



OPEN ACCESS

EDITED BY

Rodomiro Ortiz,
Swedish University of Agricultural
Sciences, Sweden

REVIEWED BY

Zhengwen Sun,
Hebei Agricultural University, China
Zunaira Anwar,
Nuclear Institute for Agriculture and
Biology, Pakistan

*CORRESPONDENCE

Ibrokhim Y. Abdurakhmonov
✉ i.y.abdurakhmonov@gmail.com;
✉ ibrokhimabdurakhmonov@
genomics.uz

RECEIVED 09 March 2026

REVISED 20 April 2026

ACCEPTED 21 April 2026

PUBLISHED 12 May 2026

CITATION

Abdurakhmonov IY (2026)
Cotton genetic mapping for plant
biotechnology: from markers to graph
pan-genomes and sustainable breeding.
Front. Plant Sci. 17:1825852.
doi: 10.3389/fpls.2026.1825852

COPYRIGHT

© 2026 Abdurakhmonov. This is an open-
access article distributed under the terms
of the [Creative Commons Attribution
License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use, distribution or
reproduction in other forums is
permitted, provided the original
author(s) and the copyright owner(s) are
credited and that the original publication
in this journal is cited, in accordance
with accepted academic practice. No
use, distribution or reproduction is
permitted which does not comply with
these terms.

Cotton genetic mapping for plant biotechnology: from markers to graph pan-genomes and sustainable breeding

Ibrokhim Y. Abdurakhmonov*

Center of Genomics and Bioinformatics, Academy of Sciences of Uzbekistan, Tashkent, Uzbekistan

Cotton improvement remains a major challenge in plant biotechnology because stable yield, fiber quality, and stress resilience must be delivered from complex allopolyploid genomes under variable environmental conditions. Genetic mapping has progressively transformed this challenge into deployable breeding knowledge, advancing from sparse marker systems to high-density SNP arrays, reference genomes, and, more recently, structural-variation-aware pan-genomes and graph genomes. This review argues that cotton's genomic complexity, homoeolog redundancy, and strong genotype-by-environment interactions have driven methodological innovation rather than the simple transfer of approaches developed in other crops. It synthesizes advances in linkage mapping, GWAS, eQTL analysis, fine mapping, functional validation, and marker deployment, with emphasis on how these tools have enabled breeder-ready assays, marker-assisted selection, and emerging predictive breeding frameworks. It further examines how cotton mapping contributes to sustainable intensification by enhancing disease resistance, abiotic resilience, fiber value, and trait-stacking efficiency. Finally, this review highlights unresolved challenges in homoeolog-aware inference, structural variant genotyping, phenotyping throughput, and environment-aware prediction. It outlines a next-phase agenda in which graph genomes, scalable validation, and climate-informed models become central to cotton biotechnology.

KEYWORDS

cotton, genetic mapping, GWAS, marker-assisted selection, pan-genome, plant biotechnology, polyploidy

1 Introduction

Cotton is a globally important fiber and oilseed crop whose improvement remains unusually challenging because economically important traits must be optimized within complex allotetraploid genomes under variable environments. Its two dominant cultivated species, *Gossypium hirsutum* and *G. barbadense*, are allotetraploids (AADD, $2n = 4x = 52$) that formed approximately 1–2 million years ago through hybridization between A- and D-genome diploids (Jiang et al., 1998). The crop's defining commercial trait, spinnable fiber, arises from single-celled epidermal trichomes that follow a developmental program that is both robust and highly sensitive to environmental conditions. Breeders have long exploited visible variation in lint yield, fiber length, strength, and micronaire, but converting these phenotypes into reliable genetic decisions has remained difficult.

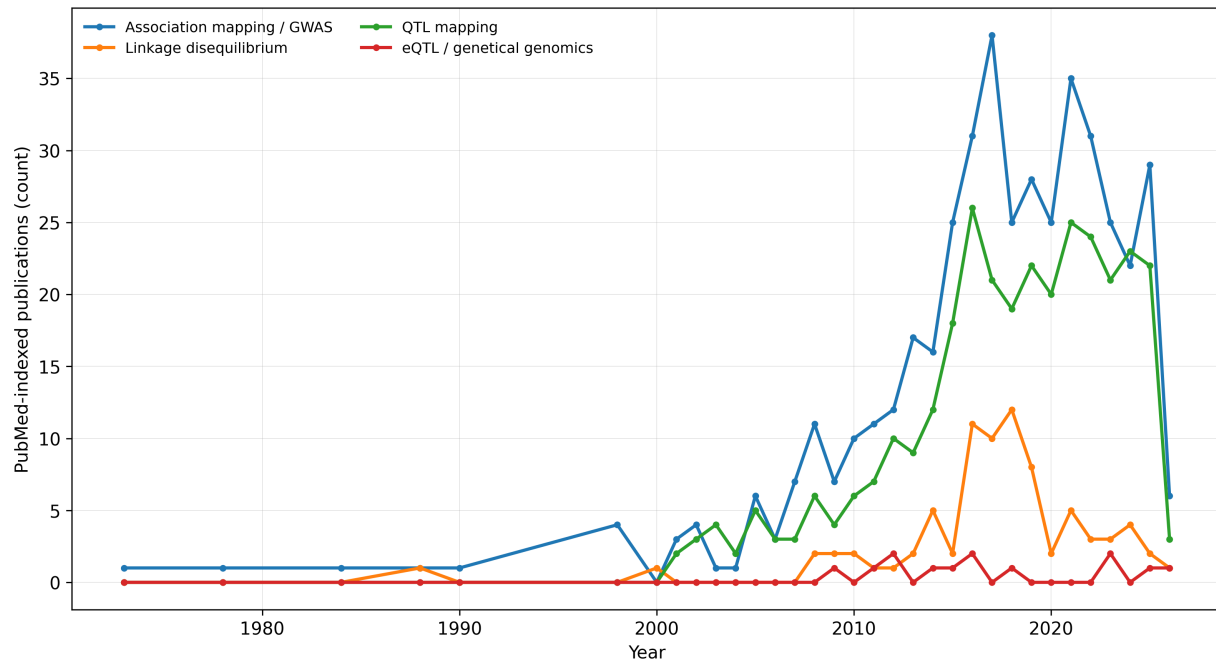


FIGURE 1

Cotton genetic mapping literature over time. PubMed-indexed publication counts per year for cotton QTL mapping, association mapping/GWAS, eQTL/genetical genomics, and linkage disequilibrium studies. The curves visualize phase changes: the QTL era, the rise of population-structure-aware association, and the emergence of genomics-enabled regulatory and pan-genome frameworks.

Polyloid genomes complicate mapping by introducing homoeologous loci that can confound marker assignment while still contributing distinct trait effects. Early work by Jiang et al. (1998) showed that polyploid formation created new opportunities for response to selection, not simply through gene duplication, but through regulatory redundancy that selection could subsequently repartition. At the same time, fiber traits are often determined within narrow developmental windows, whereas field performance reflects season-long interactions among genotype, stress, and management. This combination made cotton a stringent test case for genetic mapping. In response, the field progressed through successive improvements in tools and study design to increase resolution, robustness, and reproducibility.

This review focuses on cotton-specific constraints, particularly polyploid-aware inference, field-relevant reproducibility across environments, and the validation-to-deployment pipeline that converts loci into breeder-ready assays and cultivars. In doing so, it highlights how cotton differs from broader plant-wide mapping trajectories in its constraints, resources, and routes to breeding application.

Two organizing frameworks anchor this review. First, publication chronologies show that cotton mapping has shifted from linkage and QTL studies to population-scale association and regulatory mapping, and more recently to structural-variation and pan-genome frameworks (Figure 1). Second, progress in cotton mapping has consistently depended on aligning technological advances with breeder-relevant trait priorities and deployment pathways. The major milestones in this five-decade trajectory are summarized in Table 1, which provides a roadmap for the sections that follow.

2 Classical genetics before dense markers

Before the advent of dense molecular markers, cotton genetics during the 1970s–1990s was shaped by two central realities. First, many agronomic traits were clearly heritable, but heritability alone was not actionable at locus-level resolution. Second, polyploidy both expanded and obscured variation: when multiple homoeologous loci contributed to a trait, classical segregation analysis often provided only limited guidance on where selection should act.

A representative pre-marker trait was seed gossypol, whose inheritance could be analyzed quantitatively even within early genetic frameworks (Lee, 1973). However, this insight did not easily translate into breeding pipelines because the relevant loci could not yet be identified or tracked. This gap between recognizing that variation was genetic and being unable to identify the responsible chromosomal segments defined the central limitation of the era.

Cotton's evolutionary biology offered a deeper explanation for this difficulty. Polyploid formation could create new avenues for response to selection, not simply by duplicating genes, but by generating regulatory and functional redundancy that selection could later repartition (Jiang et al., 1998). Even so, conceptual understanding alone was insufficient. By the late twentieth century, the path forward had become clear: the field needed polymorphic markers, linkage maps, and population designs capable of converting polyploid complexity into interpretable recombination patterns.

TABLE 1 Milestones in cotton genetic mapping, framed as challenges and solutions.

Era	Period	Core challenge	Mapping innovation	Cotton-specific outcome	Key references
Foundations	1970s–1990s	Quantitative traits + polyploid complexity obscure inheritance	Early quantitative/cytogenetic framing; recognition of polyploid selection dynamics	Trait dissection problems defined; rationale for marker-based mapping	Lee, 1973; Jiang et al., 1998
Early linkage mapping	~2000–2006	Limited polymorphism in elite pools; sparse maps	RFLP/SSR linkage maps; first-generation QTL scans	Initial QTL regions for fiber strength/yield; first MAS targets	Zhang et al., 2002; Zhang et al., 2003; Park et al., 2005
Multi-environment QTL era	~2006–2012	QTL instability across environments; polygenic traits	Multi-environment QTL; meta-QTL; higher density	Robust QTL clusters; clearer trait modules	Yu et al., 2013; Said et al., 2013
Population-aware association mapping	~2008–2017	Confounding by structure; LD heterogeneity	Linkage disequilibrium (LD) characterization; mixed models; diversity panels	Reproducible marker–trait associations; germplasm-aware targets	Abdurakhmonov and Abdurakimov, 2008; Abdurakhmonov et al., 2008, Abdurakhmonov et al., 2009; Abdullaev et al., 2017
Standardized SNP/GWAS era	~2015–2021	Harmonizing loci across studies	63K/80K arrays; panel GWAS	Comparable discoveries across programs	Hulse-Kemp et al., 2015; Cai et al., 2017; Sun et al., 2017
Genomics and regulatory mapping	~2012–2021	Homoeolog ambiguity; regulatory variation	Reference genome; transcriptomes; eQTL	From loci to mechanisms in fiber/stress	Zhang et al., 2015; Wang et al., 2017a; Claverie et al., 2012
SV/pan-genome era	~2023–present	Reference bias; SV/ Presence-Absence Variant (PAV) as causal variants	SV-based pan-genomes; graph genomes	SV-aware trait discovery; better diversity representation	Jin et al., 2023; Zhang et al., 2026; Yang et al., 2026
Deployment era	ongoing	Multi-trait assembly + G×E	Marker pipelines; environment-smart testing; early GS frameworks	Cultivar delivery with quality + adaptation gains	Malik et al., 2014; Darmanov et al., 2022

3 Molecular linkage maps and first-generation QTL analysis

During 2000–2006, cotton entered the era of molecular linkage mapping. Zhang et al. (2002) constructed a molecular linkage map of allotetraploid cotton using a haploid population, establishing a framework for locating quantitative trait loci. With linkage maps, complex traits could begin to be analyzed as collections of genomic segments rather than as undifferentiated breeding values.

Early QTL studies were particularly influential when they identified loci with relatively large effects. Zhang et al. (2003) molecularly tagged a major QTL for fiber strength, demonstrating that a complex, economically important trait could be anchored to a genomic interval suitable for follow-up analysis. Additional QTL studies of fiber-related traits reinforced the view that lint yield and fiber quality were not only highly polygenic but also partly structured into genomic intervals that could be mapped and, eventually, combined in breeding programs (Mei et al., 2004).

Marker systems also became denser and more biologically informative during this period. Park et al. (2005) developed EST-derived microsatellites, providing an early bridge between genetic mapping and expressed genes, an important advance in cotton, where fiber traits arise from stage-specific transcriptional programs. As a result, mapping intervals could increasingly be linked to candidate pathways involving cell wall biosynthesis, cytoskeletal organization, and sugar transport, rather than remaining anonymous coordinates.

At the same time, these first-generation maps exposed a defining limitation of cotton genetics: QTL instability across

environments. Many loci were conditionally expressed, and similar phenotypes could arise through different physiological pathways under different field conditions. The next phase, therefore, required higher marker density, broader multi-environment replication, and study designs capable of distinguishing robust QTL from environmentally fragile signals.

4 Multi-environment mapping and meta-analysis

To address QTL instability, cotton genetics increasingly integrated evidence across studies and environments during 2006–2012. Multi-environment QTL mapping for lint yield and fiber quality made reliability an explicit criterion: mapping was no longer concerned only with detection, but also with reproducibility under realistic production conditions (Yu et al., 2013).

Shen et al. (2006) mapped resistance to root-knot nematodes, illustrating a translational pattern that would become increasingly important in cotton: loci could be identified and then deployed to reduce pesticide dependence and stabilize yield. Disease- and pest-resistance loci were especially attractive because they often showed comparatively large effects and offered direct sustainability benefits.

Meta-QTL analysis emerged as a practical response to heterogeneity across studies. Said et al. (2013) synthesized QTL reports for fiber quality, yield, morphological traits, drought tolerance, and disease resistance, identifying consensus regions for prioritization.

When individual studies lacked the power to define stable loci, aggregated evidence could reveal more robust signals.

Adaptation constraints also became more explicitly genetic during this period. Photoperiod dependence, a major barrier to expanding upland cotton into new latitudes, was addressed through induced mutation and subsequent genetic analysis, showing that key developmental switches could be reconfigured and mapped (Abdurakhmonov et al., 2007). This foreshadowed a broader lesson that remained important throughout later work: in cotton, mapping has the greatest translational value when it targets upstream regulators of phenology, disease resilience, and resource-use efficiency that can reshape downstream agronomic outcomes.

5 Linkage disequilibrium, association mapping, and population structure

During 2008–2017, cotton mapping expanded from biparental populations to broader germplasm panels, bringing population structure to the center of analysis. Cotton breeding has long been shaped by regional adaptation, market class, and intense selection for fiber standards; as a result, genetic relatedness is non-random, and linkage disequilibrium (LD) patterns vary across populations. Association mapping, therefore, required explicit treatment of both LD and population structure through careful measurement, modeling, and correction.

Cotton adopted the mixed-model framework widely used in crop association mapping, but its application in this species was constrained by features specific to allotetraploid genomes. Markers had to discriminate homoeologous loci, and inference had to remain robust despite uneven allele frequencies and subgenome-specific biases. Abdurakhmonov and Abdurakimov (2008) provided a broad synthesis of association-mapping frameworks for crop improvement, clarifying methodological principles relevant to cotton and other crops.

Early surveys of LD in both cultivated and exotic germplasm established the feasibility of association mapping in cotton by defining the population structure and LD-decay patterns that would guide later study design (Abdurakhmonov et al., 2008, Abdurakhmonov et al., 2009). Within this transition, association mapping of fiber-quality traits in upland cotton germplasm showed that LD-informed models could convert natural variation into interpretable marker–trait relationships with translational value for breeding (Abdurakhmonov et al., 2009). Similar analyses in *Gossypium barbadense* germplasm further demonstrated the utility of panel-based inference in a species where extra-long-staple fiber quality and adaptation are strongly shaped by population history (Abdullaev et al., 2017). Across these studies, a practical principle emerged: discoveries made in diverse germplasm are most useful when supported by replication and targeted validation.

6 SNP arrays and GWAS at scale

If association mapping was the conceptual shift, high-density SNP arrays were the engineering shift that made the approach scalable

during 2015–2021. Hulse-Kemp et al. (2015) developed a 63K SNP array that provided standardized, high-throughput genotyping, effectively creating a common yarn that could be woven into comparable results across labs and breeding programs. This standardization changed cotton mapping from an artisanal craft into a cumulative science: loci discovered in one program could be interrogated in another without rebuilding the entire genotyping apparatus.

GWAS studies using dense marker maps began to identify favorable SNP alleles and candidate genes for traits related to fiber, yield, and other agronomic outcomes (Su et al., 2016). Cotton GWAS also expanded to address how population structure relates to agronomic trait architecture at broader scales, linking genotype to trait distributions across breeding contexts (Huang et al., 2017) and identifying genetic variation and candidate genes for important traits in structured panels (Sun et al., 2017).

Further platform development, including an 80K SNP array (Cai et al., 2017), strengthened the standardization ecosystem and made large-scale studies more routine. The cumulative effect was a new kind of resolution: rather than producing isolated QTL maps, cotton genetics began producing reusable genotype–phenotype maps that could be compared, meta-analyzed, and prioritized.

Multi-parent resources added another dimension during this era. Huang et al. (2021) developed an 8-way Upland cotton MAGIC population, creating recombination-rich panels that bridge linkage and association mapping. These resources accelerate discovery in breeding-relevant backgrounds by combining the diversity of multiple founders with controlled recombination.

7 Genomes add a new dimension: polyploid assemblies and regulatory mapping

Reference genomes transformed the meaning of cotton genomic coordinates during 2012–2021. The sequencing of allotetraploid cotton (TM-1) by Zhang et al. (2015) provided a scaffold for converting mapped loci into candidate genes and mechanistic hypotheses, ushering in a new phase of polyploid crop genetics. Once reference assemblies became available, QTL and GWAS signals could be interpreted in the context of gene models, synteny, and biological pathways, enabling stronger links between fiber traits and the underlying processes of cell wall formation, carbohydrate partitioning, and stress response.

Genome-scale resources also clarified that the two cotton subgenomes did not contribute equally to domestication and improvement. Wang et al. (2017a) presented evidence for asymmetric subgenome selection and cis-regulatory divergence, providing a mechanistic explanation for why some loci show greater consistency across mapping contexts: recurrent selection has shaped both regulatory architecture and coding sequence. In fiber biology, this distinction was especially important because many quality traits depend on coordinated transcriptional programs and developmental timing rather than on variation in a single structural gene.

TABLE 2 Key genomic resources for cotton mapping.

Resource	What it enables	Polyploid-specific notes	Key references
Interspecific mapping populations (<i>hirsutum</i> × <i>barbadense</i> RILs)	QTL discovery + map-based cloning	High polymorphism; homoeolog resolution required	Lacape et al., 2009
SSR/EST marker systems	Early high-utility linkage maps	Transferable but limited density	Zhang et al., 2002; Park et al., 2005
SNP genotyping arrays (63K, 80K)	Standardized high-density genotyping	Subgenome specificity matters	Hulse-Kemp et al., 2015; Cai et al., 2017
GBS/BSA-seq mapping	Dense maps; rapid fine mapping	Needs homoeolog-aware calling	Zhao et al., 2017; Zhao et al., 2021
Reference genome (TM-1) and comparative assemblies	Candidate genes; syntenic	Annotation and subgenome bias	Zhang et al., 2015
eQTL/genetical genomics in fiber	Regulatory mapping; causal nomination	Stage-specific programs; cis/trans logic	Claverie et al., 2012
SV-based pan-genomes and graph pan-genomes	SV-aware association; reduces reference bias	PAV and rearrangements become first-class variants	Jin et al., 2023; Zhang et al., 2026; Yang et al., 2026

Transcriptomic resources added another layer of interpretation. Zhang et al. (2013a) analyzed the *Gossypium arboreum* transcriptome under water stress, demonstrating that sequencing-based approaches can reveal candidate pathways and temporal dynamics. In cotton, however, expression resources were most informative when integrated directly with genetic mapping.

This integration emerged through genetical genomics and eQTL mapping in developing cotton fibers. Claverie et al. (2012) used cDNA-AFLP-based approaches to link genotype to expression variation during fiber development, enabling candidate-gene nomination by colocalizing trait and regulatory loci. Interspecific resources, including recombinant inbred line populations derived from *G. hirsutum* × *G. barbadense* crosses (Lacape et al., 2009), provided essential infrastructure for this integrative work by supplying the recombination and polymorphism needed to map both phenotypes and expression traits. The key genomic resources that enabled these advances are summarized in Table 2; Supplementary Table 1.

8 From loci to cultivars: fine mapping, validation, and deployment

As marker density and population resources improved, cotton fine-mapping studies increasingly combined dense recombination analysis with transcript profiling to move from broad intervals to prioritized candidate genes. Liu et al. (2016) delimited a major multi-trait fiber QTL, the classical T1 region, to a narrow genetic interval and used RNA-seq and RT-PCR to prioritize likely causal genes. This study illustrated how expression data can strengthen the connection between statistical association and biological mechanism. Related work by Xu et al. (2017) localized introgressed fiber-length loci, including qFL-*chr1*, to sub-centimorgan intervals using near-isogenic backgrounds and substitution mapping, clarifying candidate pathways involving hormone signaling and transport processes while generating markers suitable for deployment. Similar studies integrating differential gene expression with QTL mapping for yield further emphasized that agronomic phenotypes reflect dynamic developmental processes and that mapping

becomes more informative when intermediate biological layers are considered (Liu et al., 2011).

Disease resistance provided a direct bridge between genetic mapping and sustainability. Wang et al. (2009) mapped Fusarium wilt resistance loci, illustrating how genetic information can reduce dependence on chemical control and increase yield stability, particularly when validated and pyramided. Later work strengthened this translational pathway in two complementary ways. First, chromosome-scale localization by Jiang et al. (2009) showed that resistance could be conferred by multiple linked factors, with Verticillium wilt resistance QTL clustering on D-subgenome chromosomes D7 and D9. Second, LD-aware strategies combined association evidence with localized recombination to resolve clustered QTL into major and minor components, narrowing intervals and identifying candidate defense genes (Zhao et al., 2017).

At the population scale, GWAS in diverse upland cotton panels identified repeatable loci for resistance to Verticillium wilt and Fusarium wilt, including race-specific contexts, helping breeders prioritize markers with more consistent effects across genetic backgrounds and trials (Abdelraheem et al., 2020). Integrative pipelines that combined GWAS with QTL-seq and transcriptome sequencing further translated these signals into breeder-ready assays, including Kompetitive Allele-Specific PCR (KASP) markers for Verticillium resistance that can be used in routine selection (Zhao et al., 2021). Transcriptome profiling of *G. barbadense* under Verticillium challenge complemented this mapping framework by identifying responsive pathways and candidate mechanisms that could be aligned with mapped loci (Zhang et al., 2013b).

Studies of domestication and geographic adaptation added another dimension to translation by linking genome-scale selection patterns with classical trait dissection across the transition from wild to cultivated cotton (Grover et al., 2020). Understanding the genomic basis of differentiation and fiber improvement helped connect historical selection with modern breeding priorities and provided a framework for designing future crosses and introgressions (He et al., 2021). In parallel, trait mapping extended to architectural and phenological determinants of yield potential, such as the node of the first fruiting branch, reflecting a more mechanistic understanding of how development shapes productivity (Zhang et al., 2021).

TABLE 3 Functionally validated genes from cotton mapping studies.

Gene	Trait	Mapping approach	Validation method	Key finding	Reference
GhSAD1	Cold tolerance	GWAS (200 accessions)	VIGS; Arabidopsis OE	Regulates ABA signaling; HapB increases cold tolerance	Ge et al., 2022
GhCHS	Fiber elongation	eQTL mapping	RNAi in cotton	Flavonoid pathway; ROS accumulation affects elongation	Chen et al., 2026
GhDFR	Secondary wall synthesis	eQTL mapping	RNAi in cotton	Flavonoid pathway; cellulose biosynthesis	Chen et al., 2026
GhPIN3	Plant height	QTL mapping (BIL population)	VIGS	Auxin efflux carrier; silencing increases height	Ma et al., 2019
GhPAP	Fiber strength	SLAF-BSA-seq + InDel analysis	VIGS; Arabidopsis OE	Plastid lipid-associated protein; affects helix formation	Zhang et al., 2025a
GhABH	Fiber quality (multi-effect)	Fine mapping + KASP	VIGS	α/β -hydrolase; regulates cell wall thickness	Zhang et al., 2025b
GhTPS11	Flowering time	QTL mapping + transcriptome	VIGS; Arabidopsis OE	Trehalose-6-phosphate synthase; age pathway	Feng et al., 2026
GhE6	First fruiting branch height	RTM-GWAS + meta-QTL	VIGS	Early flowering regulation	Su et al., 2024
GhUBP15	Plant height	UAV-based GWAS	VIGS	Ubiquitin protease; plant architecture	Ye et al., 2023
GhCUL1	Plant height	UAV-based GWAS	VIGS	Cullin protein; plant architecture	Ye et al., 2023
GhCPR30	Verticillium wilt resistance	QTL mapping + BSA-seq	VIGS	Disease resistance; silencing increases susceptibility	Wang et al., 2023
miR477b	Fiber length	QTL co-localization	VIGS	Regulates DELLA via HOX3	Song et al., 2024
GbCYP72A1	Verticillium wilt resistance	QTL mapping	VIGS	Cytochrome P450; hormone signaling	Xu et al., 2023
GhALDH7B4	Fiber strength	SLAF-BSA-seq	VIGS; Arabidopsis OE	Aldehyde dehydrogenase; cell wall components	Tang et al., 2024
GhRBB1_A07	Fiber quality	MAGIC GWAS	Gene expression + SNP	Regeneration of bulb biogenesis; superior fiber	Islam et al., 2016

VIGS, virus-induced gene silencing; OE, overexpression; BIL, backcross inbred line; BSA-seq, bulk segregant analysis sequencing; GWAS, genome-wide association study; UAV, unmanned aerial vehicle; RTM-GWAS, restricted two-stage multi-locus GWAS.

Cotton now also provides examples of mapping-derived markers translated into practical selection. Darmanov et al. (2022) showed that marker-assisted selection can deliver superior fiber-quality cultivars in the Ravnaq series, providing a concrete example of how mapping moves from publication to field deployment (see Box 1). More broadly, syntheses of molecular markers and cotton genetic improvement have emphasized that future sustainability will depend on integrating mapping with predictive tools and realistic breeding workflows (Malik et al., 2014). Table 3 catalogs functionally validated genes identified through cotton mapping studies, illustrating the field's progression from statistical association to biological understanding; a more comprehensive list is provided in Supplementary Table 2.

Several factors underpinned this translational success. Long-term institutional commitment over more than 15 years provided the continuity needed to sustain a complex marker-to-cultivar pipeline. Equally important was the close integration of discovery research with applied breeding, which enabled fundamental findings to be rapidly tested in breeding-relevant genetic backgrounds. Early investment in linkage disequilibrium characterization, undertaken before GWAS became routine, established a foundation for understanding population structure and marker-trait relationships that informed subsequent stages of the program (Abdurakhmonov

et al., 2008, Abdurakhmonov et al., 2009). The validation strategy combined the breadth of diversity panels with the precision of structured mapping populations, using both approaches to confirm associations before marker deployment (Kushanov et al., 2016). Associated SSR markers were then converted into breeder-friendly, higher-throughput formats suitable for routine screening in applied breeding programs.

Sustained international collaboration, particularly the Uzbekistan-USDA partnership, further strengthened this effort by combining complementary expertise, environments, and breeding resources (Darmanov et al., 2022). Importantly, trait-marker associations were evaluated across three distinct environments, Uzbekistan, the United States (College Station, Texas), and Mexico (USDA-ARS Winter Nursery), which helped identify the most stable associations across contrasting production settings (Abdurakhmonov et al., 2008, Abdurakhmonov et al., 2009). Collectively, these features show that successful translation in cotton genetics depends not only on robust science but also on long-time horizons, coordinated validation, and sustained collaboration across the discovery-to-deployment continuum. More broadly, this case suggests that, in a polyploid crop with modest polymorphism, the transition from LD-based discovery to cultivar release is feasible but requires sustained integration across the discovery, validation, and breeding phases.

BOX 1 Case study: the Ravnaq cultivars as a marker-to-deployment pipeline in cotton.

The Ravnaq cultivar series provides an instructive example of long-horizon translation from LD-informed discovery to cultivar release in cotton. The translation pipeline unfolded over approximately 15 years:

Phase	Activities	Key outputs	Duration	References
LD characterization	Survey of LD patterns in 335 <i>G. hirsutum</i> accessions; population structure analysis; identification of LD blocks associated with fiber quality	First comprehensive LD map in cotton; established the feasibility of association mapping in cotton germplasm	2005-2009	Abdurakhmonov et al., 2008, Abdurakhmonov et al., 2009
Marker-trait association	Multi-environment phenotyping (Uzbekistan, Mexico); MLM-based association mapping with Q+K correction	20–25 SSR markers associated with fiber length, strength, and micronaire across environments	2008-2012	Abdurakhmonov et al., 2008, Abdurakhmonov et al., 2009
Validation and refinement	Cross-validation in independent panels; QTL co-localization; meta-analysis	Confirmed stable associations; narrowed candidate intervals	2010-2014	Abdullaev et al., 2017
Marker conversion	Development of breeder-friendly PCR assays; conversion of associated SSRs to high-throughput formats	Deployable markers for MAS	2012-2015	Darmanov et al., 2022; Kushanov et al., 2016
Marker-assisted backcrossing	Five generations of backcrossing with foreground selection (BC ₅); background recovery using genome-wide markers	Near-isogenic lines with introgression of favorable LD blocks	2014-2019	Darmanov et al., 2022
Multi-location trials	Testing across diverse environments in Uzbekistan: yield and fiber quality assessment	Consistent superiority in fiber strength (38.3 vs 32.5 g tex ⁻¹) and length (1.18 vs 1.12 inch)	2017-2021	Darmanov et al., 2022
Cultivar release	Distinctness, uniformity, stability testing; certification	'Ravnaq-1' and 'Ravnaq-2' released for commercial cultivation	2021-2022	Darmanov et al., 2022

9 Structural variation and graph pan-genomes: beyond SNP-centric mapping

From 2023 onward, cotton mapping has increasingly moved beyond SNP-centric, single-reference frameworks based primarily on biallelic variation. Cotton increasingly challenges these assumptions because structural variation, presence/absence variation, and genomic rearrangements, especially across wild, landrace, and geographically diverse germplasm, can have large functional effects while remaining invisible to SNP-only approaches.

Jin et al. (2023) provided direct evidence for this shift by combining an SV-based cotton pan-genome with GWAS. Their study showed that structural variants are not merely genomic curiosities, but trait-associated variants that can reshape mapping outcomes and alter which alleles are realistically deployable in breeding. This marked a qualitative change in the questions cotton genetics could address. Instead of asking only which SNP tags a QTL, researchers could begin to ask which structural haplotype is likely causal and how to track it in breeding populations.

Graph-based genomes represent the next major advance in this framework because they enable more comprehensive incorporation of structural variation into discovery and breeding pipelines. A newly published complementary study by Zhang et al. (2026) further sharpened this transition by generating a telomere-to-telomere (T2T) genome for the elite cultivar NDM13 together with near-T2T assemblies for 27 additional representative *Gossypium hirsutum*

accessions. Using these resources, along with transcriptomic profiling across 15 tissues and population-scale analyses, the authors constructed a T2T-reference-based pangenome and identified 761,536 structural variants across 1,671 worldwide accessions evaluated in 22 environments, linking previously hidden SVs with breeding-relevant traits. Together with the pan-genome graph study by Yang et al. (2026), which used 107 assemblies spanning the wild-to-domesticated continuum and resolved six large-scale structural variants, these results show that modern cotton mapping is moving toward complementary T2T-reference, population-scale SV, and graph-based frameworks. For cotton improvement, the implication is immediate: some of the most useful alleles may not be fully represented or genotyped in conventional linear-reference, SNP-centric pipelines, and future mapping efforts will increasingly need to capture structural haplotypes, copy-number differences, and presence/absence variation directly.

Recent work has also expanded cotton mapping beyond sequence variation alone. Huang et al. (2024) integrated epigenomic and three-dimensional genomic analyses, revealing developmental dynamics and subgenome asymmetry in the transcriptional regulatory architecture of allotetraploid cotton. This study showed that mapping now extends beyond DNA sequence polymorphism to include chromatin conformation and regulatory topology, adding new mechanistic dimensions to the interpretation of complex traits.

Further comparative genomics in elite cotton lines further strengthened this point. Sreedasyam et al. (2024) generated high-quality genomes for three modern cultivars. They updated TM-1, showing that substantial sequence and structural variation persists even among modern breeding materials and overlaps with

TABLE 4 Trait exemplars: where mapping most clearly touches sustainability (abridged version; full table in Supplementary Table 3).

Trait domain	Typical targets	What made it hard	Mapping-to-mechanism bridge	Breeding leverage (sustainability link)	Key references
Fiber quality	QTL clusters; pleiotropy; stage-specific control; allele discovery via GWAS	Single-cell development; strong G×E; homoeolog redundancy	Fine-mapping + expression support; eQTL co-localization; SV-aware scans	Higher value per hectare; quality stability reduces waste and energy in processing	Mei et al., 2004 ; Abdurakhmonov et al., 2009 ; Sun et al., 2017 ; Liu et al., 2016 ; Xu et al., 2017 ; Zhang et al., 2025a , Zhang et al., 2025b ; Tang et al., 2024 ; Islam et al., 2016
Lint yield and stability	Polygenic loci; environment-responsive QTL; yield-component dissection	Strong field heterogeneity; correlated traits; management effects	Multi-environment QTL + transcriptome-linked loci	Stable yield reduces land expansion pressure and input volatility	Yu et al., 2013 ; Liu et al., 2011
Earliness/phenology and architecture	Maturity SNPs; fruiting branch traits; flowering/earliness loci	Trade-offs with yield/fiber; photoperiod sensitivity; breeding-region stratification	GWAS + candidate genes; QTL-seq for key architectural traits	Earlier harvest windows reduce risk, irrigation needs, and pest pressure	Su et al., 2016 ; Huang et al., 2017 ; Zhang et al., 2021 ; Abdurakhmonov et al., 2007 ; Ma et al., 2019 ; Feng et al., 2026 ; Su et al., 2024
Verticillium wilt resistance	Major-effect loci plus quantitative background; resistance QTL clusters	Pathogen diversity; durability challenge; environment interactions	QTL clustering + fine mapping; integrative GWAS/QTL-seq/transcriptome → KASP markers	Durable resistance reduces fungicide use and yield losses	Jiang et al., 2009 ; Zhao et al., 2017 ; Abdelraheem et al., 2020 ; Zhao et al., 2021 ; Wang et al., 2023 ; Xu et al., 2023
Fusarium wilt resistance	Major resistance genes/loci; panel-wide QTL	Race structure and geographic spread	Gene mapping + GWAS confirmation in breeding germplasm	Reduced chemical inputs and replanting; resilience in infested soils	Wang et al., 2009 ; Abdelraheem et al., 2020
Nematode resistance	Resistance QTL for root-knot nematodes	Phenotyping is destructive and environment-dependent	Interval/QTL mapping with validated markers	Reduced nematicide reliance; safer soils and water	Shen et al., 2006
Abiotic resilience	Meta-QTL for tolerance; expression-responsive loci	Strong G×E; drought timing effects; complex physiology	Meta-QTL synthesis; stress transcriptomes; eQTL and prediction	Water productivity and yield stability under climate volatility	Said et al., 2013 ; Zhang et al., 2013a ; Abdelraheem et al., 2017 ; Ge et al., 2022
Domestication and geographic differentiation	Selection sweeps; cis-regulatory divergence; improvement loci	Bottlenecks; subgenome asymmetry; introgression vs local adaptation	Population genomics + regulatory divergence; mapping of improvement signals	Broadens adaptive diversity; guides pre-breeding for resilience	Wang et al., 2017a ; He et al., 2021 ; Grover et al., 2020
Pan-genomes, SV, and graph coordinates	SV-based GWAS; graph pan-genomes; structural haplotypes	Reference bias; SV genotyping; subgenome-aware coordinates	SV-aware association on pan-genomes; epigenome/3D-genome context	Captures hidden diversity and reduces false negatives; enables durable stacking	Jin et al., 2023 ; Zhang et al., 2026 ; Yang et al., 2026 ; Huang et al., 2024
Multi-parent resources (MAGIC)	Allele-series mapping; recombination-rich panels	Construction cost; phenotyping scale	MAGIC, GWAS, bridging linkage and association	Accelerates discovery and validation in breeding-relevant backgrounds	Huang et al., 2021

transcriptional differences relevant to fiber development. Together with graph-based population resources, these results show that SV analysis becomes more informative when paired with expression and other functional datasets, because the practical question is not only where an SV occurs but whether it changes gene activity in ways that matter for breeding.

10 Deployment pipelines for sustainability: how mapping reduces inputs, not only time

Cotton mapping delivers its greatest practical value not simply by identifying loci, but by reducing the economic and environmental costs of crop improvement. Its contribution to sustainability lies in reducing breeding trial burden, enabling the assembly of multiple favorable traits, and improving performance stability under stress.

Disease-resistance loci provide some of the clearest examples. Once mapped, resistance loci can be pyramided to reduce dependence on pesticides and improve yield stability (Wang et al., 2009; Jiang et al., 2009; Zhao et al., 2017; Abdelraheem et al., 2020; Zhao et al., 2021). Abiotic-resilience traits become equally important sustainability targets when mapping enables selection for tolerance that stabilizes productivity under drought, heat, or salinity without increasing input demand. Abdelraheem et al. (2017) provided a comprehensive meta-analysis of QTL associated with abiotic and biotic stress resistance in tetraploid cotton, synthesizing decades of discovery into a more actionable framework for breeding.

Fiber-quality improvement represents a distinct sustainability pathway. Greater fiber quality and improved stability increase value per hectare and reduce waste across the production and textile value chain. Wild introgression studies that identify fiber-quality QTL alleles from *Gossypium tomentosum* (Zhang et al., 2011) show how biodiversity can be reintroduced into elite breeding when mapped loci guide the recovery of agronomically suitable backgrounds. In this sense, sustainability is not limited to input reduction alone, but also includes improving output value and resource-use efficiency. Table 4 summarizes how mapping has contributed to sustainability across major trait domains; a more detailed version is provided in Supplementary Table 3.

Taken together, these studies show that cotton mapping now supports a more environmentally smart breeding model, in which selection is guided not only by yield potential but also by stability under stress, disease pressure, and resource limitations. In this way, cotton genetic mapping functions increasingly as an applied plant biotechnology platform for sustainable crop improvement.

11 Cotton's journey comparisons with other crops

Cotton's mapping trajectory is clearer when compared with other major crops. In maize, rapid decay of linkage disequilibrium

enables high-resolution mapping but requires very dense marker coverage (Romay et al., 2013). By contrast, cotton's predominantly self-pollinating habit and domestication bottlenecks produced broader LD blocks, which facilitated early QTL discovery with sparse markers but later required larger populations and stricter statistical control for fine mapping and association analysis (Abdurakhmonov et al., 2009; Abdurakhmonov and Abdurakhmonov, 2008).

Rice benefited from the earlier availability of a complete reference genome, which accelerated the conversion of QTL intervals into candidate genes and provided useful models for multiparent population design and stress mapping (International Rice Genome Sequencing Project, 2005; Feng et al., 2016; Huang et al., 2021). Wheat offers a different polyploid comparison: although its larger, more repetitive genome initially slowed progress, international genomic resources helped accelerate discovery, and the integration of GWAS with genomic selection for quality traits provides a strategy increasingly relevant to cotton improvement (International Wheat Genome Sequencing Consortium, 2018).

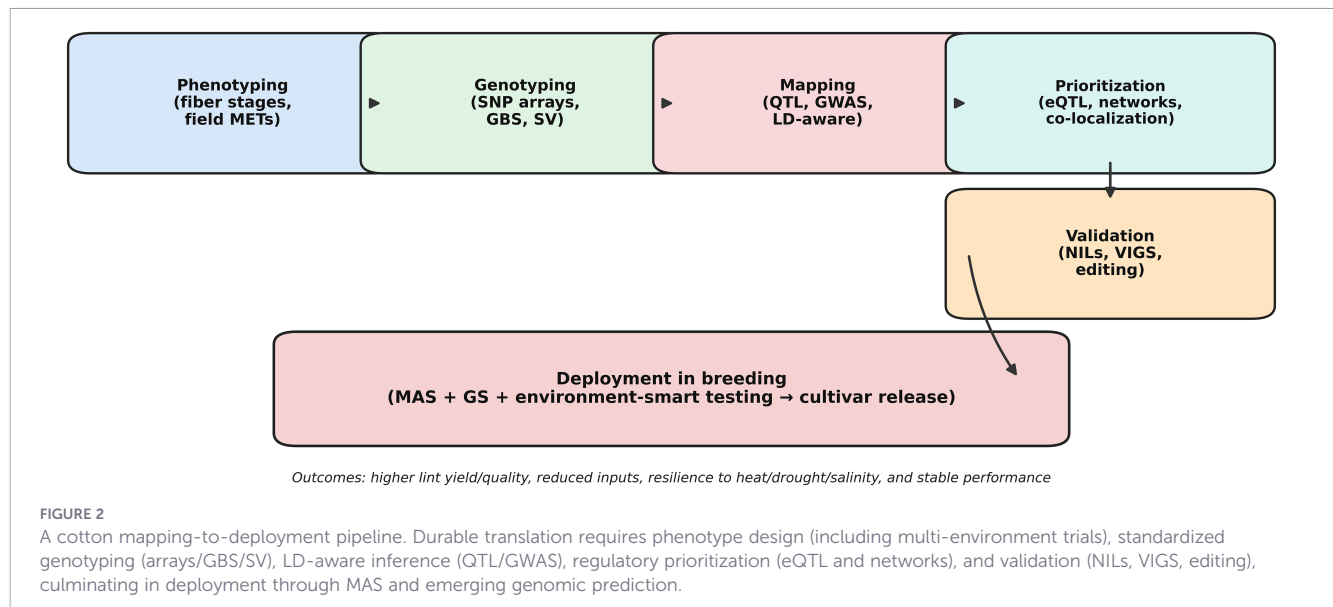
Soybean presents a closer parallel in the structure of elite germplasm, but its greater ease of transformation has generally enabled faster functional validation of candidate genes. In cotton, by contrast, validation has depended more heavily on approaches such as virus-induced gene silencing and near-isogenic line analysis (Xu et al., 2017; Zhao et al., 2021; Ge et al., 2022). Model systems such as Arabidopsis have also provided methodological templates, particularly for regulatory mapping and for integrating eQTL analysis with functional validation.

What distinguishes cotton is that polyploidy repeatedly drove methodological innovation rather than simply complicating implementation. Homoeolog-aware marker development, subgenome-specific expression analysis, and, more recently, SV-based pan-genomes and graph-based frameworks were advanced in response to cotton's distinctive genomic architecture. In this sense, cotton has not only adopted methods from other crops but has also helped push them toward greater analytical flexibility.

12 Gene regulatory networks: from single genes to systems

The transition from mapping loci to understanding gene function in cotton has increasingly shifted attention toward regulatory networks. Early eQTL studies in developing cotton fibers showed that expression variation for many genes mapped to transacting hotspots, indicating coordinated regulatory control during fiber development (Claverie et al., 2012; Liu et al., 2011). Chromosome 21 emerged as a notable hotspot for fiber-expressed genes, with multiple eQTLs colocalizing with phenotypic QTLs for fiber-quality traits (Claverie et al., 2012).

The integration of transcriptomics with QTL mapping has been especially powerful in cotton. Liu et al. (2016) combined fine mapping of the T1 fiber-quality QTL with RNA-seq to identify candidate genes whose expression differences between parental lines



were associated with the phenotype. This strategy, in which expression data are used to prioritize positional candidates, has become an important component of post-GWAS validation in cotton (Xu et al., 2017; Zhao et al., 2021; Huang et al., 2024).

More recently, epigenomic and three-dimensional genomic analyses have shown that the two cotton subgenomes are not transcriptionally equivalent. Huang et al. (2024) reported that the Dt subgenome exhibits more dynamic chromatin accessibility during development than the At subgenome, offering one possible explanation for why many fiber-quality QTLs map to the D subgenome even though its diploid progenitor did not produce spinnable fibers (Jiang et al., 1998). This asymmetry also extends to regulatory-element distribution: cis-regulatory elements showing signatures of selection during domestication are enriched in the D subgenome (Huang et al., 2024; Song et al., 2024).

Small RNAs add a further layer of regulatory complexity. For example, miR477b, which colocalizes with a fiber-length QTL, exhibits differential expression between long- and short-fiber lines, and silencing experiments support its role in fiber elongation by regulating DELLA and downstream targets (Song et al., 2024). Similarly, flavonoid-pathway genes such as GhCHS and GhDFR, identified through eQTL mapping, appear to influence fiber elongation and secondary wall synthesis through distinct regulatory mechanisms (Chen et al., 2026).

Together, these systems-level findings show that cotton phenotypic variation arises not only from coding-sequence polymorphisms but also from regulatory variation distributed asymmetrically across subgenomes and modulated by small RNAs, chromatin accessibility, and three-dimensional genome architecture. Future advances in cotton mapping will therefore depend on methods that can capture this regulatory complexity more directly.

13 A cotton-native mapping pipeline: conceptual synthesis

Cotton's mapping trajectory increasingly resembles an integrated pipeline (Figure 2). Phenotyping remains the major source of cost and irreproducibility, whereas genotyping has become increasingly standardized. At the same time, inference has become more sensitive to linkage disequilibrium and population structure, and faster functional assays and genome-editing tools are improving the capacity for validation. A central lesson from recent work is that cotton mapping is most durable when it integrates three complementary forms of evidence: (1) genetic evidence, including stable QTL or GWAS signals supported by multi-environment replication; (2) regulatory evidence, including expression modules and eQTL colocalization that help prioritize candidate genes; and (3) direct validation, including near-isogenic backgrounds, virus-induced gene silencing, or CRISPR-based editing. This integrative pipeline provides the practical foundation through which cotton genetic mapping contributes to plant biotechnology.

14 Persistent challenges in cotton mapping

Despite remarkable progress, several challenges persist in cotton genetic mapping. Homoeolog discrimination remains non-trivial even with short-read sequencing; many SNPs initially identified as polymorphisms between parents prove to be homoeologous sequence variants rather than true allelic differences (Hulse-Kemp et al., 2015; Page et al., 2016). This requires careful validation, often through subgenome-specific PCR or long-read sequencing.

Phenotyping throughput for fiber traits remains a bottleneck. Unlike above-ground traits that can be scored by UAVs (Ye et al., 2023), fiber quality assessment requires destructive harvesting and laboratory analysis (HVI, AFIS). This limits population sizes and replication, reducing power to detect small-effect QTLs. The gap between genotyping (millions of markers) and phenotyping (hundreds of lines) continues to widen.

G×E complexity is particularly acute in cotton, grown across diverse latitudes, water regimes, and management systems. QTLs identified in one environment frequently fail to replicate elsewhere (Yu et al., 2013; Said et al., 2013). Multi-environment trials are essential but costly, and statistical models that accurately predict G×E remain underdeveloped.

Validation throughput lags discovery. While GWAS and QTL mapping routinely identify dozens of loci, functional validation (VIGS, transgenics, genome editing) remains labor-intensive and low-throughput. The gap between candidate genes and validated genes widens each year.

Structural variation is only beginning to be systematically captured at the levels required for breeding, despite rapid recent progress in SV-based pan-genomes, T2T-reference population studies, and graph pan-genomes (Jin et al., 2023; Zhang et al., 2026; Yang et al., 2026). Presence-absence variation, copy number variants, and inversions likely contribute substantially to trait variation but are missed by SNP-centric approaches. Graph-based pan-genomes address this but require computational infrastructure not yet routine in breeding programs.

While gene editing in cotton is advancing, it still faces challenges related to efficiency and genotype dependence. Translating validated genes into edited cultivars requires optimization for each target and background. These challenges define the research agenda ahead: better homoeolog-aware analysis tools, high-throughput phenomics, environment-aware prediction models, scalable validation platforms, and routine incorporation of structural variation.

15 A practical workflow for cotton GWAS and marker deployment

For applied cotton GWAS, study design should begin with a panel matched to the breeding objective: elite-only panels for within-program deployment, broader diversity panels for allele discovery, and multi-parent or introgression resources when recombination or exotic diversity is limiting. Phenotyping should span multiple years or representative target environments, with common checks and standardized fiber assays, because single-season associations are rarely sufficient in cotton given strong genotype-by-environment interaction and phenotyping noise (Abdurakhmonov et al., 2008; Abdurakhmonov et al., 2009; Yu et al., 2013; Said et al., 2013; Hulse-Kemp et al., 2015; Huang et al., 2021). Marker datasets should then be filtered in a polyploid-aware manner to remove ambiguously mapped loci, control missingness and allele frequency, and evaluate population structure, kinship, and LD decay before association testing.

Model choice should follow the breeding question rather than convention alone. Mixed-model GWAS remains appropriate for detecting interpretable loci. Still, post-GWAS prioritization should emphasize replication across environments, local LD support, biological plausibility, and convertibility into robust assays such as KASP, CAPS, or other breeder-friendly markers. Markers become breeder-usable only when their effects are sufficiently large, stable, and easy to score in routine germplasm; loci that are small, background-dependent, or highly environment-sensitive are better treated as inputs to multi-locus prediction than as stand-alone selection tools (Liu et al., 2016; Xu et al., 2017; Zhao et al., 2021; Darmanov et al., 2022).

In practice, cotton breeders therefore face a decision point after discovery. Marker-assisted selection is most realistic for major or moderately strong loci with repeated support across populations and trials, especially for disease resistance, phenology, or fiber loci that can be validated in breeding backgrounds. When many loci are individually weak, or when phenotyping is expensive and genotype-by-environment interaction is large, genomic selection or combined marker-plus-prediction strategies are often more appropriate than locus-by-locus deployment. In this sense, marker discovery and marker usability should be treated as connected but distinct stages: GWAS identifies candidate leverage points, whereas breeding pipelines require evidence that those points are stable, cost-effective, and biologically credible enough to influence routine selection decisions.

After five decades of cotton genetic mapping, several practical lessons emerge for investigators entering the field. An important early step is to characterize linkage disequilibrium in the target germplasm, because LD structure directly affects marker-density requirements, mapping resolution, and the interpretability of association signals. In cotton, even a modest initial investment in LD assessment can generate substantial downstream benefits for study design, locus prioritization, and validation (Abdurakhmonov et al., 2008; Abdurakhmonov et al., 2009).

Polyploidy should not be regarded as a complication to be filtered out, but rather as a defining feature of cotton biology. Homoeologous gene pairs often retain functional relevance, and subgenome-specific effects remain central to trait interpretation (Wang et al., 2017a; Claverie et al., 2012). In particular, the D subgenome repeatedly contributes loci associated with fiber-quality variation in cultivated cotton, despite the limited contribution of its diploid progenitor to spinnable fiber (Jiang et al., 1998; Mei et al., 2004).

A further lesson is that statistical associations should be treated as hypotheses rather than endpoints. Robust translation depends on validation across environments and on convergence among multiple lines of evidence, including genetic association, regulatory support, and functional testing (Abdelraheem et al., 2017; Liu et al., 2016; Darmanov et al., 2022). Long-lived community resources, such as recombinant inbred lines, MAGIC populations, NAM populations, and chromosome segment substitution lines, often repay their development costs by enabling older questions to be revisited as genomic tools improve (Huang et al., 2021; Lacape et al., 2009).

As genotyping has become cheaper and more routine, high-quality phenotyping has emerged as the principal rate-limiting step. Multi-environment trials are expensive but often indispensable for distinguishing durable effects from transient associations (Ye et al., 2023). Given the magnitude of genotype-by-environment interaction in cotton, loci that appear strong in one season or location may weaken or disappear in another, making repeated evaluation across years and environments essential for identifying stable, deployable effects (Yu et al., 2013; Said et al., 2013; Abdelraheem et al., 2017).

Wild relatives preserved in germplasm collections are also likely to harbor valuable alleles that remain underused in modern breeding pools (Zhang et al., 2011; Wang et al., 2012, Wang et al., 2016; Wang et al., 2017b; Chen et al., 2020). At the same time, movement from discovery to cultivar release typically requires long timelines, often a decade or more. It proceeds most effectively when discovery researchers, breeders, and deployment partners remain coordinated across the full pipeline (Darmanov et al., 2022). Progress will also accelerate when the international cotton community shares germplasm resources, markers, and both successes and negative results, thereby improving reproducibility, increasing comparability across studies, and reducing redundant effort.

Taken together, these lessons provide a practical foundation for the next phase of cotton mapping and for its continuing contribution to plant biotechnology.

16 Future directions for cotton genetic mapping

The next phase of cotton mapping is likely to be defined by three priorities: a richer representation of genomic variation, stronger validation frameworks, and closer integration of discovery, prediction, and deployment. First, graph-based genomes and SV-aware pan-genomes may become the standard framework for cotton discovery because they reduce reference bias and capture variation increasingly relevant to breeding (Jin et al., 2023; Zhang et al., 2026; Yang et al., 2026). Recent cotton studies clarify why this shift matters. Comparative assemblies of modern cotton lines revealed substantial sequence, structural, and gene-content variation even among elite cultivars, together with transcriptional differences linked to fiber development and likely trait divergence (Sreedasyam et al., 2024). More recently, complementary advances have appeared at two scales: Zhang et al. (2026) produced a T2T reference for elite cultivar NDM13 plus near-T2T assemblies for 27 additional upland cotton accessions and used these resources to identify 761,536 SVs across 1,671 accessions with phenotypes from 22 environments, whereas Yang et al. (2026) built a graph pan-genome from 107 assemblies across the wild-to-domesticated continuum and resolved six large-scale structural variations shaping evolutionary history and agronomic trait architecture. Together, these studies suggest that future cotton mapping will increasingly require integration of T2T assemblies, population-scale SV catalogs,

graph coordinates, transcriptomics, regulatory data, and phenomics to prioritize structurally complex loci for breeding.

Second, cotton improvement will need to integrate mapping more closely with multi-environment prediction. Climate variability makes genotype-by-environment interaction an unavoidable component of trait architecture, so mapping results must increasingly be judged by their stability, transferability, and predictive value. Dense and standardized genotyping platforms are already available (Hulse-Kemp et al., 2015), but harmonized phenotyping and environment-aware models will determine whether predictive breeding becomes routine at scale.

Third, validation will determine which discoveries prove durable. In a polyploid crop, causal inference remains difficult without direct testing because of homoeolog redundancy and regulatory compensation, which can mask functional effects. Genomic analyses of subgenome divergence further underscore the need for validation strategies that are themselves homoeolog-aware (Wang et al., 2017a).

Beyond sequence variation alone, epigenomic and three-dimensional genomic context will increasingly inform mapping. The work of Huang et al. (2024) shows that regulatory architecture varies across developmental stages and between subgenomes, and that this variation will be essential for understanding how alleles function within their biological context.

Several technical bottlenecks will determine how quickly these advances become usable in breeding. SV discovery still varies across sequencing platforms, read lengths, callers, and graph-construction strategies, making cross-study comparisons and coordinate transfer difficult. Pan-genome resources also lack fully standardized conventions for graph construction, SV representation, and benchmark truth sets. At the same time, multi-omics integration remains limited by tissue specificity, developmental timing, and uneven sampling across environments. Practical breeding use adds another layer: most programs still need low-cost ways to genotype priority SVs, impute them accurately from short reads, and convert them into portable assays for routine screening. Near-term progress will therefore depend on harmonized SV standards, shared benchmark panels, graph-aware imputation pipelines, and marker-conversion workflows that reduce complex haplotypes to breeder-usable tests.

More broadly, cotton's history supports an important conclusion: technically difficult crops often drive methodological progress. Across successive eras, each major challenge generated tools and analytical frameworks that later proved useful beyond cotton itself. The long-term outcome, cultivars that are more productive, resilient, and resource-efficient, will depend on sustained scientific cooperation and on the effective integration of genomic discovery with breeding application.

Author contributions

IA: Conceptualization, Validation, Writing – review & editing, Resources, Visualization, Writing – original draft, Formal analysis.

Funding

The author(s) declared that financial support was received for this work and/or its publication. This work was supported by the Center of Genomics and Bioinformatics, Academy of Sciences of Uzbekistan

Acknowledgments

The author acknowledges the global cotton research community whose cumulative efforts over decades made this synthesis possible.

Conflict of interest

The author(s) declared that this work was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Generative AI statement

The author(s) declared that generative AI was not used in the creation of this manuscript.

Any alternative text (alt text) provided alongside figures in this article has been generated by Frontiers with the support of artificial intelligence and reasonable efforts have been made to ensure

accuracy, including review by the authors wherever possible. If you identify any issues, please contact us.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fpls.2026.1825852/full#supplementary-material>

SUPPLEMENTARY TABLE 1

Additional genomic resources for cotton mapping (expanded from Table 2).

SUPPLEMENTARY TABLE 2

Additional functionally validated genes from cotton mapping studies (expanded from Table 3). VIGS, virus-induced gene silencing; OE, overexpression; BIL, backcross inbred line; BSA-seq, bulk segregant analysis sequencing; GWAS, genome-wide association study; UAV, unmanned aerial vehicle; RTM-GWAS, restricted two-stage multi-locus GWAS; CSSL, chromosome segment substitution line; LP, lint percentage; FL, fiber length; ROS, reactive oxygen species; TF, transcription factor.

SUPPLEMENTARY TABLE 3

Complete trait exemplars: where mapping most clearly touches sustainability (expanded from Table 4).

References

- Abdelraheem, A., Elabbli, H., Zhu, Y., Kuraparthi, V., Hinze, L., Stelly, D., et al. (2020). A genome-wide association study uncovers consistent quantitative trait loci for resistance to Verticillium wilt and Fusarium wilt race 4 in the US Upland cotton. *Theor. Appl. Genet.* 133, 563–577. doi: 10.1007/s00122-019-03487-x
- Abdelraheem, A., Liu, F., Song, M., and Zhang, J. F. (2017). A meta-analysis of quantitative trait loci for abiotic and biotic stress resistance in tetraploid cotton. *Mol. Genet. Genomics* 292, 1221–1235. doi: 10.1007/s00438-017-1342-0
- Abdullaev, A. A., Salakhutdinov, I. B., Egamberdiev, S. S., Khurshut, E. E., Rizaeva, S. M., Ulloa, M., et al. (2017). Genetic diversity, linkage disequilibrium, and association mapping analyses of *Gossypium barbadense* L. germplasm. *PLoS One* 12, e0188125. doi: 10.1371/journal.pone.0188125
- Abdurakhmonov, I. Y., and Abdurakhimov, A. (2008). Application of association mapping to understanding the genetic diversity of plant germplasm resources. *Int. J. Plant Genomics* 2008, 574927. doi: 10.1155/2008/574927
- Abdurakhmonov, I. Y., Kohel, R. J., Yu, J. Z., Pepper, A. E., Abdullaev, A. A., Kushanov, F. N., et al. (2008). Molecular diversity and association mapping of fiber quality traits in exotic *G. hirsutum* L. germplasm. *Genomics* 92, 478–487. doi: 10.1016/j.ygeno.2008.07.013
- Abdurakhmonov, I. Y., Kushanov, F. N., Djaniqulov, F., Buriev, Z. T., Pepper, A. E., Fayzieva, N., et al. (2007). The role of induced mutation in conversion of photoperiod dependence in cotton. *J. Hered.* 98, 258–266. doi: 10.1093/jhered/esm007
- Abdurakhmonov, I. Y., Saha, S., Jenkins, J. N., Buriev, Z. T., Shermatov, S. E., Scheffler, B. E., et al. (2009). Linkage disequilibrium based association mapping of fiber quality traits in *G. hirsutum* L. variety germplasm. *Genetica* 136, 401–417. doi: 10.1007/s10709-008-9337-8
- Cai, C., Zhu, G., Zhang, T., and Guo, W. (2017). High-density 80 K SNP array is a powerful tool for genotyping *G. hirsutum* accessions and genome analysis. *BMC Genomics* 18, 654. doi: 10.1186/s12864-017-4062-2
- Chen, Z., Li, Z., Qi, Z., Qiao, L., Zhang, X., Tu, L., et al. (2026). Flavonoid metabolism enzymes GhCHS and GhDFR influence fiber elongation and secondary cell wall synthesis in cotton. *Plant Physiol.* 218, kiag031. doi: 10.1093/plphys/kiag031
- Chen, Q., Wang, W., Wang, C., Zhang, M., Yu, J., Zhang, Y., et al. (2020). Validation of QTLs for fiber quality introgressed from *Gossypium mustelinum* by selective genotyping. *G3* 10, 2377–2384. doi: 10.1534/g3.120.401125
- Claverie, M., Souquet, M., Jean, J., Forestier-Chiron, N., Lepitre, V., Pré, M., et al. (2012). cDNA-AFLP-based genetical genomics in cotton fibers. *Theor. Appl. Genet.* 124, 665–683. doi: 10.1007/s00122-011-1738-x
- Darmanov, M. M., Makamov, A. K., Ayubov, M. S., Khusenov, N. N., Buriev, Z. T., Shermatov, S. E., et al. (2022). Development of superior fibre quality upland cotton cultivar series 'Ravnaq' Using marker-assisted selection. *Front. Plant Sci.* 13, 906472. doi: 10.3389/fpls.2022.906472
- Feng, Y., Lu, Q., Zhai, R., Zhang, M., Xu, Q., Yang, Y., et al. (2016). Genome wide association mapping for grain shape traits in indica rice. *Planta* 244, 819–830. doi: 10.1007/s00425-016-2548-9
- Feng, J., Pan, A., Gao, Y., Yuan, Y., Lv, W., Du, Z., et al. (2026). GhTPS11, underlying an early-maturity QTL cluster on Chr. D08, positively regulates flowering through the age pathway in cotton. *Theor. Appl. Genet.* 139, 43. doi: 10.1007/s00122-026-05156-2
- Ge, C., Wang, L., Yang, Y., Liu, R., Liu, S., Chen, J., et al. (2022). Genome-wide association study identifies variants of GhSAD1 conferring cold tolerance in cotton. *J. Exp. Bot.* 73, 2222–2237. doi: 10.1093/jxb/erab555
- Grover, C. E., Yoo, M. J., Lin, M., Murphy, M. D., Harker, D. B., Byers, R. L., et al. (2020). Genetic analysis of the transition from wild to domesticated cotton (*Gossypium hirsutum* L.). *G3* 10, 731–754. doi: 10.1534/g3.119.400909
- He, S., Sun, G., Geng, X., Gong, W., Dai, P., Jia, Y., et al. (2021). The genomic basis of geographic differentiation and fiber improvement in cultivated cotton. *Nat. Genet.* 53, 916–924. doi: 10.1038/s41588-021-00844-9

- Huang, C., Nie, X., Shen, C., You, C., Li, W., Zhao, W., et al. (2017). Population structure and genetic basis of the agronomic traits of upland cotton in China revealed by a genome-wide association study using high-density SNPs. *Plant Biotechnol. J.* 15, 1374–1386. doi: 10.1111/pbi.12722
- Huang, C., Shen, C., Wen, T., Gao, B., Zhu, D., Li, D., et al. (2021). Genome-wide association mapping for agronomic traits in an 8-way Upland cotton MAGIC population by SLAF-seq. *Theor. Appl. Genet.* 134, 2459–2468. doi: 10.1007/s00122-021-03835-w
- Huang, X., Wang, Y., Zhang, S., Pei, L., You, J., Long, Y., et al. (2024). Epigenomic and 3D genomic mapping reveals developmental dynamics and subgenomic asymmetry of transcriptional regulatory architecture in allotetraploid cotton. *Nat. Commun.* 15, 10721. doi: 10.1038/s41467-024-55309-4
- Hulse-Kemp, A. M., Lemm, J., Plieske, J., Ashrafi, H., Buyyarapu, R., Fang, D. D., et al. (2015). Development of a 63K SNP array for cotton and high-density mapping of intraspecific and interspecific populations of *Gossypium* spp. G3 5, 1187–1209. doi: 10.1534/g3.115.018416
- International Rice Genome Sequencing Project (2005). The map-based sequence of the rice genome. *Nature* 436, 793–800. doi: 10.1038/nature03895
- International Wheat Genome Sequencing Consortium (2018). Shifting the limits in wheat research and breeding using a fully annotated reference genome. *Science* 361, eaar7191. doi: 10.1126/science.aar7191
- Islam, M. S., Thyssen, G. N., Jenkins, J. N., Zeng, L., Delhom, C. D., McCarty, J. C., et al. (2016). A MAGIC population-based genome-wide association study reveals functional association of GhRBB1_A07 gene with superior fiber quality in cotton. *BMC Genomics* 17, 903. doi: 10.1186/s12864-016-3249-2
- Jiang, C., Wright, R. J., El-Zik, K. M., and Paterson, A. H. (1998). Polyploid formation created unique avenues for response to selection in *Gossypium* (cotton). *Proc. Natl. Acad. Sci. U.S.A.* 95, 4419–4424. doi: 10.1073/pnas.95.8.4419
- Jiang, F., Zhao, J., Zhou, L., Guo, W., and Zhang, T. (2009). Molecular mapping of Verticillium wilt resistance QTL clustered on chromosomes D7 and D9 in upland cotton. *Sci. China C. Life. Sci.* 52, 872–884. doi: 10.1007/s11427-009-0110-8
- Jin, S., Han, Z., Hu, Y., Si, Z., Dai, F., He, L., et al. (2023). Structural variation (SV)-based pan-genome and GWAS reveal the impacts of SVs on the speciation and diversification of allotetraploid cottons. *Mol. Plant* 16, 678–693. doi: 10.1016/j.molp.2023.02.004
- Kushanov, F. N., Pepper, A. E., Yu, J. Z., Buriev, Z. T., Shermatov, S. E., Saha, S., et al. (2016). Development, genetic mapping and QTL association of cotton PHYA, PHYB, and HY5-specific CAPS and dCAPS markers. *BMC Genet.* 17, 141. doi: 10.1186/s12863-016-0448-4
- Lacape, J. M., Jacobs, J., Arioli, T., Derijcker, R., Forestier-Chiron, N., Llewellyn, D., et al. (2009). A new interspecific, *Gossypium hirsutum* × *G. barbadense*, RIL population: towards a unified consensus linkage map of tetraploid cotton. *Theor. Appl. Genet.* 119, 281–292. doi: 10.1007/s00122-009-1037-y
- Lee, J. (1973). The inheritance of gossypol level in *Gossypium*. II. Inheritance of seed gossypol in two strains of cultivated *Gossypium barbadense* L. *Genetics* 75, 259–264. doi: 10.1093/genetics/75.2.259
- Liu, R., Wang, B., Guo, W., Wang, L., and Zhang, T. (2011). Differential gene expression and associated QTL mapping for cotton yield based on a cDNA-AFLP transcriptome map in an immortalized F2. *Theor. Appl. Genet.* 123, 439–454. doi: 10.1007/s00122-011-1597-5
- Liu, D., Zhang, J., Liu, X., Wang, W., Liu, D., Teng, Z., et al. (2016). Fine mapping and RNA-Seq unravels candidate genes for a major QTL controlling multiple fiber quality traits at the T1 region in upland cotton. *BMC Genomics* 17, 295. doi: 10.1186/s12864-016-2605-6
- Ma, J., Pei, W., Ma, Q., Geng, Y., Liu, G., Liu, J., et al. (2019). QTL analysis and candidate gene identification for plant height in cotton based on an interspecific backcross inbred line population of *Gossypium hirsutum* × *G. barbadense*. *Theor. Appl. Genet.* 132, 2663–2676. doi: 10.1007/s00122-019-03380-7
- Malik, W., Ashraf, J., Iqbal, M. Z., Khan, A. A., Qayyum, A., Ali Abid, M., et al. (2014). Molecular markers and cotton genetic improvement: current status and future prospects. *Sci. World J.* 2014, 607091. doi: 10.1155/2014/607091
- Mei, M., Syed, N. H., Gao, W., Thaxton, P. M., Smith, C. W., Stelly, D. M., et al. (2004). Genetic mapping and QTL analysis of fiber-related traits in cotton (*Gossypium*). *Theor. Appl. Genet.* 108, 280–291. doi: 10.1007/s00122-003-1433-7
- Page, J. T., Liechty, Z. S., Alexander, R. H., Clemons, K., Hulse-Kemp, A. M., Ashrafi, H., et al. (2016). DNA sequence evolution and rare homoeologous conversion in tetraploid cotton. *PLoS Genet.* 12, e1006012. doi: 10.1371/journal.pgen.1006012
- Park, Y. H., Alabady, M. S., Ulloa, M., Sickler, B., Wilkins, T. A., Yu, J., et al. (2005). Genetic mapping of new cotton fiber loci using EST-derived microsatellites in an interspecific recombinant inbred line cotton population. *Mol. Genet. Genomics* 274, 428–441. doi: 10.1007/s00438-005-0037-0
- Romay, M. C., Millard, M. J., Glaubitz, J. C., Peiffer, J. A., Swarts, K. L., Casstevens, T. M., et al. (2013). Comprehensive genotyping of the USA national maize inbred seed bank. *Genome Biol.* 14, R55. doi: 10.1186/gb-2013-14-6-r55
- Said, J. I., 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, 776. doi: 10.1186/1471-2164-14-776
- Shen, X., Van Becelaere, G., Kumar, P., Davis, R. F., May, O. L., and Chee, P. (2006). QTL mapping for resistance to root-knot nematodes in the M-120 RNR Upland cotton line (*Gossypium hirsutum* L.) of the Auburn 623 RNR source. *Theor. Appl. Genet.* 113, 1539–1549. doi: 10.1007/s00122-006-0401-4
- Song, J., Liu, G., Jin, C., Pei, W., Zhang, B., Jia, B., et al. (2024). Co-localization and analysis of miR477b with fiber length quantitative trait loci in cotton. *Physiol. Plant* 176, e14303. doi: 10.1111/ppl.14303
- Sreedasyam, A., Lovell, J. T., Mamidi, S., Khanal, S., Jenkins, J. W., Plott, C., et al. (2024). Genome resources for three modern cotton lines guide future breeding efforts. *Nat. Plants* 10, 1039–1051. doi: 10.1038/s41477-024-01713-z
- Su, J., Li, D., Yuan, W., Li, Y., Ju, J., Wang, N., et al. (2024). Integrating RTM-GWAS and meta-QTL data revealed genomic regions and candidate genes associated with the first fruit branch node and its height in upland cotton. *Theor. Appl. Genet.* 137, 207. doi: 10.1007/s00122-024-04703-z
- Su, J., Pang, C., Wei, H., Li, L., Liang, B., Wang, C., et al. (2016). Identification of favorable SNP alleles and candidate genes for traits related to early maturity via GWAS in upland cotton. *BMC Genomics* 17, 687. doi: 10.1186/s12864-016-2875-z
- Sun, Z., Wang, X., Liu, Z., Gu, Q., Zhang, Y., Li, Z., et al. (2017). Genome-wide association study discovered genetic variation and candidate genes of fibre quality traits in *Gossypium hirsutum* L. *Plant Biotechnol. J.* 15, 982–996. doi: 10.1111/pbi.12693
- Tang, L., Liu, C., Li, X., Wang, H., Zhang, S., Cai, X., et al. (2024). An aldehyde dehydrogenase gene, GhALDH7B4_A06, positively regulates fiber strength in upland cotton (*Gossypium hirsutum* L.). *Front. Plant Sci.* 15, 1377682. doi: 10.3389/fpls.2024.1377682
- Wang, B., Draye, X., Zhuang, Z., Zhang, Z., Liu, M., Lubbers, E. L., et al. (2017b). QTL analysis of cotton fiber length in advanced backcross populations derived from a cross between *Gossypium hirsutum* and *G. mustelinum*. *Theor. Appl. Genet.* 130, 1297–1308. doi: 10.1007/s00122-017-2889-1
- Wang, B., Liu, L., Zhang, D., Zhuang, Z., Guo, H., Qiao, X., et al. (2016). A Genetic Map Between *Gossypium hirsutum* and the Brazilian Endemic *G. mustelinum* and Its Application to QTL Mapping. G3 6, 1673–1685. doi: 10.1534/g3.116.029116
- Wang, B., Nie, Y., Lin, Z., Zhang, X., Liu, J., and Bai, J. (2012). Molecular diversity, genomic constitution, and QTL mapping of fiber quality by mapped SSRs in introgression lines derived from *Gossypium hirsutum* × *G. darwinii*. *Theor. Appl. Genet.* 125, 1263–1274. doi: 10.1007/s00122-012-1911-x
- Wang, P., Su, L., Qin, L., Hu, B., Guo, W., and Zhang, T. (2009). Identification and molecular mapping of a Fusarium wilt resistant gene in upland cotton. *Theor. Appl. Genet.* 119, 733–739. doi: 10.1007/s00122-009-1084-4
- Wang, M., Tu, L., Lin, M., Lin, Z., Wang, P., Yang, Q., et al. (2017a). Asymmetric subgenome selection and cis-regulatory divergence during cotton domestication. *Nat. Genet.* 49, 579–587. doi: 10.1038/ng.3807
- Wang, Y., Zhao, J., Chen, Q., Zheng, K., Deng, X., Gao, W., et al. (2023). Quantitative trait locus mapping and identification of candidate genes for resistance to Verticillium wilt in four recombinant inbred line populations of *Gossypium hirsutum*. *Plant Sci.* 327, 111562. doi: 10.1016/j.plantsci.2022.111562
- Xu, P., Gao, J., Cao, Z., Chee, P. W., Guo, Q., Xu, Z., et al. (2017). Fine mapping and candidate gene analysis of qFL-chr1, a fiber length QTL in cotton. *Theor. Appl. Genet.* 130, 1309–1319. doi: 10.1007/s00122-017-2890-8
- Xu, J., Zhao, J., Liu, J., Dong, C., Zhao, L., Ai, N., et al. (2023). GbCYP72A1 improves resistance to Verticillium wilt via multiple signaling pathways. *Plant Dis.* 107, 3198–3210. doi: 10.1094/PDIS-01-23-0033-RE
- Yang, Z., Yang, Z., Gao, C., Zhang, M., Hu, G., Yang, L., et al. (2026). Graph pan-genome illuminates evolutionary trajectories and agronomic trait architecture in allotetraploid cotton. *Nat. Genet.* 58, 218–229. doi: 10.1038/s41588-025-02462-1
- Ye, Y., Wang, P., Zhang, M., Abbas, M., Zhang, J., Liang, C., et al. (2023). UAV-based time-series phenotyping reveals the genetic basis of plant height in upland cotton. *Plant J.* 115, 937–951. doi: 10.1111/tpj.16272
- Yu, J., Zhang, K., Li, S., Yu, S., Zhai, H., Wu, M., et al. (2013). Mapping quantitative trait loci for lint yield and fiber quality across environments in a *Gossypium hirsutum* × *Gossypium barbadense* backcross inbred line population. *Theor. Appl. Genet.* 126, 275–287. doi: 10.1007/s00122-012-1980-x
- Zhang, J., Guo, W., and Zhang, T. (2002). Molecular linkage map of allotetraploid cotton (*Gossypium hirsutum* L. × *Gossypium barbadense* L.) with a haploid population. *Theor. Appl. Genet.* 105, 1166–1174. doi: 10.1007/s00122-002-1100-4
- Zhang, T., Hu, Y., Jiang, W., Fang, L., Guan, X., Chen, J., et al. (2015). Sequencing of allotetraploid cotton (*Gossypium hirsutum* L. acc. TM-1) provides a resource for fiber improvement. *Nat. Biotechnol.* 33, 531–537. doi: 10.1038/nbt.3207
- Zhang, J., Jia, X., Guo, X., Wei, H., Zhang, M., Wu, A., et al. (2021). QTL and candidate gene identification of the node of the first fruiting branch (NFFB) by QTL-seq in upland cotton (*Gossypium hirsutum* L.). *BMC Genomics* 22, 882. doi: 10.1186/s12864-021-08164-2
- Zhang, S., Liu, C., Tang, L., Li, X., Cai, X., Wang, H., et al. (2025a). A plastid lipid-associated protein-encoding gene (GhPAP) that positively regulates fiber strength was

- identified via genetic mapping and transcriptomic analysis of a stable QTL on chromosome D06 of upland cotton. *Theor. Appl. Genet.* 138, 134. doi: 10.1007/s00122-025-04922-y
- Zhang, Z., Rong, J., Waghmare, V. N., Chee, P. W., May, O. L., Wright, R. J., et al. (2011). QTL alleles for improved fiber quality from a wild Hawaiian cotton, *Gossypium tomentosum*. *Theor. Appl. Genet.* 123, 1075–1088. doi: 10.1007/s00122-011-1649-x
- Zhang, Y., Sun, Z., Tian, S., Wu, L., Gu, Q., Ke, H., et al. (2026). A pangenome reference and population studies link structural variants with breeding traits in *Gossypium hirsutum*. *Nat. Genet.* 58, 928–939. doi: 10.1038/s41588-026-02523-z
- Zhang, Y., Wang, X. F., Ding, Z. G., Ma, Q., Zhang, G. R., Zhang, S. L., et al. (2013a). Transcriptome profiling of *Gossypium barbadense* inoculated with *Verticillium dahliae* provides a resource for cotton improvement. *BMC Genomics* 14, 637. doi: 10.1186/1471-2164-14-637
- Zhang, S., Wei, J., Tang, L., Liu, C., Li, X., Cai, X., et al. (2025b). An alpha/beta-hydrolase-like gene (GhABH) is identified to be responsible for fiber quality from a multi-effect quantitative trait locus (QTL) on chromosome A06 of upland cotton. *Theor. Appl. Genet.* 138, 268. doi: 10.1007/s00122-025-05049-w
- Zhang, X., Yao, D., Wang, Q., Xu, W., Wei, Q., Wang, C., et al. (2013b). mRNA-seq analysis of the *Gossypium arboreum* transcriptome reveals tissue selective signaling in response to water stress during seedling stage. *PLoS One* 8, e54762. doi: 10.1371/journal.pone.0054762
- Zhang, T., Yuan, Y., Yu, J., Guo, W., and Kohel, R. J. (2003). Molecular tagging of a major QTL for fiber strength in Upland cotton and its marker-assisted selection. *Theor. Appl. Genet.* 106, 262–268. doi: 10.1007/s00122-002-1101-3
- Zhao, Y., Chen, W., Cui, Y., Sang, X., Lu, J., Jing, H., et al. (2021). 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.* 134, 1063–1081. doi: 10.1007/s00122-020-03752-4
- Zhao, Y., Wang, H., Chen, W., Zhao, P., Gong, H., Sang, X., et al. (2017). Regional association analysis-based fine mapping of three clustered QTL for *verticillium* wilt resistance in cotton (*G. hirsutum* L.). *BMC Genomics* 18, 661. doi: 10.1186/s12864-017-4074-y