Journal of Cotton Research

RESEARCH

Open Access

QTL mapping associated with Verticillium wilt resistance in cotton based on MAGIC population

AYYAZ Muhammad[1†], CHANG Zewei[1†], DING Shugen[1], HAN Peng[1], XU Lin[2,3], ABUDUKEYOUMU Abudurezike[2,3], SIDDHO Irfan Ali[1], LI Zhibo[1], LIN Hairong[1], XU Jianwei[1*], WU Yuanlong[1*] and NIE Xinhui[1*]

Abstract

Background Cotton is an important cash crop in China and a key component of the global textile market. Verticillium wilt is a major factor affecting cotton yield. Single nucleotide polymorphism (SNP) markers and phenotypic data can be used to identify genetic markers and loci associated with cotton resistance to Verticillium wilt. We used eight upland cotton parent materials in this study to construct a multiparent advanced generation inter-cross (MAGIC) population comprising 320 lines. The Verticillium wilt resistance of the MAGIC population was identified in the greenhouse in 2019, and the average relative disease index (ARDI) was calculated. A genome-wide association study (GWAS) was performed to discover SNP markers/genes associated with Verticillium wilt resistance.

Results ARDI of the MAGIC population showed wide variation, ranging from 16.7 to 79.4 across three replicates. This variation reflected a diverse range of resistance to Verticillium wilt within the population. Analysis of distribution patterns across the environments revealed consistent trends, with coefficients of variation between 12.25% and 21.96%. Families with higher ARDI values, indicating stronger resistance, were more common, likely due to genetic diversity and environmental factors. Population structure analysis divided the MAGIC population into three subgroups, with Group I showing higher genetic variation and Groups II and III displaying more uniform resistance performance. Principal component analysis (PCA) confirmed these divisions, highlighting the genetic diversity underlying Verticillium wilt resistance. Through GWAS, we identified 19 SNPs significantly associated with Verticillium wilt resistance, distributed across three chromosomes. The screening of candidate genes was performed on the transcriptome derived from resistant and susceptible cultivars, combined with gene annotation and tissue expression patterns, and two key candidate genes, *Ghir_A01G006660* and *Ghir_A02G008980,* were found to be potentially associated with Verticillium wilt resistance. This suggests that these two candidate genes may play an important role in responding to Verticillium wilt.

Conclusion This study aims to dissect the genetic basis of Verticillium wilt resistance in cotton by using a MAGIC population and GWAS. The study seeks to provide valuable genetic resources for marker-assisted breeding and enhance the understanding of resistance mechanisms to improve cotton resilience against Verticillium wilt.

Keywords Upland cotton, Verticillium wilt, MAGIC population, Quantitative trait loci, Association analysis

†Muhammad Ayyaz and Zewei Chang contributed equally to this work.

*Correspondence:
Xu Jianwei
xjnxh2004130@126.com
Wu Yuanlong
wyl0322@shzu.edu.cn
Nie Xinhui
438120660@qq.com
Full list of author information is available at the end of the article

Background

Cotton (*Gossypium* spp.) is an economically important crop cultivated worldwide for fiber and oil production (Campbell et al. 2010). Global demand for cotton continues to increase yearly, even though global production has recently declined. High-end textile products made from natural fibers are also growing in popularity, making it essential to simultaneously improve cotton yield and fiber quality. The genus *Gossypium* L. (cotton), the largest and most widely distributed in the *Gossypieae* tribe, includes 52 species: 46 diploid species (2n = 2x = 26) and seven tetraploid species (five recognized and two recently discovered, 2n = 4x = 52) (Cai et al. 2023). Research has demonstrated a significant increase in the pathogenicity of *Verticillium dahliae* in the cotton-growing regions of Xinjiang, China (Wagner et al. 2021). The physiological categorization of host pathogenicity is the main focus of research in China on the differentiation of *Verticillium dahliae* pathogenicity in cotton. Verifying the biological and molecular functions of genes is the best strategy for improving target traits in cotton Verticillium wilt (VM) (Li et al. 2023). Verticillium wilt is an important disease of cotton with the causative agent being the soilborne hemi biotrophic fungus *Verticillium dahliae* Kleb (Wilson et al. 2024).

During infection, the *Verticillium dahliae* secretome supplies a range of molecules, such as toxins to manipulate the host responses and aid its growth. This process can result in vascular occlusion, which prevents the transfer of water and other mineral substances from roots to the leaves and tissues. Consequently, it causes wilting, drying a reduction in photosynthesis, shedding of immature bolls, and importantly a significant reduction in fiber yield (Wilson et al. 2024). Cotton-defoliating Verticillium wilt was first reported in Jiangsu Province, China, in 1983 (Yang et al. 2022). Deciduous and non-deciduous cotton Verticillium wilt bacterial strains have been identified in China and the pathogenicity of the former is greater than that of the latter (Yang et al. 2023). A TIR-NBS-LRR gene was identified to be the likely candidate gene conferring the VW resistance for the QTL on A10. Some cotton genes associated with responses to VW have been reported based on gene expression studies (Zhang et al. 2020), but no direct genetic relationship with any VW resistance gene or QTL has been established through molecular linkage mapping (Zhang et al. 2020).

Previous studies have identified QTLs linked to Verticillium wilt resistance dispersed across 26 pairs of cotton chromosomes using various mapping populations and genetic markers; however, these have not yet been validated and implemented in breeding procedures. The primary causes are a low rate of disease resistance QTL interpretation, great genetic separation between the QTL and target gene, and insufficient placement accuracy (Palanga et al. 2017). In recent years, numerous QTL mapping studies have been conducted on cotton Verticillium wilt resistance (Zhang et al. 2013), which has partially shown that cotton Verticillium wilt resistance is a quantitative trait. Only a few major genes have been cloned and used in breeding, although many QTLs associated with VW resistance have been identified through linkage mapping or genome-wide association study (GWAS). This is because the desirable trait is complex and influenced by a variety of factors, such as environmental conditions, phenotypes, genotypes, and population structure (Zhang et al. 2014).

GWAS is a technique that screens for markers linked to target trait variations by utilizing high-density markers (such as SNP marker) across many populations, based on the laws of linkage disequilibrium. Understanding the relationship between the genotype and phenotype is a goal (Khaskheli et al. 2013). Many trait variation loci in Arabidopsis, cotton, maize, wheat, and other crops have been identified using GWAS and it is crucial to examine their regulatory mechanisms (Pincot et al. 2020). GWAS was used to detect SNP associated with VW resistance in cotton and identified 17 SNPs that had a significant association with *Verticillium dahliae* response. To further confirm these loci for VW resistance, a comparison of the GWAS was performed with QTLs identified in previous studies (Li et al. 2017).

Many complex populations have been used in mapping analysis to detect more reliable QTLs, such as multi-parent advanced generation intercross (MAGIC) populations, which are made genetically rich by combining alleles of all parent lines (Cavanagh et al. 2008). MAGIC population consists of various parents and a stable population comprising multiple families is generated after several generations of self-crossing and purification. This population incorporates the best traits from several parents. In comparison to natural populations, a MAGIC population breaks linkage disequilibrium, has a higher number of recombinations and richer mutations, and is more resilient to false-positive results (Cavanagh et al. 2008). Currently, there are limited studies of MAGIC populations of cotton but research on MAGIC populations has been applied to other crops like maize (Giraud et al. 2017), rice (Bandillo et al. 2013), tomato (Pascual et al. 2015), sorghum (Ongom et al. 2018), barley (Novakazi et al. 2020), wheat (Huang et al. 2012), kidney bean (Diaz et al. 2020). Based on MAGIC populations, few studies have identified cotton resistance loci against Verticillium wilt. Verticillium wilt resistance was observed in

indoor MAGIC populations and five cotton lines with better disease resistance levels were discovered (Zhang et al. 2021).

In this study, eight cotton parent materials with outstanding qualities and varying degrees of resistance to Verticillium wilt were selected after a thorough assessment. The foundational population of cotton disease-resistant MAGIC was created by aggregating crossbreeding resources consisting of 320 families with consistent agronomic features. The genotypes used in this study were obtained from theMAGIC population, specifically developed by intercrossing eight parental cotton varieties selected from the cotton-growing regions of Northern Xinjiang, China. This served to clarify the genetic makeup of cotton that is resistant to Verticillium wilt. This study aims to investigate the genetic basis of Verticillium wilt resistance in cotton by using a MAGIC population and GWAS to identify significant SNP markers and candidate genes. The study seeks to provide valuable genetic resources for marker-assisted breeding and enhance the understanding of resistance mechanisms to improve cotton resilience against Verticillium wilt.

Materials and methods

Plant materials

This study used a MAGIC population consisting of 320 families to map the QTLs for Verticillium wilt resistance in cotton. The MAGIC population used in this study was developed using eight parental lines, as previously described by Tian et al. (2023). These parental lines were selected to maximize genetic diversity and represent a broad spectrum of traits. Details of the parent lines, including their genotypic and phenotypic characteristics, can be found in the article by Tian et al. (2023). Xinluzao 36 and Zhongzhimian 2 were used as disease-resistant controls. Two replications of the experiment (the first replication designed as R1; the second replication designed as R2) were carried out in the greenhouse at Shihezi University in 2019, and the average of the R1 and R2 were labeled R3.

Quantitative root dip inoculation method

A mixture of nutrient soil and vermiculite was prepared at a precise volume ratio of 3:2 (v/v), which is known to be efficient in stimulating plant growth and root development, to provide an ideal growth medium. The seeds were covered with a 1.5–2 cm of vermiculite nutrient soil mixture. To prepare the Czapek medium, a specific mixture of substances was used to strengthen the culture media: 60 g of sucrose, 0.04 g of ferrous sulfate, 6 g of sodium nitrate, 2 g of magnesium sulfate, 2 g of potassium chloride, and 2 g of sodium dihydrogen phosphate (Fonseca et al. 2022). The deciduous strain V991 was the focus of this experiment. A hemocytometer was used to carefully count spores. When the cotton seedlings reached the "two true leaves and one heart" stage, approximately ten days after seeding, they were inoculated with *Verticillium dahliae* (V991). In detail, 10 mL of spore suspension was typically applied to the roots. As the pathogen typically enters the plant through the roots and spreads through the vascular system, this technique replicates the pathogen's natural infection pathway. The conidia concentration was 1×10^6 mL^{-1} (Ibarra-Cortés et al. 2018). After the V991 was completely absorbed the roots, the pots were placed in a greenhouse at a temperature of 25 °C to 28 °C and a relative humidity of 70%.

Resistance identification methods

The cotton seedlings began to show symptoms approximately 7 d after inoculation, with more noticeable symptoms usually appearing approximately 15 d later. The survey employed a 5-level classification scheme as shown in Table 1 (Karademir et al. 2010).

Result calculation

The average relative disease index (ARDI) provides a strong metric that is in good agreement with resistance traits found in QTL analysis by integrating disease severity and resistance throughout time points. The resistance of each variety was assessed based on the ARDI ratings as shown in Table 2. ARDI is calculated as follows,

$$\mathrm{ARDI} = \frac{\sum_i^n \mathrm{RDI}_i}{n}$$

where RDI_i represents the relative disease index (RDI) for the ith time, and n represents the number of times to calculate the *RDI* throughout the reproductive period.

Phenotypic data statistics and analysis

To understand the effects of Verticillium wilt on cotton seedlings, it is essential to carefully monitor and analyze disease-related factors such as symptoms, severity, and progression. The advancement of Verticillium wilt in cotton was quantified using ARDI, which reflects both disease severity and its progression over time. Data related to disease symptoms were systematically collected, and ARDI values were computed using Microsoft Excel 2019 to ensure accurate tracking of the disease's impact. IBM SPSS Statistics 26 was employed to perform variance analysis on the ARDI values to assess genetic variability among the eight cotton parent lines. Additionally, the correlation of disease indices across various environments was analyzed using R Studio with the ggplot2 package and base functions.

Table 1 The identification standard of Verticillium wilt resistance of cotton

Incidence level	Symptom description
Level 0	Healthy, disease-free leaves
Level 1	Part of 1–2 cotyledons or the entire leaf turns yellow and soft, but the true leaves show no symptoms
Level 2	Cotyledons and one true leaf showed symptoms
Level 3	Two true leaves showed symptoms
Level 4	All leaves show symptoms. In severe cases, the leaves fall off or the top center dies

Table 2 Standard for classification of resistance

Level	Resistance type	Abbreviations	ARDI
1	Immunity	I	0.0
2	High resistance	HR	0.1–10.0
3	Disease resistance	R	10.1–20.0
4	Disease-resistant	T	20.1–35.0
5	Susceptible to disease	S	35.1–80.0
6	High susceptible	HS	80.1–100.0

Genotyping of MAGIC population

The 328 MAGIC lines were used for DNA sampling. A sequencing library was constructed from high purity DNA from eight parents and 320 family members. After the library quality was qualified, the Illumina HiSeq2500 sequencer was used for high-throughput sequencing (Zhao et al. 2024). The sequence data were aligned to the reference genome using the Burrows-Wheeler Aligner (BWA) software. SNP calling was performed using the Genome Analysis Toolkit (GATK) software. Finally, resequencing data of eight parents and 319 families were obtained. The sequencing depth of the eight parents was about 50×, and the average sequencing depth of the 319 families was 7×. A total of about 6 TB of sequencing data was obtained. To obtain high quality SNP, SNP with a high missing rate > 0.1, locations with highly rare mutations, a minimum allele frequency > 0.05, and a heterozygosity rate < 0.5 were all eliminated using Plink software (Micheletti et al. 2015).

Population structure analysis

Population structure and principal component analysis (PCA) were performed on the MAGIC population using high-quality SNP markers. Using Admixture software conducted population structure analysis, set the K value to 1–10, and repeated the run 5 times for each K value. Each run had 10 000 iterations (Pritchard et al. 2000). Evolutionary trees were constructed using Treebest software. PCA was performed using the GCTA software (Li et al. 2010). The genetic population structure was also estimated using the Structure software, and subpopulations that help explain differences in Verticillium wilt resistance within the cotton MAGIC population were discovered. Genetic contributions to resistance traits were quantified by calculating genomic heritability using GCTA, and the results of both population structure and heritability analyses were visually presented by the ggplot2 package in R software.

GWAS in the MAGIC population

Using mixed linear model (MLM) in Tassel software, combined the phenotypic data from the three environments (R1, R2, and R3) with the filtered genotyping data, GWAS was conducted. PCA and the K matrix were used to correct false positive results caused by population structure (PCA + K, K = 3), and the SNPs with $-\log_{10}P$ values above 4 were selected as significant trait-associated SNP (Tian et al. 2023). The Manhattan and Q-Q plots were produced using ggplot2 package in R software.

Identification of candidate genes

Based on the reference genome of the upland cotton [*G. hirsutum* (AD1) 'TM-1' genome HAU v1.1] (https://www.cottongen.org/find/genes), this study first identified regions 500 kb upstream and 500 kb downstream of the SNPs with the lowest *P*-values as candidate QTL regions (Li et al. 2017; Zheng et al. 2017; Su et al. 2020). The regions were used for candidate gene identification according to their intragenic SNP variation and annotated with the *Arabidopsis thaliana* database (https://www.arabidopsis.org/) (Tian et al. 2023).

Expression analysis of candidate genes

The expression patterns of the candidate genes associated with resistance to Verticillium wilt were identified in TM-1, selected disease-resistant MAGIC line (M138), and disease-susceptible parent line (P2) under each experimental treatment (Zhang et al. 2023). For the *Verticillium dahliae* disease assays, TM-1 seedlings were cultivated for up to two weeks in Hoagland media. After that, the roots were coiled upon the full expansion of the second leaf, and the seedlings were infected by defoliating

them into a suspension of V991 spores (1×10^6 mL^{-1}) for 2 min (Tian et al. 2023). In the same way, control seedlings were treated with a sterile water dip.

Results

Phenotypic analysis of Verticillium wilt resistance traits in MAGIC population

The ARDI exhibited significant phenotypic variation in the resistance of the MAGIC population to Verticillium wilt, ranging from 16.7 to 79.4 across three replicates (R1: 16.7–79.4, R2: 32.1–76.2, R3: 28.9–69.9), indicating both highly susceptible and highly resistant lines as indicated in Table 3. The ARDI showed a consistent distribution across replicates, with a coefficient of variation (CV) ranging from 14.14 to 21.96 and a standard deviation (SD) ranging from 7.74 to 10.15. According to the distribution analysis, the ARDI in R1 data exhibited a positive skewness of 0.41, indicating that it was skewed to the right with a predominance of lower ARDI values. However, the ARDI in R2 had a left-skewed distribution with a skewness of – 0.14. This variance points to a wide, roughly typical distribution of resistance features in the population. By lowering variability from the two independent biological replicates, the ARDI in R3 along with the average R1 and R2 offers a more reliable metric and can be utilized in genetic mapping studies.

Significant phenotypic variation in ARDI values was observed (Table 3), highlighting a spectrum of resistance and susceptibility to Verticillium wilt. Families with higher ARDI values, indicative of greater resistance, were more prevalent, likely due to genetic diversity and environmental adaptations. Families with lower ARDI values may reflect environmental influences or experimental variability. The nearly normal distribution in R3 (kurtosis: 0.57) suggests integrated responses across environments, with substantial correlation to individual replicates, making it a reliable predictor of ARDI trends. Statistical metrics like skewness and kurtosis capture the interplay of genetic and environmental factors shaping Verticillium wilt resistance (Fig. 1).

In the investigation of Verticillium wilt resistance in cotton, the correlation analysis indicates a substantial correlation between ARDI values in the three environments (R1, R2, and R3), as shown in Fig. 2. Specifically, ARDI in R1 and R3 ($r=0.77$) and R2 and R3 ($r=0.55$) showed a significantly positive correlation. The ARDI between R1 and R2, on the other hand, revealed a slight, non-significant connection ($r=-0.081$).

Initial screening of high-quality SNPs

Through SNP screening and quality control, a total of 1 782 719 high-quality SNP markers were obtained, which were distributed on 26 chromosomes of upland cotton (Table 4). Chromosome D04 had the highest SNP density (3 326 SNP per Mb) and lowest markers (16 966). While A06 had the most markers (179 504) and low density (691SNP per Mb). The Chromosome D09 had the lowest density of SNP.

Population structure and PCA

The MAGIC population was divided into three subgroups (K = 3) by the population structure analysis, as shown in Fig. 3A. Group I (green) exhibits a high degree of genetic variation, whereas Groups II (orange) and III (yellow) exhibit greater homogeneity and consistent resistance features to Verticillium wilt (Fig. 3A). Genetic diversity and the distribution of resistance to Verticillium wilt varied among these subgroups. According to the result of PCA, the cotton MAGIC population was also split into three subgroups. The results confirmed that the cotton population was divided into three groups, further validating the results of the population structure.

Phylogenetic tree of the MAGIC population

The phylogenetic tree was constructed using TreeBest software to visualize the evolutionary relationships within the cotton MAGIC population. Each branch of the phylogenetic tree represented a distinct population group, providing a clear visual representation of the genetic links within the population as shown in Fig. 4. The grouping criteria were guided by the covariate K = 3, previously identified through population structure analysis. This covariate played a crucial role in delineating the genetic clusters observed in the phylogenetic tree. By incorporating the K = 3 grouping, the tree not only reflected the population's genetic diversity but also highlighted its underlying structure. The phylogenetic tree includes 328 MAGIC population lines, which are grouped into three subgroups. The number of lines in

Table 3 Descriptive statistics of ARDI across three replicates

Replicate	Mean ± SD	Maximum	Minimum	CV/%	Skewness	Kurtosis
R1	46.24 ± 10.15	79.4	16.7	21.96	0.41	0.62
R2	54.71 ± 7.74	76.2	32.1	14.14	−0.14	−0.03
R3	50.55 ± 6.19	69.9	28.9	12.25	0.11	0.57

SD standard deviation, *CV* coefficient of variation

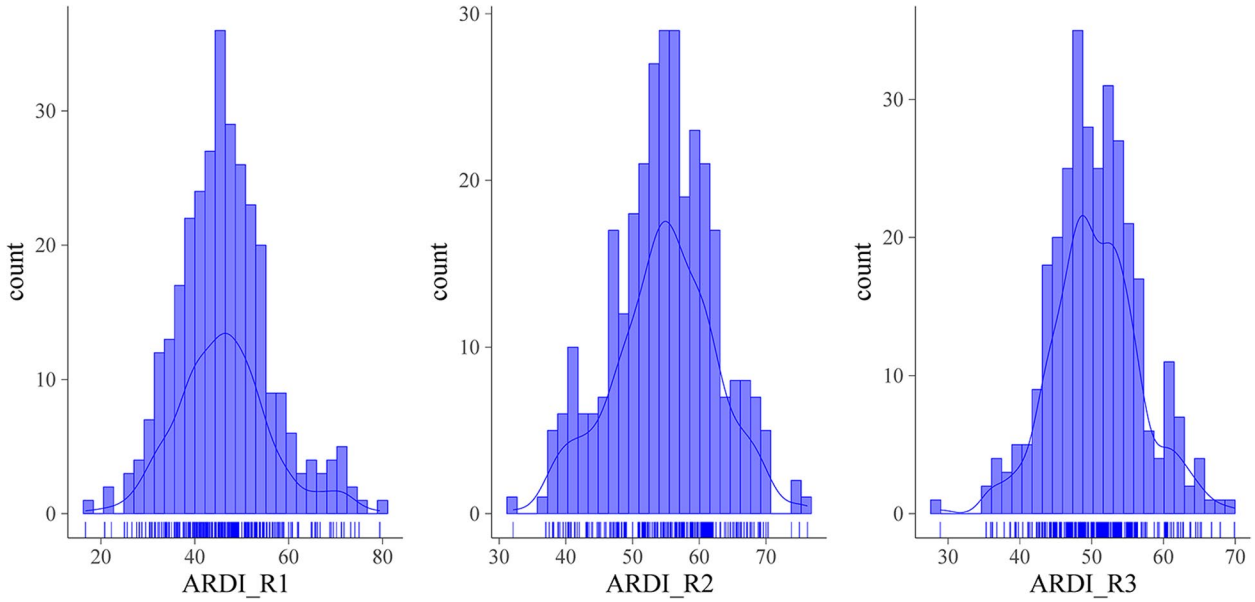

Fig. 1 Histogram of phenotypic frequency distribution for ARDI in three replications

each subgroup is as follows: subgroup 1 (green) includes 71 lines, subgroup 2 (orange) contains 91 lines, and subgroup 3 (yellow) comprises 166 lines. Additionally, the eight parents are distributed among the subgroups as follows: subgroup 1 includes P4, P6, and P7, subgroup 2 contains P2 and P8, and subgroup 3 includes P1, P3, and P5. This distribution highlights the genetic relationships between the MAGIC lines and their parental genotypes.

Identification of SNP loci for Verticillium wilt resistance in MAGIC population by using GWAS

A MLM model was used to conduct correlation analysis in the visualization interface, and Manhattan plots and Q-Q plots of three sets of data were obtained by GWAS as shown in Fig. 5. The SNPs shown in Fig. 5 were based on the combined analysis of data from all three environments (R1, R2, and R3). Used $-\log_{10}P>4$ as the screening criterion to detect significant Verticillium wilt resistance SNP loci, a total of 19 significant SNP loci associated with Verticillium wilt resistance were identified in the MAGIC population (Table 5). These 19 SNP loci were distributed on three chromosomes, *i.e.* sixteen SNP loci on A01 chromosome, two SNP loci on A02 chromosome, and one on D01 chromosome. Among these, seven SNPs are in the 5' UTR, six in the 3' UTR, with two missense and four synonymous variants. For example, on Chr A01, SNP 1_10496127 showed a synonymous variant (A to T, Gly), and SNP 1_10505144 had a missense mutation (A to T, Gln to Leu). At this threshold level, two highly stable sites were obtained at their specific position of 10 858 740 and 24 848 832 on Chromosomes A01 and A02 as shown in Table 5.

Screening of candidate genes by gene annotation and transcriptome

Through GWAS analysis, significant SNPs associated with the trait of interest were identified on chromosomes A01, A02, and D01, the related candidate genes were listed in Supplementary Tables S1-S3. Nineteen candidate genes were highlighted, as shown in Table 6. To further explore their potential roles, the tissue-specific expression patterns of these candidate genes in *G. hirsutum* TM-1 were examined using publicly available transcriptomic data. The results revealed that two genes, *Ghir_A01G006660* and *Ghir_A02G008980*, were consistently expressed, while the remaining genes exhibited diverse expression patterns that were categorized into three groups: Group-1 genes (including *Ghir_A01G006410*, *Ghir_A01G006450*, *Ghir_A01G006650*, and *Ghir_A01G006660*) were predominantly expressed in leaf or root tissues; Group-2 genes (including *Ghir_A01G006680*, *Ghir_A01G006540*, *Ghir_A01G006620*, *Ghir_A01G006670*, *Ghir_A01G006600*, *Ghir_A01G006710* and *Ghir_A01G006720*) were mainly expressed in reproductive tissues; and Group-3 genes (including *Ghir_A01G006520*, *Ghir_A01G006530*, *Ghir_A01G006640*, *Ghir_D01G014360* and *Ghir_A01G006500*, *Ghir_A02G008980*) were expressed primarily in root and stem tissues as shown in Fig. 6A. To further refine the selection of candidate genes, the expression patterns of 17 functional genes were analyzed in *Verticillium dahliae* resistant (M138) and *Verticillium dahliae* susceptible (P2) cotton cultivars following inoculation with *Verticillium dahliae* at 0, 1, 6, 12, 24, and 48 h. The results shown in Fig. 6B demonstrated differential expression of these genes

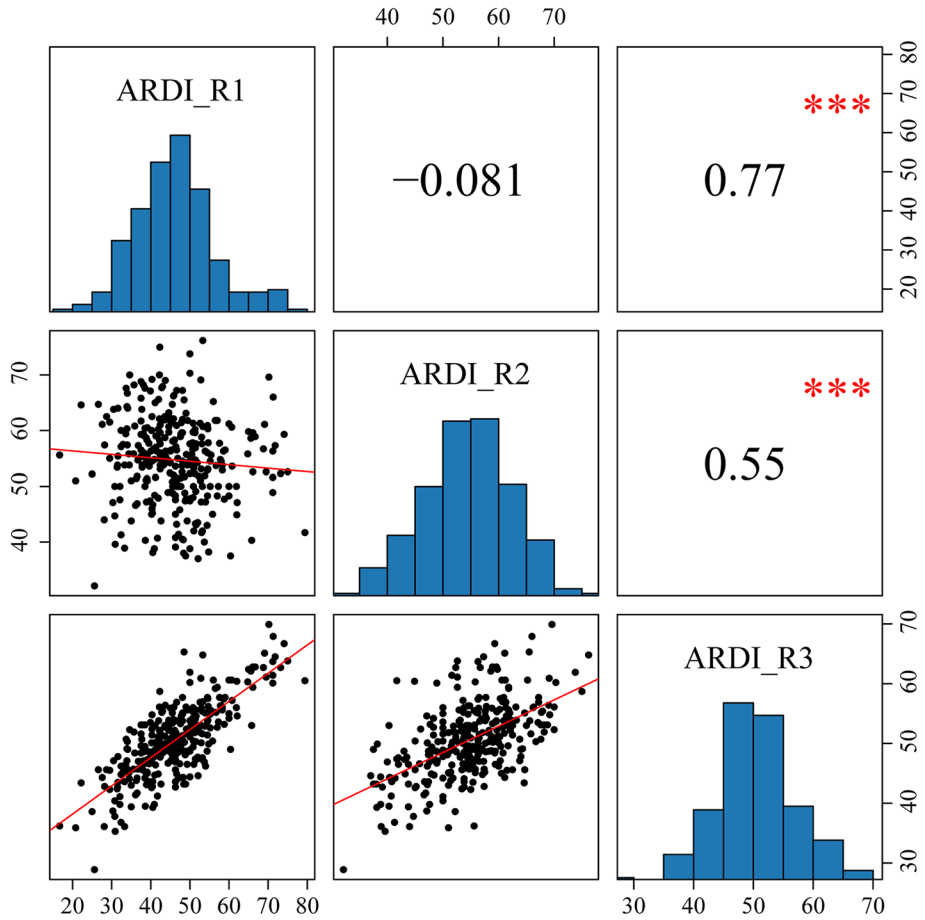

Fig. 2 Correlation analysis of ARDI across three replicates. *** represents significant at the 0.001 level

Table 4 High-quality SNPs genotyped in the MAGIC population

Chr	SNP number	SNP density /Mb	Chr	SNP number	SNP density /Mb
A01	88 640	1 328	D01	82 245	768
A02	43 714	2 472	D02	74 816	933
A03	35 561	3 179	D03	35 055	1 503
A04	28 489	2 988	D04	16 966	3 326
A05	67 686	1 616	D05	35 033	1 796
A06	179 504	691	D06	45 901	1 456
A07	141 684	690	D07	43 860	1 351
A08	98 781	1 238	D08	77 047	896
A09	50 783	1 616	D09	103 982	507
A10	84 023	1 366	D10	58 215	1 168
A11	76 752	1 605	D11	45 541	1 601
A12	56 207	1 915	D12	43 221	1 450
A13	135 183	801	D13	33 830	1 872
Total	1 087 007		Total	695 712	

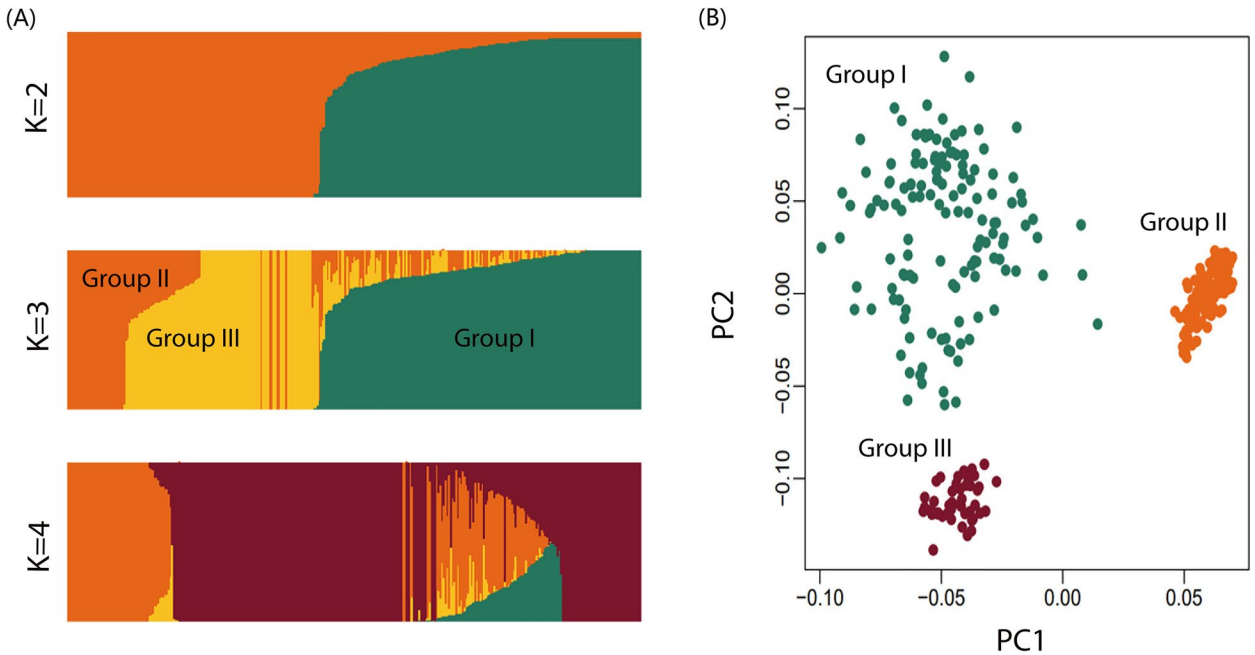

Fig. 3 Analysis of population structure of the 328 cotton lines. **A** Population structure of the 328 cotton lines at K = 2, K = 3, and K = 4. **B** PCA of the MAGIC population

between the resistant and susceptible cultivars at various time points. Based on these results, two candidate genes, *Ghir_A01G006660* and *Ghir_A02G008980*, were identified as the most likely associated with resistance to Verticillium wilt. These genes are homologous to *AT5G54770* and *AT1G05230* in *Arabidopsis thaliana* and belong to the thiamine biosynthetic gene family and the homeobox-leucine zipper protein family, respectively, suggesting their critical roles in resistance mechanisms against *Verticillium dahliae*.

Discussion

In this study, we identified 19 SNPs significantly associated with Verticillium wilt resistance in cotton using a GWAS approach. In a previous study, GWAS was also used to detect SNP associated with Verticillium wilt resistance in cotton and identified 17 SNPs that had significant association with *Verticillium dahliae* response (Li et al. 2017). To further confirm these loci for Verticillium wilt resistance, a comparison of the GWAS was performed with QTLs identified in previous studies. A total of 85 QTLs for Verticillium wilt resistance containing 139 simple sequence repeat (SSR) markers were selected from 11 QTL mapping reports (Li et al. 2017). The physical locations of these SSR primer sequences were mapped to the reference genome sequence via electronic polymerase chain reaction (e-PCR) (Zhang et al. 2015a). Some peak SNPs on chromesone A02 identified in another study overlapped with previous QTLs, although they were lower than the threshold ($P < 1.17 \times 10^{-5}$,$-\log_{10}P = 4.93$) (Li et al. 2017). For example, the SNP (A02_45851582) located on the A02 peak in RDIF2015 was positioned between MUSS294 and NAU1072 in QTL *qFDI711-30–-0.01* (Li et al. 2017). Peak SNPs A13_2361609 identified in RDIG2015 and A13_2361299 in RDIF2015 were positioned between NAU2730 and NAU5110 in *qVWI-09-c13-1* (Zhang et al. 2015c). Peak SNP D05_54737235 in RDIG2015 was mapped to the regions of QTLs *qFDI711-27–26.01* and *qVV-D5- 1BC1S2592* (Yang et al. 2008). Peak SNP D12_11092664 identified in RDIF2015 was positioned between BNL3867 and BNL1605 in *q7.22–2* (Wang et al. 2008).

In another previous study gene annotation showed that most candidate genes were related to disease resistance in plant. For qRDI-A01-1, one positively regulated gene and two negatively regulated genes were identified (Zhao et al. 2021), the negative regulated gene *Ghir_A01G022110* coding for 2-oxoglutarate-dependent dioxygenase involved in the flavonoid biosynthesis pathway, which is demonstrated to be related to Verticillium wilt resistance by the fact that the enrichment of flavonoids in a spontaneous cotton mutant with red coloration results in a significantly increase in resistance to *Verticillium dahliae* (Long et al. 2019). For qRDI-D01-1, one negative regulated gene, *Ghir_D01G023650*, was identified to encode a dioxygenase (DOX) involved in the flavanol synthase, which is a homologue to the negative regulated gene *Ghir_A01G022110* in

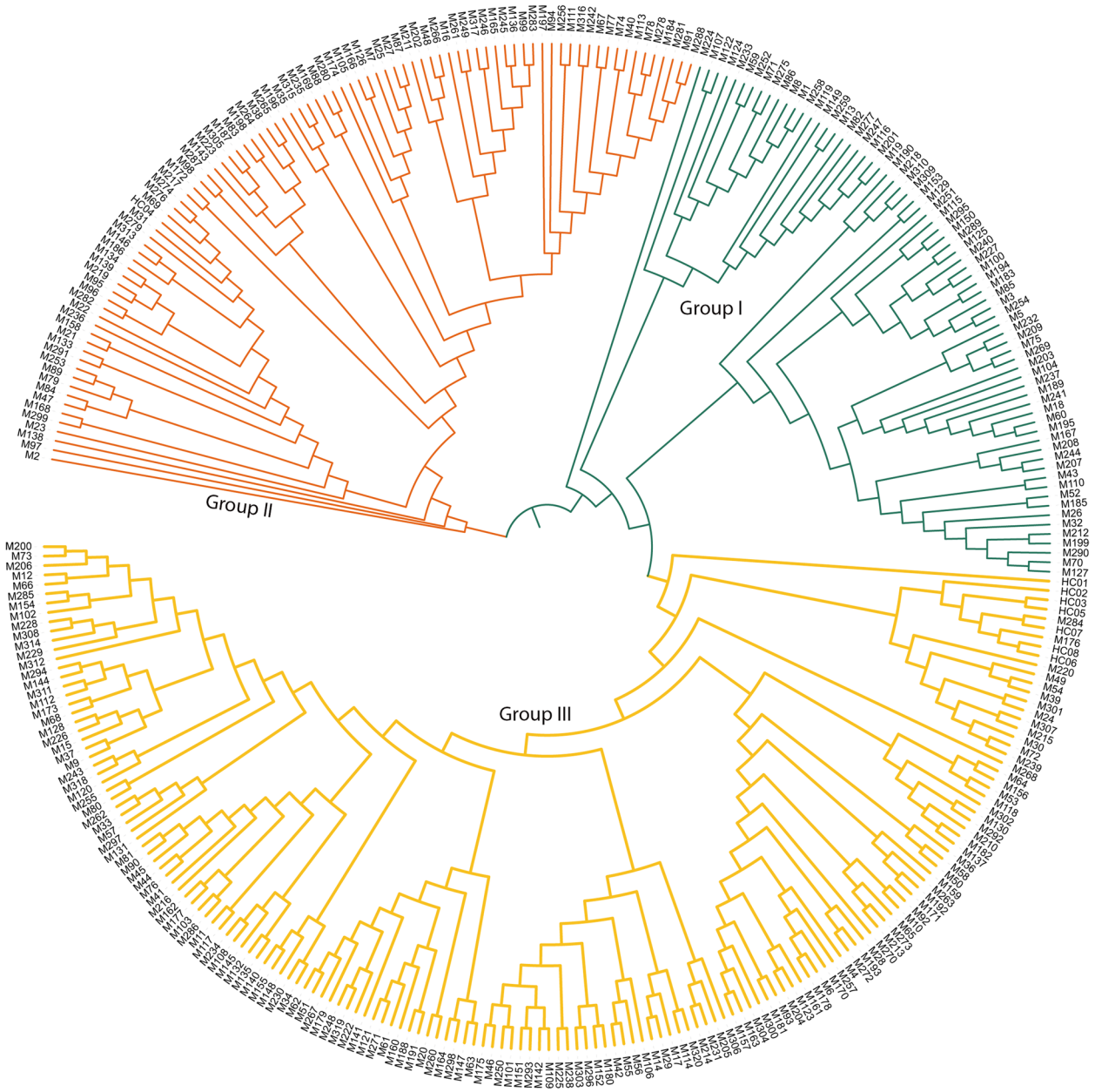

Fig. 4 Unrooted phylogenetic tree of the MAGIC population using multisequence alignment of *G. hirsutum*. Members in the clade were shown in the same color to represent their subgroup

qRDI-A01-1, and is related to the flavonoid biosynthesis pathway involved in Verticillium wilt resistance in cotton (Hernández-Vega et al. 2017).

These results show that the loci for Verticillium wilt resistance identified in this study had considerable overlaps with previously reported loci, suggesting that a GWAS with a panel of 328 cotton lines is suitable for the identification of significant SNPs associated with Verticillium wilt resistance. The analysis provided critical information for breeding programs targeting cotton resistance to Verticillium wilt. In addition to assisting with population classification, these analyses provide a solid foundation for advancing future genetic and statistical research. Understanding the innate structure of this MAGIC population is important to improving the accuracy and consistency of genetic studies. Complex quantitative traits in plants have been mapped using GWAS (Crowell et al. 2016). The power of a GWAS can be determined by four primary elements: marker density, statistical methodologies, trait acquisition veracity, and genetic diversity richness. There

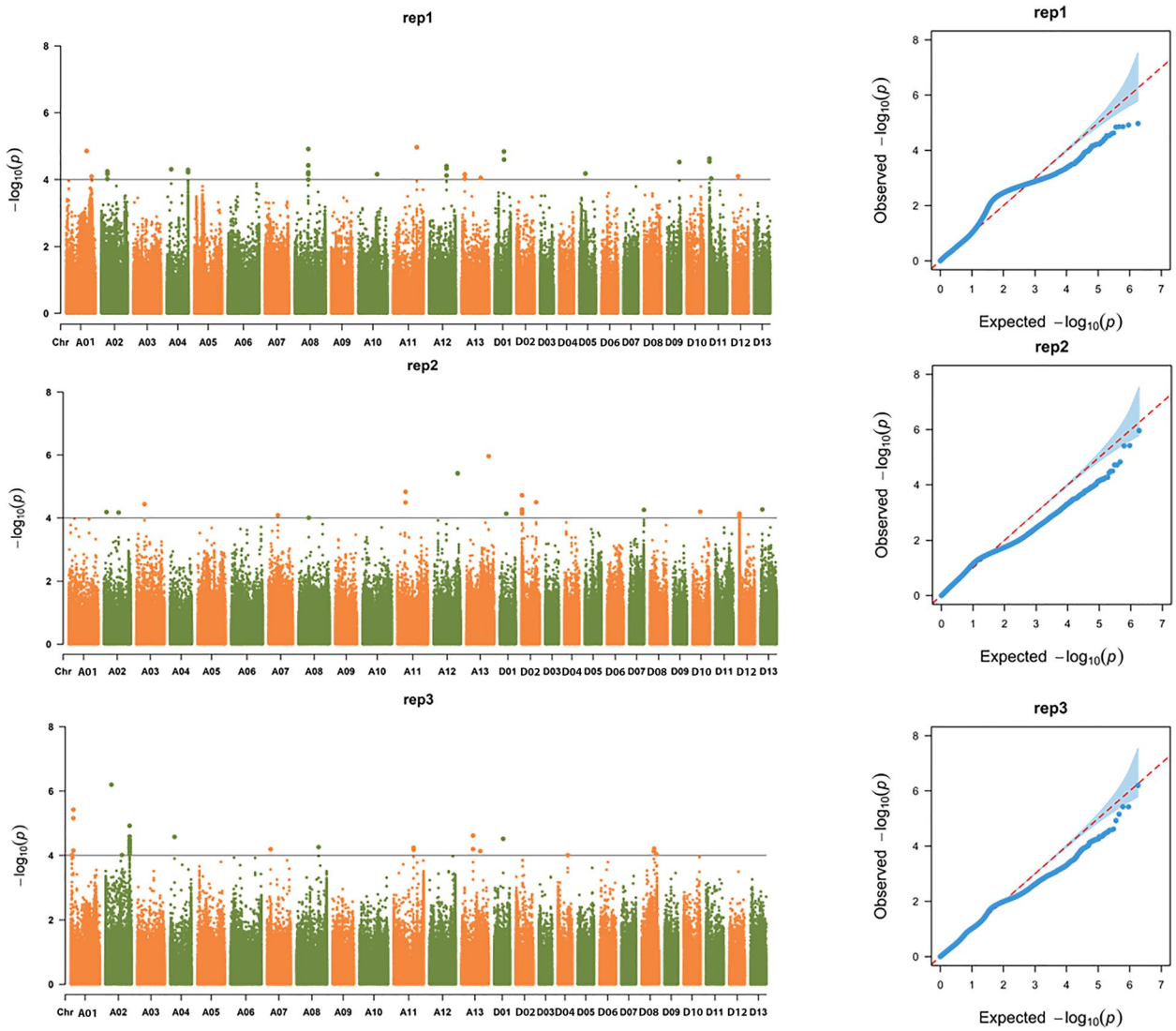

Fig. 5 The result of GWAS of Verticillium wilt resistance in the MAGIC population. The right figures represent the Manhattan plot and the left figures represent the QQ plot of the three environments, respectively

is a significant degree of genotypic and phenotypic variation in the Chinese cultivar collection of *G. hirsutum*. In addition to ensuring adequate genetic variation, the sample size was comparable to the GWAS population sizes used for *Zea mays* (Wen et al. 2014), *Arabidopsis thaliana* (Zhao et al. 2007), and *Brassica napus* (Xu et al. 2016). The influence of the environment on phenotypic variance compromises the validity of QTL mapping.

Key candidate genes associated with VW resistance

Based on the expression patterns of candidate genes and functional annotation, two candidate genes, *Ghir_A01G006660* and *Ghir_A02G008980*, were identified in this study. Whether these particular genes influence cotton resistance to Verticillium wilt is one of the questions that raises for further research. This work highlights the need for further experimental validation to determine the functionality and efficacy of related genes and genetic markers in preventing cotton Verticillium wilt, even though these tools offer valuable information. Importantly, the suggested experimental validation will direct breeding initiatives backed by molecular markers. These recently discovered genetic markers could be used in breeding processes to improve the accuracy and efficacy of cotton breeding programs aimed at transferring resistance to Verticillium wilt.

Table 5 The significant SNPs associated with Verticillium wilt resistance

SNP name	Chrom	Position/ bp	Ref/Alt	*P*-value	SNP type	Gene	Distance from gene	Amino Acid change
1_10272467	A01	10 272 467	G/A	0.628038	5_prime_UTR_variant	*Ghir_A01G006410*	c.–401G > A	
1_10357404	A01	10 357 404	T/C	0.482081	3_prime_UTR_variant	*Ghir_A01G006450*	c. + 306A > G	
1_10455757	A01	10 455 757	C/T	0.332541	3_prime_UTR_variant	*Ghir_A01G006500*	c. + 276G > A	
1_10474409	A01	10 474 409	T/A	0.489779	3_prime_UTR_variant	*Ghir_A01G006520*	c. + 363A > T	
1_10496127	A01	10 496 127	A/T	0.498727	synonymous_variant	*Ghir_A01G006530*	c.720A > T	p.Gly240Gly
1_10505144	A01	10 505 144	A/T	0.68326	missense_variant	*Ghir_A01G006540*	c.56A > T	p.Gln19Leu
1_10823589	A01	10 823 589	C/T	0.327431	missense_variant	*Ghir_A01G006600*	c.1132C > T	p.Arg378Cys
1_10835614	A01	10 835 614	A/G	0.248498	5_prime_UTR_variant	*Ghir_A01G006620*	c.–10A > G	
1_10839243	A01	10 839 243	T/A	0.412287	3_prime_UTR_variant	*Ghir_A01G006630*	c. + 489 T > A	
1_10844785	A01	10 844 785	G/C	0.480341	5_prime_UTR_variant	*Ghir_A01G006640*	c.–232G > C	
1_10851769	A01	10 851 769	C/G	0.292882	3_prime_UTR_variant	*Ghir_A01G006650*	c. + 18C > G	
1_10858740	A01	10 858 740	G/A	0.220158	5_prime_UTR_variant	*Ghir_A01G006660*	c.–10G > A	
1_10862677	A01	10 862 677	T/C	0.267366	5_prime_UTR_variant	*Ghir_A01G006670*	c.–349 T > C	
1_10871429	A01	10 871 429	C/T	0.691852	synonymous_variant	*Ghir_A01G006680*	c.648G > A	p.Ser216Ser
1_10928341	A01	10 928 341	C/T	0.281593	5_prime_UTR_variant	*Ghir_A01G006710*	c.–166C > T	
1_10936720	A01	10 936 720	C/T	0.381335	5_prime_UTR_variant	*Ghir_A01G006720*	c.–6G > A	
2_24848832	A02	24 848 832	C/T	0.231726	synonymous_variant	*Ghir_A02G008980*	c.1180G > A	p.Ile394Ile
2_25263551	A02	25 263 551	G/A	0.110986	synonymous_variant	*Ghir_A02G008990*	c.90C > T	p.Ser30Ser
14_37452750	D01	37 452 750	G/A	0.175257	3_prime_UTR_variant	*Ghir_D01G014360*	c. + 315C > T	

A Adenine, *G* Guanine, *C* Cytosine, *T* Thymine

Ghir_A01G006660 encodes thiamine which is essential for cotton resistance to Verticillium wilt. Thiamine-induced resistance in *Arabidopsis* has been shown in a study to be effective against a variety of pathogens, including bacteria and fungi (Ahn et al. 2005). The thiazole moiety of thiamine, or vitamin B1, is produced by this gene and it is an essential component of cellular metabolism. It has been demonstrated that thiamine, is involved in plant defense systems and is necessary for plant metabolism. Rice resistance to abiotic stress was increased by overexpressing a gene involved in thiamine biosynthesis (Samanta et al. 2020). Enhancing thiamine biosynthesis and mitotic DNA damage tolerance. It is crucial for enhancing cotton resilience to diseases because of its role in important metabolic pathways and stress reactions. The foundation for creating disease-resistant cotton varieties was established by the identification and validation of this gene through MAGIC population research.

Ghir_A02G008980 encodes a protein belonging to the homeobox leucine zipper (HD-Zip) family. The regulatory involvement of this gene in growth, development, and stress response pathways makes it important for cotton resistance to Verticillium wilt. HD-Zip proteins help plants respond to stress, such as pathogen attacks. This study found that the HD-Zip gene, *Ghir_A02G008980,* is effective in the resistance to Verticillium wilt. An investigation in *Arabidopsis* showed that HD-Zip proteins can control the expression of genes linked to defense, strengthening plant resistance against a range of infections (Zhang et al. 2014). Understanding the involvement of HD-Zip proteins in stress responses is possible through functional studies of these proteins in model plants, such as *Arabidopsis*. This study can be concluded by comprehending their roles in cotton and how they affect disease resistance (Henriksson et al. 2005). The HD-Zip family protein encoded by *Ghir_A02G008980* is a major factor in the ability of cotton plants to withstand Verticillium wilt. Further studies are required to examine these potential candidate genes. To establish a link between gene expression and phenotypic variation, additional studies are needed to validate the candidate genes. This visual representation offered valuable insights into the genetic architecture of the MAGIC population, enabling a better understanding of its evolutionary dynamics and facilitating targeted breeding strategies for cotton improvement.

The advantage of MAGIC populations and their application in gene mapping

Although a large number of SNP markers (1 782 719) were used in this study, only 19 SNPs were found to be significantly associated with Verticillium wilt resistance. This lower detection ratio may be due to several reasons. First, Verticillium wilt resistance is a complex trait controlled by many small-effect genes, making it harder to

Table 6 The candidate genes and their annotations derived from GWAS (Chromesome A01, A02, and D01)

Chrome	Gene ID	Direction	Start site	End site	Arabidopsis homologous genes	Annotation
A01	*Ghir_A01G006410*	+	10 265 036	10 267 499	*AT5G54600*	Translation protein SH3-like family protein (source: Araport11)
A01	*Ghir_A01G006450*	–	10 350 924	10 353 615	*AT5G54630*	Zinc finger protein-like protein (source: Araport11)
A01	*Ghir_A01G006500*	–	10 447 518	10 452 586	*AT5G54690*	Encodes a protein with putative galacturonosyl transferase activity. Mutants defective in this gene displayed a notable reduction in xylose (> 50%) in the cell walls from stems and roots and a reduction in cellulose (~ 25%)
A01	*Ghir_A01G006520*	–	10 467 876	10 473 157	*AT5G54730*	Yeast autophagy 18 F-like protein (source: Araport11)
A01	*Ghir_A01G006530*	+	10 486 951	10 495 367	*AT3G20720*	Amino-terminal region of chorein (source: Araport11)
A01	*Ghir_A01G006540*	+	10 498 661	10 499 242	*AT1G03670*	Ankyrin repeat-containing protein
A01	*Ghir_A01G006600*	+	10 814 956	10 818 733	*AT3G23980*	Encodes a protein that interacts with the Polycomb-group (Pc-G) histone methyltransferase CLF (CURLY LEAF). It colocalizes with CLF to the nucleus and represses a subset of Pc-G target genes. The pleiotropic developmental mutant phenotype suggests that BLI prevents premature differentiation
A01	*Ghir_A01G006620*	+	10 828 187	10 830 408	*AT1G54290*	Translation initiation factor SUI1 family protein (source: Araport11)
A01	*Ghir_A01G006630*	+	10 831 342	10 832 673	*AT1G01380*	ETC1 is involved in trichome and root hair patterning in Arabidopsis
A01	*Ghir_A01G006640*	+	10 837 937	10 841 189	*AT4G27120*	ER-resident adaptor protein. Part of a complex with C53 and UFL1, the E3 ligase that mediates simulation. Involved in the pathway that links ribosome-associated quality control with selective autophagy at the ER
A01	*Ghir_A01G006650*	+	10 843 201	10 845 603	*AT1G54290*	Translation initiation factor SUI1 family protein (source: Araport11)
A01	*Ghir_A01G006660*	+	10 852 083	10 853 731	*AT5G54770*	Encodes a thiamine biosynthetic gene that has a dual function in thiamine biosynthesis and mitochondrial DNA damage tolerance. It appears to be involved in producing the thiazole portion of thiamine (vitamin B1). A crystal structure of the protein reveals that it forms a 2-ring homo-octamer. The mRNA is cell-to-cell mobile
A01	*Ghir_A01G006670*	+	10 855 850	10 859 035	*AT3G20580*	COBRA-like protein 10 precursor (source: Araport11)
A01	*Ghir_A01G006680*	–	10 862 315	10 872 311	*AT5G54780*	Ypt/Rab-GAP domain of gyp1p superfamily protein (source: Araport11)
A01	*Ghir_A01G006710*	+	10 921 624	10 925 300	*AT1G50940*	Encodes the electron transfer flavoprotein ETF alpha, a putative subunit of the mitochondrial electron transfer flavoprotein complex (ETF beta is At5g43430.1) in Arabidopsis. Mutations of the ETF beta gene result in accelerated senescence and early death compared with wild-type during extended darkness
A01	*Ghir_A01G006720*	–	10 929 111	10 930 096	*AT5G54790*	CTD small phosphatase-like protein (source: Araport11)
A02	*Ghir_A02G008980*	–	24 837 008	24 838 741	*AT1G05230*	Encodes a homeobox-leucine zipper family protein belonging to the HD-ZIP IV family. Mutants have trichomes that appear glass-like under a dissecting microscope as compared with the wild-type trichomes. The mutations do not affect trichome growth or branch number
A02	*Ghir_A02G008990*	–	25 251 432	25 252 157	*AT1G30760*	Encodes a BBE-like enzyme that acts in monolignol metabolism by catalyzing the oxidation of aromatic allylic alcohols, such as coumarin-, sinapyl-, and coniferyl alcohol, to the corresponding aldehydes. The mRNA is cell-to-cell mobile
D01	*Ghir_D01G014360*	–	37 436 687	37 440 699	*AT1G57680*	Plasminogen activator inhibitor (source: Araport11)

detect a large number of significant associations. Second, while the MAGIC population provides more recombination and reduced linkage disequilibrium compared with other populations, the moderate population size (328 lines) may have limited the ability to detect SNPs with small effects. Additionally, the strict significance threshold used in the analysis might have excluded SNPs with weaker associations. Lastly, environmental variability could have influenced the disease response, potentially masking some of the genetic associations. Despite these

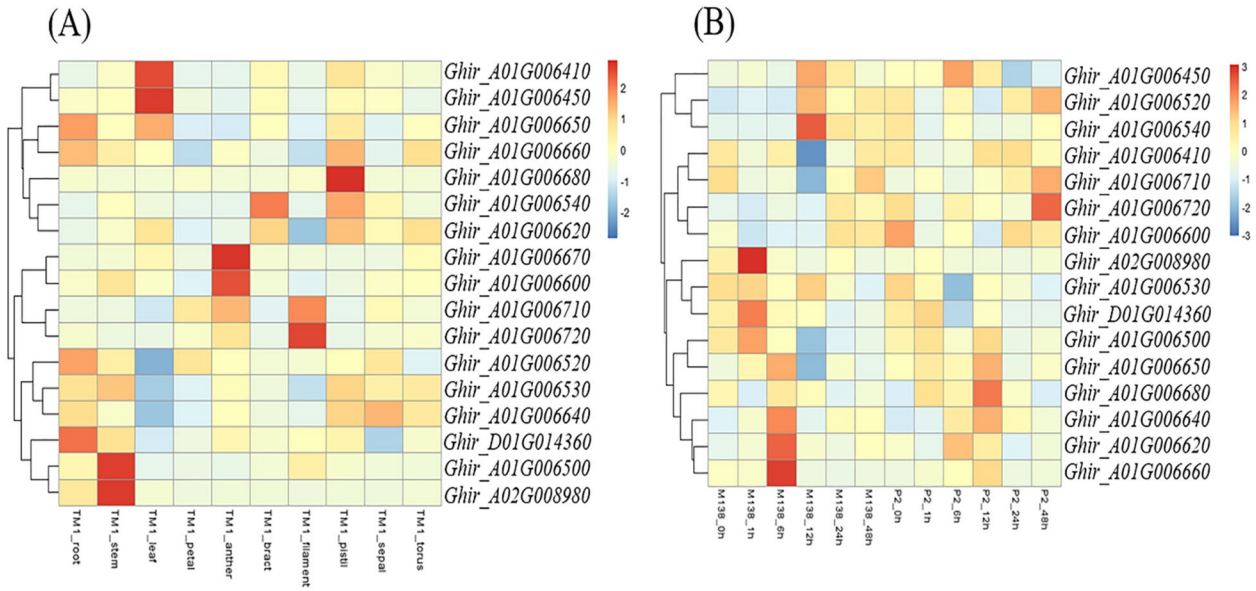

Fig. 6 Expression pattern of candidate genes in *Gossypium hirsutum*. **A** Gene expression profile of candidate genes for various tissues in TM-1. **B** Gene expression profile of candidate genes against M138 and P2 cotton cultivar

challenges, the identification of these 19 SNPs shows that the MAGIC population is still a powerful method for identifying key genetic loci associated with disease resistance. MAGIC populations have been used extensively in QTL mapping and crop breeding due to their strong performance, low false-positive rate, and robustness (Cavanagh et al. 2008; Huang et al. 2018). To improve the possibility of effective breeding through the polymerization of advantageous alleles, a MAGIC population was created for breeding purposes by combining several parents with desired features in a variety of traits (Cavanagh et al. 2008). For QTL mapping, a MAGIC population is more favorable for the genetic analysis of complex traits since it has a simpler structure than natural populations and more diversified genetic diversity and recombination rates than a bi-population (Cavanagh et al. 2008; Huang et al. 2015). The eight parents in this study showed distinct fiber yield, fiber quality, environmental adaptation, and resistance to Verticillium wilt. They were developed from breeding lines and varieties. As a result, the MAGIC population developed for this study aids in the genetic investigation of other features, cotton breeding, and the underlying genetic basis of Verticillium wilt resistance.

A new perspective on the genetic basis of VW resistance in cotton

The deadliest crop disease impacting cotton yield and fiber quality is cotton Verticillium wilt (Xu et al. 2011). Extensive research has been conducted on mapping Verticillium wilt resistance genes/QTLs in an attempt to find alternative remedies to the issue (Zhang et al. 2015c). However, because of many issues with populations, phenotype identification, and environment, only a few important genes have been found and cloned (Zhang et al. 2020). For example, the empirical irreproducibility of early segregating populations (e.g., F_2, BC_1, and $F_{2:3}$) makes it impossible to reliably identify the Verticillium wilt resistance trait in any plant (Zhang et al. 2015b). This limitation can be addressed by substituting permanent populations, like MAGIC populations, recombinant or backcross inbred lines (Fang et al. 2014). Furthermore, it might be challenging to identify Verticillium wilt resistance because the resistance is a complicated mosaic phenotype. The typical technique for determining Verticillium wilt resistance is an artificial grading scale (Zhang et al. 2021). Breeding resistant varieties are the most economical and effective method to control Verticillium wilt (Wu et al. 2021).

Conclusion

This study carried out an extensive GWAS to identify significant SNPs that are substantially associated with cotton resistance to Verticillium wilt. Nineteen SNP were found to have significant correlations with ARDI in response to Verticillium wilt. Additionally, the screening of candidate genes was performed on the transcriptomic data and two candidate genes were identified, i.e. *Ghir_A01G006660* and *Ghir_A02G008980*, which were associated with the resistance to Verticillium wilt. The research highlights the necessity for further experimental validation to determine the functionality and efficacy of these genes in cotton Verticillium wilt resistance.

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s42397-025-00211-7.

Supplementary Material 1.

Acknowledgements
We would like to thank the anonymous reviewers for their valuable comments and helpful suggestions; these have greatly helped us improve the quality of the manuscript.

Authors' contributions
Nie XH designed the experiments. Ayyaz M and Chang ZW performed the experiments. Ayyaz M wrote the main manuscript and prepared all the figures. Ding SG, Han P and Xu L performed the data analysis. Wu YL, Xu JW, Abudukeyoumu A, Siddho IA, Li ZB, and Lin HR revised and polished the manuscript. All authors contributed to the interpretation of results and have read and approved the final manuscript.

Funding
This work was supported by funding from the fund for National Key Research and Development Program of China (2023YFD2301203-05), National Natural Science Foundation of China (32260510), Special Financial Project for Seed Industry Development in the Autonomous Region (BNZJ2024-10, BNZJ2024-30), Key Project for Science and Technology Development of Shihezi city, Xinjiang Production and Construction Crops (2022NY01), Shihezi University high-level talent research project (RCZK202337), and Science and Technology Planning of Shuanghe city, Xinjiang Production and Construction Crops (2021NY02).

Data availability
The datasets used and analyzed in the current study are available from the corresponding author upon reasonable request.

Declarations

Ethics approval and consent to participate
Not applicable.

Consent for publication
Not applicable.

Competing interests
The authors declare that they have no competing interests. Author Nie XH is a member of the Editorial Board of Journal of Cotton Research. Author Nie XH was not involved in the journal's review of, or decision related to this manuscript.

Author details
[1]Key Laboratory of Oasis Ecology Agricultural of Xinjiang Production and Construction Corps, Agricultural College, Shihezi University, Shihezi, Xinjiang 832003, China. [2]Institute of Crop Variety Resources, Xinjiang Academy of Agricultural Sciences, Urumqi 830091, China. [3]National Central Asian Characteristic Crop Germplasm Resources Medium-Term Gene Bank (Urumqi), Urumqi 830091, China.

Received: 2 July 2024 Accepted: 17 January 2025
Published online: 26 February 2025

References

Ahn IP, Kim S, Lee YH. Vitamin B_1 functions as an activator of plant disease resistance. Plant Physiol. 2005;138(3):1505–15. https://doi.org/10.1104/pp.104.058693.
Bandillo N, Raghavan C, Muyco PA, et al. Multi-parent advanced generation inter-cross (MAGIC) populations in rice: progress and potential for genetics research and breeding. Rice. 2013;6:1–15. https://doi.org/10.1186/1939-8433-6-11.
Cai X, Hou Y, Muhammad UJ, et al. The morphological diversity of pollen in the genus *Gossypium*. J Cotton Res. 2023;6(1):6. https://doi.org/10.1186/s42397-023-00143-0.
Campbell B, Saha S, Percy R, et al. Status of the global cotton germplasm resources. Crop Sci. 2010;50(4):1161–79. https://doi.org/10.2135/cropsci2009.09.0551.
Cavanagh C, Morell M, Mackay I, et al. From mutations to MAGIC: resources for gene discovery, validation and delivery in crop plants. Curr Opin Plant Biol. 2008;11(2):215–21. https://doi.org/10.1016/j.pbi.2008.01.002.
Crowell S, Korniliev P, Falcao A, et al. Genome-wide association and high-resolution phenotyping link *Oryza sativa* panicle traits to numerous trait-specific QTL clusters. Nat Commun. 2016;7(1):10527. https://doi.org/10.1038/ncomms10527.
Diaz S, Ariza-Suarez D, Izquierdo P, et al. Genetic mapping for agronomic traits in a MAGIC population of common bean (*Phaseolus vulgaris* L.) under drought conditions. BMC Genomics. 2020;21(1):799. https://doi.org/10.1186/s12864-020-07213-6.
Fang H, Zhou H, Sanogo S, et al. Quantitative trait locus analysis of Verticillium wilt resistance in an introgressed recombinant inbred population of upland cotton. Mol Breed. 2014;33:709–20. https://doi.org/10.1007/s11032-013-9987-9.
Fonseca CS, da Silva NR, Ballesteros LF, et al. *Penicillium brevicompactum* as a novel source of natural pigments with potential for food applications. Food Bioprod Process. 2022;132(1):188–99. https://doi.org/10.1016/j.fbp.2022.01.007.
Giraud H, Bauland C, Falque M, et al. Linkage analysis and association mapping QTL detection models for hybrids between multiparental populations from two heterotic groups: application to biomass production in maize (*Zea mays* L.). G3 (Bethesda). 2017;7(11):3649–57. https://doi.org/10.1534/g3.117.300121.
Henriksson E, Olsson AS, Johannesson H, et al. Homeodomain leucine zipper class I genes in *Arabidopsis*. Expression patterns and phylogenetic relationships. Plant Physiol. 2005;139 (1):509–18. https://doi.org/10.1104/pp.105.063461.
Hernández-Vega JC, Cady B, Kayanja G, et al. Detoxification of polycyclic aromatic hydrocarbons (PAHs) in *Arabidopsis thaliana* involves a putative flavonol synthase. J Hazard Mater. 2017;321:268–80. https://doi.org/10.1016/j.jhazmat.2016.08.058.
Huang BE, George AW, Forrest KL, et al. A multiparent advanced generation inter-cross population for genetic analysis in wheat. Plant Biotechnol J. 2012;10(7):826–39. https://doi.org/10.1111/j.1467-7652.2012.00702.x.
Huang BE, Verbyla KL, Verbyla AP, et al. MAGIC populations in crops: current status and future prospects. Theor Appl Genet. 2015;128:999–1017. https://doi.org/10.1007/s00122-015-2506-0.
Huang C, Shen C, Wen T, et al. SSR-based association mapping of fiber quality in upland cotton using an eight-way MAGIC population. Mol Genet Genomics. 2018;293:793–805. https://doi.org/10.1007/s00438-018-1419-4.
Ibarra-Cortés K, Guzmán-Franco A, González-Hernández H, et al. Susceptibility of *Diaphorina citri* (Hemiptera: *Liviidae*) and its parasitoid *Tamarixia radiata* (Hymenoptera: *Eulophidae*) to entomopathogenic fungi under laboratory conditions. Neotrop Entomol. 2018;47:131–8. https://doi.org/10.1007/s13744-017-0539-6.
Karademir E, Karademir Ç, Ekinci R, et al. Assessment of tolerance level of some Cotton (*Gossypium hirsutum* L.) varieties against Verticillium wilt (*Verticillium dahliae* Kleb.) disease. Not Bot Horti Agrobot Cluj-Napoca. 2010;38(1):196–202. https://doi.org/10.15835/nbha3814574.
Khaskheli MI, Sun JL, He SP, et al. Screening of cotton germplasm for resistance to *Verticillium dahliae* Kleb. under greenhouse and field conditions. Eur J Plant Pathol. 2013;137(2):259–72. https://doi.org/10.1007/s10658-013-0235-2.
Li H, Durbin RJB. Fast and accurate long-read alignment with burrows–wheeler transform. Bioinformatics. 2010;26(5):589–95. https://doi.org/10.1093/bioinformatics/btp698.
Li T, Ma X, Li N, et al. Genome-wide association study discovered candidate genes of Verticillium wilt resistance in upland cotton (*Gossypium hirsutum* L.). Plant Biotechnol J. 2017;15(12):1520–32. https://doi.org/10.1111/pbi.12734.

Li R, Ma XY, Zhang YJ, et al. Genome-wide identification and analysis of a cotton secretome reveals its role in resistance against *Verticillium dahliae*. BMC Biol. 2023;21(1):166. https://doi.org/10.1186/s12915-023-01650-x.

Long L, Liu J, Gao Y, et al. Flavonoid accumulation in spontaneous cotton mutant results in red coloration and enhanced disease resistance. Plant Physiol Biochem. 2019;143:40–9. https://doi.org/10.1016/j.plaphy.2019.08.021.

Micheletti D, Dettori MT, Micali S, et al. Whole-genome analysis of diversity and SNP-major gene association in peach germplasm. PLoS ONE. 2015;10(9):e0136803. https://doi.org/10.1371/journal.pone.0136803 .

Novakazi F, Krusell L, Jensen JD, et al. You had me at "MAGIC"!: four barley MAGIC populations reveal novel resistance QTL for *powdery mildew*. Genes. 2020;11(12):1512. https://doi.org/10.3390/genes11121512.

Ongom PO, Ejeta G. Mating design and genetic structure of a multi-parent advanced generation intercross (MAGIC) population of sorghum (*Sorghum bicolor* (L.) Moench). G3 (Bethesda). 2018;8(1):331–41. https://doi.org/10.1534/g3.117.300248.

Palanga KK, Jamshed M, Rashid MHo, et al. Quantitative trait locus mapping for Verticillium wilt resistance in an upland cotton recombinant inbred line using SNP-based high density genetic map. Front Plant Sci. 2017;8:382. https://doi.org/10.3389/fpls.2017.00382 .

Pascual L, Desplat N, Huang BE, et al. Potential of a tomato MAGIC population to decipher the genetic control of quantitative traits and detect causal variants in the resequencing era. Plant Biotechnol J. 2015;13(4):565–77. https://doi.org/10.1111/pbi.12282.

Pincot DD, Hardigan MA, Cole GS, et al. Accuracy of genomic selection and long-term genetic gain for resistance to Verticillium wilt in strawberry. Plant Genome. 2020;13(3):e2005. https://doi.org/10.1002/tpg2.20054 .

Pritchard JK, Stephens M, Donnelly PJG. Inference of population structure using multilocus genotype data. Genetics. 2000;155(2):945–59. https://doi.org/10.1093/genetics/155.2.945.

Samanta S, Singh A, Roychoudhury A. Involvement of sulfur in the regulation of abiotic stress tolerance in plants. In: Roychoudhury A, Tripathi DK, editors. Protective chemical agents in the amelioration of plant abiotic stress: biochemical and molecular perspectives. Hoboken: Wiley; 2020. p. 437–66. https://doi.org/10.1002/9781119552154.ch22.

Su J, Wang C, Ma Q, et al. An RTM-GWAS procedure reveals the QTL alleles and candidate genes for three yield-related traits in upland cotton. BMC Plant Biol. 2020;20:1–15. https://doi.org/10.1186/s12870-020-02613-y.

Tian XM, Peng H, Jing W, et al. Association mapping of lignin response to Verticillium wilt through an eight-way MAGIC population in upland cotton. J Integr Agric. 2023;22(5):1324–37. https://doi.org/10.1016/j.jia.2022.08.034.

Wagner TA, Gu A, Duke SE, et al. Genetic diversity and pathogenicity of *Verticillium dahliae* isolates and their co-occurrence with *Fusarium oxysporum* f. sp. vasinfectum causing cotton wilt in Xinjiang, China. Plant Dis. 2021;105(4):978–85. https://doi.org/10.1094/PDIS-09-20-2038-RE .

Wang HM, Lin ZX, Zhang XL, et al. Mapping and quantitative trait loci analysis of Verticillium wilt resistance genes in cotton. J Integr Plant Biol. 2008;50(2):174–82. https://doi.org/10.1111/j.1744-7909.2007.00612.x.

Wen W, Li D, Li X, et al. Metabolome-based genome-wide association study of maize kernel leads to novel biochemical insights. Nat Commun. 2014;5(1):3438. https://doi.org/10.1038/ncomms4438.

Wilson IW, Moncuquet P, Yuan Y, et al. Genetic mapping and characterization of Verticillium wilt resistance in a recombinant inbred population of upland cotton. Int J Mol Sci. 2024;25(4):2439. https://doi.org/10.3390/ijms25042439.

Wu Y, Zhang L, Zhou J, et al. Calcium-dependent protein kinase GhCDPK28 was identified and involved in Verticillium wilt resistance in cotton. Front Plant Sci. 2021;12:772649. https://doi.org/10.3389/fpls.2021.772649 .

Xu L, Hu K, Zhang Z, et al. Genome-wide association study reveals the genetic architecture of flowering time in rapeseed (*Brassica napus* L.). DNA Res. 2016;23(1):43–52. https://doi.org/10.1093/dnares/dsv035 .

Xu L, Zhu L, Tu 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. 2011;62(15):5607–21. https://doi.org/10.1093/jxb/err245.

Yang C, Guo W, Li G, et al. QTLs mapping for Verticillium wilt resistance at seedling and maturity stages in *Gossypium barbadense* L. Plant Sci. 2008;174(3):290–8. https://doi.org/10.1016/j.plantsci.2007.11.016.

Yang M, Huang C, Kang X, et al. Early monitoring of cotton Verticillium wilt by leaf multiple "Symptom" characteristics. Remote Sens. 2022;14(20):5241. https://doi.org/10.3390/rs14205241.

Yang J, Liu L, Yang L, et al. High-quality genome assembly of *Verticillium dahliae* VD991 allows for screening and validation of pathogenic genes. Front Microbiol. 2023;14:1177078. https://doi.org/10.3389/fmicb.2023.1177078.

Zhang T, Hu Y, Jiang W, et al. Sequencing of allotetraploid cotton (*Gossypium hirsutum* L. acc. TM-1) provides a resource for fiber improvement. Nat Biotechnol. 2015a;33(5):531–37. https://doi.org/10.1038/nbt.3207 .

Zhang Y, Wang XF, Ding ZG, et al. Transcriptome profiling of *Gossypium barbadense* inoculated with *Verticillium dahliae* provides a resource for cotton improvement. BMC Genomics. 2013;14:1–18. https://doi.org/10.1186/1471-2164-14-637 .

Zhang J, Fang H, Zhou H, et al. Genetics, breeding, and marker-assisted selection for Verticillium wilt resistance in cotton. Crop Sci. 2014;54(4):1289–303. https://doi.org/10.2135/cropsci2013.08.0550.

Zhang J, Sanogo S, Ma Z, et al. Breeding, genetics, and quantitative trait locus mapping for Fusarium wilt resistance in cotton. Crop Sci. 2015b;55(6):2435–52. https://doi.org/10.2135/cropsci2015.01.0056.

Zhang J, Yu J, Pei W, et al. Genetic analysis of Verticillium wilt resistance in a backcross inbred line population and a meta-analysis of quantitative trait loci for disease resistance in cotton. BMC Genomics. 2015c;16:1–13. https://doi.org/10.1186/s12864-015-1682-2.

Zhang J, Abdelraheem A, Thyssen GN, et al. Evaluation and genome-wide association study of Verticillium wilt resistance in a MAGIC population derived from intermating of eleven upland cotton (*Gossypium hirsutum*) parents. Euphytica. 2020;216:1–13. https://doi.org/10.1007/s10681-019-2547-6.

Zhang Y, Chen B, Sun Z, et al. A large-scale genomic association analysis identifies a fragment in Dt11 chromosome conferring cotton Verticillium wilt resistance. Plant Biotechnol J. 2021;19(10):2126–38. https://doi.org/10.1111/pbi.13650.

Zhang Q, Zhang C, Pan Z, et al. Genome-wide identification and analysis of the WNK kinase gene family in upland cotton. Plants. 2023;12(23):4036. https://doi.org/10.3390/plants12234036.

Zhao K, Aranzana MJ, Kim S, et al. An *Arabidopsis* example of association mapping in structured samples. PLoS Genet. 2007;3(1):e4. https://doi.org/10.1371/journal.pgen.0030004 .

Zhao Y, Chen W, Cui Y, et al. Detection of candidate genes and development of KASP markers for Verticillium wilt resistance by combining genome-wide association study, QTL-seq and transcriptome sequencing in cotton. Theor Appl Genet. 2021;134:1063–81. https://doi.org/10.1007/s00122-020-03752-4.

Zhao T, Guan X, Hu Y, et al. Population-wide DNA methylation polymorphisms at single-nucleotide resolution in 207 cotton accessions reveal epigenomic contributions to complex traits. Cell Res. 2024;34:1–14. https://doi.org/10.1038/s41422-024-01027-x.

Zheng M, Peng C, Liu H, et al. Genome-wide association study reveals candidate genes for control of plant height, branch initiation height and branch number in rapeseed (*Brassica napus* L.). Front Plant Sci. 2017;8:1246. https://doi.org/10.3389/fpls.2017.01246 .