



OPEN

DATA DESCRIPTOR

# Transcriptomic datasets of Verticillium wilt resistant and non-resistant *Gossypium barbadense* varieties during pathogen inoculation

Xianpeng Xiong<sup>1,5</sup>, Cong Sun<sup>2,5</sup>, Bin Chen<sup>3</sup>, Jie Sun<sup>3</sup>, Cong Fei<sup>4</sup>✉ & Fei Xue<sup>3</sup>✉

Cotton is a significant cash crop and the primary source of natural fiber globally. Among the numerous diseases encountered in cotton production, Verticillium wilt is one of the most serious, caused by the pathogen *Verticillium dahliae* (*V. dahliae*). Unfortunately, there are no effective targeted methods to combat this disease. Genomic resources for Verticillium wilt resistance primarily exist in *Gossypium barbadense* (*G. barbadense*). Regrettably, there have been limited transcriptomic comparisons between *V. dahliae*-resistant and -susceptible varieties of *G. barbadense* due to the scarcity of susceptible resources. In this study, we conducted a transcriptome analysis on both *V. dahliae*-resistant and -susceptible varieties of *G. barbadense* at the 0, 12, 24 and 48 hours after *V. dahliae* inoculation. This comparative transcriptome analysis yielded high-quality data and offered new insights into the molecular mechanisms underlying cotton's resistance against this destructive pathogen.

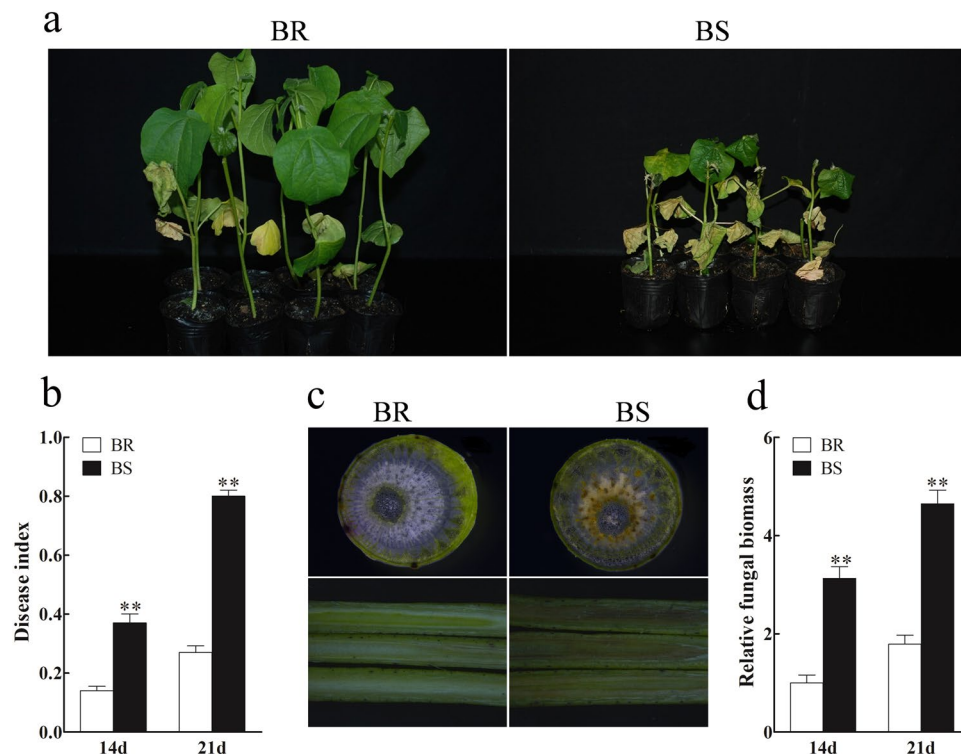
## Background & Summary

Cotton, as one of the prominent fiber crops, serves as vital renewable natural fiber source, which contributes to approximately 35% of global fiber production<sup>1,2</sup>. In 2022, China becomes the second-largest cotton-producing country worldwide, yielding a total production of 5.977 million tons<sup>3</sup>. However, various diseases cause the loss of cotton production, with cotton Verticillium wilt (CVW) emerging as one of the most widespread and great threat to cotton production<sup>4</sup>. Over 40% of the cotton planting area in China being affected by *V. dahliae*, and this disease has become a main threat to the sustainable development of cotton production<sup>5</sup>.

CVW is a soil-borne vascular fungal disease, which is primarily caused by the pathogen of *V. dahliae*<sup>6,7</sup>. Root exudates induces *V. dahliae* germination and *V. dahliae* initiates invasion of the host plant through the root system. The pathogen then penetrates the root epidermal gaps via the root tip or natural wounds, ultimately growing vertically into the root vascular tissue. Subsequently, the spores of *V. dahliae* rapidly spreads to the cotton stems and leaves, resulting in gradual wilting and potential death of the cotton plant<sup>8,9</sup>. The microsclerotia of *V. dahliae* can survive up to 10 years, coupled with its wide range of hosts and the lack of effective fungicides for prevention or eradication, poses significant challenges in controlling the disease<sup>5</sup>. Consequently, the most effective, economical, safe, and environmentally friendly integrated disease management approach for controlling Verticillium wilt is planting and breeding resistant cultivars.

*Gossypium barbadense* (*G. barbadense*) and *Gossypium hirsutum* (*G. hirsutum*) are two major cultivated allopolyploid species, originating from natural hybridization and polyploidization events between diploid A0

<sup>1</sup>Shenzhen Branch, Guangdong Laboratory of Lingnan Modern Agriculture, Genome Analysis Laboratory of the Ministry of Agriculture and Rural Affairs, Agricultural Genomics Institute at Shenzhen, Chinese Academy of Agricultural Sciences, Shenzhen, 518120, China. <sup>2</sup>State Key Laboratory of Cotton Biology, Institute of Cotton Research, Chinese Academy of Agricultural Sciences, Anyang, 455000, China. <sup>3</sup>Key Laboratory of Oasis Eco-Agriculture, College of Agriculture, Shihezi University, Shihezi, 832000, China. <sup>4</sup>Department of Life Sciences, Yuncheng University, Yuncheng, 044000, China. <sup>5</sup>These authors contributed equally: Xianpeng Xiong, Cong Sun. ✉e-mail: [15299959897@163.com](mailto:15299959897@163.com); [xuefei@shzu.edu.cn](mailto:xuefei@shzu.edu.cn)



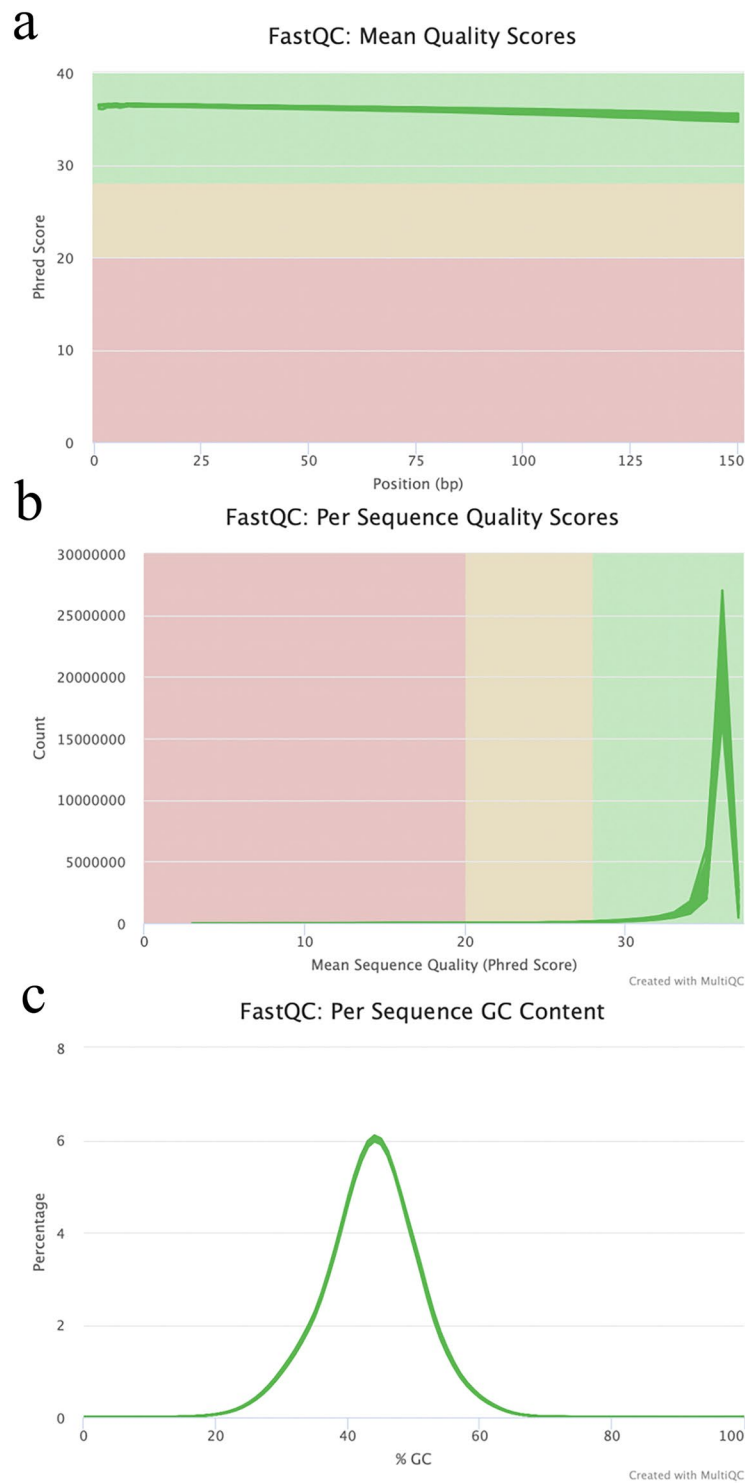
**Fig. 1** Comparison of *Verticillium dahliae* (*V. dahliae*) resistance. **(a)** The phenotype of BR and BS at 14 days post-infection (dpi) with *V. dahliae*, visual observations were made. **(b)** The disease index of both varieties was assessed at 14 and 21 dpi. **(c)** The phenotype of stems from BR and BS plants was examined 14 days after inoculation with *V. dahliae*. **(d)** The relative fungal biomass in the stems of BR and BS at 14 days post-inoculation (dpi) was determined. All data presented are expressed as mean  $\pm$  standard deviation based on three independent biological replicates. Statistical analysis was performed using the T-test ( $P < 0.05$ ,  $*P < 0.01$ ).

and D5 genomes approximately 1–2 million years ago (MYA)<sup>10</sup>. These two cotton species exhibit significant differences in traits, such as stress tolerance, fiber quality, and lint yield<sup>11</sup>. These variations can be attributed to different processes of natural selection and human domestication. *G. hirsutum* is extensively cultivated due to its superior yield performance and better adaptability, contributing to over 90% of the global cotton fiber production<sup>12,13</sup>. *V. dahliae* impacts more than half of the cotton acreage, and *G. hirsutum* is notably susceptible to *V. dahliae*. This susceptibility leads to a substantial reduction in both fiber quality and yield<sup>14</sup>. Notably, *G. barbadense* demonstrates high resistance or immunity to cotton *V. dahliae*, making it a valuable resource for identifying key genes associated with *Verticillium* wilt resistance and uncovering the corresponding regulatory mechanisms<sup>15,16</sup>. However, the lack of *V. dahliae*-susceptible *G. barbadense* germplasm resources has limited the research focus to comparing *V. dahliae*-resistant *G. barbadense* with either *V. dahliae*-resistant *G. hirsutum* or *V. dahliae*-susceptible *G. hirsutum* varieties<sup>17–22</sup>. Currently, there is only one study comparing transcriptomic differences between *V. dahliae*-resistant and -susceptible varieties of *G. barbadense*, but the transcriptome data has not been released<sup>16</sup>. The rare bioinformatic data limited the deep understanding of how immune network of *G. barbadense* play role in defeating the invasion of *V. dahliae*.

In this study, we performed a comparison of the dynamic changes in the transcriptomes of resistant and susceptible *G. barbadense* varieties before and after inoculation with *V. dahliae* using RNA-seq. Our study demonstrates that the disease-resistant variety exhibits higher expression levels of genes involved in the suberin biosynthetic process, phenylpropanoid biosynthetic process, and other biological pathways associated with the establishment of cell physical barriers under normal growth conditions. Furthermore, we observed that the transcriptome response of the resistant variety is less sensitive even after invasion by *Verticillium* wilt, suggesting that its stronger physical barrier may play a significant role in conferring resistance against *V. dahliae*.

## Methods

**Plant material.** Two *Gossypium barbadense* varieties, namely *V. dahliae*-resistant *G. barbadense* ‘Xinhai 47’ (BR) and the *V. dahliae*-susceptible *G. barbadense* ‘Shidahaigan 1’ (BS) were selected for pathogenicity assays and RNA-seq analysis. To facilitate proper growth of the seedlings, the germinated seeds were incubated in controlled conditions in an incubator with a photoperiod of 16 hours of light and 8 hours of darkness; maintaining day and night temperatures of 25°C and 23°C respectively.



**Fig. 2** Evaluation of sequence quality in raw FASTQ data. The quality scores of all 48 raw RNA-seq reads were assessed using FastQC, and the results were compiled and summarized using MultiQC. **(a)** Displays the distribution of read counts for mean sequence quality. **(b)** Represents the distribution of mean quality scores. **(c)** Illustrates the per-sequence GC content.

**Disease resistance identification.** Two-leaf-stage cotton seedlings were infected with the highly aggressive defoliating *V. dahliae* strain “V991” by using the root irrigation method<sup>23</sup>. The strain was kindly provided by Associate Professor Yanfei Sun (Shihezi University). The concentration of “V991” spore liquid was adjusted to 106 spores/mL by using a hemocytometer, followed by the application of 20 mL per plant into the nutrient bowls. Subsequently, the plants were incubated for further cultivation. At 14 days post-infection (dpi), the BR plants exhibited mild symptoms of Verticillium wilt, characterized by slight yellowing of a few lower leaves.

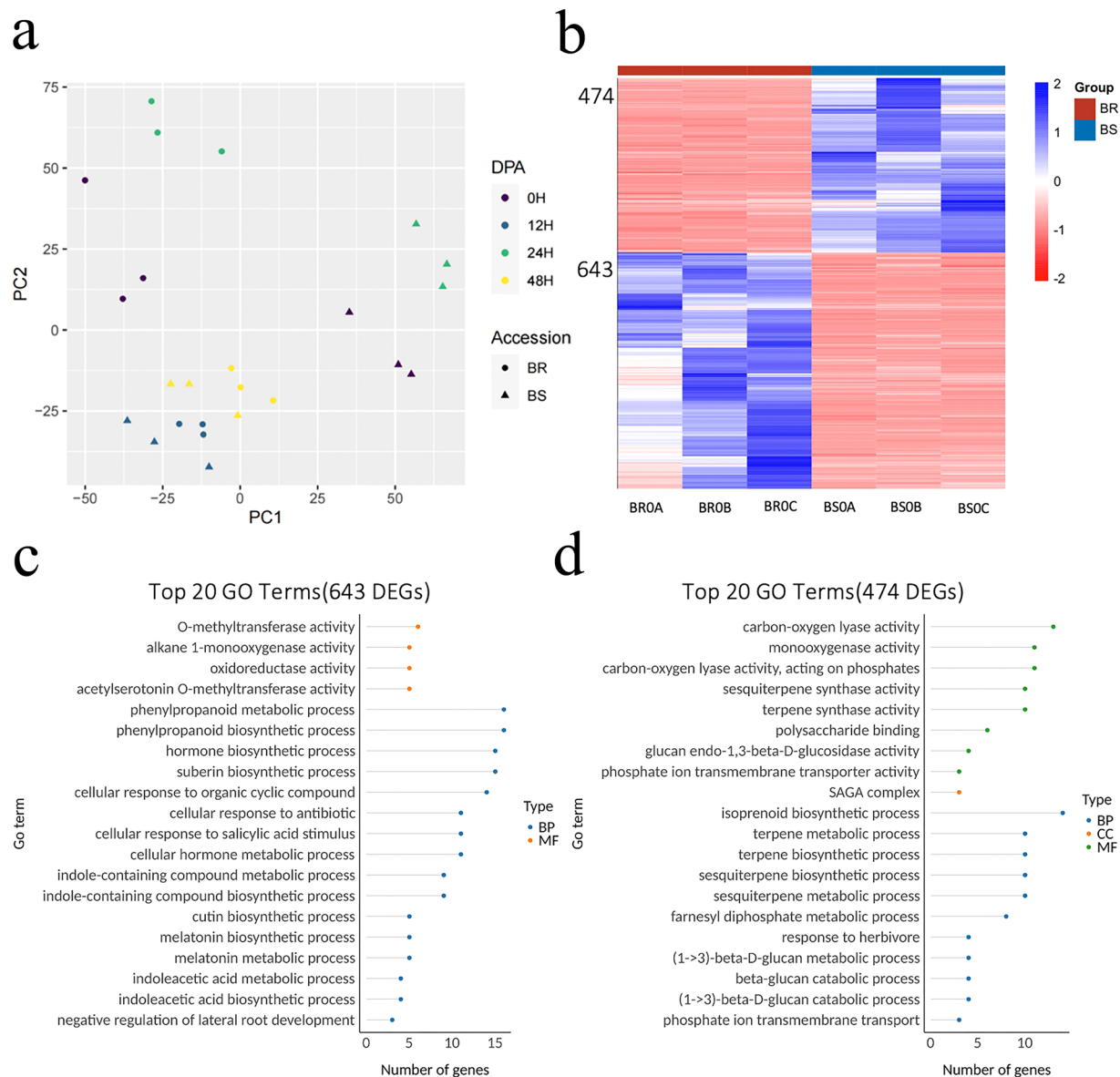
Sample_name	Accession	Clean_base (bp)	Raw reads	Clean reads	Effective rate(%)	Q20 (%)	Q30 (%)	GC content(%)
BR0A	CRR841041	7340324000	49509550	49228234	99.43	98.43	94.95	43.19
BR0B	CRR841042	7904142000	53177188	52872874	99.43	98.30	94.53	43.19
BR0C	CRR841043	10309525000	69464440	68978266	99.30	98.34	94.68	43.12
BR12A	CRR841044	6892841000	46425442	46129546	99.36	98.20	94.44	43.04
BR12B	CRR841045	9201922000	61987652	61601610	99.38	98.39	94.78	43.15
BR12C	CRR841046	8085703000	54372360	54072498	99.45	98.45	94.96	43.16
BR24A	CRR841047	7254628000	48880126	48576874	99.38	98.37	94.77	43.22
BR24B	CRR841048	8807664000	59285298	58956642	99.45	98.39	94.78	43.24
BR24C	CRR841049	7788918000	52521184	52213186	99.41	98.40	94.90	43.15
BR48A	CRR841050	7662713000	51597542	51274828	99.37	98.28	94.48	43.22
BR48B	CRR841051	9007112000	60665338	60296608	99.39	98.36	94.76	43.28
BR48C	CRR841052	7716154000	51952982	51636162	99.39	98.39	94.82	43.18
BS0A	CRR841053	7805688000	52561304	52246472	99.40	98.05	93.99	43.23
BS0B	CRR841054	9606783000	64661964	64266908	99.39	98.35	94.73	43.24
BS0C	CRR841055	9418788000	63448496	63078938	99.42	98.37	94.73	43.26
BS12A	CRR841056	8680751000	58390302	58060342	99.43	98.06	93.98	43.07
BS12B	CRR841057	10597593000	71386004	70992320	99.45	98.10	94.12	43.23
BS12C	CRR841058	10775798000	72558132	72116712	99.39	98.36	94.73	43.31
BS24A	CRR841059	8786492000	59146386	58814490	99.44	98.38	94.85	43.24
BS24B	CRR841060	9342497000	62871924	62502326	99.41	98.35	94.72	43.24
BS24C	CRR841061	10422598000	70139402	69753650	99.45	98.41	94.87	43.31
BS48A	CRR841062	8523364000	57428882	57030050	99.31	98.07	93.96	43.33
BS48B	CRR841063	8485533000	57163326	56799548	99.36	98.30	94.57	43.24
BS48C	CRR841064	9422100000	63381432	63029702	99.45	98.33	94.65	43.17

**Table 1.** RNA-seq data quality summary. All data above were counted for read1 + read2. Raw bases: (Raw reads)\* (sequence length). Clean bases: (Clean reads) \* (sequence length). Effective Rate (%): (Clean reads/ Raw reads) \*100%. Error rate: base error rate. Q20, Q30: (Base count value > 20 or 30)/(Total base count). GC content: (G and C base count)/(Total base count).

In contrast, the BS plants displayed more severe symptoms, including significant yellowing and leaf loss, with some plants eventually succumbing to the infection (Fig. 1a). The investigation of disease index and quantification of fungal biomass was conducted with reference to previous studies<sup>23,24</sup>. The disease index of BS was notably higher than that of BR at both 14 and 21 dpi (Fig. 1b). Assessing the cotton stem sections revealed browning in both transverse and longitudinal stem sections of BS plants after 14 days of *V. dahliae* inoculation, indicating a more severe condition compared to BR plants (Fig. 1c). Additionally, quantification of fungal biomass in the stems illustrated significantly higher accumulation in BS plants compared to BR plants at 14 and 21 dpi (Fig. 1d). Overall, these results convincingly demonstrated the significantly higher resistance of BR to *V. dahliae* than to BS.

**RNA Extraction, Library Preparation and Sequencing.** Root samples from both the BR and BS genotypes were collected at 0, 12, 24 and 48 hours after *V. dahliae* inoculation separately. The samples were immediately frozen in liquid nitrogen. All samples were harvested from five cotton plants with three independent biological replicates. In total, 24 samples (2 genotypes × 4-time points × 3 biological replicates) were subjected to RNA-seq. For total RNA extraction, 0.1 g of the finely ground samples was accurately weighed and using the plant polysaccharide polyphenol total RNA extraction kit from TianGen Biotech Co., Ltd. (Beijing). From the extracted total RNA, 3 µg of qualified RNA was used for further processing. Specifically, mRNA was purified using magnetic beads with Okigo (dT) oligo for enrichment. Subsequently, the mRNA was fragmented, and first-strand cDNA synthesis was performed using a random primer with a six-base sequence. The second-strand cDNA synthesis was carried out by adding buffer, dNTPs, and DNA polymerase I. Subsequently, the double-stranded cDNA was purified using AMPure XP beads. To prepare the cDNA library, the purified double-stranded cDNA underwent end repair, A-tailing, and sequencing adapter ligation. Fragment size selection was performed using AMPure XP beads, and PCR amplification was conducted to obtain the cDNA library. The resulting library was subjected to RNA sequencing on the Illumina HiSeq 4000 platform, generating 150 bp paired-end reads. A total of 48 clean reads were obtained from the 24 cotton samples. The quality assessment of these raw RNA-seq reads (48 in total) was conducted using FastQC18 software, and the summarized results were generated using MultiQC software tool<sup>25</sup>. The raw sequencing data underwent data quality control with the fastp program, which involved the removal of sequences containing adapters, poly-N sequences, and low-quality data<sup>26</sup>.

**Differential expression genes (DEGs) identification and GO enrichment.** The clean reads were aligned to the *Gossypium barbadense* L. (Hai7124) genome<sup>27</sup> using hisat2 with default parameters<sup>28</sup>. FeatureCounts



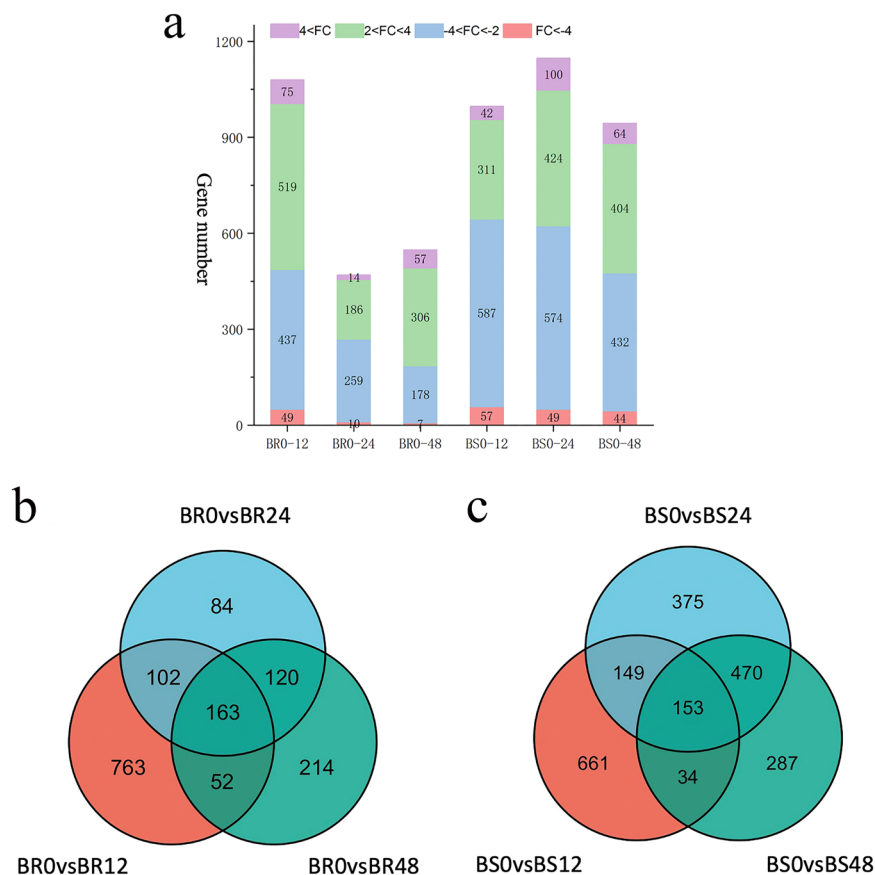
**Fig. 3** Basal transcriptomes of Xinhai 47 (BR), a resistant variety, and shidahaigan 1 (BS), a susceptible variety. **(a)** Principal component analysis (PCA) based on TPM values for 24 samples. **(b)** Heatmap showing the expression patterns of 1117 differentially expressed genes (DEGs) between BR and BS without *Verticillium dahliae* (*V. dahliae*) inoculation. **(c)** GO enrichment of highly expressed DEGs in uninfected BR plants compared to BS plants. **(d)** GO enrichment of highly expressed DEGs in uninfected BS plants compared to BR plants.

(Version 2.0.1) was used to calculate the read count of each gene, and gene expression levels were quantified using TPM<sup>29</sup> (transcripts per kilobase of exon model per million mapped reads). Differentially expressed genes (DEGs) were identified using the DESeq R package (1.18.0)<sup>30</sup> based on the criteria of  $|\log_2(\text{fold-change})| \geq 2$  and an adjusted P value  $< 0.05$ . The gene Ontology (GO) enrichment analysis of DEGs was performed using the clusterProfiler<sup>31</sup> R package, with a significance threshold set at an adjusted p-value  $< 0.05$ .

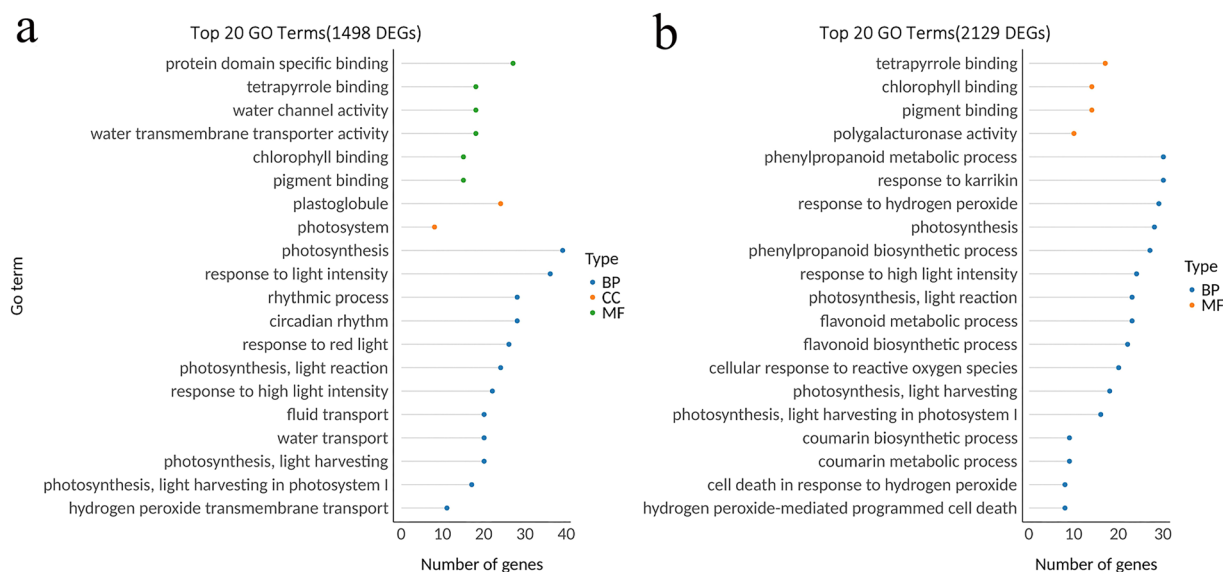
**Validation of DEGs by quantitative real-time PCR (qRT-PCR).** We selected 6 DEGs and designed gene specific primers online using the qPrimerDB website (<https://biodb.swu.edu.cn/qprimerdb/>), with amplified fragments ranging in length from 80 to 250 bp. The primers used in qRT-PCR were listed in Supplementary Table S1.

We performed the qRT-PCR experiment on Roche LightCycle 480II, using SYBR Green Mix. Additionally, *GbUBQ7* (DQ116441.1) was used as the internal reference gene and the relative expression level of the target gene is determined through Livak's  $2^{-\Delta\Delta CT}$  method<sup>32</sup>.

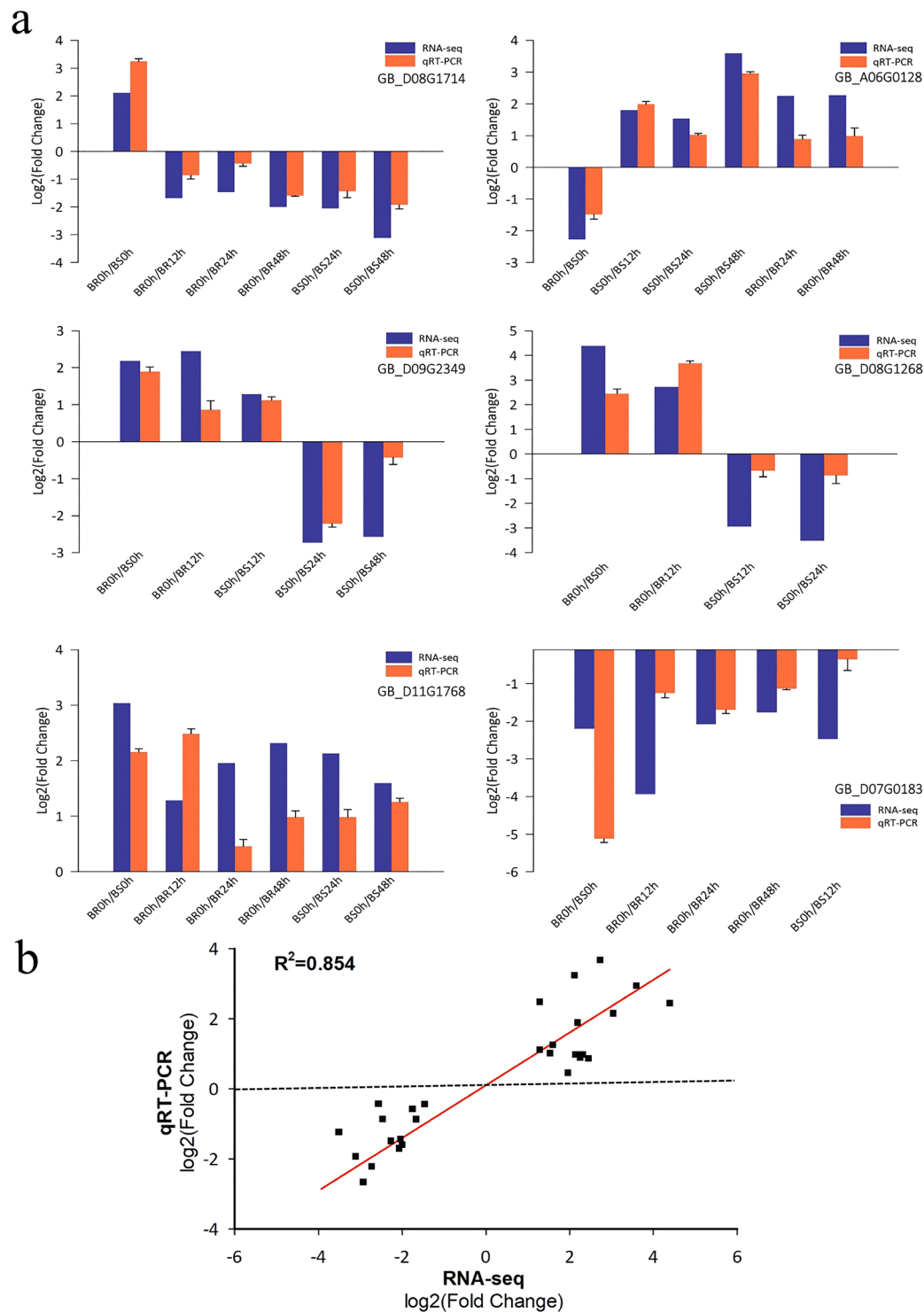




**Fig. 4** Identification of differentially expressed genes (DEGs) in Xinhai 47 (BR) and shidahaigan 1 (BS) at 12, 24, and 48 h after *Verticillium dahliae* infection. **(a)** The number of DEGs in BR and BS after *V. dahliae* infection at 12, 24, 48 hpi. **(b)** Venn diagram of DEGs at each time point in BR after inoculation with *V. dahliae*. **(c)** Venn diagram of DEGs at each time point in BS after inoculation with *V. dahliae*.



**Fig. 5** Gene Ontology (GO) analysis was performed on the 1498 and 2129 differentially expressed genes (DEGs) identified in the BR and BS varieties, respectively, after infection with *Verticillium dahliae* (*V. dahliae*). **(a)** GO analysis of DEGs in the BR variety following inoculation with *V. dahliae*. **(b)** GO analysis of DEGs in the BS variety following inoculation with *V. dahliae*.



**Fig. 6** Validation of RNA-seq results by qRT-PCR analysis. **(a)** Fold change of six DEGs through qRT-PCR and RNA-seq. **(b)** Pearson correlation of the expression fold change of six DEGs between the RNA-seq and RT-qPCR assays.

### Data Records

24 raw RNA-seq data reported in this paper have been deposited in the Genome Sequence Archive (Genomics, Proteomics & Bioinformatics 2021) in National Genomics Data Center (Nucleic Acids Res 2021), China National Center for Bioinformation/Beijing Institute of Genomics, Chinese Academy of Sciences (GSA: CRA012121)<sup>33</sup> that are publicly accessible at <https://ngdc.cncb.ac.cn/gsa/browse/CRA012121>. Other data, such as the breakdown of each sample file on GSA, lists of raw gene counts, TPM of the 24 samples, the information of DEGs and results of GO enrichment analysis are also available at figshare<sup>34</sup>. (<https://doi.org/10.6084/m9.figshare.24097941.v2>).

## Technical Validation

**Quality assessment of sequencing data.** The MultiQC analysis revealed that the quality scores of all sequences exceeded 30 (Fig. 2a). Moreover, the per-sequence quality scores predominantly concentrated within the range of 35–40 (Fig. 2b), indicating a base error rate of less than 0.03%. The GC distribution of all sequences exhibited a normal distribution closely resembling the theoretical distribution (Fig. 2c), providing further evidence of the high quality of the reads. Throughout the data processing, an average of 5.9 million adapter sequences were filtered out from each sample, resulting in an average sample size of clean reads of 8.83 Gb (range 6.96–10.88 Gb). The percentage of clean reads was calculated as 99.4% (Table 1).

**Principal component analysis (PCA) and different basal transcriptome identification.** To evaluate the transcriptional differences between the BR and BS varieties, we performed principal component analysis (PCA) using the TPM values from 24 samples. The results demonstrated a clear distinction between the BR and BS varieties, regardless of *V. dahliae* inoculation (Fig. 3a, Figshare File 2<sup>34</sup>). Remarkably, even in the absence of *V. dahliae* inoculation, the BR and BS samples exhibited distinct separation based on PC1. This observation suggests a potential association between the transcriptional dissimilarities of the two varieties and their distinct levels of disease resistance. A total of 1117 DEGs were identified between the BR and BS varieties at 0 hours post-inoculation (hpi), prior to *V. dahliae* infection. Of these DEGs, 643 exhibited significantly higher expression levels in BR than in BS, whereas 474 displayed significantly lower expression levels in BR (Fig. 3b, Figshare File 3<sup>34</sup>). Among the top 20 GO terms in 643 DEGs, genes associated with physical barrier formation, such as *GbCYP86A1* (GB\_A07G1113), *GbCER1* (GB\_A10G0724) in suberin biosynthetic process were identified. Furthermore, a GO term related to Verticillium wilt resistance, which included cellular response to antibiotic pathways and melatonin biosynthetic process, was also significantly enriched in 643 DEGs (Fig. 3c, Figshare File 4<sup>34</sup>). The most enriched processes among the 474 DEGs with higher expression in BS were related to carbon-oxygen lyase activity and isoprenoid biosynthesis (Fig. 3d, Figshare File 5<sup>34</sup>).

**Analysis of differentially expressed genes (DEGs) involved in the response to *V. dahliae* inoculation.** To investigate the molecular mechanisms underlying the resistance of different varieties of *G. barbadense* to *V. dahliae* infection, we compared the gene expression changes at 12, 24, and 48 hpi with an uninfected sample (0 hpi). A total of 1080, 469, 548 DEGs were identified at 12, 24, and 48 hpi in the BR variety respectively. Similarly, in the BS variety, we found 997, 1147, and 944 DEGs at the same time points (Fig. 4a, Figshare File 6<sup>34</sup>). Interestingly, BR demonstrated a high number of induced genes at all three time points, whereas BS displayed the opposite pattern. Analysis revealed 1498 nonredundant DEGs in BR and 2129 in BS, in response to *V. dahliae* at all three time points post-inoculation (Fig. 4b–c, Figshare File 7<sup>34</sup>).

Among the top 20 GO terms, 8 terms were shared between BR and BS. The most significantly enriched terms included photosynthesis, light reaction, and response to high light intensity (Fig. 5a, Figshare File 8<sup>34</sup>). Additionally, BS exhibited specific enrichment in phenylpropanoid metabolic process, response to hydrogen peroxide, and flavonoid metabolic process. These GO terms have previously been reported to play a role in cotton resistance to *V. dahliae* (Fig. 5b, Figshare File 9<sup>34</sup>). The findings suggest that BR displayed insensitivity to the pathogen, whereas BS exhibited a rapid activation of resistance-related transcriptome responses upon inoculation with *V. dahliae*.

**Validation results of RNA-seq data by qRT-PCR.** Although the log<sub>2</sub>(fold-change) value from qRT-PCR did not precisely match that of RNA-seq, there was a strong positive correlation ( $R^2 = 0.854$ ) observed between the results of RNA-seq and qRT-PCR expression analyses, indicating the reliability of the RNA-seq results (Fig. 6).

In summary, we have presented a collection of high-quality transcriptome data illustrating the response of *G. barbadense* to *V. dahliae*. Our preliminary analysis indicates a potential association between the resistance of *G. barbadense* to *V. dahliae* and the presence of physical barriers. We believe that this transcriptome data offers valuable insights into the molecular mechanisms that underlie cotton's resistance to *V. dahliae*, providing a novel approach for identifying key genes involved in such resistance.

## Code availability

The following software used in the study was run with default parameters if not specifically stated in the Methods paragraph and no custom code was used for the purposes of this study.

FastQC: <https://www.bioinformatics.babraham.ac.uk/projects/fastqc> (version 0.11.9).

fastp: <https://github.com/OpenGene/fastp> (version 0.20.0).

hisat2: <https://daehwankimlab.github.io/hisat2/download> (version 2.2.1).

FeatureCounts: <https://sourceforge.net/projects/subread> (version 2.0.4).

Differential analysis: R package DESeq2 1.38.3.

clusterProfile: R package clusterProfiler v4.6.2.

Venn: <https://hiplot.cn/basic/venn>.

Received: 11 September 2023; Accepted: 12 December 2023;

Published online: 02 January 2024

## References

- Xiao, G., Zhao, P. & Zhang, Y. A pivotal role of hormones in regulating cotton fiber development. *Front Plant Sci* **10**, 87, <https://doi.org/10.3389/fpls.2019.00087> (2019).
- Zhu, Y. *et al.* Interactions between *Verticillium dahliae* and cotton: pathogenic mechanism and cotton resistance mechanism to Verticillium wilt. *Front Plant Sci* **14**, 1174281, <https://doi.org/10.3389/fpls.2023.1174281> (2023).



3. Chen, X. *et al.* Long-term increases in continuous cotton yield and soil fertility following the application of cotton straw and organic manure. *Agronomy* **13**, 2133, <https://doi.org/10.3390/agronomy13082133> (2023).
4. Shaban, M. *et al.* Physiological and molecular mechanism of defense in cotton against *Verticillium dahliae*. *Plant Physiol Biochem* **125**, 193–204, <https://doi.org/10.1016/j.plaphy.2018.02.011> (2018).
5. Zhang, X. *et al.* Transcriptomic analysis of gene expression of *Verticillium dahliae* upon treatment of the cotton root exudates. *BMC Genomics* **21**, 155, <https://doi.org/10.1186/s12864-020-6448-9> (2020).
6. Fradin, E. F. & Thomma, B. P. Physiology and molecular aspects of *Verticillium* wilt diseases caused by *V. dahliae* and *V. albo-atrum*. *Mol Plant Pathol* **7**, 71–86, <https://doi.org/10.1111/j.1364-3703.2006.00323.x> (2006).
7. Klimes, A., Dobinson, K. F., Thomma, B. P. & Klosterman, S. J. Genomics spurs rapid advances in our understanding of the biology of vascular wilt pathogens in the genus *Verticillium*. *Annu Rev Phytopathol* **53**, 181–198, <https://doi.org/10.1146/annurev-phyto-080614-120224> (2015).
8. Wang, J. *et al.* The *Verticillium dahliae* small cysteine-rich protein VdSCP23 manipulates host immunity. *Int J Mol Sci* **24**, <https://doi.org/10.3390/ijms24119403> (2023).
9. Zhao, F. & Dai, X. Infection process of *Verticillium dahliae* klebahn in cotton. *Genomics and Applied Biology* **28**, 786–792 (2009).
10. Wen, X. *et al.* A comprehensive overview of cotton genomics, biotechnology and molecular biological studies. *Sci China Life Sci* <https://doi.org/10.1007/s11427-022-2278-0> (2023).
11. Viot, C. R. & Wendel, J. F. Evolution of the cotton genus, *Gossypium*, and its domestication in the Americas. *Crit. Rev. Plant Sci.* **42**, 1–33, <https://doi.org/10.1080/07352689.2022.2156061> (2023).
12. Lacape, J. M. *et al.* A new interspecific, *Gossypium hirsutum* x *G. barbadense*, RIL population: towards a unified consensus linkage map of tetraploid cotton. *Theor Appl Genet* **119**, 281–292, <https://doi.org/10.1007/s00122-009-1037-y> (2009).
13. Park, Y. H. *et al.* Genetic mapping of new cotton fiber loci using EST-derived microsatellites in an interspecific recombinant inbred line cotton population. *Mol Genet Genomics* **274**, 428–441, <https://doi.org/10.1007/s00438-005-0037-0> (2005).
14. Li, R. *et al.* Genome resource for the *Verticillium* wilt resistant *Gossypium hirsutum* cultivar Zhongzhimian No. 2. *Mol Plant Microbe Interact* **36**, 68–72, <https://doi.org/10.1094/mpmi-10-22-0205-a> (2023).
15. He, L. *et al.* Advanced genes expression pattern greatly contributes to divergence in *Verticillium* wilt resistance between *Gossypium barbadense* and *Gossypium hirsutum*. *Front Plant Sci* **13**, 979585, <https://doi.org/10.3389/fpls.2022.979585> (2022).
16. Wang, F. *et al.* Multicopper oxidases GbAO and GbSKS are involved in the *Verticillium dahliae* resistance in *Gossypium barbadense*. *J Plant Physiol* **280**, 153887, <https://doi.org/10.1016/j.jplph.2022.153887> (2023).
17. Gao, W. *et al.* Proteomic and virus-induced gene silencing (VIGS) Analyses reveal that gossypol, brassinosteroids, and jasmonic acid contribute to the resistance of cotton to *Verticillium dahliae*. *Mol Cell Proteomics* **12**, 3690–3703, <https://doi.org/10.1074/mcp.M113.031013> (2013).
18. Sun, Q. *et al.* Analysis of sea-island cotton and upland cotton in response to *Verticillium dahliae* infection by RNA sequencing. *BMC Genomics* **14**, 852, <https://doi.org/10.1186/1471-2164-14-852> (2013).
19. Xiong, X. P. *et al.* The cotton lignin biosynthetic gene *Gh4CL30* regulates lignification and phenolic content and contributes to *Verticillium* wilt resistance. *Mol Plant Microbe Interact* **34**, 240–254, <https://doi.org/10.1094/mpmi-03-20-0071-r> (2021).
20. Xu, L. *et al.* 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. *J Exp Bot* **62**, 5607–5621, <https://doi.org/10.1093/jxb/err245> (2011).
21. Zhang, W. *et al.* Large-scale identification of *Gossypium hirsutum* genes associated with *Verticillium dahliae* by comparative transcriptomic and reverse genetics analysis. *PLoS One* **12**, e0181609, <https://doi.org/10.1371/journal.pone.0181609> (2017).
22. Zhang, Y. *et al.* Transcriptome profiling of *Gossypium barbadense* inoculated with *Verticillium dahliae* provides a resource for cotton improvement. *BMC Genomics* **14**, 637, <https://doi.org/10.1186/1471-2164-14-637> (2013).
23. Xiong, X. *et al.* The cotton WRKY transcription factor *GhWRKY70* negatively regulates the defense response against *Verticillium dahliae*. *Crop Journal* **7**, 393–402, <https://doi.org/10.1016/j.cj.2018.10.005> (2019).
24. Atallah, Z. K., Bae, J., Jansky, S. H., Rouse, D. I. & Stevenson, W. R. Multiplex real-time quantitative PCR to detect and quantify *Verticillium dahliae* colonization in potato lines that differ in response to *Verticillium* wilt. *Phytopathology* **97**, 865–872, <https://doi.org/10.1094/phyto-97-7-0865> (2007).
25. Ewels, P., Magnusson, M., Lundin, S. & Kaller, M. MultiQC: summarize analysis results for multiple tools and samples in a single report. *Bioinformatics* **32**, 3047–3048, <https://doi.org/10.1093/bioinformatics/btw354> (2016).
26. Chen, S. F., Zhou, Y. Q., Chen, Y. R. & Gu, J. Fastp: an ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics* **34**, 884–890, <https://doi.org/10.1093/bioinformatics/bty560> (2018).
27. Hu, Y. *et al.* *Gossypium barbadense* and *Gossypium hirsutum* genomes provide insights into the origin and evolution of allotetraploid cotton. *Nat Genet* **51**, 739–748, <https://doi.org/10.1038/s41588-019-0371-5> (2019).
28. Pertea, M., Kim, D., Pertea, G. M., Leek, J. T. & Salzberg, S. L. Transcript-level expression analysis of RNA-seq experiments with HISAT, StringTie and Ballgown. *Nat Protoc* **11**, 1650–1667, <https://doi.org/10.1038/nprot.2016.095> (2016).
29. Zhao, Y. *et al.* TPM, FPKM, or normalized counts? A comparative study of quantification measures for the analysis of RNA-seq data from the NCI patient-derived models repository. *J Transl Med* **19**, 269, <https://doi.org/10.1186/s12967-021-02936-w> (2021).
30. Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* **15**, 550, <https://doi.org/10.1186/s13059-014-0550-8> (2014).
31. Wu, T. *et al.* ClusterProfiler 4.0: A universal enrichment tool for interpreting omics data. *Innovation (Camb)* **2**, 100141, <https://doi.org/10.1016/j.xinn.2021.100141> (2021).
32. Livak, K. J. & Schmittgen, T. D. Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) method. *Methods* **25**(4), 402–408, <https://doi.org/10.1006/meth.2001.1262> (2001).
33. *Genome Sequence Archive* <https://ngdc.cncb.ac.cn/gsa/browse/CRA012121> (2023).
34. Xiong, X. & Sun, C. Transcriptome data of samples from two *Gossypium barbadense* varieties before and after *Verticillium dahliae* inoculation. *Figshare*. <https://doi.org/10.6084/m9.figshare.24097941.v2> (2023).

## Acknowledgements

This project was supported by the China Postdoctoral Science Foundation (Grant No. 2021M703552), the Natural Science Foundation of Hubei Provincial Department of Science and Technology (2023AFB931), and the Natural Science Foundation of Hubei Provincial Department of Education (Q20224301).

## Author contributions

F.X. and C.F. designed the study, revised the manuscript. J.S. revised the manuscript. X.X. and C.S. performed the experiments and wrote the original draft. B.C. performed the sample collection, RNA isolation.

## Competing interests

The authors declare no competing interest.

### Additional information

**Supplementary information** The online version contains supplementary material available at <https://doi.org/10.1038/s41597-023-02852-2>.

**Correspondence** and requests for materials should be addressed to C.F. or F.X.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2024