Transcriptome analysis of *Nicotiana benthamiana* infected by *Tobacco curly shoot virus*

Ke Li, Gentu Wu, Mingjun Li, Mingge Ma, Jiang Du, Miao Sun, Xianchao Sun and Ling Qing*

**Abstract**

**Background:** *Tobacco curly shoot virus* (TbCSV) is a monopartite begomovirus associated with betasatellite (*Tobacco curly shoot betasatellite*, TbCSB), which causes serious leaf curl disease on tomato and tobacco in China. It is interesting that TbCSV induced severe upward leaf curling in *Nicotiana benthamiana*, but in the presence of TbCSB, symptoms changed to be downward leaf curling. However, the mechanism of interactions between viral pathogenicity, host defense, viral-betasatellite interactions and virus-host interactions remains unclear.

**Methods:** In this study, RNA-seq was used to analyze differentially expressed genes (DEGs) in *N. benthamiana* plants infected by TbCSV (Y35A) and TbCSV together with TbCSB (Y35AB) respectively.

**Results:** Through mapping to *N. benthamiana* reference genome, 59,814 unigenes were identified. Transcriptome analysis revealed that a total of 4081 and 3196 DEGs were identified in Y35AB vs CK (control check) and Y35A vs CK, respectively. Both GO and KEGG analyses were conducted to classify the DEGs. Ten of the top 15 GO terms were enriched in both DEGs of Y35AB vs CK and Y35A vs CK, and these enriched GO terms mainly classified into three categories including biological process, cellular component and molecular function. KEGG pathway analysis indicated that 118 and 111 pathways were identified in Y35AB vs CK and Y35A vs CK, respectively, of which nine and six pathways were significantly enriched. Three major pathways in Y35AB vs CK involved in metabolic pathways, carbon metabolism and photosynthesis, while those in Y35A vs CK were related to Ribosome, Glyoxylate and dicarboxylate metabolism and DNA replication. We observed that 8 PR genes were significantly up-regulated and 44 LRR-RLK genes were significantly differentially expressed in Y35A treatment or in Y35AB treatment. In addition, 7 and 13 genes were identified to be significantly changed in biosynthesis and signal transduction pathway of brassinosteroid (BR) and jasmonic acid (JA) respectively.

**Conclusions:** These results presented here would be particularly useful to further elucidate the response of the host plant against virus infection.

**Keywords:** Begomovirus, *Tobacco curly shoot virus*, Betasatellite, *Nicotiana benthamiana*, RNA-seq

---

**Background**

*Tobacco curly shoot virus* (TbCSV), a monopartite begomovirus (genus *Begomovirus*, family *Geminiviridae*), was isolated from tobacco in Yunnan province, China in 2002 [1]. Some TbCSV isolates were identified associating with a betasatellite molecule (*Tobacco curly shoot betasatellite*, TbCSB) [2]. TbCSV is one of the most important pathogens causing leaf curl disease on tomato and tobacco, and is a severe constraint to crop yields. TbCSV consists of a circular single-stranded DNA genome that encodes 6 proteins, including AV1 (coat protein, CP) and AV2 in virion-sense strand whilst AC1 (replication-related, Rep), AC2 (transcriptional activator, TrAP), AC3 (replication enhancer, Ren) and AC4 in complementary-sense strand [1]. TbCSB is a small circular single-stranded DNA molecule which encodes a sole protein $\beta C1$ [2].

Viral infection is a complicated procedure involving in the interaction between viruses and host plants. Understanding host responses to viral infection is advantageous...
in the development of effective strategies for virus control. In recent years, the interactions between viruses and host plants were clarified through analysis of transcriptomics [3–5]. RNA-seq is a recently developed approach to transcriptome profiling via deep-sequencing technologies, which provides a far more precise measurement of levels of transcripts and their isoforms than other methods [6]. Thus, this technique was widely applied in interpretation of the interaction between virus infection and host plant. The multiple resistance mechanisms against cotton leaf curl disease (CLCuD) were revealed by the transcriptome analysis based on an RNA-seq in a naturally immune cotton species (Gossypium arboeum) caused by CLCuD [7]. The technology of transcriptomics and proteomics were employed to study the differential host gene expression during Mungbean yellow mosaic India virus (MYMIV) and Mungbean yellow mosaic virus (MYMV) infection under natural conditions [8]. The differential regulations of host genes revealed to be involved in cell cycle, cell-wall biogenesis, chloroplast, photosynthesis, hormone and sulphur assimilation pathways which may contribute to symptom development in soybean plants [8]. Recently, a research performed RNA-seq-based transcriptome sequencing of Jatropha curcas mosaic virus-infected and healthy leaf tissues of J. curcas, which provides a repertoire of molecular components after viral infection [9]. Similarly, several host genes were identified to be involved in different cellular and physiological processes during the infection of other viruses, including South African cassava mosaic virus (SACMV) [10], Tomato yellow leaf curl Sardinia virus (TYLCSV), Mungbean yellow mosaic India virus (MYMIV) [11, 12], and Chilli leaf curl virus (ChiLCV) [13].

At present, it has been reported that the infectious clone of TbCSV induced severe upward leaf curling in N. benthamiana, but in the presence of TbCSB the symptom changed to a downward leaf curl [14]. Thus, TbCSV may represent an evolutionary intermediate between the truly monopartite begomoviruses and those that require TbCSB. Moreover, studies showed that AC2 and AC4 proteins of TbCSV mediate suppression of RNA silencing [15], and βC1 protein of TbCSB could bind single- and double-stranded DNA to suppress host RNA silencing activities [16]. However, it remains unclear that the molecular mechanism of viral pathogenicity, host defense and viral-host interactions. In this study, in order to get insights into the molecular mechanisms in response to TbCSV infection in N. benthamiana plants, the transcriptome analysis of N. benthamiana infected by TbCSV or TbCSV/TbCSB was performed, including differentially expressed genes, GO enrichment, KEGG enrichment, pathogenesis-related (PR) protein genes, LRR-RLK genes, brassinosteroid and jasmonic acid pathways. These data contribute to a better knowledge on the molecular mechanisms of TbCSV- and TbCSV/TbCSB- host interaction, and serve as a basis for devising new strategies to control TbCSV/TbCSB disease complex.

Methods
Virus sources and agroinoculation
The infectious clones of TbCSV isolate Y35 (pBinY35A-1.9) (Y35A) and its betasatellite (pBinY35β-2.0) (Y35B) were provided by Professor Xueping Zhou in the Biotechnology Institute of Zhejiang University. Both viral infectious clones were introduced into Agrobacterium tumefaciens strain EHA105. N. benthamiana (accession: Nb-1) plants were grown in a greenhouse under a 16 h light and 8 h dark cycle at 26 °C. The Agrobacterium with infectious clones was infiltrated into N. benthamiana leaves at 4–6 leaf stage as previously described [2].

DNA extraction and PCR detection
Leaf samples were collected at 20 days-post inoculation (dpi) and total DNA were extracted using CTAB method [17]. The TbCSV-specific primers (TbCSV-F, 5′-ATGC CTCAGCCAAGAAAACTTTT-3′; TbCSV-R, 5′-TCAA CACGACGACGTCTGTCC-3′) and TbCSB-specific primers (TbCSB-F, 5′-ATGACAATTAATACAACA ACAAG-3′; TbCSB-R, 5′-TCAATACATGGTATTGTCC-3′) were designed to amplify fragments of 1086 bp and 357 bp in size, respectively. PCR reactions were performed in a 20 µL volume with reaction mixtures containing 10 µL of 2× Taq Master Mix (Novoprotein Scientific Inc., Shanghai, China), 0.5 µL of DNA, 0.25 µM of each primer and a proper volume of ddH2O. The PCR reaction conditions were as follows: 94 °C for 3 min, followed by 35 cycles of denaturation for 30 s at 94 °C, annealing for 30 s at 52 °C, extension for 30–65 s at 72 °C (depending on primer pairs used in distinct reactions), and a final extension for 10 min at 72 °C.

Materials for RNA sequencing
Leaf tissues of N. benthamiana plants were sampled at 20 dpi, and were immediately frozen and ground in liquid nitrogen and stored at −80 °C. After PCR detection, RNA-seq was performed for stored samples. Three treatments were setting as follows: plants infected with TbCSV alone (Y35A treatment); plants infected with TbCSV and TbCSB (Y35AB treatment); uninfected plants as control (CK). There are three biological replicates per treatment respectively.

RNA extraction and cDNA synthesis
Total RNA was extracted from N. benthamiana leaf tissues with RNAiso Plus (TAKARA Bio, Inc) as manufacturer’s protocol. All RNA samples were first treated with gDNA Eraser (TAKARA Bio, Inc) and reverse-transcribed using a
Prime Script RT reagent Kit (Perfect Real Time) according to manufacturer’s instructions.

**Library preparation for transcriptome sequencing**

A total amount of 3 μg RNA per sample was used as input for RNA sample preparations. Sequencing libraries were generated using NEBNext Ultra RNA Library Prep Kit for Illumina (NEB, USA) following manufacturer’s recommendations and index codes were added to attribute sequences to each sample. Briefly, mRNA was purified by using poly-T oligo-attached magnetic beads. After fragmentation, the cDNA was synthesized, and NEBNext Adaptors with hairpin loop structures were ligated to prepare for hybridization. PCR products were purified (AMPure XP system), and library quality was assessed by Agilent Bioanalyzer 2100 system [18]. Finally, nine libraries were successfully constructed, and were sequenced using HiSeq™ 2500 equipment (Illumina, SanDiego, CA, USA) by 50 bp single-end (SE50) methods.

**Quality control**

For sequence quality control, raw data (raw reads) of fastq format were firstly processed through in-house perl scripts. In this step, clean data (clean reads) were obtained by removing reads containing adapter, reads containing poly-N and low quality reads from raw data. Meanwhile, Q20, Q30, GC-content and sequence duplication level of clean data were calculated. All the downstream analyses were based on the clean data with high quality [19].

**Reads mapping to the reference genome**

The draft sequence of the *N. benthamiana* reference genome have been downloaded from the SGN ftp site (ftp://ftp.solgenomics.net/genomes/Nicotiana_benthamiana/assemblies) directly [20]. Index of the reference genome was built using Bowtie v2.2.3 (Broad Institute, Cambridge, MA, USA) [21] and single-end clean reads were aligned to the reference genome using TopHat v2.0.12 (Broad Institute, Cambridge, MA, USA) (mismatch = 2) [22]. The TopHat was selected as the mapping tool for that TopHat can generate a database of splice junctions based on the gene model annotation file and thus a better mapping result than other respective assemblies directly [20]. Index of the reference genome was constructed using Bowtie v2.2.3 (Broad Institute, Cambridge, MA, USA) [21] and single-end clean reads were aligned to the reference genome using TopHat v2.0.12 (Broad Institute, Cambridge, MA, USA) (mismatch = 2) [22]. The TopHat was selected as the mapping tool for that TopHat can generate a database of splice junctions based on the gene model annotation file and thus a better mapping result than other non-splice mapping tools [23].

**Quantification of gene expression level**

HTSeq v0.6.1 was used to count reads numbers mapped to each gene [24]. And then FPKM (Fragments per kilobase of transcript sequence per millions base pairs sequenced) of each gene was calculated based on the length of the gene and reads count mapped to this gene. FPKM, expected number of Fragments Per Kilobase of transcript sequence per Millions base pairs sequenced, considers the effect of sequencing depth and gene length for the reads count at the same time, and is currently the most commonly used method for estimating gene expression levels [25].

**Differential gene expression analysis**

Differential expression analysis of Y35AB vs CK and Y35A vs CK was performed using the DESeq R package (1.18.0) (parameters: negative binomial distribution-based statistic, BH-FDR corrected *p*-value < 0.05) [26], which provides statistical routines for determining differential expression in digital gene expression data using a model based on the negative binomial distribution. The resulting *p*-values were adjusted using the Benjamini and Hochberg’s approach for controlling the false discovery rate (FDR).

**GO term and KEGG pathway enrichment analysis**

Gene Ontology (GO; [http://www.geneontology.org](http://www.geneontology.org)) enrichment analysis of differentially expressed genes (DEGs) was implemented by the GOseq R package (corrected *p*-value < 0.05) [27], in which gene length bias was corrected. GO terms with corrected *p*-value less than 0.05 were considered significantly enriched by differential expressed genes. REVIGO was used for analysis of the enriched GO terms ([http://revigo.irb.hr](http://revigo.irb.hr)). KEGG ([http://www.genome.jp/kegg](http://www.genome.jp/kegg)) is a database resource for understanding high-level functions and utilities of the biological system from molecular-level information, especially large-scale molecular datasets generated by genome sequencing and other high-throughput experimental technologies. And KOBAS software was used to test the statistical enrichment of differential expression genes in KEGG pathways (corrected *p*-value < 0.05) [28].

**Quantitative real-time PCR validation**

To validate the RNA-seq data, quantitative real-time PCR (RT-qPCR) was conducted to examine the selected pattern of DEGs. For reverse transcription, 1 μg of total RNA was treated with gDNA Eraser (TAKARA Bio, Inc) and reverse-transcribed using a Prime Script RT reagent Kit (Perfect Real Time) and oligo (dT) as the primer according to the manufacturer’s protocol. Primer sets were designed using AlleleID software (v6.0) (Additional file 1: Table S1). RT-qPCR was conducted by using NovoStart™ SYBR qPCR Super Mix Plus (Novoprotein, Shanghai, China) on CFX 96 Real-Time System (Bio-Rad). Actin was used as a reference for calculating relative abundances using the 2^△△Ct method [29]. All RT-qPCR experiments were performed in triplicate.

**Results**

**Plants with typical viral symptoms in Y35A and Y35AB treatments**

At 20 dpi, Y35A inoculated plants showed slight upward leaf curling and Y35AB inoculated plants showed severe downward leaf curl (Fig. 1a-c). PCR detection showed that they are consistent with expectations (Fig. 1d-e).
Symptom observation and virus detection results showed that these plants were reliable for follow-up RNA-seq. Leaf tissues at same positions were collected from every treatment for RNA-seq.

**Overview of transcriptome sequencing**

To profile gene expression during virus infection, RNA-seq libraries were constructed for the control- and virus-inoculated *N. benthamiana* plants. 5.76 GB clean bases data and 59,814 unigenes were obtained from nine samples. There were 118,344,960–13,118,731, 12,537,134–15,503,750 and 11,290,015–12,633,203 clean reads in the three treatments of Y35AB, Y35A, and CK. The proportion of total reads that mapped to *N. benthamiana* genome of each sample was more than 96.09%, and that of uniquely mapped was above 72.69%. More than 95.58% of the clean reads had quality scores at the Q30 (an error probability for base calling of 0.1%) level. Furthermore, the GC content ranges from 43.58% to 44.38%. The sequencing data are summarized in Table 1.

![Fig. 1 Comparison of *N. benthamiana* plant symptoms of virus-infected with healthy control at 20 dpi.](image)

**Table 1** Summary of the results of RNA-seq data

| Sample name | Y35AB-1 | Y35AB-2 | Y35AB-3 | Y35A-1 | Y35A-2 | Y35A-3 | CK-1 | CK-2 | CK-3 | mean |
|-------------|---------|---------|---------|--------|--------|--------|------|------|------|------|
| Raw reads   | 121,664,53 | 119,654,70 | 132,728,80 | 155,273,72 | 125,602,10 | 142,147,40 | 126,460,00 | 113,027,05 | 117,940,17 | 128,277,61 |
| Clean reads | 120,327,20 | 118,344,96 | 131,187,31 | 155,037,50 | 125,371,34 | 141,921,35 | 126,332,03 | 112,900,15 | 117,825,62 | 127,694,16 |
| Clean bases | 0.60G   | 0.59G   | 0.66G   | 0.78G   | 0.63G   | 0.71G   | 0.63G   | 0.56G   | 0.59G   | 0.64   |
| Total mapped| 116,207,70 | 114,386,90 | 126,052,97 | 149,344,55 | 120,757,73 | 137,323,42 | 123,461,63 | 111,395,53 | 115,242,48 | 123,796,99 |
| 96.58% | 96.66% | 96.09% | 96.33% | 96.32% | 96.76% | 97.73% | 98.67% | 98.71% | 96.99% |
| Uniquely mapped | 896,132 | 884,753 | 960,976 | 113,897,20 | 91,765,12 | 103,208,07 | 92,386,44 | 83,160,46 | 85,646,17 | 93,805,53 |
| 74.47% | 74.76% | 73.25% | 73.46% | 73.19% | 72.72% | 73.13% | 73.66% | 72.69% | 73.48% |
| Q30 (%) | 95.74 | 95.68 | 95.58 | 96.20 | 96.08 | 96.14 | 95.76 | 95.70 | 95.80 | 95.85 |
| GC contents (%) | 43.58 | 43.75 | 44.11 | 44.15 | 44.23 | 44.38 | 43.93 | 43.80 | 43.92 | 43.98 |

Note: Clean reads: Reads from sequencing after filtering low-quality reads. Clean bases: The number of clean reads is multiplied by the length and converted to G. Q30: The percentage of bases with a Phred value >30.
Analysis of DEGs
An adjusted \( p \)-value < 0.05 by DESeq was used to identify DEGs. In the Y35AB vs CK, 4081 DEGs (Fig. 2a, Additional file 2: Table S2), including 1775 up-regulated and 2306 down-regulated genes, were identified (Fig. 2c). And 3196 transcripts (1391 up-regulated and 1805 down-regulated genes) showed significant changes in Y35A vs CK (Fig. 2b, c and Additional file 3: Table S3). However, a total of 1232 DEGs in Y35AB vs CK overlap with those in Y35A vs CK, including 445 up-regulated and 786 down-regulated genes (Fig. 2c).

GO enrichment analysis
Gene Ontology, an internationally standardized gene function classification system, was used to classify the DEGs. The results were enriched in Y35AB vs CK and Y35A vs CK, and 124 and 76 of them were significantly enriched, respectively. Further, the top 15 GO functional annotation terms were listed, and 10 of them were shared between Y35AB vs CK and Y35A vs CK, associated with biological process, metabolic process, cellular process, organic substance metabolic process, primary metabolic process, cellular metabolic process, nitrogen compound metabolic process, biosynthetic process, cellular nitrogen compound metabolic process and organic substance biosynthetic process (Fig. 3, Additional files 4 and 5: Tables S4 and S5). Furthermore, it was found that these enriched GO terms mainly classified into biological process.

KEGG pathway enrichment analysis
In order to further clarify molecular and biological functions of the genes, these DEGs were mapped to the KEGG database. KEGG pathway analysis indicated that 118 and 111 pathways were identified in Y35AB vs CK and Y35A vs CK, respectively, of which nine and six were significantly enriched with \( p \)-value < 0.05. The significantly enrichment analysis showed that DEGs of Y35AB vs CK involved in metabolic pathways, carbon metabolism, photosynthesis, carbon fixation in photosynthetic organisms, glyoxylate and dicarboxylate metabolism, porphyrin and chlorophyll metabolism, DNA replication, pentose phosphate pathway and nitrogen metabolism. While, DEGs involved in ribosome, glyoxylate and dicarboxylate metabolism, DNA replication, circadian rhythm-plant, photosynthesis-antenna proteins and nitrogen metabolism were significantly enriched in Y35A vs CK (Fig. 4, Additional files 6 and 7: Tables S6 and S7).

Metabolic pathway was the major pathway which contains the largest number of DEGs in both Y35AB vs CK and Y35A vs CK, but DEGs involved in metabolic pathway were not significantly different in Y35A vs CK, it further indicated that the presence of TbCSB differentially changed metabolic pathways in N. benthamiana. The rich factor of porphyrin and chlorophyll metabolism pathway in Y35AB vs CK was near 1.00, the results showed that almost all genes in this pathway differentially expressed in TbCSV associated with betasatellite infected plants. Nevertheless, ribosome pathway might play an important role in response to TbCSV infection for the number of enriched DEGs in ribosome pathway was near to that enriched in metabolic pathways, and the rich factor was more than 0.75.

Screening of Y35AB vs CK and Y35A vs CK common highly expressed DEGs
FPKM reflecting both the effect of sequencing depth and gene length for the read count, and is a commonly used method for estimating gene expression levels [31]. The fold change in gene expression was used to distinguish differentially expressed genes between samples because of the significance of digital gene expression profiles [32]. The intersection of highly expressed DGEs in Y35AB vs CK and Y35A vs CK was considered. Finally, 17 DEGs were identified (|log2(Fold change)| ≥ 1 and FPKM≥50) from significantly enriched KEGG pathways, these 17 genes shared between Y35AB vs CK and Y35A vs CK all down-regulated, and involved in six pathways, including
Fig. 3 Most of the top 15 GO terms of three categories significantly enriched in DGEs of Y35AB vs CK and Y35A vs CK. MF, molecular function; BP, biological process; CC, cellular component

Fig. 4 KEGG pathways of the more or main significantly enriched in DGEs. The rich factor reflects the degree of enriched DGEs in a given pathway. The number of enriched DGEs in the pathway is indicated by the circle area, and the circle color represents the ranges of the corrected p-value.
photosynthesis, porphyrin and chlorophyll metabolism, metabolic pathways, carbon metabolism, carbon fixation in photosynthetic organisms, and glyoxylate and dicarboxylate metabolism. Additionally, six DEGs that were not enriched in any pathway were screened (|log$_2$ (Fold change)| ≥ 2 and FPKM≥50), including four up-regulated and two down-regulated. In total, 23 candidate genes were screened as shown in Table 2.

**Table 2** Screened candidate genes in Y35AB and Y35A

| Gene symbol | Unigene ID          | Protein properties                                           | log$_2$FC$^a$ | log$_2$FC$^b$ | pathway                                      |
|-------------|---------------------|--------------------------------------------------------------|---------------|---------------|----------------------------------------------|
| PsaFa       | Niben101Scf 00271g04024 | photosystem I reaction center subunit III, chloroplastic-like | -1.4180       | -1.1769       | Photosynthesis                               |
| PsaFb       | Niben101Scf 02156g05001 | photosystem I reaction center subunit III, chloroplastic-like | -1.5190       | -1.6335       | Photosynthesis                               |
| PsaFc       | Niben101Scf 04964g00002 | photosystem I reaction center subunit III, chloroplastic-like | -1.5257       | -1.2865       | Photosynthesis                               |
| PsaNa       | Niben101Scf 17701g01020 | photosystem I reaction center subunit N, chloroplastic-like  | -1.6702       | -1.5537       | Photosynthesis                               |
| PsaNb       | Niben101Scf 35628g00001 | photosystem I reaction center subunit N, chloroplastic-like  | -1.3129       | -1.0396       | Photosynthesis                               |
| ChlHa       | Niben101Scf 04388g00011 | magnesium-chelatase subunit ChlH, chloroplastic-like        | -1.6623       | -2.1794       | Porphyrin and chlorophyll metabolism         |
| ChlHb       | Niben101Scf 07214g00015 | magnesium-chelatase subunit ChlH, chloroplastic-like        | -1.4638       | -1.8585       | Porphyrin and chlorophyll metabolism         |
| BchP        | Niben101Scf 06249g03002 | geranylgeranyl diphosphate reductase, chloroplastic-like    | -1.8576       | -1.0521       | Porphyrin and chlorophyll metabolism         |
| HemA        | Niben101Scf 10063g00003 | glutamyl-tRNA reductase 1, chloroplastic-like               | -1.4101       | -2.0903       | Porphyrin and chlorophyll metabolism         |
| DXS         | Niben101Scf 00246g04005 | probable 1-deoxy-D-xylulose-5-phosphate synthase, chloroplastic-like | -1.0040       | -1.0126       | Metabolic pathways                           |
| GME1-X1     | Niben101Scf 01696g06050 | GDP-mannose 3,5-epimerase 1-like, transcript variant X1     | -1.6022       | -1.1137       | Metabolic pathways                           |
| CAB3A       | Niben101Scf 03961g00004 | chlorophyll a-b binding protein 50, chloroplastic-like     | -1.1418       | -1.0719       | Metabolic pathways                           |
| GDCSa       | Niben101Scf 02480g02012 | glycine dehydrogenase (decarboxylating), mitochondrial-like | -1.5732       | -1.1698       | Carbon metabolism                            |
| GDCSb       | Niben101Scf 03839g04019 | glycine dehydrogenase (decarboxylating), mitochondrial-like | -1.1991       | -1.0395       | Carbon metabolism                            |
| rbcS        | Niben101Scf 02381g04022 | ribulose bisphosphate carboxylase small chain S41, chloroplastic-like | -1.2691       | 0.7603        | Carbon fixation in photosynthetic organisms  |
| FBA1        | Niben101Scf 02864g04008 | fructose-bisphosphate aldolase 1, chloroplastic-like       | -1.3167       | -1.0072       | Carbon fixation in photosynthetic organisms  |
| gisF        | Niben101Scf 04198g01002 | ferredoxin-dependent glutamate synthase, chloroplastic-like | -1.1161       | -1.0704       | Glyoxylate and dicarboxylate metabolism      |
| GRP         | Niben101Scf 03052g00011 | glycine-rich protein-like                                    | 2.1379        | 1.4647        | No pathway enriched                          |
| GRP3a       | Niben101Scf 01084g05013 | glycine-rich protein 3-like                                  | 3.9423        | 3.1735        | No pathway enriched                          |
| GRP3b       | Niben101Scf 01084g03004 | glycine-rich protein 3-like                                  | 2.0529        | 1.8836        | No pathway enriched                          |
| CAX3        | Niben101Scf 01329g00002 | vacuolar cation/proton exchanger 3-like                     | -2.2799       | -1.2535       | No pathway enriched                          |
| FSD1        | Niben101Scf 03679g03009 | Fe superoxide dismutase , chloroplastic-like                | -2.1527       | -1.9924       | No pathway enriched                          |
| PLI         | Niben101Scf 10524g02015 | putative uncharacterized protein                            | 2.0616        | 0.9243        | No pathway enriched                          |

**Validation of the selected genes by RT-qPCR**

To validate the RNA-seq data, 12 genes were selected randomly for RT-qPCR analysis from the Table 2, including genes GDCSa, GDCSb, PsaFa, PsaFb, PsaFc, PsaNa, PsaNb, gisF, CAX3, GRP, GRP3a and PLI. The results indicated that the genes expression patterns by RT-qPCR were consistent with that by transcriptome sequencing (Fig. 5), the fold changes between RNA-seq and RT-qPCR were

*FC* Fold change
*Y35AB vs. CK*
*Y35A vs. CK*
different which maybe caused by the sensitive of methodology or the used samples between RNA-seq and RT-qPCR were not the same one.

Analysis of pathogenesis-related (PR) genes
An important feature of the plant defense in response to pathogen attack is the induction and accumulation of various PR proteins which are also a part of systematic acquired resistance (SAR). There were 29 PR genes identified by a local BLAST in RNA-seq dataset and online BLAST analysis based on reported PR family genes in *N. tabacum*. Among them, eight genes (27.6%) were significantly up-regulated, including PR-2 (Glucan endo-1,3-beta-glucosidase-like), PR-3 (Chitinase 8, glycoside hydrolase), PR-5 (Pathogenesis-related thaumatin superfamily protein), PR-11 (Chitinase-3-like protein 2), PR-17 (Plant basic secretory protein family protein), of them four genes were found in Y35AB vs CK, and five genes were observed in Y35A vs CK (Table 3).

Analysis of LRR-RLKs genes
Leucine-rich repeat receptor-like protein kinases (LRR-RLKs) are the largest of receptor-like kinases in plants and play crucial roles in development and stress responses. In the present study, We combined RNA-seq dataset obtained LRR-RLK sequences with previously

Fig. 5 Comparison of relative expression levels determined by RNA-seq and qPCR on 12 genes selected. All qPCR reactions were used three biological replica samples, repeated three times, and the data are presented as the mean ± SD. Relative expression levels are calculated from Ct values according to the $2^{-ΔΔCt}$ method. Actin was the reference gene for these qPCR experiments. Asterisks indicate statistically significant differences compared with CK, ** indicate significant difference (0.01 ≤ p < 0.05), *** indicate extremely significant difference (p < 0.01)
NCBI reported sequences to perform analysis, results showed that a total of 71 unigenes were identified potential LRR-RLK genes in RNA-seq dataset, however, 44 significantly different unigenes were observed. Among them, 18 unigenes were found in Y35AB vs CK with 10 unigenes down-regulated and the other 8 unigenes up-regulated, and 32 unigenes expressed in Y35A vs CK, with 20 and 12 unigenes were down-regulated and up-regulated, respectively. Five unigenes were commonly regulated in both Y35AB vs CK and Y35A vs CK, including one up-regulated and four down-regulated unigenes (Table 4).

**Table 3** Screened PRP family genes

| PR family | Unigene ID          | Protein properties                                      | log₂FC<sub>a</sub> | qvalue<sub>a</sub> | log₂FC<sub>b</sub> | qvalue<sub>b</sub> |
|-----------|---------------------|---------------------------------------------------------|---------------------|-------------------|---------------------|-------------------|
| PR1       | Niben101Scf13926g01014 | Cysteine-rich secretory protein, allergen V5/Tpx-1-related |                     |                   |                     |                   |
|           | Niben101Scf03376g03004 | Cysteine-rich secretory protein, allergen V5/Tpx-1-related |                     |                   |                     |                   |
|           | Niben101Scf00107g03008 | Cysteine-rich secretory protein, allergen V5/Tpx-1-related |                     |                   |                     |                   |
|           | Niben101Scf01999g07002 | Cysteine-rich secretory protein, allergen V5/Tpx-1-related |                     |                   |                     |                   |
| PR2       | Niben101Scf01001g00005 | Glucan endo-1,3-beta-glucosidase-like, Glycoside hydrolase, family 17 |                     |                   |                     |                   |
|           | Niben101Scf01001g00004 | Glucan endo-1,3-beta-glucosidase-like, Glycoside hydrolase, family 17 |                     |                   |                     |                   |
|           | Niben101Scf01001g00003 | Glucan endo-1,3-beta-glucosidase-like, Glycoside hydrolase, family 17 |                     |                   |                     |                   |
|           | Niben101Ctg13736g00004 | Glucan endo-1,3-beta-glucosidase-like, Glycoside hydrolase, family 17 |                     |                   |                     |                   |
|           | Niben101Scf04869g03002 | Glucan endo-1,3-beta-glucosidase-like, Glycoside hydrolase, family 17 |                     |                   |                     |                   |
|           | Niben101Scf01001g00006 | Glucan endo-1,3-beta-glucosidase-like, Glycoside hydrolase, family 17 |                     |                   |                     |                   |
| PR3       | Niben101Scf02041g00002 | Chitinase 8, Glycoside hydrolase, family 19 |                     |                   |                     |                   |
| PR4       | Niben101Scf01400g00014 | Thaumatin-like protein |                     |                   |                     |                   |
|           | Niben101Scf03436g01016 | Thaumatin-like protein |                     |                   |                     |                   |
| PR5       | Niben101Scf00126g00008 | Pathogenesis-related thaumatin superfamily protein | 0.9949              | 0.027102          |                     |                   |
|           | Niben101Scf05554g05006 | Pathogenesis-related thaumatin superfamily protein | 1.4385              | 3.24E-08          |                     |                   |
| PR6       | Niben101Scf00953g05001 | Cysteine-rich secretory protein, allergen V5/Tpx-1-related |                     |                   |                     |                   |
|           | Niben101Scf04053g01004 | Cysteine-rich secretory protein, allergen V5/Tpx-1-related |                     |                   |                     |                   |
| PR9       | Niben101Scf03460g04004 | Peroxidase 53, Haem peroxidase |                     |                   |                     |                   |
|           | Niben101Scf07182g05012 | Peroxidase 53, Haem peroxidase |                     |                   |                     |                   |
| PR10      | Niben101Scf03526g00006 | Major pollen allergen Bet v 1-M/N, Bet v I type allergen |                     |                   |                     |                   |
|           | Niben101Scf10735g00016 | Major pollen allergen Bet v 1-M/N, Bet v I type allergen |                     |                   |                     |                   |
|           | Niben101Scf02474g01024 | Major pollen allergen Bet v 1-M/N, Bet v I type allergen |                     |                   |                     |                   |
|           | Niben101Scf01938g04007 | Major pollen allergen Bet v 1-M/N, Bet v I type allergen |                     |                   |                     |                   |
| PR11      | Niben101Scf06295g04023 | Chitinase-3-like protein 2, Glycoside hydrolase superfamily | 4.4843              | 0.007133          |                     |                   |
|           | Niben101Scf01789g04010 | Chitinase-3-like protein 2, Glycoside hydrolase superfamily |                     |                   |                     |                   |
| PR17      | Niben101Scf03385g02011 | Plant basic secretory protein family protein, uncharacterised protein family |                     |                   |                     |                   |
|           | Niben101Scf03385g01006 | Plant basic secretory protein family protein, uncharacterised protein family |                     |                   |                     |                   |
|           | Niben101Ctg10643g00004 | Plant basic secretory protein family protein, uncharacterised protein family |                     |                   |                     |                   |

log₂FC > 0, up-regulated, log₂FC < 0, down-regulated. Inf: the readcount of control (CK) was zero. Blank lattices show genes expression has no significant difference.

FC Fold change

*Y35AB vs. CK

Y35A vs. CK
| Unigene ID | Protein properties | log 2FC<sup>a</sup> | qvalue<sup>a</sup> | log 2FC<sup>b</sup> | qvalue<sup>b</sup> |
|------------|--------------------|---------------------|-------------------|---------------------|-------------------|
| Niben101Scf13404g00002 | Brassinosteroid LRR receptor-like kinase, BRI1 | -0.8361 | 0.000272 | | |
| Niben101Scf07123g01015 | Elicitor-inducible leucine-rich repeats receptor-like protein | 5.5080 | 0.000289 | | |
| Niben101Scf00104g02013 | F-box/LRR-repeat protein 17-like | 0.8664 | 0.021066 | | |
| Niben101Scf04252g00009 | F-box/LRR-repeat protein 17-like | 1.8551 | 0.003213 | | |
| Niben101Scf09559g00005 | F-box/LRR-repeat MAX2 homolog A-like | 0.8581 | 0.018415 | | |
| Niben101Scf15394g00006 | F-box/LRR-repeat MAX2 homolog A-like | -0.7116 | 0.011214 | | |
| Niben101Scf05301g03002 | LRR receptor-like serine/threonine-protein kinase ERL1 | -0.7516 | 0.013687 | | |
| Niben101Scf00247g01007 | LRR receptor-like serine/threonine-protein kinase ERL1 | -0.5633 | 0.024694 | | |
| Niben101Scf11723g02003 | LRR receptor-like serine/threonine-protein kinase ERL1 | -0.6679 | 0.026512 | | |
| Niben101Scf03619g03004 | LRR receptor-like serine/threonine-protein kinase GSO1 | -0.5426 | 0.048845 | | |
| Niben101Scf03925g01010 | LRR receptor-like serine/threonine-protein kinase GSO1 | 2.3343 | 0.007479 | | |
| Niben101Scf02203g04006 | LRR receptor-like serine/threonine-protein kinase GSO1 | -0.6640 | 0.048845 | | |
| Niben101Scf00073g04006 | probable LRR receptor-like serine/threonine-protein kinase At1g07650 | -1.4546 | 0.005161 | | |
| Niben101Scf00418g02007 | probable LRR receptor-like serine/threonine-protein kinase At4g37250 | -0.5934 | 0.047527 | | |
| Niben101Scf00920g09027 | probable LRR receptor-like serine/threonine-protein kinase At5g48740 | -1.0192 | 0.001468 | | |
| Niben101Scf01061g08014 | probable LRR receptor-like serine/threonine-protein kinase At5g20940 | 0.9239 | 0.032952 | | |
| Niben101Scf02357g08010 | probable LRR receptor-like serine/threonine-protein kinase At1g07650 | -0.6424 | 0.001803 | | |
| Niben101Scf02665g15003 | probable LRR receptor-like serine/threonine-protein kinase At1g06840 | -0.5769 | 0.049996 | | |
| Niben101Scf02827g07005 | probable LRR receptor-like serine/threonine-protein kinase At3g45750 | 2.3756 | 0.027087 | | |
| Niben101Scf03098g00011 | probable LRR receptor-like serine/threonine-protein kinase At1g34110 | -1.2224 | 3.35E-06 | | |
| Niben101Scf03262g02006 | probable LRR receptor-like serine/threonine-protein kinase At2g24230 | -0.6865 | 0.030882 | | |
| Niben101Scf03445g05005 | probable LRR receptor-like serine/threonine-protein kinase At4g26540 | -0.8802 | 0.006699 | | |
| Niben101Scf03735g07010 | probable LRR receptor-like serine/threonine-protein kinase At4g36180 | -1.3373 | 3.02E-05 | | |
| Niben101Scf03448g01026 | probable LRR receptor-like serine/threonine-protein kinase At1g12460 | -1.2826 | 8.11E-05 | | |
| Niben101Scf05105g07003 | probable LRR receptor-like serine/threonine-protein kinase At3g45750 | 2.0599 | 0.003076 | | |
| Niben101Scf05405g07006 | probable LRR receptor-like serine/threonine-protein kinase At3g45750 | 1.5589 | 0.022376 | | |
| Niben101Scf05767g05006 | probable LRR receptor-like serine/threonine-protein kinase At4g20940 | -0.8478 | 0.000497 | | |
| Niben101Scf05977g01008 | probable LRR receptor-like serine/threonine-protein kinase At1g12460 | -1.2826 | 8.11E-05 | | |
| Niben101Scf06144g00013 | probable LRR receptor-like serine/threonine-protein kinase At4g20940 | -0.9557 | 7.81E-05 | | |
| Niben101Scf06215g02014 | probable LRR receptor-like serine/threonine-protein kinase At1g67720 | 0.7683 | 0.046840 | | |
| Niben101Scf06562g02013 | probable LRR receptor-like serine/threonine-protein kinase At4g26540 | -0.9943 | 2.01E-05 | | |
| Niben101Scf06789g03005 | probable LRR receptor-like serine/threonine-protein kinase At1g67720 | 0.9483 | 0.042093 | | |
| Niben101Scf07034g06018 | probable LRR receptor-like serine/threonine-protein kinase At4g36180 | -0.8501 | 0.034812 | | |
| Niben101Scf07515g02006 | probable LRR receptor-like serine/threonine-protein kinase At1g56140 | 0.6848 | 0.00559 | | |
| Niben101Scf07812g01013 | probable LRR receptor-like serine/threonine-protein kinase At4g20940 | -0.7450 | 0.014076 | | |
| Niben101Scf07969g01023 | probable LRR receptor-like serine/threonine-protein kinase At4g30520 | 0.7628 | 0.001146 | | |
| Niben101Scf07736g01007 | probable LRR receptor-like serine/threonine-protein kinase At1g63430 | 0.6791 | 0.006612 | | |
| Niben101Scf08134g02008 | probable LRR receptor-like serine/threonine-protein kinase At4g32750 | -0.7817 | 0.000797 | | |
| Niben101Scf1101g00023 | probable LRR receptor-like serine/threonine-protein kinase At4g30520 | 0.6651 | 0.002453 | | |
| Niben101Scf11047g00001 | probable LRR receptor-like serine/threonine-protein kinase At2g16250 | 0.7640 | 0.047280 | | |

log 2FC > 0, up-regulated, log 2FC < 0, down-regulated. Blank lattices show genes expression has no significant difference.

<sup>a</sup>FC: Fold change
<sup>b</sup>Y35AB vs. CK
<sup>c</sup>Y35A vs. CK
Analysis of brassinosteroid and jasmonic acid pathway

Plant hormones are not only essential for growth and development, but also play crucial roles in plant defense responses [33, 34]. The transcriptome data showed that only one brassinosteroid (BR) synthesis-related gene, 3-epi-6-deoxocathasterone 23-monooxygenase (CYP90C1/D1) was significantly decreased upon virus infection (Fig. 6a), it is likely result in a reduction of brassinosteroid. By contrast, expressions of jasmonic acid (JA) -related biosynthesis genes DOX1, HPL1, SAMT-X1 were decreased and PLA1/A2, luxF, KAT2, SAMT-X2 were significantly increased upon virus infection (Fig. 6b). We conjecture that the expression level of jasmonic acid might was increase. In addition, part of BR and JA signaling genes expression level were also altered, BRI1, BZR2, CYCD3-X2 were down-regulated and BZR1, CYCD3-X1 were up-regulated in BR signal transduction pathway, and JAR1, COI1, MYC2 were down-regulated and JAZ1, JAZ2 were up-regulated in JA signal transduction pathway (Fig. 6c). All synthesis- and signal transduction- related significantly difference expressed unigenes of BR and JA pathway shown in Table 5.
Discussion

TbCSV is a serious threat to many crops in China. In this study, RNA-seq-based transcriptomic characterization and comparative analysis of TbCSV-infected *N. benthamiana* with control plants was used to shed some light on understanding the molecular interactions of this pathosystem. 4081 and 3196 DEGs were screened from Y35AB vs CK and Y35A vs CK, respectively. Among the top 15 enriched physiological pathways, the porphyrin and chlorophyll metabolism, photosynthesis and pentose phosphate pathway were significantly affected in Y35AB vs CK. It suggested that, the disorder of these pathways might be responsible to the more serious leaf curling symptom on Y35AB-infected *N. benthamiana* plants. In addition, previous reports showed that βC1, an essential pathogenicity protein encoded by betasatellite, is required for symptom induction and RNA silencing suppression [2, 16, 35–37]. Thus, βC1 protein might contribute to the specific changes of those genes or biological pathways in Y35AB-infected plants.

In our study, the DEGs expression with high levels that mainly concerned with photosynthesis-related genes including photosynthesis, porphyrin and chlorophyll metabolism and carbon fixation in photosynthetic organisms were down-regulated in *N. benthamiana* plants infected with TbCSV. Meanwhile, the KEGG pathway analysis showed significant enrichments of the photosynthesis, and the porphyrin and chlorophyll metabolism pathway in Y35AB-infected plants, with higher rich factor. It has been reported that photosynthesis-related genes could be affected upon virus infection such as *Tobacco mosaic virus* (TMV) [38], *Cucumber mosaic virus* (CMV) [39], *Rice stripe virus* (RSV) [40, 41] and *Alfalfa mosaic virus* (AMV) [42]. And the chloroplast, one of the most dynamic organelles of plant cell, plays an active part in the defense response and is crucial for inter organelle signaling [43]. Therefore, viruses need to suppress the chloroplast-mediated defense by employing pathogenicity factors, such as effector proteins, for successful infections [44, 45]. Furthermore, a previous study showed that a Radish leaf curl betasatellite (RaLCB) damages the structural and functional integrity of chloroplasts, leading to inhibition of photosynthesis and symptom formation [46]. Some photosynthesis-related genes were down-regulated during both Y35A and Y35AB

| Table 5 Significant difference expression unigenes of BR and JA pathway |
|---------------------------------------------|
| Gene symbol | Unigene ID | Protein properties | log2FC<sup>a</sup> | qvalue<sup>a</sup> | log2FC<sup>b</sup> | qvalue<sup>b</sup> |
|---------------|-------------|-------------------|-----------------|----------------|-----------------|----------------|
| BR CYP90C1/D1 | Niben101Scf05764g04013 | 3-epi-6-deoxocathasterone 23-monoxygenase | -0.8193 | 0.019172 | -0.8361 | 2.72E-04 |
| BRI1          | Niben101Scf13404g00002 | Brassinosteroid LRR receptor kinase-like, BRI1 | -0.8361 | 2.72E-04 | -0.8361 | 2.72E-04 |
| BZR1          | Niben101Scf03110g05009 | Brassinosteroid resistant 1 | 0.5592 | 0.025611 | -1.2947 | 7.36E-07 |
| BZR2          | Niben101Scf06112g01006 | Brassinosteroid resistant 2 | 1.3753 | 1.44E-07 | 1.2863 | 0.002511 |
| CYCD3-X1      | Niben101Scf05643g01012 | Cyclin D3, 3 protein, plant | 1.3753 | 1.44E-07 | 1.2863 | 0.002511 |
| CYCD3-X1      | Niben101Scf02445g10017 | Cyclin D3, 3 protein, plant | 1.3753 | 1.44E-07 | 1.2863 | 0.002511 |
| CYCD3-X2      | Niben101Scf00107g02002 | Cyclin D3, 2 protein, plant | -0.8767 | 0.017675 | -1.2925 | 2.43E-04 |
| JA PLA1/A2    | Niben101Scf00711g02004 | Triacylglycerol lipase SDP1-like; Phospholipase A1/A2 | 1.8227 | 0.003036 | -3.9498 | 7.95E-14 |
| DOX1          | Niben101Scf07070g02002 | Alpha-dioxxygenase | -3.8126 | 7.60E-13 | -3.9498 | 7.95E-14 |
| loxF          | Niben101Scf05407g00001 | Linoleate 9S-lipoxygenase-like | 0.9407 | 0.021540 | 0.6201 | 0.007053 |
| HPL1          | Niben101Scf00313g08016 | Fatty acid hydroperoxide lyase, HPL1 | -1.2170 | 1.14E-09 | -0.9629 | 7.37E-07 |
| HPL1          | Niben101Scf02207g11009 | Fatty acid hydroperoxide lyase, HPL1 | -1.4966 | 1.06E-05 | 0.6201 | 0.007053 |
| KAT2          | Niben101Scf10189g02009 | 3-ketoacyl-CoA thiolase 2, peroxisomal | -1.4545 | 2.97E-11 | -0.8505 | 2.96E-04 |
| SAMT-X1       | Niben101Scf05122g00005 | Salicylate carboxymethyltransferase-like | -1.4545 | 2.97E-11 | -0.8505 | 2.96E-04 |
| SAMT-X2       | Niben101Scf01146g12006 | Salicylate carboxymethyltransferase-like | -1.4545 | 2.97E-11 | -0.8505 | 2.96E-04 |
| JAR1          | Niben101Scf01076g00005 | Jasmonic acid-amino synthetase | -0.6477 | 0.003456 | -0.8340 | 0.017053 |
| COI1          | Niben101Scf02280g08005 | Coronatine-insensitive 1, COI1 | -0.6123 | 0.024317 | -0.8990 | 2.81E-04 |
| JAZ1          | Niben101Scf00298g01006 | Jasmonate ZIM-domain protein 1 | 0.7577 | 0.039555 | 0.7728 | 3.95E-04 |
| JAZ2          | Niben101Scf07798g03026 | Jasmonate ZIM domain-containing protein | 0.7577 | 0.039555 | 0.7728 | 3.95E-04 |
| MYC2          | Niben101Scf15224g00002 | Transcription factor MYC2 | -0.6238 | 0.042856 |

log2FC > 0, up-regulated, log2FC < 0, down-regulated. Blank lattices show genes expression has no significant difference

FC: Fold change

<sup>a</sup>Y35AB vs. CK

<sup>b</sup>Y35A vs. CK
infections, indicating that these genes might participate in the interaction between TbCSV and N. benthamiana.

An important feature of the plant defense response to pathogen attack is the induction and accumulation of various PR proteins, which are also a part of systematic acquired resistance (SAR) [47]. Previously, PR proteins were divided into five families based on their localization, isoelectric point, molecular mass, and biological activity [48]. PR proteins are currently categorized into 17 families. At present, besides PR-6 is predicted by automated computational analysis based on a genomic sequence (gene ID: LOC107768382), nine out of 17 families, including PR-1, PR-2 [49], PR-3, PR-4, PR-5 [50], PR-9 [51], PR-10 [52], PR-11 [53], PR-17 [54] families were reported that associated with the acquired resistance to pathogen infections. Among these PR proteins, PR-1 is generally considered as a marker gene of disease resistance [55]. Based on NCBI retrieval and local BLAST analysis, 29 PR genes in N. benthamiana were identified. In the present study, the expression levels of 8 PR genes were found to be significantly regulated in virus-infected N. benthamiana plants (Table 3). It was previously reported that in pathogens-infected wheat, the genes encoding peroxidase, PR-1, PR-2, PR-3, PR-4, and PR-5 were induced after 6–12 h and reached the highest levels at 36–48 h [56, 57]. In our research, leaf tissues of N. benthamiana plants for RNA-seq assays were collected at 20 dpi, and the transcript levels of expression levels of 8 PR genes were identified. In the present study, the expression of these genes can be determined in a time course study. Based on our research, we speculated that some types of PR proteins might play roles in defense response to TbCSV infection.

In plants, many cellular signaling transductions are mediated by receptor-like kinases (RLKs). Leucine-rich repeat receptor-like protein kinases (LRR-RLKs) are the largest group of receptor-like kinases in plants and play crucial roles in development and stress responses [58]. LRR-RLKs regulate a wide range of plant growth and development, including meristem development [59, 60], secondary growth [61], microsporogenesis and embryo-genesis [62]. Besides, some LRR-RLKs mediated plant resistance against bacterial [63, 64] or viral pathogens [65, 66]. In this study, 71 LRR-RLKs genes were identified. Among them, 44 were differentially expressed in virus-infected N. benthamiana plants. The results showed that these LRR-RLKs genes might play important roles in N. benthamiana responses to TbCSV infection. By contrast, in the present of TbCSB, less number of LRR-RLKs genes were regulated (Table 4).

BR and JA are two important plant hormones that play crucial roles in plant growth and development [67] and also involved in plant defense responses to pathogen infections [68–70]. In the BR signaling pathway, BRASSINAZOLE RESISTANT (BZR) is a transcription factor, which orchestrating plant developmental and physiological processes by regulating BR-target gene expression [71, 72]. BRASSINOSTEROID INSENSITIVE 1 (BRI1) is a brassinosteroid receptor, of which ubiquitously expresses plasma membrane-localized protein kinase [73], belonging to an LRR-RLK subfamily [74]. For the JA signaling transduction, the perception of JA is mediated by the co-receptor CORONATINE INSENSITIVE 1 (COI1), which is an F-box protein cooperated with a family of the JASMONATE ZIM-domain (JAZ) transcriptional proteins. Upon the pathogens infection, bioactive JAs promote direct interaction between COI1 and JAZ proteins and then trigger the SCF-COI1 complex, resulting in poly-ubiquitination and degradation of the JAZ proteins. Ultimately, a variety of transcription factors could be activated for the expression of the JA-responsive genes [75]. Besides, MYC2 is a bHLH (basic helix-loop-helix) transcription factor and a positive regulator of JA responses [76]. A study has shown that rhizobacterium-mediated induction of JA reduces the symptoms of CMV infection in Arabidopsis thaliana ecotype Columbia plants [77]. In this study, BRI1, BZR1, BZR2, CYCD3, JAR1, COI1, JAZ and MYC2 genes were significantly regulated under the infection of TbCSV (Fig. 6). It suggested that BR and JA pathways played a role in response to TbCSV infection.

Conclusion

In this study, a genome-wide transcript profile was obtained for N. benthamiana plant infected by TbCSV or plus TbCSB using RNA-seq approaches, and genes involved in plant defense system were found to be significantly regulated after TbCSV and TbCSV/TbCSB infection. The information provided in our study would be particularly useful for investigating the molecular mechanisms concerning begomovirus-host interaction and excavate resistance genes.

Additional files

Additional file 1: Table S1. Primers used for RT-qPCR. (DOCX 18.9 kb)
Additional file 2: Table S2. List of DEGs between Y35A vs CK. (XLS 796 kb)
Additional file 3: Table S3. List of DEGs between Y35A vs CK. (XLS 629 kb)
Additional file 4: Table S4. Enriched GO terms of DEGs between Y35AB vs CK. (XLS 3.15 mb)
Additional file 5: Table S5. Enriched GO terms of DEGs between Y35A vs CK. (XLS 2.83 mb)
Additional file 6: Table S6. Enriched KEGG pathways of DEGs between Y35AB vs CK. (XLS 216 kb)
Additional file 7: Table S7. Enriched KEGG pathways of DEGs between Y35A vs CK. (XLS 181 kb)

Acknowledgements

We are grateful to Professor Xueping Zhou in the Biotechnology Institute of Zhejiang University for providing the infectious clones of TbCSV isolate Y35 (Y35A) and its betasatellite (Y35B).
Funding
This research work was supported by the Fundamental Research Funds for the Central Universities (Grant No. XDJK2017A006 and XDJK2015G039), National Natural Science Foundation of China (Grant No. 31772127), China Postdoctoral Science Foundation (2015M572431), and Chongqing Postdoctoral Science special Foundation (Xm2015130).

Availability of data and materials
All data and materials described in the manuscript are available in the Additional files 2, 3, 4, 5, 6 and 7.

Authors’ contributions
LQ conceived and designed the study and revised the paper. KL, GW, ML and MM performed the experiments and made analysis of the data. JD, MS and XS participated preparation including discussion and editing. All authors read and approved the final manuscript.

Ethics approval and consent to participate
Not applicable.

Consent for publication
All the authors consent to publish.

Competing interests
The authors declare that they have no competing interests.

Publisher’s Note
Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Received: 4 January 2018 Accepted: 14 August 2018
Published online: 03 September 2018

References
1. Xie Y, Zhou XP, Zhang ZK, Qi YJ. Tobacco curly shoot virus isolated in Yunnan is a distinct species of Begomovirus. Chin Sci Bull. 2002;47:197–200.
2. Zhou XP, Xie Y, Tao XR, Zhang ZK, Li ZH, Fauquet CM. Characterization of DNAβ associated with begomoviruses in China and evidence for co-evolution with their cognate viral DNA-A. J Gen Virol. 2003;84:237–47.
3. Pacheco R, Garcia-Marcos A, Manzano A, Garcia De Lacoboa M, Camanes G, Garcia-Agustín P, Ramón-Díaz-Ruíz J, Tenllado F. Comparative analysis of transcriptionic and hormonal responses to compatible and incompatible plant-virus interactions that lead to cell death. Mol Plant-Microbe Interact. 2012;25:709–23.
4. Kamitani M, Nagano AJ, Honjo MN, Kudoh H, Kummerli R. RNA-Seq reveals virus-virus and virus-plant interactions in nature. Front Microbiol. 2016;7:921–11.
5. Mozzé L, Napoli C, Sardo L, Accotto GP. Transcriptomics of the interaction between the monopartite phiL0m0-lim0tinivirus Tomato Yellow Leaf Curl Sardinia Virus and Solanum lycopersicum highlights a role for plant hormones, autophagy and plant immune system fine tuning during infection. PLoS One. 2014;9:e89511.
6. Wang Z, Genest M, Snyder M. RNA-Seq: a revolutionary tool for transcriptomics. Nat Rev Genet. 2009;10:57–63.
7. Naqv RZ, Zaidi SS, Akhtar KP, Stricker S, Woldemariam M, Mishra B, Mukhtar MS, Scheffler BE, Scheffler JA, Jander G, et al. Transcriptomics reveals multiple resistance mechanisms against cotton leaf curl disease in a naturally immune cotton species, Gossypium arboreum. Sci Rep. 2017;7:15580.
8. Pavan Kumar BK, Kanakala S, Malathi VG, Gopal P, Usha R. Transcriptomic and proteomic analysis of yellow mosaic diseased soybean. J Plant Biochem Biotechnol. 2016;26:224–34.
9. Sood A, Chauhan RS. Comparative NGS transcriptomics unravels molecular components associated with mosaic virus infection in a bioenergy plant species, Jatropha curcas L. BioEnergy Res. 2017;10:129–45.
10. Perce EJ, Roy-MEC. Assessing global transcriptome changes in response to South African Cassava Mosaic Virus (ZA-99) infection in susceptible Arabidopsis thaliana. PLoS One. 2013;8:e67534.
11. Yadav RK, Chattopadhyay D. Differential soybean gene expression during early phase of infection with Mungbean yellow mosaic India virus. Mol Biol Rep. 2014;41:5123–34.
12. Kundu A, Patel A, Paul S, Pal A. Transcript dynamics at early stages of molecular interactions of MYMV with resistant and susceptible genotypes of the leguminous host, vigna mungo. PLoS One. 2015;10:e0124687.
13. Kushwaha N, Sahu PP, Prasad M, Chakraborty S. Chili leaf curl virus infection highlights the differential expression of genes involved in protein homeostasis and defense in resistant chili plants. Appl Microbiol Biotechnol. 2015;99:4757–70.
14. Li ZH, Xie Y, Zhou XP. Tobacco curly shoot virus DNAB is not necessary for infection but intensifies symptoms in a host-dependent manner. Phytopathology. 2005;95:902–8.
15. Cui XF, Zhou XP. AC2 and AC4 proteins of Tomato yellow leaf curl China virus and Tobacco curly shoot virus mediate suppression of RNA silencing. Chin Sci Bull. 2004;49:2607–12.
16. Cui XF, Li GX, Wang DW, Hu DW, Zhou XP. A Begomovirus DNAβ-encoded protein binds DNA, functions as a suppressor of RNA silencing, and targets the cell nucleus. J Virol. 2005;79:10764–75.
17. Abarshi MM, Mohammed IU, Warswa P, Hillocks RJ, Holt J, Legg JP, Seal SE, Maruthi MN. Optimization of diagnostic RT-PCR protocols and sampling procedures for the reliable and cost-effective detection of Cassava brown streak virus. J Virol Methods. 2010;163:353–9.
18. Xiong HC, Guo HJ, Xie YD, Zhao LS, Gu YJ, Zhao SR, Li JH, Liu LX. RNAseq analysis reveals pathways and candidate genes associated with salinity tolerance in a salt-tolerant wheat mutant. Sci Rep. 2017;7:27231.
19. Chen JW, Hou K, Qin P, Liu H, Bi N, Yang WT, Wu W. RNA-Seq for gene identification and transcript profiling of three Stevia rebaudiana genotypes. BMC Genomics. 2014;15:571–81.
20. Deng XG, Zhu T, Peng XJ, Xi DH, Guo H, Yin Y, Zhang DW, Lin HH. Role of brassinosteroid signaling in modulating tobacco mosaic virus resistance in Nicotiana benthamiana. Sci Rep. 2016;6:20579.
21. Langmead B, Salzberg SL. Fast gapped-read alignment with bowtie2. Nat Methods. 2012;9:357.
22. Kim D, Pertea G, Trapnell C, Pimentel H, Kelley R, Salzberg SL. TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. Genome Biol. 2013;14:R36.
23. Trapnell C, Roberts A, Goff L, Pertea G, Kim D, Kelley DR, Pimentel H, Salzberg SL, Rinn JL, Pachter L. Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. Nat Protoc. 2012;7:562–78.
24. Anders S, Pyl PT, Huber W. HTSeq—a Python framework to work with high-throughput sequencing data. Bioinformatics. 2015;31:166–9.
25. Trapnell C, Pachter L, Salzberg SL. TopHat: discovering splice junctions with RNA-Seq. Bioinformatics. 2009;25:1105–11.
26. Anders S, Huber W. Differential expression of RNA-Seq data at the gene level—the DESeq package. Heidelberg: Embi; 2012.
27. Young MD, Wakefield MJ, Smyth GK, Oshlack A. Gene ontology analysis for RNA-seq: accounting for selection bias. Genome Biol. 2010;11:R14.
28. Mao X, Cai T, Olyarchuk JG, Wei L. Automated genome annotation and pathway identification using the KEGG Orthology (KO) as a controlled vocabulary. Bioinformatics. 2005;21:787–93.
29. Livak KJ, Schmittgen TD. Analysis of relative gene expression data using real-time quantitative PCR and the 2^{-ΔΔCT} method. Methods. 2002;25:402–8.
30. Peng H, He XJ, Gao J, Ma HX, Zhang ZM, Shen Y, Pan GT, Lin HJ. Transcriptomic changes during maize roots development responsive to Cadmium (Cd) pollution using comparative RNAseq-based approach. Biochem Biophys Res Commun. 2015;464:1040–7.
31. Trapnell C, Williams BA, Pertea G, Mortazavi A, Kwan G, Van Baren MJ, Salzberg SL, Wold BJ, Pachter L. Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. Nat Biotechnol. 2010;28:511–5.
32. Audic S, Claverie JM. The significance of digital gene expression profiles. Genome Res. 1997;7:986–95.
33. Whatham SA, Yang C, Goodin MM. Global impact: elucidating plant responses to viral infection. Mol Plant-Microbe Interact. 2006;19:1207.
34. Alazem M, Lin NS. Roles of plant hormones in the regulation of host–virus interactions. Mol Plant Pathol. 2015;16:529.
35. Saeed M, Behjatnia S, Mansoor S, Zafar Y, Hasnain S, Rezaian MA. A single complementarity-sense transcript of a geminiviral DNA beta satellite is determinant of pathogenicity. Mol Plant-Microbe Interact. 2005;18:7–14.
36. Cui X, Tao X, Xie Y, Fauquet CM, Zhou X. A DNAβ associated with Tomato yellow leaf curl China virus is required for symptom induction. J Virol. 2004;78:13966–74.
47. Saunders K, Norman A, Gucciardo S, Stanley J. The DNA beta satellite component associated with ageratum yellow vein disease includes an essential pathogenicity protein (beta C1). Virology. 2004;324:37–47.

48. Bhat S, Folimonova SY, Cole AB, Ballard KD, Lez I, Watson BS, Summer LW, Nelson RS. Influence of host chloroplast proteins on Tobacco mosaic virus accumulation and intercellular movement. Plant Physiol. 2013;161:154–47.

49. Mochizuki T, Ogata Y, Hirata Y, Ohto S. Quantitative transcriptional changes associated with chlorosis severity in mosaic leaves of tobacco plants infected with Cucumber mosaic virus. Mol Plant Pathol. 2014;15:242–54.

50. Xu Y, Zhou XP. Role of rice stripe virus NSvc4 in cell-to-cell movement and symptom development in Nicotiana benthamiana. Front Plant Sci. 2012;3:269.

51. Lagrimini LM, Burkhart W, Moyer M, Rothstein S. Molecular cloning of proteins in infected plants. Annu Rev Phytopathol. 2006;44:135–95.

52. Zhang Y, Zhang ZL, Jiang CH, Chang AX, Yang AG, Luo CG, Wang SM, Wang YY. Cloning and expression analysis of a pathogenesis related protein in infected with TMV. J Gen Virol. 1980;47:79–88.

53. Melchers LS, Apotheker-De Groot M, Van Der Knaap JA, Ponstein AP, Sela-Meijer G, Chini A, Fonseca S, Fernández G, Adie B, Chico JM, Lorenzo O, García-Casado G, López-Ortiz I, Lozano FM, Ponce MR. The JAZ family of repressors is the missing link in jasmonate signalling. Nature. 2007;448:666–70.

54. Chini A, Fonseca S, Fernández G, Adie B, Chico JM, Lorenzo O, García-Casado G, López-Ortiz I, Lozano FM, Ponce MR. The JAZ family of repressors is the missing link in jasmonate signalling. Nature. 2007;448:666–70.

55. Ryals JA, Neuenschwander UH, Willits MG, Molina A, Jones JD, Boller T, Felix G. Perception of the bacterial PAMP EF-Tu by the receptor EFR restricts Agrobacterium-mediated transformation. Cell. 2006;125:749.

56. Albrecht C, Rusinova E, Hecht V, Bajajens E, De Vries S. The Arabidopsis thaliana SOMATIC EMBRYOGENESIS RECEPTOR-LIKE KINASE51 and 2 control male gametogenesis. Plant Cell. 2005;17:3337–49.

57. Gómez-Gómez L, Boller T. FLS2: an LRR receptor-like kinase involved in the perception of the bacterial elicitor flagellin in Arabidopsis. Mol Cell. 2005;20:1003.

58. Pritsch C, Kunze G, Chinchilla D, Ciancia A, Jones JD, Boller T, Felix G. Identification of a putative receptor kinase that controls shoot and floral meristem size in Arabidopsis. Cell. 1997;89:575–85.

59. Saunders K, Norman A, Gucciardo S, Stanley J. The DNA beta satellite component associated with ageratum yellow vein disease includes an essential pathogenicity protein (beta C1). Virology. 2004;324:37–47.

60. Schoof H, Lenhard M, Haecker A, Mayer KF, Jürgens G, Laux T. The stem cell population of Arabidopsis shoot meristems in maintained by a regulatory loop between the CLAVATA and WUSCHEL genes. Cell. 2000;100:635.

61. Agusti J, Lichtenberger R, Schwarz M, Nehlin L, Greb T. Characterization of transcriptome remodeling during cambium formation identifies MX1 and RUL1 as opposing regulators of secondary growth. PLoS Genet. 2011;7:e1001312.

62. Albrecht C, Rusinova E, Hecht V, Bajajens E, De Vries S. The Arabidopsis thaliana SOMATIC EMBRYOGENESIS RECEPTOR-LIKE KINASE51 and 2 control male gametogenesis. Plant Cell. 2005;17:3337–49.

63. Yusuf MA, Chakraborty S. A geminivirus betasatellite damages the structural chloroplast and suppresses defenses. J Exp Bot. 2015;66:5881–90.

64. Zhang Y, Zhang ZL, Jiang CH, Chang AX, Yang AG, Luo CG, Wang SM, Wang YY. Cloning and expression analysis of a pathogenesis related protein in infected with TMV. J Gen Virol. 1980;47:79–88.

65. Bruyne LD, Hofte M. Making sense of hormone-mediated defense networking: from rice to Arabidopsis. Front Plant Sci. 2014;5:611.

66. Santos AA, Lopes RV, Afpita JA, Fontes EP. NGR-interacting kinase, NIK, a transducer of plant defence signalling. J Exp Bot. 2010;61:1389.

67. Kim TW, Wang ZY. Brassinosteroid signal transduction from receptor kinases to transcription factors. Annu Rev Plant Biol. 2010;61:681.

68. De Vleesschauwer D, Xu J, Hoffer M. Making sense of hormone-mediated defense networking: from rice to Arabidopsis. Front Plant Sci. 2014;5:611.

69. Santos AA, Lopes RV, Afpita JA, Fontes EP. NGR-interacting kinase, NIK, a transducer of plant defence signalling. J Exp Bot. 2010;61:1389.

70. Clouse SD. Brassinosteroid signal transduction: from receptor kinase activation to transcriptional networks regulating plant development. Plant Cell. 2011;23:1219–30.

71. Wang WF, Bai MY, Wang ZY. The brassinosteroid signaling network - a paradigm of signal integration. Curr Opin Plant Biol. 2014;21:147.

72. Chini A, Fonseca S, Fernández G, Adie B, Chico JM, Lorenzo O, García-Casado G, López-Ortiz I, Lozano FM, Ponce MR. The JAZ family of repressors is the missing link in jasmonate signalling. Nature. 2007;448:666–70.

73. Schumacher K, Chory J. Brassinosteroid signal transduction: still casting the paradigm of signal integration. Curr Opin Plant Biol. 2014;21:147.

74. Li J, Chory J. A putative leucine-rich repeat receptor kinase involved in brassinosteroid signal transduction. Cell. 1997;90:928–39.

75. Kim H, Park D, Hahn Y. Identification of novel RNA viruses in alfalfa (Medicago sativa) an alphafratipivirus, a Deltafratipivirus, and a Marafivirus. Gene. 2017;638:57–9.

76. Chini A, Fonseca S, Fernández G, Adie B, Chico JM, Lorenzo O, García-Casado G, López-Ortiz I, Lozano FM, Ponce MR. The JAZ family of repressors is the missing link in jasmonate signalling. Nature. 2007;448:666–70.

77. Ryu CM, Murphy JF, Mysore KS, Klopper JW. Plant growth-promoting rhizobacteria systemically protect Arabidopsis thaliana against Cucumber mosaic virus by a salicylic acid and NPR1-independent and jasmonic acid-dependent signaling pathway. Plant J. 2004;39:381–92.