

Genome Sequencing Advances in *Gossypium*: Implications for Cotton Breeding

Jinhua Chen, Mengting Luo ✉

Institute of Life Science, Jiyang College of Zhejiang AandF University, Zhuji, 311800, Zhejiang, China

✉ Corresponding author: mengting.luo@jicat.org

Cotton Genomics and Genetics, 2024, Vol.15, No.2 doi: [10.5376/cgg.2024.15.0007](https://doi.org/10.5376/cgg.2024.15.0007)

Received: 15 Jan., 2024

Accepted: 22 Feb., 2024

Published: 07 Mar., 2024

Copyright © 2024 Chen and Luo, This is an open access article published under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Preferred citation for this article:

Chen J.H., and Luo M.T., 2024, Genome sequencing advances in *Gossypium*: implications for cotton breeding, Cotton Genomics and Genetics, 15(2): 66-80 (doi: [10.5376/cgg.2024.15.0006](https://doi.org/10.5376/cgg.2024.15.0006))

Abstract The advancement of genome sequencing technology has completely changed the field of cotton (*Gossypium*) breeding, providing unprecedented insights into the genetic structure of this important crop. This study systematically evaluates and integrates the latest advances in cotton genome sequencing, with a focus on historical milestones, technological innovations, and actual impacts on breeding plans. This study emphasizes key achievements such as sequencing of *Gossypium arboreum* and assembly of *Gossypium hirsutum* and *Gossypium barbadense* reference genomes. The integration of high-throughput sequencing technology and multi omics methods significantly enhances the understanding of cotton genetics, promoting the development of high-quality cotton varieties with improved fiber quality, disease resistance, and environmental adaptability. This study also discusses the challenges and limitations related to genome sequencing, and provides a forward-looking outlook on the future directions of cotton genomics research, including the potential of synthetic biology and precision agriculture. By elucidating the transformative potential of genome sequencing in cotton breeding, this provides a reference for continuing to invest in genome technology to achieve sustainable development of cotton agriculture.

Keywords Cotton breeding; Genomics; High-throughput sequencing; Multi-omics integration; Synthetic biology

1 Introduction

Cotton (*Gossypium*) stands as one of the most vital crops globally, playing a pivotal role in both agriculture and industry. As a primary source of natural fiber, cotton supports the textile industry, which is a significant economic driver in many countries. The plant's by-products, such as cottonseed oil and meal, are valuable for food and livestock feed. Cotton cultivation provides livelihoods for millions of farmers, particularly in developing regions, underscoring its socioeconomic importance. Enhancing cotton's yield, fiber quality, and resistance to pests and diseases can therefore have substantial economic and social benefits (Yang et al., 2020).

Genome sequencing has revolutionized the field of plant breeding by providing comprehensive insights into the genetic makeup of crops. This technology enables the identification of genes associated with desirable traits, facilitating targeted breeding strategies. In the context of cotton, genome sequencing has been instrumental in understanding the complex genome of different *Gossypium* species. By decoding the genetic information, researchers can identify markers for traits such as fiber quality, yield potential, and resistance to biotic and abiotic stresses (Huang et al., 2021). The integration of genome sequencing with traditional breeding methods accelerates the development of superior cotton varieties, ultimately enhancing agricultural productivity and sustainability (Ma and Cao, 2018).

This study comprehensively analyzes the advancements in cotton species genome sequencing and its impact on cotton breeding. The research covers historical milestones in cotton genome research, the latest technological advancements, and achievements in sequencing different cotton genomes. Additionally, the study explores the practical applications of genome sequencing in cotton breeding, including case studies of successful breeding strategies. The challenges and limitations faced by cotton genome research are also discussed, along with suggestions for future improvements. By elucidating the transformative potential of genome sequencing in enhancing cotton breeding practices and meeting the global demand for high-quality cotton, this study provides a reference for continued investment in genomic technologies to achieve sustainable development in cotton agriculture.

2 Historical Perspective on Cotton Genome Sequencing

2.1 Early efforts and milestones in cotton genome research

Early research on the cotton (*Gossypium*) genome primarily focused on deciphering the basic structure and function of the genome. Initial genome sequencing efforts used techniques such as Restriction Fragment Length Polymorphism (RFLP) and Simple Sequence Repeat (SSR) markers to construct preliminary maps of the cotton genome. These markers were essential tools for constructing genetic maps and locating genes within the cotton genome. For example, the study by Li et al. (2018) utilized high-throughput sequencing to reveal extensive repetitive sequences in the cotton genome, providing valuable data for subsequent genome assembly and functional genomics research (Li et al., 2018).

Early genome research also included preliminary investigations into the genetic basis of key agronomic traits in cotton. For instance, studies on fiber quality, yield, and related traits led to the identification of several Quantitative Trait Loci (QTL). These studies not only highlighted critical regions of the cotton genome that control these important traits but also provided a theoretical foundation and technical support for subsequent molecular breeding efforts (Huang et al., 2021). Additionally, early research focused on the structural features of the cotton genome, such as gene density, repetitive sequences, and chromosomal structural variations. These findings were instrumental in guiding future genome assembly and annotation efforts (Du et al., 2018).

With the advancement of sequencing technologies in the 21st century, significant breakthroughs were made in cotton genome research. High-throughput sequencing technologies enabled deep sequencing of the cotton genome, revealing extensive gene sequences and structural variations. This provided crucial data for genome assembly and functional annotation, laying a solid foundation for understanding the complexity and evolutionary processes of the cotton genome. These early studies set the stage for subsequent genomic research, offering valuable resources for future studies (Ma and Cao, 2018).

2.2 Key Discoveries and their impact on cotton genetics

Key discoveries in cotton genome research have significantly advanced the field of cotton genetics. One major discovery is the revelation of whole-genome duplication events. Polyploidization is a crucial process in the evolution of the cotton genome, with research uncovering multiple whole-genome duplication events that have profoundly impacted the structure and function of the cotton genome. Through comprehensive genome sequencing and comparative genomics analyses, scientists identified the timing and extent of these duplication events, providing essential insights into the evolutionary history of the cotton genome (Wang et al., 2016; Pan et al., 2020).

Another significant discovery is the identification of genes and regulatory mechanisms involved in fiber development. Cotton fiber is a key economic trait, with fiber quality directly affecting the market value of cotton. Genomic studies have identified several critical genes associated with fiber length, strength, and maturity. Functional analyses of these genes have elucidated the molecular mechanisms underlying fiber development, providing a theoretical basis and technical support for improving cotton fiber quality through molecular breeding (Huang et al., 2021).

Cotton genome research has uncovered many important metabolic pathways and regulatory networks. For example, studies have identified genes related to stress resistance, which play crucial roles in cotton's response to biotic and abiotic stresses. Functional analyses of these genes have provided deeper insights into the mechanisms by which cotton adapts to environmental stresses such as pests, drought, and salinity. These discoveries have enriched our understanding of cotton genome function and offered new strategies for improving cotton resistance through molecular breeding (He et al., 2021).

2.3 Evolution of sequencing technologies and their adoption in cotton research

The continuous advancement of genome sequencing technologies has brought unprecedented opportunities for cotton genome research. Early genome sequencing efforts primarily relied on Sanger sequencing technology, which, despite its high accuracy, was limited by low throughput and high costs, restricting large-scale genome sequencing applications. The emergence of high-throughput sequencing technologies has dramatically increased

the speed and efficiency of genome sequencing while significantly reducing costs, making whole-genome sequencing feasible. These technological advancements have greatly propelled cotton genome research.

With the advent of high-throughput sequencing technologies, scientists have been able to deeply sequence and assemble the entire cotton genome. Using platforms such as PacBio and Illumina, researchers have successfully completed the whole-genome sequencing of multiple cotton varieties and related species. These datasets have provided high-quality genome sequences and revealed complex structures within the cotton genome, such as repetitive sequences, gene family expansions, and gene recombination events. These findings are crucial for understanding the evolution and function of the cotton genome.

The integration of multi-omics data is another significant advancement in cotton genome research. By combining genome sequencing with transcriptomics, proteomics, and metabolomics data, scientists can gain a comprehensive understanding of gene expression regulation and metabolic pathway dynamics. These studies have revealed gene expression patterns and regulatory mechanisms in cotton under various developmental stages and environmental conditions. For example, integrating transcriptomics and metabolomics data has identified key regulatory genes and metabolic pathways involved in fiber development and stress resistance (Li et al., 2021).

3 Technological Advances in Genome Sequencing

3.1 Development of high-throughput sequencing technologies

High-throughput sequencing (HTS) technologies have revolutionized the field of genomics, allowing for the rapid and comprehensive sequencing of complex genomes such as those of cotton (*Gossypium*). These technologies have evolved significantly since the advent of Sanger sequencing, providing higher speed, accuracy, and cost-effectiveness. The development of HTS platforms like Illumina, PacBio, and Oxford Nanopore has enabled the detailed sequencing of both diploid and tetraploid cotton species, revealing intricate details of their genome structures (Wang et al., 2018).

One of the landmark achievements in this domain was the sequencing of *Gossypium hirsutum* and *Gossypium barbadense*, which provided high-quality reference genomes. This was achieved by integrating single-molecule real-time sequencing, BioNano optical mapping, and high-throughput chromosome conformation capture techniques. These efforts have significantly improved the contiguity and completeness of the genome assemblies, especially in regions with high repeat content like centromeres.

The assembly of the *Gossypium hirsutum* TM-1 genome, which integrated whole-genome shotgun reads, BAC-end sequences, and genetic maps, highlighted the asymmetric evolution between A and D subgenomes. This comprehensive sequencing effort provided critical insights into the genomic signatures of selection and domestication associated with fiber improvement and stress tolerance (Zhang et al., 2015).

3.2 Innovations in data analysis and bioinformatics

The rapid advancements in sequencing technologies have been paralleled by significant innovations in data analysis and bioinformatics. The complexity and volume of data generated by HTS necessitate sophisticated tools and methods for data processing, analysis, and interpretation. Innovations in bioinformatics have enabled the efficient handling of large datasets, facilitating the assembly, annotation, and comparative analysis of cotton genomes.

Bioinformatics platforms such as COTTONOMICS integrate vast amounts of genomic, transcriptomic, and epigenetic data, providing a comprehensive database for cotton research. This platform allows for the retrieval and analysis of data concerning cotton genomes, genomic variations, gene expression, and epigenetic regulation, thereby enabling researchers to decipher complex genetic traits and their regulatory networks (Dai et al., 2022).

The development of novel computational pipelines such as IGIA for reconstructing accurate gene structures from integrated data has been pivotal. These pipelines facilitate the exploration of transcriptional landscapes in cotton, revealing dynamic gene expression patterns, alternative splicing events, and regulatory mechanisms involved in fiber development and stress responses (Wang et al., 2019).

3.3 Integration of multi-omics approaches

The integration of multi-omics approaches has become a cornerstone in advancing cotton genomics. By combining genomics, transcriptomics, proteomics, and metabolomics, researchers can achieve a holistic understanding of the molecular mechanisms underlying important agronomic traits. This integrative approach has led to significant discoveries in gene function, regulatory networks, and metabolic pathways.

For instance, integrating transcriptomic and metabolomic data has provided valuable insights into the regulation of fiber development and stress resistance. Studies have identified key regulatory genes and metabolic pathways that play crucial roles in these processes, offering new targets for genetic improvement (Huang et al., 2021).

Advancements in multi-omics platforms have facilitated the detailed mapping of quantitative trait loci (QTL) associated with fiber quality and yield. These efforts have resulted in the identification of candidate genes and genetic variants that can be used in marker-assisted selection and genome editing to enhance cotton breeding programs (Thyssen et al., 2018).

4 Recent Achievements in *Gossypium* Genome Sequencing

4.1 Sequencing of the *Gossypium arboreum* genome

The sequencing of the *Gossypium arboreum* genome represents a significant milestone in cotton genomics, providing valuable insights into the genetic makeup of this diploid cotton species. *Gossypium arboreum*, also known as tree cotton, is one of the progenitors of the cultivated tetraploid cotton species and has been a critical resource for understanding cotton genetics and evolution. Recent advancements in sequencing technologies have enabled the generation of a high-quality genome assembly for *Gossypium arboreum*. This assembly has revealed important information about the structure, function, and evolution of cotton genomes.

One of the important achievements of kapok genome sequencing is the identification of a large number of structural variations and repetitive sequences, which constitute the foundation of its genomic complexity. Huang et al. (2020) revealed multiple key findings through comprehensive genomic analysis, including phylogenetic relationships, nucleotide variations, and chromosomal differences between different *Gossypium* species (Figure 1) (Huang et al., 2020).

Comparative genomics analyses using the *Gossypium arboreum* genome have provided insights into the polyploidization events that have shaped the evolution of cotton. These analyses have identified conserved and divergent genomic regions between diploid and tetraploid cotton species, revealing the genetic mechanisms underlying cotton domestication and adaptation (Wang et al., 2018).

4.2 Assembly of reference genomes for *Gossypium hirsutum* and *Gossypium barbadense*

The assembly of high-quality reference genomes for *Gossypium hirsutum* (upland cotton) and *Gossypium barbadense* (Pima cotton) marks a significant advancement in cotton genomics. These two species are the most widely cultivated cotton varieties, known for their superior fiber quality and yield. The reference genomes provide a comprehensive framework for understanding the genetic basis of important traits and for guiding cotton breeding programs.

The reference genome assemblies for *Gossypium hirsutum* and *Gossypium barbadense* were achieved by integrating multiple sequencing technologies, including single-molecule real-time sequencing, BioNano optical mapping, and high-throughput chromosome conformation capture techniques. These assemblies have significantly improved the contiguity and completeness of the genome sequences, especially in regions with high repeat content such as centromeres (Wang et al., 2018).

The improved genome assemblies have facilitated the identification of extensive structural variations, such as large paracentric and pericentric inversions, that have occurred after polyploidization. These structural variations are associated with important agronomic traits, including fiber quality and disease resistance. Additionally, the reference genomes have enabled the construction of introgression lines, allowing researchers to identify quantitative trait loci (QTL) associated with superior fiber quality and other desirable traits (Li et al., 2015).

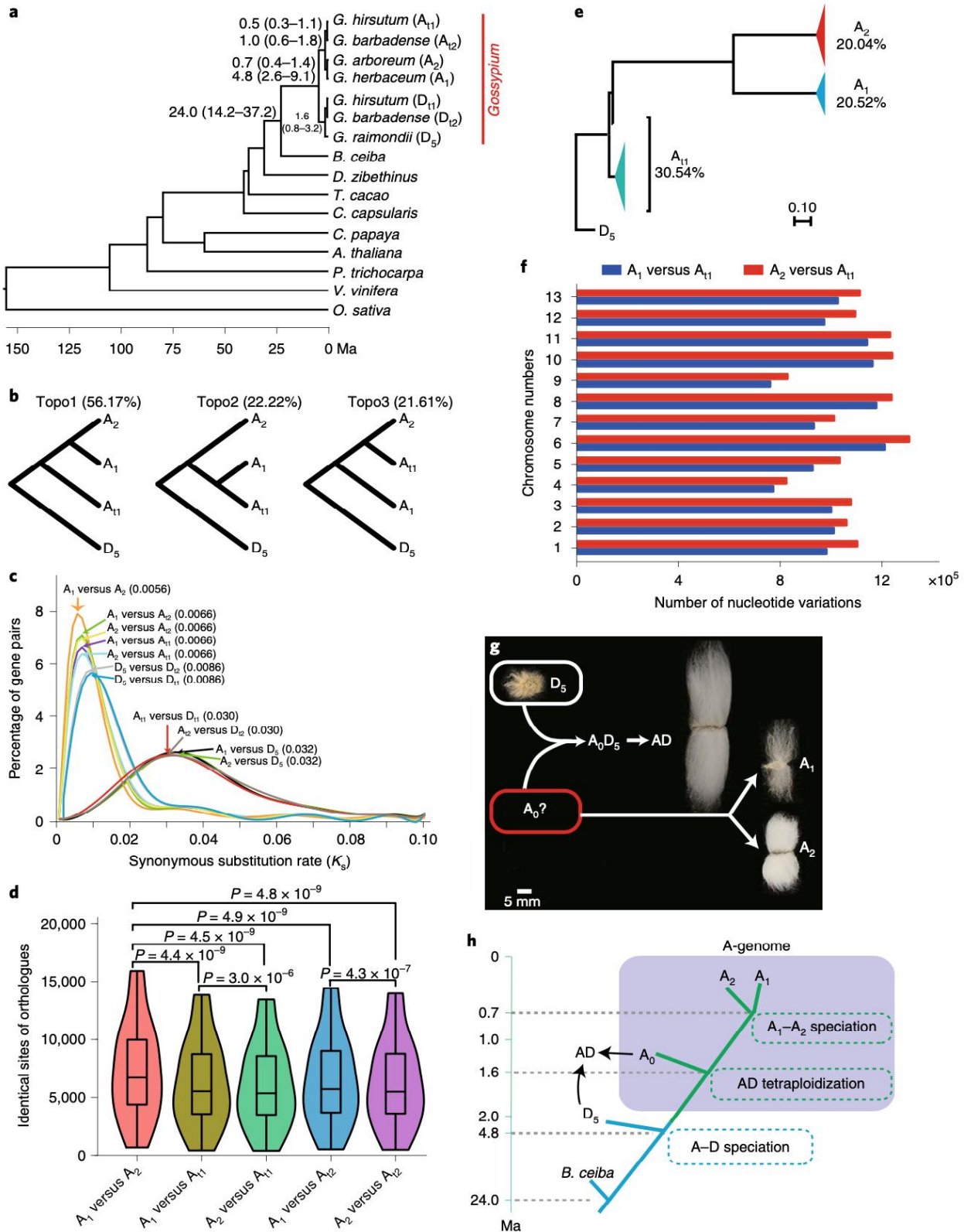


Figure 1 The evolution of the allotetraploid cotton genome (Adapted from Huang et al., 2020)

Image caption: (a) Inferred phylogenetic analysis among *Gossypium* and other eudicot plants. (b) Summary of phylogenetic analysis with the approximately unbiased test in 10-kb windows. (c) Distribution of K_s values for orthologous genes among cotton genomes. Peak values for each comparison are indicated in the parentheses. (d) Comparisons of identical sites in orthologous genes. Violin plots summarize the distribution of identical sites. The center line in each box indicates the median, and the box limits indicate the upper and lower quartiles of divergence ($n = 20$ types of synonymous mutation). P values were derived with Student's t -test. (e) Phylogenetic and ancestral allele analysis based on SNPs. The red, blue and green triangles represent the collapsed 21 A₂ accessions,

14 A1 accessions and 30 (AD)1 accessions, respectively. The percentage value indicates the percentage of ancestral alleles for each species that were identical to those of the D5-genome. (f) Number of nucleotide variations in A1 or A2 compared with At1 across the chromosomes. (g) A model for the formation of allotetraploid cotton showing fiber phenotypes from the (AD)1 (accession TM-1), the D5, the A1 (var. africanum) and the A2 (cv. Shixiya1). Scale bar, 5 mm. h, A schematic map of the evolution of cotton genomes. Major evolutionary events are shown in dashed boxes (Adapted from Huang et al., 2020)

Comparative genomics studies using these reference genomes have provided valuable insights into the evolution and domestication of cotton. These studies have revealed the asymmetric evolution between the A and D subgenomes of tetraploid cotton, highlighting the genetic basis of fiber improvement and stress tolerance (Zhang et al., 2015).

4.3 Advances in CRISPR/Cas9 genome editing

The advent of CRISPR/Cas9 genome editing technology has opened new avenues for genetic improvement in cotton. This technology allows for precise, targeted modifications of the cotton genome, enabling the introduction of desirable traits and the elimination of undesirable ones. The high efficiency and specificity of CRISPR/Cas9 make it an invaluable tool for cotton breeding.

In recent years, significant progress has been made in applying CRISPR/Cas9 to edit the genomes of allotetraploid cotton species. Researchers have successfully used CRISPR/Cas9 to target and modify multiple sites within the cotton genome, achieving high rates of gene editing efficiency. For example, studies have demonstrated the successful editing of genes involved in fiber development and stress response, resulting in improved fiber quality and enhanced resistance to biotic and abiotic stresses (Lu et al., 2019).

CRISPR/Cas9 technology has also been used for gene knockout and introduction of specific mutations in cotton, providing valuable insights into gene function and regulation. The ability to accurately edit cotton genes accelerates functional genomics research, enabling researchers to analyze complex genetic pathways and identify key regulatory genes. Wang et al. (2017) introduced a mutation in the *DsRed2* gene in cotton variety YZ1 using the CRISPR/Cas9 system. Figure 2 shows the experimental process and results. The upper row (a) displays wild-type cotton seeds (upper row) and *DsRed2* overexpressing lines (lower row). By comparing the two, it can be observed that the mutant exhibits differences in performance under different lighting conditions. In addition, the development of multiple CRISPR/Cas9 systems has made it possible to edit multiple genes simultaneously, further improving the efficiency of cotton genome editing. This method performs well in targeting gene families and regulatory networks, promoting the engineering improvement of complex cotton traits (Wang et al., 2017).

5 Implications for Cotton Breeding

5.1 Enhancements in genetic diversity and trait selection

Genome sequencing technology provides powerful tools for enhancing genetic diversity and trait selection in cotton. High-throughput sequencing technologies enable comprehensive analysis of the complex structure and variation within the cotton genome. These technologies make it possible to dissect the genome in detail, revealing genes and regulatory networks associated with important agronomic traits. For example, a study on the cotton genome revealed extensive structural variations that significantly impact fiber quality and disease resistance (Wang et al., 2018).

Genome sequencing has highlighted the loss of genetic diversity during cotton domestication. Through comprehensive genomic variation analysis, researchers have identified significant differences between cultivated and wild cotton, particularly in gene loss and retention. These findings provide new perspectives on the selective pressures during domestication and improvement processes, aiding in the development of more effective breeding strategies to increase genetic diversity in cotton (Li et al., 2021).

5.2 Marker-Assisted selection and its benefits

Marker-assisted selection (MAS) is a significant application of genome sequencing in cotton breeding. MAS uses genomic markers to track and select genes associated with specific traits, making the breeding process more efficient and precise. Genome sequencing has provided a vast array of single nucleotide polymorphism (SNP)

markers, widely used to construct high-density genetic maps and conduct quantitative trait loci (QTL) analysis. For instance, a study used a high-density genetic map for QTL analysis of boll weight in cotton, identifying multiple stable QTLs that showed significant phenotypic variation across different environments (Zhang et al., 2016).

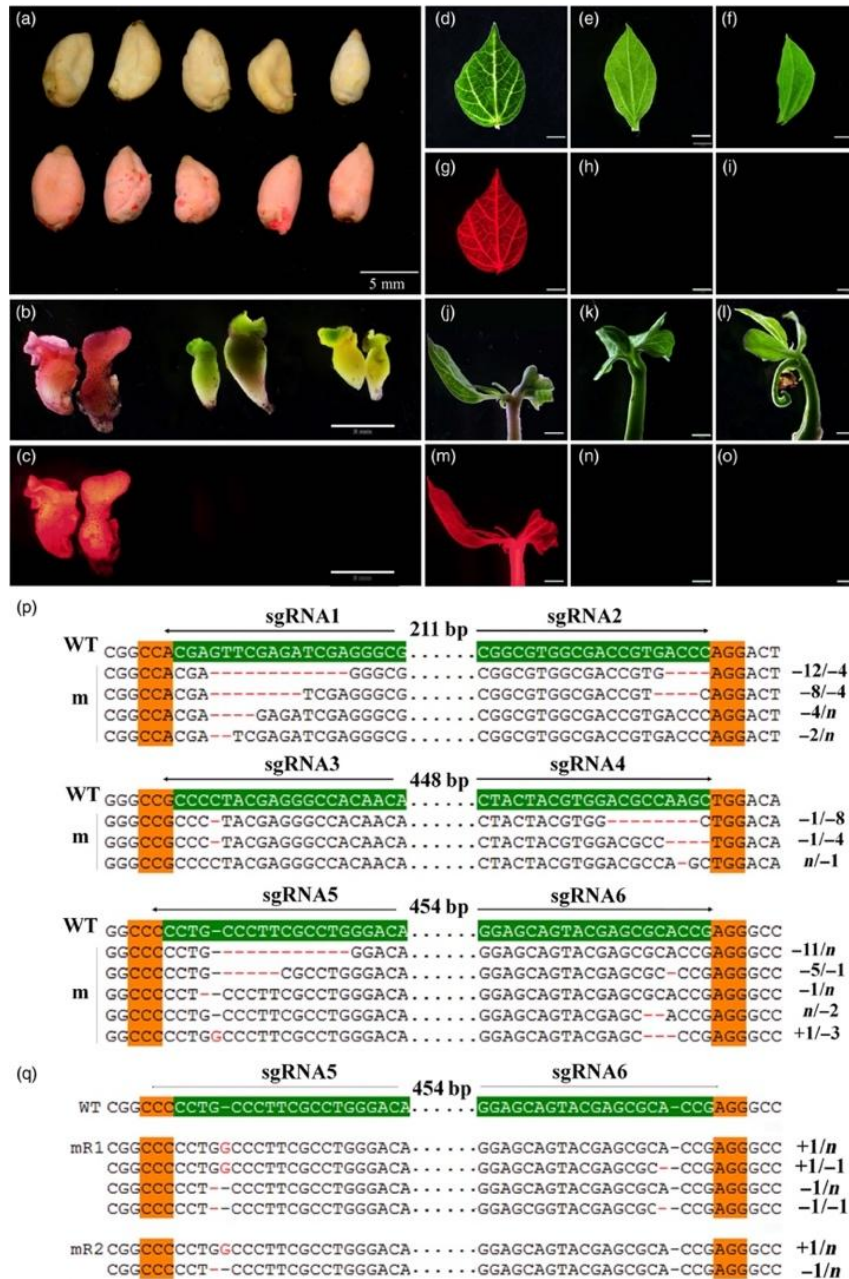


Figure 2 DsRed2 mutation is induced by the CRISPR/Cas9 system (Adapted from Wang et al., 2017)
 Image caption: (a) Seeds of wild-type cotton YZ1 (upper row) and a DsRed2 overexpression line (bottom row). (b) and (c) Regenerated somatic embryos of the control line and two mutants (mR1 and mR2) in the white light field (b) and a red fluorescence field at an excitation wavelength of 530 to 550 nm (c). (d) to (o) Leaves and young seedlings from corresponding plants in (b) were observed in the white light field (d, e, f, j, k, l) and the red fluorescence field (g, h, i, m, n, o). Bar in (a) is 5 mm, in (b) to (o) is 2 mm. Sanger sequencing of somatic embryos (p) and two independent mutants (q) at the DsRed2 target sites are exhibited. The sgRNA target sites are highlighted in green background. PAM regions are highlighted in orange. Nucleotide deletions or insertions are shown in red, with details labelled at right. The gaps between the paired sgRNAs are in dotted line, and their lengths are labelled above. WT, the wild type. m, mutation clones (Adapted from Wang et al., 2017)

The application of MAS significantly increases breeding efficiency, enabling the rapid accumulation and fixation of target traits within a shorter timeframe. By using MAS, researchers can effectively select individuals with

desirable traits, accelerating the development of new varieties. For example, MAS has been used to select cotton varieties with improved disease resistance and higher yields, enhancing both productivity and resistance to environmental stresses.

5.3 Development of disease-resistant and high-yielding varieties

Genome sequencing offers new possibilities for developing disease-resistant and high-yielding cotton varieties. By deeply analyzing the cotton genome, researchers can identify and edit key genes associated with disease resistance and yield. For example, CRISPR/Cas9 genome editing technology has been successfully applied to the cotton genome for targeted mutations, creating new cotton varieties with improved traits (Li et al., 2017).

The application of genome editing technologies enables precise modifications to the cotton genome, enhancing the accuracy and efficiency of breeding. By knocking out or introducing specific genes through genome editing, significant improvements in disease resistance and yield can be achieved. Moreover, genome sequencing has revealed genes associated with fiber quality, plant architecture, and stress tolerance, providing new strategies for comprehensive trait improvement in cotton (Peng et al., 2020).

6 Case Study: Application of Genome Sequencing in Cotton Breeding

6.1 Case study on *Verticillium wilt* resistance

Verticillium wilt, caused by the soil-borne pathogen *Verticillium dahliae*, is a significant disease affecting cotton, leading to substantial yield losses globally. Traditional breeding methods for identifying and utilizing resistance genes are time-consuming and complex. With the advancement of genome sequencing technologies, researchers can precisely locate and identify genes associated with disease resistance, thereby accelerating the development of resistant varieties.

Li et al. (2017) conducted a genome-wide association study (GWAS) on 299 cotton accessions using 85 630 single nucleotide polymorphism (SNP) markers. They identified 17 SNP markers significantly associated with *Verticillium wilt* resistance. Further haplotype block structure analysis predicted 22 candidate genes linked to the significant SNP A10_99672586 on chromosome A10. Among these genes, *CG02* was significantly upregulated in resistant genotypes and downregulated in susceptible ones. Quantitative real-time PCR and virus-induced gene silencing (VIGS) analyses revealed that silencing *CG02* increased the susceptibility of cotton plants to *Verticillium wilt*, indicating *CG02* as a crucial resistance gene (Li et al., 2017).

Zhao et al. (2021) further validated these results by combining GWAS, QTL-seq, and transcriptome sequencing. Using the Cotton 63K Illumina Infinium SNP array on 120 core cotton accessions, they identified five significant QTLs that overlapped with previously reported QTLs. Integrating GWAS, QTL-seq, and transcriptome sequencing, they identified eight candidate genes with genomic DNA sequence variations and expression differences between resistant and susceptible accessions. Most of these genes were involved in transcription factor activity, flavonoid biosynthesis, and plant innate immunity. Additionally, they developed 10 KASP markers, which were successfully validated in different cotton varieties and can be used for marker-assisted selection (MAS) to enhance *Verticillium wilt* resistance (Zhao et al., 2021).

6.2 Case study on fiber quality improvement in *Gossypium barbadense*

Gossypium barbadense is widely cultivated for its superior fiber quality, with key attributes such as length, strength, and fineness. Advances in genome sequencing have enabled researchers to gain deeper insights into the genetic factors influencing these fiber traits, providing new breeding strategies for fiber quality improvement.

Liu et al. (2015) sequenced the genome of *G. barbadense*, revealing genes associated with fiber development and secondary cell wall biosynthesis. Through genome sequencing and transcriptome analysis, they identified candidate genes playing key roles in fiber elongation and cell wall thickening. These genes included enzymes involved in cellulose and lignin synthesis, as well as transcription factors regulating fiber development. Functional validation of these genes showed that overexpression or silencing of specific genes could significantly impact fiber length and strength (Liu et al., 2015).

Wang et al. (2017) conducted a study using high-throughput sequencing technology and GWAS to analyze fiber quality traits in *G. barbadense* across multiple environments. They identified several QTLs significantly associated with fiber length and strength, explaining major variations in these traits. Further analysis of these QTL regions identified a series of candidate genes that were highly expressed in high-quality fiber varieties. Using gene editing technologies such as CRISPR/Cas9, researchers can target these genes to directly improve fiber traits (Wang et al., 2017).

6.3 Case study on boll weight QTL mapping

Boll weight is a critical trait directly affecting the total yield of cotton. Traditional breeding methods for improving boll weight often rely on phenotypic selection, which is less efficient and time-consuming. Advances in genome sequencing technologies allow researchers to more accurately locate and identify QTLs associated with boll weight, providing support for efficient breeding.

Fan et al. (2018) used specific locus amplified fragment sequencing (SLAF-seq) technology to construct a high-density genetic map and identified 18 stable QTLs for boll weight in upland cotton (*Gossypium hirsutum*). These QTLs were consistently detected across multiple environments, explaining a significant portion of the phenotypic variation in boll weight. And a comprehensive analysis was conducted on QTLs related to fiber quality and lint yield traits, which are crucial for cotton breeding. Figure 3 shows the QTLs for various traits identified in their study. The marker information from these QTLs provides important support for marker-assisted selection (MAS), enabling breeders to select individuals with desirable boll weight traits earlier and more accurately in the breeding process (Fan et al., 2018).

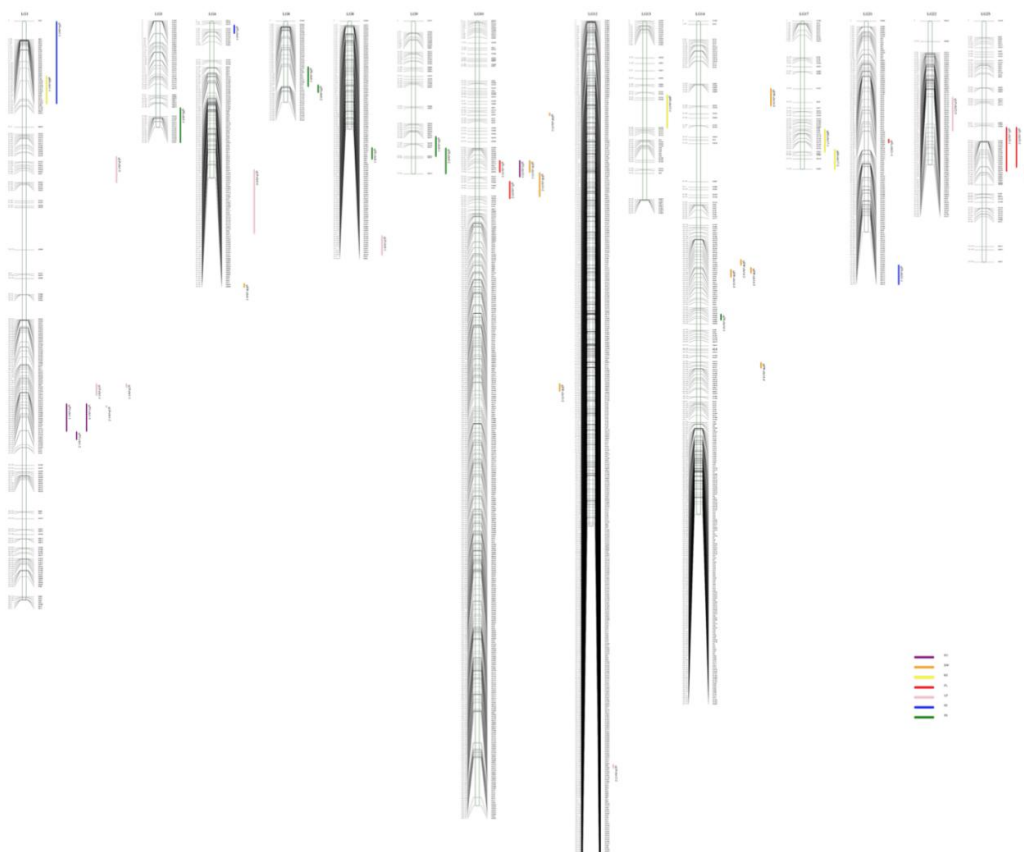


Figure 3 QTL associated with fiber quality and lint yield traits (Adapted from Fan et al., 2018)

Image caption: The figure demonstrates the genomic locations of QTLs associated with these traits, providing a visual representation of the genetic architecture underlying fiber quality and yield in cotton. These insights facilitate a more targeted approach to cotton breeding, allowing breeders to combine multiple desirable traits in new cultivars (Adapted from Fan et al., 2018)

Abdelraheem et al. (2019) conducted a study using high-throughput genome sequencing and GWAS to explore boll weight and yield traits in Upland cotton varieties in the USA. They evaluated the resistance and yield traits of

376 Upland cotton varieties in six independent replicated greenhouse trials, identifying 30 significant SNP markers associated with boll weight and disease resistance. These markers were distributed on five chromosomes and showed consistency across multiple environments. The results indicated that these QTL regions are rich in NBS-LRR genes, which play important roles in plant disease resistance (Abdelraheem et al., 2019).

7 Challenges and Limitations

7.1 Technical challenges in sequencing and assembly

The process of sequencing and assembling the cotton genome presents several technical challenges due to its complex and highly repetitive nature. Cotton genomes, particularly those of allotetraploid species such as *Gossypium hirsutum*, contain two sets of homologous chromosomes (At and Dt subgenomes), which complicates the sequencing and assembly process. For instance, the presence of extensive structural rearrangements and sequence divergence between the subgenomes requires advanced sequencing technologies and sophisticated assembly algorithms to accurately reconstruct the genome (Wang et al., 2018).

Another significant challenge is the high content of repetitive elements, which constitute a large portion of the cotton genome. These repetitive sequences can lead to assembly errors and gaps, making it difficult to achieve high-quality and complete genome assemblies. Techniques such as single-molecule real-time (SMRT) sequencing, BioNano optical mapping, and high-throughput chromosome conformation capture have been employed to address these issues, but they still require substantial computational resources and expertise (Zhang et al., 2015).

Moreover, the integration of different sequencing platforms and data types to produce a coherent and contiguous genome assembly remains a technical challenge. The combination of short-read sequencing, long-read sequencing, and physical mapping techniques necessitates advanced bioinformatics tools and pipelines to accurately merge these datasets and resolve complex genomic regions (Peng et al., 2020).

7.2 Data management and bioinformatics hurdles

Managing and analyzing the vast amounts of data generated by genome sequencing projects is another major challenge. The storage, processing, and interpretation of high-throughput sequencing data require significant computational infrastructure and bioinformatics expertise. For example, handling large datasets from various sequencing technologies, such as Illumina, PacBio, and Oxford Nanopore, necessitates robust data management systems and efficient computational pipelines (Wang et al., 2019).

Bioinformatics hurdles include the development and maintenance of software tools capable of accurately assembling genomes, annotating genes, and identifying structural variants. The complexity of the cotton genome, with its high level of polyploidy and repetitive content, requires specialized algorithms to correctly assemble and annotate genomic regions. Additionally, the integration of multi-omics data, such as transcriptomics, proteomics, and metabolomics, poses further challenges in terms of data compatibility and analysis (Thyssen et al., 2018).

Furthermore, the annotation of functional elements within the genome, such as genes, regulatory elements, and non-coding RNAs, relies heavily on comparative genomics and the availability of well-annotated reference genomes. The ongoing refinement of bioinformatics tools and the development of standardized protocols for genome assembly and annotation are essential to overcome these hurdles (Fan et al., 2018).

7.3 Ethical and environmental considerations

The advancements in genome sequencing and editing technologies raise several ethical and environmental concerns. One of the primary ethical considerations is the potential for unintended consequences of genome editing, such as off-target effects and ecological impacts. The use of CRISPR/Cas9 and other genome editing tools to modify cotton genomes must be carefully regulated to prevent unintended genetic alterations that could affect non-target species or lead to unforeseen ecological disruptions (Li et al., 2017).

Environmental considerations include the potential impact of genetically modified cotton on biodiversity and ecosystem health. The introduction of genetically engineered cotton varieties with traits such as pest resistance or herbicide tolerance could lead to changes in agricultural practices and pest management strategies, with potential

consequences for non-target organisms and soil health. The development and deployment of genetically modified crops must be accompanied by rigorous environmental impact assessments and monitoring programs to mitigate potential risks (Yang et al., 2020).

The ethical implications of intellectual property rights and access to genetic resources must be addressed. The proprietary nature of certain genome editing technologies and the ownership of genetic sequences can restrict access to these resources for researchers and farmers in developing countries. Policies promoting equitable access to genetic resources and the benefits derived from genome sequencing are essential to ensure that advancements in cotton genomics benefit all stakeholders (Peng et al., 2020).

8 Future Directions in Cotton Genome Research

8.1 Potential of Synthetic biology in cotton improvement

Synthetic biology offers transformative potential for cotton improvement by enabling the design and construction of new biological parts, devices, and systems that do not naturally exist. Through synthetic biology, researchers can engineer cotton plants with desired traits such as enhanced fiber quality, disease resistance, and environmental stress tolerance. For example, the synthesis of novel metabolic pathways can lead to the production of high-value compounds or improve photosynthetic efficiency (Huang et al., 2021).

Recent advancements include the development of synthetic promoters and regulatory elements to fine-tune gene expression in cotton. This precise control over gene expression can optimize the production of desired traits while minimizing unintended effects. Additionally, the integration of synthetic biology with CRISPR/Cas9 genome editing technology has further enhanced the ability to make targeted modifications in the cotton genome (Yang et al., 2022).

Synthetic biology also holds promise for creating cotton varieties with improved environmental sustainability. For instance, synthetic pathways can be designed to enhance nitrogen use efficiency, reducing the need for synthetic fertilizers and mitigating environmental pollution. As synthetic biology continues to evolve, its application in cotton improvement is expected to bring about significant advancements in crop productivity and sustainability (Ashraf et al., 2018).

8.2 Integrating genomics with precision agriculture

The integration of genomics with precision agriculture represents a forward-looking approach to optimize cotton cultivation. Precision agriculture involves the use of technology to monitor and manage field variability in crops, leading to more efficient use of resources and increased crop yields. By combining genomic data with precision agriculture techniques, it is possible to develop cotton varieties that are specifically tailored to local environmental conditions and management practices (Wang et al., 2019).

One example of this integration is the use of genomic information to inform variable rate applications of inputs such as fertilizers, water, and pesticides. By understanding the genetic makeup of cotton plants and their responses to different inputs, farmers can apply resources more precisely, reducing waste and improving crop performance. Advanced genomic tools such as genome-wide association studies (GWAS) and genomic selection can identify key genetic markers associated with desirable traits, enabling the development of precision breeding programs (Zhang et al., 2019).

Additionally, integrating genomics with precision agriculture can help address challenges such as climate change and soil degradation. Genomic insights can guide the development of cotton varieties that are more resilient to abiotic stresses like drought and heat, while precision agriculture techniques can optimize the management of these varieties in the field. This holistic approach is poised to enhance the sustainability and profitability of cotton farming (Yu et al., 2015).

8.3 Future prospects for genome editing technologies

Genome editing technologies, particularly CRISPR/Cas9, have revolutionized plant genetics by providing precise tools to modify specific genes within the genome. The future prospects for genome editing in cotton are vast, with

potential applications in improving fiber quality, enhancing disease resistance, and increasing yield. CRISPR/Cas9 has already been successfully used to edit genes associated with important traits in cotton, demonstrating its efficacy and potential (Gao et al., 2017).

One promising area is the use of genome editing to develop cotton varieties with reduced gossypol content in seeds. Gossypol is a toxic compound that limits the use of cottonseed as a protein source for humans and non-ruminant animals. By precisely knocking out the genes involved in gossypol biosynthesis, researchers can create cotton varieties with low gossypol content in seeds while retaining the compound in other parts of the plant to deter pests.

Another future prospect is the enhancement of cotton's tolerance to biotic and abiotic stresses through genome editing. By targeting specific genes that confer resistance to pests, diseases, and environmental stresses, it is possible to develop robust cotton varieties that require fewer chemical inputs and are better adapted to changing climates (Li et al., 2018).

Furthermore, advancements in multiplex genome editing, where multiple genes can be edited simultaneously, hold promise for rapidly stacking desirable traits in cotton. This approach can accelerate the breeding process and enable the development of cotton varieties with a combination of traits that enhance productivity and quality (Wang et al., 2017).

9 Concluding Remarks

The advancements in genome sequencing have significantly enhanced our understanding of the cotton genome, paving the way for more effective cotton breeding programs. High-density genetic maps and the identification of quantitative trait loci (QTLs) for important traits such as fiber quality and boll weight have been crucial. For example, the construction of a high-density genetic map using specific locus amplified fragment sequencing (SLAF-seq) identified 18 stable QTLs for boll weight, providing valuable information for marker-assisted selection. Additionally, sequencing efforts in allotetraploid cotton have revealed structural rearrangements and gene loss, which are critical for understanding fiber improvement and stress tolerance.

The integration of multi-strategic RNA-seq analyses has further enriched our knowledge of the cotton transcriptome, revealing tissue-specific gene expression and regulatory mechanisms involved in fiber development. These findings collectively contribute to the development of superior cotton varieties through precise genetic modifications and informed breeding strategies.

Genome sequencing has transformative potential in agriculture, particularly in the enhancement of crop traits and the sustainability of farming practices. By providing detailed genetic information, genome sequencing enables the identification of beneficial alleles and the creation of genetically superior crops. For cotton, this includes improved fiber quality, increased yield, and enhanced resistance to diseases and environmental stresses.

The ability to sequence and analyze entire genomes allows for a more precise and comprehensive approach to crop improvement. This not only accelerates the breeding process but also reduces the reliance on trial-and-error methods, leading to more predictable and reliable outcomes. In the broader context of agriculture, genome sequencing can contribute to food security by developing crops that are resilient to climate change and capable of thriving in diverse environmental conditions.

The future of cotton genomics and breeding is poised for remarkable advancements, driven by continuous improvements in sequencing technologies and bioinformatics tools. The integration of synthetic biology and precision agriculture with genomics holds immense potential for creating high-performance cotton varieties tailored to specific growing conditions and market demands.

Looking forward, the focus will likely be on refining genome editing techniques, such as CRISPR/Cas9, to achieve more targeted and efficient modifications. This will enable the stacking of multiple beneficial traits, enhancing both productivity and quality. Additionally, the development of comprehensive genomic databases and collaborative platforms will facilitate the sharing of knowledge and resources, accelerating the pace of innovation

in cotton breeding. The combination of these advanced techniques and collaborative efforts will likely lead to the realization of "super cotton" varieties that are not only high-yielding and high-quality but also resilient to environmental challenges, ensuring a sustainable and profitable cotton industry for the future.

Acknowledgments

We would like to thank the peer reviewers for their valuable feedback and suggestions on my research.

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

- Abdelraheem A., Ellassbli H., Zhu Y., Kuraparthi V., Hinze L., Stelly D., Wedegaertner T., and Zhang J., 2019, 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, *Theoretical and Applied Genetics*, 133: 563-577.
<https://doi.org/10.1007/s00122-019-03487-x>
- Ashraf J., Zuo D., Wang Q., Malik W., Zhang Y., Abid M., Cheng H., Yang Q., and Song G., 2018, Recent insights into cotton functional genomics: progress and future perspectives, *Plant Biotechnology Journal*, 16: 699-713.
<https://doi.org/10.1111/pbi.12856>
- Dai F., Chen J., Zhang Z., Liu F., Li J., Zhao T., Hu Y., Zhang T., and Fang L., 2022, COTTONOMICS: a comprehensive cotton multi-omics database, *Database: The Journal of Biological Databases and Curation*, 2022: 1-8.
<https://doi.org/10.1093/database/baac080>
- Du X., Huang G., He S., Yang Z., Sun G., Ma X., Li N., Zhang X., Sun J., Liu M., Jia Y., Pan Z., Gong W., Liu Z., Zhu H., Ma L., Liu F., Yang D., Wang F., Fan W., Gong Q., Peng Z., Wang L., Wang X., Xu S., Shang H., Lu C., Zheng H., Huang S., Lin T., Zhu Y., and Li F., 2018, Resequencing of 243 diploid cotton accessions based on an updated a genome identifies the genetic basis of key agronomic traits, *Nature Genetics*, 50: 796-802.
<https://doi.org/10.1038/s41588-018-0116-x>
- Fan L., Wang L., Wang X., Zhang H., Zhu Y., Guo J., Gao W., Geng H., Chen Q., and Qu Y., 2018, A high-density genetic map of extra-long staple cotton (*Gossypium barbadense*) constructed using genotyping-by-sequencing based single nucleotide polymorphic markers and identification of fiber traits-related QTL in a recombinant inbred line population, *BMC Genomics*, 19: 489.
<https://doi.org/10.1186/s12864-018-4890-8>
- Gao W., Long L., Tian X., Xu F., Liu J., Singh P., Botella J., and Song C., 2017, Genome editing in cotton with the CRISPR/Cas9 system, *Frontiers in Plant Science*, 8: 1364.
<https://doi.org/10.3389/fpls.2017.01364>
- He S., Sun G., Geng X., Gong W., Dai P., Jia Y., Shi W., Pan Z., Wang J., Wang L., Xiao S., Chen B., Cui S., You C., Xie Z., Wang F., Sun J., Fu G., Peng Z., Hu D., Wang L., Pang B., and Du X., 2021, The genomic basis of geographic differentiation and fiber improvement in cultivated cotton, *Nature Genetics*, 53: 916-924.
<https://doi.org/10.1038/s41588-021-00844-9>
- Huang G., Huang J., Chen X., and Zhu Y., 2021, Recent advances and future perspectives in cotton research, *Annual Review of Plant Biology*, 72: 437-462.
<https://doi.org/10.1146/annurev-arplant-080720-113241>
- Huang G., Wu Z., Percy R., Bai M., Li Y., Frelichowski J., Hu J., Wang K., Yu J., and Zhu Y., 2020, Genome sequence of *Gossypium herbaceum* and genome updates of *Gossypium arboreum* and *Gossypium hirsutum* provide insights into cotton A-genome evolution, *Nature Genetics*, 52: 516-524.
<https://doi.org/10.1038/s41588-020-0607-4>
- Li C., Fu Y., Sun R., Wang Y., and Wang Q., 2018, Single-Locus and multi-locus genome-wide association studies in the genetic dissection of fiber quality traits in upland cotton (*Gossypium hirsutum* L.), *Frontiers in Plant Science*, 9: e86308.
<https://doi.org/10.3389/fpls.2018.01083>
- Li C., Unver T., and Zhang B., 2017, A high-efficiency CRISPR/Cas9 system for targeted mutagenesis in Cotton (*Gossypium hirsutum* L.), *Scientific Reports*, 7: 43902.
<https://doi.org/10.1038/srep43902>
- Li C., Wang Y., Ai N., Li Y., and Song J., 2018, A genome-wide association study of early-maturation traits in upland cotton based on the Cotton SNP80K array, *Journal of Integrative Plant Biology*, 60(10): 970-985.
<https://doi.org/10.1111/jipb.12673>
- Li F., Fan G., Lu C., Xiao G., Zou C., Kohel R., Ma Z., Shang H., Ma X., Wu J., Liang X., Huang G., Percy R., Liu K., Yang W., Chen W., Du X., Shi C., Yuan Y., Ye W., Liu X., Zhang X., Liu W., Wei H., Wei S., Huang G., Zhang X., Zhu S., Zhang H., Sun F., Wang X., Liang J., Wang J., He Q., Huang L., Wang J., Cui J., Song G., Wang K., Xu X., Yu J., Zhu Y., and Yu S., 2015, Genome sequence of cultivated Upland cotton (*Gossypium hirsutum* TM-1) provides insights into genome evolution, *Nature Biotechnology*, 33: 524-530.
<https://doi.org/10.1038/nbt.3208>

- Li J., Yuan D., Wang P., Wang Q., Sun M., Liu Z., Si H., Xu Z., Ma Y., Zhang B., Pei L., Tu L., Zhu L., Chen L., Lindsey K., Zhang X., Jin S., and Wang M., 2021, Cotton pan-genome retrieves the lost sequences and genes during domestication and selection, *Genome Biology*, 22: 119.
<https://doi.org/10.1186/s13059-021-02351-w>
- Li T., Ma X., Li N., Zhou L., Liu Z., Han H., Gui Y., Bao Y., Chen J., Dai X., 2017, Genome-wide association study discovered candidate genes of *Verticillium wilt* resistance in upland cotton (*Gossypium hirsutum* L.), *Plant Biotechnol. J.*, 15(12): 1520-1532.
<https://doi.org/10.1111/pbi.12734>
- Liu X., Zhao B., Zheng H., Hu Y., Lu G., Yang C., Chen J., Chen J., Chen D., Zhang L., Zhou Y., Wang L., Guo W., Bai Y., Ruan J., Shangguan X., Mao Y., Shan C., Jiang J., Zhu Y., Jin L., Kang H., Chen S., He X., Wang R., Wang Y., Chen J., Wang L., Yu S., Wang B., Wei J., Song S., Lu X., Gao Z., Gu W., Deng X., Ma D., Wang S., Liang W., Fang L., Cai C., Zhu X., Zhou B., Chen Z., Xu S., Zhang Y., Wang S., Zhang T., Zhao G., and Chen X., 2015, *Gossypium barbadense* genome sequence provides insight into the evolution of extra-long staple fiber and specialized metabolites, *Scientific Reports*, 5: 14139.
<https://doi.org/10.1038/srep14139>
- Lu H., Cui X., Zhao Y., Magwang R., Li P., Cai X., Zhou Z., Wang X., Liu Y., Xu Y., Hou Y., Peng R., Wang K., and Liu F., 2019, Identification of a genome-specific repetitive element in the *Gossypium* D genome, *Peer J.*, 8: e8344.
<https://doi.org/10.7717/peerj.8344>
- Ma X., and Cao X., 2018, Cotton variant genomes—a breakthrough in population genetics analysis, *Science China Life Sciences*, 61: 869-870.
<https://doi.org/10.1007/s11427-018-9326-4>
- Pan Y., Meng F., and Wang X., 2020, Sequencing multiple cotton genomes reveals complex structures and lays foundation for breeding, *Frontiers in Plant Science*, 11: 560096.
<https://doi.org/10.3389/fpls.2020.560096>
- Peng R., Jones D.C., Liu F., and Zhang B., 2021, From sequencing to genome editing for cotton improvement, *Trends Biotechnol.*, 39(3): 221-224.
<https://doi.org/10.1016/j.tibtech.2020.09.001>
- 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: 989-999.
<https://doi.org/10.1007/s00122-018-3254-8>
- Wang K., Wang D., Zheng X., Qin A., Zhou J., Guo B., Chen Y., Wen X., Ye W., Zhou Y., and Zhu Y., 2019, Multi-strategic RNA-seq analysis reveals a high-resolution transcriptional landscape in cotton, *Nature Communications*, 10: 4714.
<https://doi.org/10.1038/s41467-019-12575-x>
- Wang K., Wang D., Zheng X., Qin A., Zhou J., Guo B., Chen Y., Wen X., Ye W., Zhou Y., and Zhu Y., 2019, Multi-strategic RNA-seq analysis reveals a high-resolution transcriptional landscape in cotton, *Nature Communications*, 10: 4714.
<https://doi.org/10.1038/s41467-019-12575-x>
- Wang M., Tu L., Lin M., Lin Z., Wang P., Yang Q., Ye Z., Shen C., Li J., Zhang L., Zhou X., Nie X., Li Z., Guo K., Ma Y., Huang C., Jin S., Zhu L., Yang X., Min L., Yuan D., Zhang Q., Lindsey K., and Zhang X., 2017, Asymmetric subgenome selection and cis-regulatory divergence during cotton domestication, *Nature Genetics*, 49: 579-587.
<https://doi.org/10.1038/ng.3807>
- Wang M., Tu L., Yuan D., Zhu D., Shen C., Li J., Liu F., Pei L., Wang P., Zhao G., Ye Z., Huang H., Yan F., Ma Y., Zhang L., Liu M., You J., Yang Y., Liu Z., Huang F., Li B., Qiu P., Zhang Q., Zhu L., Jin S., Yang X., Min L., Li G., Chen L., Zheng H., Lindsey K., Lin Z., Udall J., and Zhang X., 2018, Reference genome sequences of two cultivated allotetraploid cottons, *Gossypium hirsutum* and *Gossypium barbadense*, *Nature Genetics*, 51: 224-229.
<https://doi.org/10.1038/s41588-018-0282-x>
- Wang P., Zhang J., Sun L., Ma Y., Xu J., Liang S., Deng J., Tan J., Zhang Q., Tu L., Daniell H., Jin S., and Zhang X., 2017, High efficient multisites genome editing in allotetraploid cotton (*Gossypium hirsutum*) using CRISPR/Cas9 system, *Plant Biotechnology Journal*, 16: 137-150.
<https://doi.org/10.1111/pbi.12755>
- Wang X., Guo H., Wang J., Lei T., Liu T., Wang Z., Li Y., Lee T., Li J., Tang H., Jin D., and Paterson A., 2016, Comparative genomic de-convolution of the cotton genome revealed a decaploid ancestor and widespread chromosomal fractionation, *The New Phytologist*, 209(3): 1252-1263.
<https://doi.org/10.1111/nph.13689>
- Yang Z., Gao C., Zhang Y., Yan Q., Hu W., Yang L., Wang Z., and Li F., 2022, Recent progression and future perspectives in Cotton genomic breeding, *Journal of Integrative Plant Biology*, 65(2): 548-569.
<https://doi.org/10.1111/jipb.13388>
- Yang Z., Qanmber G., Wang Z., Yang Z., and Li, F., 2020, *Gossypium* genomics: trends, scope, and utilization for cotton improvement, *Trends in Plant Science*, 25(5): 488-500.
<https://doi.org/10.1016/j.tplants.2019.12.011>
- Yu J., Jung S., Cheng C., Lee T., Zheng P., Buble K., Crabb J., Humann J., Hough H., Jones D., Campbell J., Udall J., and Main D., 2015, Cotton Gen: The community database for cotton genomics, genetics, and breeding research, *Plants*, 10(12): 2805.
<https://doi.org/10.3390/plants10122805>
- Zhang K., Kuraparthi V., Fang H., Zhu L., Sood S., and Jones D., 2019, High-density linkage map construction and QTL analyses for fiber quality, yield and morphological traits using CottonSNP63K array in upland cotton (*Gossypium hirsutum* L.), *BMC Genomics*, 20: 889.
<https://doi.org/10.1186/s12864-019-6214-z>

Zhang T., Hu Y., Jiang W., Fang L., Guan X., Chen J., Zhang J., Saski C., Scheffler B., Stelly D., Hulse-Kemp A., Wan Q., Liu B., Liu C., Wang S., Pan M., Wang Y., Wang D., Ye W., Chang L., Zhang W., Song Q., Kirkbride R., Chen X., Dennis E., Llewellyn D., Peterson D., Thaxton P., Jones D., Wang Q., Xu X., Zhang H., Wu H., Zhou L., Mei G., Chen S., Tian Y., Xiang D., Li X., Ding J., Zuo Q., Tao L., Liu Y., Li J., Lin Y., Hui Y., Cao Z., Cai C., Zhu X., Jiang Z., Zhou B., Guo W., Li R., and Chen Z., 2015, Sequencing of allotetraploid cotton (*Gossypium hirsutum* L. acc. TM-1) provides a resource for fiber improvement, *Nature Biotechnology*, 33: 531-537.

<https://doi.org/10.1038/nbt.3207>

Zhang Z., Shang H., Shi Y., Huang L., Li J., Ge Q., Gong J., Liu A., Chen T., Wang D., Wang Y., Palanga K., Muhammad J., Li W., Lu Q., Deng X., Tan Y., Song W., Cai J., Li P., Rashid H., Gong W., and Yuan Y., 2016, Construction of a high-density genetic map by specific locus amplified fragment sequencing (SLAF-seq) and its application to quantitative trait loci (QTL) analysis for boll weight in upland cotton (*Gossypium hirsutum*), *BMC Plant Biology*, 16: 1-18.

<https://doi.org/10.1186/s12870-016-0741-4>

Zhao Y., Chen W., Cui Y., Sang X., Lu J., Jing H., Wang W., Zhao P., and Wang H., 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, *Theoretical and Applied Genetics*, 134: 1063-1081.

<https://doi.org/10.1007/s00122-020-03752-4>



Disclaimer/Publisher's Note

The statements, opinions, and data contained in all publications are solely those of the individual authors and contributors and do not represent the views of the publishing house and/or its editors. The publisher and/or its editors disclaim all responsibility for any harm or damage to persons or property that may result from the application of ideas, methods, instructions, or products discussed in the content. Publisher remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.
