






In vivo assembly of the sorgoleone biosynthetic pathway and its impact on agroinfiltrated leaves of *Nicotiana benthamiana*

Zhiqiang Pan¹ , Joanna Bajsa-Hirschel¹ , Justin N. Vaughn² , Agnes M. Rimando^{1†}, Scott R. Baerson¹  and Stephen O. Duke¹ 

¹Natural Products Utilization Research Unit, US Department of Agriculture, Agricultural Research Service, University, MS 38677, USA; ²Genomics and Bioinformatics Research Unit, USDA, ARS, Athens, GA 30605, USA

Summary

Author for correspondence:
Zhiqiang Pan
Email: zhiqiang.pan@usda.gov

Received: 30 September 2020
Accepted: 17 December 2020

New Phytologist (2021) 230: 683–697
doi: 10.1111/nph.17213

Key words: allelochemical, biosynthetic pathway, differential gene expression, *Nicotiana benthamiana*, *Sorghum bicolor*, Sorgoleone.

- Sorgoleone, a hydrophobic compound exuded from root hair cells of *Sorghum* spp., accounts for much of the allelopathic activity of the genus. The enzymes involved in the biosynthesis of this compound have been identified and functionally characterized. Here, we report the successful assembly of the biosynthetic pathway and the significant impact of *in vivo* synthesized sorgoleone on the heterologous host *Nicotiana benthamiana*.
- A multigene DNA construct was prepared for the expression of genes required for sorgoleone biosynthesis *in planta* and deployed in *N. benthamiana* leaf tissues via *Agrobacterium*-mediated transient expression. RNA-sequencing was conducted to investigate the effects of sorgoleone, via expression of its biosynthesis pathway, on host gene expression.
- The production of sorgoleone in agroinfiltrated leaves as detected by gas chromatography/mass spectrometry (GC/MS) resulted in the formation of necrotic lesions, indicating that the compound caused severe phytotoxicity to these tissues. RNA-sequencing profiling revealed significant changes in gene expression in the leaf tissues expressing the pathway during the formation of sorgoleone-induced necrotic lesions.
- Transcriptome analysis suggested that the compound produced *in vivo* impaired the photosynthetic system as a result of downregulated gene expression for the photosynthesis apparatus and elevated expression of proteasomal genes which may play a major role in the phytotoxicity of sorgoleone.

Introduction

Sorghum species are known to possess allelopathic properties which result from the biosynthesis and release of biologically active compounds that repress the growth of weeds (Putnam *et al.*, 1983; Forney *et al.*, 1985; Einhellig & Rasmussen, 1989). Early studies on the exudates from the roots of *Sorghum bicolor* demonstrated that the root exudates play a role in the growth inhibition of lettuce seedlings (*Lactuca sativa*), as well as a number of important invasive weed species (Putnam *et al.*, 1983; Forney *et al.*, 1985; Netzly & Butler, 1986). The major component of the hydrophobic root exudate from *S. bicolor* of various accessions was identified as 2-hydroxy-5-methoxy-3-[(Z,Z)-8',11',14'-pentadecatriene]-*p*-benzoquinone, referred to as sorgoleone (Netzly & Butler, 1986). Sorgoleone appears to affect multiple molecular targets (Weston *et al.*, 2013). It is a potent inhibitor of photosystem II (Einhellig *et al.*, 1993; Gonzalez *et al.*, 1997), as it is a close analog of plastoquinone (Czarnta *et al.*, 2001). Sorgoleone also strongly inhibits the enzyme HPPD (*p*-hydroxyphenylpyruvate dioxygenase), which is involved in the

formation of plastoquinone (Meazza *et al.*, 2002) and has been shown to act as a respiratory inhibitor (Rasmussen *et al.*, 1992). These phytotoxic and allelopathic properties of sorgoleone with multiple molecular target sites make it promising for development as a natural product alternative to synthetic herbicides (Duke, 2003).

The biosynthesis of sorgoleone only occurs in root hair cells, from which it is exuded in oily droplets (Dayan *et al.*, 2009). At the transmission electron microscope level, sorgoleone is observed as cytoplasmically dense osmiophilic globules in root hair cells, which are deposited between the plasmalemma and cell wall, and are presumably associated with sorgoleone rhizosecretion (Czarnta *et al.*, 2003). The biosynthetic pathway that leads to the formation of sorgoleone was proposed according to *in vivo* labeling studies (Fate & Lynn, 1996; Dayan *et al.*, 2003). It was later demonstrated that the biosynthesis of sorgoleone (Fig. 1) involves the production of an alkylresorcinolic intermediate, 5-[(Z,Z)-8',11',14'-pentadecatrienyl]resorcinol by SbARS2 (Cook *et al.*, 2010) that utilizes a 16:3 $\Delta^{9,12,15}$ fatty acyl-CoA starter unit, which is generated from palmitoleoyl-CoA catalyzed consecutively by desaturases SbDES2 (Pan *et al.*, 2007) and SbDES3 (Yang *et al.*, 2004; Pan *et al.*, 2007). This intermediate is then

†Deceased.

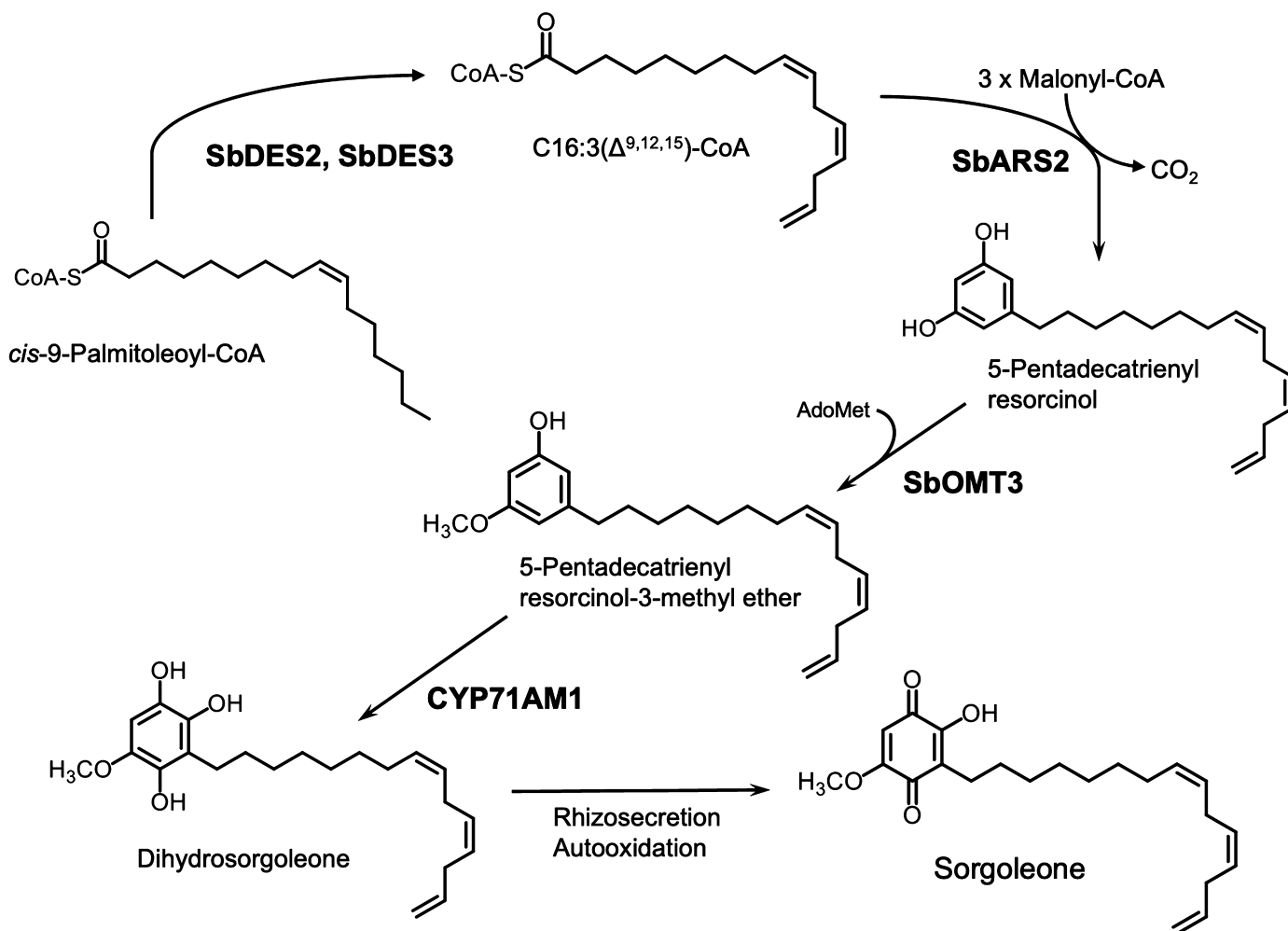


Fig. 1 Biosynthesis of sorgoleone. The enzymatic reactions starting from palmitoleoyl-CoA are shown. Dihydrosorgoleone, the hydroquinone produced *in vivo*, undergoes auto-oxidation once secreted into the soil to yield sorgoleone, a more stable benzoquinone. CYP71AM1, cytochrome P450; SbARS, alkylresorcinol synthase; SbDES, fatty acid desaturase; SbOMT, *O*-methyltransferase.

methylated by SAM-dependent *O*-methyltransferases SbOMT3 (Baerson *et al.*, 2008) and dihydroxylated by cytochrome P450 enzyme CYP71AM1, an enzyme that was recently identified and functionally characterized (Pan *et al.*, 2018), yielding dihydrosorgoleone, which then rapidly undergoes oxidation upon rhizosecretion to the benzoquinone sorgoleone (Fig. 1).

Upon completing the identification and functional characterization of genes encoding the enzymes catalyzing all of the biosynthetic steps leading to the production of dihydrosorgoleone, questions arose as to whether *in vivo* expression of all these genes simultaneously could produce dihydrosorgoleone, the precursor of sorgoleone, and, if it does, how the heterologous host cells would respond to the synthesized compound physiologically and how the host gene expression would be impacted. In this report, we described the assembly of the pathway using *Agrobacterium*-mediated transient expression assays with *Nicotiana benthamiana*, which has been widely used to transiently express genes to study their functions (Sainsbury & Lomonosoff, 2014; Bally *et al.*, 2018). Agroinfiltrated *N. benthamiana* leaf tissues using a multi-gene vector containing all of the genes required for sorgoleone biosynthesis caused necrotic lesions, evidence of phytotoxicity.

Chemical analysis results obtained from these agroinfiltrated leaves were consistent with the plasmid-dependent accumulation of sorgoleone. Furthermore, transcriptome analysis revealed that sorgoleone significantly impacts gene expression on *N. benthamiana* leaf tissues.

Materials and Methods

Chemicals and plant materials

Standard laboratory reagents were purchased from Sigma unless specified otherwise. Seeds of *Nicotiana benthamiana* (TW 16) were obtained from the United States Department of Agriculture–Agricultural Research Service (USDA-ARS), National Genetic Resources Program Germplasm Resources Information Network (GRIN), and plants were maintained in a growth chamber (Conviron, Pembina, ND, USA) at 24°C, under a 16 h : 8 h, light : dark photoperiod with 150 $\mu\text{mol m}^{-2} \text{s}^{-1}$ light). All tissues were collected and then flash-frozen in liquid nitrogen and kept at -80°C before being used for RNA extraction and sorgoleone extraction.

Plasmid construction

A multigene vector, pLH-Sorg, for transient expression experiments was assembled to contain transgene cassettes directing the expression of *S. bicolor* *SbDES2*, *SbDES3*, *SbARS2*, *SbOMT3*, and *CYP71AM1* (see Supporting Information Methods S1 for details on the construction of the plasmid). Binary vectors containing either one expression cassette (i.e. *CYP71AM1* (pLH-G)) or multiple expression cassettes (i.e. *SbDES2-SbDES3* (pLH-DES), *SbARS2-SbOMT3* (pLH-PQ), *SbARS2-SbOMT3-CYP71AM1* (pLH-GPQ), and *SbDES2-SbDES3-SbARS2-SbOMT3* (pLH-PQD)) were also assembled to test the possible effects of intermediates in the sorgoleone biosynthetic pathway (see Methods S1 for details). To avoid the possibility of interference – for example, homologous recombination in *E. coli*, or transgene silencing in plants (Chung *et al.*, 2005) – we made the constructs using a combination of different promoters/terminators for the expression of these sorgoleone genes.

Transient expression in leaves of *Nicotiana benthamiana*

The binary vectors for transient expression experiments (pLH-Sorg, pLH-G, pLH-DES, pLH-PQ, pLH-GPQ and pLH-PQD), as well as pCB404-P19 (Pan *et al.*, 2018), were mobilized into the *Agrobacterium tumefaciens* strain EHA105 as described previously (Hofgen & Willmitzer, 1988). The cultures were maintained in LB (Luria-Bertani) medium supplemented with appropriate antibiotics (25 mg l⁻¹ rifampicin, 100 mg l⁻¹ spectinomycin, and 300 mg l⁻¹ streptomycin for strains harboring pLH vectors; 25 mg l⁻¹ rifampicin, 300 mg l⁻¹ streptomycin, and 50 mg l⁻¹ kanamycin for the pCB vector). Agroinfiltration for transient expression in the leaf tissues of *N. benthamiana* was carried out as described by Wydro *et al.* (2006). Briefly, bacteria were grown in 10 ml YEB medium (5 g l⁻¹ beef extract, 1 g l⁻¹ yeast extract, 5 g l⁻¹ bacteriological peptone, 5 g l⁻¹ sucrose, and 0.5 g l⁻¹ MgSO₄) containing appropriate antibiotics. Following overnight growth at 28°C, bacteria were pelleted by centrifugation for 20 min at 4000 g, washed once with water, and resuspended in infiltration buffer (10 mM 2-(N-morpholino) ethanesulfone (MES), 10 mM MgCl₂, pH 5.7, 100 μM acetosyringone) to OD₆₀₀ = 0.6. The bacterial suspensions were mixed in a ratio of 3 : 1 (v/v; genes of interest/p19) and infiltrated with a 1-ml syringe without a needle into the abaxial side of c. 4-wk-old *N. benthamiana* leaves. In all agroinfiltration experiments, the *Agrobacterium* strain harboring pCB404-P19 for the expression of P19 silencing repressor (Garabagi *et al.*, 2012) was used to co-infiltrate *N. benthamiana* leaves with strains containing constructs for expressing *S. bicolor* genes. A minimum of 12 plants and 3 leaves per plant were infiltrated for each construct. The infiltrated plants were maintained in the Conviron growth chamber.

Gas chromatography/mass spectrometry (GC/MS) analysis of sorgoleone

To determine the sorgoleone levels in agroinfiltrated *N. benthamiana* leaves, the infiltrated leaf areas were collected from

six plants (18 plants total for three replicates) for each time point. Leaf tissues (5 g ± 0.2 fresh weight) stored at -80°C were powdered in liquid nitrogen using a pestle and mortar. Pulverized tissues were first extracted by gentle swirling in 50 ml of methanol/chloroform (1 : 1) in 250 ml Erlenmeyer flasks for 1 min and sonicated for 30 min (Branson ultrasonic laboratory bath, model 2510; Thermo Fisher Scientific, Waltham, MA). The samples then were filtered through filter paper (VWR 5.5 cm; Missouri City, TX, USA, Qualitative 413) under vacuum. The precipitates were extracted twice more with two additional applications of 25 ml methanol/chloroform (1 : 1). The filtrates then were combined and dried using a rotary evaporator under vacuum. Dried samples were dissolved in 10 ml methanol/chloroform (1 : 1) and impregnated onto 1.5 g silica gels. The extract-silica gels were then dried by stirring in a beaker under a nitrogen stream and placed on the top of the pre-rinsed (using 20 ml 9 : 1 methanol/chloroform) Supelco Sep-Pak C18 Column (10 g; cat. no. WAT043345; Waters, Milford, MA, USA). The extracts were consecutively fractionated using solvents: (1) 40 ml methanol/chloroform (9 : 1); (2) 80 ml methanol/chloroform (9 : 1); (3) 80 ml methanol/chloroform (96 : 4); and (4) 60 ml chloroform. The fractions were dried using a rotary evaporator under vacuum. Next, 1.5–2 mg of column fractions 2 and 3 were transferred to tared DP vials. N, O-bis(trimethylsilyl)trifluoroacetamide (BSTFA) was added to the dried sample to achieve a final concentration of 50 mg ml⁻¹. The solution was heated at 100°C for 30 min and then cooled to room temperature. The silylated fractions were then analyzed by GC-MS according to previously described methods (Pan *et al.*, 2018). Sorgoleone quantification was performed according to the methods described previously by Cook *et al.* (2010).

RNA isolation

For each sample, agroinfiltrated leaf areas were collected from five plants (three leaves per plant), for a total of 15 plants, with three replicates for each timepoint. Total RNAs were isolated from flash-frozen tissues using an RNeasy plant mini kit (Qiagen) according to the manufacturer's instructions. RNAs were then treated with RNase-free DNase I kit to remove residual DNA contamination and were re-purified with an RNeasy MinElute Cleanup Kit (Qiagen) according to the procedure provided by the manufacturer. RNA recovery and purity were determined spectrophotometrically using a NanoDrop 2000 Spectrophotometer (Thermo Fisher Scientific) for these samples, and sample integrity was also assessed by agarose gel electrophoresis. The quality and quantity of prepared total RNA for real-time quantitative reverse transcription-polymerase chain reaction (RT-qPCR) (see Methods S2 for details) were assessed according to the Minimum Information for Publication of Quantitative Real-Time PCR Experiments (MIQE) guidelines (Bustin *et al.*, 2009, 2010).

RNA sequencing and data processing

For RNA-seq library construction, RNA integrity was evaluated using an RNA Nano 6000 Assay Kit and the Bioanalyzer 2100 system (Agilent Technologies, Santa Clara, CA) according to the

manufacturer's instructions. The library preparation was carried out using a TruSeq Stranded mRNA Library Prep Kit (Illumina, San Diego, CA) following the manufacturer's instructions. RNA-seq libraries were sequenced on an Illumina NextSeq 500 Sequencer to generate 75-bp paired-end reads.

All libraries in a sequencing batch were run across all four sequencing lanes. Resultant read files were concatenated, and quality and nucleotide distribution were assessed using FASTX tools. FASTX trimmer was used to remove the last base from all reads. Transcript counts were estimated using KALLISTO with default parameters (kmer = 31) and the transcriptome sequences available at Sol Genomics (ftp://ftp.solgenomics.net/genomes/Nicotiana_benthamiana/annotation/Niben101/). Because KALLISTO performs pseudo-mapping, we cross-validated mapping coverage for four libraries using BOWTIE2 with default. Counts were corrected for transcript length in KALLISTO and used by DESEQ2 for further differential expression analysis. Counts for genes on the infiltrated construct were removed for all analyses except those characterizing the expression of the eight genes on the construct (*SbDES2*, *SbDES3*, *SbARS2*, *SbOMT3*, *CYP71AM1*, *NPTII*, *P19*, and *bar* gene). Two separate differential expression analyses were performed. In the first, day + condition + day: condition (an interaction term) was tested against the reduced model of day + condition (without interaction). Adjusted *P*-values reflect the probability of a day + condition interaction after correcting for multiple tests. In the second, time and condition were combined into a single 'group' factor, and contrasts were reported between conditions separately for each day. In this approach, the log₂ fold change is directly interpretable as the condition effect for that day (construct vs vector control). The RNA-seq data have been deposited in the NCBI Sequence Read Archive (SRA) under the BioProject accession no. PRJNA660791.

Nicotiana benthamiana sequences were then assigned to the Arabidopsis genes (Arabidopsis Columbia-0 reference genome, TAIR10) using the MERCATOR automated protein function annotation pipeline (v.4) on PlaBi database web server (<https://plabi.ipd.de/>). Gene Ontology (GO) analysis of the differentially expressed genes (DEGs) was then performed with the CLASSIFICATION SUPERVIEWER tool at the Bio-Analytic Resource for Plant Biology (<http://bar.utoronto.ca/>) with default settings and a significance of *P*-value < 0.05 (threshold), the G:PROFILER web server (<https://biit.cs.ut.ee/gprofiler>) with default settings and adjusted *P*-value < 0.05, and The Arabidopsis Information Resource (TAIR) website (<https://www.arabidopsis.org/>). For KEGG (Kyoto Encyclopedia of Genes and Genomes) pathway enrichment analysis, the R-based program using the KEGG database in G:PROFILER (Raudvere *et al.*, 2019) was used with default settings and adjusted enrichment *P*-values of 0.05.

Results

Physiological response to the expression of the sorgoleone biosynthetic pathway

Sorgoleone is one of the most extensively studied allelochemicals, and the enzymes that comprise its biosynthetic pathway have

been identified and functionally characterized. To test whether the expression of the pathway in heterologous systems is capable of producing sorgoleone, we constructed a multigene vector, termed pLH-Sorg (Figs 2a, S1), which contains all the genes (*SbDES2*, *SbDES3*, *SbARS2*, *SbOMT3*, *CYP71AM1*) required for sorgoleone biosynthesis from palmitoleoyl-CoA. In this binary vector, cDNAs of these genes were used to assemble the expression cassettes, in which each gene expression cassette consists of a constitutive promoter, an open reading frame (ORF) and a terminator (see Methods S1 for details on vector construction). To avoid potential interference caused by repetitive sequences and transgene silencing in plants (Chung *et al.*, 2005), we selected different promoters and terminators for the construction of the plasmid vector (Figs 2a, S1). *Nicotiana benthamiana* leaves from *c.* 4-wk-old plants were then co-infiltrated with *A. tumefaciens* harboring pLH-Sorg and a strain containing a plasmid expressing the P19 silencing suppressor that has been shown to suppress post-transcriptional gene silencing (Garabagi *et al.*, 2012). Leaves were also co-infiltrated with *Agrobacterium* harboring the empty vector and the strain containing the plasmid expressing P19 as a control. Three leaves from each plant were selected for agroinfiltration (Fig. S2) in order to ensure consistency across the experiments. Remarkably, the areas infiltrated with the multigene construct developed necrotic lesions (Fig. 2b) by 5 d post-infiltration (dpi), in contrast to the infiltrated leaves with the empty vector control. This result indicated that a compound had accumulated to toxic concentrations in the infiltrated leaves as a result of expression of the sorgoleone biosynthetic pathway.

To determine further whether the necrotic lesions were caused by synthesized sorgoleone or the intermediates of the biosynthetic pathway (Fig. 1), we made a series of vectors to express a single gene or combinations of genes within the pathway (Fig. 3a–e). Three leaves from each *N. benthamiana* plant were then infiltrated with *Agrobacterium* harboring one of these constructs. As shown in Fig. 3, expression of *CYP71AM1* by a single gene construct (Fig. 3a,f) or combinations of 2–4 genes (Fig. 3b–e,g–j) of the pathway did not cause significant necrotic lesions; however, expression of three genes (*SbARS2*, *SbOMT3*, and *CYP71AM1*; Fig. 3d,i) resulted in a slightly lesion-like symptom in the infiltrated area, particularly on the younger leaf (leaf on the left of the panel, Fig. 3i) as compared with the empty vector control. We speculated that, by acting upon the endogenous substrates, the chemical products from this construct might be similar to the sorgoleone analogs isolated from *S. bicolor*, that is, 2-hydroxy-5-methoxy-3-pentadecyl-p-benzoquinone (sorgoleone-364) and 2-hydroxy-5-methoxy-3-[8'-pentadecene]-p-benzoquinone (sorgoleone-362), possessing C-15:0 and C-15:1 aliphatic side chains, respectively (Kagan *et al.*, 2003), and lacking the double bonds at C-8' (sorgoleone-364) or C-11' and C-14' (sorgoleone-362) since genes coding for the desaturases (*SbDES2* and *SbDES3*) were not included in the construct. Further, only leaf areas infiltrated with *Agrobacterium* harboring pLH-Sorg with the entire pathway had pronounced necrotic lesions (Fig. 3k). It is worth noting that of three leaves expressing the entire sorgoleone pathway, the older leaf displayed slightly decreased

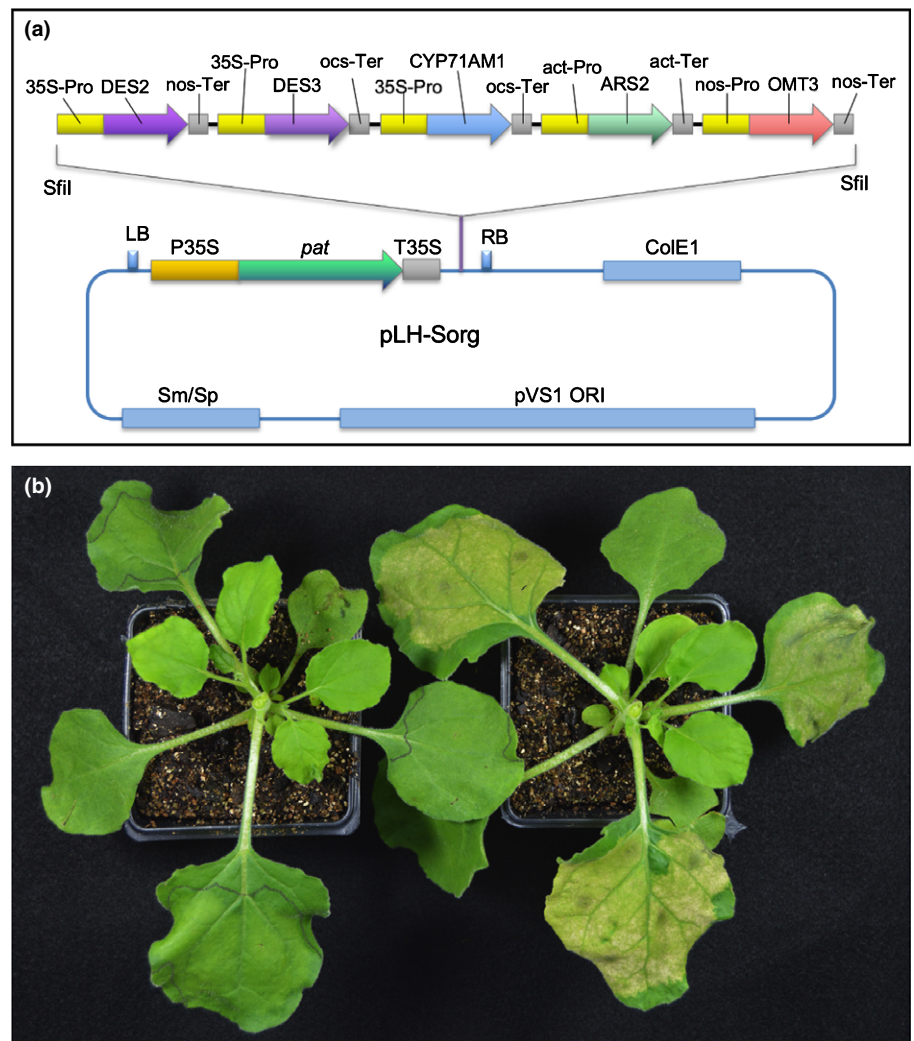


Fig. 2 Physiological response to the transient expression of the sorgoleone biosynthetic pathway in *Nicotiana benthamiana* leaves. (a) A schematic diagram illustrates the details of the construct (pLH-Sorg) containing five gene expression cassettes. (b) Necrotic lesions observed in leaf tissues infiltrated with *Agrobacterium* harboring the plasmid pLH-Sorg at 5 d post-infiltration (dpi) vs *Agrobacterium* bearing vector only as control (left). The younger leaves and stems above the infiltrated leaves were removed at the time the photos were taken. 35S-Pro, 35S CaMV promoter; act-Pro, Arabidopsis actin 2 promoter; act-Ter, Arabidopsis actin 2 terminator; ARS2, *S. bicolor* alkyresorcinol synthases 2; CYP71AM1, *S. bicolor* cytochrome P450; DES2, *Sorghum bicolor* fatty acid desaturase 2; DES3, *S. bicolor* fatty acid desaturase 3; nos-Pro, nopaline synthase promoter; nos-Ter, nopaline synthase terminator; ocs-Ter, octopine synthase terminator; OMT3, *S. bicolor* O-methyltransferase 3.

symptom severity, possibly because older leaves are less transcriptionally active. The expression of these *S. bicolor* genes in *N. benthamiana* infiltrated with these constructs was also confirmed by RT-qPCR (Fig. S3). These results suggest that the necrotic lesions observed in the infiltrated leaves expressing all five genes are most likely caused by the synthesized sorgoleone.

Necrotic lesion development in response to the production of sorgoleone

To examine the possible physiological relevance of sorgoleone being produced by the expression of the biosynthetic pathway, we monitored the development of the necrotic lesions daily on the agroinfiltrated leaves. *Nicotiana benthamiana* leaves were infiltrated with *Agrobacterium* harboring pLH-Sorg for the expression of the pathway or the strain containing the empty vector as control. The initial lesion in the infiltrated leaf area was not visible (Fig. 4a–c) until 4 d after infiltration (Fig. 4d) and increased progressively (Fig. 4e–f), although a slight change in pigmentation or color change was observed at 3 d after infiltration (Fig. 4c). The leaf infiltrated with the *Agrobacterium* bearing the empty vector did not develop necrosis (plant on the left;

Fig. 4) although there was slight chlorosis at 6 dpi (Fig. 3f) which was probably due to the agrobacterial infection. To confirm the expression of the transgenes in the agroinfiltrated leaves, we performed RT-qPCR experiments. The expression of these *S. bicolor* genes in infiltrated *N. benthamiana* leaves was detectable as early as 2 dpi by RT-qPCR, which is also consistent with the RNA-sequencing results (below and Fig. S4). To test whether the necrosis that developed in the agroinfiltrated leaves corroborated the biosynthesis of sorgoleone, the infiltrated leaf areas were collected and subjected to sorgoleone extraction and GC/MS analysis (see the Materials and Methods section and Fig. S5). As shown in Fig. 5, sorgoleone can be detected at an early time point of 3 d after infiltration, suggesting that the lesions may have been the result of the production of sorgoleone. Therefore, in this study we focused on the period from 3 dpi and onwards.

Transcriptome responses to the expression of the sorgoleone biosynthetic pathway

Transcriptome sequencing has been widely used to study genomic responses to many biotic, abiotic, and environmental stresses. To assess how gene expression in *N. benthamiana* was

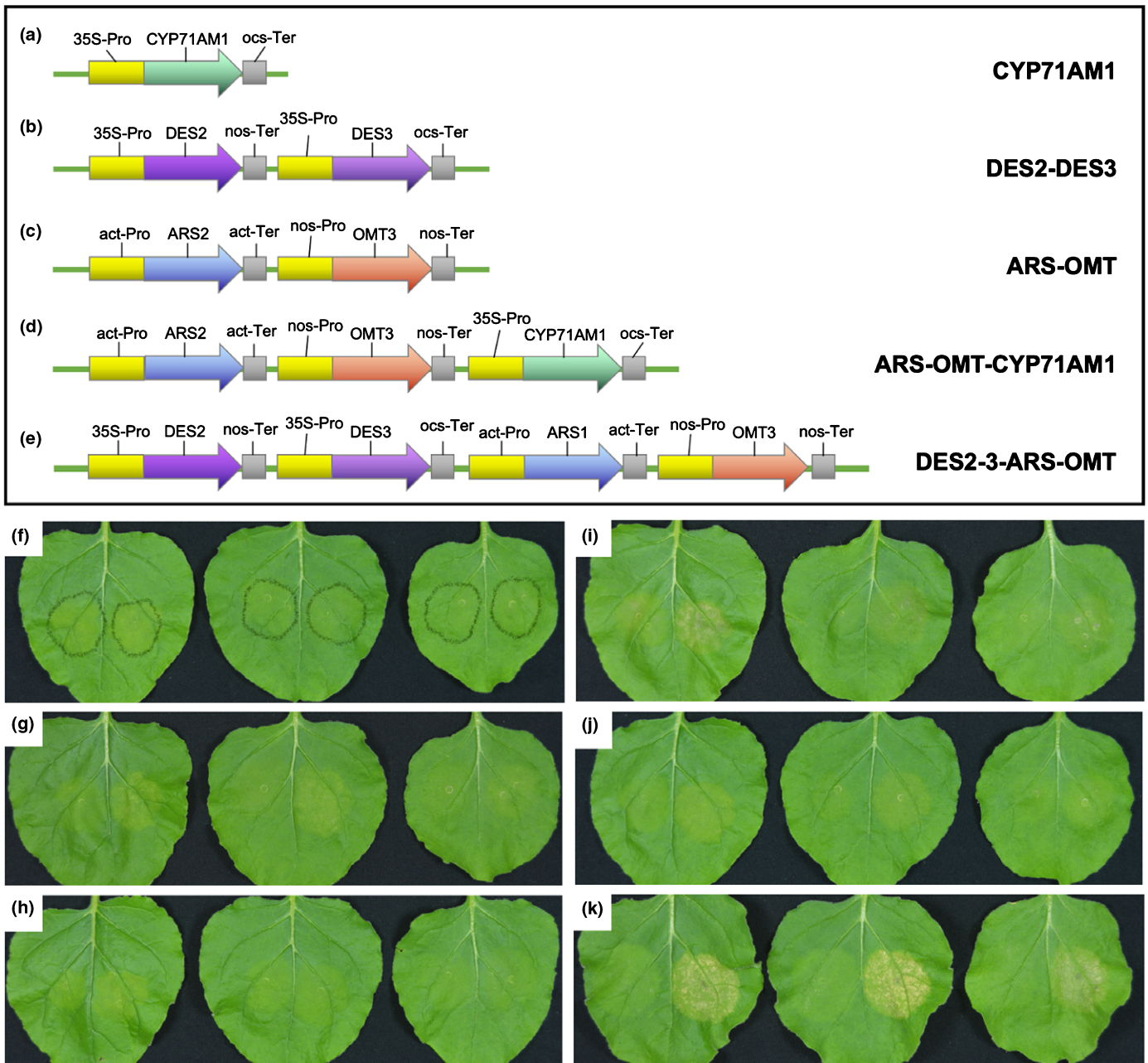


Fig. 3 Gene expression cassettes for a set of constructs and the responses of *Nicotiana benthamiana* leaves to the expression of *Sorghum bicolor* genes. The binary vector pLH7000 was used to make all constructs for gene expression (see Supporting Information Methods S1 for details). (a) The cytochrome P450 CYP71AM1 expression cassette which was inserted into the binary vector to generate pLH-G. (b) The expression cassettes for expressing desaturase 2 (*SbDES2*) and desaturase 3 (*SbDES3*) in the vector pLH-DES. (c) The expression cassettes for the expression of alkylresorcinol synthase 2 (*SbARS2*) and O-methyltransferase 3 (*SbOMT3*) in the vector pLH-PQ. (d) The expression cassettes for expressing *SbARS2*, *SbOMT3*, and *CYP71AM1* in the vector pLH-GPQ. (e) The expression cassettes for expressing *SbDES2*, *SbDES3*, *SbARS2* and *SbOMT3* in the vector pLH-PQD. (f–j) *Nicotiana benthamiana* leaves infiltrated with *Agrobacterium* harboring pLH-G (f), pLH-DES (g), pLH-PQ (h), pLH-GPQ (i), and pLH-PQD (j), respectively. (k) *Nicotiana benthamiana* leaves infiltrated with *Agrobacterium* harboring pLH-Sorg, which contains all five gene expression cassettes for sorgoleone biosynthesis. The leaves depicted in the panels on the left were infiltrated with *Agrobacterium* bearing the empty vector pLH7000 as control. Photos were taken at 5 d post-infiltration (dpi).

impacted by sorgoleone and the possible molecular mechanisms behind such impacts, we performed RNA sequencing (RNA-seq) analysis. *Nicotiana benthamiana* leaves were infiltrated with *Agrobacterium* harboring pLH-Sorg containing all five gene expression cassettes for the sorgoleone biosynthesis. As a control,

N. benthamiana leaves were infiltrated with *Agrobacterium* harboring the empty vector. The leaf tissues were harvested at the time points detailed in Fig. 4. RNAs were extracted from these leaf tissues and subjected to RNA-seq analysis. Leaf tissues were selected for DEG identification from 3 dpi – the earliest time

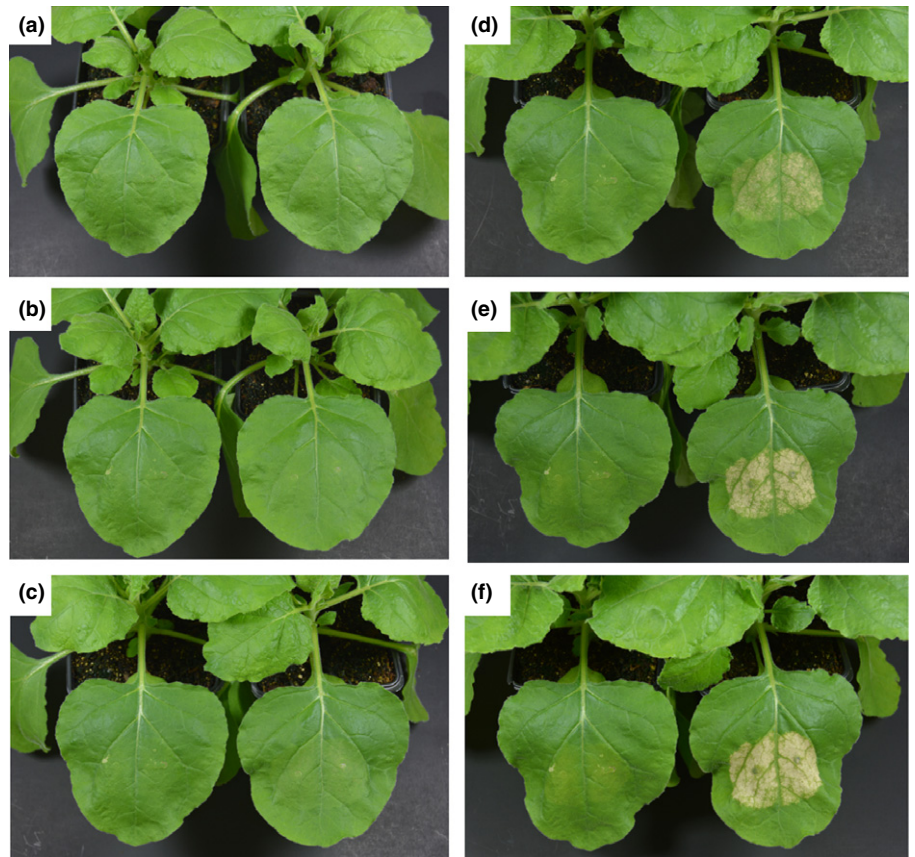


Fig. 4 Development of leaf necrotic lesions. Leaves of *Nicotiana benthamiana* were infiltrated with *Agrobacterium* harboring pLH-Sorg that contains five gene expression cassettes for sorgoleone biosynthesis (pictured in the images on the right) and the empty vector control (left). The development of the lesions was monitored daily: (a) 1 d post infiltration (dpi), (b) 2 dpi, (c) 3 dpi, (d) 4 dpi, (e) 5 dpi, and (f) 6 dpi.

point at which sorgoleone accumulation was evident – onwards. Of all detected transcripts, those with significantly altered expression levels (i.e. more than twofold; $\log_2\text{foldchange} > 1$ and adjusted $P < 0.01$) vs control samples were considered to be differentially expressed genes and were thus used for further analysis. After applying a fold change cut-off, a total of 4822 DEGs (Arabidopsis homologs) were identified at one time (dpi) or another

(Table S1). In detail, there are 3378 DEGs for 3 dpi, 2840 DEGs for 4 dpi, 2143 DEGs for 5 dpi, and 2026 DEGs for 6 dpi (Fig. 6a; Table S1). It is noteworthy that the number of downregulated genes is greater than the number of upregulated genes. Of the 4822 DEGs, 816 genes were differentially expressed across all time points (Fig. 6b; Table S2). These common genes appeared to be the core sorgoleone responsive genes. A comprehensive heat

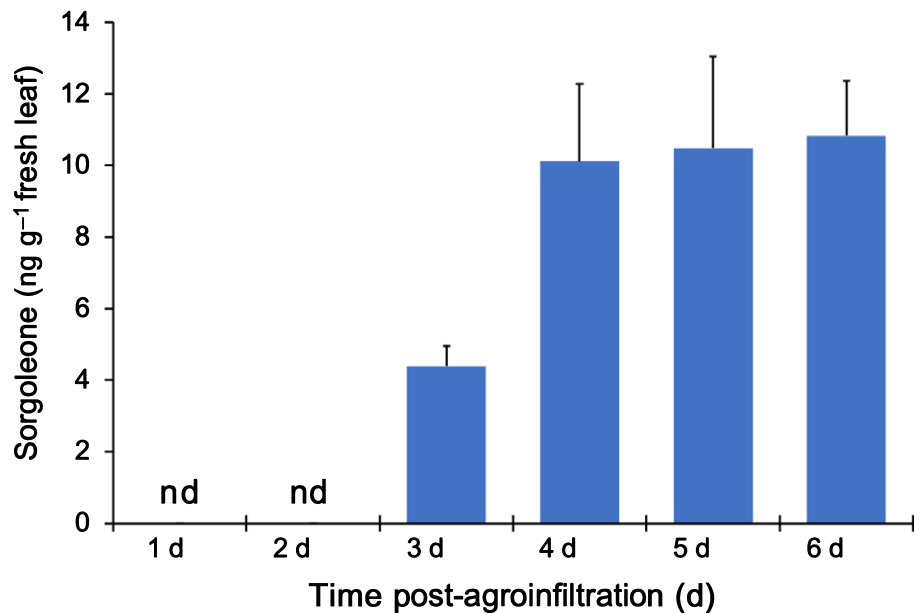


Fig. 5 Sorgoleone content of agroinfiltrated *Nicotiana benthamiana* leaves. Sorgoleone concentrations were determined by gas chromatography/mass spectrometry (GC/MS) analysis of the extracts prepared from agroinfiltrated leaf areas expressing the sorgoleone pathway. Bars represent the means \pm SD from assays performed in triplicate. nd, not detected.

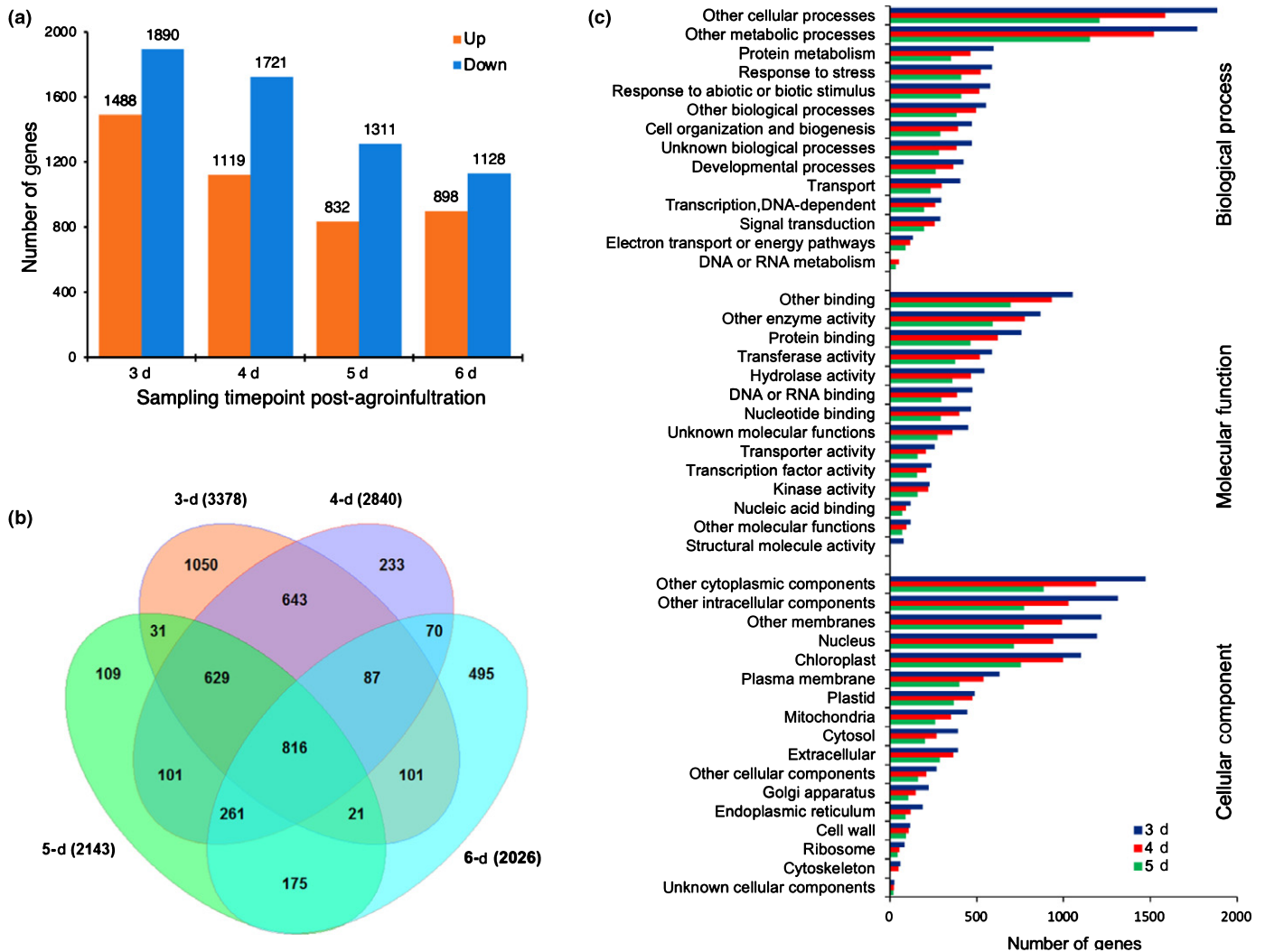


Fig. 6 The global transcriptome response to the expression of the sorgolesone biosynthetic pathway in *Nicotiana benthamiana* leaves. (a) The number of differentially expressed genes (DEGs) at each time point. (b) Venn diagram showing the uncommon and the shared DEGs. (c) Gene Ontology (GO) annotation of differentially expressed genes. Classification of the DEGs was performed using the CLASSIFICATION SUPERVIEWER tool at the Bio-Analytic Resource for Plant Biology (BAR; <http://bar.utoronto.ca>) with the default setting ($P < 0.05$). The classification of DEG sets of 3 through 5 d post infiltration (dpi) is shown.

map was produced (Fig. S6), and hierarchical clustering was performed using the Euclidean distance similarity metric; the map indicated that the transcript abundance of DEGs at 3 dpi is clearly different from that at 4 dpi and onwards, with unique gene groups being upregulated or downregulated in sorgolesone pathway-expressing leaves. To explore these DEGs further, we performed gene ontology (GO) analysis using the CLASSIFICATION SUPERVIEWER tool at the Bio-Analytic Resource for Plant Biology (<http://bar.utoronto.ca>) for enriched GO terms ($P < 0.05$). Within the macro-functional class Biological Processes, the most enriched GO-terms of DEGs were 'metabolic process' and 'cellular process'; among the Molecular Functions, the most frequent terms were 'binding', 'enzyme activity', and 'transferase activity'; the Cellular Component class showed 'cytoplasmic components', 'intracellular components', 'membranes', 'nucleus' and 'chloroplast' as major terms (Fig. 6c). To further clarify the molecular and biological functions of the DEGs, these DEGs were mapped

to KEGG pathways for enrichment analysis. The results showed that several pathways for DEGs of 3 to 6 dpi were significantly enriched, with P -value < 0.05 , including metabolic pathways, biosynthesis of secondary metabolites, photosynthesis, proteasome, and carbon fixation on photosynthetic organisms (Fig. 7).

Differentially expressed genes associated with photosynthesis

Based on GO term enrichment analysis using the CLASSIFICATION SUPERVIEWER tool at the Bio-Analytic Resource for Plant Biology (BAR), the term 'chloroplast' was the most significantly overrepresented (P -value = 4.231×10^{-95} at 3 dpi, 3.991×10^{-106} at 4 dpi, 3.064×10^{-79} at 5 dpi, and 4.245×10^{-19} at 6 dpi). Of these DEGs, 32% at 3 dpi, 35% at 4 dpi, 35% at 5 dpi, and 24% at 6 dpi were associated with the term 'chloroplast' (Fig. 6c; Table S3). Further analysis using G:PROFILER software (Raudvere

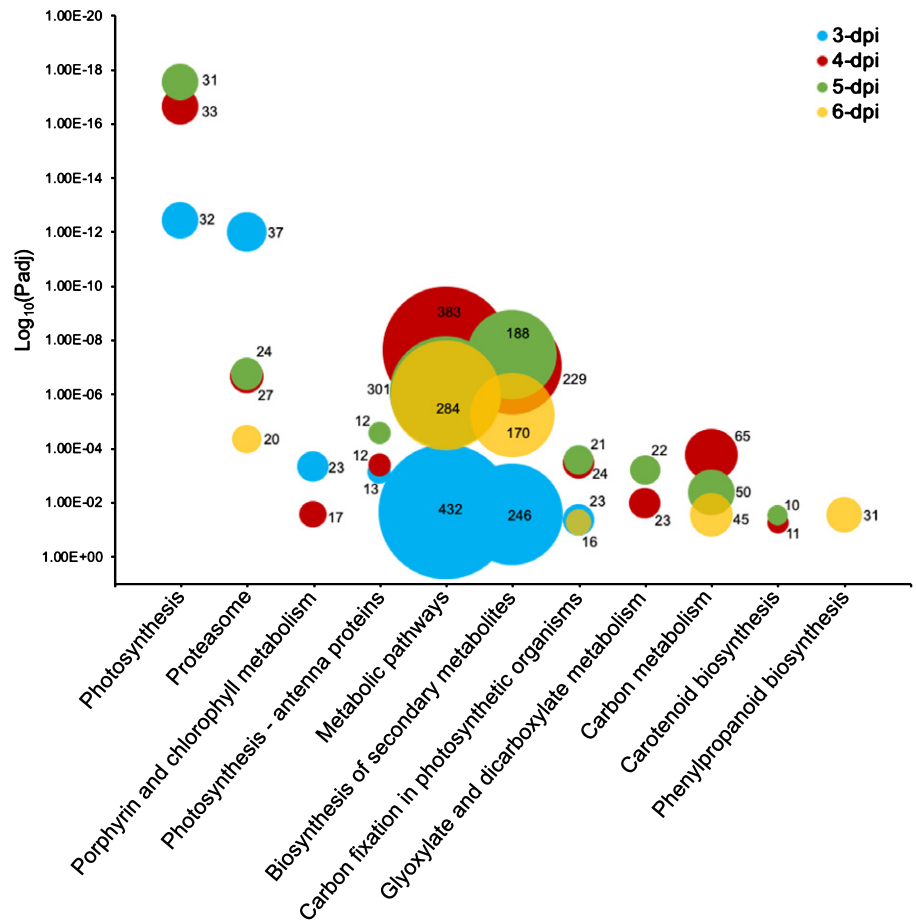


Fig. 7 KEGG pathway classification of significantly enriched differentially expressed genes (DEGs). The statistical enrichment of differential expression genes in KEGG pathways was based on the results of analysis at G:PROFILER (Raudvere *et al.*, 2019) with default settings and threshold *P*-values set at 0.05. The data source for biological pathways was set to the KEGG database (KEGG FTP release 2020-02-03). The numbers of genes are indicated. dpi, days post infiltration.

et al., 2019) indicated that among these chloroplast-associated DEGs, 284 transcripts at one time-point or another from 3 to 5 dpi were assigned to the term ‘photosynthesis’ (this term did not appear at 6 dpi) and almost all of them were downregulated (Table S4). Of these photosynthesis-associated DEGs, 39 genes were mapped to the whole sets of 54 photosynthesis-associated nuclear genes in KEGG pathways (map00195 and map00196). The pathway term ‘photosynthesis’ was the most significantly overrepresented, with a *P*-value of 3.33×10^{-13} for 3 dpi, 2.16×10^{-17} for 4 dpi, and 2.71×10^{-18} for 5 dpi (Table S5). These data sets are composed mostly of genes encoding proteins for photosystem II (PS II), photosystem I (PSI), light-harvesting antennae, cytochrome *b_{6/f}*, and ATP synthase. Their expressions were all downregulated in response to the expression of the sorgoeone biosynthetic pathway (Table 1).

Analyses of other significant differentially expressed genes

As mentioned above, the most affected gene sets are those associated with photosynthesis. Other subsets of the DEGs are also considered here to provide a broader view of the molecular mechanisms involved in the response to the expression of the sorgoeone biosynthetic pathway. The results of the GO term enrichment analyses revealed a substantial number of transcription factors (Fig. 6c; Table S6). We further analyzed the DEGs by comparing the whole set of DEGs with 1716 Arabidopsis

transcription factors (TFs) retrieved from the Plant Transcription Factor Database (Tian *et al.*, 2020). A total of 341 *N. benthamiana* DEGs were identified as being either upregulated or downregulated TF-related genes, with 230, 219, 148 and 175 DEGs for 3, 4, 5 and 6 dpi, respectively (Table S6). Among these TFs, the top 10 families accounted for more than half of the affected genes, and were associated with stress-responsive genes (Baillio *et al.*, 2019; Das *et al.*, 2019), including ERF (ethylene response factor), MYB domain (myeloblastosis), bHLH (basic helix-loop-helix), WRKY (WRKY-domain), NAC (NAM/ATAF/CUC), bZIP, C2H2 (C2H2 zinc finger), GATA/GRAS, ARF (auxin response), and MADS (Fig. 8).

KEGG analysis of the DEGs indicated that metabolic pathways and biosynthesis of secondary metabolites were enriched and comprised a large portion of affected transcripts (Fig. 7). It also revealed that 37 DEGs were associated with proteasomes, which can be mapped to the KEGG proteasome pathway (map03050, consisting of 60 genes) at one time point or another (Table 2). All of them were upregulated and appeared to decrease in abundance from 4 dpi onwards. Of these DEGs, 8 belong to 20S proteasome subunits, 7 to regulatory ATPase, and 4 to 26S regulatory subunits. In addition, the pathway of porphyrin and Chl metabolism (PCM) was also over-represented at 3 and 4 dpi – nearly half of the PCM-associated DEGs were affected (24 out of 53 genes in KEGG pathways; map00860), and almost all of them were downregulated (Table S7). When the analysis was

Table 1 List of photosynthesis-associated differentially expressed genes (DEGs) showing log₂FoldChange.

Protein-accession	Gene ID	Transcript description	3 dpi	4 dpi	5 dpi
Photosystem II (PS II)					
Niben101Scf00879g05002.1	AT1G03600	PS II family protein (PSB27)	-2.053	-3.033	-2.488
Niben101Scf02840g01013.1	AT1G14150	PsbQ-like 2 (PnsL2)	-2.210	-2.677	
Niben101Scf05304g05008.1	AT1G44575	Chl A-B binding family protein (NPQ4)	-1.449	-2.774	-3.608
Niben101Scf05437g04013.1	AT1G67740	PS II BY (PSBY)	-2.004	-2.679	-2.352
Niben101Scf01116g01004.1	AT1G79040	PS II subunit R (PSBR)	-1.145	-2.178	-2.762
Niben101Scf00104g02012.1	AT2G30570	PS II reaction center W (PSBW)	-1.827	-3.290	-2.444
Niben101Scf00578g00008.1	AT3G01440	PsbQ-like 1 (PnsL3)	-2.712	-3.575	-2.210
Niben101Scf02029g00004.1	AT3G50820	PS II subunit O-2 (PSBO2)	-1.922	-2.905	-3.286
Niben101Scf00337g07012.1	AT3G55330	PsbP-like protein 1 (PPL1)	-1.957	-2.180	-1.915
Niben101Scf00176g05004.1	AT4G05180	PS II subunit Q-2 (PSBQ-2)	-1.678	-2.840	-3.467
Niben101Scf03365g01008.1	AT4G28660	PS II reaction center PSB28 protein (PSB28)	-1.543	-2.245	-2.045
Niben101Scf05943g00002.1	AT5G66570	PS II oxygen-evolving complex 1 (PSBO1)	-2.187	-3.059	-3.182
Photosystem I (PS I)					
Niben101Scf02147g01005.1	AT1G03130	PS I subunit D-2 (PSAD-2)	-1.102	-2.429	-2.131
Niben101Scf00851g01001.1	AT1G08380	PS I subunit O (PSAO)	-1.626	-2.541	-2.801
Niben101Scf01797g03034.1	AT1G30380	PS I subunit K (PSAK)	-1.821	-2.844	-3.422
Niben101Scf00271g04024.1	AT1G31330	PS I subunit F (PSAF)	-1.901	-2.768	-2.445
Niben101Scf01147g01007.1	AT1G52230	PS I subunit H2 (PSAH2)	-1.641	-2.553	-3.077
Niben101Scf02293g03024.1	AT1G55670	PS I subunit G (PSAG)	-2.145	-2.940	-2.731
Niben101Scf00597g06001.1	AT2G20260	PS I subunit E-2 (PSAE-2)	-2.180	-2.845	-2.866
Niben101Scf00220g04015.1	AT3G16140	PS I subunit H-1 (PSAH-1)	-1.976	-2.632	-3.225
Niben101Scf01506g02013.1	AT4G02770	PS I subunit D-1 (PSAD-1)	-1.982	-2.502	-1.365
Niben101Scf06919g00006.1	AT4G12800	PS I subunit I (PSAL)	-1.839	-2.579	-3.075
Niben101Scf17701g01020.1	AT5G64040	PS I reaction center subunit (PSAN)	-1.920	-2.965	-3.551
Light-harvesting (LH) antenna in PS II and PS I					
Niben101Scf02128g00025.1	AT1G15820	LH complex PS II subunit 6 (LHCB6)	-2.195	-3.677	-3.676
Niben101Scf00298g02018.1	AT1G19150	PSI type II Chl a/b-binding protein (LHCA6)	-2.591	-2.779	-2.882
Niben101Ctg16210g00008.1	AT1G29920	Chl A/B-binding protein 2 (CAB2)	-2.313	-2.926	-3.520
Niben101Scf01784g05005.1	AT1G45474	PS I LH complex protein 5 (Lhca5)	-1.173		
Niben101Scf01328g01012.1	AT1G61520	PSI type III Chl a/b-binding protein (LHCA3)	-1.982	-2.964	-3.731
Niben101Scf01548g00002.1	AT2G05100	PS II LH complex protein 2.1 (LHCB2.1)	-1.794	-2.720	-3.413
Niben101Scf08975g01010.1	AT3G08940	LH complex PS II (LHCB4.2)	-2.188	-2.996	-3.145
Niben101Scf01020g00024.1	AT3G27690	PS II LH complex protein 2.3 (LHCB2.3)	-1.938	-2.539	-2.245
Niben101Scf00406g03001.1	AT3G54890	Chl a-b binding protein 6 (LHCA1)	-2.165	-3.006	-3.724
Niben101Ctg06879g00004.1	AT3G61470	PS I LH complex protein (LHCA2)	-2.254	-2.820	-3.482
Niben101Scf02459g00024.1	AT4G10340	LH complex of PS II 5 (LHCB5)	-1.999	-2.853	-2.993
Niben101Scf02971g05016.1	AT5G54270	LH Chl B-binding protein 3 (LHCB3)	-1.792	-2.752	-3.372
Cytochrome <i>b6f</i>					
Niben101Scf03929g04002.1	AT4G03280	Photosynthetic electron transfer C (PETC)	-1.390	-2.403	-2.697
ATP synthase					
Niben101Scf01399g00012.1	AT4G04640	F1 complex (ATPC1)	-1.836	-2.648	-2.101
Niben101Scf15372g00003.1	AT4G09650	F-type H ⁺ -transporting subunit (ATPD/PDE332)	-1.693	-2.318	-1.625
Niben101Scf04786g01004.1	AT4G32260	F0 complex, subunit B/B' (PDE334)	-1.420	-2.342	-2.596

These genes were mapped to the pathways (map00195 and map00196) in the KEGG database. dpi, days post-infiltration.

extended to the genes related to electron transport, via the term 'photosynthetic electron transport chain', 31 DEGs were identified and all of them were significantly downregulated (Table S8). Finally, DEGs involved in carbon metabolism were also affected significantly, with *c.* 24% of total term-size (Table S5) at 4 dpi, and over half of them were downregulated (Fig. 7; Table S9).

Discussion

Although the candidate genes encoding the enzymes required for the biosynthetic pathway of sorgoleone, a lipid resorcinol-type allelochemical, have been identified and functionally characterized (Pan *et al.*, 2007; Baerson *et al.*, 2008; Cook *et al.*, 2010;

Pan *et al.*, 2018), the simultaneous expression of all of the genes required for the functional assembly of the entire pathway *in planta* has not been reported. In this study, we successfully performed *Agrobacterium*-mediated transient expression assays with *N. benthamiana*, using a multigene vector containing all of the genes required for sorgoleone biosynthesis from palmitoleoyl-CoA (Fig. 1). The visible onset of necrotic lesions in *N. benthamiana* leaves expressing the sorgoleone biosynthetic genes started at 3 d after agroinfiltration. Analysis of the transiently infected *N. benthamiana* leaves by GC/MS revealed the presence of sorgoleone in the leaf samples, consistent with the plasmid-dependent accumulation of this compound. The accumulation of sorgoleone in *N. benthamiana* leaves apparently caused the

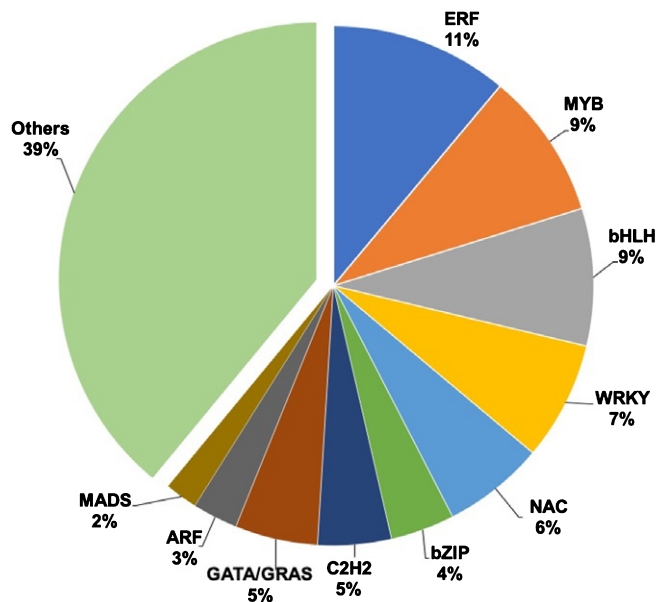


Fig. 8 Major families of transcription factors (TFs) found in differentially expressed genes (DEGs). The area of the pie chart is scaled by total TF-associated DEGs. ARF, auxin response; bHLH, basic helix-loop-helix; bZIP, basic leucine zipper domain proteins; C2H2, Cys2His2-like zinc finger proteins; GATA/GRAS, GATA/GRAS domain family proteins; MADS, MADS-domain proteins; MYB, myb domain proteins; NAC, NAM/ATAF/CUC domain proteins; WRKY, WRKY-domain proteins.

observed necrotic lesions, indicating its phytotoxicity in plant cells.

In plants, fatty acid biosynthesis mainly takes place in chloroplasts and the plastids of nonphotosynthetic tissues. Some acyl-carrier protein-bound C16:0 fatty acid is then released and exported to the ER as a pool of fatty acids for further elongation, acyl editing, and lipid assembly (N. Li *et al.*, 2016; W. Li *et al.*, 2016; LaBrant *et al.*, 2018). In the case of biosynthesis of sorgoleone in sorghum root-hair cells, the lipid tail of sorgoleone is likely synthesized in plastids and transported to the endoplasmic reticulum (ER) where it is further processed by downstream enzymes (Czarnota *et al.*, 2003; Weston *et al.*, 2013). The first desaturation of C16:0 fatty acid takes place either in plastids or the ER – it is currently unclear which – by the action of the Δ^9 stearoyl-ACP desaturase to produce oleoyl-ACP, which is subsequently released from acyl-carrier proteins as CoA esters (e.g. palmitoleoyl-CoA (C16:1 Δ^9 fatty acid)). The synthesis of sorgoleone in this transient expression system may have taken place in the ER since none of the five genes possess chloroplast transit peptide signals (Pan *et al.*, 2007, 2018; Baerson *et al.*, 2008; Cook *et al.*, 2010). The synthesis and mechanism of trafficking/transiting of the palmitoleoyl-CoA starter between the chloroplasts and the ER, as well as the site(s) of enzymatic reactions in *N. benthamiana* leaf cells, remain unknown. Nevertheless, our results show that the pathway was successfully assembled *in planta*, in which it utilized the endogenous C16:1 fatty acid as a starter. Importantly, the *in vivo*-produced sorgoleone causes necrotic lesions, indicating that the compound possesses cytotoxicity, one of the characteristics of sorgoleone and other allelochemicals.

To gain insight into the molecular basis of the impact by expressing the sorgoleone pathway, we performed RNA-sequencing and compared the changes in global gene expression between the multigene vector containing all the sorgoleone genes and the empty vector control. Transient expression of the pathway in *N. benthamiana* leaves markedly affected the expression of a large number of genes (Table S1). Based on the annotation and classification of *N. benthamiana* DEGs that were assigned to the Arabidopsis genome, gene set enrichment analyses using GO ontologies and KEGG pathways resulted in a large number of significantly over-represented terms, including ‘metabolic process’, ‘chloroplast’, ‘transcription regulators’, ‘photosynthesis’, ‘proteasome’, etc. (Figs 6, 7).

Photosynthesis is one of the most important biological processes in plants. One of the major responses of *N. benthamiana* leaves to sorgoleone produced by expressing its biosynthesis pathway is a significant decrease in the abundance of the transcripts for the protein assembly of the photosynthesis apparatus. It has been demonstrated that PsbP protein is essential for assembly, regulation, and stability of PSII (Ifuku *et al.*, 2005; Yi *et al.*, 2006, 2007); PsbQ protein is required for the assembly/stability of PSII under prolonged low light conditions (Yi *et al.*, 2006); and PsbO is required for the accumulation and stabilization of the PSII core complex (Murakami *et al.*, 2002). Furthermore, knockdown of PsbP in PSII significantly affects the assembly and accumulation of the PSI supercomplex and alters the thylakoid grana stacking (Ido *et al.*, 2009). Transcriptome analysis of *N. benthamiana* transiently expressing the sorgoleone pathway showed that these PSII-related genes were downregulated in response to sorgoleone (Table 1). Moreover, many transcripts encoding members of the photosynthesis machinery and assimilatory metabolism were significantly enriched among the DEGs and decreased in abundance in the presence of sorgoleone (Tables 1, S4, S9). Taken together, this may explain the lesion phenotype of leaves agroinfiltrated with the multigene construct. Sorgoleone may therefore contribute negatively to the differential expression of photosynthetic genes in a way that can potentially compromise the assembly of the photosynthesis apparatus, leading to diminished photosynthetic activity and ultimately causing leaf necrosis.

Most of the autotoxic effects observed in plant cells producing sorgoleone are probably secondary or tertiary effects of sorgoleone acting on a primary target such as the PsbA (D1) protein of PSII, the most likely target protein of sorgoleone as an allelochemical (Weston & Czarnota, 2001; Dayan *et al.*, 2009). Sorgoleone is as active on isolated chloroplasts in stopping photosynthetic electron flow as the PSII-inhibiting herbicide DCMU (3-(3,4-dichlorophenyl)-1,1-dimethylurea; Gonzalez *et al.*, 1997). Inhibition of electron flow of PSII, which would reduce or stop all photosynthetic processes, would be expected to cause a corresponding alteration of the expression of genes associated with these processes. Our transcriptome results support PSII as at least one primary cause of many of the effects seen on the expression of genes associated with photosynthesis (e.g. Table 1). Other effects on gene expression could be more directly

Table 2 List of proteasome-associated differentially expressed genes (DEGs).

Protein-accession	Gene ID	Gene name and description	3 dpi	4 dpi	5 dpi	6 dpi
Niben101Scf01738g04008.1	AT3G22110	20S proteasome alpha subunit C1 (PAC1)	2.272	1.456	1.189	1.284
Niben101Scf00764g03003.1	AT3G14290	20S proteasome alpha subunit E2 (PAE2)	1.510			
Niben101Scf06716g00006.1	AT2G27020	20S proteasome alpha subunit G1 (PAG1)	2.720	1.800	1.530	1.412
Niben101Scf02025g00002.1	AT3G51260	20S proteasome alpha subunit PAD1 (PAD1)	2.936	1.854	1.493	1.492
Niben101Scf09547g02002.1	AT3G22630	20S proteasome beta subunit D1 (PBD1)	2.587	1.366	1.095	
Niben101Scf04549g09016.1	AT1G56450	20S proteasome beta subunit G1 (PBG1)	2.377	1.445	1.075	1.148
Niben101Scf00231g02003.1	AT5G40580	20S proteasome beta subunit PBB2 (PBB2)	2.501	2.314	1.941	
Niben101Scf02460g03001.1	AT2G05840	20S proteasome subunit PAA2 (PAA2)	2.317			
Niben101Scf01171g08017.1	AT1G04810	26S proteasome regulatory complex, Rpn2/Psmd1	4.331	4.238	3.522	3.326
Niben101Scf01146g02016.1	AT2G32730	26S proteasome regulatory complex, Rpn2/Psmd1	1.501			
Niben101Scf01052g06004.1	AT4G24820	26S proteasome regulatory subunit Rpn7	1.974	1.166		
Niben101Scf01462g02022.1	AT4G28470	26S proteasome regulatory subunit S2 1B (RPN1B)	2.222	1.645	1.357	1.210
Niben101Scf04861g00014.1	AT1G45000	AAA-type ATPase family protein	2.400	1.730	1.276	1.191
Niben101Scf07364g00017.1	AT5G23540	Mov34/MPN/PAD-1 family protein	1.729			
Niben101Scf03035g05011.1	AT1G79210	Ntn hydrolases superfamily protein	2.611	1.454	1.199	1.253
Niben101Scf02487g00001.1	AT3G26340	Ntn hydrolases superfamily protein	1.108			
Niben101Scf00017g00005.1	AT4G31300	Ntn hydrolases superfamily protein (PBA1)	2.961	1.782	1.212	1.023
Niben101Scf06600g00001.1	AT3G27430	Ntn hydrolases superfamily protein (PBB1)	3.338	2.249	1.748	1.542
Niben101Scf02864g01014.1	AT3G60820	Ntn hydrolases superfamily protein (PBF1)	2.239	1.152	1.102	1.217
Niben101Scf00193g01010.1	AT1G29150	Non-ATPase subunit 9	1.067			
Niben101Scf01718g02010.1	AT1G75990	PAM domain (PCI/PINT) protein	1.509			
Niben101Scf00780g03004.1	AT1G20200	PAM domain (PCI/PINT) protein (EMB2719)	2.402	1.783	1.318	1.121
Niben101Scf00654g04006.1	AT3G13330	Proteasome activating protein 200 (PA200)	2.335	2.103	1.883	1.383
Niben101Scf02594g06002.1	AT5G42790	Proteasome alpha subunit F1 (PAF1)	2.160	1.330		
Niben101Scf00986g02003.1	AT1G21720	Proteasome beta subunit C1 (PBC1)	2.803	1.879	1.538	1.498
Niben101Scf01671g06021.1	AT4G19006	Proteasome component (PCI) domain protein	2.021	1.476	1.265	1.018
Niben101Scf00612g02003.1	AT5G45620	Proteasome component (PCI) domain protein	2.882	2.090	1.770	1.658
Niben101Scf03147g07002.1	AT3G53970	Proteasome inhibitor-like protein	2.978	1.868	1.253	1.189
Niben101Scf01068g05007.1	AT1G16470	Proteasome subunit PAB1 (PAB1)	3.675	2.640	2.340	2.181
Niben101Scf00597g03003.1	AT4G29040	Regulatory particle AAA-ATPase 2A (RPT2a)	1.091			
Niben101Scf00674g06001.1	AT4G38630	Regulatory particle non-ATPase 10 (RPN10)	2.138	1.578	1.008	
Niben101Scf00132g03012.1	AT1G64520	Regulatory particle non-ATPase 12A (RPN12a)	2.749	2.011	1.715	1.502
Niben101Scf00182g03011.1	AT2G26590	Regulatory particle non-ATPase 13 (RPN13)	1.237	1.040		
Niben101Scf00747g12004.1	AT1G53750	Regulatory particle triple-A 1A (RPT1A)	1.813	1.191		
Niben101Scf00386g00002.1	AT5G58290	Regulatory particle triple-A ATPase 3 (RPT3)	1.212		1.013	
Niben101Scf00369g09017.1	AT3G05530	Regulatory particle triple-A ATPase 5A (RPT5A)	2.498	1.765	1.402	1.457
Niben101Scf00574g11010.1	AT5G05780	RP non-ATPase subunit 8A (RPN8A)	1.478			

dpi, days post-infiltration.

associated with other protein targets of sorgoleone (e.g. HPPD) or to secondary effects of disrupted photosynthesis.

The ubiquitin–proteasome system plays vital roles in diverse plant developmental and environmental responses, stress responses, and the maintenance of protein homeostasis (Gladman *et al.*, 2016; Marshall & Vierstra, 2019). It also functions as a critical transcriptional co-regulator to coordinate the expression of developmental and stress-responsive genes (Geng *et al.*, 2012; Kelley & Estelle, 2012). The proteasome is a multisubunit and multicatalytic protease complex responsible for degrading a wide range of intracellular proteins, an essential process in eukaryotes. In our transcriptome data, we identified a large number of transcripts corresponding to the proteasome subunit core complex. These transcripts were all upregulated in response to the expression of the sorgoleone biosynthetic pathway (Table 2). This finding suggests that sorgoleone may activate the transcription factors involved in the expression of genes required for the assembly of the proteasome supercomplex through an as-yet unidentified mechanism. The involvement of proteasomes in the response to

sorgoleone is new and requires further investigation. One hypothesis is that the presence of sorgoleone in the cells might coordinately regulate gene expression, specifically the set of genes encoding proteins for proteasome assembly, which could cause impaired chloroplast development or stability and, ultimately, necrosis, since excess proteasome subunits apparently do not typically accumulate within cells in free forms. Nevertheless, studies are needed to address the concerted transcriptional regulation of these proteasome genes during the development of the response to sorgoleone.

Transcription factors play central roles in all aspects of plant function, including cell differentiation, organ development, responses to environmental factors/stresses, etc., and coordinate gene expression patterns that give rise to specific phenotypic outputs. Transcription factors modulate gene expression by binding to specific DNA sequences, interact with various proteins in transcriptional complexes of a given gene, and regulate gene expression in certain biological contexts. In *Arabidopsis*, > 50 TF families are predicted to contain 1716 TFs (Jin *et al.*, 2017). In surveying our

N. benthamiana DEG datasets, we found 299 TF homologs covering 42 families that were significantly affected by the expression of the sorgoleone biosynthetic pathway, of which approx. two-thirds were downregulated (Table S6). Several TF families, including WRKY, MYB, NAC, bZIP, bHLH and ERF, have been implicated in stress responses across many plant species (Das *et al.*, 2019). About 50% of the TFs affected by the expression of the sorgoleone genes belong to these families. These results could indicate that synthesized sorgoleone in agroinfiltrated *N. benthamiana* leaves triggers a massive stress response. Further, the expression of genes for the development of chloroplasts and the components of photosynthesis has been shown to require coordinated control of sets of transcription factors (Wang *et al.*, 2017). Transcription factors can positively or negatively regulate the expression of their target genes. For example, transcription factor Golden2-like (GLK; Niben101Scf01462g01010.1) was significantly downregulated at 3 and 4 dpi (Table S6), which could positively regulate both chloroplast development and expression of photosynthesis genes by binding directly to promoters of the genes encoding light-harvesting complex proteins, as well as genes encoding key enzymes of the Chl biosynthesis pathway (Waters *et al.*, 2009; Wang *et al.*, 2017). Downregulation of such TFs could therefore lead to defects in chloroplast development and deterioration of photosynthetic function, and ultimately necrotic lesions. However, this regulatory mechanism remains to be elucidated. Similarly, upregulated genes encoding proteasome proteins by the set of specific TFs may also contribute to leaf necrotic lesions upon exposure to sorgoleone, although the connections between proteasome activity and the lesions are currently unknown and need further investigation.

In summary, as a proof of concept, we have successfully assembled the sorgoleone biosynthetic pathway by transient expression in *N. benthamiana*. Sorgoleone, produced *in vivo*, not only caused a severe necrotic phenotype but also significantly impacted gene expression, including genes associated with photosynthetic machinery, TFs, and the proteasome. Our results showed that the degree of necrosis was correlated with transcription profiling (i.e. gene expression during the development of the symptom). The current study strongly suggests that stably transformed plants harboring transgenes encoding the sorgoleone biosynthetic enzymes (SbDES2, SbDES3, SbARS1, SbOMT3 and CYP71AM1) will likely be capable of performing the *de novo* biosynthesis of sorgoleone. Work is underway to generate Gramineae cereal species expressing the genes for this compound using in-house cloned root-hair specific expression promoters for weed management.

Acknowledgements






We thank Dr Ammeeta Agarwal for assistance with RNA sequencing, and Marilyn Ruscoe and Gloria Hervey for providing technical support for this work.

Author contributions

ZP, SRB and SOD designed the study; ZP, JB-H and AMR performed all experiments; AMR performed the chemical analysis;

JNV performed RNA-seq data processing; ZP analyzed the data and wrote the manuscript.

ORCID

Scott R. Baerson  <https://orcid.org/0000-0002-0756-6379>
 Joanna Bajsa-Hirschel  <https://orcid.org/0000-0001-7607-5808>
 Stephen O. Duke  <https://orcid.org/0000-0001-7210-5168>
 Zhiqiang Pan  <https://orcid.org/0000-0003-0771-8261>
 Justin N. Vaughn  <https://orcid.org/0000-0002-5261-5096>

References

- Baerson SR, Dayan FE, Rimando AM, Nanayakkara NP, Liu CJ, Schroder J, Fishbein M, Pan Z, Kagan IA, Pratt LH *et al.* 2008. A functional genomics investigation of allelochemical biosynthesis in *Sorghum bicolor* root hairs. *Journal of Biological Chemistry* 283: 3231–3247.
- Baillo EH, Kimotho RN, Zhang Z, Xu P. 2019. Transcription factors associated with abiotic and biotic stress tolerance and their potential for crops improvement. *Genes (Basel)* 10: 771.
- Bally J, Jung H, Mortimer C, Naim F, Philips JG, Hellens R, Bombarely A, Goodin MM, Waterhouse PM. 2018. The rise and rise of *Nicotiana benthamiana*: a plant for all reasons. *Annual Review of Phytopathology* 56: 405–426.
- Bustin SA, Beaulieu JF, Huggett J, Jaggi R, Kibenge FS, Olsvik PA, Penning LC, Toegel S. 2010. MIQE precis: practical implementation of minimum standard guidelines for fluorescence-based quantitative real-time PCR experiments. *BMC Molecular Biology* 11: 74.
- Bustin SA, Benes V, Garson JA, Hellemans J, Huggett J, Kubista M, Mueller R, Nolan T, Pfaffl MW, Shipley GL *et al.* 2009. The MIQE guidelines: minimum information for publication of quantitative real-time PCR experiments. *Clinical Chemistry* 55: 611–622.
- Chung SM, Frankman EL, Tzfira T. 2005. A versatile vector system for multiple gene expression in plants. *Trends in Plant Science* 10: 357–361.
- Cook D, Rimando AM, Clemente TE, Schroder J, Dayan FE, Nanayakkara NP, Pan Z, Noonan BP, Fishbein M, Abe I *et al.* 2010. Alkylresorcinol synthases expressed in *Sorghum bicolor* root hairs play an essential role in the biosynthesis of the allelopathic benzoquinone sorgoleone. *Plant Cell* 22: 867–887.
- Czarnota MA, Paul RN, Dayan FE, Nimbal CI, Weston LA. 2001. Mode of action, localization of production, chemical nature, and activity of sorgoleone: a potent PSII inhibitor in *Sorghum* spp. root exudates. *Weed Technology* 15: 813–825.
- Czarnota MA, Paul RN, Weston LA, Duke SO. 2003. Anatomy of sorgoleone-secreting root hairs of sorghum species. *International Journal of Plant Sciences* 164: 861–866.
- Das A, Pramanik K, Sharma R, Gantait S, Banerjee J. 2019. In-silico study of biotic and abiotic stress-related transcription factor binding sites in the promoter regions of rice germin-like protein genes. *PLoS ONE* 14: e0211887.
- Dayan FE, Howell J, Weidenhamer JD. 2009. Dynamic root exudation of sorgoleone and its *in planta* mechanism of action. *Journal of Experimental Botany* 60: 2107–2117.
- Dayan FE, Kagan IA, Rimando AM. 2003. Elucidation of the biosynthetic pathway of the allelochemical sorgoleone using retrobiosynthetic NMR analysis. *Journal of Biological Chemistry* 278: 28607–28611.
- Duke SO. 2003. Weeding with transgenes. *Trends in Biotechnology* 21: 192–195.
- Einhellig FA, Rasmussen JA. 1989. Prior cropping with grain sorghum inhibits weeds. *Journal of Chemical Ecology* 15: 951–960.
- Einhellig FA, Rasmussen JA, Hejl AM, Souza IF. 1993. Effects of root exudate sorgoleone on photosynthesis. *Journal of Chemical Ecology* 19: 369–375.
- Fate GD, Lynn DG. 1996. Xenogonin methylation is critical in defining the chemical potential gradient that regulates the spatial distribution in *Striga* pathogenesis. *Journal of American Chemical Society* 118: 11369–11376.

- Forney DR, Foy CL, Wolf DD. 1985. Weed suppression in no-till alfalfa (*Medicago sativa*) by prior cropping of summer-annual forage grasses. *Weed Science* 33: 490–497.
- Garabagi F, Gilbert E, Loos A, McLean MD, Hall JC. 2012. Utility of the P19 suppressor of gene-silencing protein for production of therapeutic antibodies in *Nicotiana* expression hosts. *Plant Biotechnology Journal* 10: 1118–1128.
- Geng F, Wenzel S, Tansey WP. 2012. Ubiquitin and proteasomes in transcription. *Annual Review of Biochemistry* 81: 177–201.
- Gladman NP, Marshall RS, Lee KH, Vierstra RD. 2016. The proteasome stress regulon is controlled by a pair of NAC transcription factors in *Arabidopsis*. *Plant Cell* 28: 1279–1296.
- Gonzalez VM, Kazimir J, Nimal CI, Weston LA, Cheniae GM. 1997. Inhibition of a photosystem II electron transfer reaction by the natural product sorgoleone. *Journal of Agricultural and Food Chemistry* 45: 1415–1421.
- Hofgen R, Willmitzer L. 1988. Storage of competent cells for *Agrobacterium* transformation. *Nucleic Acids Research* 16: 9877.
- Ido K, Ifuku K, Yamamoto Y, Ishihara S, Murakami A, Takabe K, Miyake C, Sato F. 2009. Knockdown of the PsbP protein does not prevent assembly of the dimeric PSII core complex but impairs accumulation of photosystem II supercomplexes in tobacco. *Biochimica et Biophysica Acta* 1787: 873–881.
- Ifuku K, Yamamoto Y, Ono TA, Ishihara S, Sato F. 2005. PsbP protein, but not PsbQ protein, is essential for the regulation and stabilization of photosystem II in higher plants. *Plant Physiology* 139: 1175–1184.
- Jin J, Tian F, Yang DC, Meng YQ, Kong L, Luo J, Gao G. 2017. PlantTFDB 4.0: toward a central hub for transcription factors and regulatory interactions in plants. *Nucleic Acids Research* 45: D1040–D1045.
- Kagan IA, Rimando AM, Dayan FE. 2003. Chromatographic separation and *in vitro* activity of sorgoleone congeners from the roots of *Sorghum bicolor*. *Journal of Agricultural and Food Chemistry* 51: 7589–7595.
- Kelley DR, Estelle M. 2012. Ubiquitin-mediated control of plant hormone signaling. *Plant Physiology* 160: 47–55.
- LaBrant E, Barnes AC, Roston RL. 2018. Lipid transport required to make lipids of photosynthetic membranes. *Photosynthesis Research* 138: 345–360.
- Li N, Xu C, Li-Beisson Y, Philippar K. 2016. Fatty acid and lipid transport in plant cells. *Trends in Plant Science* 21: 145–158.
- Li W, Herrera-Estrella L, Tran LP. 2016. The yin-yang of cytokinin homeostasis and drought acclimation/adaptation. *Trends in Plant Science* 21: 548–550.
- Marshall RS, Vierstra RD. 2019. Dynamic regulation of the 26S proteasome: from synthesis to degradation. *Frontiers in Molecular Biosciences* 6: 40.
- Meazza G, Scheffler BE, Tellez MR, Rimando AM, Romagni JG, Duke SO, Nanayakkara D, Khan IA, Abourashed EA, Dayan FE. 2002. The inhibitory activity of natural products on plant p-hydroxyphenylpyruvate dioxygenase. *Phytochemistry* 60: 281–288.
- Murakami R, Ifuku K, Takabayashi A, Shikanai T, Endo T, Sato F. 2002. Characterization of an *Arabidopsis thaliana* mutant with impaired *psbO*, one of two genes encoding extrinsic 33-kDa proteins in photosystem II. *FEBS Letter* 523: 138–142.
- Netzly DH, Butler LG. 1986. Roots of sorghum exude hydrophobic droplets containing biologically active components. *Crop Science* 26: 775–778.
- Pan Z, Baerson SR, Wang M, Bajsa-Hirschel J, Rimando AM, Wang X, Nanayakkara NPD, Noonan BP, Fromm ME, Dayan FE *et al.* 2018. A cytochrome P450 CYP71 enzyme expressed in *Sorghum bicolor* root hair cells participates in the biosynthesis of the benzoquinone allelochemical sorgoleone. *New Phytologist* 218: 616–629.
- Pan Z, Rimando AM, Baerson SR, Fishbein M, Duke SO. 2007. Functional characterization of desaturases involved in the formation of the terminal double bond of an unusual 16:3 $\Delta^{9,12,15}$ fatty acid isolated from *Sorghum bicolor* root hairs. *Journal of Biological Chemistry* 282: 4326–4335.
- Putnam AR, Defrank J, Barnes JP. 1983. Exploitation of allelopathy for weed control in annual and perennial cropping systems. *Journal of Chemical Ecology* 9: 1001–1010.
- Rasmussen JA, Hejl AM, Einhellig FA, Thomas JA. 1992. Sorgoleone from root exudate inhibits mitochondrial functions. *Journal of Chemical Ecology* 18: 197–207.
- Raudvere U, Kolberg L, Kuzmin I, Arak T, Adler P, Peterson H, Vilo J. 2019. g:Profiler: a web server for functional enrichment analysis and conversions of gene lists (2019 update). *Nucleic Acids Research* 47: W191–W198.
- Sainsbury F, Lomonossoff GP. 2014. Transient expressions of synthetic biology in plants. *Current Opinion in Plant Biology* 19: 1–7.
- Tian F, Yang DC, Meng YQ, Jin J, Gao G. 2020. PlantRegMap: charting functional regulatory maps in plants. *Nucleic Acids Research* 48: D1104–D1113.
- Wang P, Hendron RW, Kelly S. 2017. Transcriptional control of photosynthetic capacity: conservation and divergence from *Arabidopsis* to rice. *New Phytologist* 216: 32–45.
- Waters MT, Wang P, Korkaric M, Capper RG, Saunders NJ, Langdale JA. 2009. GLK transcription factors coordinate expression of the photosynthetic apparatus in *Arabidopsis*. *Plant Cell* 21: 1109–1128.
- Weston LA, Alsaadawi IS, Baerson SR. 2013. Sorghum allelopathy – from ecosystem to molecule. *Journal of Chemical Ecology* 39: 142–153.
- Weston LA, Czarnota MA. 2001. Activity and persistence of sorgoleone, a long-chain hydroquinone produced by *Sorghum bicolor*. *Journal of Crop Production* 4: 363–377.
- Wydro M, Kozubek E, Lehmann P. 2006. Optimization of transient *Agrobacterium*-mediated gene expression system in leaves of *Nicotiana benthamiana*. *Acta Biochimica Polonica* 53: 289–298.
- Yang X, Scheffler BE, Weston LA. 2004. *SORI*, a gene associated with bioherbicide production in sorghum root hairs. *Journal of Experimental Botany* 55: 2251–2259.
- Yi X, Hargett SR, Frankel LK, Bricker TM. 2006. The PsbQ protein is required in *Arabidopsis* for photosystem II assembly/stability and photoautotrophy under low light conditions. *Journal of Biological Chemistry* 281: 26260–26267.
- Yi X, Hargett SR, Liu H, Frankel LK, Bricker TM. 2007. The PsbP protein is required for photosystem II complex assembly/stability and photoautotrophy in *Arabidopsis thaliana*. *Journal of Biological Chemistry* 282: 24833–24841.

Supporting Information

Additional Supporting Information may be found online in the Supporting Information section at the end of the article.

Fig. S1 Assembly of expression cassettes.

Fig. S2 A representative *c.* 4-wk-old *Nicotiana benthamiana* plant used in the experiments.

Fig. S3 The expression of *Sorghum bicolor* genes in agroinfiltrated *N. benthamiana* leaves as detected by quantitative reverse transcription-polymerase chain reaction (RT-qPCR).

Fig. S4 Expression of all transgenes of the sorgoleone pathway in agroinfiltrated *N. benthamiana* leaves.

Fig. S5 A representative image of gas chromatography-mass spectroscopy (GC-MS) analysis of agroinfiltrated *N. benthamiana* leaves.

Fig. S6 Cluster analysis of differentially expressed genes.

Methods S1 Construction of binary vectors for transient expression of genes for the biosynthesis of sorgoleone in *N. benthamiana*.

Methods S2 Real-time quantitative reverse transcription-polymerase chain reaction.

Table S1 List of DEGs that were significantly changed (more than twofold) at all time points.

Table S2 List of DEGs common to all time points.

Table S3 GO term enrichment analysis.

Table S4 List of DEGs associated with photosynthesis.

Table S5 KEGG pathway analysis of the DEGs.

Table S6 List of DEGs associated with transcription factors.

Table S7 List of DEGs associated with the term 'porphyrin and chlorophyll metabolism'.

Table S8 List of DEGs associated with the term 'photosynthetic electron transfer chain'.

Table S9 List of DEGs associated with KEGG term 'carbon metabolism'.

Please note: Wiley Blackwell are not responsible for the content or functionality of any Supporting Information supplied by the authors. Any queries (other than missing material) should be directed to the *New Phytologist* Central Office.



About *New Phytologist*

- *New Phytologist* is an electronic (online-only) journal owned by the New Phytologist Foundation, a **not-for-profit organization** dedicated to the promotion of plant science, facilitating projects from symposia to free access for our Tansley reviews and Tansley insights.
- Regular papers, Letters, Viewpoints, Research reviews, Rapid reports and both Modelling/Theory and Methods papers are encouraged. We are committed to rapid processing, from online submission through to publication 'as ready' via *Early View* – our average time to decision is <26 days. There are **no page or colour charges** and a PDF version will be provided for each article.
- The journal is available online at Wiley Online Library. Visit **www.newphytologist.com** to search the articles and register for table of contents email alerts.
- If you have any questions, do get in touch with Central Office (np-centraloffice@lancaster.ac.uk) or, if it is more convenient, our USA Office (np-usaoffice@lancaster.ac.uk)
- For submission instructions, subscription and all the latest information visit **www.newphytologist.com**