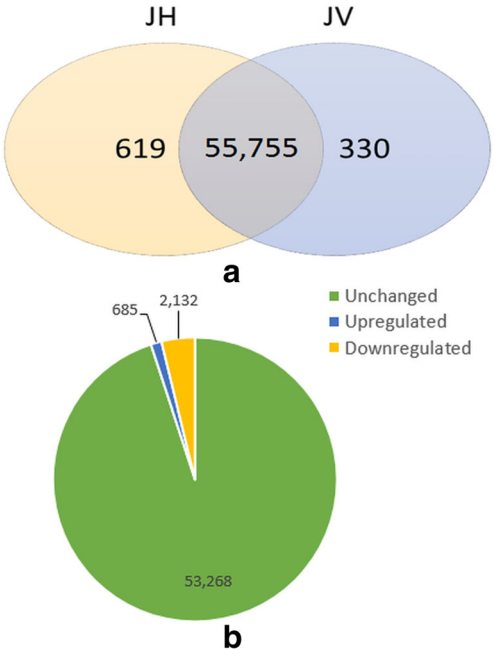

Fig. 2 **a** Venn diagram showing distribution of genes expressed in JH and JV. **b** Pie chart showing distribution of total upregulated and downregulated genes in JV

KEGG annotation, it was observed that a total of 23 different pathways were upregulated in response to virus infection in JV as revealed by the genes associated to them. Out of these, majority of upregulated genes mapped to pathways such as oxidative phosphorylation, endocytosis, arginine and proline metabolism, terpenoid biosynthesis, ascorbate metabolism, amino sugar and nucleotide sugar metabolism, and lipid metabolism which implied that these were the main pathways upregulated in JV (Fig. 3). The most affected process was the metabolism process, as most of the annotated genes were correlated to metabolism processes.

Oxidative Phosphorylation

It was observed that 80 genes involved in "oxidative phosphorylation" were upregulated in response to viral infection (JV) (Fig. 5). Genes such as ATPase, oxidoreductase, oxidase, and dehydrogenase were significantly overexpressed (Supplementary Table S1).

Endocytosis

A total of 40 genes associated with "endocytosis" process showed higher transcript abundance in response to viral infection (JV) (Fig. 5) out of which charged multivesicular body protein 5, Ras-related protein Rab-11A, epsin, and DnaJ homolog subfamily C member were significantly overexpressed in response to viral infection (Supplementary Table S1).

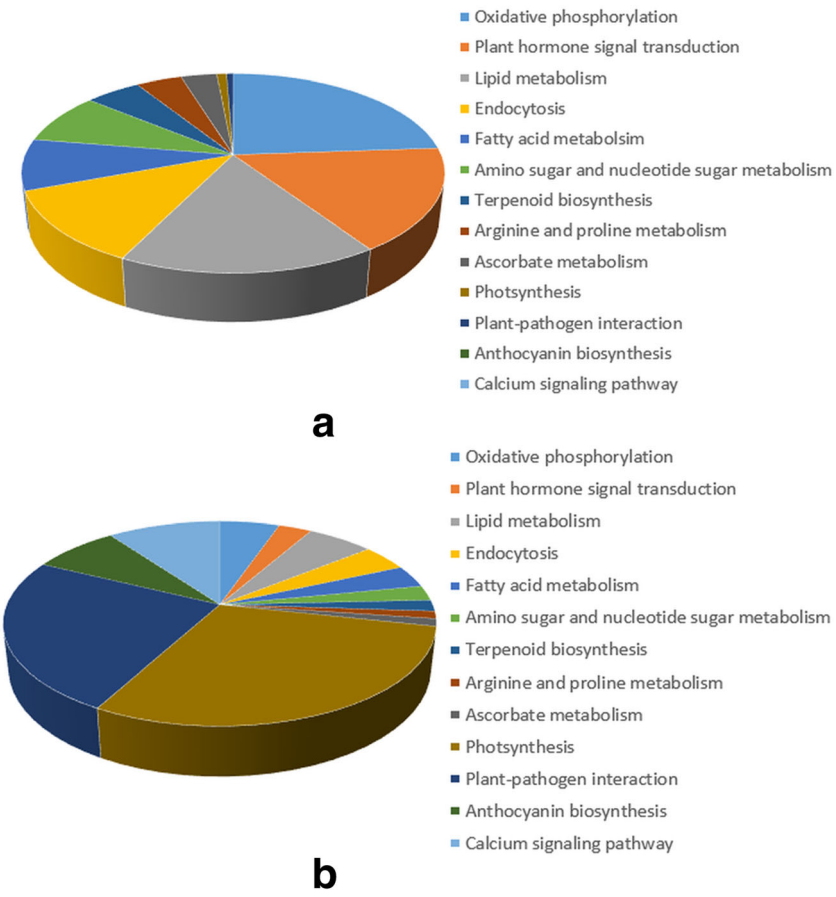

Fig. 3 Pathways upregulated (**a**) and downregulated (**b**) in response to virus infection

Metabolism of Amino Acids and Vitamins

Genes involved in metabolism of arginine and proline (amino acids) and ascorbate (vitamins) were upregulated in response to viral infection (JV) as compared to healthy (JH). Fourteen genes linked to "arginine and proline metabolism" showed upregulation in response to viral infection where genes such as nitric-oxide synthase and prolyl 4-hydroxylase were significantly overexpressed. For "ascorbate metabolism," 11 genes were upregulated in JV (Fig. 5). Genes such as GDP-L-galactose phosphorylase and L-ascorbate peroxidase were significantly upregulated in response to viral infection (Supplementary Tables S3 and S4).

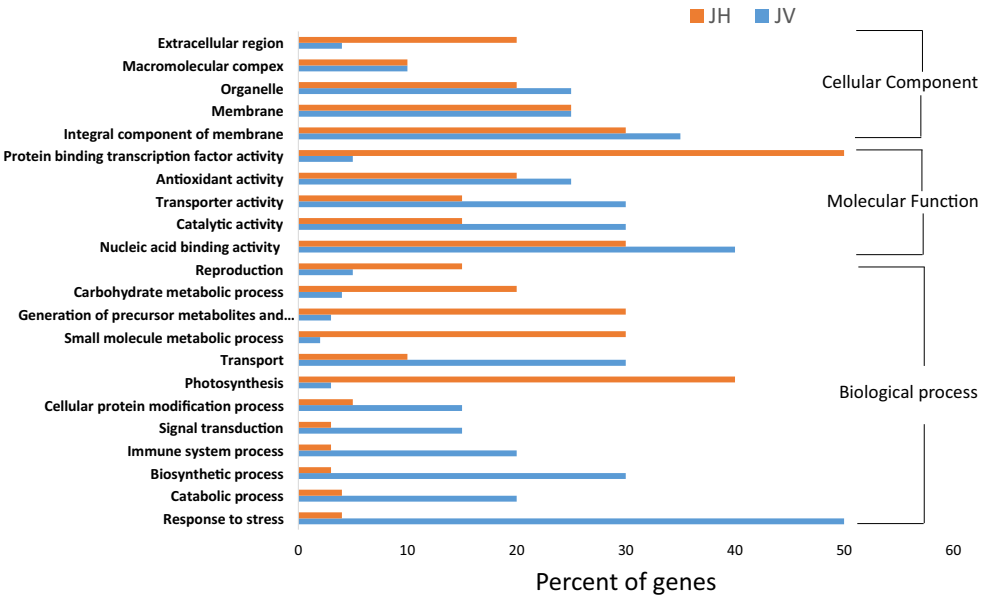

Fig. 4 GO classification and distribution of GO annotated genes in JV and JH

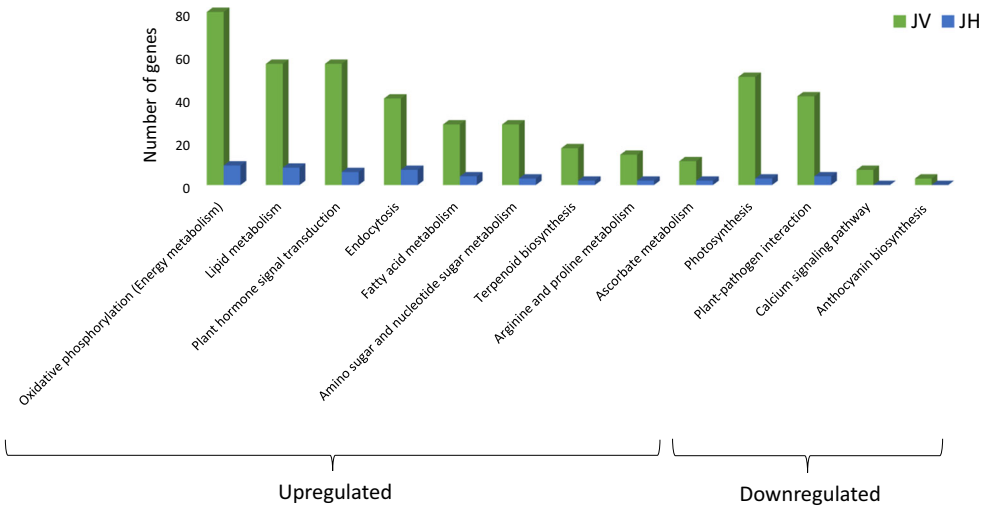

Fig. 5 Number of upregulated and downregulated genes in JV and JH

Fatty Acid and Lipid Catabolism

A total of 56 and 28 genes related to lipid catabolism and fatty acid degradation, respectively, were found to be upregulated in JV (Fig. 5). For lipid catabolism, the significantly overexpressed genes in JV were lipase and kinase. In case of fatty acid catabolism, genes such as acetyl-CoA acyltransferase, long-chain acyl-CoA synthetase, very-long-chain enoyl-CoA reductase, and acyl-carrier-protein desaturase showed significantly higher expression in response to viral infection (Supplementary Tables S5 and S6).

Amino Sugar and Nucleotide Sugar Metabolism

It was observed that 28 genes related to sugar metabolism (synthesis) were upregulated in JV as per transcript abundance (Fig. 5). In response to viral infection, the expression level of genes such as UDP-apiose/xylose synthase and L-arabinokinase were significantly elevated (Supplementary Table S7).

Terpenoid Biosynthesis

Seventeen genes related to biosynthesis of monoterpenoid, diterpenoid, and triterpenoid biosynthesis showed higher transcript abundance in response to viral infection (JV) (Fig. 5). Genes such as synthases, oxidases, and dehydrogenases were significantly enriched in response to viral infection (Supplementary Table S8).

Signal Transduction of Hormones

Further, upon KEGG-based functional annotation, we identified genes upregulated in "plant hormone signal transduction" in response to virus infection. Total of 56 genes were upregulated in JV in response to virus infection (Fig. 5). The majority of genes upregulated belonged to the families involved in hormone signaling (Table 3). This implied that signaling of various plant hormones such as salicylic acid (SA), ethylene, jasmonic acid (JA), abscisic acid (ABA), and auxin is activated during viral infection (Supplementary Table S9).

Table 3 Gene families of "plant hormone signal transduction" activated in JV

Gene family	Upregulated/ downregulated
SAUR family protein	+
Auxin- responsive GH3 gene family	+
Transport inhibitor response 1	−
Pathogenesis-related protein 1	−
Auxin-responsive protein IAA	+
DELLA protein	−
F-box protein GID2	−
Two-component response regulator ARR-A family	+
Auxin influx carrier (AUX1 LAX family)	+
Arabidopsis histidine kinase 2/3/4 (cytokinin receptor)	+
Abscisic acid receptor PYR/PYL family	+
Jasmonic acid-amino synthetase	+
Protein phosphatase 2C	−
BR-signaling kinase	−
Jasmonate ZIM domain-containing protein	+
Regulatory protein NPR1	−
Two-component response regulator ARR-B family	+
Ethylene-insensitive protein	+
BRI1 kinase inhibitor 1	−

+upregulated, − downregulated)

Identification of Pathways Downregulated in Response to Viral Infection

The pathways which were downregulated in response to viral infection (JV) were photosynthesis, anthocyanin biosynthesis, plant-pathogen interaction, and calcium signaling pathway (Fig. 3).

Photosynthesis

Fifty genes involved in "photosynthesis" showed low transcript abundance in JV (Fig. 5). Significantly repressed genes in response to viral infection were genes associated with PSI and PSII, genes related to light-harvesting complex, LHC I and LHC II, ferredoxin, and cytochrome b6 (Supplementary Table S10).

Anthocyanin Biosynthesis

Only three genes linked to "anthocyanin biosynthesis" were identified on the basis of low transcript abundance in response to viral infection (Fig. 5). The genes were anthocyanidin 3-O-glucoside 2-O-xylosyltransferase and anthocyanidin 3-O-glucoside 5-O-glucosyltransferase (Supplementary Table S11).

Plant-Pathogen Interaction

Based on KEGG pathway assignments, we observed that total of 41 genes linked with "plant-pathogen interaction" were downregulated in JV (Fig. 5). Genes related to innate immunity such as cyclic nucleotide-gated ion channels (CNGCs), calcium-binding protein (CML) disease resistance proteins, i.e.,RPM1, RPS2, kinases, and defense-related transcription factor genes such as WRKY were significantly repressed in JV as compared to JH (Supplementary Table S12).

Calcium Signaling Pathway

A total of seven genes linked to calcium signaling pathway showed low transcript abundance in response to viral infection (JV) (Fig. 5). Genes such as solute carrier family 25 (mitochondrial adenine nucleotide translocator) and voltage-dependent anion channel protein 2 were significantly repressed in JV (Supplementary Table S13).

Identification of Transcription Factors

Upon KEGG-based functional annotation, we observed that 21 transcription factors linked to signal transduction of plant hormones were upregulated in JV. The upregulated transcription factors were of MYC2 (3), TGA (7), ABA-responsive element binding factor (9), and ethylene-responsive transcription factor (2) TF families. Furthermore, three transcription factors related to WRKY family linked with plant-pathogen interaction were significantly downregulated in JV.

Coexpression Network Analysis

For coexpression network analysis, genes from one upregulated pathway and one downregulated pathway in response to viral infection were considered. The genes linked to plant hormone signal transduction and plant-pathogen interaction were analyzed to identify genes coexpressed and showing interactions with identified reference genes on the basis of scoring function (Figs. 6, 7 and 8). From plant hormone signal transduction pathway, HAB1 had interaction with ORE14, RKP, REF4, ZIG4A, SIZ1, LRS1, UPL3, DegP7, KEA2, IBR3, and RR1 whereas closely related to ZIG4A and SIZ1. Reference gene CML13 from plant-pathogen interaction showed interaction with genes such as SCN1, NTF2A, PRA1.E, RABA2c, PRF1, UBC22, FAH2, RAB6A, SBH2, WPP1, WPP2, and VHA-G2. SCN1, NTF2A, PRF1, and WPP2 were closely related to reference gene, CML13. The interaction of other identified reference genes from both the pathways to other coexpressed genes are shown in Supplementary Figs. S2, S3, S4, S5, S6, S7, S8, S9, S10, S11, and S12.

RT-qPCR-Based Validation

For experimental validation, a set of seven genes involved in plant hormone signal transduction showing about and more than threefold for FPKM values in JV as compared to JH, was examined using RT-qPCR to confirm the transcriptome data. The genes from plant hormone signal transduction were selected for experimental validation as maximum identified genes belonged to this process. Identified transcription factors showing upregulation and downregulation were also analyzed using RT-qPCR to confirm the transcriptome data. All seven genes showed higher expression in JV as compared to JH. For example, fold expression of 16 was observed in JV for gene PHST_1 (Jasmonate ZIM domain-containing protein) as compared to fourfold expression value in JH. Maximum fold expression value of 20 was observed for gene PHST_2 (auxin-responsive GH3 gene family) in case of JV whereas the fold expression value was 7.5 in JH (Fig. 6a). For transcription factors, MYC2 showed fold expression value of 3.5 in JV whereas 1.2 in JH. TGA exhibited fold expression of 4.5 in JV and 2 in JH. ABA-responsive element binding factor showed fold expression of 3.3 in JV followed by 1.4 in JH. Similarly, ethylene-responsive transcription factor had fold expression of 2.2 in JV and onefold expression value in JH. On the other hand, WRKY showed higher expression in JH (6.5) and low fold expression of 2.2 in JV (Fig. 6c). Computational identification of transcript

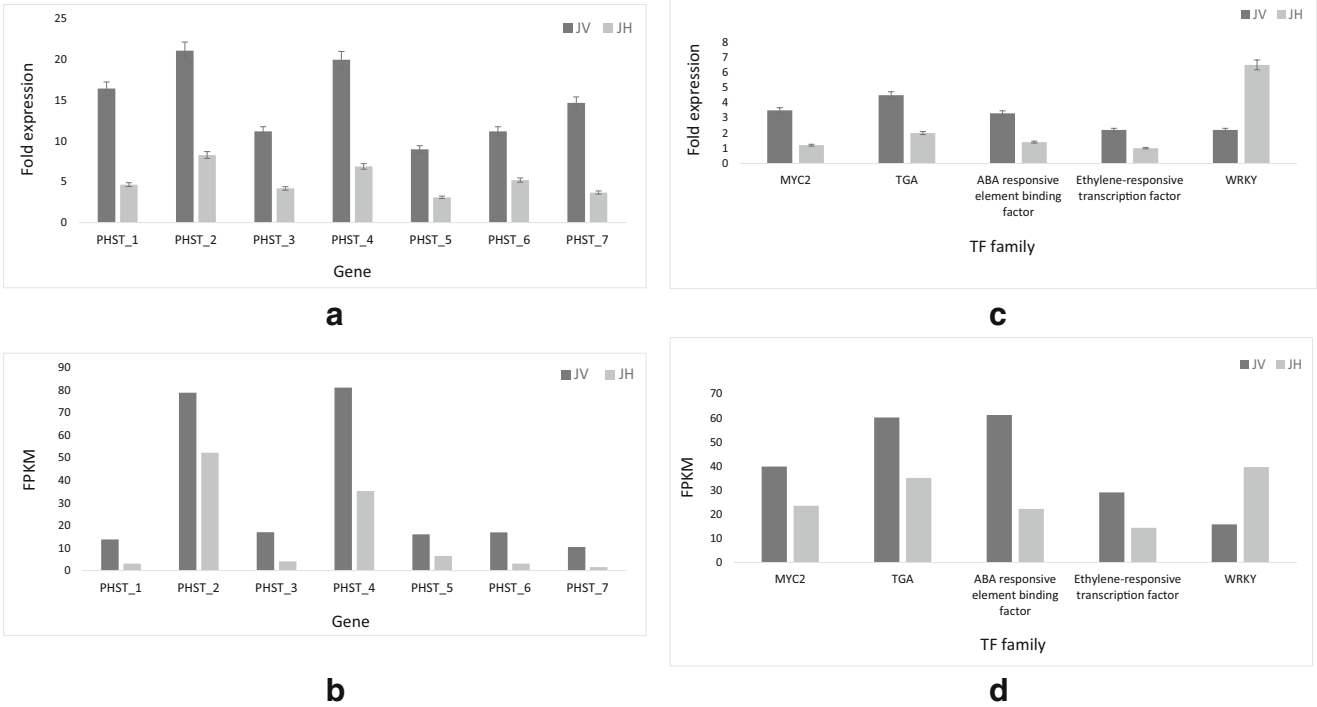

Fig. 6 **a** RT-qPCR-based fold expression pattern of genes involved in signal transduction of hormones showing about and more than two fold in JV as compared to JH. **b** In silico abundance (FPKM values) of genes involved in signal transduction of hormones showing about and more than two fold in JV as compared to JH (PHST_1: Jasmonate ZIM domain-containing protein, PHST_2: auxin-responsive GH3 gene family, PHST_3: two-component response regulator ARR-A family, PHST_4: auxin-responsive protein IAA, PHST_5: abscisic acid receptor PYR/PYL family, PHST_6: two-component response regulator ARR-B family, PHST_7: SAUR family protein). **c** RT-qPCR-based fold expression pattern of transcription factors upregulated and downregulated during virus infection. **d** In silico abundance (FPKM values) of transcription factors upregulated and downregulated during virus infection

abundance and validation with RT-qPCR are in conjunction with each other (Fig. 6b, d).

Accessibility to Transcriptome Data

The raw reads and annotated data, i.e., CDSs for JH and JV are available at http://14.139.240.55/NGS/download.php.

Discussion

Promising genotypes of Jatropha are being commercially cultivated on large scale which lead to predisposition to biotic stresses. As Jatropha is becoming vulnerable to various diseases, specifically diseases caused by viruses, identification of molecular insights to comprehend disease response is crucial

Reference Genes	Coexpressed Genes												
HAB1	ORE14	RKP	REF4	ZIGA4	SIZ1	LRS1	UPL3	DegP7	KEA2	IBR3	RR1		
IAA3	XCP1	FAF3	FLA11	TRX1	ATAO1	XYP2	SAH7	SBTI1.1	AtAGAL2	TBL3	820169		
CML42	TEM2	PRA8	EDL2	ZHD6									
CML30	UBC17	MCP1a	RTM1	PP2-A1	WDY								
IAA13	TSD1	WAT1	IAA4	ACL5	PAP2	IAA12	IAA9	BRH1	MYC4	SAMDC4	DOF2	SHY2	TMO5
BES1	MPK6	BON2	SBI1	RSW9	WAV2	NADK1	UBP14	GNTI	bHLH121	RFI2			
FCA	SCC2	RIK	FC1	PEX1	SAC3B	REV1	SET29	EDM2					
RR17	BGLU32	AGL78											
CML13	SCN1	NTF2A	PRA1.E	RABA2c	PRF1	UBC22	FAH2	RAB6A	SBH2	842627	WPP1	WPP2	VHA-G2
ABF1	HSFA8	STH3	GI	CYCT1.3	HSFA1B	PSD3	WRKY3	TAG1	PRR5	BET9			
RIN4	ECH2	GG1	ZAR1	SOBIR1	ERD6	PEN3	XBAT34	PMR2	BIR1	CRK19			
PBS1	SYP123	RSH3	CYP89A7	RHS10	RHS18	RHS8	IRE	EXO70C2	TET12				
CML41	SYP123	RSH3	CYP89A7	RHS10	RHS12	RHS18	RHS8	IRE	EXO70C2	TET12			
CPK1	DIS1	PME41	MLO8	CPK4	XI-6	ACA2	PEPR1	MUR3	AtKINUb				
CML38	ATAGP1	WRKY33	CML39	SCABP5	WRKY40	SRO5	ZAT10	ERF11	ERF2	DIC2	SZF1	AR781	PP2-A11

Plant Hormone Signal Transduction

Plant Pathogen Interaction

Fig. 7 List of coexpressed genes in "plant pathogen interaction" and "plant hormone signal transduction" pathways. Genes in red colour blocks showing close association with reference genes. Reference genes in *bold letters* representing homologs

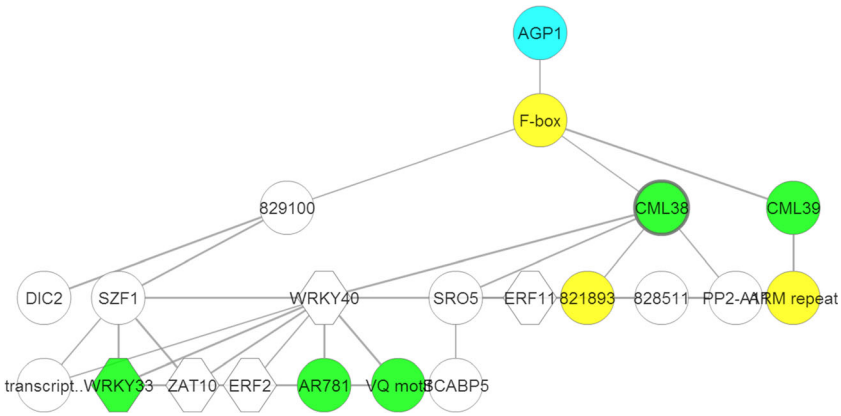

Fig. 8 Coexpression network analysis of gene CML38 (plant-pathogen interaction)

to understand which biological processes are affected. In this study, we gave new dimensions to virus response in Jatropha through transcriptomic-based comparative analysis of healthy and virus-infected leaves. On characterization and comparison of transcriptomes of healthy versus virus-infected leaves, we found genes and transcription factors linked to specific biological processes being upregulated and downregulated in response to mosaic virus infection in *J. curcas*.

Reduction in Seed Yield and Oil Content due to Mosaic Disease in *J. curcas*

Data pertaining to seed yield and oil content parameters was recorded consecutively for two years in mosaic virus infected (JV) and healthy plant (JH). It was observed that different parameters related to seed yield, i.e., seeds per fruit, number of seeds per plant, weight of seeds per plant, and 100-seed weight, showed overall reduction in plants infected with virus as compared to healthy one. On oil content estimation, there was 5 % reduction in total oil content of virus-infected plants in comparison to healthy plants, on an average. These observations for reduction in seed yield and oil content are in conjunction to the previous studies of effects of virus infection in *J. curcas* [28, 29]. The possible reason for this reduction in yield and oil content due to virus infection is the reduced rate of photosynthesis which has been discussed later.

Enhanced Energy Metabolism and Activation of Endocytotic Genes During Viral Infection

Upon functional annotation via pathway mapping with KEGG, we observed that the metabolism processes are affected the most in response to viral infection as most of the genes related to metabolism processes were upregulated or downregulated. In our analysis, we found that energy metabolism (oxidative phosphorylation) was upregulated in response to viral infection. In photosynthesis, light energy is captured and converted into ATP and NADPH. These ATP and NADPH act as energy sources for various physiological processes. We observed upregulation of genes such as NAD(P)H-quinone oxidoreductase, ubiquinol-cytochrome c reductase, NAD(P)H-quinone oxidoreductase, F-type H+-transporting ATPase, NADH dehydrogenase, and cytochrome c oxidase during virus infection which suggested that organelles like mitochondria produce energy to drive cellular processes, which has been observed in virus infection in rice and tobacco [10, 30]. However, our observations are deviated from the assumption which implied termination of processes for generation of plant energy in response to disease and infection [31]. Further, endocytosis was found to be significantly enriched in response to viral infection. Endocytosis is a cellular process in which cells internalize extracellular material or foreign particles for recycling or degradation [32]. During virus infection, the host cells also destroy pathogens by engulfing them [32] which has been evidenced by the upregulation of genes involved in endocytosis in our analysis. Genes regulating endocytosis such as charged multivesicular body protein 5, Ras-related protein Rab-11A, epsin, and DnaJ homolog subfamily C member 6 were upregulated in response to mosaic viral infection [33–35]. These results are consistent with previous report of transcriptional response for virus infection in maize [32].

Metabolism of Amino Acids and Vitamins Is Induced in Response to Viral Infection

Synthesis of amino acids, i.e., arginine and proline was induced in response to viral infection as genes linked to arginine and proline metabolism were significantly overexpressed. On exposure to specific infection or pathogen, plants produce reactive oxygen species (ROS) to programmed cell death and to terminate the disease process [36]. Proline and arginine acts as potential scavengers of ROS and thus prevent the function of ROS. We observed that genes involved in the synthesis of these amino acids were upregulated in response to mosaic virus infection which indicates potential role of these amino acids in infection. In response to viral infection, genes nitric-oxide synthase and prolyl 4-hydroxylase were significantly

overexpressed. Nitric-oxide synthase has been reported to catalyze the production of arginine whereas prolyl 4-hydroxylase has been associated to synthesize proline [37, 38]. The genes related to arginine and proline metabolism have been previously shown to be linked with biotic and abiotic stresses in various plant species like Arabidopsis to tobacco etch virus infection, cotton to aphid infestation, and eucalyptus to cold stress [39, 40]. Biosynthesis of vitamins such as ascorbic acid was also induced in response to mosaic viral infection as evidenced by the upregulation of genes involved in ascorbate metabolism. Deficiency of ascorbic acid leads to the activation of cell death and disease resistance response in plants [41]. Genes involved in biosynthesis of ascorbic acid such as GDP-L-galactose phosphorylase and L-ascorbate peroxidase were significantly overexpressed in response to mosaic virus infection suggesting potential role of ascorbic acid in disease induction in plants.

Catabolism of Fatty Acids and Lipids Is Associated to Sugar Biosynthesis in Response to Viral Infection

Lipids and fatty acids regulate plant defense response against various pathogens as they act as signaling molecules [42, 43]. We observed upregulation for fatty acid and lipid catabolism upon mosaic viral infection as the genes involved in lipid metabolism and fatty acid metabolism showed higher transcript abundance. These results are supported by the fact that the *J. curcas* mosaic virus disease (JcMD) reduces the overall oil content in *J. curcas* [5]. Genes involved in fatty acid catabolism such as acetyl-CoA acyltransferase, long-chain acyl-CoA synthetase, acyl-carrier-protein desaturase, and in lipid catabolism, i.e., lipase were significantly upregulated in response to mosaic viral infection. These results are in line with previous report by Freitas-Astúa et al. [44] suggesting that fatty acid and lipid metabolism are important for the susceptibility of virus infection and diseases. In plants, starch (sugar) is accumulated during the day and used in dark hours to provide energy for key metabolic processes [45]. There is a reciprocal relationship between sugar biosynthesis and oil biosynthesis [46]. Genes regulating sugar metabolism were found to be upregulated in response to mosaic virus infection where UDP-apiose/xylose synthase and L-arabinokinase were mainly overexpressed. Both UDP-apiose/xylose synthase and L-arabinokinase are involved in sugar biosynthesis [47, 48].

Terpenoids Functions as Plant Growth Regulators During Viral Infection

In plant kingdom, terpenoids function in defense mechanisms against a broad range of pathogens [49]. Plants interact with pathogens through some signaling molecules such as terpenoid metabolites. We observed higher transcript abundance of diterpenoid biosynthesis genes such as gibberellin 2-oxidase and gibberellin 3-beta-dioxygenase in response to mosaic virus infection. Our observations are in agreement with assumption that during pathogen infection, formation of some diterpenoid and sesquiterpene metabolites is induced as plant growth regulators (gibberellins) and phytoalexins [49, 50]. Genes encoding synthase enzymes catalyzing the formation of monoterpenoid, sesquiterpenoid, and triterpenoid biosynthesis were upregulated in mosaic viral infection. These observations are similar to what has been reported previously for virus infestation in tobacco and cassava [10, 51].

Hormones Signaling Is Enhanced During Virus Infection

During *Jc*MD infection in *J. curcas*, various physiological abnormalities occurs such as leaves get curled, blistered, and mottled [5]. Appearance of these physiological abnormalities and symptoms upon virus infection is due to the interplay of various phytohormones which regulate plant growth and development. During the course of virus infection, some cytopathic effects occur which are supposed to alter the normal plant growth [52], which may be due to alterations in plant hormone metabolism. During mosaic virus infection, we observed higher transcript abundance of genes, i.e., SAUR family, auxin-responsive GH3 gene, and auxin-responsive protein IAA, which regulate signaling of auxins. For auxins, our results supported the fact that substantial rise in activity of auxins sometimes causes severe symptomatology during infection [53]. Abscisic acid gets more accumulated during infection with viruses [54]. The upregulation of genes such as abscisic acid receptor PYR/PYL family involved in the signaling of abscisic acid, in response to mosaic viral infection, was observed. Ethylene-insensitive protein 3, a gene involved in the ethylene signaling, was also overexpressed in response to mosaic virus infection, suggesting possible role of ethylene signaling in virus accumulation and infection [55]. Another plant hormone, Jasmonic acid, was observed to be a negative regulator of infection and positive regulator of resistance against mosaic virus as evidenced by the upregulation of gene, i.e., jasmonate ZIM domain-containing protein repressing the signaling of Jasmonic acid [56]. Thus, our results implied that signaling of various plant hormones such as abscisic acid (ABA), auxin, and ethylene is activated during viral infection.

Repressed Photosynthesis During Virus Infection Is Associated with Reduction in Seed Yield and Oil Content

We also observed downregulation of major pathways related to overall growth and development in *J. curcas* due to mosaic virus infection. Photosynthesis, a major developmental process, was significantly repressed in response to mosaic virus infection in *J. curcas*. In photosynthesis, light energy is captured and converted into ATP and reducing power (NADPH). During mosaic virus infection, the overall chlorophyll in the

leaves gets degraded due to the induction of chlorophyllase activity [11]. Also, there is the reduction in leaf area followed by chlorotic spots which also correlate with the degradation of chlorophyll. Protein complexes regulating photosynthesis have photosystems (PSI and PSII) as functional and structural units. In photosynthesis, these photosystems help to regulate the primary photochemistry of photosynthesis, absorption of light, and energy transfer. PSI and PSII absorbs photons of a wavelength of 700 and 680 nm, respectively. Electrons flow from PSII to PSI through cytochrome b6 (a membrane bound protein). Genes related to PSI, PSII, and cytochrome b6 were downregulated which indicates less photosynthesis during viral infection in *J. curcas*. Along with chlorophyll degradation, there is also deficiency of light-harvesting complex (LHC) [11]. LHC gather light energy and transfer this energy to the photosynthetic reaction centers [57]. LHC is composed of LHC proteins that bind light-harvesting pigments. During mosaic virus infection, we observed less transcript abundance of light-harvesting complex I chlorophyll a/b binding protein and light-harvesting complex II chlorophyll a/b binding protein genes. We also observed downregulation of ferredoxin as this gene functions principally in photosynthesis. Electrons are transferred from photoreduced PSI to ferredoxin NADP(+) reductase by ferredoxin [58]. Further repression of photosynthesis in virus infection condition was correlated with the reduction in fruit size, seed yield, and oil content as per our observations. ATP and NADPH, the photosynthesis products, are utilized in CO_2 fixation that provides carbon skeletons for all cellular reactions [59]. Light reactions in photosynthetic reactions feed ATP and NADPH to carbon fixation. Since photosynthesis and carbon fixation is repressed, there is less partitioning of carbon toward lipid accumulation which might be responsible for reduction in oil content during virus infection. The reduction in fruit size and seed yield is due to the fact that regulation of photosynthetic reactions is essential for the metabolic reactions. Further nitrogen can improve photosynthetic parameters, increase maximal photochemical efficiency, and reduce fluorescent and nonphotochemical quenching coefficiency and as a result, increase fruit and seed yield with higher seed filling rate [60]. Also, solar radiation is a major factor which affects the uptake of nutrient solution and growth processes during photosynthesis.

Degradation of Anthocyanin in Viral Infection

Anthocyanins (flavonoids) are water-soluble pigments which are synthesized in the cytosol and localized in the vacuoles. During mosaic virus infection, the leaves get curled, reduced in size, and become chlorotic which lead to a significant degradation of pigmentation. Genes, anthocyanidin 3-O-glucoside 2-O-xylosyltransferase and anthocyanidin 3-O-glucoside 5-O-glucosyltransferase, involved in anthocyanin biosynthesis showed less abundance, supporting the degradation of anthocyanin during infection. These observations are in accordance with previous studies for tobacco and grapes [10, 61]. Our results further supported the fact that anthocyanin also regulate defense response in plants [62].

Repression of Defense Mechanism and Contribution of Host Factors Toward Replication and Multiplication of Virus During Infection

We further observed that various pathways related to defense response were significantly repressed during mosaic virus infection in *J. curcas*. Genes linked to plant-pathogen interaction, a process related to basal defense response, showed more expression in healthy leaf tissue as compared to mosaic virus-infected leaf tissue which directed its repression during infection. Various genes in the plant-pathogen interaction pathway were downregulated during mosaic virus infection. Genes such as CNGCs, CMLs, disease resistance proteins, i.e., RPM1, RPS2, kinases, and WRKY transcription factor genes were significantly repressed in viral infection, which have been implicated in defense response and innate immunity in other plants. CNGCs have been reported for their role in immunity and have been characterized in plant species [63–65]. CMLs are also involved in providing immunity against various biotic stresses [66, 67]. Role of various kinases in providing innate immunity has also been described in plants [68–70]. Also various reports exist for WRKY transcription factors providing innate immunity against biotic stresses [71, 72]. Genes involved in calcium signaling pathway also showed low transcript abundance in response to infection clearly indicating the role of calcium signaling in basal defense response. Also, calcium signaling could be controlled by other signaling systems like ubiquitin-proteasome system to mark immunity in plants against pathogens [73]. In order to complete life cycle, viruses are evolutionarily able to capture and manipulate cellular pathways and cellular components. Fatty acids are considered as vital constituents of cellular membrane lipids which regulate the fluidity and integrity of cell membranes. Also, various lipid groups can be covalently joined to proteins, a process known as protein lipidation. As viruses depend on intracellular membranes for replication, the process of protein lipidation is likely to influence the replication of viruses [74]. Fatty acid synthases are the enzymes having role in the replication of viruses [75]. We observed downregulation of genes encoding fatty acid synthases in response to virus infection. The subunits of gene "3-oxoacyl-[acyl-carrier-protein] synthase" involved in the condensation step of fatty acid biosynthesis showed less transcript abundance during virus infection, suggesting their possible role in viral replication. Also, the presence of gene-encoding enzymes catalyzing the formation of lipids was downregulated during infection which supports the role of lipids in causing an infection. Initial steps of viral infection include the attachment of the virus particle to a specific receptor located on the cell surface, in some cases a specific

lipid. For multiplication, viruses use various cellular factors including lipids involved in their metabolism. More proportion of genes belonged to generation of energy process in healthy plant as compared to virus infection. Further upregulation of genes involved in energy metabolism was observed during virus infection. These results are in accordance with the fact that the energy produced by host is being used by viruses for polymerization of their proteins and n-RNA synthesis as nucleoside triphosphate. As per GO analysis, more number of genes related to transport mechanism were in healthy plant as compared to virus infected. It supports the assumption that during local movement, virus initially moves from one infected cell to adjoining cells and move through vascular tissue to cause a systemic infection in the plant [76]. On the basis of subcellular location, it was observed that percent of genes associated to membrane part was higher in healthy condition as compared to virus-infected condition. These observations are in line with the fact that various viral encoded proteins are involved in membrane targeting of the replication components during replication of viruses [77].

Further, the overall observations of GO analysis are consistent with the supposition that biotic stresses in plants mark a change from growth and reproduction to physiological and metabolic processes designed for defense-related responses [78].

Upregulation and Downregulation of Transcription Factors in Response to Viral Infection

Transcriptional regulation is also an integral component toward overall understanding of disease response in plants. Immunity regulation and response to stresses is generally driven by transcription factors. Our analysis indicated that along with genes, these regulatory proteins are also activated in response to virus infection. There are many reports describing the activation of various transcription factors in response to virus infection in plants [10, 12]. We observed upregulation and downregulation of transcription factors in plant hormone signal transduction and plant-pathogen interaction, respectively, during virus infection. Upregulation of transcription factors involved in signaling of phytohormones during virus infection was observed. TF families, i.e., MYC2, TGA, ABA-responsive element binding factor, and ethylene-responsive transcription factor, were associated toward hormone signaling in response to virus infection. In plant-pathogen interaction, transcription factors related to WRKY family were downregulated during infection as they are involved in defense response mechanism [71].

Coexpression Network Analysis of Genes

GCNs are graphic illustrations that represent the synchronized transcription of genes in response to a particular stimulus. Gene coexpression analysis revealed the presence of various genes which are coexpressed with the target or reference genes. Few genes showed close association or interaction with the identified reference gene. These genes can be primarily targeted with identified disease inducing or resistance genes for understanding the molecular responses to develop resistant genotypes in Jatropha. For example, we observed that most of the coexpressed genes identified with defense response pathways are involved in secretory pathways which are responsible for providing immune response to plants [79]. It was inferred that the genes which are coexpressed with the identified reference genes indicated that there is some degree of conservation of their function. Similar approach of constructing gene coexpression networks in response to disease resistance and immune responses have been applied for immunity expression data of *Arabidopsis*, rice, soybean, tomato, and cassava which shed light on global patterns of events activated throughout plant immune responses [80]. Coexpression networks have also been developed for resistance genes identified from transcriptome data in a number of plant species [81, 82].

RT-qPCR-Based Experimental Validation

For experimental validation, selected genes on the basis of more transcript abundance in virus infection from transcriptome data were examined using RT-qPCR approach. The genes from plant hormone signal transduction were chosen for experimental validation as most number of identified genes were associated to this process. Also, the foremost symptoms during mosaic virus infection are linked with changes in metabolism of major plant hormones. Also, the transcription factors upregulated and downregulated during virus infection were validated. RT-qPCR-based experimental validation showed conjunction with the computational approach based on FPKM expression. However, slight variations in expression levels between computational approach and RT-qPCR might be due to the bioinformatics process executed in the transcriptome analysis, which includes many factors that can affect the reproducibility of measureable expression profiles, comprising alignment choices, transcript expression estimation, etc. [83].

Conclusions

Overall, the transcriptome-based characterization and comparative analysis of healthy versus virus-infected leaves has shed new dimensions in understanding the molecular perspective of plant-pathogen interaction in Jatropha. The current study revealed that mosaic virus symptom expression in *J. curcas* is a complex process which can be tackled on the basis of changes in gene expression. To our knowledge, this is the first comprehensive transcriptome-based analysis with respect to elucidation of molecular mechanisms associated with disease response in this bioenergy crop, *J. curcas*. Comparative transcriptomic investigation between symptomatic infected and healthy leaves

revealed major differences in gene expression levels. The present study led to the identification of genes and corresponding biological processes being upregulated and downregulated in response to mosaic virus infection in *J. curcas*. Further investigations of the comprehensive mechanisms of plant responses to viral infection will be expedited.

Acknowledgments The authors are thankful to Mr. Ankush Bansal (Jaypee University of Information Technology) for assisting in the coexpression network analysis and Mr. Tarun Pal (Jaypee University of Information Technology) for assisting in maintaining transcriptome data in the in-house web server. Financial support from the Department of Biotechnology (DBT), Govt. of India, to RSC in the form of a R&D project on *Jatropha curcas* is also acknowledged. The authors are also thankful to Dr. Sandeep Sharma, Scientist, Himalayan Forest Research Institute (HFRI), Shimla, India, for providing experimental farm facilities.

Availability of Transcriptome Data The raw reads and annotated data, i.e., CDSs for JH and JV are available at http://14.139.240.55/NGS/download.php.

Author Contribution AS and RSC defined the research theme and designed experiments. AS performed experiments and in silico analysis. AS and RSC wrote the manuscript and discussed analysis. Both authors read and approved the final manuscript.

Compliance with Ethical Standards

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

References

1. Ricci M, Bellaby P, Flynn R (2007) Stakeholders' and publics' perceptions of hydrogen energy technologies. In: Flynn R, Bellaby P (eds) Risk and the public acceptance of new technologies. Palgrave Macmillan, New York, pp. 175–197
2. Wulff BB, Horvath DM, Ward ER (2011) Improving immunity in crops: new tactics in an old game. Curr Opin Plant Biol 14(4):468–476
3. Tewari JP, Dwivedi HD, Pathak M, Srivastava SK (2007) Incidence of a mosaic disease in *Jatropha curcas* L. from eastern Uttar Pradesh. Curr Sci 93:1048–1049
4. Aswatha Narayana DS, Rangaswamy KS, Shankarappa MN, Maruthi MN, Reddy CNL, Rekha AR, Murthy KVK (2007) Distinct Begmoviruses closely related to cassava mosaic viruses causes Indian Jatropha mosaic disease. Int J Virol 3:1–11
5. Raj SK, Kumar S, Snehi SK, Pathre U (2008) First report of cucumber mosaic virus on *Jatropha curcas* L. In India. Plant Dis 92:171
6. Ramkat RC, Calari A, Maghuly F, Laimer M (2011) Biotechnological approaches to determine the impact of viruses in the energy crop plant *Jatropha curcas*. Virol J 8:386
7. Kashina BD, Alegbejo MD, Banwo OO, Nielsen SL, Nicolaisen M (2013) Molecular identification of a new begomovirus associated with mosaic disease of *Jatropha curcas* L. in Nigeria. Arch Virol 158:511–514
8. Góngora-Castillo E, Ibarra-Laclette E, Trejo-Saavedra DL, Rivera-Bustamante RF (2012) Transcriptome analysis of symptomatic and recovered leaves of geminivirus-infected pepper (*Capsicum annuum*). Virol J 9:295
9. Allie F, Pierce EJ, Okoniewski MJ, Rey MEC (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
10. Lu J, Du ZX, Kong J, Chen LN, Qiu YH, Li GF, Meng XH, Zhu SF (2012) Transcriptome analysis of *Nicotiana tabacum* infected by *Cucumber mosaic virus* during systemic symptom development. PLoS One 7:e43447
11. Liu J, Yang J, Bi H, Zhang P (2014) Why mosaic? Gene expression profiling of *African cassava mosaic virus*-infected cassava reveals the effect of chlorophyll degradation on symptom development. J Integr Plant Biol 56(2):122–132
12. Choi H, Jo Y, Lian S, Jo KM, Chu H, et al. (2015) Comparative analysis of chrysanthemum transcriptome in response to three RNA viruses: cucumber mosaic virus, tomato spotted wilt virus and potato virus X. Plant Mol Biol 88(3):233–248
13. Cho WK, Lian S, Kim SM, Seo BY, Jung JK, Kim KH (2015) Time-course RNA-Seq analysis reveals transcriptional changes in rice plants triggered by rice stripe virus infection. PLoS One 10(8):e0136736
14. Postnikova OA, Hult M, Shao J, Skantar A, Nemchinov LG (2015) Transcriptome analysis of resistant and susceptible alfalfa cultivars infected with root-knot nematode *Meloidogyne incognita*. PLoS One 10(2):e0118269
15. Fan H, Zhang Y, Sun H, Liu J, Wang Y, Wang X, 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(7):e0132277
16. Ishihara T, Mitsuhara I, Takahashi H, Nakaho K (2012) Transcriptome analysis of quantitative resistance-specific response upon *Ralstonia solanacearum* infection in tomato. PLoS One 7(10):e46763
17. Liu JJ, Sturrock RN, Benton R (2013) Transcriptome analysis of *Pinus monticola* primary needles by RNA-seq provides novel insight into host resistance to *Cronartium ribicola*. BMC Genomics 14:884
18. Costa GGL, Cardoso KC, Del Bem LEV, Lima AC, Cunha MAS, de Campos-Leite L, Vicentini R, Papes F, Moreira RC, Yunes JA, Campos FAP, Da Silva MJ (2010) Transcriptome analysis of the oil-rich seed of the bioenergy crop *Jatropha curcas* L. BMC Genomics 11:462
19. King AJ, Li Y, Graham IA (2011) Profiling the developing *Jatropha curcas* L. seed transcriptome by pyrosequencing. Bioenerg Res 4: 211–221
20. Wang H, Zou Z, Wang S, Gong M (2013) Global analysis of transcriptome responses and gene expression profiles to cold stress of *Jatropha curcas* L. PLoS One 8:e82817
21. Juntawong P, Sirikhachornkit A, Pimjan R, et al. (2014) Elucidation of the molecular responses to waterlogging in Jatropha roots by transcriptome profiling. Front Plant Sci 5:658
22. Pan BZ, Chen MS, Ni J, Xu ZF (2014) Transcriptome of the inflorescence meristems of the biofuel plant *Jatropha curcas* treated with cytokinin. BMC Genomics 15(1):974
23. Polston JE, Londoño MA, Capobianco H (2014) The complete genome sequence of new world Jatropha mosaic virus. Arch Virol 159(11):3131–3136
24. Sato S, Hirakawa H, Isobe S, et al. (2011) Sequence analysis of the genome of an oil-bearing tree, *Jatropha curcas* L. DNA Res 18:65–76
25. Sood A, Jaiswal V, Chanumolu SK, Malhotra N, Pal T, Chauhan RS (2014) Mining whole genomes and transcriptomes of Jatropha (*Jatropha curcas*) and Castor bean (*Ricinus communis*) for NBS-LRR genes and defense response associated transcription factors. Mol Biol Rep 41(11):7683–7695
26. Conesa A, Gotz S, Garcia-Gomez JM, Terol J, Talon M, Robles M (2005) Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research. Bioinformatics 21(18): 3674–3676
27. Stuart J, Segal E, Koller D, Kim S (2003) A gene co-expression network for global discovery of conserved genetic modules. Science 302(5643):249–255

28. Gao S, Qu J, Chua NH, Ye J (2010) A new strain of Indian cassava mosaic virus causes a mosaic disease in the biodiesel crop *Jatropha curcas* L. Arch Virol 155:607–612
29. Jayanna K (2006) Studies on Jatropha mosaic virus disease. M.Sc. Dissertation, University of Agricultural Sciences, Dharwad
30. Satoh K, Kondoh H, Sasaya T, Shimizu T, Choi IR, Omura T, Kikuchi S (2010) Selective modification of rice (*Oryza sativa*) gene expression by rice stripe virus infection. J Gen Virol 91(1):294–305
31. Mardi M, Farsad LK, Gharechahi J, Salekdeh GH (2015) In-depth transcriptome sequencing of Mexican lime trees infected with candidatus *Phytoplasma aurantifolia*. PLoS One 10(7):e0130425
32. Cassone BJ, Wijeratne S, Michel AP, et al. (2014) Virus-independent and common transcriptome responses of leafhopper vectors feeding on maize infected with semi-persistently and persistent propagatively transmitted viruses. BMC Genomics 15(1):133
33. Nielsen E, Cheung AY, Ueda T (2008) The regulatory RAB and ARF GTPases for vesicular trafficking. Plant Physiol 147:1516–1526
34. Haupt S, Cowan GH, Ziegler A, Roberts AG, Oparka KJ, Torrance L (2005) Two plant-viral movement proteins traffic in the endocytic recycling pathway. Plant Cell 17:164–181
35. Sen A, Madhivanan K, Mukherjee D, et al. (2012) The epsin protein family: coordinators of endocytosis and signaling. Biomol Concepts 3: 117–126
36. Signorelli S, Coitiño EL, Borsani O, Monza J (2014) Molecular mechanisms for the reaction between ·OH radicals and proline: insights on the role as reactive oxygen species scavenger in plant stress. J Phys Chem 118:37–47
37. Guo FQ, Crawford NM (2005) *Arabidopsis* nitric oxide synthase1 is targeted to mitochondria and protects against oxidative damage and dark-induced senescence. Plant Cell 17:3436–3450
38. Vlad F, Spano T, Vlad D, Daher FB, Ouelhadj A, Kalatizis P (2007) *Arabidopsis* prolyl 4- hydroxylase are differentially expressed in response to hypoxia, anoxia and mechanical wounding. Physiol Plant 130:471–483
39. Agudelo-Romero P, Carbonell P, de la Iglesia F, Carrera J, Rodrigo G, Jaramillo A, Perez-Amador MA, Elena SF (2008) Changes in the gene expression profile of *Arabidopsis thaliana* after infection with *Tobacco etch virus*. Virol J 5:92
40. Dubey NK, Goel R, Ranjan A, Idris A, Singh SK, Bag SK, Chandrashekar K, Pandey KD, Singh PK, Sawant SV (2013) Comparative transcriptome analysis of *Gossypium hirsutum* L. In response to sap sucking insects: aphid and whitefly. BMC Genomics 14:241
41. Pavet V, Olmos E, Kiddle G, Mowla S, Kumar S, Antoniw J, Alvarez ME, Foyer CH (2005) Ascorbic acid deficiency activates cell death and disease resistance responses in *Arabidopsis*. Plant Physiol 139: 1291–1303
42. Laxalt AM, Munnik T (2002) Phospholipid signalling in plant defence. Curr Opin Plant Biol 5:332–338
43. Thomma BP, Penninckx IA, Broekaert WF, Cammue BP (2001) The complexity of disease signaling in *Arabidopsis*. Curr Opin Immunol 13:63–68
44. Freitas-Astúa J, Bastianel M, Locali-Fabris EC, et al. (2007) Differentially expressed stress-related genes in the compatible citrus–citrus leprosis virus interaction. Genet Mol Biol 30:980–1018
45. Breuer G, de Jaeger L, Artus VP, Martens DE, Springer J, Draaisma RB, Eggink G, Wijffels RH, Lamers PP (2014) Superior triacylglycerol (TAG) accumulation in starchless mutants of *Scenedesmus obliquus*: (II) evaluation of TAG yield and productivity in controlled photobioreactors. Biotech Biofuels 7:1–11
46. Fan J, Yan C, Andre C, Shanklin J, Schwender J, Xu C (2012) Oil accumulation is controlled by carbon precursor supply for fatty acid synthesis in *Chlamydomonas reinhardtii*. Plant Cell Physiol 53: 1380–1390
47. Molhoj M, Verma R, Reiter WD (2004) The biosynthesis of d-galacturonate in plants. Functional cloning and characterization of a membrane-anchored UDP-d-glucuronate 4-epimerase from Arabidopsis. Plant Physiol 135:1221–1230
48. Burget EG, Verma R, Mølhøj M, Reiter WD (2003) The biosynthesis of l-arabinose in plants: molecular cloning and characterization of a Golgi-localized UDP-d-xylose 4-epimerase encoded by the *MUR4*-gene of *Arabidopsis*. Plant Cell 15:523–531
49. Li R, Tee CS, Jiang YL, Jiang XY, Venkatesh PN, Sarojam R, Ye J (2015) A terpenoid phytoalexin plays a role in basal defense of *Nicotiana benthamiana* against potato virus X. Sci Rep 5:9682
50. Pallas V, Garcia JA (2011) How do plant viruses induce disease? Interactions and interference with host components. J Gen Virol 92: 2691–2705
51. Maruthi MN, Bouvaine S, Tufan HA, Mohammed IU, Hillocks RJ (2014) Transcriptional response of virus-infected cassava and identification of putative sources of resistance for cassava brown streak disease. PLoS One 9(5):e96642
52. Whenham RJ, Fraser RSS (1980) Stimulation by abscisic acid of RNA synthesis in discs from healthy and tobacco mosaic virus-infected tobacco leaves. Planta 150:349–353
53. Pennazio S, Roggero P (1996) Plant hormones and plant virus diseases. The auxins Microbiologica 19:369–378
54. Alazem M, Lin K-Y, Lin N-S (2014) The abscisic acid pathway has multifaceted effects on the accumulation of *Bamboo mosaic virus*. Mol Plant-Microbe Interact 27:177–189
55. Casteel C, De Alwis M, Bak A, Dong H, Steven A, Jander G (2015) Disruption of ethylene responses by turnip mosaic virus mediates suppression of plant defense against the aphid vector, *Myzus persicae*. Plant Physiol 169:209–218
56. Qi T, Song S, Ren Q, Wu D, Huang H, Chen Y, Fan M, Peng W, Ren C, Xie D (2011) The jasmonate-ZIM-domain proteins interact with the WD-repeat/bHLH/MYB complexes to regulate jasmonate-mediated anthocyanin accumulation and trichome initiation in *Arabidopsis thaliana*. Plant Cell 23:1798–1814
57. Tokutsu R, Teramoto H, Takahashi Y, Ono T, Minagawa J (2003) The light-harvesting complex of photosystem I in *Chlamydomonas reinhardtii*: protein composition, gene structures and phylogenic implications. Plant Cell Physiol 45:138–145
58. Fukuyama K (2004) Structure and function of plant-type ferredoxins. Photosynth Res 81:289–301
59. Johnson X, Alric J (2013) Central carbon metabolism and electron transport in *Chlamydomonas reinhardtii*: metabolic constraints for carbon partitioning between oil and starch. Eukaryot Cell 12(6): 776–793
60. Wang HZ, Zhao PJ, Xu JC, et al. (2003) Virus resistance in transgenic watermelon plants containing a WMV-2 coat protein gene. Acta Genet Sin 30:70–75
61. Weng K, Li ZQ, Liu RQ, Wang L, Wang YJ, Xu Y (2014) Transcriptome of Erysiphe necator-infected *Vitis pseudoreticulata* leaves provides insight into grapevine resistance to powdery mildew. Horti Res 1:14049
62. Lev-Yadun S, Gould KS (2009) Role of anthocyanins in plant defense. In: Gould KS, Davies KM, Winefield C (eds) Life's colorful solutions: the biosynthesis, functions, and applications of anthocyanins. Springer, Berlin, pp. 21–48
63. Kaplan B, Sherman T, Fromm H (2007) Cyclic nucleotide-gated channels in plants. FEBS Lett 581(12):2237–2246
64. Moeder W, Urquhart W, Ung H, Yoshioka K (2011) The role of cyclic nucleotide gated ion channels in plant immunity. Mol Plant 4:442–452
65. Chin K, DeFalco TA, Moeder W, Yoshioka K (2013) The *Arabidopsis* cyclic nucleotide-gated ion channels AtCNGC2 and AtCNGC4 work in the same signaling pathway to regulate pathogen defense and floral transition. Plant Physiol 163:611–624
66. Reddy ASN, Ali GS, Celesnik H, Day IS (2011) Coping with stresses: roles of calcium- and calcium/calmodulin-regulated gene expression. Plant Cell 23:2010–2032

67. Cheval C, Aldon D, Galaud JP, Ranty B (2013) Calcium/calmodulin-mediated regulation of plant immunity. Biochim Biophys Acta 1833: 1766–1771
68. Pedley KF, Martin GB (2005) Role of mitogen-activated protein kinases in plant immunity. Curr Opin Plant Biol 8:541–547
69. Tena G, Boudsocq M, Sheen J (2011) Protein kinase signaling networks in plant innate immunity. Curr Opin Plant Biol 14:519–529
70. Singh P, Zimmerli L (2013) Lectin receptor kinases in plant innate immunity. Front Plant Sci 4:1–4
71. Pandey SP, Somssich IE (2009) The role of WRKY transcription factors in plant immunity. Plant Physiol 150:1648–1655
72. Eulgem T (2005) Regulation of the Arabidopsis defense transcriptome. Trends Plant Sci 10:71–78
73. Zhang L, Du L, Poovaiah BW (2014) Calcium signaling and biotic defense responses in plants. Plant Signal Behav 9(11):e973818
74. Resh MD (2006) Palmitoylation of ligands, receptors, and intracellular signaling molecules. Sci STKE 2006:re14
75. Wang RY, Li K (2012) Host factors in the replication of positive-strand RNA viruses. Chang Gung Med J 35:111–124
76. Agbeci M, Grangeon R, Nelson RS, et al. (2013) Contribution of host intracellular transport machineries to intercellular movement of *Turnip mosaic virus*. PLoS Pathog 9:e1003683
77. Jakubiec A, Notaise J, Tournier V, Hericourt F, Block MA, Drugeon G, van Aelst L, Jupin I (2004) Assembly of turnip yellow mosaic virus replication complexes: interaction between the proteinase and polymerase domains of the replication proteins. J Virol 78:7945–7957
78. Bilgin D, Zavala J, Zhu J, Clough S, Ort D, DeLucia E (2010) Biotic stress globally downregulates photosynthesis genes. Plant Cell Environ 33(10):1597–1613
79. Kwon C, Bednarek P, Schulze-Lefert P (2008) Secretory pathways in plant immune responses. Plant Physiol 147:1575–1583
80. Leal LG, L'opez C, L'opez-Kleine L (2014) Construction and comparison of gene co-expression networks shows complex plant immune responses. Peer J 2:e610
81. Zheng ZL, Zhao Y (2013) Transcriptome comparison and gene co-expression network analysis provide a systems view of citrus response to 'Candidatus Liberibacter asiaticus' infection. BMC Genomics 14:27
82. Tully JP, Hill AE, Ahmed HM, Whitley R, Skjellum A, Mukhtar MS (2014) Expression-based network biology identifies immune-related functional modules involved in plant defense. BMC Genomics 15:421
83. Łabaj P, Leparc G, Linggi B, Markillie L, Wiley H, Kreil D (2011) Characterization and improvement of RNA-seq precision in quantitative transcript expression profiling. Bioinformatics 27(13):i383–i391