Journal of
Experimental
Botany
www.jxb.oxfordjournals.org

RESEARCH PAPER

Lignin metabolism has a central role in the resistance of cotton to the wilt fungus *Verticillium dahliae* as revealed by RNA-Seq-dependent transcriptional analysis and histochemistry

Li Xu, Longfu Zhu, Lili Tu, Linlin Liu, Daojun Yuan, Li Jin, Lu Long and Xianlong Zhang*

National Key Laboratory of Crop Genetic Improvement, Huazhong Agricultural University, Wuhan, Hubei, PR China

* To whom correspondence should be addressed. E-mail: xlzhang@mail.hzau.edu.cn

Abstract

The incompatible pathosystem between resistant cotton (*Gossypium barbadense* cv. 7124) and *Verticillium dahliae* strain V991 was used to study the cotton transcriptome changes after pathogen inoculation by RNA-Seq. Of 32 774 genes detected by mapping the tags to assembly cotton contigs, 3442 defence-responsive genes were identified. Gene cluster analyses and functional assignments of differentially expressed genes indicated a significant transcriptional complexity. Quantitative real-time PCR (qPCR) was performed on selected genes with different expression levels and functional assignments to demonstrate the utility of RNA-Seq for gene expression profiles during the cotton defence response. Detailed elucidation of responses of leucine-rich repeat receptor-like kinases (LRR-RLKs), phytohormone signalling-related genes, and transcription factors described the interplay of signals that allowed the plant to fine-tune defence responses. On the basis of global gene regulation of phenylpropanoid metabolism-related genes, phenylpropanoid metabolism was deduced to be involved in the cotton defence response. A closer look at the expression of these genes, enzyme activity, and lignin levels revealed differences between resistant and susceptible cotton plants. Both types of plants showed an increased level of expression of lignin synthesis-related genes and increased phenylalanine-ammonia lyase (PAL) and peroxidase (POD) enzyme activity after inoculation with *V. dahliae*, but the increase was greater and faster in the resistant line. Histochemical analysis of lignin revealed that the resistant cotton not only retains its vascular structure, but also accumulates high levels of lignin. Furthermore, quantitative analysis demonstrated increased lignification and cross-linking of lignin in resistant cotton stems. Overall, a critical role for lignin was believed to contribute to the resistance of cotton to disease.

Key words: Cotton, defence response, phenylpropanoid, lignin, RNA-Seq, signal transduction, *Verticillium dahliae*.

Introduction

Deep-sequencing technologies have become a revolutionary tool to better understand the complicated eukaryote transcriptomes, and they include three widely used deep-sequencing platforms throughout the world, the Roche 454 Life Sciences, the Illumina Genome Analyzer, and Applied Biosystems' SOLiD (Ansorge, 2009). Highly specific, sensitive, and quantitative measurements enable the deep-sequencing technologies to overcome the shortcomings of traditional hybridization-based approaches. Different from the traditional hybridization-based approaches, high-throughput sequencing technologies, referring to as RNA-Seq, have much greater power to distinguish between paralogous genes, detect low or high abundance transcripts, and allow replicate quantification based on the number of sequences obtained (Wang *et al.*, 2009). Furthermore, RNA-Seq can identify transcript sequence polymorphisms and novel trans-splicing and splice isoforms; in addition, there is no strict requirement for a reference genome

sequence (Bethel *et al.*, 2009). These features make RNA-Seq available to study the non-model species with genomic sequences that are yet to be determined.

Several reports based on RNA-Seq in human, mouse, yeast, *Drosophila melanogaster*, *Arabidopsis thaliana*, and *Vitis vinifera* have been published (Cloonan *et al.*, 2008; Mortazavi *et al.*, 2008; Nagalakshmi *et al.*, 2008; Gan *et al.*, 2010; Otto *et al.*, 2010; Zenoni *et al.*, 2010). Recently, RNA-Seq was also reported to analyse non-model species by mapping to a sequence database or a characterized genome of related organisms (Xu *et al.*, 2009; Birzele *et al.*, 2010). Millions of short tags were generated using the platform for interpreting the genome and revealing the transcriptomic changes during development.

Cotton (*Gossypium* spp.) is widely cultivated for the important economic value of its fibre. The cotton acreage in China reached 4.85 million ha and the yield reached 5.97 million tons in 2010. Cotton Verticillium wilt caused by *Verticillium dahliae* is a soil-borne vascular disease (Sal'kova and Guseva, 1965). The representative symptoms caused by *V. dahliae* in the susceptible cotton include leaf curl, necrosis and defoliation, and vascular tissue wilt and discolouration (Sink and Grey, 1999). Verticillium wilt was spread to China by cotton introduction from America in 1935 and was responsible for the significant losses in the 1970s and 1980s (Bugbee, 1970; Cai *et al.*, 2009). In China, an area of >2 million ha of cotton is subject to Verticillium wilt, and the disease has become the most economically important disease of cotton. A severe outbreak of this disease causes reduction in fibre quality and significant yield losses, which may be up to 30% (Cai *et al.*, 2009). The symptoms of diseased cotton seedlings in the field are shown in Supplementary Fig. S1A available at *JXB* online. Anatomical analysis of the diseased cotton stem exhibited wilt of the vascular structure, shown in Supplementary Fig. S1B. However, no efficient management has been developed for the control of this disease. First, conventional breeding for improvement of cotton resistance to Verticillium wilt has not been successful (Cai *et al.*, 2009). There are few available commercial Upland cotton (*Gossypium hirsutum*) varieties with resistance to *V. dahliae*. The cotton species *G. barbadense* is immune to *V. dahliae*, but it is not the main cultivar worldwide. Secondly, there is no chemical pesticide available for cotton Verticillium wilt as it is a soil-borne vascular disease. Lastly, cultural practices such as seeding, irrigation, fertilization, and crop rotation are known to impact disease development but are not able to control Verticillium wilt effectively (Kamal *et al.*, 1985).

Recent development of molecular biology and bioinformatics technology has made it possible to better understand the molecular mechanism during plant–pathogen interaction and provide the possibility to improve plant disease resistant through plant genetic engineering. Plants use a series of defence mechanisms to protect themselves against pathogen attack through a complex perception, transduction, and exchange of signals (Verhage *et al.*, 2010). As an initial step in responding to pathogen challenge, plants sense the non-self signals in a timely fashion, and then signal transduction is conducted by a phytohormone signalling pathway and other signalling systems. After signal perception and transduction, subsequent transcriptional activation or repression of transcription factors (TFs) eventually leads to the induction of downstream signal transduction, the expression of plant defence-related genes, the production of phytoalexin, and the reinforcement of the cell wall (Yang *et al.*, 1997).

Some previous studies have emphasized the molecular mechanism of the Verticillium wilt disease caused by *V. dahliae*. Resistant species can also be colonized by the fungus, followed by a serious basal defence response and quick induction of phytoalexin production, including terpenoids and phenylpropanoid substances (Cu *et al.*, 2000; Tan, *et al.*, 2000; Mcfadden, *et al.*, 2001). Cotton terpenoid phytoalexins are synthesized as a defence response, induced by fungal and bacterial infection or other environmental stimuli. Some reports have elucidated the cotton terpenoid pathway, which may contribute greatly to the cotton defence response to pathogens (Tan *et al.*, 2000; Luo *et al.*, 2001; Xu *et al.*, 2004). Phenylpropanoids function as pre-formed and inducible antimicrobial compounds in plant–microbe interactions (Naoumkina *et al.*, 2010). Previous studies have reported that the phenylpropanoid pathway plays a critical role during the plant defence response to *V. dahliae* (Smit and Dubery, 1997; Pomar *et al.*, 2004; Gayoso *et al.*, 2010). The only *Verticillium* resistance genes, *Ve1* and *Ve2*, coded for receptor-mediated endocytosis, and with leucine zipper motifs and PEST sequences, were cloned in tomato (*Solanum lycopersicum*) by map-based cloning (Kawchuk *et al.*, 2001). The resistant tomato line LA3038 (*Ve/Ve*) exhibited induction of H_2O_2 production and an increase in peroxidase (POD) and phenylalanine-ammonia lyase (PAL) activity in inoculated roots, as well as an increase in the synthesis of lignins (Gayoso *et al.*, 2010). Cotton (*G. hirsutum*) hypocotyl tissue exhibited increased synthesis and deposition of lignin and lignin-like phenolic polymers after treatment with the elicitor from *V. dahliae* (Smit and Dubery, 1997). Some important enzymes of the phenylpropanoid pathway including chalcone synthase (CHS) and PAL have been characterized in cotton after infection with *V. dahliae* (Cu *et al.*, 2000).

Several studies have characterized transcriptomic changes in the cotton defence response by hybridization- or sequence-based approaches (Hill *et al.*, 1999; Zuo *et al.*, 2005). In the authors' laboratory, 211 unique genes were differentially identified in the *G. barbadense* response to *V. dahliae* by combining the use of suppression subtractive hybridization and microarray (Xu *et al.*, 2011). However, because of the complicated cotton genome, the limited information cannot provide a comprehensive understanding of the cotton defence response to *V. dahliae*. The recent development of deep-sequencing technology has made it possible to monitor the gene expression profiles of the defence response systematically and offers a great potential to enrich our understanding of how cotton has elaborated various chemical defence mechanisms against pathogen attacks.

In this study, the first global analysis of the cotton defence transcriptome dynamics was performed using the RNA-Seq method. On the basis of global gene regulation of phenylpropanoid metabolism-related genes, phenylpropanoid metabolism was deduced to be involved in the cotton defence response to *V. dahliae*. A closer look at the expression of these genes, enzyme activity, and the lignin level revealed obvious differences between resistant and susceptible cotton plants.

Materials and methods

Preparation of material

A highly aggressive defoliating fungus, *V. dahliae* V991, was incubated on a potato–dextrose agar plate for 1 week and then was inoculated into Czapek broth on a shaker at 120 rpm at 25 °C for 3–4 d, until the concentration of spores reached ~10^8–10^9 spores ml^{-1}. The suspension liquid was adjusted to 10^7 spores ml^{-1} with sterile distilled water for inoculation (Sink and Grey, 1999).

The *G. barbadense* L. variety 7124 (resistant) and *G. hirsutum* L. variety YZ-1 (susceptible) seeds were grown in commercial sterilized soil (a complex of soil, peat, and composted pine bark) at 24 °C/20 °C day/night temperatures with a photoperiod of 14 h light and 10 h dark for 2–3 weeks.

The cotton seedlings were infected with *V. dahliae* by root dip inoculation into a suspension of fungal spores for 1 min and then were returned to their original pots. Control plants were not inoculated but were otherwise treated and sampled with distilled water in the same way. Roots of four individual seedlings were collected for each treatment at each sampling time point.

Illumina sequencing and data processing

Total RNA was isolated from the collected roots using a modified guanidine thiocyanate method (Tu *et al.*, 2007). The five RNA samples, including four samples isolated from *G. barbadense* cotton roots inoculated after 4, 12, 24, and 48 h and a mixed sample of mock-inoculated plants after 4, 12, 24, and 48 h, were sent to Beijing Genomics Institute (Shenzhen) where the libraries were produced and sequenced using the Illumina Genome Analyzer (Solexa). Briefly, the cDNA was digested with *Nla*III and then ligated with the first adaptor containing the recognition site of *Mme*I, a type II endonuclease that cleaves at sites 21 bp from the recognition site. After digestion by *Mme*I, the transcripts were ligated with the second adaptor. With the sequencing primers designed based on the two adaptors, each tunnel will generate millions of raw reads with a sequencing length of 35 bp. Raw sequences were transformed into clean tags after certain steps of data processing, including removal of the 3' adaptor sequence, empty reads, low-quality tags, and tags that are too long or too short, leaving tags of 21 bp.

All clean tags were mapped to the cotton contigs assembly using the Bowtie mapping algorithm and only allowed no more than a 1-nucleotide mismatch (Langmead *et al.*, 2009). Clean tags mapped to the reference contigs assembly from multiple genes were filtered. The remaining clean tags were designed as un-ambiguous clean tags. The number of unambiguous clean tags for each gene was calculated and then normalized to TPM (number of transcripts per million clean tags), which associated the tag number with gene expression levels (Morrissy *et al.*, 2009). Differential gene expression between pathogen-inoculated and mock samples was determined by taking the log2 ratio of TPM.

Identification of differentially expressed genes

A rigorous algorithm was used to identify differentially expressed genes in the cotton defence response. The false discovery rate (FDR) was set at 5% to determine the threshold of *P*-value in multiple tests and analyses by manipulating the FDR value (Audic and Claverie, 1997). $P \leqslant 0.001$ and the absolute value of log2Ratio $\geqslant 1$ were used as the threshold to judge the significance of gene expression difference according to Audic and Claverie (1997). Cluster analysis of gene expression patterns was performed by Genesis based on the K-means method (http://genome.tugraz.at/) (Sturn *et al.*, 2002). Gene ontology (GO) analysis was applied to predict gene function and calculate the functional category distribution frequency.

Data validation by qPCR and RT-PCR

First-strand cDNA was generated from 3 μg of total RNA isolated from roots of both mock-treated and pathogen-treated *G. barbadense* and *G. hirsutum* plants by using the Superscript first-strand synthesis system (Invitrogen, Foster City, CA, USA). The quantitative real-time PCR (qPCR) experiment was conducted according to the guidelines of the Minimum Information for Publication of Quantitative Real-Time PCR Experiments (Bustin *et al.*, 2009). Primers for qPCR were designed using Primer Express 3.0 software (Applied Biosystems, Foster City, CA, USA), synthesized commercially (Genscript Biotechnology, Nanjing, China), and are shown in Supplementary Table S1 at *JXB* online. For qPCR, 20 μl qPCRs were run in three technical replicates on an ABI 7500 Real Time PCR System (Applied Biosystems), using 5 μl of first-strand cDNAs and SYBR Green PCR Master Mix (Applied Biosystems). PCR cycles were as follows: one cycle of 1 min at 94 °C, followed by 40 cycles at 94 °C for 15 s and 58 °C for 45 s. Following amplification, all products were subjected to melt curve analysis. A negative control without a cDNA template was run with each analysis to evaluate the overall specificity. The cotton *ubiquitin* gene was used as the reference gene to normalize the total amount of cDNA in each reaction. Relative fold changes were calculated with $2^{-\Delta\Delta Ct}$ using the SDS software from the 7500 Real Time PCR System (Applied Biosystems) (Rieu and Powers, 2009). To assess the correlation between different platforms, the Pearson correlations were calculated by SPSS 16.0 to compare the mRNA expression levels measured by RNA-Seq and qPCR.

Primers for reverse transcription-PCR (RT-PCR) were designed using Primer premier 5.0 software and synthesized commercially (Genscript Biotechnology), as shown in Supplementary Table S2 at *JXB* online. Standard PCR cycles were as follows: one cycle of 5 min at 94 °C, followed by 32 cycles at 94 °C for 30 s, 57 °C for 30 s, and 72 °C for 30 s, and a final step at 72 °C for 5 min. The cotton *ubiquitin* gene was used as a control. The PCR products were separated on 1.2% agarose gel and stained with ethidium bromide.

Measurement of enzyme activities

The root and stem samples were collected at 1, 4, 12, 24, 48, 96, 144, and 240 h from both *V. dahliae*- and mock-inoculated cotton plants. Four plants were collected for each sample at different time intervals and were frozen immediately in liquid nitrogen and stored at -70 °C. Samples of 100 mg were homogenized in extraction buffer (sodium acetate 50 mM, pH 5.0) using a mortar and pestle. The lysate was then centrifuged at 14 000 g for 15 min at 4 °C. The supernatant was collected as crude enzyme for the estimation of enzyme activities.

The determination of PAL activity was performed according to the method of Dubery and Smit (1994). The extract was incubated for 2 h at 37 °C in 10 mM l-phenylalanine, 0.1 M sodium borate, pH 8.8. The reaction was stopped by adding 5 M HCl. The mixture was centrifuged and the amount of trans-cinnamic acid formed in the supernatant was measured spectrophotometrically at 290 nm. PAL activity was expressed as micrograms of cinnamic acid formed per microgram of protein. The determination of POD activity was assayed at 470 nm by using guaiacol as hydrogen donor (Smit and Dubery, 1997). The reaction mixture contained 0.1 M NaPi, 50 mM guaiacol, 10 mM H_2O_2, and crude extracts.

The assay was carried out for 5 min at 25 °C and enzyme activity was determined over the linear part of the reaction spectrophotometrically at 470 nm; POD activity was expressed as nanokatal (nKat) per gram of sample.

Histochemical test

Hand-cut cross-sections were made from the base of the stem of both the inoculated and mock-treated cotton plants at 14 d after treatment. Lignin histochemistry was examined using Wiesner reagent (Pomar *et al.*, 2004). Cross-sections were incubated for 10 min in a phloroglucinol solution (2% in 95% ethanol) or 95% ethanol (staining control), treated with 18% HCl for 5 min, and directly observed with a bright-field microscope using a Leica fluorescence microscope (DM2500, Leica, Wetzlar, Germany).

Determination of total lignin content and lignin monomer composition

Total lignin content was determined by the Klason method with modifications (Eudes *et al.*, 2006). The lignin content was determined in three biological replicates and from 300 mg of extract-free dry stem. The Klason lignin content was calculated as weight percent of dry extract-free stem.

Alkaline nitrobenzene oxidation and high-performance liquid chromatography analyses were performed to analyse the lignin monomer composition (Lapierre *et al.*, 1995; Billa *et al.*, 1996). Quantification of the monomer composition was carried out at 280 nm and 313 nm, using the corresponding standards.

Results

The highly aggressive defoliating fungus *V. dahliae* strain V991 was used to inoculate *G. barbadense* cv. 7124 (resistant) and *G. hirsutum* cv. YZ-1 (susceptible). Successful infection was verified by the appearance of visual symptoms on inoculated susceptible plants (Fig. 1). The susceptible plants exhibited cotyledon wilting, leaf chlorosis, and severe dwarfing. However, the resistant plants also showed slight stunting compared with the control plants. The incompatible pathosystem between *G. barbadense* cv. 7124 and *V. dahliae* V991 was developed to study the transcriptome changes after pathogen inoculation by RNA-Seq.

Analyses of RNA-Seq data

A total of 27 947 550 tags were generated that were 21 bp in length. Each sample was represented by at least 5.5 million tags, which is sufficient for quantitative analysis of gene expression. For each sample, ~50% of the tags were not mapped to the assembly cotton contigs. Of the total tags, 44.8% mapped tags included 38.9% matched unique reference sequence. About 2 million unambiguous clean tags per sample for each gene were calculated and then normalized to TPM, which associated the tag numbers with gene expression levels. The summary of the tag information is shown in Table 1.

Cluster and functional assignments of differentially expressed genes

A total of 32 774 genes were detected during the cotton defence response to *V. dahliae*. Among these, 3442 differentially expressed genes were filtered with a cut-off of $P \leq 0.001$ and the absolute value of log2Ratio ≥ 1 based on the FDR <0.05 (Supplementary Data S1 at *JXB* online). The number of genes up- or down-regulated at different time points after pathogen inoculation is shown in Fig. 2. With an increase in time after inoculation, the number of the differentially expressed genes also increased. More down-regulated genes were identified than up-regulated genes at all time points after pathogen inoculation. Genes were divided into three groups based on their expression patterns after pathogen inoculation by Genesis based on K-means method (Fig. 3). Type I and II included genes positively and negatively regulated during the cotton defence response both including five clusters. Type III included genes showing complex expression patterns.

To facilitate the global analysis of gene expression profiles and evaluate the potential functions of genes that showed significant transcriptional changes after pathogen
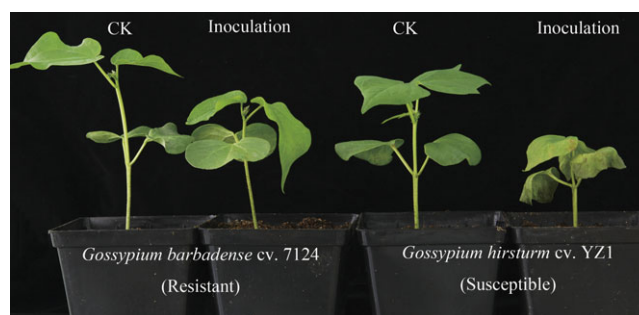

Fig. 1. Growth and disease symptoms on both resistant and susceptible cotton plants 14 d after *V. dahliae* inoculation.

Table 1. Summary of tag numbers

	Time after inoculation (h)					Total
	CK	4	12	24	48	
Clean tag	5 609 148	4 829 828	5 539 569	6 141 380	5 827 625	27 947 550
All tag mapping to gene	2 538 783	2 117 140	2 447 219	2 724 901	2 688 605	12 516 648
% all tag mapping to gene	45.26%	43.83%	44.18%	44.37%	46.14%	44.79%
Unambiguous tag mapping to gene	2 210 526	1 839 038	2 133 954	2 353 331	2 319 546	10 856 395
% Unambiguous tag mapping to gene	39.41%	38.08%	38.52%	38.32%	39.80%	38.85%
Unknown tag	3 070 365	2 712 688	3 092 350	3 416 479	3 139 020	15 430 902
% Unknown tag	54.74%	56.17%	55.82%	55.63%	53.86%	55.21%

challenge, GO-based classification was conducted by sequences subjected to InterPro and GO annotation. A total of 2128 genes were annotated from 3442 differentially expressed genes based on sequence similarity to those in the public databases. The GO functional classification analysis on each type was performed to calculate the functional category distribution frequency based on biological processes level 3 (Fig. 4). Analysis of categories revealed
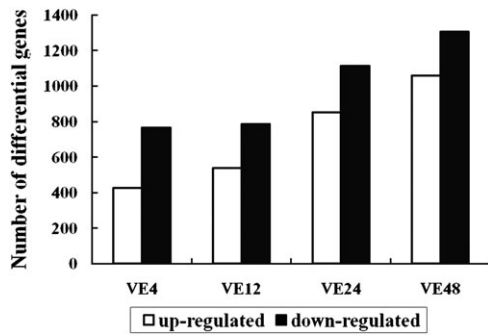

Fig. 2. Number of differentially expressed genes at different times after inoculation. Ve4, Ve12, Ve24, and Ve48 refer to 4, 12, 24, and 48 h after inoculation with *V. dahliae*.

that the defence-responsive genes were classified into 22 biological processes, including metabolism, transport, catabolic process, response to stress and stimulus, biosynthetic process, regulation, cell communication, and so on. It implied that normal developmental processes were strongly affected after pathogen invasion. Type I, with genes mainly up-regulated during the defence response, included more genes assigned to secondary metabolism process than type II, which also correlated with the following results regarding the contribution of secondary metabolism to disease resistance.

Validation of RNA-Seq-based gene expression by qPCR

To validate the RNA-Seq-based gene expression levels that correlated with the number of tags obtained, qPCR was performed on selected genes with different expression levels and functional assignments. Transcriptional regulation revealed by RNA-Seq was confirmed in a biologically independent experiment using qPCR. Of the 361 candidates, 51 genes showed either no specific amplification or unexpected size amplification.

The Pearson correlation coefficient was calculated by SPSS to assess the correlation between different platforms.
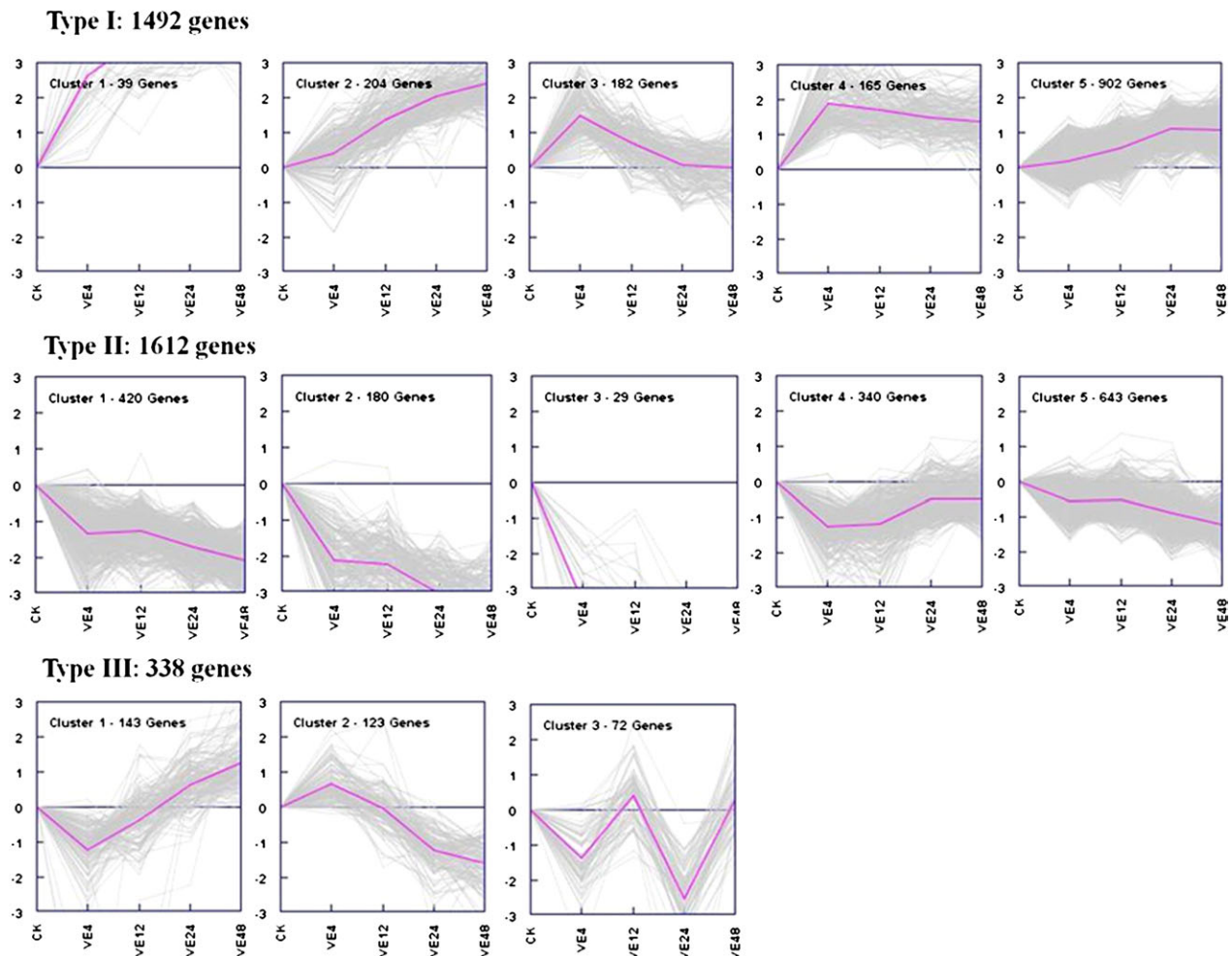

Fig. 3. Cluster analysis was developed by the K-means method on the gene expression profiles. CK indicates the mock control, and Ve4, Ve12, Ve24, and Ve48 refer to 4, 12, 24, and 48 h after inoculation with *V. dahliae*.
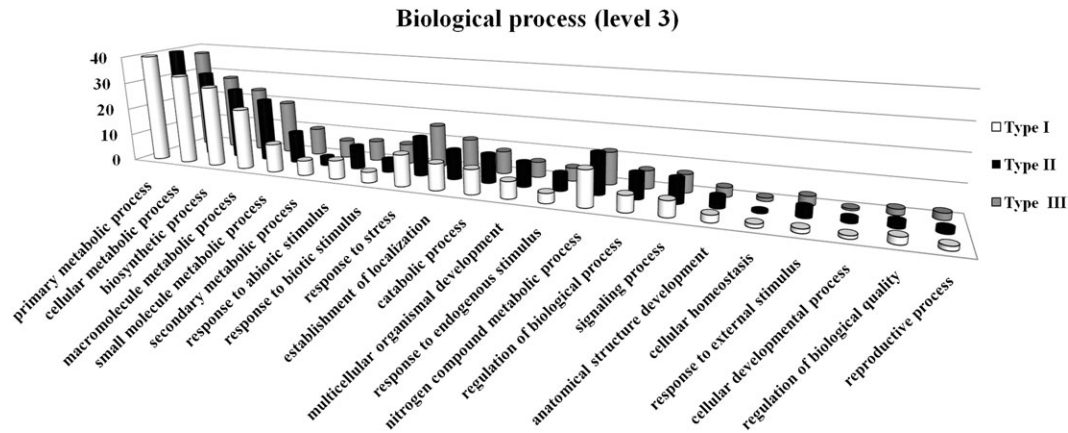
Biological process (level 3)

Fig. 4. GO functional classification analysis on each type of gene. Histograms represent the functional distribution, which is expressed as a percentage of the amount of genes. Type I, II, and III are indicated as in Fig. 3.
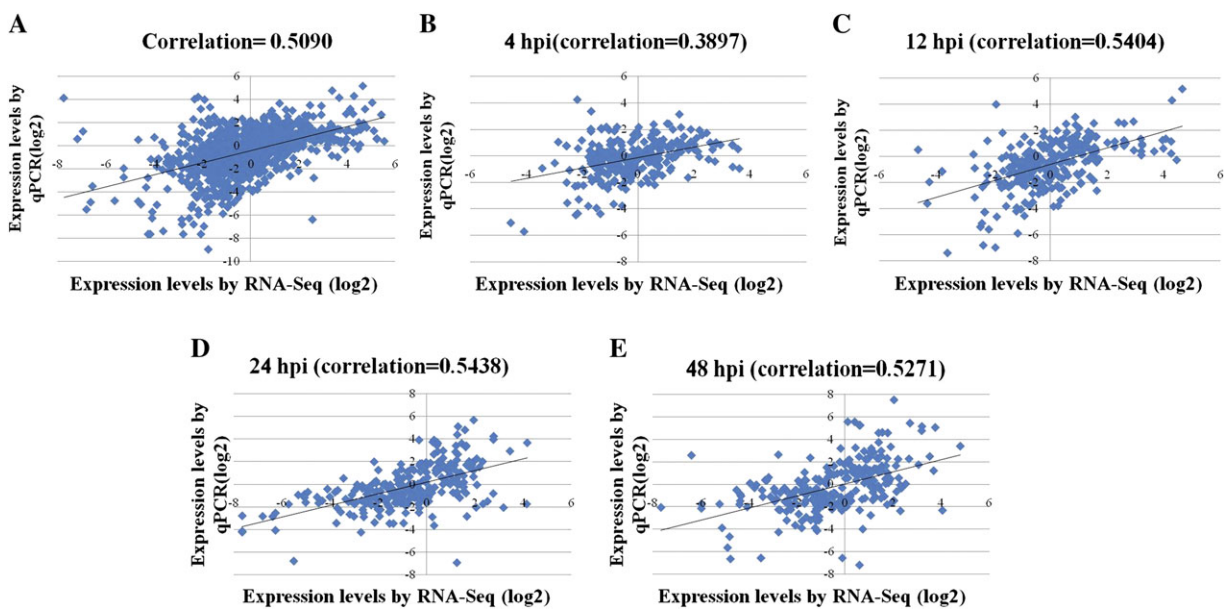

Fig. 5. Correlations between RNA-Seq and qPCR for different time points. hpi, hours post-inoculation.

Overall, the qPCR measurements were moderately correlated with RNA-Seq results (Fig. 5A, R=0.509; correlation is significant at the 0.01 level). Gene expression levels estimated by qPCR at 4 h post-inoculation (hpi) showed low correlation with them by RNA-Seq (Fig. 5B, R=0.3897; correlation is significant at the 0.01 level). In contrast, the correlations were higher at 12, 24, and 48 hpi and ranged from 0.52 to 0.54 (Fig. 5C–E).

Global gene regulation in response to V. dahliae

Several subsets of genes, including receptor-like kinases (RLKs), TFs, and phytohormone-responsive genes, were identified.

RLKs: Of the 19 RLKs obtained from differentially expressed genes, classified as leucine-rich repeat RLKs

(LRR-RLKs), seven were induced genes and 12 were repressed genes (Supplementary Fig. S2 at *JXB* online).

Phytohormone signalling-related genes: Groups of genes involved in different phytohormone signalling pathways during the cotton defence response were characterized by the RNA-Seq approach. The detailed gene expression patterns are shown in Supplementary Fig. S3 at *JXB* online. The largest group, assigned as ethylene (ET) signal transduction-related genes encoding ethylene receptor (ETR), ethylene response factor (ERF), and ethylene response element-binding protein (EREBP), showed distinct expression patterns. Two jasmonate ZIM-domain (JAZ) genes, with most similarity to JAZ1 and JAZ10 in *Arabidopsis*, which acted as a repressor in the JA signalling pathway, were identified as suppressed based on RNA-Seq. Another two genes encoding non-expressor of pathogenesis-related

protein 1 (NPR1), the key regulator of the salicylic acid (SA)-mediated defence response, were detected as suppressed during the cotton defence response. Two types of abscisic acid (ABA) signalling pathway genes were identified, including six induced ABA-deficient genes and four suppressed ABA-insensitive genes. A group of genes encoding enzymes involved in the gibberellin (GA) biosynthetic pathway has been isolated, including GA 3-β hydroxylase, GA 20-oxidase, and GA 2-oxidase. Some auxin-related genes were identified with different expression patterns.

TFs: A total of 172 regulated TFs were identified, including WRKY, AP2/EREBP, bZIP/HD-ZIP, MYB, C_xH_y zinc-finger, WD-40 repeat family protein, BTB/POZ domain-containing protein, bHLH, NAC, MYC, and the remainder represented different classes of TFs with unknown roles, some of which have been reported to be critical components of plant adaptive response to biotic and abiotic stresses. The majority of TFs were suppressed based on the RNA-Seq analysis after pathogen inoculation, which correlated moderately with the results by qPCR (Supplementary Fig. S4 at *JXB* online).

Phenylpropanoid pathway-related genes: A subset of genes participating in multiple branches of the phenylpropanoid pathway, including lignins, flavonoids, and phenolic compounds, were identified based on RNA-Seq analyses. Global transcriptome changes focused on 15 types of phenylpropanoid pathway-related genes during defence responses of cotton to *V. dahliae* have been examined. The expression patterns of the characterized genes are shown in

Supplementary Fig. S5 at *JXB* online. The results were also verified by qPCR and correlated well with the results of qPCR (Supplementary Fig. S5, $R=0.677$; correlation is significant at the 0.01 level). An overview of the phenylpropanoid pathway and the expression profiles of identified genes involved in these pathways during cotton and *V. dahliae* interactions are shown in Fig. 6. The red box contains enzymes that function in the core phenylpropanoid pathway, including PAL, cinnamate 4-hydroxylase (C4H), and 4-coumarate:CoA ligase (4CL). After that, the pathway leads to three specific branch pathways for the formation of lignins, anthocyanins, and flavonoids. Most of the identified genes involved in this pathway were up-regulated after pathogen inoculation, except flavanone 3-hydroxylase (F3H) and anthocyanidin synthase (ANS). Some of the genes exhibited both up- and down-regulated expression patterns, including 4CL, CHS, caffeoyl-CoA O-methyl-transferase (CCoAOMT), and cinnamyl alcohol dehydrogenase (CAD). Overall, the results indicated the activation of the phenylpropanoid pathway in the cotton defence response to *V. dahliae*.

To better characterize the gene function and the difference between resistant and susceptible cotton after pathogen inoculation, a subset of genes was subjected to qPCR and RT-PCR analysis during the interactions between *V. dahliae* and both cotton lines, including the genes functioning in the core phenylpropanoid pathway and the synthesis of lignin. The relative expression level obtained by RNA-Seq after taking the equation and logarithmic transformations of TPM and also by qPCR for data verification is shown in Fig. 7. To understand the absolute gene
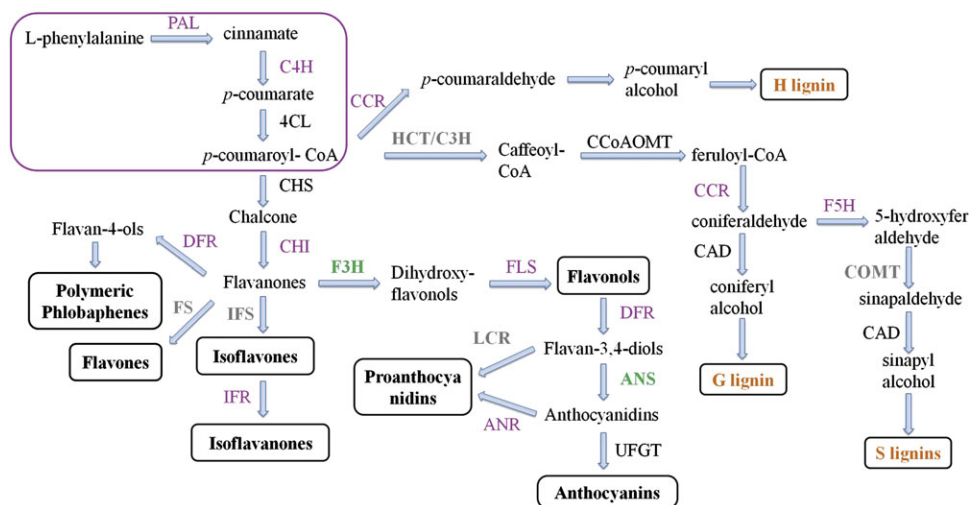

Fig. 6. Overview of putative phenylpropanoid pathway involved in cotton and expression profiles of genes involved in this pathway (Ferrer *et al.*, 2008). The box includes the critical enzymes comprising the entry pathway. Enzymes coloured in purple or green indicate the induction or suppression of the genes, respectively. Enzymes in black are those with both increased and decreased transcripts, whereas enzymes shown in grey are those that have not been identified in the present study. PAL, phenylalanine ammonia-lyase; C4H, cinnamate 4-hydroxylase; 4CL, 4-coumarate:CoA ligase; CHS, chalcone synthase; CHI, chalcone isomerase; IFS, isoflavone synthase; IFR, isoflavone reductase; FS, flavone synthases; F3H, flavanone 3-hydroxylase; FLS, flavonol synthase; DFR, dihydroflavonol-4-reductase; ANS, anthocyanidin synthase; ANR, anthocyanidin reductase; UFGT, UDP-flavonoid glucosyltransferase; C3H, *p*-coumarate 3 hydroxylase; HCT, hydroxycinnamoyl transferase; CCR, cinnamoyl CoA reductase; CAD, cinnamyl alcohol dehydrogenase; CCoAOMT, caffeoyl-CoA O-methyltransferase.

expression level, RT-PCR results and the original TPM obtained by RNA-Seq are also shown in Supplementary Fig. S6 at *JXB* online. Overall, most of the genes examined were induced after *V. dahliae* inoculation in both susceptible and resistant cotton plants, with the exception of *CHS-4* and *CCoAOMT-1*, both by qPCR and by RT-PCR. Moreover, the genes displayed a lower expression level and delayed increase in the inoculated susceptible cotton plants.

Measurement of enzyme activity in the cotton defence response

The activities of enzymes involved in the phenylpropanoid pathway leading to the synthesis of lignin, including PAL and POD, were measured in both *G. hirsutum* (susceptible) and *G. barbadense* (resistant) after *V. dahliae* inoculation (Fig. 8). The activity of PAL appeared to be stimulated at
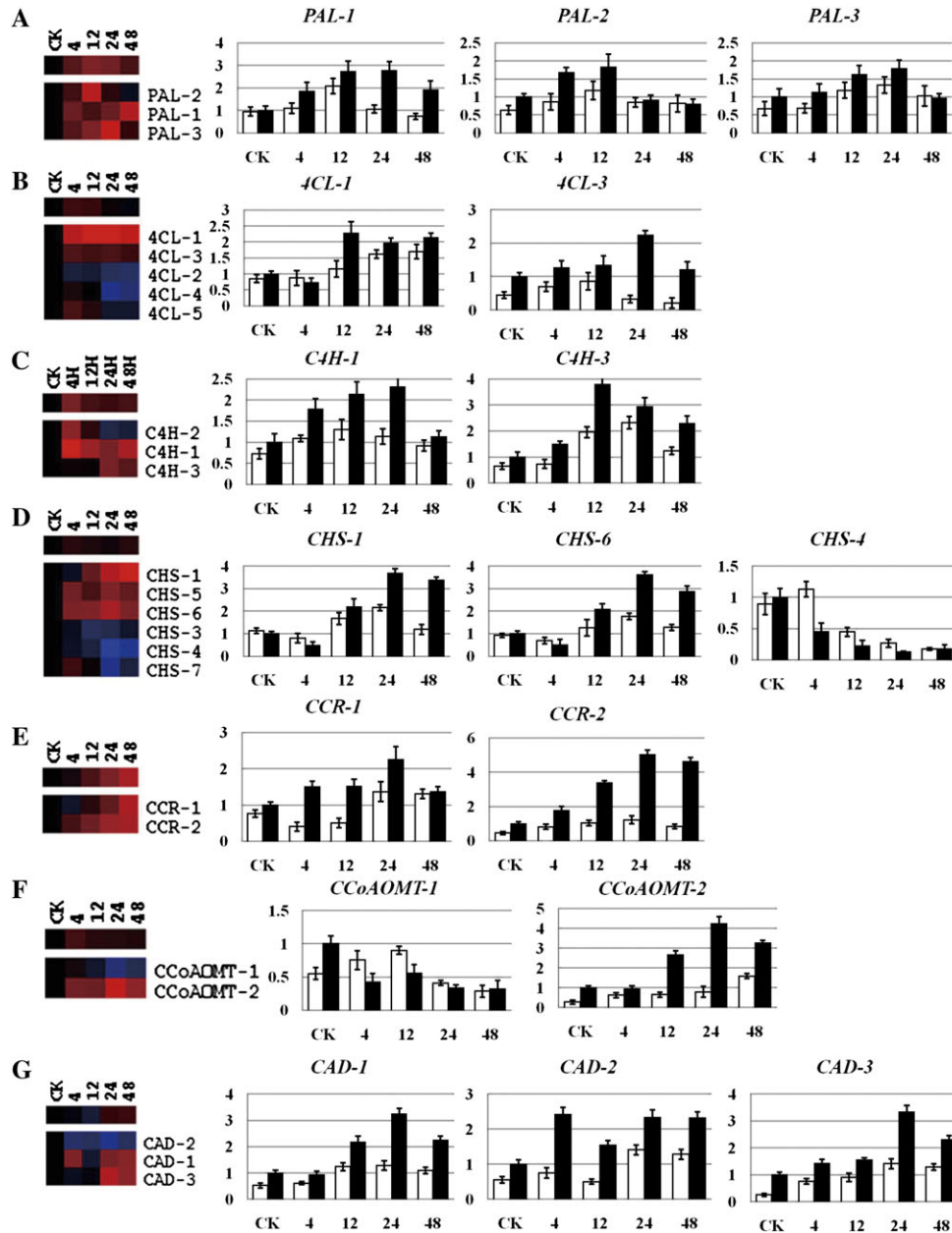
Fig. 7. Detailed expression profiles of genes in the major lignin biosynthetic pathway. The relative expression level was obtained by RNA-Seq after taking the equation and logarithmic transformations of TPM, and also by qPCR for data verification. White and black columns refer to *G. hirsutum* (susceptible) and *G. barbadense* (resistant), respectively. Error bars represent the SD for three independent experiments, and three technical replicates were analysed. CK, 4, 12, 24, and 48 on the *x*-axis refer to control and 4, 12, 24, and 48 h after inoculation with *V. dahliae*. The *y*-axis represents the relative expression level. (A) PAL, phenylalanine ammonia-lyase; (B) 4CL, 4-coumarate:CoA ligase; (C) C4H, cinnamate 4-hydroxylase; (D) CHS, chalcone synthase; (E) CCR, cinnamoyl CoA reductase; (F) CCoAOMT, caffeoyl-CoA O-methyltransferase; (G) CAD, cinnamyl alcohol dehydrogenase.

1 h after *V. dahliae* inoculation in roots from both suscepti-ble and resistant cotton plants and decreased at 144 h after inoculation. PAL activity of resistant plants was maintained at a high level from 12 h to 144 h after inoculation, whereas it accumulated to a relatively low degree in susceptible cotton plants compared with resistant plants. PAL activity of stems showed patterns distinct from its activity in roots. The inoculation of resistant plants led first to an increase that peaked at 12 h and then a second lesser increase detected after 48 h and peaking at 96 h. The inoculated susceptible plants only showed a slight increase at 4 h, which coincided with the first peak in resistant plants, and then decreased at 12 h. Interestingly, PAL activity of roots from both control plants also showed a slight increase, but showed no appreciable changes in the stems.

Similar POD activities were observed in root and stems from both susceptible and resistant plants, which increased continuously after pathogen inoculation and slightly in-creased in both control plants. In agreement with the enzyme activity of PAL, the resistant plants accumulated a higher level of POD activity than susceptible plants both in roots and in stems.

Histochemical localization and quantitative analysis of lignin

The phloroglucinol-HCl (Wiesner) reaction was used to detect the lignin on stem cross-sections. The analyses of both resistant and susceptible cotton plant lines were carried out before inoculation and 14 d after inoculation, which may reflect different stages of stem growth. All lines showed a deep magenta stain in the xylem. Little difference

was seen between the untreated resistant and susceptible cotton plants as well as mock-inoculated resistant and susceptible cotton plants, with the exception that the pith diameter of susceptible plants was a little bigger than that of resistant plants (Fig. 9A–H). However, the cross-sections of the inoculated susceptible cotton stems showed delayed development of xylem and decreased lignified xylem bun-dles, vessels, and interfascicular fibres compared with the inoculated resistant cotton (Fig. 9I–L).

These distinct features prompted the performance of additional analyses to better define the lignin content and monomeric composition of the two different cotton lines. Lignin contents were estimated in cell wall residues (CWRs) obtained by solvent extraction of ground freeze-dried stem tissues by the improved Klason method. Inoculation of the two cotton lines with *V. dahliae* resulted in increases in the total stem lignin content at 14 d after treatment, but the increase in resistant plants reaching 5% was more significant than in susceptible plants at 3%. At 25 d after inoculation, the resistant cotton line also responded to *V. dahliae* inoculation with a 5% increase in stem lignin compared with the control (Table 2). The susceptible cotton plants were not analysed at 25 d after inoculation because they died quickly once they showed the symptoms of disease.

To better determine the nature of cotton stem lignins, nitrobenzene oxidation was adopted to analyse the mono-mer composition in the CWRs isolated from both control and inoculated cotton stems. The analysis of nitrobenzene oxidation products revealed that cotton stem lignins con-tained a higher proportion of guaiacyl (G) and syringyl (S) units and a lower proportion of *p*-hydroxyphenyl (H) units, which only accounted for 5–10%. Stem lignins of inoculated
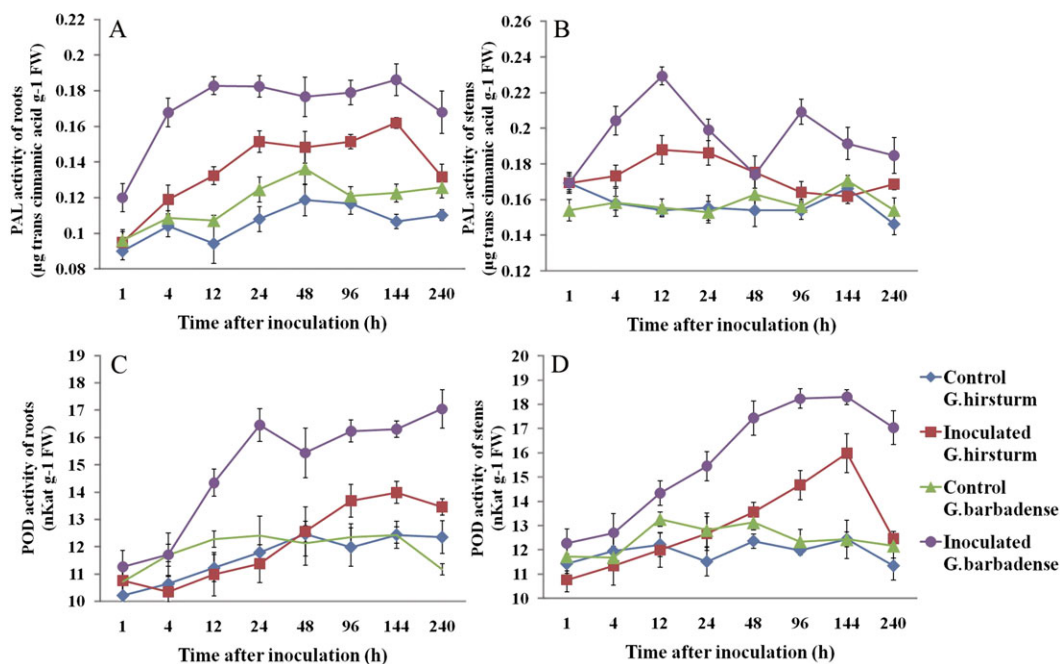

Fig. 8. Enzyme activity in roots and stems of both control and inoculated susceptible and resistant cotton plants at different time points after inoculation. (A) PAL, roots; (B) PAL, stems; (C) POD, roots; (D) POD, stems. Error bars represent the SD for three independent experiments.
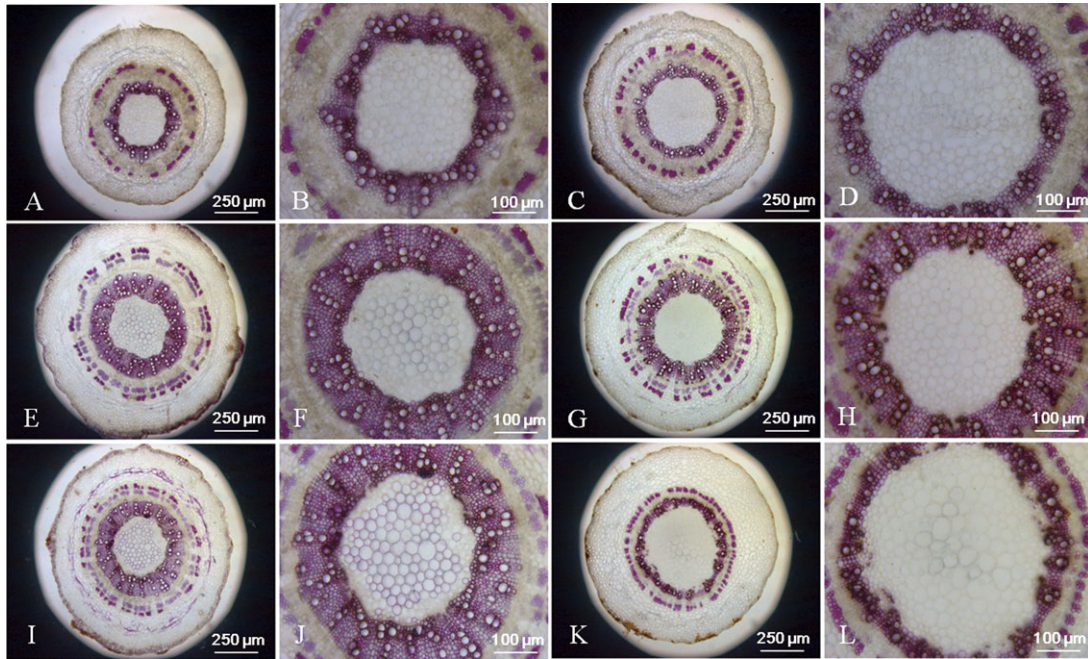
Fig. 9. Histochemical analysis of lignin in stem cross-sections of control and *V. dahliae*-inoculated cotton plants. Hand-cut cross-sections of stem were stained with Wiesner reagents for detecting lignin. Pink staining with the Wiesner reagent indicates the presence of *p*-hydroxycinnamyl aldehyde end groups in lignin. (A, B, E, F, I ,J) Sections from control and *V. dahliae*-inoculated *G. barbadense* (resistant) stems. (C, D, G, H, K, L) Sections from control and *V. dahliae*-inoculated *G. hirsutum* (susceptible) stems. (A–D) Sections from 16-day-old plants before inoculation. (E–H) Sections from control plants 14 d after treatment. (I–K) Sections from inoculated plants 14 d after treatment. Adjustments to magnification and illumination were made to allow optimal viewing of individual sections.

Table 2. Lignin content and monomer composition after inoculation of the cotton stems with *V. dahliae*
Values followed by a different letter are significantly different (*P* <0.05).

			Lignin content (%)	Monomer composition (%)			G/S ratio
				H units	G units	S units	
14 dpi	*G. hirsutum* (S)	Control	12.83±0.81a	5.44	53.55	41.01	1.31
		Inoculated	15.87±0.45 b	5.71	47.82	46.47	1.02
	G. barbadense (R)	Control	13.49±0.61 a	6.01	46.43	47.55	0.98
		Inoculated	18.46±0.78 c	9.92	58.52	31.56	1.85
25 dpi	*G. barbadense* (R)	Control	17.22±0.97 c	10.44	56.89	32.67	1.74
		Inoculated	22.38±1.25 d	9.32	65.24	25.44	2.56

resistant plants contained almost 2-fold the level of H units compared with susceptible plants. Furthermore, in both control and inoculated resistant plants, H units accounted for ~10% at 25 d after treatment. The G/S ratios of inoculated resistant cotton increased compared with the control from 0.98 to 1.85 and from 1.74 to 2.56 at 14 d and 25 d after treatment, respectively. In contrast, the G/S ratios of susceptible cotton decreased from 1.31 to 1.02 after *V. dahliae* inoculation (Table 2).

Discussion

Because of the special features of allotetraploid cotton with a large and complex genome, comprehensive sequence information describing the genome or transcriptome of cotton has not been elucidated. Moreover, previous studies provided too limited information to determine the complex transcriptome dynamics during the cotton defence response. The recent availability of high-throughput sequencing technologies provides an unprecedented opportunity to explore the defence response thoroughly by large-scale expression profile analysis, despite uncharacterized genome sequences.

An experimental approach was designed to study the resistant mechanisms during incompatible interaction between cotton and *V. dahliae*. Four different sampling time points, 4, 12, 24, and 48 h after inoculation, were chosen to coincide with crucial stages of fungal infection and to isolate the early pathogen-responsive genes. The time point, within the first 12 hpi, often corresponds to fungal spore germination and penetration of the epidermal cells (Fradin and Thomma, 2006). Complex perception, transduction, and exchange of signals usually occur in the early stages of fungal infection (Zhang and Klessig, 2001; Kunkel and Brooks, 2002; Jones and Dangl, 2006).

A thorough pre-processing of the sequencing data combined with an advanced mapping strategy was applied for the expression profiling analysis by RNA-Seq. To obtain longer cotton sequence contigs, an assembly of cotton gene sequences was performed before tag mapping. The extremely low abundance transcripts were also filtered because of the effect of sequencing error. Clean tags mapped to multiple genes were filtered as suggested by Birzele *et al.* (2010). In total, ~50% of the tags sequenced could be assigned to genes and were used for gene expression profiling. Compared with other studies, the proportion was relatively low, which may be partly explained by the fact that most of the sequences existing in the database were generated from cotton fibre. In addition, cotton has a larger number of as yet unknown genes and expressed elements. The other 50% are likely to correspond to as yet unknown transcripts or splice variants, novel exons or untranslated regions (UTRs), and RNA genes. Therefore, the result of 50% appears to be reasonable. An attempt was also made to map to the well-annotated genome of *A. thaliana* because there is no closely related species which could be made use of. As a result, just 4% of the tags can be mapped, due to the species diversity as well as the UTR specificity.

Much work was performed on validating the RNA-Seq method by qPCR to determine whether the novel technology for analysis of a non-model organism with no reference genome is available. A group of selected genes with different expression levels and functional assignments were analysed by qPCR. Compared with previous studies, the result exhibited a moderate correlation between the two different platforms, which may be explained by the lack of a reference genome. Although an improved bioinformatics pipeline should be developed to adapt RNA-Seq to measure gene expression profiles in cotton, the results still suggested that RNA-Seq can be used for gene expression profiles during the cotton defence response. As a result, the early collected sample will be more affected by the sampling process and method.

Groups of genes identified as LRR-RLKs, phytohormone signalling-related genes, and TFs were assayed by RNA-Seq during the cotton defence response. Detailed elucidation of the response of these genes described the interplay of signals that allowed the plant to fine-tune defence responses. Plant RLKs are transmembrane proteins with putative extracellular domains and intracellular protein kinase domains, which can perceive external signals and initiate signal cascades (Gou *et al.*, 2010). RLKs control a wide range of processes, including development, disease resistance, and hormone perception. Plant innate immunity depended on the timely perception of the signals, which can be accomplished by pattern recognition receptors (PRRs), including well characterized LRR-RLKs (Panstruga *et al.*, 2009). The identification of groups of LRR-RLKs in the gene expression profiles based on RNA-Seq suggested putative roles for LRR-RLKs in mediating early events in the response of cotton to *V. dahliae*. Both up-regulated and down-regulated expression patterns implicated the putative multiple pathways recognized by LRR-RLKs during the cotton immunity.

Phytohormones such as ABA, JA, ET, SA, auxin, GA, cytokinin, and brassinosteroids appeared to play critical roles in the complex signalling cascades and were integrated into current models of defence responses (Bari and Jones, 2009). In summary, the complicated expression of these genes indicated the extraordinary complexity of signalling pathways regulated by phytohormones during the cotton defence response to *V. dahliae*. However, the exact role of these hormones and the cross-talk between them to fine-tune the defence response of cotton to *V. dahliae* remains to be discovered.

TFs always act as master switches by controlling the expression of series of genes with slight changes in expression to regulate different plant development or response to biotic and abiotic stresses (Singh *et al.*, 2002). The majority of TFs were suppressed based on the RNA-Seq analysis after pathogen inoculation, which suggested that establishment of the cotton resistance to *V. dahliae* might require the down-regulation of some kinds of genes to modulate the plant immune signalling network.

Secondary metabolites play a fundamental role in the plant's ability to fight against invading pathogens, which is derived from multiple pathways, including terpenoid, phenylpropanoid, and nitrogen-containing substances based on different biosynthesis substrates and pathways (Makkar *et al.*, 2007). Systematic analyses on a global range of genes involved in multiple branches of the phenylpropanoid pathway have been reported for *Arabidopsis*, rice, tomato, and legume plants, such as *Medicago truncatula*, alfalfa, pea, and soybean (Kandan *et al.*, 2002; Liu *et al.*, 2002; Scheideler *et al.*, 2002; Zabala *et al.*, 2006; Azaiez *et al.*, 2009; Uppalapati *et al.*, 2009). Although some reports provided evidence to suggest that lignin, flavonoid phytoalexins, and phenolic compounds play important roles during the cotton defence response to *V. dahliae* (Dubery and Smit, 1994; Smit and Dubery, 1997; Cu *et al.*, 2000; Mcfadden *et al.*, 2001), none reported on the overall analysis of transcriptome changes of phenylpropanoid pathway genes in cotton. The present results revealed global gene expression profiles of the phenylpropanoid pathway in response to *V. dahliae*. The genes identified by RNA-Seq in this study almost represented the main branches of the phenylpropanoid pathway, which led to the synthesis of lignin, flavonoid, anthocyanin, and proanthocyanidin. As a result the study focused on the phenylpropanoid pathway with an emphasis on the synthesis and deposition of lignin.

Previous studies have demonstrated that the lignin pathway is activated in the characterized incompatible pathosystems because lignification and reinforcement of cell walls are important processes in the response of plants to fungal infection (Naoumkina *et al.*, 2010). In particular, some reports revealed that lignin played critical roles in plant responses to *V. dahliae*. The synthesis and deposition of lignins were examined after exposure of cotton hypocotyls to an elicitor of *V. dahliae*, which was assumed to be a physical barrier to make the cell wall more resistant to mechanical pressure during fungal penetration (Smit and Dubery, 1997). Inoculation of pepper with *V. dahliae*

increased the total stem lignin content, which suggested that lignification may be a mechanism through which pepper could restrict the growth of *V. dahliae* hyphae in the xylem (Pomar *et al.*, 2004). Moreover, inoculation with *V. dahliae* also induced a significant increase in the total amount of lignin in tomato roots in both a susceptible line and a resistant line, although at earlier times in the latter (Gayoso *et al.*, 2010). Some genes encoding the enzymes involved in the lignin pathway exhibited increased expression, which indicated the putative role of lignin in the cotton defence response. More evidence was required to elucidate the definite role of lignin in the cotton defence response to *V. dahliae*.

To investigate thoroughly the difference in these genes between *G. hirsutum* (susceptible) and *G. barbadense* (resistant) after *V. dahliae* inoculation, a subset of genes functioning in the core phenylpropanoid pathway and the synthesis of lignin were submitted to qPCR and RT-PCR analyses based on relative and absolute gene expression levels, respectively. As expected, most of the examined genes were induced after *V. dahliae* inoculation in both susceptible and resistant cotton plants, with the exception of *CHS-4* and *CCoAOMT-1*, by both qPCR and RT-PCR. Moreover, the genes displayed a lower expression level and delayed increase in the inoculated susceptible cotton plants.

The activities of enzymes involved in the core and entry pathway of the phenylpropanoid pathway and biosynthesis of lignin were also measured in this study. PAL played an essential role in the phenylpropanoid pathway and has been reported to be responsive to both biotic and abiotic stress, including pathogen attack, wounding, cold, and UV light (Huang *et al.*, 2010). POD is involved in the polymerization of monolignols into lignin and reinforcement of the cell wall after pathogen attack and wounding (Marjamaa *et al.*, 2009). In this study, both PAL and POD activity were stimulated after *V. dahliae* inoculation, in both susceptible and resistant cottons. However, compared with resistant plants, the inoculated susceptible plants showed an obviously slower rate of increase and lower enzyme activities. Together with the gene expression profiles, it was speculated that the most significant difference between susceptible and resistant plants was whether or not the plants could activate the pathway in a timely and efficient manner. Interestingly, there was also a slight increase in PAL and POD activities in the roots of both control plants. This was explained by the fact that the wounding treatment of roots is responsible for the increase in enzyme, which agreed with other studies which showed that wounding stress can induce PAL expression and activity (Huang *et al.*, 2010). However, PAL activity was induced quickly in both roots and stems of susceptible and resistant plants, compared with POD activity, which confirmed the position of the two enzymes involved in the phenylpropanoid pathway. A similar result was also observed in the cotton defence response to elicitor of *V. dahliae* and in the tomato defence response to *V. dahliae* (Smit and Dubery, 1997; Gayoso *et al.*, 2010). The stem was used to investigate lignin histochemistry, content, and monomeric composition, subsequently. Therefore, PAL

and POD activities were also detected in the stem. Enzyme activity also increased after pathogen inoculation in stems, and POD activity shared a similar pattern with that seen in roots, but PAL activity did not.

Because of the differentially expressed genes and induced enzyme activity of PAL and POD, together with the apparent phenotypes between inoculated susceptible and resistant cotton plants, histochemical analyses were performed using cross-sections of both stems by Wiesner reagent. No difference was evident compared with the control plants, and development of metaxylem and lignification of xylem progressed quickly during the 14 d after treatment. Moreover, the inoculation of *V. dahliae* only significantly affected the vascular development of susceptible cotton plants but not resistant plants, including delayed development of xylem and decreased number of lignified xylem bundles, vessels, and interfascicular fibres, which may be responsible for the disease symptoms characterized by severely stunted growth and wilting of inoculated susceptible cotton. Although no appreciable changes were seen in the histochemical analysis of stem lignins at 14 d after inoculation, the resistant cotton showed slightly stunted growth.

Direct speculation on the content of lignin deposition in the cell wall is not possible from the difference of staining colour and the degree of lignification because of the variable thickness of hand-cut cross-sections. The result of quantitative analysis of lignin correlated with the analyses of the expression profiles of related genes and enzyme activities. Inoculated resistant cotton plants can accumulate more lignin compared with susceptible cotton at 14 d after *V. dahliae* inoculation. This result, together with the gene expression profiles and enzyme activity analyses, suggested that the pathway was assigned to both basic and induced resistance mechanisms in the cotton and *V. dahliae* interaction.

Lignins are synthesized from the oxidative coupling of three monolignols, named *p*-hydroxyphenyl (H), guaiacyl (G), and syringyl (S). The proportions of the three main units in the cell wall vary according to plant species and tissue type (Bhuiyan *et al.*, 2009). It has been demonstrated that plants can adopt effective mechanisms to restrict the pathogen spread in the vascular structure, by changing the monomer composition and cross-linking to reinforce the cell wall (Pomar *et al.*, 2004; Gayoso *et al.*, 2010; Gallego-Giraldo *et al.*, 2011). In view of this fact, nitrobenzene oxidation was developed to study the lignifications of both resistant and susceptible cotton when exposed to *V. dahliae*. The analysis of nitrobenzene oxidation products revealed that cotton stem lignin contained a higher proportion of G and S units and a lower proportion of H units, in agreement with the fact that dicotyledon lignin displays mainly G and S units. An increase in the G/S ratio was noticed in the inoculated resistant cotton stems compared with that of control plants at both 14 d and 25 d after treatment, whereas the G/S ratio was decreased in susceptible cotton after *V. dahliae* inoculation. Overall, the resistant cotton displayed increased amounts of G and H units and a decreased amount of S units, and G lignin was hypothesized to play a pivotal role in the cotton defence response. The

result suggested that the resistant cotton stems were more cross-linked after pathogen inoculation and the resistant cotton stems at 25 d after treatment were more cross-linked than at 14 d. Similar results also were reported in pepper and tomato after *V. dahliae* inoculation (Pomar *et al.*, 2004; Gayoso *et al.*, 2010). In accordance with these reports, it was suggested that increased lignification and cross-linking of resistant cotton stems help them to restrict pathogen growth in the vasculature. By combining histochemical and biochemical data, it was speculated that gene expression levels, enzyme activities, total lignin levels, and the degree of putative lignification were increased after inoculation. Overall, a critical role for lignin was deduced to contribute to the cotton disease resistance.

In addition, phenylpropanoid products may have important roles not only as phytoalexins but also as signal molecules in the plant defence response (Dixon *et al.*, 2002). The rapid accumulation of flavonoids, which were measured after pathogen inoculation in both control and inoculated cotton roots, suggested the possibility of flavonoids acting as signal molecules. The exact molecular mechanism is still poorly understood. The transcriptional analysis of these genes from the multiple branches of the phenylpropanoid pathway helped to better understand the molecular mechanism of the cotton defence response to *V. dahliae*. Further characterization of these genes may provide candidates for genetic improvement of cotton. Moreover, identification of transcription factors affecting the phenylpropanoid pathway and lignin synthesis and deposition will provide new insights into the control of plant development and disease resistance.

Supplementary data

Supplementary data are available at *JXB* online.

Data S1. TPM of differential expression genes at different time after inoculation with *V. dahliae*.

Fig. S1. The representative symptoms of diseased cotton seedlings in the field (A) and anatomical analysis of the diseased cotton stem (B).

Fig. S2. Expression of LRR-RLKs identified by RNA-Seq (A) and validation by qPCR (B). CK refers to the mock control, and VE4, VE12, VE24, and VE48 refer to 4, 12, 24, and 48 h after inoculation with *V. dahliae*.

Fig. S3. Detailed expression patterns of defence-responsive genes identified in different phytohormone signalling pathways and data validation by qPCR. CK refers to the mock control, and VE4, VE12, VE24, and VE48 refer to 4, 12, 24, and 48 h after *V. dahliae* inoculation. JAZ, jasmonate ZIM-domain; NPR1, non-expressor of pathogenesis-related protein 1; ACO, 1-aminocyclopropane-1-carboxylate oxidase; ETR, ethylene receptor; ERF, ethylene response factor; ARF, auxin response factor; ABA, ABA-deficient gene; ABI, ABA-insensitive gene; GA4, GA 3-β hydroxylase; GA20O, GA 20-oxidase.

Fig. S4. Expression of TFs identified by RNA-Seq (A) and validation by qPCR (B). CK refers to the mock control,

and VE4, VE12, VE24, and VE48 refer to 4, 12, 24, and 48 h after inoculation with *V. dahliae*.

Fig. S5. Detailed expression patterns of defence-responsive genes identified in the phenylpropanoid pathway and data validation by qPCR. CK refers to the mock control, and VE4, VE12, VE24, and VE48 refer to 4, 12, 24, and 48 h after inoculation with *V. dahliae*. PAL, phenylalanine ammonia-lyase; C4H, cinnamate 4-hydroxylase; 4CL, 4-coumarate:CoA ligase; CHS, chalcone synthase; CHI, chalcone isomerase; IFR, isoflavone reductase; FS, flavone synthases; F3H, flavanone 3-hydroxylase; FLS, flavonol synthases; DFR, dihydroflavonol-4-reductase; ANS, anthocyanidin synthase; ANR, anthocyanidin reductase; UFGT, UDP-flavonoid glucosyltransferase; CCR, cinnamoyl CoA reductase; CCoAOMT, caffeoyl-CoA O-methyltransferase.

Fig. S6. Detailed expression profiles of genes in the major lignin biosynthetic pathway. The absolute gene expression levels were obtained by RT-PCR and the original TPM of RNA-Seq. CK refers to the mock control, and 4, 12, 24, and 48 refer to 4, 12, 24, and 48 h after inoculation with *V. dahliae*. PAL, phenylalanine ammonia-lyase; 4CL, 4-coumarate:CoA ligase; C4H, cinnamate 4-hydroxylase; CHS, chalcone synthase; CCR, cinnamoyl CoA reductase; CCoAOMT, caffeoyl-CoA O-methyltransferase; CAD, cinnamyl alcohol dehydrogenase.

Table S1. Primers used for qPCR

Table S2. Primers used for RT-PCR

References

Ansorge WJ. 2009. Next-generation DNA sequencing techniques. *Nature Biotechnology* **25,** 195–203.

Audic S, Claverie JM. 1997. The significance of digital gene expression profiles. *Genome Research* **7,** 986–995.

Azaiez A, Boyle B, Levee V, Seguin A. 2009. Transcriptome profiling in hybrid poplar following interactions with Melampsora rust fungi. *Molecular Plant-Microbe Interaction* **22,** 190–200.

Bari R, Jones JD. 2009. Role of plant hormones in plant defence responses. *Plant Molecular Biology* **69,** 473–488.

Bethel G, Robertson AJ, Perkins AC, et al. 2009. Next is now: new technologies for sequencing of genomes, transcriptomes, and beyond. *Current Opinion in Plant Biology* **12,** 107–118.

Bhuiyan NH, Selvaraj G, Wei YD, King J. 2009. Role of lignification in plant defense. *Plant Signaling and Behavior* **4,** 158–159.

Billa E, Tollier MT, Monties B. 1996. Characterisation of the monomeric composition of *in situ* wheat straw lignins by alkaline nitrobenzene oxidation: effect of temperature and reaction time. *Journal of the Science of Food and Agriculture* **72,** 250–256.

Birzele F, Schaub J, Rust W, Clemens C, Baum P, Kaufmann H, Weith A, Schulz TW, Hildebrandt T. 2010. Into the unknown: expression profiling without genome sequence information in CHO by next generation sequencing. *Nucleic Acids Research* **38,** 3999–4010.

Bugbee WM. 1970. Effect of Verticillium wilt on cotton yield, fiber properties and seed quality. *Crop Science* **10,** 649–652.

Bustin SA, Benes V, Garson JA, et al. 2009. The MIQE guidelines: minimum information for publication of quantitative real-time PCR experiments. *Clinical Chemistry* **55,** 611–622.

Cai YF, He XH, Mo JC, Sun Q, Yang JP, Liu JG. 2009. Molecular research and genetic engineering of resistance to Verticillium wilt in cotton: a review. *African Journal of Biotechnology* **8,** 7363–7372.

Cloonan N, Forrest ARR, Kolle G, Gardiner BBA, Faulkner GJ, Brown MK, Taylor DF, Steptoe AL, Wani S, Grimmond SM. 2008. Stem cell transcriptome profiling via massive-scale mRNA sequencing. *Nature Methods* **5,** 613–619.

Cu YX, Bell AA, Joost O, Magill CW. 2000. Expression of potential defense response genes in cotton. *Physiological and Molecular Plant Pathology* **56,** 25–31.

Dixon RA, Achnine L, Kota P, Liu CJ, Reddy MS, Wang L. 2002. The phenylpropanoid pathway and plant defence—a genomics perspective. *Molecular Plant Pathology* **3,** 371–390.

Dubery IA, Smit F. 1994. Phenylalanine ammonia-lyase from cotton (*Gossypium hirsutum*) hypocotyls: properties of the enzyme induced by a *Verticillium dahliae* phytotoxin. *Biochimica et Biophysica Acta* **1207,** 24–30.

Eudes A, Pollet B, Sibout R, Do CT, Seguin A, Lapierre C, Jouanin L. 2006. Evidence for a role of AtCAD 1 in lignification of elongating stems of *Arabidopsis thaliana*. *Planta* **225,** 23–39.

Ferrer JL, Austin MB, Stewart CJ, Noel JP. 2008. Structure and function of enzymes involved in the biosynthesis of phenylpropanoids. *Plant Physiology and Biochemistry* **46,** 356–370.

Fradin EF, Thomma BPHJ. 2006. Physiology and molecular aspects of Verticillium wilt diseases caused by V. dahliae and *V. albo-atrum*. *Molecular Plant Pathology* **7,** 71–86.

Gallego-Giraldo L, Jikumaru Y, Kamiya Y, Tang Y, Dixon RA. 2011. Selective lignin downregulation leads to constitutive defense response expression in alfalfa (*Medicago sativa* L.). *New Phytologist* **190,** 627–639.

Gan QA, Chepelev I, Wei G, Tarayrah L, Cui KR, Zhao KJ, Chen X. 2010. Dynamic regulation of alternative splicing and chromatin structure in *Drosophila* gonads revealed by RNA-seq. *Cell Research* **20,** 763–783.

Gayoso C, Pomar F, Novo-Uzal E, Merino F, Martinez de Ilarduya O. 2010. The Ve-mediated resistance response of the tomato to *Verticillium dahliae* involves H_2O_2, peroxidase and lignins and drives PAL gene expression. *BMC Plant Biology* **10,** 232–251.

Gou XP, He K, Yang H, Yuan T, Lin HH, Clouse SD, Li J. 2010. Genome-wide cloning and sequence analysis of leucine-rich repeat receptor-like protein kinase genes in *Arabidopsis thaliana*. *BMC Genomics* **11,** 19–34.

Hill MK, Lyon KJ, Lyon BR. 1999. Identification of disease response genes expressed in *Gossypium hirsutum* upon infection with the wilt pathogen. *Verticillium dahliae. Plant Molecular Biology* **40,** 289–296.

Huang J, Gu M, Lai Z, Fan B, Shi K, Zhou YH, Yu JQ, 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.

Jones JD, Dangl JL. 2006. The plant immune system. *Nature* **444,** 323–329.

Kamal ME. 1985. Integrated control of Verticillium wilt of cotton. *Plant Disease* **69,** 1025–1032.

Kandan A, Commare RR, Nandakumar R, Ramiah M, Raguchander T, Samiyappan R. 2002. Induction of phenylpropanoid metabolism by Pseudomonas fluorescens against tomato spotted wilt virus in tomato. *Folia Microbiologica (Praha)* **47,** 121–129.

Kawchuk LM, Hachey J, Lynch DR, et al. 2001. Tomato *Ve* disease resistance genes encode cell surface-like receptors. *Proceedings of the National Academy of Sciences, USA* **98,** 6511–6515.

Kunkel BN, Brooks DM. 2002. Cross talk between signaling pathways in pathogen defense. *Current Opinion in Plant Biology* **5,** 325–331.

Langmead B, Trapnell C, Pop M, Salzberg SL. 2009. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biology* **10,** R25.

Lapierre C, Pollet B, Rolando C. 1995. New insights into the molecular architecture of hardwood lignins by chemical degradative methods. *Research on Chemical Intermediates* **21,** 397–412.

Liu CJ, Blount JW, Steele CL, Dixon RA. 2002. Bottlenecks for metabolic engineering of isoflavone glycoconjugates in *Arabidopsis*. *Proceedings of the National Academy of Sciences, USA* **99,** 14578–14583.

Luo P, Wang YH, Wang GD, Essenberg M, Chen XY. 2001. Molecular cloning and functional identification of (+)-delta-cadinene-8-hydroxylase, a cytochrome P450 mono-oxygenase (CYP706B1) of cotton sesquiterpene biosynthesis. *The Plant Journal* **28,** 95–104.

Makkar HP, Siddhuraju P, Becker K. 2007. Plant secondary metabolites. *Methods in Molecular Biology* **393,** 1–122.

Marjamaa K, Kukkola EM, Fagerstedt KV. 2009. The role of xylem class III peroxidases in lignification. *Journal of Experiment Botany* **60,** 367–376.

Mcfadden HG, Chapple R, Defeyter R, Dennis E. 2001. Expression of pathogenesis-related genes in cotton stems in response to infection by *Verticillium dahliae*. *Physiological and Molecular Plant Pathology* **58,** 119–131.

Morrissy AS, Morin RD, Delaney A, Zeng T, McDonald H, Jones S, Zhao Y, Hirst M, Marra MA. 2009. Next-generation tag sequencing for cancer gene expression profiling. *Genome Research* **19,** 1825–1835.

Mortazavi A, Williams BA, McCue K, Schaeffer L, Wold B. 2008. Mapping and quantifying mammalian transcriptomes by RNA-Seq. *Nature Methods* **5,** 621–628.

Nagalakshmi U, Wang Z, Waern K, Shou C, Raha D, Gerstein M, Snyder M. 2008. The transcriptional landscape of the yeast genome defined by RNA sequencing. *Science* **320,** 1344–1349.

Naoumkina MA, Zhao Q, Gallego-Giraldo L, Dai X, Zhao PX, Dixon RA. 2010. Genome-wide analysis of phenylpropanoid defence pathways. *Molecular Plant Pathology* **11,** 829–846.

Otto TD, Wilinski D, Assefa S, et al. 2010. New insights into the blood-stage transcriptome of Plasmodium falciparum using RNA-Seq. *Molecular Microbiology* **76,** 12–24.

Panstruga R, Parker JE, Schulze-Lefert P. 2009. SnapShot: plant immune response pathways. *Cell* **136, 978.**e1–3.

Pomar F, Novo M, Bernal MA, Merino F, Barcelo AR. 2004. Changes in stem lignins (monomer composition and crosslinking) and peroxidase are related with the maintenance of leaf photosynthetic integrity during *Verticillium* wilt in *Capsicum annuum*. *New Phytologist* **163,** 111–123.

Rieu I, Powers SJ. 2009. Real-time quantitative RT-PCR: design, calculations, and statistics. *The Plant Cell* **21,** 1031–1033.

Sal'kova EG, Guseva NN. 1965. The role of pectolytic enzymes of the *Verticillium dahliae* fungus in the development of cotton wilt. *Doklady Akademii Nauk SSSR* **163,** 515–522.

Scheideler M, Schlaich NL, Fellenberg K, Beissbarth T, Hauser NC, Vingron M, Slusarenko AJ, Hoheisel JD. 2002. Monitoring the switch from housekeeping to pathogen defense metabolism in *Arabidopsis thaliana* using cDNA arrays. *Journal of Biological Chemistry* **277,** 10555–10561.

Sink KC, Grey WE. 1999. A root-injection method to assess verticillium wilt resistance of peppermint and its use in identifying resistant somaclones of cv. Black Mitcham. *Euphytica* **106,** 223–230.

Singh K, Foley RC, Onate-Sanchez L. 2002. Transcription factors in plant defense and stress responses. *Current Opinion in Plant Biology* **5,** 430–436.

Smit F, Dubery LA. 1997. Cell wall reinforcement in cotton hypocotyls in response to a *Verticillium dahliae* elicitor. *Phytochemistry* **44,** 811–815.

Sturn A, Quackenbush J, Trajanoski Z. 2002. Genesis: cluster analysis of microarray data. *Bioinformatics* **18,** 207–208.

Tan XP, Liang WQ, Liu CJ, Luo P, Heinstein P, Chen XY. 2000. Expression pattern of (+)-delta-cadinene synthase genes and biosynthesis of sesquiterpene aldehydes in plants of *Gossypium arboreum* L. *Planta* **210,** 644–651.

Tu LL, Zhang XL, Liang SG, Liu DQ, Zhu LF, Zeng FC, Nie YC, Guo XP, Deng FL, Tan JF, Xu L. 2007. Genes expression analyses of sea-island cotton (*Gossypium barbadense* L.) during fiber development. *Plant Cell Reports* **26,** 1309–1320.

Uppalapati SR, Marek SM, Lee HK, Nakashima J, Tang Y, Sledge MK, Dixon RA, Mysore KS. 2009. Global gene expression profiling during *Medicago truncatula– Phymatotrichopsis omnivora* interaction reveals a role for jasmonic acid, ethylene, and the flavonoid pathway in disease development. *Molecular Plant-Microbe Interaction* **22,** 7–17.

Verhage A, van Wees SCM, Pieterse CMJ. 2010. Plant immunity: it's the hormones talking, but what do they say? *Plant Physiology* **154,** 536–540.

Wang Z, Gerstein M, Snyder M. 2009. RNA-Seq: a revolutionary tool for transcriptomics. *Nature Reviews Genetics* **10,** 57–63.

Xu L, Zhu LF, Tu LL, Guo XP, Long L, Sun LQ, Gao W, Zhang XL. 2011. Differential gene expression in the cotton defense response to *Verticillium dahliae* by SSH. *Journal of Phytopathology* (in press).

Xu Q, Yu K, Zhu AD, Ye JL, Liu Q, Zhang JC, Deng XX. 2009. Comparative transcripts profiling reveals new insight into molecular processes regulating lycopene accumulation in a sweet orange (*Citrus sinensis*) red-flesh mutant. *BMC Genomics* **10,** 540–555.

Xu YH, Wang JW, Wang S, Wang JY, Chen XY. 2004. Characterization of *GaWRKY1*, a cotton transcription factor that regulates the sesquiterpene synthase gene (+)-delta-cadinene synthase-A. *Plant Physiology* **135,** 507–515.

Yang Y, Shah J, Klessig DF. 1997. Signal perception and transduction in plant defense responses. *Genes and Development* **11,** 1621–1639.

Zabala G, Zou J, Tuteja J, Gonzalez DO, Clough SJ, Vodkin LO. 2006. Transcriptome changes in the phenylpropanoid pathway of *Glycine max* in response to *Pseudomonas syringae* infection. *BMC Plant Biology* **6,** 26–44.

Zenoni S, Ferrarini A, Giacomelli E, Xumerle L, Fasoli M, Malerba G, Bellin D, Pezzotti M, Delledonne M. 2010. Characterization of transcriptional complexity during berry development in *Vitis vinifera* using RNA-Seq. *Plant Physiology* **152,** 1787–1795.

Zhang S, Klessig DF. 2001. MAPK cascades in plant defense signaling. *Trends in Plant Science* **6,** 520–527.

Zuo K, Wang J, Wu W, Chai Y, Sun X, Tang K. 2005. Identification and characterization of differentially expressed ESTs of *Gossypium barbadense* infected by *Verticillium dahliae* with suppression subtractive hybridization. *Molecular Biology* **39,** 214–223.