Early transcriptome analysis of the brown streak virus–cassava pathosystem provides molecular insights into virus susceptibility and resistance

Ravi B. Anjanappa[1]*, Devang Mehta[1]*, Michal J. Okoniewski[2,3], Alicja Szabelska[3,4], Wilhelm Gruissem[1] and Hervé Vanderschuren[1,5]

[1]Department of Biology, ETH Zurich, Switzerland; [2] ID Scientific IT Services, ETH Zurich, Switzerland; [3]Functional Genomics Center Zurich, Zurich, Switzerland; [4]Department of Mathematical and Statistical Methods, Poznan University of Life Sciences, Poland; [5]AGROBIOCHEM Department, Gembloux Agro-Bio Tech, University of Liège, Belgium

*These authors contributed equally to this work.

Author for correspondence:

Hervé Vanderschuren

Tel: +32 81 62 25 71/ +41 44 632 87 25

Email: herve.vanderschuren@ulg.ac.be/ hvanderschuren@ethz.ch

Running Title: RNA-seq analysis of the cassava-CBSV pathosystem

SUMMARY

Cassava brown streak viruses (CBSVs) are responsible for significant cassava yield losses in eastern sub–Saharan Africa. In the present work, we inoculated CBSV–susceptible and –resistant cassava varieties with a mixed infection of CBSVs using top-cleft grafting. Virus titres in grafted scions were monitored in a time course experiment in both varieties. We performed RNA-seq of the two cassava varieties at the earliest time-point of full infection in the susceptible scions. Genes encoding proteins in RNA silencing and salicylic acid pathways were regulated in the susceptible cassava variety but transcriptional changes were limited in the resistant variety. After infection, genes related to callose deposition at plasmodesmata were regulated and callose deposition was significantly reduced in the susceptible cassava variety. We also show that β–1,3–glucanase enzymatic activity is differentially regulated in the susceptible and resistant varieties. The differences in transcriptional responses to CBSV infection indicate that resistance involves callose deposition at plasmodesmata but does not trigger typical anti-viral defence responses. A meta-analysis of the current RNA-seq dataset and selected, previously reported, host–potyvirus and virus-cassava RNA-seq datasets revealed comparable host responses across pathosystems only at similar time points after infection or infection of a common host.

Key words: cassava, potyvirus, RNA-Seq, transcriptome, virus resistance, salicylic acid, RDR1, callose

HIGHLIGHT

Our results suggest that resistance to CBSV in cassava involves callose deposition at the plasmodesmata and our meta-analysis of multiple virus-crop RNA-seq studies suggests that conserved responses across different host-virus systems are limited and depend greatly on time after infection.

INTRODUCTION

Cassava brown streak disease (CBSD) is one of the most damaging cassava virus diseases in Africa causing harvest losses of up to 70% in susceptible varieties (Hillocks *et al.*, 2001; Legg *et al.*, 2011). Ongoing CBSD outbreaks are threatening cassava production in sub–Saharan Africa (Legg et al., 2011; Legg et al., 2014) and the recent spread of CBSD into central Africa (Bigirimana *et al.*, 2011; Mulimbi *et al.*, 2012) indicates that it could soon become a major constraint to cassava production in all sub–Saharan cassava growing regions.

The characteristic symptoms of CBSD include leaf chlorosis and dry hard necrosis in roots, thus affecting both the quality and yield of the edible storage roots. CBSD is caused by two phylogenetically distinct cassava brown streak virus species: *Cassava brown streak virus* (CBSV) and *Ugandan cassava brown streak virus* (UCBSVs) (Mbanzibwa *et al.*, 2011), collectively referred to as CBSVs. CBSVs are (+) ssRNA viruses belonging to the genus *Ipomovirus*, family Potyviridae (Monger *et al.*, 2001; Mbanzibwa *et al.*, 2009), and are transmitted by whiteflies (Maruthi *et al.*, 2005). They consist of a ~9 kb genome with a single open reading frame encoding a large poly-protein, which is co- and post–translationally cleaved into ten viral proteins (Winter et al., 2010). CBSV and UCBSV differ in virulence in both controlled and field conditions, with CBSV generally accumulating to higher titres in most host genotypes (Winter *et al.*, 2010; Mohammed *et al.*, 2012; Kaweesi *et al.*, 2014; Ogwok *et al.*, 2015). Two recent field analyses of different cassava genotypes for the presence of CBSV and UCBSV found that both viruses co–occurred in a majority of tested genotypes (Ogwok *et al.*, 2015; Kaweesi *et al.*, 2014).

The release of a cassava reference genome (Prochnik et al., 2012) has opened up new opportunities to perform large–scale characterization of the cassava transcriptome and proteome (Allie et al., 2014; Maruthi et al., 2014; Vanderschuren et al., 2014). Cassava varieties differing in resistance against cassava geminiviruses (CMGs) were recently analysed in a time course experiment to identify pathways that are differentially regulated during South African cassava mosaic virus (SACMV) infection (Allie et al., 2014). Similarly, a late (one year post–infection) time point comparative transcriptome analysis was reported of transcripts differentially expressed in a

susceptible variety (Albert) and a variety (Kaleso) classified as moderately resistant but displaying mild CBSD symptoms after CBSV infection (Maruthi et al., 2014). The late time point study found no regulation in cassava homologs of known R-gene analogues, any NBS-LRR genes, nor genes encoding proteins of the anti-viral RNA silencing pathway. We recently characterised KBH 2006/18 and KBH 2006/26, which are two resistant elite breeding lines with immunity to a mixed infection of CBSV and UCBSV (Anjanappa et al., 2016). Interestingly, KBH 2006/18 and KBH 2006/26 show no symptoms or hypersensitive response upon infection, even at 16-weeks post graft-inoculation (Anjanappa et al., 2016).

Here we used two cassava varieties contrasting for CBSV resistance (i.e. 60444 and KBH2006/18) to characterize the transcriptome response at an early time-point of CBSV and UCBSV co-infection. We report an in-depth characterisation of pathways and genes regulated in compatible and incompatible cassava–ipomovirus interactions and present biochemical evidence to explain the inhibition of virus intercellular movement in the resistant breeding line KBH 2006/18. Our data provide new insights into the cassava-CBSV interaction and identify commonalities in plant responses to virus infection.

METHODS

Plant material and virus inoculation

The elite CBSV-resistant cassava breeding variety KBH 2006/18 was obtained from IITA (Tanzania) and the susceptible variety 60444 from the ETH cassava germplasm collection. Cassava plants were grown under greenhouse conditions (27°C, 16h light, 60% humidity). Infected 60444 stem cuttings with mixed infection of the CBSV isolate TAZ-DES-01 (KF878104) and the UCBSV isolate TAZ-DES-02 (KF878103) were used to inoculate KBH 2006/18 and 60444 plants using a previously described top-cleft grafting method (Moreno *et al.*, 2011; Anjanappa *et al.*, 2016). KBH 2006/18 and 60444 plants were also top-cleft grafted onto virus-free 60444 rootstocks as mock controls. Individual

leaves (the second leaf from the point of graft union) were collected from scions of CBSV- and mock-infected plants at 28 days after grafting (dag) and at 60 dag (Fig. S1).

RT-qPCR

RNA was extracted from cassava leaves using a modified CTAB RNA extraction protocol (Moreno *et al.*, 2011). One µg of DNaseI-treated total RNA was reverse-transcribed using the RevertAid First Strand cDNA Synthesis Kit (Thermo-Fisher) according to the manufacturer's instructions. RT-qPCR was performed with SYBR Green dye (Thermo-Fisher) using a 7500 Fast Real Time PCR System (Applied Biosystems). Three independent biological replicates were analysed for each time-point. The relative expression level of individual genes in each sample was calculated as previously described (Moreno *et al.*, 2011; Vanderschuren *et al.*, 2012). Primers used to validate differentially expressed genes and to quantitate virus titres are listed in Table S1.

Illumina RNA-Seq and data analysis

Library preparation and sequencing were performed as follows: 5 µg of total RNA extract was treated with DNase according to manufacturer instructions (Qiagen, RNase-free DNase Kit) and subsequently column purified using the RNeasy Plant Mini Kit (Qiagen). RNA quality was assessed using a Qubit® (1.0) Fluorometer (Thermo-Fisher) as well as a Bioanalyzer 2100 (Agilent). Samples with a RNA integrity number (RIN) above 6 were used for sequencing. One µg of total RNA was polyA enriched and mRNA libraries were synthesized using a TruSeq RNA Sample Prep Kit v2 (Illumina). Cluster generation was done using 10 pM of pooled normalized libraries on the cBOT with TruSeq PE Cluster Kit v3–cBot–HS (Illumina) and subsequently Illumina HiSeq 2000 sequencing was performed to generate the reads.

Sequence assembly, mapping to the cassava genome and identification of differentially expressed genes (DEGs)

The raw read files were trimmed using the trim sequence tool in the CLC Genomics Workbench v8.5 and settings declared in Notes S1. Unique trimmed reads were mapped to the *M. esculenta*

reference genome AM560-2 v6.0 available on Phytozome (www.phytozome.jgi.doe.gov) using the transcriptomics toolbox in the CLC Genomics Workbench v8.5. Empirical analysis of differentially expressed reads was performed using default parameters to identify differentially expressed genes (DEGs) between control non-infected and CBSVs-infected scions. DEGs were then filtered based on a FDR-corrected P-value of <0.01 and a fold–change of ≥2. DEGs were annotated in the CLC Workbench using the Phytozome GFF3 annotation file (www.phytozome.jgi.doe.gov).

Identification of de novo assembled transcripts (dnATs)

Read mapping of trimmed reads was performed using the transcriptomics toolbox in the CLC Genomics Workbench v8.5 with settings described in Notes S1 but with "Maximum number of hits for a read = 10", allowing for more stringent detection of unmapped reads. Unmapped reads for each cassava variety were pooled and assembled using Trinity (Grabherr *et al.*, 2011) with default settings. Assembled contigs were then size-filtered and contigs greater than 1000 nt were classified as *de novo* assembled transcripts (*dn*ATs). Differential expression of *dn*ATs between infected and mock-infected samples was determined using STAR (Dobin *et al.*, 2013).

Determination of full-length virus genomes from RNA-Seq analysis

Unmapped reads (described previously) were assembled into contigs using the CLC Genomics Workbench *de novo* assembly tool. The assembled contigs were BLASTed against published CBSV (HG965221.1) and UCBSV (HG965222.1) genomes to identify full length virus genomes. Two full length genomes corresponding to CBSV (TAZ-DES-01; KF878104) and UCBSV (TAZ-DES-02; KF878103) were obtained as single assembled contigs. The two full-length viral genomes were used to map reads from each individual sample (Fig. S1) using CLC Genomics Workbench read mapping to contigs tool in order to estimate virus read counts per sample and read count distribution across the virus genomes.

β–1,3–glucanase enzymatic assay

Leaf samples from three independent cassava scions were collected and soluble proteins were extracted from 100 mg of leaf powder in 300 µl of extraction buffer (20 mM HEPES, pH 8.0, 5 mM $MgCl_2$ and 1 tablet/50 ml buffer of EDTA–free protease inhibitor) and quantitated using the BCA assay. The amount of glucose content released by the activity of β-1,3-glucanase in a known amount of soluble proteins in presence of the substrate laminarin (Laminaria digitata, Sigma–Aldrich) was then measured spectrophotometrically at 540 nm (Ramada *et al.*, 2010). β-1,3-glucanase specific activity was subsequently calculated, with 1 unit (U) of activity representing the amount of enzyme required to produce 1 µmol of reducing sugar per minute.

Quantification of plasmodesmata–associated callose

Leaf sections from uninfected and infected scions were stained overnight with Aniline blue (Merck) using a previously described method (Guenoune–Gelbart *et al.*, 2008). Stained leaf samples were then examined using a Zeiss LSM 780 laser confocal microscope with a 40X water–immersion objective. Callose accumulation was detected at an excitation of 405 nm and emission between 413 and 563nm was recorded. Callose quantification and deposit counting was performed as described by Zavaliev & Epel (2015) for eight replicate images per plant variety and treatment condition.

Pfam-based meta-analysis

For Pfam-based meta-analysis, RNA-Seq DEG data were downloaded from published reports (Allie et al., 2014; Rubio et al., 2015 a,b; Maruthi et al., 2014) and filtered using an absolute fold-change of ≥2. Pfam annotations for genomes of each species were obtained from Phytozome annotation files (www.phytozome.jgi.doe.gov). Commonly regulated Pfams were determined by constructing a 5-set Venn diagram and the resulting Pfam overlaps between pairs of sets were used to calculate a pairwise similarity matrix.

RESULTS:

CBSV inoculation induces greater transcriptome modulation in susceptible cassava

166 Following successful grafting of 60444 and KBH2006/18 virus–free scions onto rootstocks with a
167 mixed infection of CBSV (TAZ–DES–01) and UCBSV (TAZ–DES–02) (CBSV-infection in short), we
168 performed a time course sampling of the scions at 16, 22 and 28 days after grafting (dag) (Fig. S1).
169 We assessed virus titres in inoculated and mock scions at the selected time points. Virus infection
170 reached homogenous titre levels across the three biological replicates only at 28 dag in inoculated
171 60444 scions (Fig. S2). No virus was detected in KBH 2006/18 scions (Fig. S2), consistent with
172 previous results (Anjanappa *et al.*, 2016). Notably, we did not observe CBSD symptoms in any
173 inoculated plant at 28 dag. Samples from 60444 and KBH2006/18 scions at 28 dag were selected for
174 transcriptome analysis.

175 Gene expression in 60444 and KBH 2006/18 from non-inoculated and CBSV-inoculated scions
176 collected at 28 dag was analysed using Illumina HiSeq 2000 sequencing. Total RNA from the second
177 emergent leaf immediately after the graft-union was used for RNA-Seq analysis, with three
178 independent replicates per condition and plant variety. A total of 1,254,619,280 reads with 100-bp
179 paired-ends were generated from the 12 samples, ranging from 77 to 121 million reads per sample.
180 On average, 93.4% of the total reads were successfully mapped to the cassava reference genome
181 (*M. esculenta* AM560-2 v6.0). The reads not mapped to the reference genome were assembled *de*
182 *novo* using Trinity (Grabherr *et al.*, 2011). The size-selected Trinity-assembled contigs, referred to as
183 *de novo* assembled transcripts (*dn*ATs), were subsequently blasted against the NCBI nucleotide
184 collection to provide gene descriptions. The *dn*ATs were then used for differential gene expression
185 analysis using STAR (Dobin *et al.*, 2013). None of the *dn*ATs appeared to be differentially regulated
186 in compatible and incompatible cassava-CBSV pathosystems (Table S2).

187 A total of 1,292 genes in 60444 and 585 genes in KBH 2006/18 were differentially expressed at 28
188 dag, reflecting a higher number of modulated genes in the compatible host-virus interaction (Table
189 S3). Over 68% of all differentially expressed genes (DEGs) were up-regulated in variety 60444 while
190 only 30% were up-regulated in KBH 2006/18. A total of 158 DEGs were common between both
191 varieties, and a major proportion (99 DEGs) of these were down-regulated (Fig. 1a). Among the

192 DEGs found in both plant varieties, only 12 genes were differently regulated (i.e. up-regulated in
193 60444 and down-regulated in KBH 2006/18 or vice-versa) (Fig.1a).

194 **Comparison and validation of RNA-Seq data**

195 The RNA-Seq expression results of three significantly regulated genes (*PLASMODESMATA*
196 *LOCATED PROTEIN 1(PDLP1), CALRETICULIN 1B (CRT1B), HEAT SHOCK PROTEIN 17.6 (HSP17.6))*
197 and one non-significantly regulated gene (*CALRETICULIN 3 (CRT3))*, were confirmed using RT-qPCR
198 on the same RNA samples used for sequencing. Fold-change values (comparing infected and non-
199 infected samples) for these RNAs were similar to those observed in the RNA-Seq analysis (Fig. S3).
200 We used unmapped reads to *de novo* assemble the full-length virus genomes of CBSV (TAZ–DES–
201 01) and UCBSV (TAZ–DES–02). Full-length virus genome contigs were used for read counts to
202 estimate virus abundance in each sample used for RNA-Seq (Fig. S4). Read counts detected for
203 CBSV (TAZ–DES–01) and UCBSV (TAZ–DES–02) indicated that CBSV (TAZ–DES–01) was more
204 prevalent than UCBSV (TAZ–DES–02) in two of the three inoculated 60444 scions. The read count
205 analysis confirmed the absence of both CBSVs in the inoculated KBH2006/18 scions (Table S4).

206 **Gene sets enriched in resistant and susceptible cassava**

207 We performed Gene Set Enrichment Analysis using the PlantGSEA toolkit (Yi *et al.*, 2013)) to
208 identify significantly enriched gene ontology (GO) categories and KEGG pathways (Table S4).
209 Significantly enriched GO categories (Biological Process) were then summarized and visualized
210 based on semantic similarity using REVIGO (Supek *et al.*, 2011) (Fig. 1b). As expected, genes related
211 to stimulus response, immune system process and photosynthesis were significantly regulated in
212 both the susceptible and resistant plant varieties (Fig. 1b).

213 We also examined KEGG-annotated pathways enriched after virus infection in both varieties by
214 mapping DEGs onto enriched KEGG pathways using Pathview (Luo and Brouwer, 2013) (Table S4).
215 In both resistant and susceptible varieties, genes in photosystem I and II as well as the light
216 harvesting complex are down–regulated. Pathways related to phenylalanine metabolism,

217 phenylpropanoid biosynthesis and downstream products of phenylpropanoid biosynthesis were
218 enriched in the infected 60444 scions along with several genes involved in plant-pathogen
219 interactions. In KBH 2006/18, the cyanoamino metabolism, starch and sucrose metabolism
220 pathways were enriched. Interestingly, the phenylpropanoid pathway was also overrepresented in
221 KBH 2006/18, but it showed an opposite response as compared to the infected 60444 variety.

222 A recent transcriptome study identified the downregulation of nuclear genes encoding chloroplast
223 proteins (NGCPs) as a key response to PAMP (specifically, bacterial pathogen) perception (Shi *et al.*,
224 2015). Considering the overrepresentation of chloroplast-related genes (TAIR GO:0009507) in our
225 data sets, we further investigated the relative abundance of NGCP transcripts regulated by CBSV
226 infection. In 60444, 7% of all cassava NGCPs (193 out of 2682 cassava NGCPs), which represent 15%
227 of the total DEGs, were regulated upon infection. In KBH 2006/18, 19% of all DEGs are NGCPs (Fig.
228 1c). In both varieties, a majority of NGCPs were down-regulated (Table S4), indicating that CBSV
229 inoculation induces a chloroplastic response, as observed with other PAMPs in Arabidopsis (Göhre *et*
230 *al.*, 2012; Zheng *et al.*, 2012; de Torres Zabala *et al.*, 2015).

231 **CBSV infection modulates callose deposition at plasmodesmata**

232 In a previous study investigating the cassava-CBSV pathosystem we found that KBH 2006/18 allows
233 transmission of CBSVs through the stem vasculature while no virus can be detected in KBH 2006/18
234 leaves (Anjanappa *et al.*, 2016). This suggests that the resistance of KBH 2006/18 to CBSVs involves
235 at least restriction of inter-cellular virus movement from vascular tissues to mesophyll cells. Here we
236 find that transcript levels of genes involved in virus movement and callose deposition at
237 plasmodesmata were regulated in infected 60444 but not KBH 2006/18 leaves. The transcript level
238 of *β–1,3–GLUCANASE (BG3)* encoding an enzyme which degrades callose at the plasmodesmata
239 (Iglesias and Meins Jr, 2000), was 9.7-fold up-regulated in 60444 (Fig. 2a). ANKYRIN REPEAT
240 FAMILY PROTEINS (ANKs) allow viral movement through plasmodesmata by promoting callose
241 degradation (Ueki *et al.*, 2010). We also found an upregulation of two different cassava *ANKs*,
242 Manes.11G035500 (3.8-fold) in 60444 and Manes.05G056700 (2.03-fold) in KBH 2006/18. The

243 expression of *PLASMODEMATA–LOCATED PROTEIN 1 (PDLP1)*, which is a positive regulator of virus
244 cell-to-cell movement, was upregulated in infected 60444. PDLP1 binds to viral movement proteins
245 and helps the formation of viral movement tubules at the plasmodesmata (Amari *et al.*, 2010),
246 thereby facilitating virus cell-to-cell movement. In contrast, the expression of a possible negative
247 regulator of plasmodesmata permeability, *GLUCAN SYNTHASE LIKE 4 (GSL4)* (Maeda *et al.*, 2014),
248 was also up–regulated by 2.6–fold in 60444 (Fig. 2a).

249 We further characterized *BG3* expression levels at 60 dag to follow the regulation of this gene in
250 infected and symptomatic 60444 leaves. Up-regulation of *BG3* expression was attenuated at 60 dag
251 as the CBSV infection progressed. Consistently, no significant change of *BG3* expression was
252 observed in KBH 2006/18 leaves at 60 dag (Fig. 2b). The higher *BG3* transcript levels resulted in a
253 38% increase in BG3 enzymatic activity in infected 60444 leaves (Fig. 2b).

254 In order to understand the functional implications of the simultaneous upregulation of positive
255 (*PDLP1, GSL4*) and negative (*BG3, ANKs*) regulators of callose deposition upon CBSV infection, we
256 stained callose in leaf tissue from 60444 and KBH 2006/18 at 60 days after grafting (Fig. 2c). This
257 revealed a significant decrease in plasmodesmata-associated callose deposition in epidermal cells of
258 infected 60444 leaves and a strong increase in the amount of plasmodesmata–associated callose
259 deposition in KBH 2006/18.

260 **Salicylic acid defence and lignin biosynthesis pathways are regulated in response to**
261 **CBSVs**

262 Overall, we found that in the susceptible cassava variety 60444 defence pathways were more
263 extensively regulated during CBSV infection as compared to the resistant variety KBH 2006/18.
264 Transcript levels for several genes encoding the enzymes involved in lignin and SA biosynthesis were
265 up-regulated in the compatible cassava-CBSV interaction and not regulated or down–regulated in
266 the incompatible interaction (Fig. 3). Both SA and lignin biosynthetic pathways follow the
267 phenylpropanoid pathway, which was enriched in our initial KEGG pathway gene–set analysis (Table
268 S4).

269 SA biosynthesis in plants begins in the chloroplast with the shikimate pathway for synthesis of
270 aromatic amino acids, which in turn is fed by fructose–1,6–bisphosphate (Fig. 3). Although fructose-
271 1,6-bisphosphate is probably not a limiting precursor for SA biosynthesis, we did observe a 4.6-fold
272 up-regulation of the PHOSPHOFRUCTOKINASE 3, a gene involved in the conversion of fructose-6-
273 phosphate to fructose-1,6-bisphosphate in the susceptible variety. We also observed a 2.4-fold up-
274 regulation of the *SHIKIMATE KINASE 3*, responsible for the conversion of shikimate to shikimate–3–
275 phosphate. Previous genetic studies in Arabidopsis and *Nicotiana benthamiana* (reviewed by Chen *et*
276 *al.*, 2009) have indicated that pathogen-induced SA synthesis is catalysed by ISOCHORISMATE
277 SYNTHASE (ICS) to convert chorismate to isochorismate, which is then used to produce SA by an
278 unknown mechanism. An alternative pathway for SA biosynthesis from chorismate involves its
279 conversion to arogenate, which is the substrate of AROGENATE DEHYDRATASE (ADT) to produce
280 phenylalanine. *ADT1* and ADT6 were upregulated in 60444 after CBSV infection, suggesting that
281 virus-induced SA synthesis in cassava, unlike in *N. benthamiana* (Catinot *et al.*, 2008), might proceed
282 through the phenylalanine pathway rather than via isochorismate. PHENYLALANINE AMMONIA-
283 LYASE (PAL), which catalyses the conversion of phenylalanine to cinnamic acid, is a key enzyme in
284 the synthesis of both lignin and SA. Indeed, quadruple *pal* mutants with only 10% PAL activity
285 produced about 50% less SA and were more susceptible to biotic stress (Huang *et al.*, 2010). We find
286 that *PAL1* (Manes.10G047500) and *PAL2* (Manes.07G098700) were 4.4-fold and 2.7-fold up-
287 regulated, respectively, in 60444. In contrast, *PAL1* was downregulated 2.3–fold in KBH 2006/18.
288 Cinnamic acid is the substrate at the branch point of the SA and lignin biosynthesis pathways.
289 Ligation of cinnamic acid to Coenzyme A by 4-COUMARATE:COA LIGASE (4CL) leads to SA
290 biosynthesis via both peroxisomal β-oxidative and cytosolic non-oxidative pathways. The conversion
291 of cinnamic acid first to p-coumaric acid by CINNAMATE–4–HYDROXYLASE (C4H) and then ligation
292 of Coenzyme A to p-coumaric acid by 4CL produces p-coumaroyl–CoA, which is a phenylpropanoid
293 pathway precursor (Widhalm and Dudareva, 2015). In our study, *4CL1* (Manes.14G151400,
294 Manes.04G095300), *4CL2* (Manes.11G071800), *4CL3* (Manes.09G127000.1) and *C4H*
295 (Manes.02G227200) were all up-regulated in CBSV-infected 60444 (Fig. 3).

296 SA is a key regulator of plant defence responses and has been implicated in various mechanisms
297 resulting in the hypersensitive response (HR) and systemic acquired resistance (SAR) (reviewed in
298 Vlot *et al.*, 2009). One such mechanism is the NPR1-dependent response (Vlot *et al.*, 2009), which
299 involves the SA-induced activation and nuclear localisation of NPR1 to activate the transcription of
300 genes for pathogenesis-related proteins, including BG3. NPR1 interacts with various members of
301 the TGA family of transcription factors, which together activate the expression of WKRY
302 transcription factors (Vlot *et al.*, 2009). We find that *TGA9* (Manes.04G004100) expression was up-
303 regulated by 2.8-fold in CBSV-infected 60444, together with the upregulation of two cassava
304 *WRKY70* genes (Manes.07G142400 by 4.5-fold and Manes.10G002200 by 2.5-fold). In Arabidopsis,
305 WRKY70 is an SA-induced transcription factor that activates the expression of several pathogenesis-
306 related genes, including genes encoding a peroxidase and senescence-associated proteins (*SAG12*,
307 *SAG21*, *SAG24*, *SAG29*)(Li *et al.*, 2004; Ülker *et al.*, 2007). In CBSV-infected 60444, *SAG12*
308 (Manes.16G038400) was upregulated 3.6-fold (Fig. 3, Table S3). Similarly, the genes encoding 10 out
309 of 14 Class III peroxidases (Prxs, or PR-9 subfamily), which are involved in SA-induced pathogen
310 defence responses (Almagro *et al.*, 2009), were upregulated (Fig. 3).

311 Lignins are polymers of aromatic compounds and integral components of the plant cell wall. They
312 have a role in defence against microbial and fungal pathogens (Vanholme *et al.*, 2010; Miedes *et al.*,
313 2014). Lignin synthesis is often induced in plant immune responses (Malinovsky *et al.*, 2014). We find
314 several genes encoding enzymes involved in the production of phenylpropanoid intermediates and
315 their conversion to the three monolignols that were up–regulated after CBSV infection in 60444
316 (Fig. 3, Table S3).

317 **The antiviral RNA silencing pathway is up-regulated in the compatible cassava-CBSVs**
318 **interaction**

319 RNA silencing is a major defence mechanism against plant viruses (Wang *et al.*, 2012), especially
320 against RNA viruses, whose genomes and their replicative forms can be directly targeted by RNA-
321 induced Silencing Complexes (RISCs) and DICER–LIKE (DCL) proteins (Waterhouse and Fusaro,

322 2006). RNA silencing is thought to begin with the synthesis and formation of long dsRNA molecules,
323 which are substrates of DCL proteins in plants (Fig. 4) (Pumplin and Voinnet, 2013). In Arabidopsis,
324 four DCL proteins are involved in RNA silencing, of which DCL2 and DCL4 have overlapping
325 functions in antiviral defence (Garcia–Ruiz *et al.*, 2010). However, DCL4 alone is sufficient for
326 antiviral RNA silencing (Garcia-Ruiz *et al.*, 2010). DCL2 appears to be required for the production of
327 secondary small interfering RNAs (siRNAs) (Parent *et al.*, 2015). We find that the expression of three
328 cassava *DCL2* genes was increased in CBSV-infected 60444 (Manes.12G002700 by 2.4-fold,
329 Manes.12G002800 by 3.3-fold, Manes.12G003000 by 3.3-fold).

330 ARGONAUTE (AGO) are essential proteins of the RNA Induced Silencing Complexes (RISCs), which
331 are involved in the sequence-specific silencing of genes by RNA degradation, inhibition of
332 translation and epigenetic modifications. Of the 10 AGO proteins known in plants, AGO2 has a
333 broad and likely specific role in antiviral silencing (Jaubert *et al.*, 2011; Carbonell *et al.*, 2012;
334 Scholthof *et al.*, 2011). Expression of the cassava *AGO2* gene (Manes.07G023100) was 17.5-fold up-
335 regulated in CBSV-infected 60444 scions. A gene expressing AGO5 (Manes.04G011400), which is
336 another but less well characterised AGO protein, was also up–regulated 2.3-fold. The virus-induced
337 increased expression of *AGO2* and *AGO5* in cassava is similar to *N. benthamiana*, in which Potato
338 Virus X infection increases *AGO5* expression and AGO5 acts synergistically with AGO2 to suppress
339 virus infection to a greater degree than the other AGO proteins (Brosseau and Moffett, 2015).

340 RNA-DEPENDENT RNA POLYMERASE 1 (RDR1) is an important component of the post-
341 transcriptional gene silencing (PTGS) immune response to virus infections and is involved in the
342 defence of various classes of viruses, including potyviruses. (Rakhshandehroo *et al.*, 2009; Garcia–
343 Ruiz *et al.*, 2010; Wang *et al.*, 2010; Lee *et al.*, 2016). RDRs catalyse the synthesis of the antisense
344 strand in the conversion of viral RNA to a long dsRNA template that can be processed by DCLs.
345 PTGS is often initiated by siRNAs binding to a target and recruiting RDRs, which then results in the
346 production of further long dsRNA templates for the production of secondary siRNAs. This
347 mechanism, also known as transitivity or silencing amplification, is largely facilitated by RDR6, and

348 possibly also RDR1 (Wang *et al.*, 2010). We find that expression of *RDR1* (Manes.17G084800) was
349 highly up-regulated (38.5-fold) in CBSV-infected 60444 scions (Fig. 4). The steep up-regulation of
350 *RDR1* was sustained at all three early time points (Fig. S5). The precise role of RDR1 in virus
351 resistance has not yet been established, and one study in particular (in *N. benthamiana*)
352 demonstrated a repressor function of RDR1 in antiviral silencing (Ying *et al.*, 2010). It is also possible
353 that RDR1 contributes to virus resistance by targeting host genes. Cao et al. (2014) identified 369
354 genes as potential RDR1 targets and validated the *RDR1*-dependent reduction in transcription of
355 *PHOTOSYSTEM II LIGHT HARVESTING COMPLEX B1.3 (LHCB1.3)* gene in response to potyvirus
356 infection. We consistently observed a 3.4-fold downregulation of *LHCB1.3* (Manes.17G066700) in
357 CBSV-infected 60444 scions. However, of the 26 putative RDR1 target genes (out of the 369 genes
358 identified by Cao et al., (2014)) regulated in our study, only 13 were down-regulated (Table S6).

359 The expression of the PTGS-related genes *SILENCING DEFECTIVE 5 (SDE5)* and *SUPPRESSOR OF*
360 *GENE SILENCING 3 (SGS3)* were also significantly up-regulated in CBSV-infected 60444 (Fig. 3, Table
361 S3) but not changed in CBSV-infected KBH 2006/18. *AGO2* and *RDR1,* which are both up-regulated
362 in CBSV-infected 60444, are also induced by SA signalling pathways (Lewsey *et al.*, 2010; Jovel *et al.*,
363 2011; Hunter *et al.*, 2013; Lee *et al.*, 2016), which is consistent with the upregulation of several
364 cassava genes involved in SA synthesis and signalling pathways after CBSV infection.

365 **Similar transcript regulation in selected host-potyvirus and cassava-virus pathosystems**

366 We performed a comparative analysis of the DEGs we detected above with other RNA-Seq studies
367 of host-virus pathosystems to understand similarities and differences in gene expression regulation.
368 The studies we selected involved naturally occurring crop-virus interactions with detectable host
369 responses to potyvirus infections and comparable RNA–Seq data sets (late-stage cassava-CBSV
370 infection, Maruthi et al., 2014; early-stage peach and late-stage apricot infection with Plum Pox
371 Virus, Rubio *et al.*, 2015a,b). We also included an RNA-Seq study of cassava infected with South
372 African Cassava Mosaic Virus, a ssDNA geminivirus pathosystem (Allie *et al.*, 2014), to compare the
373 response of cassava to two different virus types.

374 We first filtered DEGs from all datasets according to our differential expression criteria (fold change
375 >2; FDR <0.01) and used Pfam identifiers (Finn *et al.*, 2015) matching each DEG as a common term to
376 compare gene expression between cassava and the two Prunus species. This allowed us to construct
377 a similarity matrix based on a pairwise similarity percentage score, where 100% denotes complete
378 similarity (Fig. 5a, grey boxes). The greatest pairwise similarity (25.54%) was found between peach
379 and cassava RNA-Seq datasets obtained within two-months post-infection. Interestingly, we found
380 more common Pfam identifiers between our early time-point cassava-CBSV dataset and the early
381 time-point cassava–SACMV dataset rather than between the two cassava-CBSV datasets. Also, the
382 two cassava-CBSV datasets and the two Prunus datasets had a lower degree of similarity than most
383 other pairs, indicating a strong temporal component to the modulation of host gene expression
384 after virus infection. Our meta-analysis did not reveal a strong dependence of host gene expression
385 changes on either RNA or DNA virus or host genus.

386 We next compared up- and down-regulated DEGs separately across the five datasets (Fig. 5b,c;
387 Table S5) and found five Pfam identifiers in common for up-regulated and four for down-regulated
388 DEGs. Of these, three Pfams that contained genes in both the up- and down-regulated data-sets in
389 all pathosystems included PF01190 (pollen protein Ole e I-like), PF00249 (Myb-like DNA-binding
390 domain) and PF03168 (late embryogenesis-abundant protein). While PF01190 represents a largely
391 uncharacterised protein family, the other two identifiers comprise stress response genes. Genes
392 belonging to PF00067 are up–regulated in all pathosystems and this Pfam comprises the
393 cytochrome P450 protein family, which includes members involved in the synthesis of plant defence
394 compounds. Genes encoding members of the GDSL-like lipase family (PF00657) are also up-
395 regulated in all pathosystems and are required for systemic ethylene-mediated resistance (Kwon *et*
396 *al.*, 2009; Kim *et al.*, 2013).

397 In addition to the Pfams common to all RNA-Seq datasets, genes in another three Pfams were found
398 up- and four were down-regulated in all potyvirus infection studies (Fig. 5b,c). Of these, genes

399 belonging to PF02365 (NO APICAL MERISTEM protein) and PF00847 (APETALA 2 domain protein)
400 were found in both the up- and down-regulated data sets.

401 DISCUSSION

402 We show that the transcriptional response of cassava to CBSV infection was significantly stronger in
403 the susceptible 60444 variety than the resistant KBH 2006/18 variety. This is similar to results
404 reported earlier for extreme virus resistance in potato (Goyer *et al.*, 2015). To ensure that our dataset
405 represents the earliest time-point of complete virus infection, we first measured virus titres in the
406 inoculated scions at three time-points until we found consistent infection in all 60444 plants (Fig.
407 S2). This allowed us to analyse changes in gene expression prior to symptom development, and thus
408 early in the infection process. While graft inoculation may delay infection until the graft junction is
409 fully established, we previously demonstrated consistent infection using the top-grafting method
410 (Moreno *et al.*, 2011; Vanderschuren et al., 2012; Anjanappa *et al.*, 2016). Graft virus transmission
411 using field-infected rootstocks or scions was necessary because no infectious CBSV clones are
412 available (Mohammed *et al.*, 2012; Vanderschuren *et al.*, 2012; Wagaba *et al.*, 2013; Anjanappa *et al.*,
413 2016). Although the lack of detectable virus titres in KBH 2006/18 indicates the absence of viral
414 replication in the resistant cultivar, we observed a transcriptional response to graft-inoculation. It is
415 currently unknown if this is due to the local recognition of viral RNAs and/or proteins, which are
416 mobilized in the vascular tissues after the graft was established, or to an extended systemic
417 response transmitted by the CBSV-susceptible rootstock to the CBSV-resistant scion.

418 Based on the DEGs in the susceptible variety, 60444, we found that CBSV infection increases the
419 expression of genes encoding SA synthesis enzymes, which likely increases SA levels that activate
420 SA-regulated genes involved in antiviral responses. Our data suggest that in CBSV-infected cassava
421 SA is synthesized via the phenylalanine rather than the isochorismate pathway. Additionally, we
422 found genes encoding proteins of the DCL2/AGO5/AGO2-mediated anti-viral silencing response
423 were upregulated in CBSV-infected 60444 but not KBH 2006/18. The silencing response may be

424 amplified as indicated by the strong upregulation of the gene encoding RDR1, but it is also possible
425 that RDR1 assists in the repression of RDR6–mediated silencing amplification (Ying et al., 2010).

426 The susceptible 60444 and resistant KBH 2006/18 varieties differed strongly in callose deposition at
427 plasmodesmata during CBSV infection. While DEGs associated with virus movement through
428 plasmodesmata were found only in CBSV-infected 60444, callose deposition at plasmodesmata was
429 reduced in 60444 and increased in KBH 2006/18, which could only be partially explained by the
430 detected DEGs. Nevertheless, the results are consistent with the limitation of CBSV movement in
431 KBH2006/18 from stems into leaves (Anjanappa *et al.*, 2016). An efficacious CBSV infection of
432 susceptible 60444 variety thus involves a virus-induced reduction in callose deposition at the
433 plasmodesmata that is absent in the resistant KBH2006/18 variety.

434 The meta-analysis of our RNA-Seq dataset and other plant-virus RNA-Seq datasets revealed a low
435 level of similarity in host transcriptome responses to different virus types and at different time-
436 points of infection. A previous transcriptome-based meta-analysis of plant-virus interactions
437 compared transcriptional responses in Arabidopsis after infection with different virus species
438 (Rodrigo *et al.*, 2012). Their results indicated a correlation between host transcriptome response and
439 virus phylogeny, with closely related viruses generating transcript regulation of similar host genes.
440 Since our analysis did not involve a model plant, we managed differences in genome annotations in
441 different species by using Pfam identifiers. While Pfam identifiers reduce resolution because they
442 represent protein families instead of specific genes, this approach does not require identification of
443 exact gene orthologs across multiple species. Thus, this allows the comparison of transcriptome
444 regulation between different viruses and plant species. Overall, we found only a small number of
445 common Pfam identifiers, even between different hosts of the same genus or between viruses of the
446 same family. However, based on the numbers of common Pfam identifiers, there was strong
447 temporal impact on gene regulation after virus infection, indicating that earlier transcript-level
448 changes are perhaps most diagnostic to infection by different viruses. Pfam identifiers that were
449 found in common across the selected pathosystems did not relate to classical defence pathways but

rather to protein families implicated in developmental processes. A previous review of expression-profile studies also found commonalities within developmental genes across various compatible pathosystems (Whitham *et al.*, 2006). The common regulation of developmental genes might describe similarities in symptom development or downstream signalling processes rather than in primary immune responses across different host-virus interactions.

Together, our study suggests that analysis of early time-points during virus-host interactions captures the emerging impact of virus replication on host gene expression. For example, our RNA-Seq dataset from scion leaves 28 days after grafting revealed a broad immune response (including SA and RNA silencing pathways) in the susceptible cassava 60444 variety that was not found in susceptible cassava plants after a long-term infection (Maruthi *et al.*, 2014). Our meta-analysis of different RNA-Seq datasets also showed significant differences in transcriptome responses at different times after infection. The absence of a detectable immune response in the resistant KBH 2006/18 variety suggests that even earlier time-points after infection should be analysed, preferably using Agrobacterium with infectious CBSV clones or infections of protoplasts followed by transcriptomics or proteomics. The Pfam-based meta-analysis also identified interesting protein families, such as meristem identity proteins (NAM and Apetala2), which were regulated in different host-potyvirus interactions and at different time points after infection. Understanding the functions of these proteins in host–virus interactions might reveal conserved developmental responses to virus infections across different taxa. Our Pfam-based comparison methodology also makes larger scale comparison of crop RNA–Seq studies possible, with the aim of developing a more comprehensive understanding of conserved responses of non–model plants to different biotic and abiotic stresses.

ACKNOWLEDGEMENTS

The authors would like to thank Asuka Kuwabara for assisting with the microscopy, Kamil Sklodowski for help with the image analysis, and Matthias Hirsch-Hoffman for bioinformatics data

475 management. We also thank Irene Zurkirchen for support in the glasshouse and R. Glen Uhrig for
476 helpful comments on the manuscript. We gratefully acknowledge the support of our research by
477 ETH Zurich. Ravi B. Anjanappa was supported by the Research Fellow Partnership Programme of
478 ETH Global. Devang Mehta is supported by the European Union's Seventh Framework Programme
479 for research, technological development, and demonstration (EU GA–2013–608422–IDP BRIDGES).

480 SUPPORTING INFORMATION

481 Fig. S1: Experimental scheme of the study

482 Fig. S2: CBSV quantitation at 3 different time points after grafting

483 Fig. S3: Validation of RNA-seq data by RT-qPCR

484 Fig. S4: Virus read counting from unmapped RNA-seq reads

485 Fig. S5: RT-qPCR based RDR1 expression across the three timepoints

486 Table S1: Primers used in the study

487 Table S2: Significantly differentially expressed genes

488 Table S3: List of denovo assembled transcripts and their differential expression

489 Table S4: Results of gene set enrichment analysis

490 Table S5: Comparison of multiple virus-host RNA-seq datasets

491 Table S6: DEGs in the study that match putative RDR1 targets identified by Cao et al., 2014.

REFERENCES

Allie F, Pierce EJ, Okoniewski MJ, Rey C. 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.

Almagro L, Gómez Ros L V., Belchi-Navarro S, Bru R, Ros Barceló A, Pedreño MA. 2009. Class III peroxidases in plant defence reactions. Journal of Experimental Botany **60**, 377–390.

Amari K, Boutant E, Hofmann C, *et al.* 2010. A family of plasmodesmal proteins with receptor-like properties for plant viral movement proteins. PLoS Pathogens **6**, 1–10.

Anjanappa RB, Mehta D, Maruthi MN, Kanju E, Gruissem W, Vanderschuren H. 2016. Characterization of brown streak virus-resistant cassava. Molecular Plant-Microbe Interactions **29**, 527–534.

Bigirimana S, Barumbanze P, Ndayihanzamaso P, Shirima R, Legg JP. 2011. First report of cassava brown streak disease and associated Ugandan cassava brown streak virus in Burundi. New Disease Reports **24**, 26.

Brosseau C, Moffett P. 2015. Functional and Genetic Analysis Identify a Role for Arabidopsis ARGONAUTE5 in Antiviral RNA Silencing. The Plant Cell **27**, 1742–54.

Cao M, Du P, Wang X, Yu Y-Q, Qiu Y-H, Li W, Gal-On A, Zhou C, Li Y, Ding S-W. 2014. Virus infection triggers widespread silencing of host genes by a distinct class of endogenous siRNAs in Arabidopsis. Proceedings of the National Academy of Sciences of the United States of America **111**, 14613–8.

Carbonell A, Fahlgren N, Garcia-Ruiz H, Gilbert KB, Montgomery TA, Nguyen T, Cuperus JT, Carrington JC. 2012. Functional analysis of three Arabidopsis ARGONAUTES using slicer-defective mutants. The Plant Cell **24**, 3613–29.

Catinot J, Buchala A, Abou-Mansour E, Métraux J-P. 2008. Salicylic acid production in response to

biotic and abiotic stress depends on isochorismate in Nicotiana benthamiana. FEBS Letters **582**, 473–478.

Chen Z, Zheng Z, Huang J, Lai Z, Fan B. 2009. Biosynthesis of salicylic acid in plants. Plant Signaling & Behavior **4**, 493–496.

Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, Batut P, Chaisson M, Gingeras TR. 2013. STAR: Ultrafast universal RNA-seq aligner. Bioinformatics **29**, 15–21.

Finn RD, Coggill P, Eberhardt RY, *et al.* 2015. The Pfam protein families database: towards a more sustainable future. Nucleic Acids Research **44**, D279–D285.

Garcia-Ruiz H, Takeda A, Chapman EJ, Sullivan CM, Fahlgren N, Brempelis KJ, Carrington JC. 2010. Arabidopsis RNA-dependent RNA polymerases and dicer-like proteins in antiviral defense and small interfering RNA biogenesis during Turnip Mosaic Virus infection. The Plant Cell **22**, 481–96.

Göhre V, Jones AME, Sklenář J, Robatzek S, Weber APM. 2012. Molecular Crosstalk Between PAMP-Triggered Immunity and Photosynthesis. Molecular Plant-Microbe Interactions **25**, 1083–1092.

Goyer A, Hamlin L, Crosslin JM, Buchanan A, Chang JH. 2015. RNA-Seq analysis of resistant and susceptible potato varieties during the early stages of potato virus Y infection. BMC Genomics **16**, 472.

Grabherr MG, Haas BJ, Yassour M, *et al.* 2011. Full-length transcriptome assembly from RNA-Seq data without a reference genome. Nature Biotechnology **29**, 644–52.

Guenoune-Gelbart D, Elbaum M, Sagi G, Levy A, Epel BL. 2008. Tobacco mosaic virus (TMV) replicase and movement protein function synergistically in facilitating TMV spread by lateral diffusion in the plasmodesmal desmotubule of Nicotiana benthamiana. Molecular Plant-Microbe Interactions **21**, 335–45.

Hillocks RJ, Raya MD, Mtunda K, Kiozia H. 2001. Effects of brown streak virus disease on yield and

quality of cassava in Tanzania. Journal of Phytopathology **149**, 389–394.

Huang J, Gu M, Lai Z, Fan B, Shi K, Zhou Y-H, Yu J-Q, Chen Z. 2010. Functional analysis of the Arabidopsis PAL gene family in plant growth, development, and response to environmental stress. Plant Physiology **153**, 1526–1538.

Hunter LJR, Westwood JH, Heath G, Macaulay K, Smith AG, Macfarlane S a, Palukaitis P, Carr JP. 2013. Regulation of RNA-dependent RNA polymerase 1 and isochorismate synthase gene expression in Arabidopsis. PLoS ONE **8**, e66530.

Iglesias VA, Meins Jr F. 2000. Movement of plant viruses is delayed in a b-1,3-glucanase-deficient mutant showing a reduced plasmodesmatal size exclusion limit and enhanced callose deposition. The Plant Journal **21**, 157–166.

Jaubert M, Bhattacharjee S, Mello AFS, Perry KL, Moffett P. 2011. ARGONAUTE2 mediates RNA-silencing antiviral defenses against Potato virus X in Arabidopsis. Plant Physiology **156**, 1556–64.

Jovel J, Walker M, Sanfaçon H. 2011. Salicylic acid-dependent restriction of Tomato ringspot virus spread in tobacco is accompanied by a hypersensitive response, local RNA silencing, and moderate systemic resistance. Molecular Plant-Microbe Interactions **24**, 706–718.

Kaweesi T, Kawuki R, Kyaligonza V, Baguma Y, Tusiime G, Ferguson ME. 2014. Field evaluation of selected cassava genotypes for cassava brown streak disease based on symptom expression and virus load. Virology Journal **11**, 216.

Kim HG, Kwon SJ, Jang YJ, Nam MH, Chung JH, Na Y-C, Guo H, Park OK. 2013. *GDSL Lipase 1 modulates plant immunity through feedback regulation of ethylene signaling*.

Kwon SJ, Jin HC, Lee S, Nam MH, Chung JH, Kwon S Il, Ryu CM, Park OK. 2009. GDSL lipase-like 1 regulates systemic resistance associated with ethylene signaling in Arabidopsis. The Plant Journal **58**, 235–245.

Lee W-S, Fu S-F, Li Z, Murphy AM, Dobson EA, Garland L, Chaluvadi SR, Lewsey MG, Nelson

564 **RS, Carr JP**. 2016. Salicylic acid treatment and expression of an RNA-dependent RNA polymerase 1
565 transgene inhibit lethal symptoms and meristem invasion during tobacco mosaic virus infection in
566 Nicotiana benthamiana. BMC Plant Biology **16**, 15.

567 **Legg JP, Jeremiah SC, Obiero HM, *et al.*** 2011. Comparing the regional epidemiology of the
568 cassava mosaic and cassava brown streak virus pandemics in Africa. Virus Research **159**, 161–170.

569 **Lewsey MG, Murphy AM, Maclean D, *et al.*** 2010. Disruption of two defensive signaling pathways
570 by a viral RNA silencing suppressor. Molecular Plant-Microbe Interactions **23**, 835–45.

571 **Li J, Brader G, Palva ET**. 2004. The WRKY70 Transcription Factor: A Node of Convergence for
572 Jasmonate-Mediated and Salicylate-Mediated Signals in Plant Defense. The Plant Cell **16**, 319–331.

573 **Luo W, Brouwer C**. 2013. Pathview: An R/Bioconductor package for pathway-based data integration
574 and visualization. Bioinformatics **29**, 1830–1831.

575 **Maeda H, Song W, Sage T, Dellapenna D**. 2014. Role of callose synthases in transfer cell wall
576 development in tocopherol deficient Arabidopsis mutants. Frontiers in Plant Science **5**, 46.

577 **Malinovsky FG, Fangel JU, Willats WGT**. 2014. The role of the cell wall in plant immunity. Frontiers
578 in Plant Science **5**, 178.

579 **Maruthi MN, Bouvaine S, Tufan HA, Mohammed IU, Hillocks RJ**. 2014. Transcriptional Response
580 of Virus-Infected Cassava and Identification of Putative Sources of Resistance for Cassava Brown
581 Streak Disease. PLoS ONE **9**, e96642.

582 **Maruthi MN, Hillocks RJ, Mtunda K, Raya MD, Muhanna M, Kiozia H, Rekha AR, Colvin J,**
583 **Thresh JM**. 2005. Transmission of Cassava brown streak virus by Bemisia tabaci (Gennadius).
584 Journal of Phytopathology **153**, 307–312.

585 **Mbanzibwa DR, Tian YP, Tugume a K, *et al.*** 2011. Evolution of cassava brown streak disease-
586 associated viruses. The Journal of General Virology **92**, 974–87.

587 **Mbanzibwa DR, Tian YP, Tugume AK, Mukasa SB, Tairo F, Kyamanywa S, Kullaya A, Valkonen**

588 **JPT**. 2009. Genetically distinct strains of Cassava brown streak virus in the Lake Victoria basin and
589 the Indian Ocean coastal area of East Africa. Archives of Virology **154**, 353–359.

590 **Miedes E, Vanholme R, Boerjan W, Molina A**. 2014. The role of the secondary cell wall in plant
591 resistance to pathogens. Frontiers in Plant Science **5**, 358.

592 **Mohammed IU, Abarshi MM, Muli B, Hillocks RJ, Maruthi MN**. 2012. The symptom and genetic
593 diversity of cassava brown streak viruses infecting cassava in East Africa. Advances in Virology **2012**,
594 1–10.

595 **Monger WA, Seal S, Cotton S, Foster GD**. 2001. Identification of different isolates of Cassava
596 brown streak virus and development of a diagnostic test. Plant Pathology **50**, 768–775.

597 **Moreno I, Gruissem W, Vanderschuren H**. 2011. Reference genes for reliable potyvirus quantitation
598 in cassava and analysis of Cassava brown streak virus load in host varieties. Journal of Virological
599 Methods **177**, 49–54.

600 **Mulimbi W, Phemba X, Assumani B**. 2012. First report of Ugandan cassava brown streak virus on
601 cassava in Democratic Republic of Congo. New Disease Reports **26**, 11.

602 **Ogwok E, Alicai T, Rey MEC, Beyene G, Taylor NJ**. 2015. Distribution and accumulation of cassava
603 brown streak viruses within infected cassava (Manihot esculenta) plants. Plant Pathology **64**, 1235–
604 1246.

605 **Parent JS, Bouteiller N, Elmayan T, Vaucheret H**. 2015. Respective contributions of Arabidopsis
606 DCL2 and DCL4 to RNA silencing. The Plant Journal **81**, 223–232.

607 **Pumplin N, Voinnet O**. 2013. RNA silencing suppression by plant pathogens: defence, counter-
608 defence and counter-counter-defence. Nature Reviews Microbiology **11**, 745–60.

609 **Rakhshandehroo F, Takeshita M, Squires J, Palukaitis P**. 2009. The influence of RNA-dependent
610 RNA polymerase 1 on potato virus Y infection and on other antiviral response genes. Molecular
611 Plant-Microbe Interactions **22**, 1312–1318.

612 **Ramada MH., Lopes FA., Ulhoa CJ, Silva R dN.** . 2010. Optimized microplate β-1,3-glucanase
613 assay system for Trichoderma spp. screening. Journal of Microbiological Methods **81**, 6–10.

614 **Rodrigo G, Carrera J, Ruiz-Ferrer V, del Toro FJ, Llave C, Voinnet O, Elena SF**. 2012. A meta-
615 analysis reveals the commonalities and differences in arabidopsis thaliana response to different viral
616 pathogens. PLoS ONE **7**.

617 **Rubio M, Ballester AR, Olivares PM, De Moura MC, Dicenta F, Martínez-Gómez P**. 2015*a*. Gene
618 expression analysis of plum pox virus (Sharka) susceptibility/resistance in apricot (prunus armeniaca
619 L.). PLoS ONE **10**, 1–16.

620 **Rubio M, Rodriguez-Moreno L, Ballester AR, de Moura MC, Bonghi C, Candresse T, Martínez-**
621 **Gómez P**. 2015*b*. Analysis of gene expression changes in peach leaves in response to Plum pox virus
622 infection using RNA-Seq. Molecular Plant Pathology **16**, 164–176.

623 **Scholthof HB, Alvarado VY, Vega-Arreguin JC, Ciomperlik J, Odokonyero D, Brosseau C,**
624 **Jaubert M, Zamora A, Moffett P**. 2011. Identification of an ARGONAUTE for antiviral RNA silencing
625 in Nicotiana benthamiana. Plant Physiology **156**, 1548–1555.

626 **Shi B, Lin L, Wang S, *et al.*** 2015. Identification and regulation of host genes related to Rice stripe
627 virus symptom production. New Phytologist **209**, 1106–1119.

628 **Supek F, Bosnjak M, Skunca N, Smuc T**. 2011. Revigo summarizes and visualizes long lists of gene
629 ontology terms. PLoS ONE **6**.

630 **de Torres Zabala M, Littlejohn G, Jayaraman S, *et al.*** 2015. Chloroplasts play a central role in plant
631 defence and are targeted by pathogen effectors. Nature Plants **1**, 15074.

632 **Ueki S, Spektor R, Natale DM, Citovsky V**. 2010. ANK, a host cytoplasmic receptor for the Tobacco
633 mosaic virus cell-to-cell movement protein, facilitates intercellular transport through
634 plasmodesmata. PLoS Pathogens **6**.

635 **Ülker B, Shahid Mukhtar M, Somssich IE**. 2007. The WRKY70 transcription factor of Arabidopsis

influences both the plant senescence and defense signaling pathways. Planta **226**, 125–137.

Vanderschuren H, Moreno I, Anjanappa RB, Zainuddin IM, Gruissem W. 2012. Exploiting the combination of natural and genetically engineered resistance to cassava mosaic and cassava brown streak viruses impacting cassava production in Africa. PLoS ONE **7**, e45277.

Vanholme R, Demedts B, Morreel K, Ralph J, Boerjan W. 2010. Lignin biosynthesis and structure. Plant Physiology **153**, 895–905.

Vlot AC, Dempsey DA, Klessig DF. 2009. Salicylic Acid, a Multifaceted Hormone to Combat Disease. Annual Review of Phytopathology **47**, 177–206.

Wagaba H, Beyene G, Trembley C, Alicai T, Fauquet CM, Taylor NJ. 2013. Efficient transmission of cassava brown streak disease viral pathogens by chip bud grafting. BMC Research Notes **6**, 516.

Wang M-B, Masuta C, Smith N a., Shimura H. 2012. RNA Silencing and Plant Viral Diseases. Molecular Plant-Microbe Interactions **25**, 1275–1285.

Wang X-B, Wu Q, Ito T, Cillo F, Li W-X, Chen X, Yu J-L, Ding S-W. 2010. RNAi-mediated viral immunity requires amplification of virus-derived siRNAs in Arabidopsis thaliana. Proceedings of the National Academy of Sciences of the United States of America **107**, 484–489.

Waterhouse PM, Fusaro AF. 2006. Viruses face a double defense by plant small RNAs. Science **313**, 54–55.

Whitham SA, Yang C, Goodin MM. 2006. Global impact: elucidating plant responses to viral infection. Molecular Plant-Microbe Interactions **19**, 1207–1215.

Widhalm JR, Dudareva N. 2015. A familiar ring to it: Biosynthesis of plant benzoic acids. Molecular Plant **8**, 83–97.

Winter S, Koerbler M, Stein B, Pietruszka A, Paape M, Butgereitt A. 2010. Analysis of cassava brown streak viruses reveals the presence of distinct virus species causing cassava brown streak disease in East Africa. Journal of General Virology **91**, 1365–1372.

660 **Yi X, Du Z, Su Z**. 2013. PlantGSEA: a gene set enrichment analysis toolkit for plant community.
661 Nucleic Acids Research **41**, 98–103.

662 **Ying X-B, Dong L, Zhu H, Duan C-G, Du Q-S, Lv D-Q, Fang Y-Y, Garcia JA, Fang R-X, Guo H-S**.
663 2010. RNA-dependent RNA polymerase 1 from Nicotiana tabacum suppresses RNA silencing and
664 enhances viral infection in Nicotiana benthamiana. The Plant Cell **22**, 1358–72.

665 **Zavaliev R, Epel BL**. 2015. Imaging callose at plasmodesmata using aniline blue: quantitiative
666 confocal microscopy. In: Heinlein M, ed. Plasmodesmata: Methods and Protocols, Methods in
667 Molecular Biology. New York: Springer Science+Business Media, 105–119.

668 **Zheng XY, Spivey NW, Zeng W, Liu PP, Fu ZQ, Klessig DF, He SY, Dong X**. 2012. Coronatine
669 promotes pseudomonas syringae virulence in plants by activating a signaling cascade that inhibits
670 salicylic acid accumulation. Cell Host & Microbe **11**, 587–596.

671 Fig. Legends

672 **Fig. 1:** Gene regulation in susceptible (60444) and resistant (KBH 2006/18) cassava varieties at 28
673 dag and mixed CBSV and UCBSV infection. **a.** Fold change of differentially expressed genes (DEGs)
674 (FDR<0.05, fold change ≥2) in both varieties. Fold change of DEGs in 60444 form the X–coordinate
675 and in KBH 2006/18 form the Y–coordinate. **b.** Functional characterization of DEGs based on gene
676 ontology (GO) of biological processes. Significantly enriched GO categories in 60444 and KBH
677 2006/18 were first reduced based on semantic similarity (REVIGO; revigo.irb.hr) and then plotted.
678 Blue circles represent GO categories enriched in 60444 and orange circles those in KBH 2006/18;
679 overlapping blue and orange circles represent GO categories enriched in both 60444 and KBH
680 2006/18. Circles for the top 10 significantly enriched categories are shaded. **c.** Proportion of nuclear-
681 encoded chloroplast genes (NECGs) differentially regulated after infection. 6% of the cassava
682 genome consists of NECGs while 16% of the 60444 and 20% of the KBH 2006/18 DEGs after
683 infection are NECGs.

Fig. 2: Regulation of callose deposition at plasmodesmata in susceptible (60444) and resistant (KBH 2006/18) cassava varieties after CBSV-infection. **a.** Schematic depiction of a plasmodesmata used by viruses for intercellular movement. Also shown are genes encoding key enzymes significantly regulated after infection (*BG3: GLUCANASE*, encoding a callose-degrading β-1,3-glucanase; *GSL4: GLUCAN SYNTHASE-LIKE 4*, encoding a callose synthesis enzyme; *PDLP1*: *PLASMODEMATA-LOCATED PROTEIN 1*, encoding a protein involved in virus movement across the plasmodesmata and in callose synthesis; *ANKs: ANKYRIN REPEAT FAMILY PROTEINS*, encoding proteins that are negative regulators of callose deposition) with their fold-changes based on RNA-Seq analysis at 28 dag. **b.** *(upper panel)* RT-qPCR analysis of *BG3* mRNA at 60 dag. 60444 leaves were symptomatic for virus infection at this time-point. *(lower panel)* BG3 enzymatic activity also significantly increases after infection in 60444 and is unchanged in KBH 2006/18 at 60 dag. **c.** Aniline blue staining for detection of callose deposits. Leaf samples from 60444 were stained and visualized using a confocal laser scanning microscope at 400X magnification. Callose deposits are visible as small bright dots along the cell walls (white arrowheads). **d.** (*upper panel)* Quantitation of callose deposits at plasmodesmata in 60444 and KBH 2006/18 at 60 dag, observed after aniline blue staining and microscopy. Infection results in a decrease in the amount of callose deposited at the plasmodesmata in 60444 and an increase in KBH 2006/18. (*lower panel)* The number of callose deposits found in microscopic images. Infection also decreases in the number of callose deposits in 60444. (*****, $P < 0.05$; ******, $P < 0.01$, Un-paired t-test.)

Fig. 3: Regulation of genes encoding proteins of the salicylic acid synthesis and signalling and the lignin synthesis pathways in cassava after virus infection. Significantly regulated genes are depicted together with their mRNA fold-change based on RNA-Seq analysis in susceptible (60444) and resistant (KBH 2006/18) cassava varieties after infection. *PFK3*: *PHOSPHOFRUCTOKINASE 3*; *SK1*: *SHIKIMATE KINASE 1*; *ADT1*, *ADT6*: *AROGENATE DEHYDRATASE 1 & 6*; *PAL1*, *PAL2*: *PHENYLALANINE AMMONIA–LYASE 1 & 2*; *C4H*: *CINNAMATE–4–HYDROXYLASE*; *4CL1*, *4CL2*, *4CL3*: *4–COUMARATE–CoA LIGASE 1, 2 & 3*; *HCT*: *HYDROXYCINNAMOYL–COA SHIKIMATE/QUINATE*

710 *HYDROXYCINNAMOYL TRANSFERASE*; *CCR*: *CINNAMOYL COA REDUCTASE 1*; *CAD*: *CINNAMOYL*
711 *ALCOHOL DEHYDROGENASE*; *NPR1*: *NON–EXPRESSOR OF PR GENES 1*; *NIMIN1*, *NIMIN2*, *NIMIN3*:
712 *NIM1–INTERACTING*; *TGA9*: *TGACG MOTIF–BINDING PROTEIN 9*; *WRKY70*: *WRKY DNA–BINDING*
713 *PROTEIN 70*; *SAG12*: *SENESCENCE–ASSOCIATED GENE 12*.

714 **Fig. 4:** Schematic depiction of the antiviral RNA silencing pathway with DEGs encoding proteins that
715 function in antiviral defence. Genes significantly regulated in variety 60444 after inoculation are
716 shown together with their mRNA fold-change based on RNA-Seq analysis. These genes were not
717 regulated in variety KBH 18/2006. *DCL2*: *DICER–LIKE 2*, *AGO5*: *ARGONAUTE 5*, *AGO2*: *ARGONAUTE*
718 *2*, *RDR1*: *RNA DEPENDENT RNA POLYMERASE 1*, *SGS3*: *SUPRESSOR OF GENE SILENCING 3*, *SDE5*:
719 *SILENCING DEFECTIVE 5*.

720 **Fig. 5:** Comparison of transcriptome changes in different virus-host systems. We compared our
721 RNA-Seq of CBSV-infected cassava with three other poty virus-plant transcriptome studies (late-
722 stage cassava-CBSV infection, Maruthi et al., 2014; early-stage peach and late-stage apricot
723 infection with Plum Pox Virus, Rubio *et al.*, 2015a,b). In addition, we included a RNA-Seq study of
724 cassava infected with South African Cassava Mosaic virus, a ssDNA geminivirus pathosystem (Allie
725 et al., 2014), to compare the response of cassava to two different virus types. **a.** Pairwise similarity
726 matrix of numbers of DEGs in each pathosystem, based on comparison of Pfam identifiers
727 associated with each gene. A pairwise percentage similarity score was computed by multiplying the
728 DEGs in one pathosystem with 200 and then dividing this number by the sum of the DEGs in both
729 pathosystems that were compared. **b.** and **c.** Venn diagrams were constructed using the Pfam
730 identifiers associated with DEGs from all five studies. The Fig. shows specific Pfam identifiers
731 common to all five studies as well as those common to all potyvirus studies. **b.** Up–regulated DEGs.
732 **c.** Down–regulated DEGs.

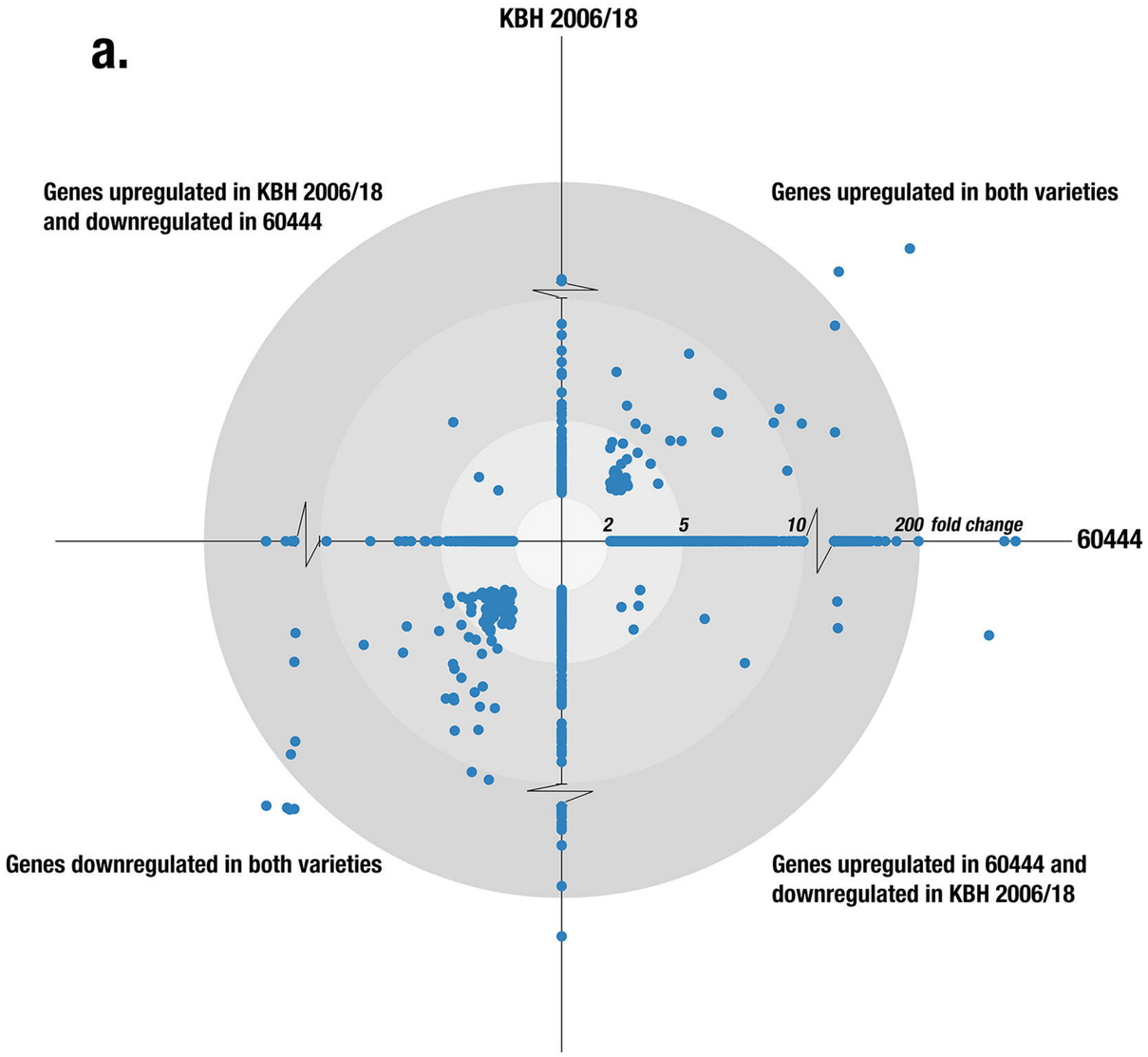
733 **Fig. S1:** Experimental scheme for the study. **a.** Schematic representation of the workflow for RNA-
734 seq study. **b.** Workflow for callose quantitation, enzymatic assay and expression analysis for β 1,3
735 Glucanase.

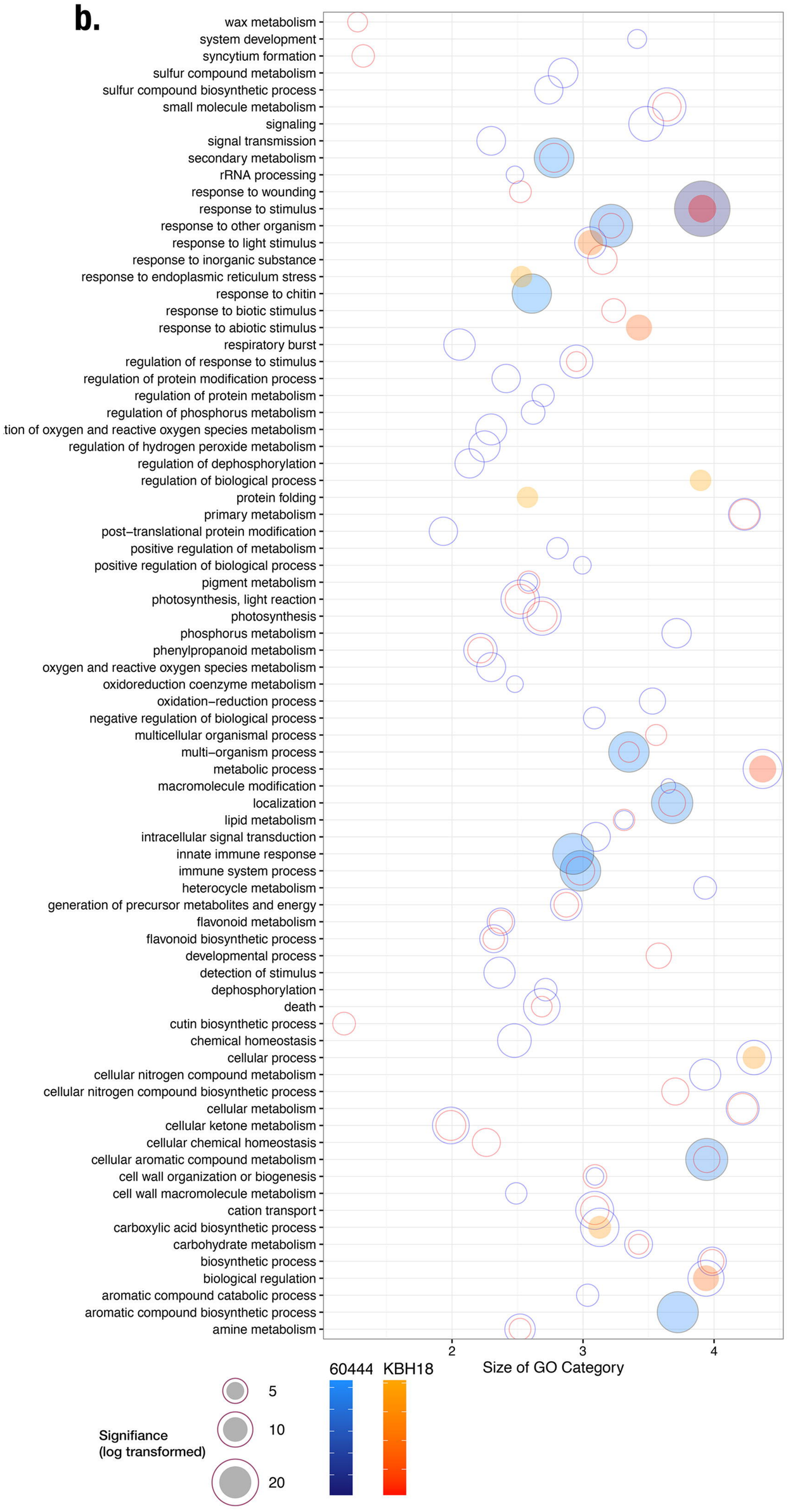
736 **Fig. S2:** CBSV quantitation at 3 different time points after grafting. RT-qPCR quantitation of virus
737 titre (log2 fold change relative to reference gene mePP2A) from individual leaves from 3
738 independent biological replicates at three different time points after grafting.

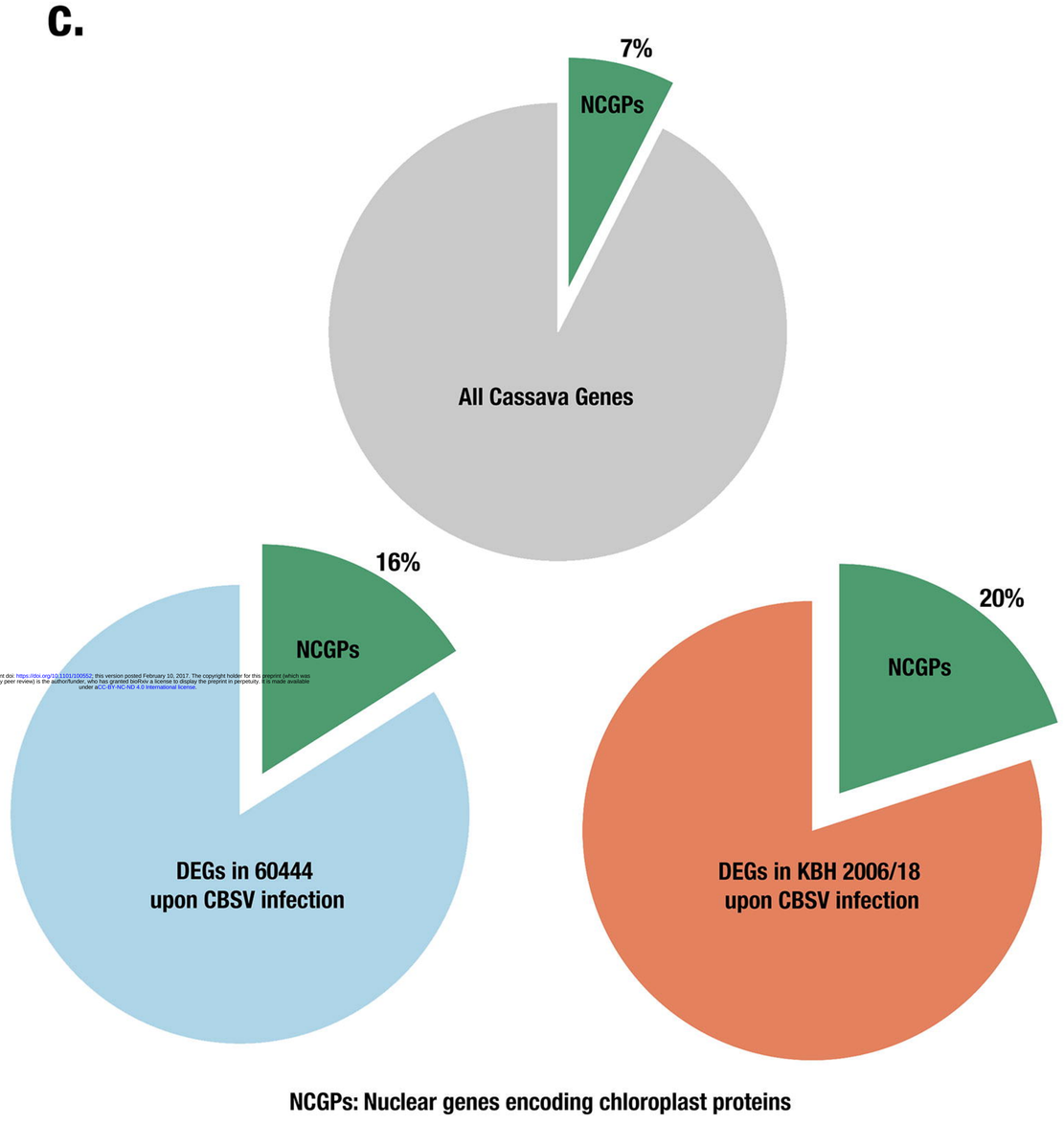
739 **Fig. S3:** Validation of RNA-seq data by RT-qPCR. RT-qPCR for four genes was performed on the
740 same samples that were used for RNA-seq analysis.

741 **Fig. S4:** Virus read counting from unmapped RNA-seq reads. Unmapped reads from each sample
742 sent for RNA-seq were mapped to the two *de novo* assembled virus genomes. **a.** Read distribution
743 across virus genomes for 60444 samples. **b.** Read counts for all samples in the compatible (60444)
744 and incompatible (KBH 2006/18) virus-host interaction.

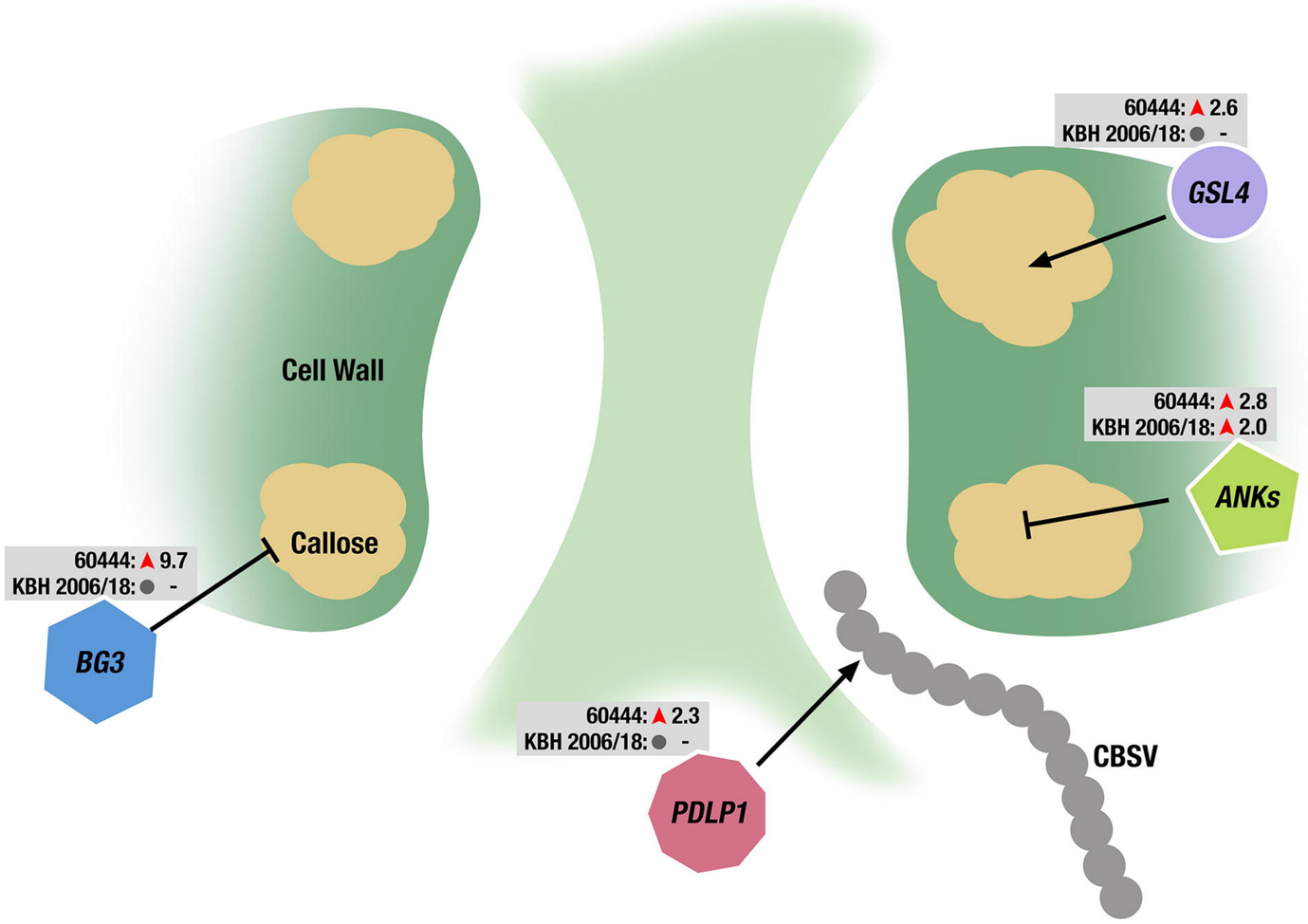
745 **Fig. S5:** RT-qPCR based RDR1 transcript expression across the three time points.



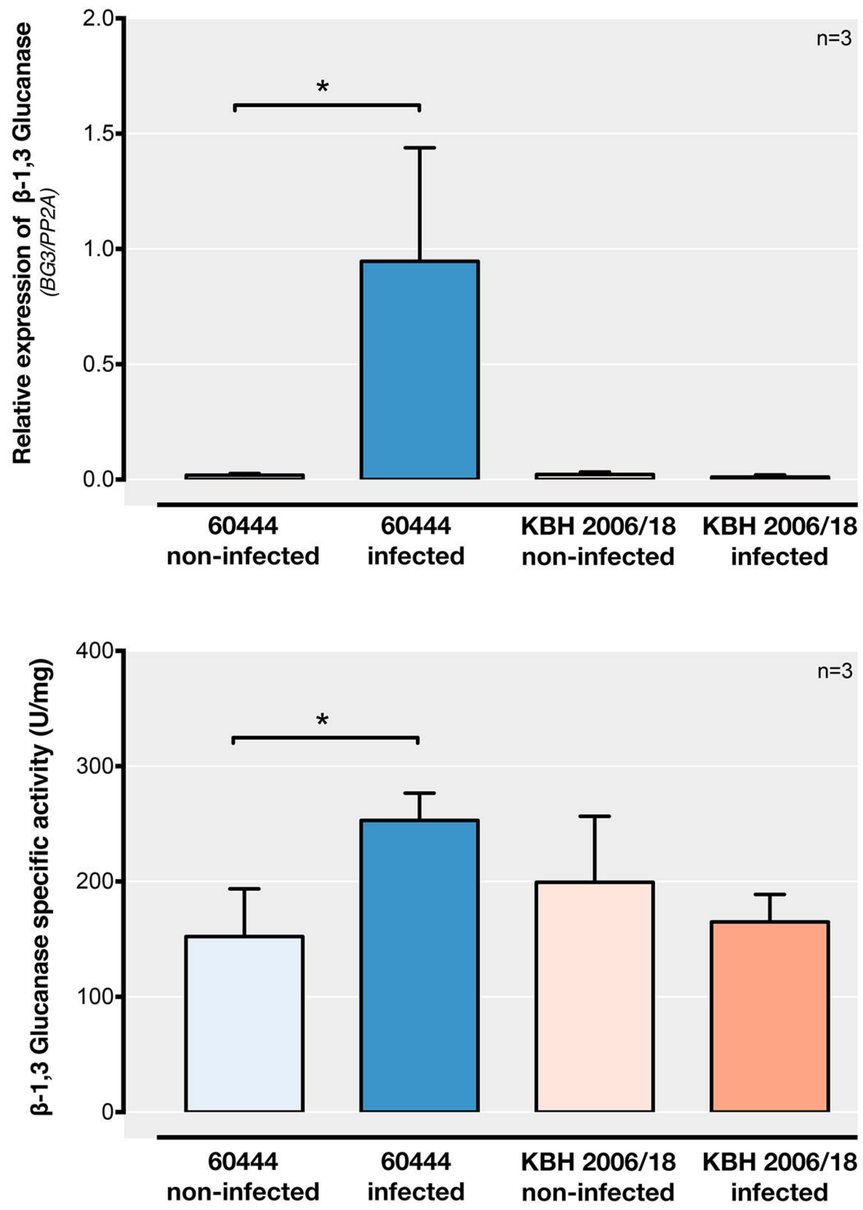




a.



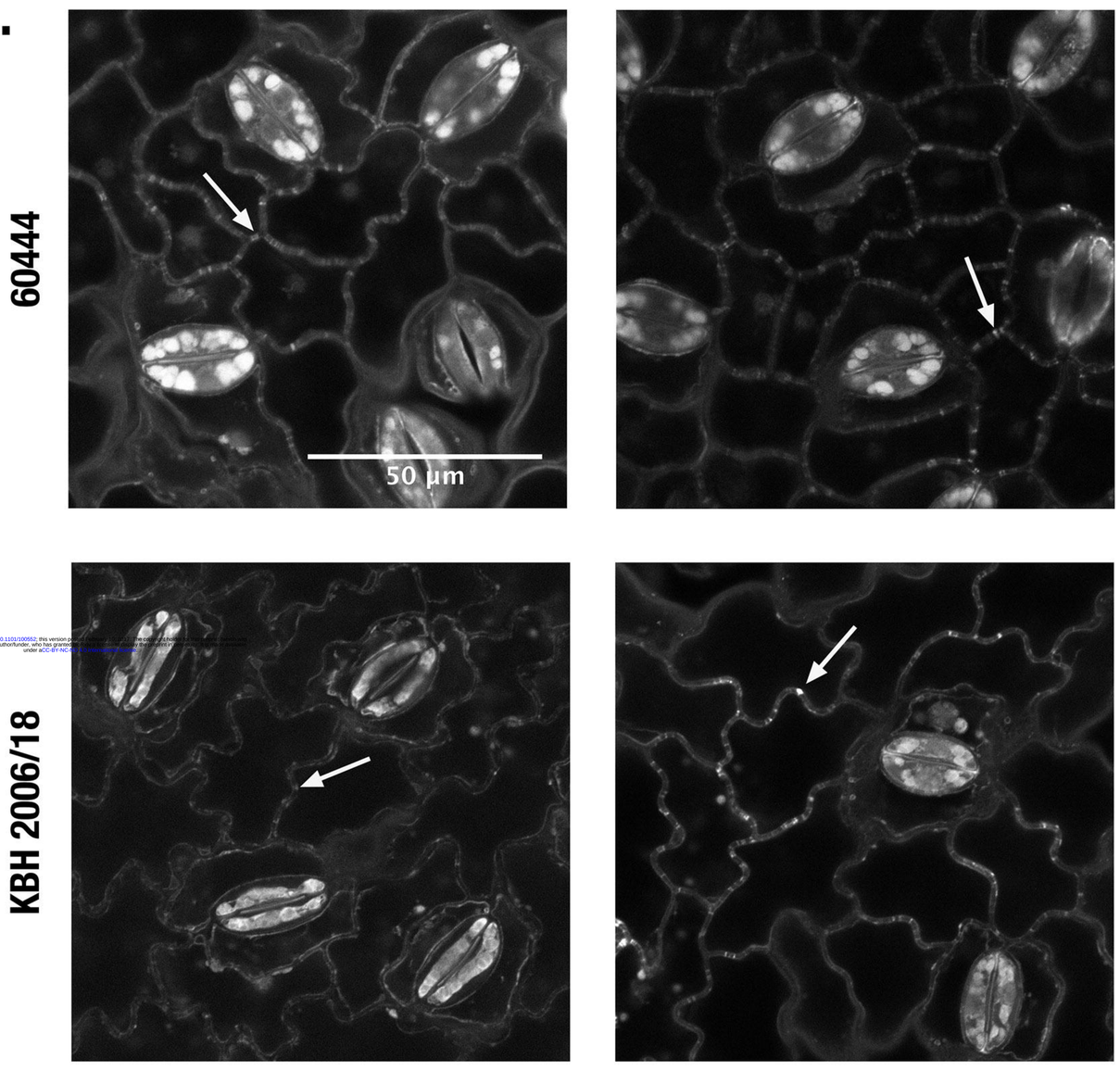
b.



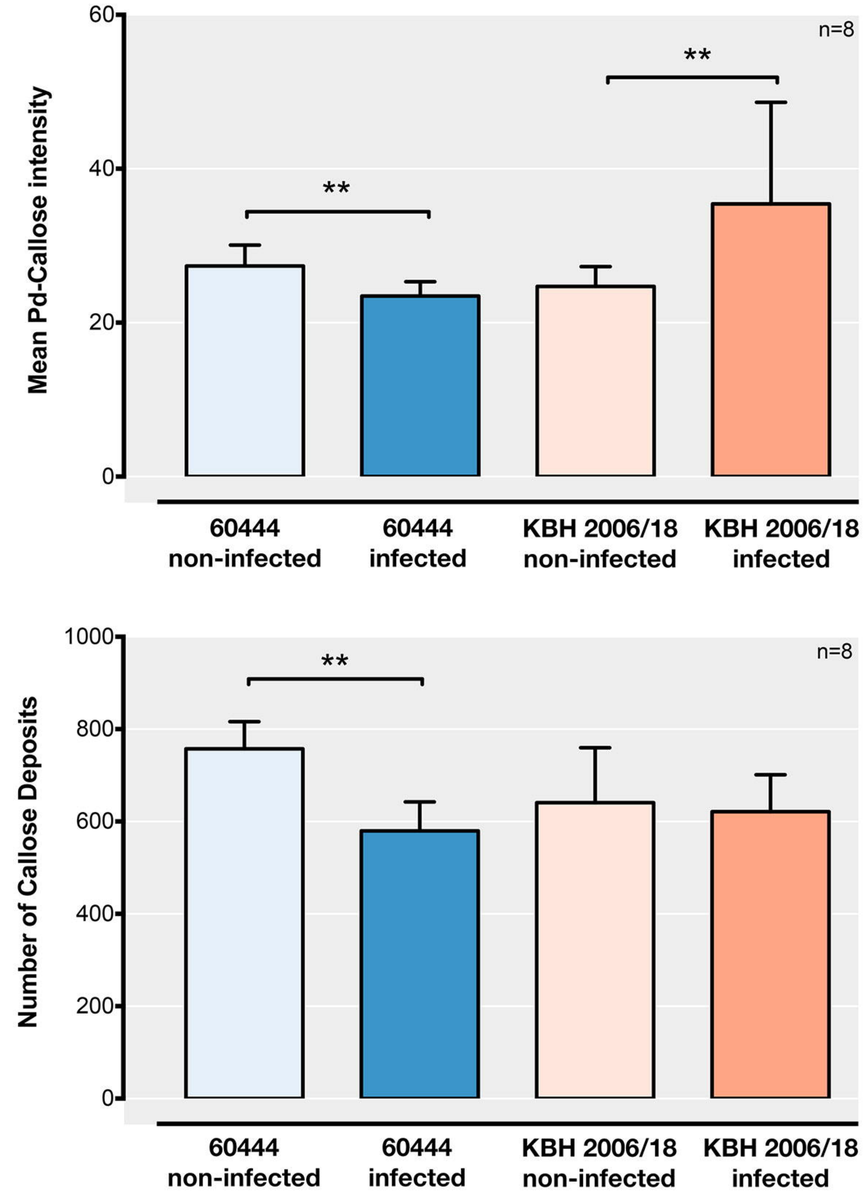
Non-infected

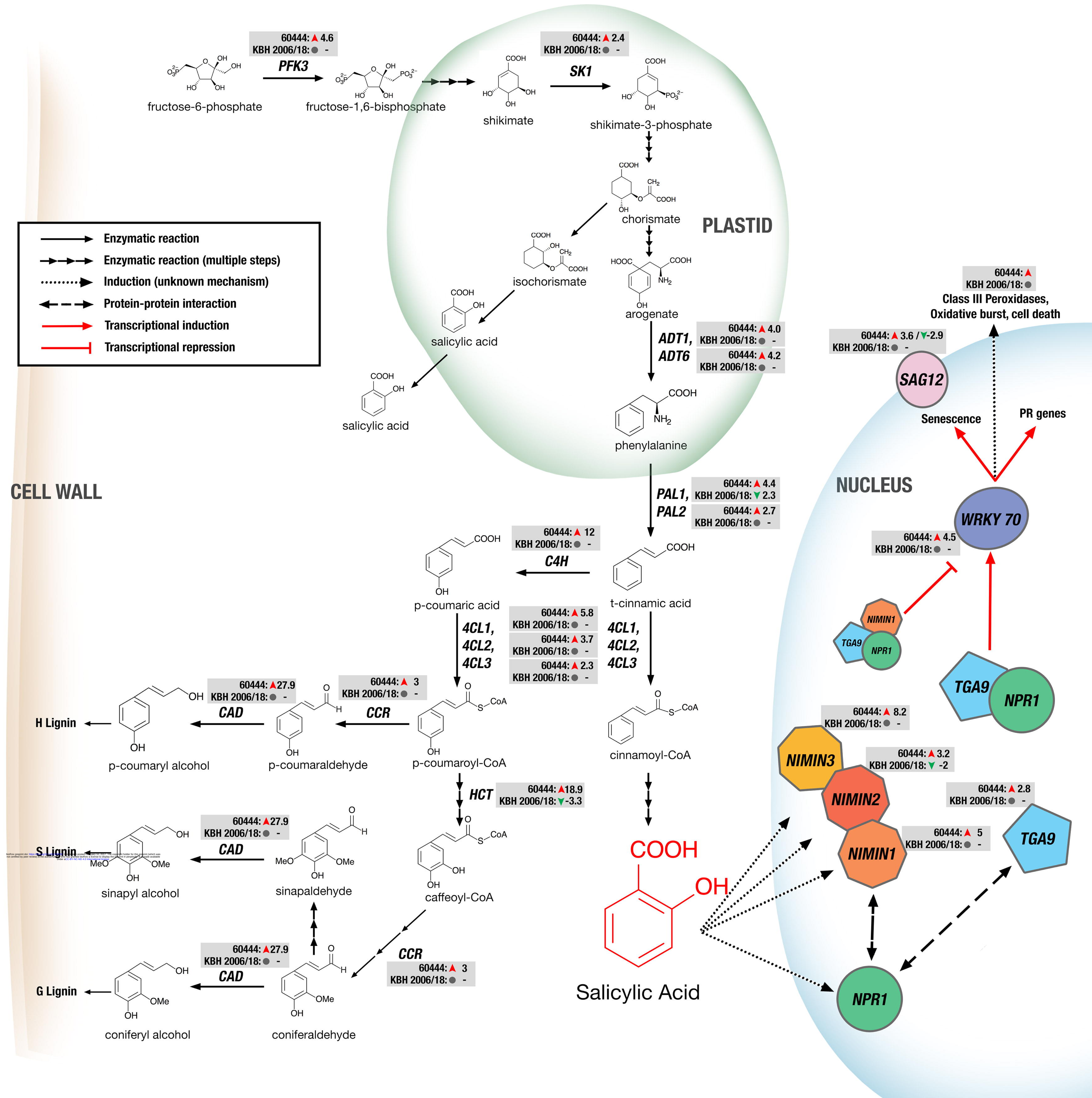
Infected

c.



d.





fructose-6-phosphate
PFK3
60444: 4.6
KBH 2006/18: -
fructose-1,6-bisphosphate
shikimate
SK1
60444: 2.4
KBH 2006/18: -
shikimate-3-phosphate
chorismate
PLASTID
isochorismate
salicylic acid
salicylic acid
arogenate
ADT1,
ADT6
60444: 4.0
KBH 2006/18: -
60444: 4.2
KBH 2006/18: -
phenylalanine
Enzymatic reaction
Enzymatic reaction (multiple steps)
Induction (unknown mechanism)
Protein-protein interaction
Transcriptional induction
Transcriptional repression
CELL WALL
PAL1,
PAL2
60444: 4.4
KBH 2006/18: 2.3
60444: 2.7
KBH 2006/18: -
t-cinnamic acid
C4H
60444: 12
KBH 2006/18: -
p-coumaric acid
4CL1,
4CL2,
4CL3
60444: 5.8
KBH 2006/18: -
60444: 3.7
KBH 2006/18: -
60444: 2.3
KBH 2006/18: -
cinnamoyl-CoA
p-coumaroyl-CoA
CCR
60444: 3
KBH 2006/18: -
p-coumaraldehyde
CAD
60444: 27.9
KBH 2006/18: -
p-coumaryl alcohol
H Lignin
HCT
60444: 18.9
KBH 2006/18: -3.3
caffeoyl-CoA
sinapaldehyde
sinapyl alcohol
S Lignin
CCR
60444: 3
KBH 2006/18: -
coniferaldehyde
coniferyl alcohol
G Lignin
Salicylic Acid
NUCLEUS
60444:
KBH 2006/18:
Class III Peroxidases,
Oxidative burst, cell death
60444: 3.6 / -2.9
KBH 2006/18: -
SAG12
Senescence
PR genes
WRKY 70
60444: 4.5
KBH 2006/18: -
NIMIN1
TGA9
NPR1
TGA9
NPR1
60444: 8.2
KBH 2006/18: -
NIMIN3
60444: 3.2
KBH 2006/18: -2
NIMIN2
60444: 5
KBH 2006/18: -
NIMIN1
60444: 2.8
KBH 2006/18: -
TGA9
NPR1

AAAAAAAAA

v. viral RNA degradation

SDE5

60444: ▲ 2
KBH 2006/18: ● -

CBSV RNA

60444: ▲ 39
KBH 2006/18: ● -

AAAAAAAAA

RDR1

SGS3

60444: ▲ 3
KBH 2006/18: ● -

i. dsRNA formation

Long dsRNA

DCL2

60444: ▲ 3.3
KBH 2006/18: ● -

AGO2

60444: ▲ 18
KBH 2006/18: ● -

vi. secondary silencing

iii. dicing

iv. RISC formation

AGO5

60444: ▲ 2.3
KBH 2006/18: ● -

ii. translation
& replication

siRNA

CBSV Viroid