



# Analysis of upland cotton (*Gossypium hirsutum*) response to *Verticillium dahliae* inoculation by transcriptome sequencing

B.X. Shao, Y.L. Zhao, W. Chen, H.M. Wang, Z.J. Guo, H.Y. Gong, X.H. Sang, Y.L. Cui and C.H. Wang

State Key Laboratory of Cotton Biology, Institute of Cotton Research of CAAS, Anyang, China

Corresponding author: H.M. Wang  
E-mail: aywhm@163.com

Genet. Mol. Res. 14 (4): 13120-13130 (2015)

Received April 2, 2015

Accepted July 23, 2015

Published October 26, 2015

DOI <http://dx.doi.org/10.4238/2015.October.26.8>

**ABSTRACT.** *Verticillium* wilt is one of the main diseases in cotton (*Gossypium hirsutum*), severely reduces yield and fiber quality, and is difficult to be controlled effectively. At present, the molecular mechanism that confers resistance to this disease is unclear. Transcriptome sequencing is an important method to detect resistance genes, explore metabolic pathways, and study resistance mechanisms. In this study, the transcriptome of a disease-resistant inbred cotton line inoculated with *Verticillium dahliae* was sequenced. A total of 126,402 unigenes were obtained using *de novo* assembly and data analysis, 99,712 (78.88%) of which were annotated into the Nr, Nt, Swiss-Prot, KEGG, COG, and GO databases. The expression patterns of 16 candidate disease-resistance genes showed that some genes were upregulated soon after *V. dahliae* inoculation and others were upregulated later, which may indicate instantaneous basal defense and lagged specific defense, respectively. We conducted a preliminary analysis of the transcriptome database, which will contribute to further research regarding the cloning of disease-resistance genes.

**Key words:** Cotton; *Verticillium* wilt; Transcriptome sequencing; Disease-resistance gene

## INTRODUCTION

Verticillium wilt is the most destructive disease in cotton production, and is caused by *Verticillium dahliae* (Sal'kova and Guseva, 1965). It can cause yellowing, wilt, defoliation, and even death (Sink and Grey, 1999). It has become one of the main factors that restrict cotton production in China, because it is difficult to obtain high resistance to this disease in upland cotton (*Gossypium hirsutum*) varieties by traditional breeding methods (Li et al., 2008), due to the lack of highly resistant germplasm resources. In recent years, the rapid development of genetic engineering technology has made research on resistance mechanisms against Verticillium wilt increasingly molecular. Considerable progress has been made on the identification of molecular markers associated with resistance, and resistance genes. These studies have provided technical support and a theoretical basis for the further study of the molecular mechanisms of cotton's response to Verticillium wilt.

Currently, a new generation of high-throughput sequencing technology has greatly reduced the cost and time needed for genome sequencing, which has brought a new efficient platform and great developing opportunities for genomics (Yue et al., 2012; Zhang et al., 2013). Transcriptome sequencing refers to sequencing the cDNA, which is reverse transcription from all mRNA in tissues or cells. This technique can analyze the structures and expression levels of transcripts, and can simultaneously detect unknown transcripts, rare transcripts, and alternative splicing. It can provide more comprehensive and accurate transcriptome information for researchers. For species that have reference genomes, transcriptome sequencing can perfect and validate genome annotation information; for species without reference genomes, it can perform *de novo* assembly (Liang et al., 2011; Li et al., 2013). At present, transcriptome sequencing is one of the most effective techniques to detect disease-resistance genes, and to understand the molecular mechanism of disease resistance in plants.

In this study, samples from a disease-resistant inbred line were used as materials; one group of samples was inoculated with *V. dahliae* and the other was inoculated with water. The infected samples were sequenced to study the mechanism of disease resistance against Verticillium wilt, and to detect important genes and biological pathways related to disease resistance in cotton.

## MATERIAL AND METHODS

### Preparation of materials

The disease-resistant inbred line 1017012 was created from a cross between upland cotton (TM-1) and disease-resistant upland cotton, Chang Kang Mian. Disease resistance was verified by testing in multiple years and multiple sites. A highly aggressive strain of *V. dahliae*, Vd080, from the Institute of Cotton Research of the Chinese Academy of Agriculture Sciences was used for inoculation.

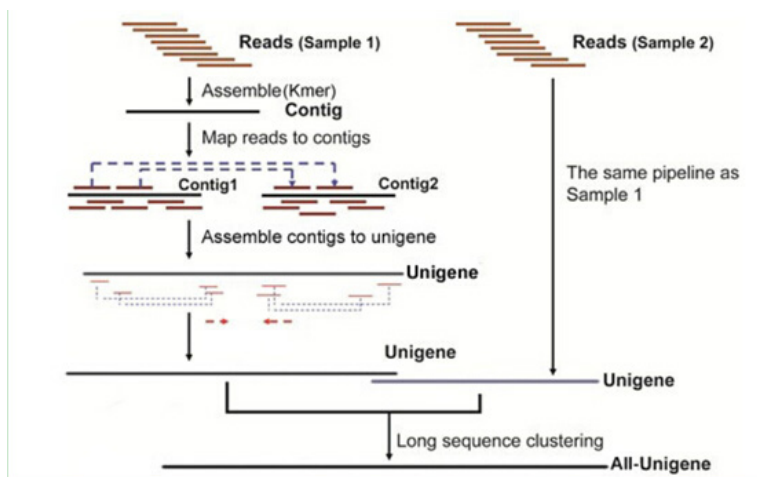
Delinted cotton seeds were immersed in water for 16 h, and then seven to eight seeds were planted separately in paper bowls filled with sterilized soil in a greenhouse. The plants were inoculated when the first true leaves had unfolded. Each seedling was inoculated with 10 mL *V. dahliae* Vd080 ( $10^7$  spore/mL) or water (control). The roots and hypocotyls of the infected plants (INF) and the control plants (CON) were collected at 0, 4, 8, 12, 24, and 48 h after inoculation.

## cDNA library preparation

Total RNA was extracted based on a modified guanidine thiocyanate method (Wang et al., 2013). Six RNA samples (from all treatment time points 0,4,8,12,24 and 48h) from the INF and were mixed in equal amount, and the treatment for the six samples of the CON. The quality and quantity of the two mixed RNA samples were checked by agarose gel electrophoresis using a NanoDrop™ 2000 micro-volume ultraviolet spectrophotometer. mRNA was purified by biotin-Oligo (dT) magnetic bead adsorption, and fragmentation buffer was added to disrupt the mRNA into 200-bp fragments. First-strand cDNA was synthesized using 200-bp fragments as templates with a random hexamer primer, and second-strand cDNA was synthesized by base pairing. The cohesive ends of the resulting short fragments were repaired, and poly(A) tails were added to the 3' ends. The sequencing adapters were then ligated. Following electrophoresis, suitable fragments were selected as templates for polymerase chain reaction (PCR). Finally, the cDNA library was constructed.

## RNA-seq and *de novo* assembly

Transcriptome sequencing was conducted after the cDNA library was constructed. The samples were loaded on an Illumina flow cell and sequenced at ultrahigh throughput using an Illumina HiSeq™ 2000. After the cDNA library was sequenced, the raw image data were transformed into sequence information using Illumina pipeline software. For *de novo* assembly the quality requirement of reads is high, because low-quality reads (reads containing ambiguous sequences “N” and adaptor sequence fragments) can influence the assembly of the short read. Therefore, a stringent filtering process was conducted and low-quality reads were cleaned out. Finally, the Q20 percentage (base quality higher than 20) was up to standard (>80%). Trinity, a short-read assembly program (Grabherr et al., 2011), was selected for *de novo* transcriptome assembly. High-quality reads with overlaps were then assembled to generate contigs, which were connected with the fewest “N” until they could not be extended at either end; such sequences were named unigenes. All of the unigenes were obtained by integrating the unigenes from the two samples. The steps of assembly are shown in Figure 1.



**Figure 1.** Assembly steps.

## Functional annotation and classification

All of the unigenes were checked in public databases to search for homologous sequences and to predict their function. Protein databases, including the non-redundant protein database (Nr), Swiss-Prot protein database, Kyoto encyclopedia of genes and genomes (KEGG), and the cluster of orthologous groups (COG), were checked by BlastX with an E-value threshold of  $10^{-5}$ . The nucleotide sequence database (Nt) was checked by BLASTn with an E-value threshold of  $10^{-5}$ . The best alignments were used to predict unigene function. Gene Ontology (GO) annotation was performed by Blast2GO (Conesa et al., 2005), and GO function classification was conducted using WEGO (<http://wego.genomics.org.cn/cgi-bin/wego/index.pl>) at the macro-scale level (Ye et al., 2006).

## Real-time PCR analysis

Unigenes homologous with R genes (resistant genes) were obtained by searching homologous sequences in the transcriptome-sequencing database with R genes cloned from other crops. In order to confirm that the unigenes were disease-resistance genes induced by *V. dahliae*, real-time PCR was performed. We estimated that 4, 12, 24, and 48 h post-inoculation were the key time points for an immune reaction, so the gene expression levels at these four time points were analyzed by real-time PCR. The cDNA was synthesized using a PrimeScript™ RT reagent kit (TaKaRa, Dalian, China). Specific primers for the 16 unigenes were designed using Primer 5.0, and their sequences are shown in Table 1. Actin (F, AATGTGCCTGCCATGTATGT; R, TCACACCATCACCAGAATCC) was used as an internal control gene for data normalization. Real-time PCR was performed using a SYBR® *Premix Ex Taq*™ kit (TakaRa) according to the manufacturer protocol, in an Applied Biosystems 7500 Fast Real-Time PCR system. The total reaction volume was 20  $\mu$ L (Table 2). The following thermal cycling profile was used: 95°C for 30 s, 40 cycles at 95°C for 3 s, and 40 cycles at 60°C for 30 s. Three replicates were included for each sample. The fold change in mRNA expression was estimated using threshold cycles by the  $\Delta\Delta$ CT method.

**Table 1.** Candidate disease resistant unigenes for real-time RT-PCR.

Gene ID	Annotation	Forward primer	Reverse primer
CL5674	TGF- $\beta$ receptor associated protein 1	TTGATGCTGCTGAAAAGAACG	GCACAAGCCGATGAAACACT
CL7420	nbs-lrr resistance protein	TGATTATCTACCTCCTCTTGTC	TCTAAACTTGAATGGGCTT
Unigene39385	NBS-type resistance protein	GACAACCCGTTCCGGAGCAT	GTCAGAGGTAGACCCGCAAT
CL1314	tir-nbs-lrr resistance protein	CCGATGTTGCCTTCTTGT	CGGGTATGTTGCTGAAACTGTT
Unigene43606	BED finger-nbs resistance protein	GAGATACCCGCTGGACTTTT	CTCCTCTGCTTTAGACTTGTT
Unigene66319	BED finger-nbs-lrr resistance protein	TTGTGAAGGGATAGAGTGTGTTGTTT	CAGACGGAGCCGATGTTGAT
Unigene76084	cc-nbs-lrr resistance protein	TGAGGCTAAAAGGCAAACCTGA	GCCCCATTGAGGACACCAT
Unigene40425	Putative disease resistance protein	TTGGCTTACGATGTGGATGA	CAGTACAGCAGGTTGGAATGAGT
CL7322	Putative disease resistance RPP13-like protein 1-like	CGTACCATCAACGAGTCTT	GGGTCCCTGCCTCCAAAC
Unigene76503	Disease resistance protein RPS2	TTTACTTGAGTGTGGGCTG	GACCTGTCGGATGCTGTT
Unigene40900	Disease resistance protein RGA2	CAAACCTGGAAGGAGTGGGACA	GGAAGGAAGATGGTAGGCAA
Unigene45880	cc-nbs-lrr resistance protein	ACGGAATAATGGAGGCACTAAA	GCTATGACAACCGAATCAAACA
Unigene74938	cc-nbs-lrr resistance protein	CTGGAGAAAGTGAACGCAT	CAAGGGACAACCCCAAATC
Unigene42468	cc-nbs-lrr resistance protein	ATGTTGTTGGGATGCTGCTT	CATTTTCGCCTTATCGGACC
Unigene71561	cc-nbs-lrr resistance protein	GGATGGTGTGAAAGTTCTGAC	GACGATGGCTTGCTTAGTTGT
CL16062	TMV resistance protein N	GGGCTCATTTTTGTGCGA	CTCCTTTTCTTCCAATCCATCA

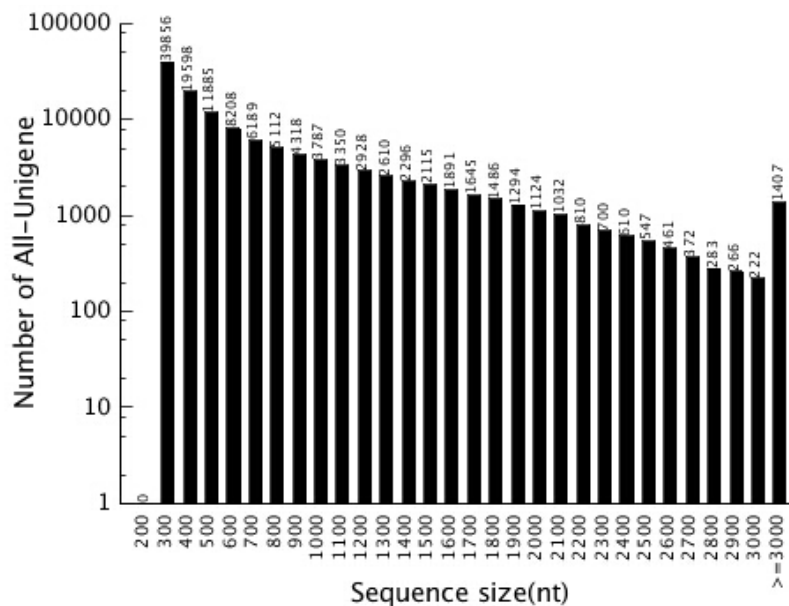
**Table 2.** Reaction volume of real-time PCR.

Reagent	Volume ( $\mu\text{L}$ )
SYBR Premix Ex Taq (Tli RNaseH Plus) (2X)	10.0
Forward Primer (10 $\mu\text{M}$ )	0.4
Reverse Primer (10 $\mu\text{M}$ )	0.4
ROX Reference Dye II (50X)	0.4
Template	2.0
dH <sub>2</sub> O	6.8
Total	20.0

## RESULTS

### *De novo* assembly

A total of 58,013,440 and 57,344,616 raw reads were generated in the INF and CON libraries, respectively. After filtering out low-quality tags (tags containing “N” and adaptor sequence fragments) by stringent quality checks and data cleaning, 53,265,960 clean reads with a Q20 of 97.4% in the INF library, and 53,179,826 clean reads with a Q20 of 97.64% in the CON library, were obtained. The *de novo* assembly of these clean reads produced 170,765 contigs in the INF library and 144,653 contigs in the CON library. The contigs were then further assembled into unigenes. A total of 83,279 unigenes with a mean size of 978 bp in the INF library and 70,669 unigenes with a mean size of 1046 bp in the CON library were created. There were 12,610 more unigenes in the INF than in the CON library; combining the two libraries produced 126,402 unigenes, and these unigenes were named All-Unigene. The sequence size distribution is presented in Figure 2; 71,339 (56.44%) unigenes were 100-500 bp long and 43.55% unigenes were longer than 500 bp.

**Figure 2.** Length distribution of unigenes.

## Annotation

To annotate the assembled unigenes, a basic local alignment search tool (BLAST) alignment with an E-value threshold of  $10^{-5}$  was conducted in Nr, Nt, Swiss-Prot, KEGG, and COG (Altschul et al., 1997; Cameron et al., 2004). The results indicated that of the 126,402 unigenes, 99,712 (78.88%) were annotated, and 93,307 (73.82%), 86,638 (68.54%), 55,955 (44.27%), 49,247 (38.96%), and 28,925 (22.88%) of which were matched to Nr, Nt, Swiss-Prot, KEGG, and COG, respectively. However, 26,690 (21.12%) unigenes were not aligned to any of the public databases, and ESTScan (<http://estscan.sourceforge.net/>) was used to predict their coding regions and decide their sequence direction.

## Functional classification by COG

COG is a database for protein orthology classification. Each COG protein is assumed to have evolved from an ancestor protein, and the database is based on bacteria, algae, and eukaryotes with complete genome-encoded proteins and phylogenetic relationships. All of the unigenes were aligned to the COG database to predict and classify their possible functions. A total of 28,925 were assigned COG classifications (Figure 3). Among the 25 COG categories, the “General function prediction only” (9241 unigenes, 31.94%) was the largest group, followed by “Transcription” (5059 unigenes, 17.49%), “Replication, recombination, and repair” (4586 unigenes, 15.85%), “Posttranslational modification, protein turnover, and chaperones” (4013 unigenes, 13.87%), “Signal transduction mechanisms” (3984 unigenes, 13.77%), and “Secondary metabolites biosynthesis” (1598 unigenes, 5.52%). A few unigenes were assigned to “Extracellular structure” and “Nuclear structure”

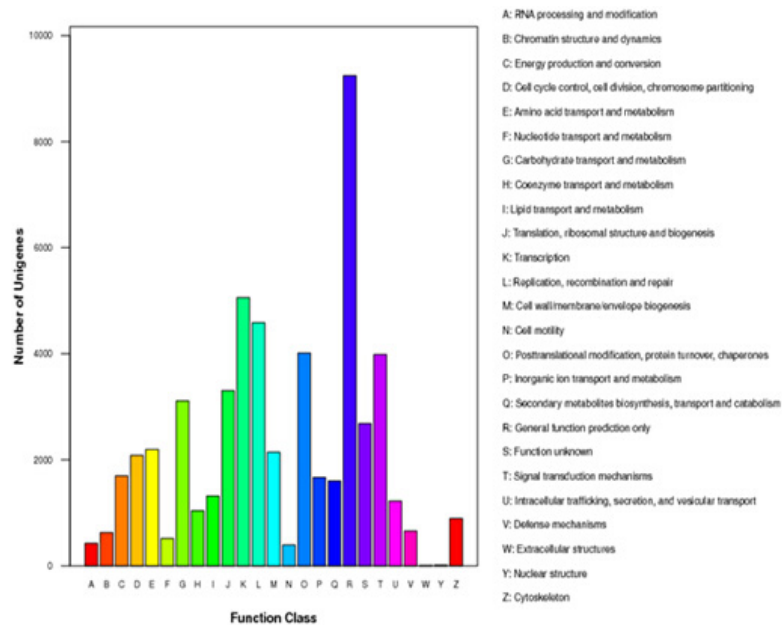


Figure 2. COG function classification of all-unigene.fa sequence.

## Functional classification by GO

GO is an international, standardized gene function classification system. It contains three ontologies: “Biological process”, “Cellular component”, and “Molecular function”. A total of 73,831 unigenes were categorized into 57 gene ontology classes (Figure 4). “Biological process” contained 24 subcategories, represented by “Cellular process” (48,224 unigenes, 65.32%) and “Metabolic process” (45,817 unigenes, 62.06%). “Cellular component” contained 17 subcategories, represented by “Cell” (57,950 unigenes, 78.49%) and “Cell part” (57,950 unigenes, 78.49%). “Molecular function” contained 16 subcategories, represented by “Binding” (36,561 unigenes, 49.52%) and “Catalytic activity” (35,692 unigenes, 48.34%). A few genes were clustered in “Carbon utilization”, “Locomotion”, “Extracellular matrix”, “Extracellular matrix part”, “Extracellular region part”, “Virion”, “Virion part”, “Channel regulator activity”, “Metallochaperone activity”, “Nutrient reservoir activity”, “Protein tag”, and “Translation regulator activity”.

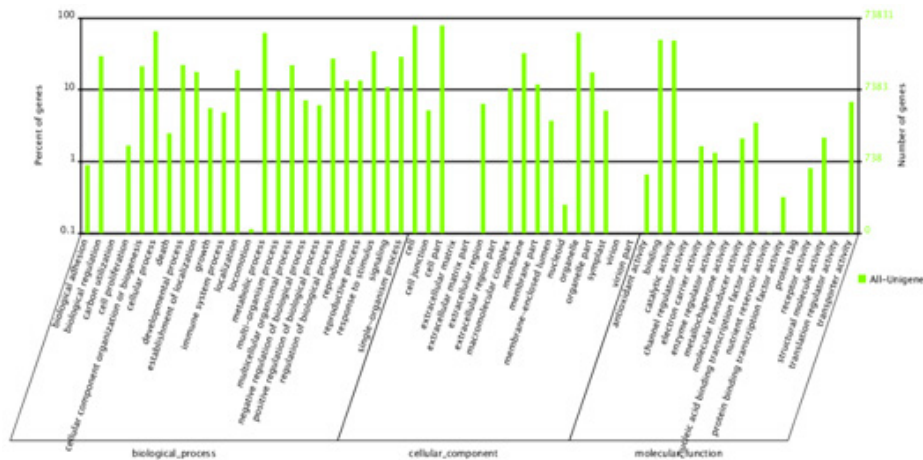


Figure 4. GO classification.

## Pathway analysis of unigenes

Pathway analysis identifies unigenes involved in biochemical metabolism pathways. A total of 49,247 unigenes could be matched to the KEGG pathway database, and were assigned to 128 pathways. The first 28 pathways are shown in Table 3. The main pathways were “Metabolic pathways” (22.16%), “Biosynthesis of secondary metabolites” (10.9%), “Plant-pathogen interaction” (6.11%), “Plant hormone signal transduction” (6.01%), “RNA transport” (3.72%), and “Spliceosome” (3.04%).

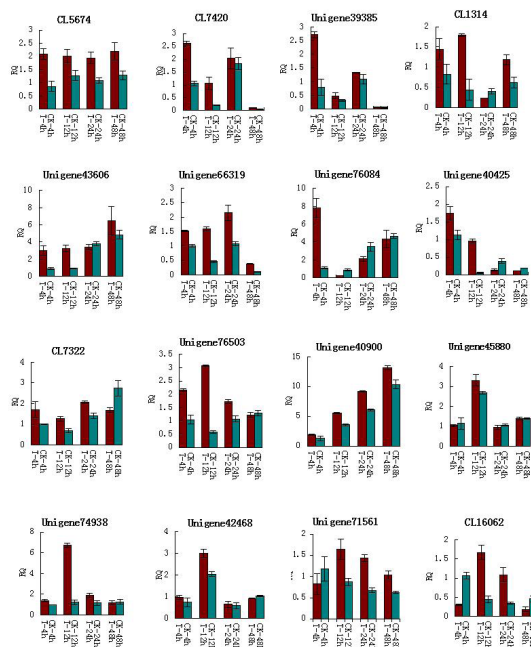
## Real-time PCR analysis

Sixteen candidate disease-resistance unigenes were homologous with R genes in other plants (Table 1). All of them contained conserved structural domains, such as nucleotide binding sites (NBSs) or leucine-rich repeats (LRRs). Two trends were observed at 4, 12, 24, and 48 h post-inoculation with *V. dahliae* (Figure 5). Firstly, 4 h after inoculation the expression levels of CL5674, CL7420, Unigene39385, CL1314, Unigene43606, Unigene66319, Unigene76084, Unigene40425, CL7322,

and Unigene76503 in the INF were significantly higher than those in the CON; secondly, 12 h after inoculation the expression levels of Unigene40900, Unigene45880, Unigene 74938, Unigene42468, Unigene71561, and CL16062 in the INF were significantly higher than those in the CON.

**Table 3.** The pathway of the annotation genes.

ID	Pathway	NO. of unigenes	Pathway ID
1	Metabolic pathways	10912 (22.16%)	ko01100
2	Biosynthesis of secondary metabolites	5368 (10.9%)	ko01110
3	Plant-pathogen interaction	3011 (6.11%)	ko04626
4	Plant hormone signal transduction	2962 (6.01%)	ko04075
5	RNA transport	1831 (3.72%)	ko03013
6	Spliceosome	1498 (3.04%)	ko03040
7	Endocytosis	1373 (2.79%)	ko04144
8	Protein processing in endoplasmic reticulum	1310 (2.66%)	ko04141
9	Starch and sucrose metabolism	1309 (2.66%)	ko00500
10	Glycerophospholipid metabolism	1278 (2.6%)	ko00564
11	Ribosome	1210 (2.46%)	ko03010
12	RNA degradation	1092 (2.22%)	ko03018
13	Purine metabolism	1057 (2.15%)	ko00230
14	mRNA surveillance pathway	1046 (2.12%)	ko03015
15	Ribosome biogenesis in eukaryotes	1045 (2.12%)	ko03008
16	Pyrimidine metabolism	992 (2.01%)	ko00240
17	Ubiquitin mediated proteolysis	930 (1.89%)	ko04120
18	Ether lipid metabolism	887 (1.8%)	ko00565
19	Phenylpropanoid biosynthesis	791 (1.61%)	ko00940
20	Pentose and glucuronate interconversions	655 (1.33%)	ko00040
21	Amino sugar and nucleotide sugar metabolism	653 (1.33%)	ko00520
22	Glycolysis / Gluconeogenesis	645 (1.31%)	ko00010
23	Oxidative phosphorylation	600 (1.22%)	ko00190
24	ABC transporters	555 (1.13%)	ko02010
25	Circadian rhythm - plant	548 (1.11%)	ko04712
26	Phagosome	541 (1.1%)	ko04145
27	RNA polymerase	507 (1.03%)	ko03020
28	Pyruvate metabolism	494 (1%)	ko00620



**Figure 5.** Real-time RT-PCR analysis for 16 candidate disease resistant Unigenes.



## DISCUSSION

Verticillium wilt poses a serious threat to the development of the cotton industry in China. At present, the most effective prevention measure is cultivating resistant varieties. Due to a lack of resistant resources in upland cotton, the progress of breeding resistant varieties using conventional breeding methods has been slow. Therefore, molecular breeding has become an important method to breed disease-resistant varieties. In recent years, researchers have investigated the mechanism of disease resistance, detected resistance genes, and screened molecular markers linked with resistance genes. Although a series of physiological and biochemical studies on cotton Verticillium wilt have been conducted, and some Verticillium wilt resistance genes have been cloned, the molecular mechanism of cotton resistance to *V. dahliae* is still unclear, because the disease-resistance mechanism is complex and there are several resistance genes involved. In this study, transcriptome sequencing was performed in order to discover disease-resistance genes and metabolic pathways related to cotton Verticillium wilt.

After being sequenced and assembled, 83,279 unigenes with a mean size of 978 bp in the INF library and 70,669 unigenes with a mean size of 1046 bp in the CON library were obtained; 12,610 more unigenes were observed in the INF than in the CON library; these may have been induced by *V. dahliae*. A total of 126,402 unigenes were obtained from the two libraries; 99,712 (78.88%) unigenes were aligned in the Nr, Nt, Swiss-Prot, KEGG, COG, and GO databases, whereas 26,690 (21.12%) unigenes were not aligned, and may represent new transcripts that greatly enrich the upland cotton genome database.

Pathway analysis revealed that 49,247 unigenes (matched to the KEGG database) were involved in 128 metabolic pathways, 10,912 (22.16%) of which were involved in "Metabolic pathways", 5,368 (10.9%) in "Biosynthesis of secondary metabolites", 3,011 (6.11%) in "Plant-pathogen interaction", 2,962 (6.01%) in "Plant hormone signal transduction", 1,309 (2.66%) "Starch and sucrose metabolism", 791 (1.61%) in "Phenylpropanoid biosynthesis", 645 (1.31%) in "Glycolysis/Gluconeogenesis", and 507 (1.03%) in "RNA polymerase". "Biosynthesis of secondary metabolites", "Plant hormone conduction", and "Plant-pathogen interaction" are closely related to plant immune defense reactions (Chen and Zhao, 2003; Jia, 2003). Other metabolic pathways may also be associated with resistance. Pathogen infection can lead to the accumulation of carbohydrates (Luo et al., 2001), and a variety of enzymes in the phenylpropanoid metabolic pathway are related to lignin synthesis, which results in cell-wall thickening (Smit and Dubery, 1997). There is evidence that RNA polymerase is highly expressed during the Verticillium wilt incubation period, and only a part of the mRNA exists after pathogen infection (Plekhanova et al., 1978). The results of these previous studies have indicated that "Starch and sucrose metabolism", "Phenylpropanoid biosynthesis", and "RNA polymerase" might contribute to pathogen defense.

Over 40 resistance genes have been cloned from *Arabidopsis*, maize, rice, and other plants (Dilbirligi et al., 2004). Sequence comparisons have shown that conserved domains exist between the amino acid sequences of R genes, such as NBSs and LRRs. The NBS domain exists in many eukaryotic proteins, can bind to ATP or GTP, and is mainly responsible for ATP hydrolysis and the releasing signal. It plays an important role in disease resistance in plants (Wang et al., 2006). The LRR domains at the amino terminal and the carboxyl terminal have different functions. The LRR at the amino terminal regulates activity, whereas the LRR at the carboxyl terminal identifies bacteria (Belkhadir et al., 2004). Dixon et al. (1996) reported that LRRs can promote the combination of R gene products and other proteins involved in disease-resistance signal transduction.

The 16 unigenes were homologous with R genes in other plants. From the annotation information, we found that most of the unigenes encoded resistant proteins and exhibited a conserved structure, except for CL5674, which was TGF- $\beta$  receptor-associated protein 1. Wan and Flavell (2008) reported that TGF- $\beta$  plays an important role in the control of immunity and autoimmunity, so this unigene might be a candidate disease-resistance gene in response to *V. dahliae*.

Some genes were upregulated at 4 h after *V. dahliae* inoculation and exhibited instantaneous induction, whereas other genes were upregulated 12 h after *V. dahliae* inoculation and exhibited lagged induction. Therefore, cotton defense to Verticillium wilt may occur in two ways: instantaneous basal defense and lagged specific defense (see Chen and Dai, 2007). Expression levels of the 16 candidate disease-resistance genes in CON changed at different time points, possibly due to variations in light and temperature.

### Conflicts of interest

The authors declare no conflict of interest.

### ACKNOWLEDGMENTS

Research supported by the National Science and Technology Supporting Plan (#2011BAD35B05).

### REFERENCES

- Altschul SF, Madden TL, Schaffer AA, Zhang J, et al. (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.* 25: 3389-3402.
- Belkhadir Y, Subramaniam R and Dangl JL (2004). Plant disease resistance protein signaling: NBS-LRR proteins and their partners. *Curr. Opin. Plant Biol.* 7: 391-399.
- Cameron M, Williams HE and Cannane A (2004). Improved gapped alignment in BLAST. *IEEE/ACM Trans. Comput. Biol. Bioinform.* 1: 116-129.
- Chen JY and Dai XF (2007). Expression analysis of cotton anti-Verticillium wilt related genes and preliminary functional research of GhPR10. Master's thesis, Chinese Academy of Agricultural Sciences, Beijing.
- Chen HQ and Zhao SQ (2003). Recent advance in plant disease resistance and acquired resistance system. *Shanxi Agric. Univ.* 23: 286-291.
- Conesa A, Gotz S, Garcia-Gomez JM, Terol J, et al. (2005). Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research. *Bioinformatics* 21: 3674-3676.
- Dilbirli M, Erayman M, Sandhu D, Sidhu D, et al. (2004). Identification of wheat chromosomal regions containing expressed resistance gene. *Genetics* 166: 461-481.
- Dixon MS, Jones DA, Keddle JS, Thomas CM, et al. (1996). The tomato Cf-2 disease resistance locus comprises two functional genes encoding leucine-rich repeat proteins. *Cell* 84: 451-459.
- Grabherr MG, Haas BJ, Yassour M, Thompson DA, et al. (2011). Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nat. Biotechnol.* Doi: 10.1038/nbt.1883.
- Jia YT (2003). Plant disease resistance signaling pathways. *Chin. Bull. Bot.* 20: 602-608.
- Li CW, Ding JP, Liu DM, Zhou RY, et al. (2008). Research progress on cotton Verticillium wilt and resistance breeding. *Cotton Sci.* 20: 385-390.
- Li ZY, Chen LP, Li Y, Han YW, et al. (2013). The next generation sequencing technology and its application in plant transcriptome. *J. Henan Agric. Sci.* 42: 1-5.
- Liang Y, Chen SY and Liu GS (2011). Application of next generation sequencing techniques in plant transcriptome. *Hereditas* 33: 1317-1326.
- Luo P, Wang YH and Wang GD (2001). Molecular cloning and function identification of (+)-delta-cadinene-8-hydroxylase, a cytochrome P450 monooxygenase (CYP706B1) of cotton sesquiterpene biosynthesis. *Plant J.* 28: 95-104.

- Plekhanova LS, Gevorkian AG, Ashirova ZK and Ibragimov AI (1978). Biosynthesis of RNA in the nuclear structure of wilted cotton plant leaves. *Biokhimiia* 43: 407-412.
- Sal'kova EG and Guseva NN (1965). The role of pectolytic enzymes of the *V. dahliae* fungus in the development of cotton wilt. *Dokl. Akad. Nauk. SSSR* 163: 515-522.
- Sink KC and Grey WE (1999). A root-injection method to assess *Verticillium* wilt resistance of peppermint (*Mentha x piperita* L.) and its use in identifying resistant somaclones of cv. Black Mitcham. *Euphytica* 106: 223-230.
- Smit F and Dubery IA (1997). Cell wall reinforcement in cotton hypocotyls in response to a *Verticillium dahliae* elicitor. *Phytochemistry* 44: 811-815.
- Wan YY and Flavell RA (2008). TGF- $\beta$  and regulatory T cell in immunity and autoimmunity. *J. Clin. Immunol.* 28: 647-659.
- Wang CH, Zhao YL, Wang HM, Chen W, et al. (2013). Screening of method to isolate high-quality total RNA from young cotton roots. *Cotton Sci.* 25: 372-376.
- Wang HY, Yang WX and Liu DQ (2006). Isolation and characterization of NBS-LRR resistance gene homology sequences from wheat. *Sci. Agric. Sin.* 39: 1558-1564.
- Ye J, Fang L, Zhang H, Zhang Y, et al. (2006). WEGO: a web tool for plotting GO annotations. *Nucleic Acids Res.* 34: W293-W297.
- Yue GD, Gao Q, Luo LH, Wang JY, et al. (2012). Application of high-throughput sequencing technology in the research field of plant and animal. *Sci. Sin. Vitae* 42: 107-124.
- Zhang QF, Li J, Fang ZX and Yang LQ (2013). Application of high-throughput sequencing technology in agricultural research. *Shandong Agric. Sci.* 45: 137-140.