Research Insight **Open Access**

Multi-Trait GWAS for Fiber Quality and Disease Resistance in Cotton

Shanjun Zhu, Mengting Luo
Institute of Life Science, Jiyang College of Zhejiang A&F University, Zhuji, 311800, China
Corresponding email: mengting.luo@jicat.org

Cotton Genomics and Genetics, 2025, Vol.16, No.5 doi: 10.5376/cgg.2025.16.0022
Received: 11 Jul., 2025
Accepted: 22 Aug., 2025
Published: 13 Sep., 2025

Preferred citation for this article:
Zhu S.J., and Luo M.T., 2025, Multi-trait GWAS for fiber quality and disease resistance in cotton, Cotton Genomics and Genetics, 16(5): 222-231 (doi: 10.5376/cgg.2025.16.0022)

Abstract Cotton is a globally important dual-purpose crop valued for its fiber yield, but both its yield and quality are severely impacted by a variety of pathogens. This study reviews the genetic architecture of fiber quality traits (such as strength, length, and fineness) and resistance to major diseases such as Verticillium wilt, Fusarium wilt, and bacterial wilt, focusing on potential genetic overlap and independence. We explore the methodological framework for multi-trait genome-wide association studies (MT-GWAS), highlighting statistical models such as multivariate linear mixed models and Bayesian methods, which outperform single-trait analyses by capturing pleiotropic loci and genetic correlations. We present key findings from cotton MT-GWAS, including the identification of co-localized QTLs, novel candidate genes, and genotype-by-environment interactions across multiple environmental datasets. We also highlight the integration of MT-GWAS with transcriptomic, metabolomic, epigenomic, and proteomic data, and the validation of functional genes using CRISPR, RNAi, and overexpression technologies. A case study demonstrates the practical application of MT-GWAS in a breeding program targeting fiber quality and disease resistance, enabling genetic validation and germplasm improvement. While MT-GWAS faces challenges such as population structure, statistical complexity, and translational gaps, advances in high-resolution phenotyping, pan-genomics, and predictive breeding strategies hold promise for broader application. This study highlights the potential of MT-GWAS to accelerate cotton improvement by revealing complex trait architecture and informing integrated breeding processes.

Keywords Multi-trait GWAS; Fiber quality; Disease resistance; Cotton breeding; Pleiotropy

1 Introduction

Globally, if one were to look for a crop that is extremely important for both the agricultural economy and the textile industry, cotton (*Gossypium* spp.) would probably be the most prominent example. It provides over 95% of the natural fiber output and builds the livelihood foundation for countless farmers and enterprises (Su et al., 2018; Thyssen et al., 2018). The continuous pursuit of high-quality fibers determines its commercial value and makes its position in the industrial chain unshakable (Sun et al., 2017). But things are not that simple. During the planting process, common diseases such as Fusarium wilt and Fusarium wilt often recur, not only affecting the yield but also directly lowering the quality of fibers, ultimately impacting profits and supply stability.

Although traditional breeding efforts have been considerable, most of them have focused on improving fiber quality or enhancing disease resistance, with few achievements that balance both (Cheng and Zhang, 2025). One of the reasons is that the genetic regulatory networks behind these two traits are rather complex, and sometimes there is a situation of "one rising and the other falling" (Wang et al., 2021). Especially in the context of constantly changing environments and evolving pathogenic bacteria, the improvement of single traits is becoming increasingly inadequate. In recent years, multitrait GWAS analysis and QTL mapping techniques have gradually demonstrated their advantages, beginning to reveal some pleomorphic loci and key candidate genes that not only affect fiber quality but also participate in disease resistance. This has opened up new ideas for molecular breeding and also made the practice of marker-assisted selection more targeted.

This study does not intend to focus merely on a single trait. On the contrary, we aim to systematically review the genetic research progress made over the years in the improvement of the dual traits of "fiber quality + disease resistance", especially the key achievements from multi-trait GWAS and QTL. In addition, the economic value and agronomic significance of such research will be briefly discussed to see how upland cotton and island cotton perform under different populations, environments and molecular means. It is hoped that it can provide some valuable references for the future cotton breeding strategies and research directions.

2 Genetic Basis of Fiber Quality and Disease Resistance in Cotton

2.1 Key quantitative traits associated with fiber strength, length, and fineness

When it comes to the quality of cotton, there are several core indicators that cannot be ignored: strength, length and fineness-especially fineness, which is often measured by the Macron value. Although these traits are also greatly influenced by the environment, their heritability is actually not low, which also means that there is some room for breeding improvement. At present, researchers have identified many genetic loci related to these traits through GWAS and QTL mapping. Even on chromosomes such as Dt11 and At07, some "hotspots" have been found, one tending towards length and the other associated with intensity (Said et al., 2013). A more detailed look shows that genes like Gh_D09G2376 and Gh_D06G1908 in these regions are considered to play a significant role in the development of fibers (Liu et al., 2020). Of course, relying solely on traditional drawing methods is not precise enough. The later introduced MAGIC and CSSL groups went a step further by uncovering the hidden gene interactions, such as which additive interactions or superior effects were influencing fibrous traits (Qi et al., 2024), thus making the research closer to the actual genetic background.

2.2 Major diseases impacting cotton yield

Yield and quality are often not technical issues but rather diseases that cause trouble first. Diseases like Fusarium wilt, Fusarium wilt and bacterial wilt have dealt a heavy blow to cotton fields. Among them, the most severe one might still be fusarium wilt. Not only will the yield decrease, but the fiber quality will also be affected (Wang et al., 2025). In recent years, researchers have screened out a number of genes related to resistance through GWAS and transcriptome methods, such as resistance sites like qVW-A01-2, as well as candidate genes like GhAMT2 and GhGT-3b_A04 (Figure 1). The mechanisms involved are not complicated either. It mainly focuses on lignin synthesis, salicylic acid signaling pathway and reactive oxygen species regulation (Mao et al., 2023). Both upland cotton and sea island cotton have actually made considerable efforts in this regard, especially in the selection of appropriate alleles or the introduction of dominant fragments from disease-resistant resources. Breeders have been making attempts (Li et al., 2023a). Although the process was not easy, the improvement in disease resistance has indeed shown initial results in some varieties.

2.3 Genetic overlap and independence between fiber and resistance traits

It is indeed not easy to achieve both disease resistance and fiber quality. Their genetic regulation is neither completely independent nor completely overlapping. In other words, some gene loci act on both traits, while others act on only one (Li et al., 2023b). For instance, in some regions of the D subgenome, there are both QTLS that improve fiber and signals that regulate wilt resistance, which shows the possibility of "balancing both ends" (Ma et al., 2021). However, real breeding cases also show that sometimes improving one trait can drag down another, such as chain burdens or genetic trade-offs, which are hard to avoid. Nowadays, multi-trait GWAS and integrated genome technologies are gradually beginning to identify key genes or chromosomal regions that have the potential to achieve bitrait improvement. This provides a new idea for cultivating cotton that is both high-quality and disease-resistant in the future. Ultimately, however, how these overlapping areas can achieve their maximum effectiveness still depends on further in-depth research.

3 Principles of Multi-Trait Genome-Wide Association Studies (MT-GWAS)

3.1 Differences between single-trait GWAS and MT-GWAS methodologies

Previous GWAS studies basically focused on only one trait at a time. Although this approach is intuitive, it is prone to missing some genes that play a role in multiple traits (Turley et al., 2018). Especially when dealing with complex traits influenced by multiple factors, the requirement for sample size becomes very high. Later, it was discovered that in fact many traits are genetically related. MT-GWAS was developed on this basis. It will analyze multiple related traits together and utilize the genetic correlations among them to enhance detection ability (Yoshida and Yanez, 2020). By this method, not only can the loci that were "missed by single trait GWAS" be identified, but also the common genetic mechanisms behind multiple traits can be revealed. However, not all traits are suitable for mixed analysis-provided that there must be a certain genetic correlation among these traits; otherwise, the advantages of MT-GWAS cannot be brought into play.

3.2 Statistical models used in MT-GWAS

MT-GWAS not only has to deal with multiple traits but also has to confront a vast amount of genotype data, so the statistical models behind it must be more powerful. At present, the most widely used model is the multilinear mixed model (mvLMM), because it can consider both fixed effects and random effects simultaneously, and also handle population structure problems and the relationships between traits (Lozano et al., 2023). Some researchers prefer Bayesian methods, such as Bayesian LASSO and EM-Bayesian LASSO. These models are more flexible in variable screening and suitable for high-dimensional data, especially in cases where the trait structure is relatively complex (Tamba et al., 2017; Wen et al., 2018). In recent years, some new models have been proposed, such as regularized versions of multi-trait hybrid models and multi-locus models. They have further enhanced computational efficiency and improved detection performance in certain scenarios.

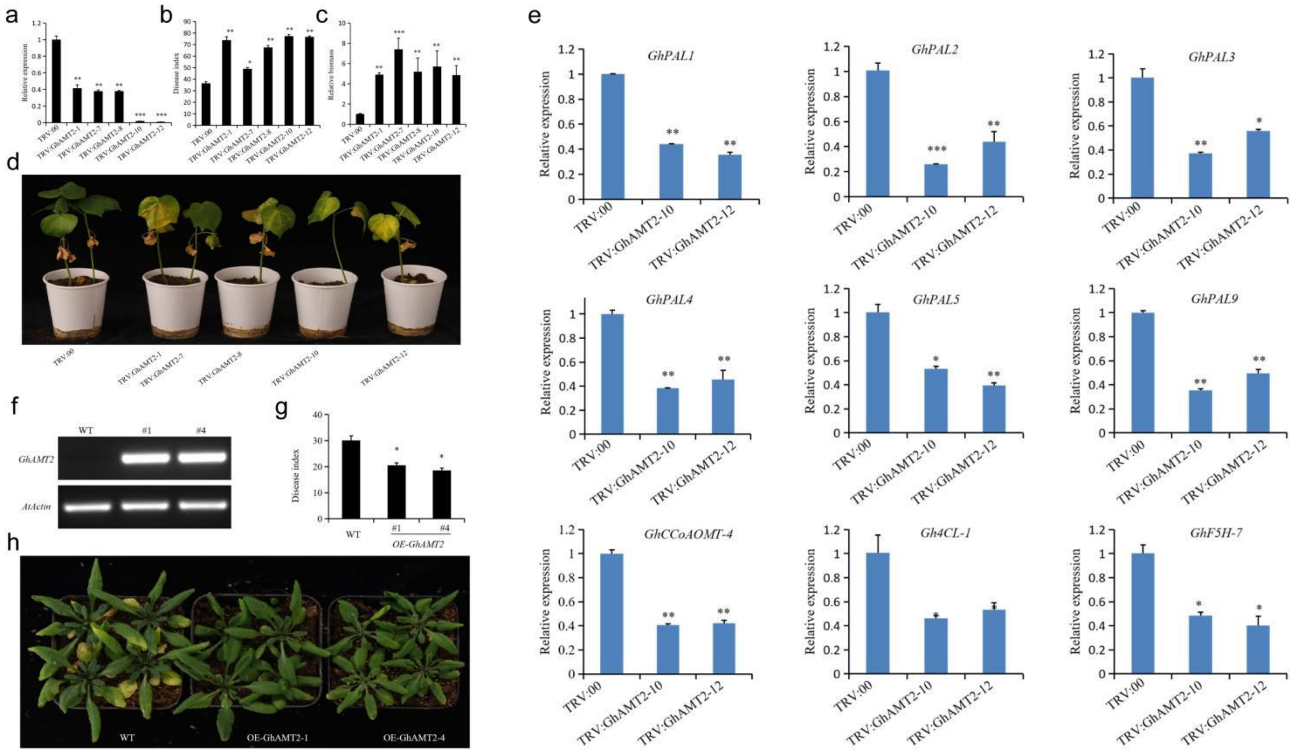

Figure 1 Functional Analysis of *GhAMT2* in enhancing resistance to *Verticillium Wilt* in Cotton and Arabidopsis. (a) Relative transcript levels of *GhAMT2* in TRV-VIGS cotton leaves for *TRV: EV* (empty vector) and *TRV: GhAMT2* plants, with *GhActin* serving as the internal reference gene. Expression levels were normalized to “1” for TRV: EV; (b) Silencing of *GhAMT2* reduces cotton resistance to *Verticillium dahliae*, as evidenced by increased disease symptoms on *TRV: GhAMT2* plants compared to *TRV: EV* plants 25 days post-infection; (c) Statistical analysis of the disease index in *GhAMT2*-silenced plants, showing a significant increase compared to controls; (d) Relative biomass of *V. dahliae* in *TRV: GhAMT2* and TRV: EV plants, illustrating higher pathogen colonization in *GhAMT2*-silenced plants; (e) *GhAMT2* promotes the expression of lignin metabolism-related genes in cotton. Expression analysis of lignin biosynthesis genes showed upregulation in the presence of *GhAMT2*; (f) Identification of transgenic Arabidopsis plants overexpressing *GhAMT2*. Semi-quantitative RT-PCR analysis demonstrates *GhAMT2* mRNA levels in three transgenic lines compared to WT (Col-0); (g) Overexpression of *GhAMT2* enhances Arabidopsis tolerance to *Verticillium dahliae*, as evidenced by reduced disease symptoms in transgenic lines compared to WT; (h) Statistical analysis of the disease index in WT and OE-*GhAMT2* Arabidopsis plants, showing significantly lower disease indices in transgenic lines. *, $0.01 < P < 0.05$; **, $P < 0.01$; ***, $P < 0.001$ (Adopted from Wang et al., 2025)

3.3 Advantages of MT-GWAS in revealing pleiotropic loci and genetic correlations

Compared with the traditional single-trait GWAS, the benefits of MT-GWAS are obvious. As long as there is a moderate or higher genetic correlation among the analyzed traits, this method can often detect the related genes more easily (Guo et al., 2024). It is particularly good at identifying the pleiotropic loci where "one gene affects multiple traits simultaneously", which is an important entry point for understanding the common genetic basis of complex traits (Sahito et al., 2024). Not only that, MT-GWAS can also quantify the genetic relationships among different traits, thereby providing clues for subsequent breeding directions, such as which traits can be improved together and which may conflict with each other. However, it should also be noted that it is, after all, just one of the tools. Whether to adopt it or not still depends on the research objective and the characteristics of the trait itself.

4 Key Discoveries in Multi-Trait GWAS for Cotton

4.1 Identification of pleiotropic QTLs controlling both fiber traits and disease resistance

Some genetic loci seem to control more than one aspect. Past multi-trait GWAS and multi-environment analyses have shown that there are many QTLS that can simultaneously regulate multiple fiber quality traits (such as length, strength, and fineness) and yield traits (Gowda et al., 2024). Some of them are even linked to production factors such as the weight of the bells, the proportion of the clothes, and the number of bells-indicating that these areas have a strong sense of "taking into account". Interestingly, these QTLS have been repeatedly detected in multiple environments, and this "stability" is highly practical for breeders (Su et al., 2020). Although most current studies focus more on the two traits of yield and fiber, it can still be observed that some QTLS are also associated with disease resistance. In other words, this kind of genetic overlap may become a breakthrough point for "simultaneous improvement of dual traits". However, it should also be noted that these QTLS do not perform consistently in all environments, and some effects are also related to conditions. This point cannot be ignored.

4.2 Novel candidate genes uncovered through MT-GWAS and their functional annotations

Not all candidate genes can be easily identified by traditional methods, especially those whose functions are "hidden behind multiple traits". Using methods such as MT-GWAS, researchers screened out some candidate genes that play a key role in fiber development and yield expression. Genes such as GH_D06G2161, which begin to be expressed in the early ovulate stage, and Ghir_D11G020400, which encodes MATE exovation protein and is related to fiber elongation, are both regarded as relatively core regulators (Wang et al., 2022). Functional annotations and transcriptome data provide more context, indicating that most of these genes are involved in cell wall synthesis, signal transduction, and developmental processes directly linked to fibrogenesis (Li et al., 2020). These new discoveries may not be immediately applicable, but they undoubtedly provide many new targets for marker-assisted selection and subsequent biotechnological interventions.

4.3 Insights into genotype-by-environment interactions from multi-environment studies

Just because a gene is stable in the laboratory doesn't mean it works equally well in the field. The research results of multi-environment GWAS remind us that the expression of QTL and candidate genes is actually greatly influenced by the environment, and sometimes this influence is quite "jumping" (Zhu et al., 2021). Some sites are very stable, and the effect is quite consistent no matter where they are planted. However, there are also many QTLS that only show obvious signs in certain regions or years, which makes it necessary to pay attention to the interaction between genes and the environment. For instance, for some loci related to fiber quality, the test results will completely change when the location is changed or the year is changed. This also explains why in the breeding process, making decisions based solely on data from a certain location may not be sufficient. Multi-environment verification is not an option but a necessary prerequisite.

5 Integration with Other Omics and Functional Genomics

5.1 Combining MT-GWAS with transcriptomics and metabolomics for gene prioritization

It's hard to tell which gene is truly the "key player" just by relying on GWAS. At this point, if the results of MT-GWAS are combined with the data of transcriptome and metabolome, it will be clearer to determine which candidate genes are worth giving priority to. For instance, if a certain gene is not only genetically related to fiber quality or disease resistance, but also active in transcriptional expression and even involved in a certain metabolic pathway, then it is not "merely related", but "may actually be involved" (Sanches et al., 2024). Through co-expression analysis, metabolic pathway enrichment, and network-based integration methods, researchers can identify truly functional genes from multiple perspectives, not just the "possibilities" in a statistical sense.

Although this method is a bit troublesome, it is much more reliable than relying solely on GWAS. Integrating multi-omics is not merely for the sake of good-looking data, but for more accurately selecting targets and biomarkers that can be used for breeding.

5.2 Validation of candidate genes using CRISPR, RNAi, and overexpression studies

Discovering the candidate gene is only the first step; verification is the key. Whether it's the quality of the fibers or their disease resistance, guessing is of no use. It depends on what the experiments say. Gene editing technologies like CRISPR/Cas9, or methods such as RNA interference (RNAi) and gene overexpression, are the "toolboxes" for verification. By knocking out, silencing or enhancing the expression of specific genes, researchers can see whether these genes are involved in certain biological processes, such as fiber development or disease resistance response (Yang et al., 2021). This kind of verification not only confirms the function of genes, but also turns "correlation" into "causality". To put it bluntly, it's about letting experiments help us put those ambiguous gene loci into practice and truly turn them into practical achievements for breeding.

5.3 Epigenomic and proteomic contributions to trait expression

Not all trait differences can be answered from DNA variations. Sometimes, epigenetic modifications and protein changes are the key. For instance, epigenetic regulations such as DNA methylation or histone modification adjust gene expression in response to environmental changes. Sometimes, the impact of such regulation on disease resistance or fiber quality is even more significant than that of the sequence itself (Morabito et al., 2025). Proteomics, on the other hand, approaches it from a different perspective-it does not focus on how much gene is expressed, but rather on how the final product (protein) performs, whether modification occurs, and whether there is interaction. If this information can be integrated and analyzed together with MT-GWAS, the genetics, epigenetics and protein regulation that affect cotton traits can be pieced together into a more complete map at the "system level". This step is indispensable for understanding the formation mechanism of complex traits.

6 Case Study

6.1 Case overview: a breeding program targeting both superior fiber and wilt resistance

When it comes to the true application of MT-GWAS in breeding, the domestic *Gossypium barbadense* project is a typical example. However, the initial difficulty of this project lies in the fact that there is more than one goal: not only to improve the quality of fibers, such as the indicators of length, strength and fabric fraction, but also to enhance the resistance to Fuswilt disease. Once this disease breaks out, it will have a great impact on both yield and quality. Traditional methods are often inefficient in dealing with such complex targets. The strategy adopted by this project is to first conduct large-scale resequencing and phenotypic analysis of different germplasms, and then further explore the genetic patterns on this basis. Ultimately, it is about identifying the genes behind complex traits and clarifying the relationships.

6.2 Deployment of MT-GWAS to detect co-localized QTLs for fiber and disease traits

There are actually quite a few traits analyzed in this project-a total of 15, including those related to fibers, diseases, and yield traits. The final screened gene loci were most related to fiber quality, followed by disease resistance and yield (Figure 2) (Zhao et al., 2022). One of the key breakthroughs is that MT-GWAS helped discover those "co-localized QTLS" that have an impact on multiple traits, meaning that the same gene region may simultaneously be responsible for both fiber and disease resistance. Such multi-potency loci are easily overlooked in traditional single-trait analysis. More importantly, some genes are not determined solely by statistics. Researchers also used gene expression data and transgenic verification methods (such as VIGS) to gradually "confirm" the functions of these candidate genes.

6.3 Outcomes: candidate genes validated, germplasm selected, and yield improvements observed

The final outcome of this breeding project is also quite clear: five key candidate genes have been functionally verified, involving not just a single trait, but multiple aspects such as disease resistance, fiber length, strength and garment size. With these achievements, subsequent lineage selection and marker-assisted breeding will have a "clear coordinate". The newly bred germplasm not only meets the fiber quality standards but also has significantly stronger disease resistance than before. The more practical benefit is that the output has gradually increased, and the overall breeding effect is closer to the market and agronomic demands.

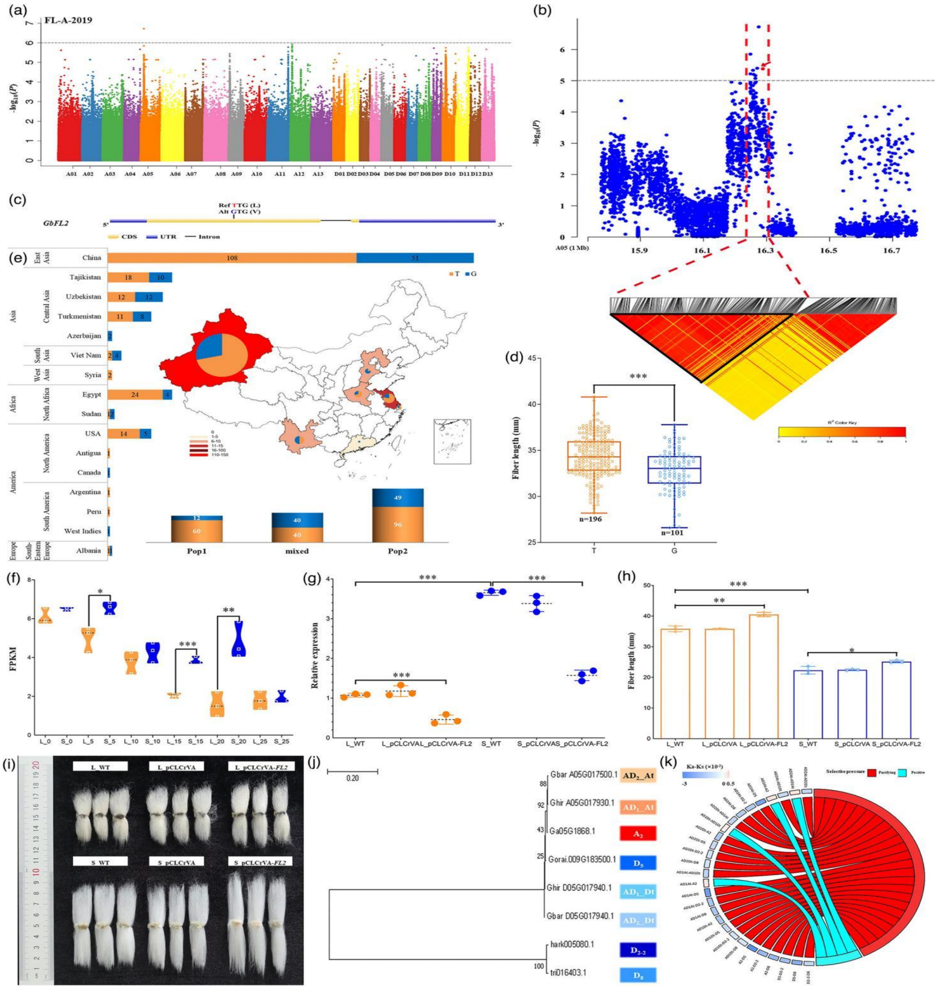

Figure 2 Identification of the FL causal gene *GbFL2* on chromosome At05. (a) Manhattan plot for FL. The dashed line represents the significance threshold (-log10$P = 6$). Effect values of genetic markers were tested using *F* tests and corrected for multiple testing using Bonferroni correction. (b) Local Manhattan plot (top) and LD heatmap (bottom) surrounding the peak on At05. The dashed line represents the significance threshold (-log10$P = 5$). Red arrows mark the position of the nonsynonymous SNP A05_16286973 in *Gbar_A05G017500* (*GbFL2*). Red dotted lines show the candidate region. (c) Structure of *GbFL2*. Blue and yellow rectangles mark UTR and CDS respectively. (d) Box plot for FL, based on the haplotypes of the two SNPs. In the box plots, the centre line indicates the median. Box limits are the upper and lower quartiles, and whiskers mark the range of the data; *n* denotes the number of accessions with the same genotype. We used a two-tailed *t*-test to perform the significance analysis. (e) Haplotype distribution in diverse geographical regions and subpopulations. The bar chart on the left shows the number of two haplotypes in distinct countries. The map in the middle displays the ratio of two haplotypes in the different provinces of China. The column diagrams below represent the number of two haplotypes in different subpopulations. (f) Expression of *GbFL2* in long-fibre accession XH58 (haplotype T) and short-fibre accession Ashi (haplotype G) at the fibre developmental stages (0, 5, 10, 15, 20, 25 DPA), detected by RNA-seq (FPKM value). Data are average values with standard deviation ($n = 3$ varieties with three technical repeats). Single (*), double (**) and triple (***) asterisks mark statistical significance levels of $P < 0.05$, 0.01 and 0.001 respectively. (g) qRT-PCR analysis of *GbFL2* expression in wild-type (WT), transgenic lines with empty VIGS vector (pCLCrVA) and target gene *GbFL2* (pCLCrVA-*FL2*) of long-fibre (L) accession XH58 and short-fibre (S) accession Ashi. The gene expression level in the long-fibre accession wild type (L_WT) was set to 1. *GbUBQ7* is an internal control. (h) Fibre length (mm) of WT, pCLCrVA, (pCLCrVA-*FL2* of long-fibre (L) accession XH58 and short-fibre (S) accession Ashi. (i) VIGS phenotypes of *GbFL2*. (j) The evolutionary origin of *GbFL2* (*Gbar_A05G017500*). We built unrooted trees using the maximum-likelihood method in MEGA7, based on complete CDS sequences. (k) Selection analysis on homologous CDS sequences of *GbFL2*. Homologous sequence in each cotton species is represented by its genome name on the left side of the circle. The difference value (Ka-Ks) of each group of homologous comparisons is indicated by coloured rectangles according to the colour bar in the upper left corner. While Ka/Ks is generally used as an indicator of selective pressure, the presence of 'Ks = 0' here precludes this; therefore, we chose another indicator, that is 'Ka-Ks' value. Two types of selection effects, purifying selection (in red) and positive selection (in blue), are shown on the right side of the circle (Adopted from Zhao et al., 2022)

7 Challenges and Limitations in Multi-Trait GWAS Applications

7.1 Population structure, environmental noise, and phenotype accuracy

When analyzed by MT-GWAS, the interference brought by the population structure is often more serious than expected. Especially when the sample itself is unbalanced, population stratification is prone to false positives, especially for data of non-European ancestry or mixed groups, the probability of error is even higher (Troubat et al., 2024). On the other hand, the environment is not an easy factor either. Even a slight fluctuation in climate or human error in field experiments may cause the measurement of traits to deviate from the right direction. Especially in crop research, environmental effects sometimes even outweigh genetic signals, making the estimation of heritability less accurate (Zhang et al., 2019). It would be best if phenotypic data could be accurate and uniform across different environments and groups. However, in practice, it often encounters limitations in terms of resources, human resources and technology, making it difficult to truly achieve large-scale implementation.

7.2 Statistical limitations: power, computational complexity, and false discovery rate

Theoretically speaking, MT-GWAS can indeed detect signals more easily than the single-trait method. However, the problem lies in that it is overly data-dependent-the correlation between traits, the genetic background of the population, and the sample size-if these variables do not cooperate properly, the detection capability will immediately decline (Suzuki et al., 2024). To find small effect sites or pleiotropic regions, the sample size needs to be astonishingly large. However, in actual crop research, samples of this scale are not common. Moreover, once the number of traits and SNPS is too large, the pressure on model calculation also increases accordingly. Multicollinearity often makes the results difficult to interpret and may also slow down the analysis efficiency (Porter and O 'Reilly, 2017). Let's talk about multiple tests. Although corrections like Bonferroni are rigorous, they can also easily "overfilter out" some genuine signals. Especially when the experimental error is large, how to balance the detection efficacy and control false positives has become a dilemma.

7.3 Translational gaps between discovery and deployment in breeding programs

Even if you identify a bunch of loci, there might not be that many that can actually be put to use. This is a common gap in MT-GWAS applications. Some sites have inherently small effects or are highly sensitive to the environment, resulting in unstable performance in the field. As a result, marker-assisted selection is unlikely to have practical significance. Another rather awkward aspect is that many candidate genes lack functional verification. The combination with other omics data is also often absent (Zhu and Luo, 2024). As a result, you have statistical associations but cannot quickly turn them into practical tools (Khatiwada et al., 2023). Traits themselves are complex enough. Coupled with the interaction between genes and the environment, even beneficial alleles may have vastly different expression results in different locations and years. Therefore, this transformation from "discovery" to "usability" remains the most challenging nut to crack in current breeding.

8 Future Directions and Concluding Remarks

Whether MT-GWAS can truly play a greater role in the future ultimately depends on whether the phenotypic data can keep up. Manual measurement alone is no longer sufficient. Technologies such as high-throughput imaging, remote sensing, and automated sensors are increasingly being used to collect field traits-not only are they highly efficient, but they also minimize human errors and enhance the accuracy of the data. However, data alone is not enough. If there is no appropriate algorithm to parse, no matter how much data there is, it will just be piled up. Nowadays, some studies have begun to attempt to incorporate deep learning and artificial intelligence methods to explore the non-linear relationships between traits that are not so easily detectable. This is very helpful for identifying QTLS with small effects and complex genetic backgrounds.

Returning to the genomic level, the traditional approach of only looking at a single reference genome has become increasingly difficult to meet the demands. Pan-genome and haplotype analyses are increasingly being adopted by more people. They can not only reveal previously overlooked structural variations and rare alleles, but also enable more precise localization, especially for those variations related to pleiotropic traits. Especially in some germplasm resources with relatively high diversity, the "supplementary information" provided by these methods may change the way the entire genetic map is understood. That is to say, what was invisible before might be due to the fact that the tools we used were too limited.

Another direction worth noting is to directly feed the results of MT-GWAS into the genomic selection (GS) process. This is actually like adding a "multi-trait brain" to predictive breeding. By integrating the identified pleiotropic loci and the genetic correlations between traits, the breeding model can optimize multiple objectives at once-such as taking into account fiber quality, disease resistance and yield. Once such models are established, both the breeding efficiency and the speed of genetic gain will be significantly enhanced. Moreover, this is also expected to shorten the "gap" between gene discovery and practical application.

Acknowledgments
The authors express their gratitude to Mr. Qi for providing valuable feedback that improved the clarity of the text.

Conflict of Interest Disclosure
The authors affirm that this research was conducted without any commercial or financial relationships that could be construed as a potential conflict of interest.

References

Cheng J.H., and Zhang J., 2025, High-yield cotton cultivation practices in arid regions, Molecular Soil Biology, 16(1): 27-36.
https://doi.org/10.5376/msb.2025.16.0003

Gowda S., Fang H., Tyagi P., Bourland F., Dever J., Campbell B., Zhang J., Abdelraheem A., Sood S., Jones D., and Kuraparthy V., 2024, Genome-wide association study of fiber quality traits in US upland cotton (*Gossypium hirsutum* L.), Theoretical and Applied Genetics, 137(9): 214.
https://doi.org/10.1007/s00122-024-04717-7

Guo H., Li T., Shi Y., and Wang X., 2024, MTML: an efficient multitrait multilocus GWAS method based on the Cauchy combination test, Biometrical Journal, 66(6): e202300130.
https://doi.org/10.1002/bimj.202300130

Khatiwada A., Yilmaz A., Wolf B., Pietrzak M., and Chung D., 2023, Multi-GPA-Tree: statistical approach for pleiotropy informed and functional annotation tree guided prioritization of GWAS results, PLoS Computational Biology, 19(12): e1011686.
https://doi.org/10.1371/journal.pcbi.1011686

Li Y., Si Z., Wang G., Shi Z., Chen J., Qi G., Jin S., Han Z., Gao W., Tian Y., Mao Y., Fang L., Hu Y., Chen H., Zhu X., and Zhang T., 2023a, Genomic insights into the genetic basis of cotton breeding in China, Molecular Plant, 16(4): 662-677.
https://doi.org/10.1016/j.molp.2023.01.012

Li Y., Zhang X., Lin Z., Zhu Q., Li Y., Xue F., Cheng S., Feng H., Sun J., and Liu F., 2023b, Comparative transcriptome analysis of interspecific CSSLs reveals candidate genes and pathways involved in verticillium wilt resistance in cotton (*Gossypium hirsutum* L.), Industrial Crops and Products, 197: 116560.
https://doi.org/10.1016/j.indcrop.2023.116560

Li Z., Wang P., You C., Yu J., Zhang X., Yan F., Ye Z., Shen C., Li B., Guo K., Liu N., Thyssen G., Fang D., Lindsey K., Zhang X., Wang M., and Tu L., 2020, Combined GWAS and eQTL analysis uncovers a genetic regulatory network orchestrating the initiation of secondary cell wall development in cotton, New Phytologist, 226(6): 1738-1752.
https://doi.org/10.1111/nph.16468

Liu W., Song C., Ren Z., Zhang Z., Pei X., Liu Y., He K., Zhang F., Zhao J., Zhang J., Wang X., Yang D., and Li W., 2020, Genome-wide association study reveals the genetic basis of fiber quality traits in upland cotton (*Gossypium hirsutum* L.), BMC Plant Biology, 20(1): 395.
https://doi.org/10.1186/s12870-020-02611-0

Lozano A., Ding H., Abe N., and Lipka A., 2023, Regularized multi-trait multi-locus linear mixed models for genome-wide association studies and genomic selection in crops, BMC Bioinformatics, 24(1): 399.
https://doi.org/10.1186/s12859-023-05519-2

Ma Z., Zhang Y., Wu L., Zhang G., Sun Z., Li Z., Jiang Y., Ke H., Chen B., Liu Z., Gu Q., Wang Z., Wang G., Yang J., Wu J., Yan Y., Meng C., Li L., Li X., Mo S., Wu N., Chen L., Zhang M., Si A., Yang Z., Wang N., Wu L., Zhang D., Cui Y., Cui J., Lü X., Li Y., Shi R., Duan Y., Tian S., and Wang X., 2021, High-quality genome assembly and resequencing of modern cotton cultivars provide resources for crop improvement, Nature Genetics, 53(9): 1385-1391.
https://doi.org/10.1038/s41588-021-00910-2

Mao H., Zhang W., Lü J., Yang J., Yang S., Jia B., Song J., Wu M., Pei W., Zhang B., Zhang J., Wang L., and Yu J., 2023, Overexpression of cotton Trihelix transcription factor *GhGT-3b_A04* enhances resistance to *Verticillium dahliae* and affects plant growth in *Arabidopsis thaliana*, Journal of Plant Physiology, 283: 153947.
https://doi.org/10.1016/j.jplph.2023.153947

Morabito A., De Simone G., Pastorelli R., Brunelli L., and Ferrario M., 2025, Algorithms and tools for data-driven omics integration to achieve multilayer biological insights: a narrative review, Journal of Translational Medicine, 23(1): 425.
https://doi.org/10.1186/s12967-025-06446-x

Porter H., and O'Reilly P., 2017, Multivariate simulation framework reveals performance of multi-trait GWAS methods, Scientific Reports, 7(1): 38837.
https://doi.org/10.1038/srep38837

Qi G., Si Z., Xuan L., Han Z., Hu Y., Fang L., Dai F., and Zhang T., 2024, Unravelling the genetic basis and regulation networks related to fibre quality improvement using chromosome segment substitution lines in cotton, Plant Biotechnology Journal, 22(11): 3135-3150.
https://doi.org/10.1111/pbi.14436

Sahito J., Zhang H., Gishkori Z., Ma C., Wang Z., Ding D., Zhang X., and Tang J., 2024, Advancements and prospects of genome-wide association studies (GWAS) in maize, International Journal of Molecular Sciences, 25(3): 1918.
https://doi.org/10.3390/ijms25031918

Said J., Lin Z., Zhang X., Song M., and Zhang J., 2013, A comprehensive meta QTL analysis for fiber quality, yield, yield-related and morphological traits, drought tolerance, and disease resistance in tetraploid cotton, BMC Genomics, 14(1): 776.
https://doi.org/10.1186/1471-2164-14-776

Sanches P., De Melo N., Porcari A., and De Carvalho L., 2024, Integrating molecular perspectives: strategies for comprehensive multi-omics integrative data analysis and machine learning applications in transcriptomics, proteomics, and metabolomics, Biology, 13(11): 848.
https://doi.org/10.3390/biology13110848

Su J., Ma Q., Li M., Hao F., and Wang C., 2018, Multi-locus genome-wide association studies of fiber-quality related traits in Chinese early-maturity upland cotton, Frontiers in Plant Science, 9: 1169.
https://doi.org/10.3389/fpls.2018.01169

Su J., Wang C., Ma Q., Zhang A., Shi C., Liu J., Zhang X., Yang D., and Ma X., 2020, An RTM-GWAS procedure reveals the QTL alleles and candidate genes for three yield-related traits in upland cotton, BMC Plant Biology, 20(1): 416.
https://doi.org/10.1186/s12870-020-02613-y

Sun Z., Wang X., Liu Z., Gu Q., Zhang Y., Li Z., Ke H., Yang J., Wu J., Wu L., Zhang G., Zhang C., and Ma Z., 2017, Genome-wide association study discovered genetic variation and candidate genes of fibre quality traits in *Gossypium hirsutum* L., Plant Biotechnology Journal, 15(8): 982-996.
https://doi.org/10.1111/pbi.12693

Suzuki Y., Ménager H., Brancotte B., Vernet R., Nerin C., Boetto C., Auvergne A., Linhard C., Torchet R., Lechat P., Troubat L., Cho M., Bouzigon E., Aschard H., and Julienne H., 2024, Trait selection strategy in multi-trait GWAS: boosting SNP discoverability, Human Genetics and Genomics Advances, 5(3): 100319.
https://doi.org/10.1016/j.xhgg.2024.100319

Tamba C., Ni Y., and Zhang Y., 2017, Iterative sure independence screening EM-Bayesian LASSO algorithm for multi-locus genome-wide association studies, PLoS Computational Biology, 13(1): e1005357.
https://doi.org/10.1371/journal.pcbi.1005357

Thyssen G., Jenkins J., McCarty J., Zeng L., Campbell B., Delhom C., Islam M., Li P., Jones D., Condon B., and Fang D., 2018, Whole genome sequencing of a MAGIC population identified genomic loci and candidate genes for major fiber quality traits in upland cotton (*Gossypium hirsutum* L.), Theoretical and Applied Genetics, 132(4): 989-999.
https://doi.org/10.1007/s00122-018-3254-8

Troubat L., Fettahoglu D., Henches L., Aschard H., and Julienne H., 2024, Multi-trait GWAS for diverse ancestries: mapping the knowledge gap, BMC Genomics, 25(1): 375.
https://doi.org/10.1101/2023.06.23.546248

Turley P., Walters R., Maghzian O., Okbay A., Lee J., Fontana M., Nguyen-Viet T., Wedow R., Zacher M., Furlotte N., Agee M., Alipanahi B., Auton A., Bell R., Bryc K., Elson S., Fontanillas P., Hinds D., Hromatka B., Huber K., Kleinman A., Litterman N., McIntyre M., Mountain J., Northover C., Sathirapongsasuti J., Sazonova O., Shelton J., Shringarpure S., Tian C., Tung J., Vacic V., Wilson C., Pitts S., Magnusson P., Oskarsson S., Johannesson M., Visscher P., Laibson D., Cesarini D., Neale B., and Benjamin D., 2018, Multi-trait analysis of genome-wide association summary statistics using MTAG, Nature Genetics, 50(2): 229-237.
https://doi.org/10.1038/s41588-017-0009-4

Wang L., Yang Y., Qin J., Ma Q., Qiao K., Fan S., and Qu Y., 2025, Integrative GWAS and transcriptomics reveal *GhAMT2* as a key regulator of cotton resistance to *Verticillium wilt*, Frontiers in Plant Science, 16: 1563466.
https://doi.org/10.3389/fpls.2025.1563466

Wang M., Qi Z., Thyssen G., Naoumkina M., Jenkins J., McCarty J., Xiao Y., Li J., Zhang X., and Fang D., 2022, Genomic interrogation of a MAGIC population highlights genetic factors controlling fiber quality traits in cotton, Communications Biology, 5(1): 60.
https://doi.org/10.1038/s42003-022-03022-7

Wang P., He S., Sun G., Pan Z., Sun J., Geng X., Peng Z., Gong W., Wang L., Pang B., Jia Y., and Du X., 2021, Favorable pleiotropic loci for fiber yield and quality in upland cotton (*Gossypium hirsutum*), Scientific Reports, 11(1): 15935.
https://doi.org/10.1038/s41598-021-95629-9

Wen Y., Zhang H., Ni Y., Huang B., Zhang J., Feng J., Wang S., Dunwell J., Zhang Y., and Wu R., 2018, Methodological implementation of mixed linear models in multi-locus genome-wide association studies, Briefings in Bioinformatics, 19(4): 700-712.
https://doi.org/10.1093/bib/bbw145

Yang Y., Saand M., Huang L., Abdelaal W., Zhang J., Wu Y., Li J., Sirohi M., and Wang F., 2021, Applications of multi-omics technologies for crop improvement, Frontiers in Plant Science, 12: 563953.
https://doi.org/10.3389/fpls.2021.563953

Yoshida G., and Yáñez J., 2020, Multi-trait GWAS using imputed high-density genotypes from whole-genome sequencing identifies genes associated with body traits in Nile tilapia, BMC Genomics, 22(1): 57.
https://doi.org/10.1186/s12864-020-07341-z

Zhang Y., Jia Z., and Dunwell J., 2019, The applications of new multi-locus GWAS methodologies in the genetic dissection of complex traits, Frontiers in Plant Science, 10: 100.
https://doi.org/10.3389/fpls.2019.00100

Zhao N., Wang W., Grover C., Jiang K., Pan Z., Guo B., Zhu J., Su Y., Wang M., Nie H., Xiao L., Guo A., Yang J., Cheng C., Ning X., Li B., Xu H., Iyaah D., Alifu A., Li P., Geng J., Wendel J., Kong J., and Hua J., 2022, Genomic and GWAS analyses demonstrate phylogenomic relationships of Gossypium barbadense in China and selection for fibre length, lint percentage and Fusarium wilt resistance, Plant Biotechnology Journal, 20(4): 691-710.
https://doi.org/10.1111/pbi.13747

Zhu G., Hou S., Song X., Wang X., Wang W., Chen Q., and Guo W., 2021, Genome-wide association analysis reveals quantitative trait loci and candidate genes involved in yield components under multiple field environments in cotton (*Gossypium hirsutum*), BMC Plant Biology, 21(1): 250.
https://doi.org/10.1186/s12870-021-03009-2

Zhu S.J., and Luo M.T., 2024, Advancements in pest management techniques for cotton crops, Bioscience Methods, 15(4): 196-206.
https://doi.org/10.5376/bm.2024.15.0020