



## OPEN ACCESS

## EDITED BY

Mehboob-ur-Rahman,  
National Institute for Biotechnology and  
Genetic Engineering, Pakistan

## REVIEWED BY

Sachin Teotia,  
Sharda University, India  
Zahoor Ahmad Mir,  
University of Manitoba, Canada  
Allah Bakhsh,  
University of the Punjab, Pakistan

## \*CORRESPONDENCE

Sultan Habibullah Khan,  
✉ sultan@uaf.edu.pk

RECEIVED 14 March 2024

ACCEPTED 27 August 2024

PUBLISHED 12 September 2024

## CITATION

Saleem MS, Khan SH, Ahmad A, Rana IA,  
Naveed ZA and Khan AI (2024) The 4Fs of  
cotton: genome editing of cotton for fiber,  
food, feed, and fuel to achieve zero hunger.  
*Front. Genome Ed.* 6:1401088.  
doi: 10.3389/fgeed.2024.1401088

## COPYRIGHT

© 2024 Saleem, Khan, Ahmad, Rana, Naveed  
and Khan. This is an open-access article  
distributed under the terms of the [Creative  
Commons Attribution License \(CC BY\)](#). 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.

# The 4Fs of cotton: genome editing of cotton for fiber, food, feed, and fuel to achieve zero hunger

Muhammad Sulyman Saleem<sup>1,2</sup>, Sultan Habibullah Khan<sup>1,2\*</sup>,  
Aftab Ahmad<sup>2,3</sup>, Iqrar Ahmad Rana<sup>1,2</sup>, Zunaira Afzal Naveed<sup>1,2</sup> and  
Azeem Iqbal Khan<sup>4</sup>

<sup>1</sup>Centre of Agricultural Biochemistry and Biotechnology (CABB), University of Agriculture Faisalabad, Faisalabad, Pakistan, <sup>2</sup>Center for Advanced Studies in Agriculture and Food Security (CAS-AFS), University of Agriculture Faisalabad, Faisalabad, Pakistan, <sup>3</sup>Department of Biochemistry, University of Agriculture Faisalabad, Faisalabad, Pakistan, <sup>4</sup>Department of Plant Breeding and Genetics, University of Agriculture Faisalabad, Faisalabad, Pakistan

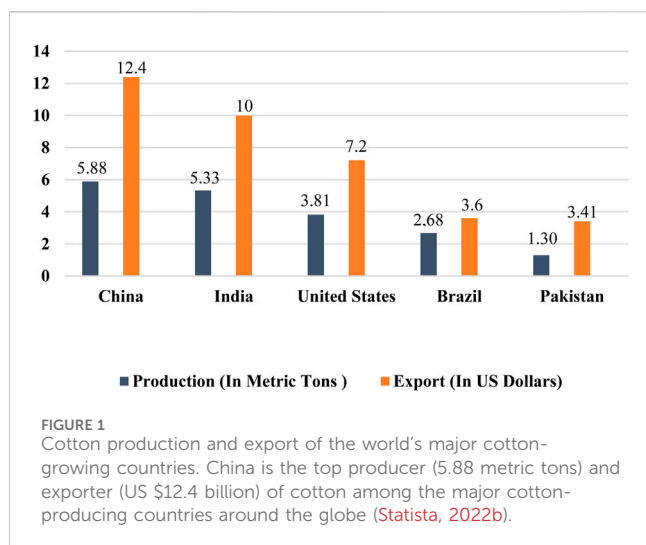
Cotton is globally known for its high-priority cellulose-rich natural fiber. In addition to providing fiber for the textile industry, it is an important source material for edible oil, livestock feed, and fuel products. Global warming and the growing population are the major challenges to the world's agriculture and the potential risks to food security. In this context, improving output traits in cotton is necessary to achieve sustainable cotton production. During the last few years, high throughput omics techniques have aided in identifying crucial genes associated with traits of cotton fiber, seed, and plant architecture which could be targeted with more precision and efficiency through the CIRPSR/Cas-mediated genome editing technique. The various CRISPR/Cas systems such as CRISPR/Cas9, CRISPR/nCas9, and CRISPR/Cas12a have been employed to edit cotton genes associated with a wide range of traits including fiber length, flowering, leaf colour, rooting, seed oil, plant architecture, gossypol content, somatic embryogenesis, and biotic and abiotic stresses tolerance, highlighting its effectiveness in editing the cotton genome. Thus, CRISPR/Cas-mediated genome editing has emerged as a technique of choice to tailor crop phenotypes for better yield potential and environmental resilience. The review covers a comprehensive analysis of cotton phenotypic traits and their improvement with the help of the latest genome editing tools to improve fiber, food, feed, and fuel-associated genes of cotton to ensure food security.

## KEYWORDS

cotton, fiber, food, feed, fuel, food security, genome editing, CRISPR/Cas systems

## 1 Economic importance of cotton in global food security

Cotton, the leading fiber-yielding crop, is commercially cultivated in approximately fifty countries across mild and hot climatic regions. Its cultivation serves to fulfill the demands of different industrial sectors worldwide. Because of the favorable climate conditions required for its natural growth, cotton is grown in various countries, including the USA, South Asia, East Asia, Central Asia, the Middle East, Australia, and Europe (Smith, 1999). The globally recognized dicotyledonous genus *Gossypium* includes approximately fifty species, but only

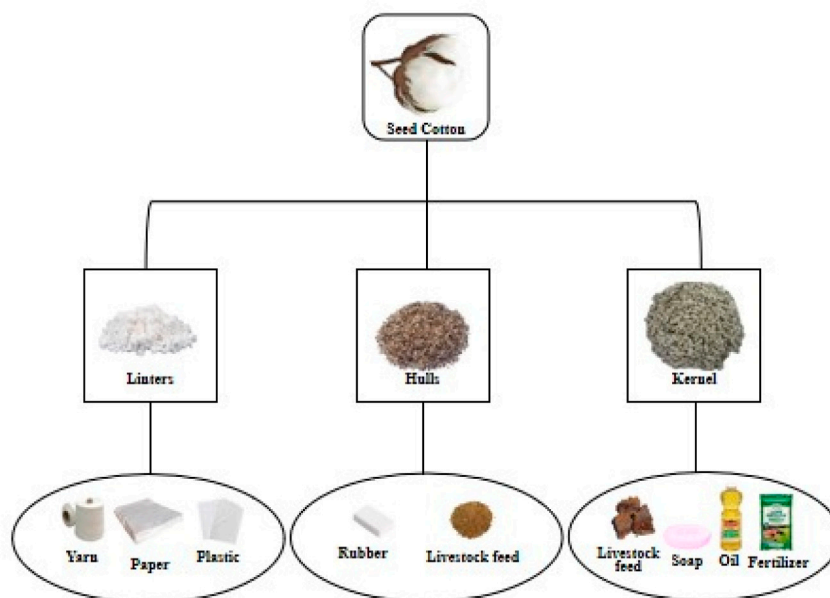


four species; *Gossypium arboreum* L., *Gossypium herbaceum* L., *Gossypium hirsutum* L., and *Gossypium barbadense* L. are commercially cultivated worldwide (Wendel and Grover, 2015). Tetraploid species (*G. hirsutum* L. and *G. barbadense* L.) dominate cultivation in over 80% of global cotton regions, while diploid species (*G. arboreum* L. and *G. herbaceum* L.) are limited to Asia and the Middle East.

The agriculture sector plays a pivotal role in the economies of several countries worldwide, with cotton commonly referred to as “white gold” because of its significant profitability for cotton-producing countries. The textile sector exerts a substantial economic impact and accounts for nearly 4% of the global market share, with the cotton part accounting for a share of 30.2% in the worldwide textile market (GITNEX, 2024). This sector is experiencing continuous expansion, mainly driven by the utilization of natural fibers supplied by the cotton crop. Cotton is a vital revenue source for nearly one billion individuals, including 250 million laborers within the cotton industry and 100 million farmers. Notably, about 90% of these farmers cultivate cotton on agricultural land of less than 2 hectares, primarily situated in developing countries (Aslam et al., 2020). Annually, around 25 million tons of cotton are produced globally, covering approximately 14 million acres of cotton-planted areas. The estimated value of this cotton production amounts to around 12 billion dollars (Ahmad and Hasanuzzaman, 2020). Furthermore, in 2021, the global export of cotton reached a total of US \$60.4 billion from all countries (Workman, 2022). In 2021, the top five cotton-producing countries were China, India, the United States, Brazil, and Pakistan according to Statista (2022b) (Figure 1). China, being a country with a massive population, holds the distinction of being the world's largest producer, consumer, and exporter of cotton. It produced approximately 5.88 million metric tons of cotton annually and exported cotton worth approximately US \$12.4 billion. Cotton cultivation and processing in China engage an estimated 300 million people, highlighting the substantial workforce dedicated to the industry. India, on the other hand, is the second-largest cotton producer globally, with an annual production of approximately 5.33 million metric tons. India's

cotton exports are valued at approximately US \$10 billion (Statista, 2022a). The USA ranks as the third-largest cotton producer globally, with an annual cotton production of about 3.81 million metric tons. The country's cotton fiber exports amount to a value of US \$7.2 billion each year. The overall worth of cotton fiber cultivated annually in the US is estimated to be around US \$6 billion, including an additional US \$500 million generated from cottonseed oil and its by-products (National Cotton Council of America, 2024). Brazil holds the position of the fourth-largest cotton producer globally, with an annual production of around 2.68 million metric tons. The export value of Brazilian cotton amounts to around US \$3.6 billion (Statista, 2022b). Pakistan is indeed recognized as the fourth-largest cotton producer globally, with an annual production of approximately 1.30 million metric tons and export earnings of around US \$3.41 billion from cotton exports. Cotton is a major cash crop in Pakistan and plays a vital role in providing essential income to the country's population. The economy significantly relies on it and the cotton-textile sectors. Pakistan's cotton textile industry, as the largest exporter of cotton yarn, contributes to 11% of the country's GDP and accounts for 60% of export income (Economic Survey of Pakistan, 2017). The progress made in the agriculture sector has shown to have a greater impact on enhancing the capability of countries to develop their economies and agricultural markets compared to development in other sectors. Cotton production generates income for nearly 250 million individuals worldwide and employs approximately 7% of the total labor force in developing countries. Hence, the agriculture sector plays a crucial role in accelerating economic growth towards achieving high-income status, while also addressing food insecurity and improving nutrition in developing countries.

In the present era, the global agricultural sector is confronted with two major challenges. Firstly, the world population is continuously growing and is predicted to surpass 9 billion by 2050 (United Nations, 2017). This population expansion poses essential concern in meeting the escalating demand for essential resources such as fiber, food, fuel, and feed (Silva, 2018; Parihar et al., 2022). Secondly, global warming poses a prominent threat to agricultural yields worldwide. Predictions indicate that the world's temperature could rise by 1.5°C by 2030 which has a considerable risk to agricultural productivity and the overall stability of the agricultural sector. Extreme global temperatures contribute to various detrimental effects on agricultural systems, leading to heat stress and exacerbating the intensity of abiotic and biotic stresses. These include the occurrence of severe and persistent droughts, soil salinization, altered rainfall patterns, increased weed growth, and a rise in the population of plant pathogens and insect pests (Arora, 2019). Indeed, the frequent climate changes, accompanied by water resource constraints and various abiotic and biotic stresses, have led to a decline in agronomic crop yield and quality. To address this challenge, it becomes imperative to enhance food production by two to three times and develop climate-resilient agronomic crops. Achieving food security requires increasing the income and productivity of farmers, reducing food prices, and improving overall nutrition. Therefore, it is crucial to prioritize the improvement in the yield and quality of agronomic crops to combat hunger and malnutrition, both in the present and the future.



**FIGURE 2**  
Illustrative overview of consumer products derived from seed cotton. Cotton fiber, cottonseed cake, and cottonseed oil are the main agricultural products, which contribute to cotton's economic value by offering key materials for various industries.

## 2 The 4Fs of cotton

Cotton stands as a primary agronomic crop cultivated worldwide due to its multifaceted uses in fiber, food, feed, and fuel (4Fs) (Yin and Ding, 2021). It serves as a valuable source of renewable resources, as ginned cottonseed comprises various components, including 16% crude oil and 46% meal in the kernel, along with 27% hull and 8% linters (Chaudry and Guitchounts, 2022). The 4Fs of cotton play a crucial role in enhancing global food security by empowering cotton-growing farmers economically. The influence of cotton production extends to multiple industrial sectors such as textiles, plastics, paper, livestock feed, soap, and oil (Figure 2).

### 2.1 Fiber

Cotton fiber is considered the purest and most vital naturally produced cellulosic fiber. Typically, a cotton fiber comprises approximately 80%–90% cellulose, 4%–6% hemicelluloses, 0.5%–1% waxes, 0%–1.5% proteins, 6%–8% water, and 8% ash, pectin, and fats (Cotton Incorporated, 2020). Because of its remarkable strength, durability, and comfort, cotton fiber is ideally suited for making clothes, housewares, and other textile products. The economic significance of cotton fiber in the global market is substantial, accounting for over 50% of textile products. The cotton fiber quality directly impacts the overall quality of cotton fiber-based garments, underscoring its significance. By 2050, it is predicted that approximately 8 billion people will require clothing made from cotton fiber, leading to a significant increase in the annual demand for cellulosic fiber, from 50 million to 130 million tons. So, the crucial role of cotton will continue to play in fulfilling the global need for natural fiber.

### 2.2 Food

Cotton also serves the edible oil industry, making about 6 million tons of cottonseed oil annually and producing around US \$4 billion in revenue (Statista, 2022a). Amid increasing edible oil costs and shortages, cottonseed oil is necessary for ensuring nutritious cooking oil, especially in developing countries. With 21% oil and 23% protein content, cottonseed can address the protein needs of almost a billion people, especially in malnourished regions such as Asia and Africa (Kumar et al., 2021). Post oil extraction, the high-protein meal can be used for flour or snacks, potentially fulfilling the daily protein needs of 575 million people. Globally, cottonseed yields 11 million tons of protein, sufficient for the annual requirements of around 600 million people. Rich in vitamins and minerals, cottonseed is significant for alleviating food crises and increasing nutrition without additional resources.

### 2.3 Feed

Cottonseed meal, a byproduct of oil extraction, is an important and cost-effective protein source for livestock feed, making up 30%–40% of the protein content, and increasing cattle production because of its high phosphorus content (Hernandez, 2016). Cottonseed hulls offer fiber-rich roughage for grain-based diets, even with low energy and digestibility (C H Knight Limited, 2024). Global animal feed production is approximately one billion tons annually, generating around US \$400 billion in revenue for the animal feed industry (International Feed Industry Federation, 2024). With increasing global requirements for livestock, fish, and dairy products driven by population growth and enhancing urban incomes, the FAO

predicts significant increases in meat (70%), dairy (55%), and aquaculture (90%) production. Cottonseed's high energy value makes it indispensable in meeting this rising demand.

## 2.4 Fuel

Energy is a fundamental requirement for human survival and economic development, and cotton-derived biofuel provides a significant renewable energy source. Now agricultural biomass accounts for around 15% of the world's primary energy (Bassam, 2001). In industrial countries, agricultural biomass produces "green electricity" and thermal energy, with crops converted into biofuels in liquid or gaseous forms (Luterbacher and Luterbacher, 2015). Cotton-derived biofuel offers renewable, eco-friendly alternatives to fossil fuels, ensuring future energy security and decreasing reliance on imports. However, the availability of cottonseed oil is restricted based on cotton production and seasonality, affecting biofuel supply consistency. The extraction and conversion process needs significant capital and technological advancements to be economically sustainable (Soomro et al., 2023). Cotton-derived biofuel competes with other biofuels and fossil fuels and faces challenges associated with land and water use, food security, biodiversity, public acceptance, and complex regulations (Soomro et al., 2023). Nonetheless, to overcome these challenges, cotton-derived biofuel holds the potential for sustainable energy.

## 3 Traits and genes associated with 4Fs of cotton

The yield of cotton cultivation encompasses two main components: the seed and the fiber, known as lint. These components have the potential to provide food, feed, fiber, and fuel, commonly referred to as the 4Fs. Understanding and characterizing cotton genes and their associated traits related to the 4Fs is crucial for comprehending the functional genetics of cotton and effectively improving traits associated with yield and quality. The advancements in molecular genetics, transgenic techniques, whole genome sequencing, and bioinformatics have revolutionized crop breeding studies, providing new tools that prioritize gene-based approaches for trait improvement. Therefore, knowledge about the genetic control of agronomical traits is indispensable for enhancing seed and fiber quality and overall yield in cotton. In this section, we discuss previous studies that have explored cotton gene function and genes with significant impacts on cotton yield and quality through upregulation or downregulation of their expression. Manipulating these genes presents an opportunity to improve cotton's desired output traits.

### 3.1 Fiber-related genes and traits

Cotton fiber is an immensely significant natural textile fiber globally. Each long cotton fiber originates from a single epidermal cell located on the surface of the ovule. Through the process of cell wall thickening, this cell transforms into an extremely elongated and durable dead fiber. The development of the fiber progresses through

distinct stages, including initiation, elongation, transitional primary wall remodeling, secondary wall synthesis, and maturation. These stages directly determine the quality properties of cotton fiber. The yield of fiber is influenced by the number of fibers originating from the outer surface of the ovule. Numerous studies have been conducted to identify and characterize genes associated with fiber development and their roles in improving fiber quality. Some genes exhibit preferential expression during the fiber initiation stage, while others show a high level of expression during the fiber elongation stage (Table 1).

The PROTODERMAL FACTOR1 gene (*GbPDF1*) is highly expressed in fiber cells during the fiber initiation stage. Its function is to maintain hydrogen peroxide homeostasis and regulate the biosynthesis of ethylene and pectin. This regulation is achieved through the interaction with a key cis-element called HDZIP2ATATHB2 (Deng et al., 2011). The *GhGAI3a* gene, encoding the DELLA protein, shows high expression at the fiber initiation stage in fiber cells. It functions as a repressor in the gibberellin (GA) signaling pathway, playing a role in regulating fiber development (Wang et al., 2012). The FASCICLIN-LIKE ARABINOGALACTAN protein encoded by the *GhFLA1* gene is involved in both the fiber initiation and elongation processes in cotton. It affects the composition of the ARABINOGALACTAN protein (AGP) and the primary cell wall (Huang et al., 2013). The *GhVIN1* gene encodes VACUOLAR INVERTASE (VIN), and it plays a crucial role in fiber initiation. It modulates the transcription of MYB transcription factors and auxin signaling constituents via VIN-based hexose signaling (Wang X. et al., 2014). The R3-MYB gene *GhCPC* encodes the caprice protein, which negatively regulates fiber initiation and early elongation. It potentially forms a complex with CPC-MYC1-TTG1/4 in cotton (Liu et al., 2015). The JASMONATE ZIM-DOMAIN 2 (JAZ2) protein, encoded by the *GhJAZ2* gene, acts as a primary transcription repressor during fiber initiation. It interacts with the *GhMYB-25* transcription factor to modulate the jasmonic acid (JA) signaling pathway (Hu et al., 2016). The *GhHDA5* gene, encoding histone deacetylase, is specifically expressed during fiber initiation and regulates the expression of fiber initiation-specific genes (Kumar et al., 2018). C<sub>2</sub>H<sub>2</sub>-zinc finger proteins encoded by C<sub>2</sub>H<sub>2</sub>-zinc finger genes play an indispensable role in regulating the fiber initiation and development process in cotton (Salih et al., 2019). The *GaHDI* gene, a HOMEODOMAIN-LEUCINE ZIPPER gene, plays a crucial role in cotton fiber initiation. It regulates the signaling cascade involved in hydrogen peroxide (H<sub>2</sub>O<sub>2</sub>) production and calcium ion (Ca<sup>2+</sup>) flux (Ding et al., 2020).

In a study conducted by Li et al. (2005), it was demonstrated that silencing the ACTIN (*GhACT1*) gene through RNA interference (RNAi) in *G. hirsutum* resulted in inhibited fiber elongation. This inhibition was attributed to the disruption of the actin cytoskeleton in fiber cells during the process of fiber development. Sun and Allen (2005) found that the shaggy-like protein kinase family (Bin 2) genes exert a negative regulatory effect on cotton fiber elongation by modulating brassinosteroid (BR) signaling. According to Wang et al. (2005), the *GhPFN1* gene in cotton was found to play a crucial role in facilitating rapid fiber elongation by triggering actin polymerization. Machado et al. (2009) revealed the significance of *GhMYB-25*, a member of the MYB family transcription factors, in regulating cotton fiber elongation. Li et al. (2011) discovered that a

TABLE 1 A list of genes and traits associated with fiber.

Gene	Specific trait	References
Actin1 ( <i>GhACT1</i> )	Fiber elongation	Li et al. (2005)
<i>GhBIN 2</i>	Fiber elongation	Sun et al. (2023)
Profilin 1 ( <i>GhPFN1</i> )	Fiber elongation	Wang et al. (2005)
<i>GhMYB-25<sup>a</sup></i>	Fiber elongation	Machado et al. (2009)
<i>GhMADS11</i>	Fiber elongation	Li et al. (2011)
Sucrose synthase ( <i>GhSUSAI</i> )	Fiber length and strength	Jiang et al. (2012)
Protodermal factor 1 ( <i>GBPDF1</i> )	Fiber initiation	Deng et al. (2011)
<i>GhGAI3A</i>	Fiber initiation	Wang et al. (2012)
Fasciclin-like arabinogalactan ( <i>GhFLA1</i> )	Fiber initiation and elongation	Huang et al. (2013)
Plasma membrane intrinsic protein ( <i>GhPIP2</i> )	Fiber elongation	Li et al. (2013)
Teosinte-branched1/cycloidea/pcf (tcp) protein ( <i>GhTCP 14</i> )	Fiber elongation	Wang et al. (2013)
Proline-rich protein ( <i>GhPRP5</i> )	Fiber length	Xu et al. (2013)
Vacuolar invertase ( <i>GhVIN1</i> )	Fiber initiation	Wang et al. (2014b)
<i>GhHOX3</i>	Fiber elongation	Shan et al. (2014)
Pagoda 1 ( <i>PAG1</i> )	Fiber elongation	Yang et al. (2014)
<i>GhKNL1</i>	Fiber elongation	Gong et al. (2014)
Phytochrome apoprotein ( <i>GhPHYA1</i> )	Fiber length	Abdurakhmonov et al. (2014)
Caprice ( <i>GhCPC</i> )	Fiber initiation and elongation	Liu et al. (2015)
Calmodulin ( <i>GhCAM7</i> )	Fiber elongation	Cheng et al. (2016)
Jasmonate zim-domain 2 ( <i>GhJAZ2</i> )	Fiber initiation	Hu et al. (2016)
B-galactosyltransferase 1 ( <i>GhGALT1</i> )	Fiber length	Qin et al. (2017)
<i>GhFSN1</i>	Fiber strength	Zhang et al. (2017b)
Histone deacetylase ( <i>GhHDA5</i> )	Fiber initiation	Kumar et al. (2018)
Ubiquitin ligase ( <i>GhHUB2</i> )	Fiber elongation	Feng et al. (2018)
C2H2-zinc finger	Fiber initiation	Salih et al. (2019)
<i>GhBHLH18</i>	Fiber strength and length	Gao et al. (2019a)
<i>GaHD1</i>	Fiber initiation	Ding et al. (2020)
Alanine rich protein ( <i>GhALARP<sup>a</sup></i> )	Fiber elongation	Zhu et al. (2021a)
Phosphoinositide-specific phospholipase C ( <i>GhPIPLC2D</i> )	Fiber length	Zhu et al. (2021b)
Transparent testa 2 ( <i>GhTT2-3A</i> )	Fiber color	Yan et al. (2018)
4-coumarate: coenzyme a ligase ( <i>Gh4CLA</i> )	Fiber color	Sun et al. (2019)
Chalcone synthases ( <i>GhCHS</i> )	Fiber color	Gao et al. (2019b)
Anthocyanidin reductases ( <i>GhANR</i> )	Fiber color	Gao et al. (2019a)
Leucoanthocyanidin reductases ( <i>GhLAR</i> )	Fiber color	Gao et al. (2019b)
Chalcone isomerase ( <i>GhCHI</i> )	Fiber color	Lv et al. (2023)
Flavanone 3-hydroxylases ( <i>GhF3H</i> )	Fiber color	Lv et al. (2023)
Flavonoid 3', 5'-hydroxylases ( <i>GhF3'5'H</i> )	Fiber color	Lv et al. (2023)
Dihydroflavonol 4-reductases ( <i>GhDFR</i> )	Fiber color	Lv et al. (2023)

<sup>a</sup>CRISPR-edited gene.

MADS-box protein encoded by the *GhMADS11* gene also exerts control over the elongation process of cotton lint fibers. A study indicated that the SUCROSE SYNTHASE (*GhSusA1*) gene is pivotal in determining cotton fiber length and strength by regulating the cell wall thickness during the secondary cell wall development stage (Jiang et al., 2012). In the study by Li et al. (2013), they identified the preferentially expressed *GhPIP2* genes in cotton fibers, which encode an aquaporin known as plasma membrane intrinsic proteins (PIPs). The suppression of *GhPIP2* gene expression through RNAi significantly inhibited fiber elongation, indicating their crucial role in this stage of cotton fiber development. Wang et al. (2013) discovered that a TEOSINTE-BRANCHED1/CYCLOIDEA/PCF (TCP) transcription factor encoded by the *GhTCP14* gene in upland cotton (*Gossypium hirsutum*) serves as a key regulator in the auxin-mediated development and rapid elongation of fiber cells. It achieves this by interacting with the promoters of *GhAUX*, *GhPIN2*, and *GhIAA3* genes. In a previous study, Xu et al. (2013) reported that the anti-sense suppression of the fiber-specific *GhPRP5* gene, which encodes a proline-rich protein, resulted in increased fiber length in cotton. Shan et al. (2014) demonstrated that *GhHOX3* encoded homeodomain-leucine zipper (HD-ZIP) transcription factor, which interacts with *GhHD1* in the gibberellic acid (GA) signaling pathway, contributing to the regulation of fiber elongation. Yang et al. (2014) identified the PAGODA1 (*PAG1*) gene, which plays a role in brassinosteroid (BR) catabolism and fiber elongation by modulating the levels of endogenous brassinosteroid. Additionally, Gong et al. (2014) found that the KNOX protein encoded by the KNOTTED-LIKE (*GhKNL1*) gene is specifically expressed in growing fiber cells during the secondary cell wall biosynthesis stage in cotton, thereby regulating fiber elongation. According to Abdurakhmonov et al. (2014), RNAi-mediated suppression of the *GhPHYA1* gene, which encodes the phytochrome red photoreceptor, led to a reduction in fiber length in cotton plants. Calcium signaling cascades involved in fiber elongation are regulated by *GhCaM7*, which encodes calmodulin, a calcium sensor, and influences calcium ion ( $\text{Ca}^{2+}$ ) influx (Cheng et al., 2016). Qin et al. (2017) reported that the *GhGalT1* gene in cotton, belonging to the CAZy glycosyltransferase 31 family, encodes B-1,3-GALACTOSYLTRANSFERASE. It negatively regulates fiber length and is involved in the biogenesis of the  $\beta$ -1,3-galactan backbone of type-II arabinogalactan glycans. Zhang J. et al. (2017) found that the *GhFSN1* gene, a member of the NAC transcription factors family in cotton, is specifically expressed during the secondary cell wall development stage. It acts as a positive regulator by activating secondary cell wall-associated genes, thereby modulating secondary cell wall thickening in cotton fibers. The *GhHUB2* gene encodes a ubiquitin ligase that regulates fiber length and secondary cell wall thickness in cotton and promotes the degradation of the *GhKNL1* protein (a repressor) through the ubiquitin-26 S proteasome mechanism (Feng et al., 2018). Gao J. et al. (2019) conducted an expression study on the *GhbHLH18* gene, revealing its specific expression during the fiber elongation stage. They found that *GhbHLH18* negatively modulates secondary cell wall development and fiber elongation by interacting with the lignin-specific *GhPER8* gene. In a study by Zhu L. et al. (2021), it was reported that the *GhAlaRP* (alanine-rich protein) gene plays a crucial role in regulating cotton fiber elongation. *GhAlaRP*

interacts with *GhAnnexin* and *GhEXPA* genes, and the RNAi-mediated suppression of *GhAlaRP* resulted in reduced fiber length through co-suppression of *GhAnnexin* and *GhEXPA* gene expression in cotton. Furthermore, Zhu S. et al. (2021) demonstrated that the *GhPIPLC2D* (PHOSPHOINOSITIDE-SPECIFIC PHOSPHOLIPASE C) gene is particularly expressed in fibers during the elongation stage. It acts as a positive regulator in fiber elongation by enhancing ethylene biogenesis. Silencing the *GhPIPLC2D* gene led to shorter lint fibers and reduced biosynthesis of Inositol-1, 4, 5-trisphosphate and ethylene.

Naturally colored cotton fibers, such as brown or green, acquire their non-white coloration through the accumulation of pigments during the fiber development process. This eliminates the need for dyeing processes in fabric manufacturing, resulting in reduced expenses and a more environmentally friendly approach by avoiding the disposal of harmful dye waste (Dutt et al., 2004). Plants produce flavonoids, which comprise the largest class of secondary metabolites, including colored compounds that are synthesized and accumulated in developing cotton fibers, giving rise to their natural pigmentation (Mikhailova et al., 2019). The transcriptional regulator *GhTT2-3A* (TRANSPARENT TESTA 2) is predominantly expressed in brown fibers and plays a crucial role in controlling brown pigmentation by regulating proanthocyanidins (PA) biosynthesis (Yan et al., 2018). Furthermore, the *Gh4CLA* gene, encoding 4-coumarate: Coenzyme A ligase, has been found to contribute to the biosynthesis of green pigments in cotton (Sun et al., 2019). Additionally, several genes involved in flavonoid biosynthesis, such as *GhCHS*, *GhANR*, *GhLAR*, *GhCHI*, *GhF3H*, *GhF3'5'H*, and *GhDFR*, have been identified and shown to be significantly expressed in developing fibers, thereby contributing to the pigmentation in naturally colored cotton (Gao Z. et al., 2019; Lv et al., 2023).

Hence, the cotton fiber-related traits are controlled by the genes involved in plant growth, development, and defensive hormones such as gibberellic acid, jasmonic acid, brassinosteroid, ethylene, abscisic acid, flavonoid, etc., pathways.

### 3.2 Seed-related genes and traits

Cottonseed is a by-product of cotton, characterized by its unique seed structure and high natural pigment content. It serves as a major source of unsaturated edible oil, free from cholesterol, with a composition consisting of 65%–70% polyunsaturated fatty acids, including 18%–24% oleic acid and 42%–52% linoleic acid. It also contains 26%–35% saturated fatty acids, such as palmitic and stearic acids, making it stand out among other seed oil-producing plants (Liu et al., 2002). The relatively high level of saturated fatty acids, particularly palmitic acid, contributes to the oxidative stability of cottonseed oil, making it a reliable choice for frying at high temperatures compared to oils rich in polyunsaturated fatty acids like linoleic acid and oleic acid. However, it is worth noting that saturated fatty acids are not considered reliable and healthy due to their association with an increase in low-density lipoprotein (LDL) cholesterol levels. Researchers have downregulated two essential fatty acid desaturase genes, namely *GhSAD-1* (STEAROYL-ACYL-CARRIER PROTEIN  $\Delta$ 9-DESATURASE) and *GhFAD2-1* (OLEOYL-PHOSPHATIDYLCHOLINE  $\omega$ 6-DESATURASE), through RNA-mediated gene silencing. This genetic modification

aims to enhance the oleic acid content in cottonseed oil (Liu et al., 2022). Additionally, Chen et al. (2020) utilized CRISPR/Cas9 mediated editing to knock out the *GhFAD2* genes, resulting in the production of non-transgenic cotton (*Gossypium hirsutum* L.) with a higher level of oleic acid content. Cottonseed consists of a hull and kernel. The hull is used for fiber or linters, while the kernel contains essential nutritional components such as oil, protein, carbohydrates, vitamins, minerals, lecithin, sterols, and more. Wang N. et al. (2017) conducted a genome-wide association study (GWAS) on the cotton LYSOPHOSPHATIDIC ACID ACYLTRANSFERASE (*LPAAT*) gene family, revealing their association with the biosynthesis of cotton seed oil components like palmitic acid, oleic acid, and triacylglycerol.

In a study conducted by Yao et al. (2018), bio-fortification of cottonseed with Pro-vitamin A was achieved through targeted upregulation of the PHYTOENE SYNTHASE (*GhPSY2D*) gene. This was accomplished by introducing a seed-specific promoter, which effectively enhanced the levels of  $\beta$ -carotene in transgenic cottonseeds. This innovative approach to bio-fortify cotton with pro-vitamin A offers a promising solution to address vitamin A deficiency on a global scale, benefiting populations worldwide.

Cottonseed contains certain anti-nutritional or toxic compounds such as gassypurpurin, gossyaerulin, gossyfulvin, gossyverdurin, and gossypol. Among these, gossypol is a yellow terpenoid pigment that plays a vital role in defending cotton plants against insect predation (Scheffler, 2016). Gossypol is present in higher quantities in raw cottonseed compared to cooked cottonseed, posing challenges in processing and consuming cottonseed as a by-product. It is toxic to non-ruminant animals since it exists in a free state within the seed. However, during cooking, gossypol binds to the free amino or free carboxyl groups of cottonseed protein, forming "bound gossypol." This reduces the nutritional value of the protein and limits the availability of essential amino acids like lysine (Rathore et al., 2020). The consumption of cottonseed containing gossypol can lead to various chronic effects such as liver damage, reproductive and immune toxicity, reduced iron bioavailability, disturbance in iron utilization, and lysine deficiency (Gadelha et al., 2014). To address this issue, researchers have employed seed tissue-specific RNAi to suppress the expression of the  $\delta$ -CADINENE SYNTHASE (*CAD1*) gene, which disrupts the gossypol biosynthesis pathway during seed development. This approach has significantly reduced gossypol levels in cottonseed without affecting its levels in other plant tissues where it serves as a defense against pests (Sunilkumar et al., 2006). Gao et al. (2020) reported the *CGPI* gene, which encodes a transcription factor preferentially expressed in the black glands of *Gossypium* spp. These glands play a crucial role in storing large amounts of gossypol and other secondary metabolites, providing defense against insects and pathogens (Gao et al., 2020). By using RNAi and CRISPR tools to suppress the expression of the *CGPI* gene, the researchers successfully created a *CGPI* mutant with reduced gossypol levels and a glandless-like phenotype (Gao et al., 2020). Additionally, they suggested that the *CGPI* transcription factor interacts with *GOPGF* in the nucleus, which is a key transcription factor for black gland formation and gossypol biosynthesis (Gao et al., 2020).

We could increase the nutritional content including polyunsaturated fatty acid, and tocopherol, and decrease or

eliminate the gossypol and cyclopropane-free fatty content in cottonseed oil through modern biotechnological tools to enhance or suppress the expression of genes that regulate their biosynthesis in cottonseed (Table 2).

### 3.3 Plant architecture-related genes and traits

Plant architecture plays a fundamental role in the productivity and management of cotton. The structure of a cotton plant is determined by meristems, which undergo differentiation to form a terminal structure during vegetative growth. The primary meristems give rise to branches (monopodial and sympodial), inflorescence, and flowers, while the secondary meristems contribute to woody growth, and organ-specific meristems shape the final structure of organs.

Roots serve as vital organs in crops, responsible for nutrient and water absorption from the soil, starch storage, physical support, and defense against biotic and abiotic stresses. Lateral root development is regulated by nitric oxide, a molecule synthesized from arginine in the presence of the nitric oxide synthase enzyme. However, increased arginase activity can inhibit the function of nitric oxide synthase due to competition for the same arginine substrate by both enzymes. To overcome this limitation, researchers utilized the CRISPR tool to edit the *GhARG* gene, which encodes ARGINASE in cotton plants. By knocking out the function of the *GhARG* gene, lateral root development was enhanced, resulting in increased crop yield (Wang N. et al., 2017).

Branching plays a crucial role in the regulation of above-ground organs in crops, directly influencing their development and yield. The intricate regulatory network responsible for branching involves phytohormones and transcription factors, serving as a fundamental mechanism for plant survival and occupation of space. Several genes involved in branch regulation have been identified and characterized in cotton, presenting potential targets for genetic manipulation of branch architecture (Table 3). McGarry et al. (2016) proposed that the expression levels of cotton SINGLE FLOWER TRUSS (*GhSFT*)-like and SELF-PRUNING (*GhSP*)-like genes dynamically modulate the development of monopodial and sympodial branches. Mutation of the SELF-PRUNING-like genes resulted in the termination of both branch systems in cotton. The cotton CUP-SHAPED COTYLEDON 2 (*GhCUC2*) genes (Zhan et al., 2021), members of the NAC family, participate in auxin and abscisic acid signaling pathways, while the BRANCHED 1 (*GhBRC1*) gene (Sun et al., 2022) from the TCP family acts as a negative regulator of branch development. In a previous study, suppression of the cotton DEHYDRATION-RESPONSIVE ELEMENT-BINDING (*GhDREB*) gene through virus-induced gene silencing (VIGS) significantly increased plant height, branch angle, and height (Ji et al., 2020). The *GhTIE1* gene, encoding a TCP interactor containing an EAR motif protein 1, also promotes shoot branch development in cotton by suppressing the expression of TCPs such as *GhBRC1*, *GhBRC2*, and *GhTCP13* genes (Diao et al., 2019). Furthermore, Zhao et al. (2017) identified and functionally characterized the *GhMAX2* gene in cotton, which encodes an F-box/LRR family protein and acts as a repressor in shoot lateral branch development.

TABLE 2 A list of genes and traits associated with seed.

Gene	Specific trait	References
Stearoyl-acyl-carrier protein $\delta 9$ -desaturase ( <i>GhSAD-1</i> )	Fatty acid desaturation	Liu et al. (2002)
Oleoyl-phosphatidylcholine $\omega 6$ -desaturase ( <i>GhFAD2-1</i> ) <sup>a</sup>	Fatty acid desaturation	Liu et al. (2002) Chen et al. (2020)
$\Delta$ -cadinene synthase ( <i>CAD1</i> )	Gossypol biosynthesis	Sunilkumar et al. (2006)
Lysophosphatidic acid acyltransferase ( <i>LPAAT</i> )	Oil biosynthesis	Wang et al. (2017b)
Phytoene synthase ( <i>GhPSY2D</i> )	Pro-vitamin A biosynthesis	Yao et al. (2018)
<i>CGPI</i>	Gossypol biosynthesis	Gao et al. (2020)

<sup>a</sup>CRISPR-edited gene.

TABLE 3 A list of genes and traits associated with plant architecture.

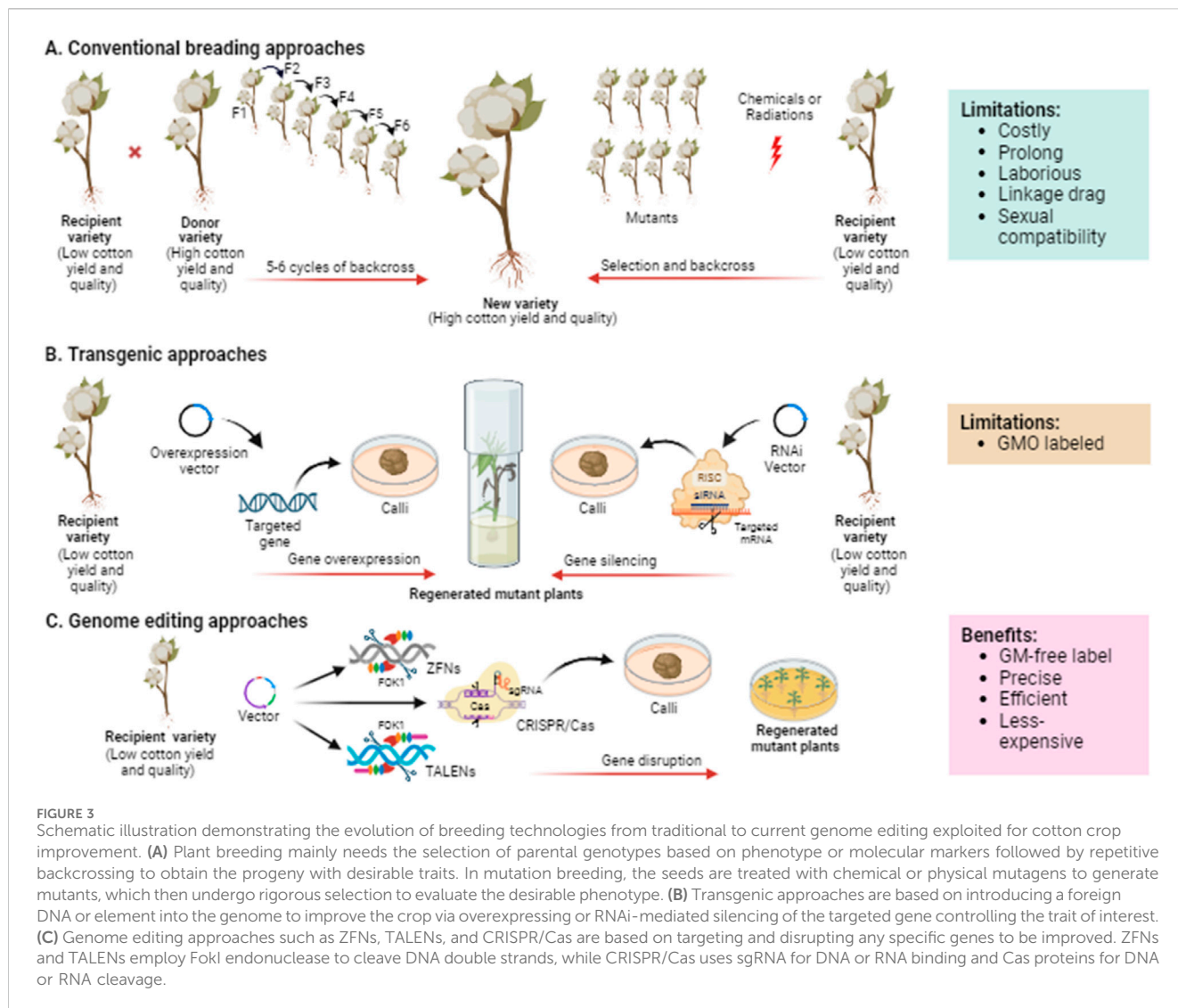
Gene	Specific trait	References
Flowering promoting factor 1 ( <i>GhFPF1</i> )	Flowering	Wang et al. (2014b)
Self-pruning ( <i>GhSP</i> )	Branching	McGarry et al. (2016)
Single flower truss ( <i>GhSFT</i> )	Branching	McGarry et al. (2016)
Phosphatidylethanolamine-binding protein ( <i>GhPEBP</i> ) <sup>a</sup>	Flowering	Zhang et al. (2016)
Late meristem identity 1 ( <i>GhLMI1</i> )	Leaf shape	Andres et al. (2016)
More axillary growth 2 ( <i>GhMAX2</i> )	Branching	Zhao et al. (2017)
Arginase ( <i>GhARG</i> ) <sup>a</sup>	Lateral rooting	Wang et al. (2017b)
Constans/flowering locus T ( <i>GhCOL</i> )	Flowering	Cai et al. (2017)
<i>GhHB12</i>	Flowering	He et al. (2018)
Tcp interactor containing ear motif protein 1 ( <i>GhTIE1</i> )	Branching	Diao et al. (2019)
$\alpha$ -amylase inhibitor 6 ( <i>GhAA16</i> )	Flowering	Qanmber et al. (2019)
Dehydration-responsive element-binding ( <i>GhDREB1B</i> )	Branch height, length, and angle	Ji et al. (2020)
<i>GhCAL</i> <sup>a</sup>	Flowering	Cheng et al. (2020)
Auxin response factor ( <i>GhARF16-1</i> )	Leaf shape	He et al. (2021)
Apetala1 ( <i>GhAP1</i> )	Flowering	Cheng et al. (2021)
Cup-shaped cotyledon 2 ( <i>GhCUC2</i> )	Branching	Zhan et al. (2021)
Branched 1 ( <i>GhBRC1</i> )	Branching	Sun et al. (2022)

<sup>a</sup>CRISPR-edited gene.

Leaves serve as the primary site for photosynthesis, making leaf shape a significant trait that influences crop productivity. In cotton, the leaf shape is regulated by various factors. One such factor is an HD-Zip transcription factor encoded by the LATE MERISTEM IDENTITY1 (*GhLMI1*)-like gene, as demonstrated by Andres et al. (2016). This transcription factor plays a role in determining the leaf shape in cotton. Additionally, He et al. (2021) proposed that the cotton AUXIN-RESPONSIVE FACTOR (*GhARF16-1*) specifically binds to the upstream region of *GhKNOX2-1*, another gene involved in leaf development. This binding interaction leads to the modulation of leaf shape by inducing the expression of *GhKNOX2-1*. Overexpression of *GhARF16-1* and *GhKNOX2-1* resulted in increased serrations in cotton leaves.

Crop plants, including cotton, undergo a critical stage transition from vegetative to reproductive development, which significantly impacts early maturity and flowering time (FTi). Early maturity is crucial for lint fiber quality, yield, and mechanical harvesting of cotton. In recent years, several transcriptomic studies have been conducted to identify and characterize genes associated with flowering that regulate flowering time and floral organs in cotton (Table 3). Wang X. et al. (2014) proposed that the FLOWERING PROMOTING FACTOR 1 (*GhFPF1*) gene plays a role in controlling flowering time in cotton. The PHOSPHATIDYLETHANOLAMINE-BINDING PROTEIN (*GhPEBP*)-like gene family also plays a critical role in modulating flowering time and photoperiod responses in cotton (Zhang et al., 2016). In contrast to *GhPEBP2*, *CONSTANS/*





FLOWERING LOCUS T (*GhFT*) (Cai et al., 2017) binds with FD-like bZIP (*GhFD*) to rescue the late-flowering phenotype in photoperiod-sensitive and wild cotton (Zhang et al., 2016). He et al. (2018) reported that the cotton *GHB12* gene, a member of the HD-ZIP I-class transcription factor family, is preferentially expressed in axillary buds. It negatively regulates flowering by binding to *GhSPL10/13*, suppressing the expression of *GHSO1*, *GhFT*, and *GhFUL* genes, resulting in delayed flowering under long-day conditions. The  $\alpha$ -AMYLASE INHIBITOR (*GhAAI66*) gene, specifically expressed in floral tissue, promotes early flowering by integrating various floral signaling mechanisms, as shown by its RNAi-mediated silencing in cotton (Qanmber et al., 2019). Two genes of the MADS-box family, *GhCAL* (Cheng et al., 2020) and *APETALA1* (*GhAPI*) (Cheng et al., 2021), play a key role in positively regulating flowering time (FTi) and the development of floral organs in cotton. *GhLFY* negatively regulates the expression of *GhAPI* in the regulatory mechanisms (Cheng et al., 2021), and *GhAPI* interacts with *GhSOC1* for the transcriptional activation of multiple flowering-associated genes (Wang C. et al., 2023).

Hence, these reported studies help provide us with knowledge about potential genes for improving cotton plant architecture as well as crop yield through advanced gene editing approaches.

## 4 Conventional breeding approaches

Plant breeding, a crucial science, involves intentionally introducing advantageous and inheritable changes in plants, making it one of the most pivotal approaches to improving crops. The significant impact of the Green Revolution, which has effectively reduced hunger and poverty for millions of people (Pingali, 2012), exemplifies the importance of this field. Selective breeding has played a prominent role in producing the majority of the currently cultivated cotton genotypes, leading to substantial improvements in cotton quality and yield. In this section, we provide a concise overview of various conventional breeding approaches employed for cotton crop improvement along with their limitations.

## 4.1 Selective breeding

Plant breeding frequently uses “selection” to carefully choose plants with superior traits from a diverse population and breed them multiple generations until these desirable traits become fixed in the selected population (Allard, 2002) (Figure 3A). The Key challenges associated with selective breeding include identifying desirable phenotypes through genetic variation, ensuring sexual compatibility between plants, and linkage drag, which make the approach costly, time-consuming, and labor-intensive.

Marker-assisted selection (MAS) involves the use of various markers to select a trait of interest. These markers can be categorized into three types: morphological, biochemical, and DNA markers. Morphological markers rely on the presence or absence of specific physical traits, while biochemical markers involve the analysis of distinct forms of an enzyme known as allozymes, which are encoded by different alleles at the same locus. DNA markers can be PCR-based, such as Random Amplified Polymorphic DNA (RAPD) (Lu and Myers, 2002), Amplified Fragment Length Polymorphism (AFLP) (Álvarez and Wendel, 2006), or non-PCR-based, such as Simple Sequence Repeats (SSRs) (Li et al., 2018) and Single Nucleotide Polymorphisms (SNPs) (Dilnur et al., 2019), among others. The challenges associated with the selection method can be overcome by utilizing marker-assisted selection (MAS), particularly the DNA-based approach. DNA-based MAS is a cost-effective and time-saving breeding method as it is not influenced by environmental factors and is not dependent on the growth stage of the plants. This makes it highly effective for implementation in early generations of breeding programs. It also enables the selection of precise parental lines for backcrossing, facilitating the introgression of multigenic traits (Frisch and Melching, 2005).

## 4.2 Mutation breeding

Mutation breeding involves inducing random mutations in crop plants using mutagens such as Ethyl Methane Sulphonate (EMS) or X-rays (Pathirana, 2011) (Figure 3A). Some of these random mutations may lead to useful traits. Subsequently, mutants with desired traits are identified, self-pollinated, or backcrossed to generate desired progenies. Cotton mutants with improved yield and fiber were developed using gamma rays, EMS, and salicylic acid (Muthusamy and Jayabalan, 2011). Advancements in mutation breeding and functional genomics have facilitated identify genes associated with traits such as plant architecture, leaf colour, and fiber development. In a recent study, Wang G. et al. (2024) developed an EMS mutant library of Asiatic cotton, identifying candidate genes in M1 mutants with different phenotypic variations in traits including, fiber, flower, leaf, and plant architecture through genome-wide mutation analysis. Further, they also confirmed the functions of *GaYUCA* and *GaPDX1* using virus-induced gene silencing, which are candidate genes for bar blade and yellow leaf vein.

For leaf colour-related genes, researchers observed a particular 0.34 Mb hypermutation interval on chromosome D10, containing 31 genes in the F2 population of the Sumian 22 mutant. Among these genes, only the ABCI1 gene showed particularly lower

expression in mutants, correlating with decreased levels of chlorophyll-associated compounds. A crucial A to T mutation at -317 bp from the ABCI1 start codon likely hinders its transcription, resulting in the green mutation through impeding chlorophyll synthesis (Gao et al., 2021). For plant architecture-related genes, scientists cloned the Asiatic cotton axillary flowering (*GbAF*) mutant gene as well as the upland cotton cluster branch (cl1) mutant gene, identifying a crucial mutation in which aspartic acid is substituted by asparagine at position 73 in *GbAF*, causing cotton bolls to develop directly on the main stem of the cotton plant. These results indicate that cotton SFT and SP gene orthologs can be used to improve cotton plant architecture (Si et al., 2018).

For fiber-related genes, a short fiber mutant was discovered in the Ghir\_A12G008870 gene encoded the tetrapeptide repeat-like superfamily protein in the Ethyl Methane Sulphonate mutant library. VIGS-based silencing of this gene decreased the fiber length in the wild-type cotton line named MD15 (Fang et al., 2020). In addition, a recessive tufted-fuzzless seed mutant was also identified on chromosome D04 having a genome interval of about 411 kb, with 7 genes showing significant differential expression between the tufted-fuzzless seed mutant and wild-type in that region. Researchers discovered a chemically-induced cotton mutant with short fiber, Ligon-lintless-y (*liy*), modulated through a single recessive locus, which affected many traits such as maturity, fiber length, and plant height (Naoumkina et al., 2021). Furthermore, F-actin polymerization was affected in the mutant *GhACT17D* from Li1 plants due to the substitution of Val for Gly65 on the nucleotide-binding domain of *GhACT17D*. Actin filaments in Li1 fibers indicated decreased filament skewness, greater filament density, faster growth and shrinkage rates, and parallel arrangements compared to the WT control (Cao et al., 2021).

Despite the relative simplicity of the physicochemical mutagenesis tool and the ease of creating mutant populations, the mutagenesis method is not well controlled, and frequently a mutant has more point mutations than the original, which could be the result of multiple point mutations acting in concert to appear the phenotype. In addition, post physicochemical mutagenesis, the plant genome could experience DNA fragment rearrangements or deletions, and urge the transposition of the reverseposon, making it more challenging to characterize functional genes.

## 5 Transgenic approaches

Transgenic approaches have the potential to overcome many of the limitations associated with conventional breeding methods. These approaches allow for the transfer of desired genes into crop plants, regardless of their origin. Transgenic approaches enable the stacking or pyramiding of multiple desired traits in crops. For instance, Bt cotton, a transgenic crop, has rapidly gained acceptance and has been commercially grown in recent years following its introduction (James, 2011). Advancements in functional genomics have facilitated the identification and characterization of a large number of potential genes encoding transcription factors (TFs) that are involved in important traits in cotton. Studies in this field have typically employed two major approaches for gene manipulation: overexpression of native genes and silencing of native genes (Figure 3B).

## 5.1 Gene overexpression

Gene overexpression involves the amplification of desired protein production in plants by utilizing expression vectors that enhance the transcription of the target gene (Chakravarthy et al., 2012). The process is relatively straightforward: genes of interest are cloned into plant expression constructs containing constitutive or tissue-specific promoters based on specific requirements. These cloned genes are then randomly integrated into the plant genome using various available plant transformation methods. Transgenic plants that overexpressed the desired gene are subsequently compared with wild-type plants to assess phenotypic traits associated with the target gene (Wilkins et al., 2000). In their study, Jiang et al. (2012) overexpressed the *GhSusA1* gene, which encodes SUCROSE SYNTHASE, a crucial enzyme for cellulose biosynthesis, playing a pivotal role in secondary cell wall synthesis and fiber cell elongation. Wang et al. (2010) conducted a study in which they overexpressed the gene encoding a fiber-preferential actin-binding protein, *GhPFN2*, resulting in the termination of fiber cell elongation with a short-fiber phenotype. Zhang J. et al. (2017) overexpressed the *GhFIM2* gene encoding an actin-bundling protein, which accelerated fiber growth through boosted actin bundle formation.

## 5.2 RNA interference

RNA Interference is an intrinsic mechanism in plants that regulates gene expression by utilizing small interfering RNAs (siRNAs) to induce post-transcriptional gene silencing (Hannon, 2002). The RNAi process involves two types of small RNAs: siRNAs and microRNAs (miRNAs), which are derived from exogenous or endogenous long double-stranded RNA molecules. These small RNAs guide RNA-induced silencing complexes (RISC) to bind to complementary mRNA sequences, leading to mRNA degradation or translation inhibition, thereby reducing the expression of target proteins. Researchers have developed a powerful RNAi tool called virus-induced gene silencing (VIGS), which allows for rapid and high-throughput functional validation of genes of interest and analysis of their phenotypic effects through transient post-transcriptional gene silencing. Li et al. (2005) used RNAi to repress the *GhACTIN1* gene expression in cotton, which disrupted the actin cytoskeleton network with a significantly decreased fiber elongation while fiber initiation was unaffected. A study reported by Walford et al. (2011), in which the *GhMYB25-like* gene silenced by RNAi, resulted in fiberless cotton seeds. In a study, the silencing of the *GhHOX3* gene expression, which encodes the homeodomain-leucine zipper transcription factor, was achieved by RNAi, resulting in reduced fiber length (Shan et al., 2014). Abdurakhmonov et al. (2016) silenced the phytochrome red/far-red photoreceptor gene (*GhPHYA1*) in cotton through RNAi, which enhanced fiber quality and agronomic traits. The expression of the pigment gland formation gene (GoPGF) encoding a bHLH transcription factor was repressed through VIGS, which results in the emergence of leaves that are either glandless or possess very few glands (Ma et al., 2016). Kumar et al. (2018) developed RNAi cotton lines with silencing of *GhHDA5* gene expression, which encodes HISTONE DEACETYLASE and is involved in fiber initiation.

Despite several studies of crop improvement achieved by transgenic approaches, only some of these transgenic crops have been successfully transferred into usable products for supply and cultivation. This is because most transgenic crops have antibiotic markers and reporter genes for their effective selection. Current biosafety rules and regulations do not permit field trials of transgenic crops having markers and reporters' genes, and making marker-free and reporter-free transgenic crops is more time-consuming, tedious, and expensive. Additionally, transgenic crops are not being socially accepted as food in some countries.

## 6 Genome editing approaches

Classical plant breeding methods and the adoption of transgenic crops have significantly improved agricultural yield and quality worldwide. However, classical breeding approaches are time-consuming, while the regulatory procedures surrounding transgenic crop availability for food consumption can be complex. Genome editing offers a precise means of manipulating the DNA of cells or organisms by targeting specific DNA sequences and creating double-strand break at desired sites using sequence-specific nucleases. Currently, three main approaches are employed for genome editing, utilizing different types of artificially engineered sequence-specific nucleases: zinc finger nucleases (ZFNs), transcription activator-like effector nucleases (TALENs), and clustered regularly interspaced short palindromic repeats/CRISPR-associated protein (CRISPR/Cas) systems (Carroll, 2014) (Figure 3C).

### 6.1 ZFNs-based approach

The zinc finger nucleases (ZFNs) based approach is an early and widely used site-specific genome editing method. ZFNs are composed of two domains: a zinc finger domain responsible for recognizing specific DNA sequences and a nuclease domain for cleaving DNA (Urnov et al., 2010). These synthetic proteins have been employed as genome-editing tools to introduce various types of mutations, deletions, and insertions in many plant species (Weinthal et al., 2010). Despite the successful studies of ZFNs in plant genome editing, several limitations have hindered their broader application in crop improvement. The main challenges include the complexity and high cost of constructing ZFNs for each specific genomic target, which require extensive knowledge of the target DNA sequence and custom zinc finger protein synthesis. In addition, ZFNs can be toxic and affect the overall health and viability of the plant. Their efficacy can vary, and there's a risk of off-target mutagenesis due to imprecise protein-DNA interactions (Puchta and Höhn, 2010).

### 6.2 TALENs-based approach

A TALENs-based approach was developed by fusing transcriptional activator-like effector repeats with the FokI endonuclease, aiming to enhance the safety, approachability, and efficiency of genome editing in plants. This system comprises a DNA binding domain derived from Transcription Activator-Like Effectors (TALEs) and a DNA cleavage domain from the FokI endonuclease (Christian et al., 2010). TALENs have been

successfully employed in various crops for genome editing and improvement purposes (Khan et al., 2016). Despite TALENs being a significantly improved genome editing tool for crop improvement, they do have certain limitations. These include the necessity of a T nucleotide before the 5' end of the target sequences for selecting TALEN sites, the complexity of designing TALE repeats, the high cost associated with protein engineering, difficulties in delivering these large proteins into plant cells, and their relatively low specificity (Luo et al., 2015).

### 6.3 CRISPR/Cas-based approach

The CRISPR/Cas-based approach has emerged as a high-throughput tool in cutting-edge genomics, with recent studies highlighting its widespread use in genome editing across various plant varieties. This approach offers simplicity, cost-effectiveness, and flexibility compared to ZFNs and TALENs (Zhang et al., 2019). CRISPR/Cas genome editing systems consist of a user-defined 20-nucleotide sequence of a single guide RNA (sgRNA) with a scaffold for attaching the Cas protein, derived from bacteria, which exhibits nuclease activity. To cleave the target sequence, the Cas protein requires a short conserved sequence known as the proto-spacer motif (PAM), which must be located downstream of the target site and cleaved at 3-4 nucleotides after the adjacent PAM sequence (Cong et al., 2013). CRISPR/Cas-based approach provides several advantages over ZFNs and TALENs including RNA-based recognition and cost-effective by eliminating the requirement for protein engineering. CRISPR/Cas allows multiplex editing with multiple gRNAs targeting multiple genes using a single Cas protein. The smaller CRISPR/Cas constructs are easier to deliver into cells, and CRISPR/Cas can recognize methylated DNA, which ZFNs and TALENs cannot (Gaba et al., 2021).

Genome-edited plants are more likely to be accepted by the public than genetically engineered ones. This is because the integrated transgene such as Cas can be eliminated in subsequent generations by segregation, making these plants non-transgenic, which may not face regulatory scrutiny (Podevin et al., 2012). Unlike mutation breeding, the specific genomic mutation is known in genome-edited plants, and there is no need for backcrossing to eliminate undesired mutations. Unlike transgenic approaches, genome editing through targeted transgene integration by homologous recombination reduces risks such as disrupting gene function or producing toxic proteins. Despite the high costs of ZFNs and TALENs, genome editing provides efficient, precise, and cost-effective crop improvements, enabling gene knockout, modulation of gene expression, and multiple transgenes stacking (Curtin et al., 2012). Hence, genome editing techniques hold great promise as powerful tools for precise genomic modification, enabling researchers to study gene function and develop crops with improved agricultural traits.

### 7 Single-cell RNA sequencing approach

Single-cell RNA sequencing has made it feasible to characterize each transcript present within a single cell. Functional genomics research on plants including cotton has significantly improved with

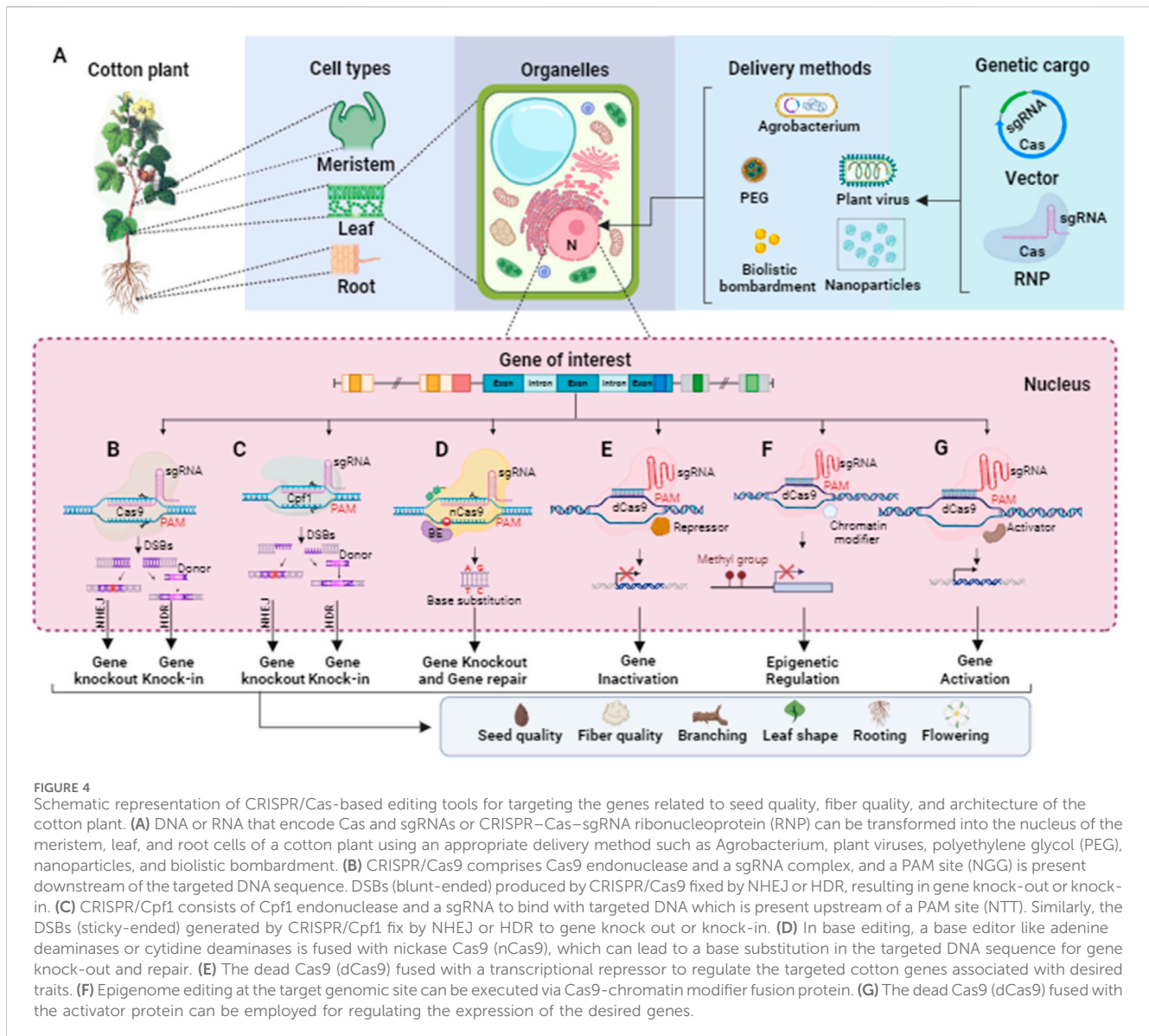
single-cell RNA sequencing, which enables an understanding of gene functions and regulatory networks behind the crucial traits (Peng et al., 2021). With the advent of single-cell RNA sequencing, significant development has been made in the cotton genome and the transcriptome landscape (Wen et al., 2023). Several studies on cotton gene expression have used single-cell RNA sequencing, which has provided important new understandings of various biological processes. Qin et al. (2022) identified core transcription factors such as *MYB25-like* and *HOX3* as crucial for cotton fiber differentiation and growth through single-cell RNA sequencing. By single-cell RNA sequencing, Wang D. et al. (2023) studied rhythmic fiber cell growth, finding that circadian processes, along with small peptide RALF1 and cis-regulatory elements such as CRE and TCPs, play important roles. In a study, Sun et al. (2023) explored gland morphogenesis in *Gossypium bickii*, finding that light and gibberellin promote pigment gland formation, affected by genes like *ERF114*, *ZAT11*, and *NTL9*. Long et al. (2023) also identified transcription factors like *PGF*, *ERF12*, and *MYB14*, as modulators of pigment gland morphogenesis. Lin et al. (2023) identified a hierarchical transcriptional network for terpenoid biosynthesis using single-cell RNA sequencing, with *HSF4a* and *NAC42* directly influencing gene expression in secretory glandular cells. In a recent study, Zhu et al. (2023) compared regenerable and recalcitrant cotton genotypes, identifying genes including *PLT3*, *LOX3*, and *LAX1/2* significant for cell fate reprogramming and plant regeneration. Guo et al. (2024) employed Single-cell RNA sequencing to map the transcriptome during somatic embryogenesis, revealing distinct cell clusters. Recently, Li P. et al. (2024) and Li Y. et al. (2024) studied transcriptome variations in cotton under environmental stress, showing gene expression deviation in particular cell types in anther and root under high temperature and salinity stress. Hence, single-cell RNA sequencing improves cotton functional genomics by offering comprehensive insights into gene functions at the single-cell level. It contributes to a deeper understanding of the genetic and epigenetic mechanisms underlying complex traits in cotton.

## 8 CRISPR/Cas-mediated genome editing in cotton

CRISPR/Cas as a cost-effective, robust, and dominant genome editing tool, has the potential to decrease the time required to develop new cotton varieties and increase sustainable cotton production by editing candidate cotton gene(s). In this section, we discussed the different types of CRISPR/Cas systems that can be used as genome editing tools for cotton improvement and overcoming the challenges in genome editing in cotton.

### 8.1 CRISPR/Cas systems as a genome-editing tool for cotton

The discovery of CRISPR in the microbial immune system has led biotechnologists to adopt CRISPR and its Cas proteins for RNA-guided genome editing tools in plants. To optimize their utility, several modifications have been introduced. One modification involves the design of a single chimeric guide RNA (sgRNA)



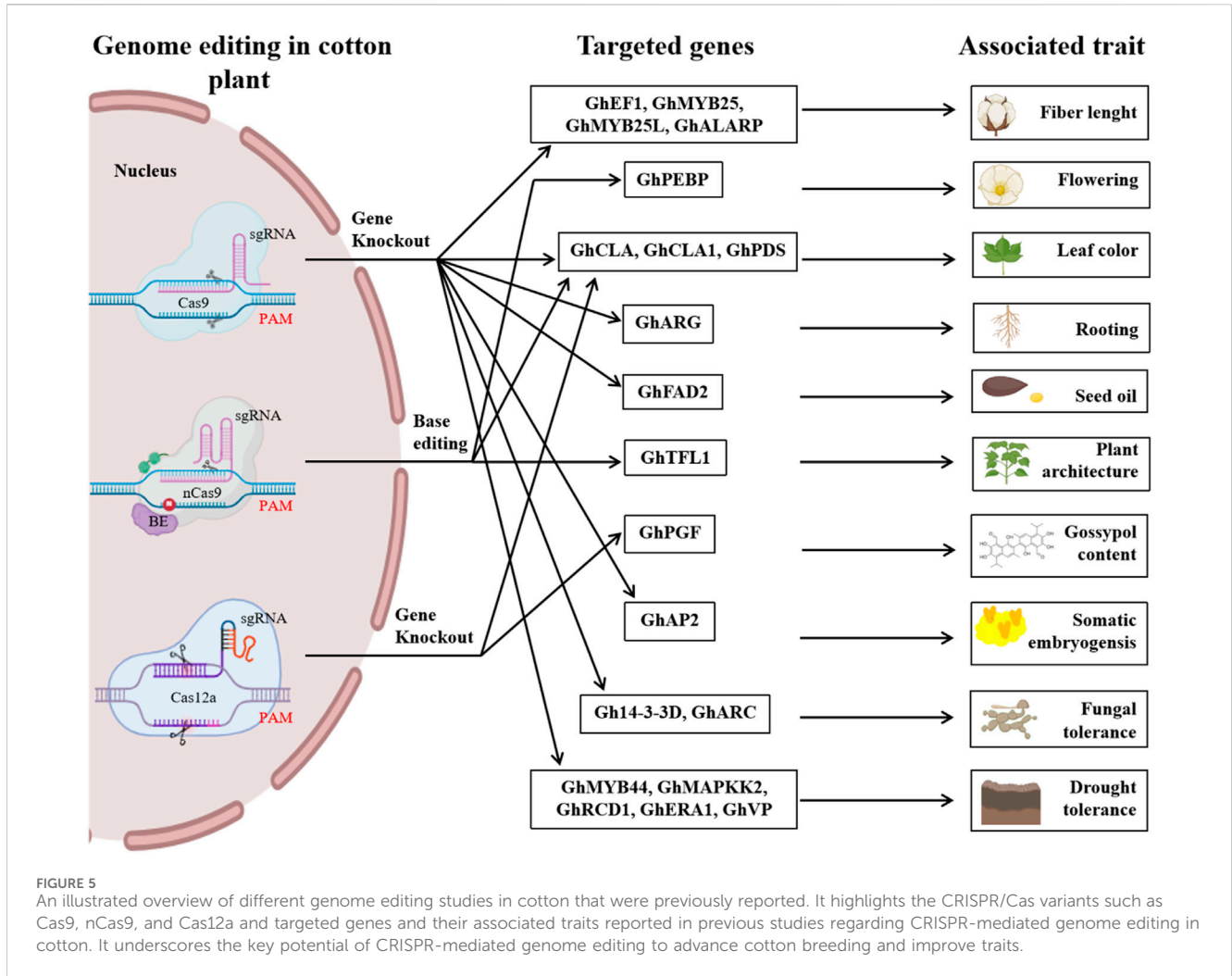
**FIGURE 4** Schematic representation of CRISPR/Cas-based editing tools for targeting the genes related to seed quality, fiber quality, and architecture of the cotton plant. **(A)** DNA or RNA that encode Cas and sgRNAs or CRISPR–Cas–sgRNA ribonucleoprotein (RNP) can be transformed into the nucleus of the meristem, leaf, and root cells of a cotton plant using an appropriate delivery method such as Agrobacterium, plant viruses, polyethylene glycol (PEG), nanoparticles, and biolistic bombardment. **(B)** CRISPR/Cas9 comprises Cas9 endonuclease and a sgRNA complex, and a PAM site (NGG) is present downstream of the targeted DNA sequence. DSBs (blunt-ended) produced by CRISPR/Cas9 fixed by NHEJ or HDR, resulting in gene knock-out or knock-in. **(C)** CRISPR/Cpf1 consists of Cpf1 endonuclease and a sgRNA to bind with targeted DNA which is present upstream of a PAM site (NTT). Similarly, the DSBs (sticky-ended) generated by CRISPR/Cpf1 fix by NHEJ or HDR to gene knock out or knock-in. **(D)** In base editing, a base editor like adenine deaminases or cytidine deaminases is fused with nickase Cas9 (nCas9), which can lead to a base substitution in the targeted DNA sequence for gene knock-out and repair. **(E)** The dead Cas9 (dCas9) fused with a transcriptional repressor to regulate the targeted cotton genes associated with desired traits. **(F)** Epigenome editing at the target genomic site can be executed via Cas9–chromatin modifier fusion protein. **(G)** The dead Cas9 (dCas9) fused with the activator protein can be employed for regulating the expression of the desired genes.

controlled by a promoter, replacing the need for expressing two non-coding RNAs (tracrRNA and pre-crRNA). Codon optimization is another modification performed on Cas proteins to ensure appropriate transcription in higher eukaryotic cells. Additionally, these proteins may be fused with nuclear-localized signals to facilitate their transport to the cell nucleus. In engineered CRISPR/Cas systems for gene editing, the gRNA typically consists of a short synthetic RNA that incorporates a trans-activating crRNA (tracrRNA) scaffold for binding with Cas proteins. The gRNA also contains a spacer region with a complementary sequence of twenty nucleotides, targeting the specific site of interest. Cas nucleases possess two nuclease domains, with one cleaving the sense strand and the other cleaving the antisense strand of a targeted gene. This flexibility allows the gRNA to target either strand, guiding the Cas enzyme to make precise cuts at specific sequences within the gene. Moreover, the presence of a protospacer-adjacent motif (PAM) sequence immediately adjacent to the target site is crucial for Cas nuclease

activity. The PAM sequence acts as a binding signal for the Cas enzyme, ensuring its specificity. The delivery of CRISPR/Cas reagents into plant cells is mediated by various methods like Agrobacterium, biolistic, polyethylene glycol (PEG), nanoparticle, and plant viruses, as shown in [Figure 4A](#).

### 8.1.1 Gene knockout

When using CRISPR/Cas9 for gene editing, the system typically cuts the double-stranded DNA at a specific site, resulting in a double-strand break (DSB). The repair of this DSB is primarily carried out through the non-homologous end joining (NHEJ) pathway (as shown in [Figure 4B](#)). During NHEJ repair, if one or more nucleotides are added or deleted at the site of the break, it can lead to a frameshift mutation. This frameshift mutation disrupts the reading frame of the targeted protein-coding gene, resulting in the loss of its normal function. This is the basis for quick and precise gene knockout without significant off-target effects. CRISPR/Cas9-based gene knockout is employed for the functional study of a



protein-coding gene and for eliminating some undesirable attributes controlled by particular genes in cotton. Gao et al. (2017) reported a rapid and effective methodology for validating the functionality of sgRNAs that target three distinct genes *GhPDS*, *GhCLA1*, and *GhEF1* in cotton, employing both transient and stable transformation. Particularly, they targeted *GhEF1* to check the efficacy of transient transformation-based sgRNA under the control of the AtU3b promoter. Moreover, they accomplished multiple gene targeting in cotton, inducing mutations in *GhPDS* and *GhEF1* at two respective target sites, in which the *GhPDS* and *GhEF1* sgRNAs were under the control of AtU6-29 and AtU3b promoters, respectively. The editing efficiency was high, with 80.6% of transgenic plants exhibiting mutations at the *GhCLA1* target site. Li et al. (2017) successfully knocked out a cotton *GhMYB-25* transcription factor gene on both A and D genomes using CRISPR/Cas9 with 98.8%–100% mutation frequency and no off-targeting. The knock-out line showed a fiber-less phenotype without changing other phenotypic characteristics. Chen et al. (2017) reported genome editing and targeted mutagenesis of the Chloroplastos alterados 1 (*GhCLA1*) and vacuolar H<sup>+</sup>-pyrophosphatase (*GhVP*) genes in upland cotton (*Gossypium hirsutum* L.) protoplast, employing the CRISPR/Cas9 system. The mutation efficiencies ranged from 47.6% to 81.8% in transgenic

cotton lines with no observed off-target mutations. Wang N. et al. (2017) knocked out the cotton *GhARG* gene on both A- and D-genome using the CRISPR/Cas9 system with the editing efficiency ranging from 10% to 98%. The knock-out line exhibited notable development of the lateral root system, enhanced lint yield, more nutrient absorption, and improved tolerance to drought. In a study by Wang Y. et al. (2017), the CRISPR/Cas9 system was utilized to conduct multiple-site genome editing of endogenous gene *GhCLA1*, with the resulting mutated phenotype and genotype transferred to their T1 progenies. The editing efficiency at each target site was 66.7%–100%, without any off-target editing at the potential off-target sites. In another reported study, a *GhALARP* gene, encoding an alanine-rich protein found in fiber cells, was edited via the CRISPR/Cas9 system, with mutation frequency ranging from 71.4% to 100%. Additionally, no off-target mutation event was observed in any predicted sites analyzed (Zhu et al., 2018). Zhang et al. (2018) knocked out the cotton *Gh14-3-3d* gene using the CRISPR/Cas9 system, and homozygous mutated plants showed resistance to *Verticillium dahlia* compared with wild-type. The edited phenotypes were stably passed on T1 generation as well and some homozygous mutants were also achieved. The first report on producing high-oleic acid, nontransgenic mutants in allotetraploid upland cotton through the CRISPR/Cas9 editing

system (76% editing efficiency) was recently published by [Chen et al. \(2020\)](#). The mutant cotton lines with a knockout of the microsomal x-6 fatty acid desaturase (*GhFAD2*) gene revealed significant increases in oleic acid along with a decrease in linoleic acid. [Lei et al. \(2022\)](#) developed an effective gene-editing system for rapidly producing cotton mutants via pollen as a transgenic receptor, with minimal off-target effects. They designed tissue-specific vectors to express Cas9 using *GhPLIMP2b* and *GhMYB24* promoters and sgRNAs targeting *GhCLA1*, *GhERA1*, and *GhGGB* genes, which mainly induced base substitutions, with editing efficiencies ranging from 3.29% to 6.45%.

### 8.1.2 Gene knock-in

In CRISPR/Cas9-based gene knock-in, the CRISPR/Cas9 system is used to target and cut the DNA at a specific position within the genome. Following the DNA cut, a foreign gene or DNA sequence is precisely incorporated into this targeted position through a process called homology-directed repair (HDR) (as shown in [Figure 4B](#)). This approach allows for the insertion of a desired gene without causing positional effects. However, CRISPR/Cas9-based gene knock-in is generally more challenging compared to gene knockout, as it often exhibits lower efficiency in genome editing.

### 8.1.3 Base editing

Genome-wide association studies have provided evidence that single nucleotide changes play a significant role in variations observed in desirable traits in crop plants ([Henikoff and Comai, 2003](#)). To address this, the CRISPR/Cas9-mediated base editing tool has emerged as a powerful approach for precisely modifying a single DNA base without the need for a DNA repair template ([Komor et al., 2016](#)). In the base editing approach, researchers have modified the Cas9 nuclease into a nickase form called nCas9, which introduces a single-stranded break in the DNA. The nCas9 is then fused with a base conversion enzyme such as cytidine deaminase or adenine deaminase (as shown in [Figure 4D](#)). CRISPR/Cas9 base editing has also been employed to disrupt gene function in plants by inducing nonsense mutations with high efficiency compared to the NHEJ process used in CRISPR/Cas9-induced knockouts ([Billon et al., 2017](#)). In a study reported by [Qin et al. \(2019\)](#), a novel *G. hirsutum*-Base Editor 3 (GhBE3) base-editing system has been developed to induce point mutations in the allotetraploid genome of cotton. This system involved fusing a cytidine deaminase with nCas9 and uracil glycosylase inhibitor into a CRISPR/Cas9 vector to target genes *GhCLA* and *GhPEBP*. The editing efficiency of GhBE3 ranged from 26.67% to 57.78% at the target sites. In addition, only <0.1% C>T substitutions were observed in the editing windows of predicted off-target sites, and the edited bases were inherited by T1 progeny. [Wang W. et al. \(2024\)](#) developed highly efficient base editors (GhABE8e) in cotton, which showed 99.9% editing efficiency with no detected off-target mutations. Further, these base editors were utilized to edit both non-coding and coding regions of cotton TERMINAL FLOWER 1 (*GhTFL1*), which are involved in cotton plant architecture. Utilizing 26 targets, they produced a comprehensive allelic population to explore the functional divergence of *GhTFL1*.

### 8.1.4 Gene regulation

CRISPR/Cas9 system has been harnessed for gene regulation purposes by targeting gene promoters, transcription factors, and enhancers, utilizing either the non-homologous end joining (NHEJ) or homology-directed repair (HDR) processes ([Gilbert et al., 2013](#)). The Cas9 endonuclease is commonly employed in CRISPR/Cas9-based gene editing and possesses multiple functional domains, including a PAM-binding domain, as well as RuvC and HNH domains. The RuvC and HNH domains are responsible for cutting the double-stranded DNA, leading to the generation of double-strand breaks (DSBs), which can subsequently be repaired through NHEJ or HDR mechanisms. By deactivating the nuclease activity of the HNH and RuvC domains in Cas9 (resulting in a deactivated Cas9 or dCas9), the binding capability of dCas9 to target sequences remains intact. This property has been exploited in the development of CRISPR/dCas9 interference (CRISPRi) and CRISPR/dCas9 activation (CRISPRa) systems for gene knockdown and gene activation, respectively ([Qi et al., 2013](#)). In these systems, dCas9 is fused with either a repressor or an activator, acting as a “gene switch” at the transcriptional level. The binding of dCas9 inhibits the binding of other proteins to the DNA sequence on which dCas9 is already bound, allowing for precise control of gene expression ([Gilbert et al., 2013](#)) (as shown in [Figures 4E, G](#)).

### 8.1.5 Epigenetic regulation

Gene expression can be epigenetically regulated through DNA methylation or demethylation at specific nucleotide sites, such as CpG, CHH, and CHG, in plants ([Adli, 2018](#)). To target and modulate DNA methylation, the CRISPR/dCas9 system has been modified by fusing the deactivated Cas9 (dCas9) protein with either a methyltransferase or a demethylase, allowing for efficient editing of DNA methylation patterns (as shown in [Figure 4F](#)). [Papikian et al. \(2019\)](#) developed a robust and efficient CRISPR/dCas9-SunTag system for plants, which incorporates the catalytic domain of the *Nicotiana tabacum* DRM methyltransferase. This system enables precise editing of DNA methylation at specific sites.

### 8.1.6 CRISPR/Cas12a

CRISPR/Cas9 system has a limitation in target site selection due to its recognition of G-rich protospacer adjacent motifs (PAMs). To overcome this limitation, the CRISPR/Cpf1 system, also known as Cas12a, was developed as an alternative to Cas9. Cpf1 recognizes T-rich PAM sequences (5'-TTN-3') and generates cohesive-end breaks instead of blunt-end breaks produced by Cas9 ([Zetsche et al., 2015](#)) (as shown in [Figure 4C](#)). In the CRISPR/Cpf1 system, a single CRISPR RNA (crRNA) of 42 nucleotides is sufficient for target site recognition and cleavage, eliminating the need for tracrRNA used in the Cas9 system. Cpf1 possesses a single nuclease domain that generates a staggered double-strand break (DSB) with a 5'overhang of 4 to 5 nucleotides. This feature makes Cpf1 particularly suitable for homology-directed repair (HDR)-based gene editing, as the ends of the DSBs become predictable. Unlike most Cas nucleases that require PAM sites at the 3'end of the target DNA sequence, Cpf1 requires the PAM site to be positioned at the 5'end. As a result, Cpf1 cleaves the target DNA sequence at the distal end from the PAM, creating the potential for subsequent rounds of cleavage ([Zetsche et al., 2015](#)). A LbCpf1 (LbCas12a) plant expression vector with a 23-nucleotide crRNA was used to target the

*GhCLA* gene in cotton (*G. hirsutum*). The findings demonstrated an editing efficiency exceeding 80% with no detected off-target effects. The edited traits were stably passed on to the T1 generation with certain homozygous mutants (Li J. et al., 2019). Li et al. (2020) knocked out the *GhPGF* gene in cotton responsible for pigment gland formation using the CRISPR/LbCpF1 (LbCas12a) system for editing under various temperatures. The results showed that the maximum temperature for CRISPR/Cas12a activity was 34°C with editing efficiencies ranging from 67.6% to 91.5%, and a homozygous gossypol-free non-transgenic line was obtained. A multiplexed crRNA-based Cas12a system has been developed to target the various ORFs of the Cotton leaf curl Multan virus (CLCuMuV) genome at multiple sites simultaneously, which showed an editing efficiency of 21.7%–55.6%. Transgenic *Nicotiana benthamiana* plants harboring multiplex LbCas12a-based construct exhibited a significant decrease in virus accumulation compared to the control plant. This multiplex LbCas12a system has the potential to develop virus resistance in cotton plants against begomoviruses (Ashraf et al., 2023).

### 8.1.7 Gene stacking

Recombinase technology is an earlier approach, in which recombinases have been used for genetic manipulation including site-specific insertion, deletion, or replacement of a target gene. In a reported study by D'Halluin et al. (2013), a re-engineered meganuclease was developed for the pyramiding of genes through the combined targeted genomic DNA cleavage and homologous recombination-based repair. By integrating up to three genes at a time and a specific genomic site, the gene stacking approach has evolved to be an adaptive strategy for improving traits and preventing varying expression of genes at various regions. Aslam et al. (2022) inserted the Cre and PhiC31-based recombination sites in the cotton genome to develop lines for recombinase-mediated gene stacking. Hence, CRISPR/Cas-based genome editing is more advantageous for gene stacking than recombinase technology due to the preciseness and targeted approach even at the first event of gene insertion in the host genome.

### 8.1.8 Mutant libraries construction

It is a significant task to critically evaluate the functions of all the sequenced genes in a cotton plant genome. By constructing a genetic library full of mutations, this issue can be resolved. The sgRNA's 20-bp target-binding site was modified to improve the capability of the CRISPR/Cas9 system to target particular genes. CRISPR/Cas9 is a viable and cost-effective tool for conducting forward genetic analysis and genome-wide mutations. This finding opened the door for the developing CRISPR/Cas9-mediated screening of plant mutant libraries in human cultured cells. Numerous mutant strains were produced when pooled sgRNA libraries were inserted into tomatoes (Jacobs et al., 2017). CRISPR/Cas9-mediated mutant libraries in rice have been created by a research group, which has produced a significant number of gene-disrupted mutations by transforming sgRNA libraries (Meng et al., 2017). In cotton, Ramadan et al. (2021) established a CRISPR/Cas9 system-mediated pooled-sgRNAs method to target multiple genes associated with male sterility. A population of mutants was produced, which would be useful to find the key genes that may improve fertility. These findings collectively showed that constructing mutant libraries with CRISPR/Cas systems

ensures speedy functional characterization of identified genes, paving the way for future improvement in the cotton genome.

## 8.2 Challenges associated with genome editing in cotton

CRISPR/Cas-mediated genome editing has increased biotic and abiotic stress tolerance, yield, and end-use quality of cotton crops, as depicted in Figure 5. However, some challenges remain for genome editing in tetraploid cotton, which we discussed in this section.

### 8.2.1 Requirement of explants with regenerative potential

The main challenge in generating CRISPR-edited cotton lines is the requirement for explants with regenerative potential. The recalcitrant to *Agrobacterium* infection and regeneration via somatic embryogenesis are the key bottlenecks in genome editing of elite cotton cultivars. Only these cotton genotypes Coker312, ZM24, JIN668, and YZ-1 have been transformed via *Agrobacterium*-mediated transformation at very low transformation efficiency. In a study reported by Chen et al. (2017), they delivered the CRISPR/Cas9 system into the tetraploid cotton genome by the vacuum infiltration-assisted *Agrobacterium*-mediated transformation using the shoot apexes as ex-plant, resulting in obtained mutation frequencies ranging from 47.6% to 81.8%. In another study, Ge et al. (2023) established an efficient and genotype-independent shoot apical meristem cell-mediated transformation (SAMT) system for different recalcitrant cotton genotypes that bypasses somatic embryogenesis. Using this SAMT, they produced edited elite cotton genotypes with inheritable targeted editing in genes (*GhRCD1* and *GhPGF*) through the CRISPR/cas9 system. This study paves a new way for producing gene-edited elite cotton varieties, which are recalcitrant to genetic transformation and regeneration.

Furthermore, several plant viruses like tobacco mosaic virus (TMV), tobacco rattle virus (TRV), and potato virus X have been genetically modified to be used as vectors in VIGS for functional genomics (Lange et al., 2013) and in delivering the ZFNs and TALENs for generating gene knockouts in both monocots and dicots (Cody and Scholthof, 2019). Now, the heritable, efficient, and transgene-free CRISPR-mediated gene editing is also achieved in germline cells of plants using the viral vectors, which are delivered into plants through *Agrobacterium*. The replicating viral vectors employed for CRISPR/Cas9 delivery can't carry the Cas9 gene due to its considerably larger size. Therefore, the Pea early-browning virus and Tobacco rattle virus have been reported to deliver small gRNAs into genetically modified plants expressing the Cas9 gene for CRISPR/Cas9-mediated gene editing (Hu et al., 2019). Sonchus Yellow Net Rhabdovirus, an RNA virus with higher cargo capacity, was engineered to deliver both the Cas9 gene and small guided RNA into the plant cell for transgene-free *In planta* CRISPR editing (Ma et al., 2020). Most of these viral vectors can't target the meristematic cells to induce mutations and give the heritable edits from seeds. To overcome this issue, the gRNA has been fused with the endogenous mobile RNA encoded by Flowering Locus T to increase mobility and ease systemic spread within plants to reach the



meristematic cells (Ellison et al., 2020). Lei et al. (2022) developed a cotton leaf crumple virus (CLCrV)-mediated VIGE system in cotton using Cas9-overexpressing cotton lines as the receptor. With the use of this VIGE system, it's feasible to simultaneously double-mutate the *GhCLA1* and *GhPDS* genes in addition to knocking out the *GhMAPKKK2*, *GhCLA1*, and *GhPDS* genes on both A and D genomes. It showed high gene editing specificity with mutation efficiencies ranging from 8.02%–52.68%. Additionally, they fused the gRNA with FT mRNA for heritable targeted editing of the *GhCLA1* and *GhPDS* genes in tetraploid cotton, which exhibited editing efficiency of 23.98%–55.43% but heritable mutant progeny was not detected. In another study reported by Gao et al. (2021), a convenient and robust genome editing strategy was developed in cotton through engineered CLCrV-mediated sgRNA delivery, in which the *GhMYB25L* gene (involved in regulating cotton fiber initiation) targeted using this strategy for mutation in stably transformed Cas9-expressing cotton lines. Further, gRNA was expressed and delivered in non-regenerative cotton variety via the grafting method, resulting in each Cas9-expressing cotton plant producing many Cas9-sgRNA cotton plants. The Hi-tom sequencing results of scions from 72 grafted plants showed that every single plant has an editing efficiency of up to 62.15% after 3 weeks of grafting. Hence, CLCrV-mediated genome editing with grafting precludes the requirement for lengthy tissue culture and laborious transformation procedures in cotton.

### 8.2.2 Requirement of specific protospacer adjacent motifs (PAMs)

The requirement of specific protospacer adjacent motifs (PAMs) has limited the range of DNA sequences that can be targeted for gene editing using the CRISPR/Cas system in cotton. However, Walton et al. (2020) have recently developed modified versions of the SpCas9 nuclease, namely SpG and SpRY, through structure-guided engineering. These modifications relax the PAM requirements, allowing for the targeting of a broader range of DNA sequences for CRISPR/Cas-mediated gene and base editing. Importantly, these modifications do not compromise the Cas nuclease activities. In a study reported by Wang G. et al. (2024), the GhABE7.10-dCpf1 and GhABE8e-dCpf1 vectors were devised and evaluated to expand the PAM sites in the cotton genome, which recognized TTTV PAM sequences. In the GhABE7.10-dCpf1 vector, dCpf1 (deactivated Cpf1) fused with adenine deaminase TadA7.10, which showed editing efficiency ranging from 0.2% to 0.5%, while in the GhABE8e-dCpf1 vector, dCpf1 fused with TadA8e, which showed editing efficiency of 1.5%. The lower editing efficiency than the GhABE8e-Cas9n system may be due to the less compatibility of TadA7.10 and TadA8e to Cpf1.

### 8.2.3 Low multiallelic editing efficiency

The low editing efficacy is a significant challenge in tetraploid cotton that sometimes needs simultaneous editing of several homoeologous gene copies to obtain the desired phenotype. The gRNA activity and efficiency could be assessed *in vivo* using protoplasts or transient co-expression assays before stable transformation as the creation of transgenic cotton lines usually takes several months. Using the transient expression approach in cotton cotyledons for high-throughput validation of gRNA, Gao et al. (2017) obtained a mutation frequency of above 80% following stable transformation. Although a single gRNA could generate

frameshift mutations, using multiple gRNAs to target a single gene could result in significant deletions across target sites and enhance the chance of producing multiallelic loss-of-function mutations. Additionally, high Cas and gRNA transcript levels could boost the effectiveness of gene editing. Thus, the promoters regulating the expression of gene editing reagents have a major effect on obtaining high mutation efficiencies. Long et al. (2023) demonstrated that the editing efficiency of the cotton CRISPR/Cas9 system was improved 4–6 times when gRNA is driven by endogenous GhU6 promoter instead of Arabidopsis AtU6-29 promoter. To enable highly efficient genome editing in tetraploid cotton, Zhang et al. (2018) employed a modified CRISPR vector that comprised the native promoter GhU6.9. Sequencing showed mutagenesis efficiencies ranging from 66.7% to 100% at four target regions. Chen et al. (2020) used the same vector system in a follow-up study and found that 76% of the recovered transgenic plants had mutations in cotton *GhFAD2* homologs. This vector was also employed by Ramadan et al. (2021) to improve their potent pooled-gRNA assembly approach. Li et al. (2022) developed the two novel genome editing vectors named pBeYDV-Cas9-KO (engineered bean yellow dwarf virus) and pRGE32-35S (Cas9 driven by CAMV35S promoter) and tested them by targeting the *GhCLA1* gene in cotton. Further, they also compared them with the ordinary CRISPR/Cas9 method, which showed that both new genome editing vectors had great efficiency (73.3% and 51.2% by pBeYDV-Cas9-KO and pRGE32-35S respectively) with no off-target effects. In cotton, Cas expression is mostly driven under the Cauliflower mosaic virus 35S promoter (CaMV 35S), But rice Ubiquitin (OsUbi) promoters are also reported (Table 4). In addition to constitutive expression, Cas can also be expressed in particular cell types. In cotton, the GhPLIMP2b and GhMYB24 promoters were used to induce Cas9 expression in pollen (Lei et al., 2022).

### 8.2.4 High off-target effects

The length of gRNA is vital to determine the off-target site within the host genome. The gRNA sequence and the occurrence of a PAM downstream to the gRNA sequence in the genome firmly control the targeting specificity of Cas9, off-target cleavage activity may still take place on DNA sequences with even 3–5 bp mismatches in the gRNA sequence (Kang et al., 2022). In the case of CRISPR/Cas9 systems, the common cause of off-target cleavage activity is the gRNA's recognition of completely or partially complementary genomic sites (Li et al., 2018). Different CRISPR/Cas variants such as HF-Cas9, eCas9, and HypaCas9 are also present with varying PAM requirements, which can be used to increase on-targeting and reduce off-target. Several approaches have been suggested to reduce off-target effects, including altering the half-time of Cas9 or reducing the amount and duration of functioning Cas9 protein in cells through selective delivery (Hajiahmadi et al., 2019). Furthermore, genome-wide off-targeting analysis can also be performed before the stable transformation of the CRISPR/Cas tool in cotton. Li et al. (2018) demonstrated a whole genome sequencing approach to identify the on and off-target mutation in edited cotton plants. Moreover, the inherent genetic mutation of wild-type can also produce new off-target sites and demolish PAMS, which signified that gRNA is designed very carefully to reduce off-target effect (Li et al., 2018).

TABLE 4 A list of reported studies of the CRISPR/Cas-mediated genome editing in cotton.

Targeted gene	Cas	Cas promoter	gRNA promoter	CRISPR reagents delivered by	Study objective	References
<i>GhEF1, GhCLA1, GhPDS</i>	Cas9	2x35S	<i>AtU6, AtU3b</i>	Agrobacterium <sup>a,b</sup>	KO; Loss of function mutation	Gao et al. (2017)
<i>GhMYB25-A, GhMYB25-D</i>	Cas9	2x35S	<i>AtU6</i>	Agrobacterium <sup>b</sup>	KO; Targeted editing of fiber quality associated gene	Li et al. (2017)
<i>GhCLA1, GhVP</i>	Cas9	CAMV35S	<i>AtU6</i>	Agrobacterium <sup>c</sup>	KO; Targeted gene editing in cotton protoplast and shoot apices	Chen et al. (2017)
<i>GhARG</i>	Cas9	CAMV35S	<i>NtU6</i>	Agrobacterium <sup>b</sup>	KO; Improve lateral root formation	Wang et al. (2017b)
<i>GhCLA1</i>	Cas9	<i>OsUbi</i>	<i>GhU6</i>	Agrobacterium <sup>b</sup>	KO; Loss of function mutation	Wang et al. (2017b)
<i>GhALARP-A, GhALARP-D</i>	Cas9	CAMV35S	<i>AtU6</i>	Agrobacterium <sup>b</sup>	KO; Editing of gene expressed in fiber	Zhu et al. (2018)
<i>Gh14-3-3D</i>	Cas9	CAMV35S	<i>AtU3b</i>	Agrobacterium <sup>b</sup>	KO; Tolerance against <i>V. dahliae</i>	Zhang et al. (2018)
<i>GhPDS</i>	Cas9	2x35S	<i>AtU6, GhU6</i>	Agrobacterium <sup>a</sup>	KO; Check the efficiency of gRNA promoters in cotton	Long et al. (2023)
<i>GhAP2, GhMYB44, GhARC</i>	Cas9	<i>OsUbi</i>	<i>GhU6</i>	Agrobacterium <sup>b</sup>	KO; Analysis of off-target activity of CRISPR/Cas9 in cotton	Li et al. (2019a)
<i>GhCLA1</i>	Cas12a (cpf1)	<i>OsUbi</i>	<i>GhU6</i>	Agrobacterium <sup>b</sup>	KO; Targeted gene mutation	Li et al. (2019b)
<i>GhPGF</i>	Cas12a (cpf1)	<i>OsUbi</i>	<i>GhU6</i>	Agrobacterium <sup>b</sup>	KO; Gossypol free cotton	Li et al. (2020)
<i>GhCLA, GhPEBP</i>	nCas9	<i>OsUbi</i>	<i>GhU6</i>	Agrobacterium <sup>b</sup>	BE; Check the efficiency of base editing in cotton	Qin et al. (2019)
<i>GhFAD2-A, GhFAD2-D</i>	Cas9	<i>OsUbi</i>	<i>GhU6</i>	Agrobacterium <sup>b</sup>	KO; Increase oleic acid content	Chen et al. (2020)
<i>GhCLA1, GhERA1, GhGGB</i>	Cas9	<i>GhPLIMP2b, GhMYB24</i>	<i>GbU6</i>	Agrobacterium <sup>c</sup>	KO; Tissue-specific editing in pollen	Lei et al. (2020)
<i>GhCLA1</i>	Cas9	CAMV35S	<i>GhU6</i>	Agrobacterium <sup>b</sup>	KO; Check the efficiency of the geminivirus-mediated gene editing vector in cotton	Li et al. (2022)
<i>GhMAPKKK2, GhCLA1, GhPDS</i>	Cas9	CAMV35S	<i>AtU6</i>	Cotton leaf crumple virus <sup>a</sup>	KO; Establish a VIGE system based on CLCrV in cotton	Lei et al. (2022)
<i>GhRCD1, GhPGF</i>	Cas9	CAMV35S	<i>AtU6</i>	Agrobacterium <sup>d</sup>	KO; Study the feasibility of SAMT for genome editing in cotton	Ge et al. (2023)
<i>GhMYB25L</i>	Cas9	CAMV35S	<i>GhU6</i>	Cotton leaf crumple virus <sup>a</sup>	KO; Establish a CLCrV-mediated CRISPR/Cas9 and grafting system in cotton	Guo et al. (2023)
<i>GhTFL1</i>	nCas9	<i>OsUbi</i>	<i>GhU6</i>	Agrobacterium <sup>b</sup>	BE; Base editing of plant architecture-related gene	Wang et al. (2024b)

<sup>a</sup>Agroinfiltration method, Ex-plants; cotyledonary leaves.

<sup>b</sup>Dip method, Ex-plants; hypocotyls.

<sup>c</sup>Vacuum infiltration method, Ex-plants; shoot apical meristem and pollen.

<sup>d</sup>Sonication method, Ex-plants; shoot apical meristem.

KO: Gene knockout.

BE: Base editing.

## 8.2.5 Assessment of multiallelic editing

Targeted editing in tetraploid cotton can result in different genotypes such as knockouts in one or more copies or all alleles. Due to this, the molecular characterization of mutations in cotton genomes can also be onerous. The medium-throughput techniques include polymerase chain reaction (PCR) to amplify the genomic

target sequence covering gRNA and gel electrophoresis for visualization of longer or shorter indels in the targeted genomic site, accordingly. These techniques reduced the expenses and labor related to genotyping tetraploid cotton plants for screening potential mutants generated by CRISPR/Cas-mediated genome editing. Cleaved amplified polymorphic sequences (CAPS/PCR-RE)

assays, T7 endonuclease 1 assay (TE71), single-strand conformation polymorphism (SSCP) method, and high-resolution melting (HRM) curve analysis have been used for the identification of mutations. However, the number and nucleotide sequence of gRNAs delivered determines which screening method should be employed (May et al., 2023). The most often employed technologies have been Illumina sequencing with high-throughput and Sanger sequencing for assessing the type and frequency of mutations in tetraploid cotton. Sanger sequencing offers a longer read length, that evaluates editing at multiple gRNA sites within a single allele and differentiates allelic variations more precisely through identifying allele-specific SNPs. In several reported studies, Sanger sequencing of multiple clones has been also used to quantify the number of co-edited alleles in a cotton line. This approach is useful for assessing either simple or chimeric mutations but it is costly and time-consuming. To avoid multicloning, various computational tools including Tracking of Indels by Decomposition (TIDE), HI-TOM, Inference of CRISPR Edits (ICE), Cas-analyzer, CRISP-ID, and Deconvolution of Complex DNA Repair (DECODR) greatly facilitates the quantitative evaluation of the extent and type of targeted edits utilizing Sanger sequencing data (May et al., 2023). In addition, third-generation long-read sequencing approaches such as Single Molecule Real-Time (SMRT) and Nanopore can be also used to quantify the number of mutated alleles in tetraploid cotton (May et al., 2023).

Furthermore, commercializing genome-edited cotton is also challenged by stringent and variable regulatory frameworks, technical issues such as off-target effects and trait complexity, and complex intellectual property landscapes. Mixed consumer perspectives and slow industry adoption restrict market acceptance. High development costs, uncertain returns on investment, limited funding, and the difficulty of conducting extensive field trials further complicate progress. Despite these obstacles, potential advantages including improved fiber and seed quality or yield make genome-edited cotton a promising study field, with future feasibility likely as regulations evolve and technical advances are made.

## 9 Key target genes for genome editing in cotton

Whole genome sequencing and multi-omic analysis have been conducted to identify key genes associated with agriculturally important traits, such as fiber and seed quality, as well as plant architecture in cotton. These genes represent potential targets for advanced CRISPR/Cas-mediated genome editing tools to enhance these specific traits (Figure 6). The identified genes can be categorized into two types: structural genes, which primarily consist of protein-coding genes directly controlling the distinctive traits of the crop. Therefore, CRISPR/Cas-mediated genome editing is an excellent choice for improving traits governed by structural genes. However, it's important to note that traits are not solely regulated by these structural genes in many plants. Numerous regulatory genes, including transcription factors and non-coding RNAs, are also involved in the underlying mechanisms. Therefore, targeting transcription factor genes using CRISPR/Cas tools holds

promise for enhancing agriculturally important traits in cotton, as these transcription factors may regulate multiple structural genes associated with these traits.

Several transcription factor genes in cotton have been discussed in the previous section as potential targets for CRISPR/Cas genome editing to improve various traits associated with cotton fiber, including yield, length, and strength. These genes include *GhFLA1*, *GhACT1*, *GhPFN1*, *GhMYB25*, *GhMADS11*, *GhSusA1*, *GhPIP2*, *GhTCP14*, *GhHOX3*, *PAG1*, *GhKNL1*, *GhCaM7*, *GhFNS1*, *GhHUB2*, *GaHD1*, *GhAlaRP*, *GhPIPLC2D*, *GhPER8*, *GhAnnexin*, *GhEXPA*, *GhAUX*, *GhIAA3*, *GbPDF1*, *GhVIN1*, *GhHDA5*, and *C2H2*. These genes serve as positive regulators in the fiber initiation and elongation processes and can be overexpressed by utilizing the CRISPRa tool, which involves the binding of activators (Table 1; Figure 6). Additionally, genes such as *GhJAZ2*, *GhGAI3a*, *GhCPC*, *GhPHYA1*, *BIN2*, *GhGalT1*, *GhPRP5*, and *GhbHLH18* act as negative regulators in the fiber initiation and elongation processes. These genes can be targeted for knockout or silencing using CRISPR/Cas9 or Cpf1 tools (Table 1; Figure 6).

Several genes have been identified as potential targets for improving cottonseed quality through the use of CRISPR/Cas tools (Table 2; Figure 6). The CGP1 and  $\Delta$ -CADINENE SYNTHASE (*CAD1*) genes, which are involved in gossypol biosynthesis, can be effectively knocked down using the CRISPRi tool by targeting the seed-specific promoter. This approach allows for the reduction of gene expression without affecting the gossypol levels in other plant tissues. The knockout of STEAROYL-ACYL-CARRIER PROTEIN  $\Delta$ 9-DESATURASE (*GhSAD-1*) and oleoylphosphatidylcholine  $\omega$ 6-DESATURASE (*GhFAD2-1*) genes, which are involved in fatty acid desaturation, through CRISPR/Cas9 or Cpf1-mediated approaches, can enhance the oleic acid content in cottonseed oil. Overexpression of the LYSOPHOSPHATIDIC ACID ACYLTRANSFERASE (*LPAAT*) gene using CRISPRa can increase the oil content in cottonseed, as this gene plays a crucial role in cottonseed oil biosynthesis. Cottonseed can be biofortified using the CRISPRa tool by targeting the PHYTOENE SYNTHASE (*GhPSY2D*) gene, resulting in enhanced Pro-vitamin A content. This biofortification strategy holds promise for addressing vitamin A deficiency on a global scale.

The previously reported genes associated with cotton plant architecture, which regulate flowering, branching in reproductive and vegetative growth, rooting pattern, and leaf shape, represent potential targets for CRISPR/Cas-based editing. These gene modifications aim to improve cotton plant architecture and increase productivity (Table 3; Figure 6). Roots play a crucial role in the plant's defense system against belowground biotic stresses and serve as sensors for water, nutrients, and environmental conditions. The ARGINASE (*GhARG*) gene can be targeted for knockout using CRISPR/Cas9 or Cpf1 tools to enhance lateral root formation in cotton. This genetic modification can lead to an increase in total root surface area, effectively stimulating both vegetative and generative growth of the entire cotton plant. Ultimately, this improvement can result in increased fiber productivity, especially under drought and low nutrient conditions.

In plants, the leaf is a vital organ responsible for generating food through photosynthesis. Leaf shape plays a crucial role in sunlight interception. Therefore, the overexpression of genes such as LATE

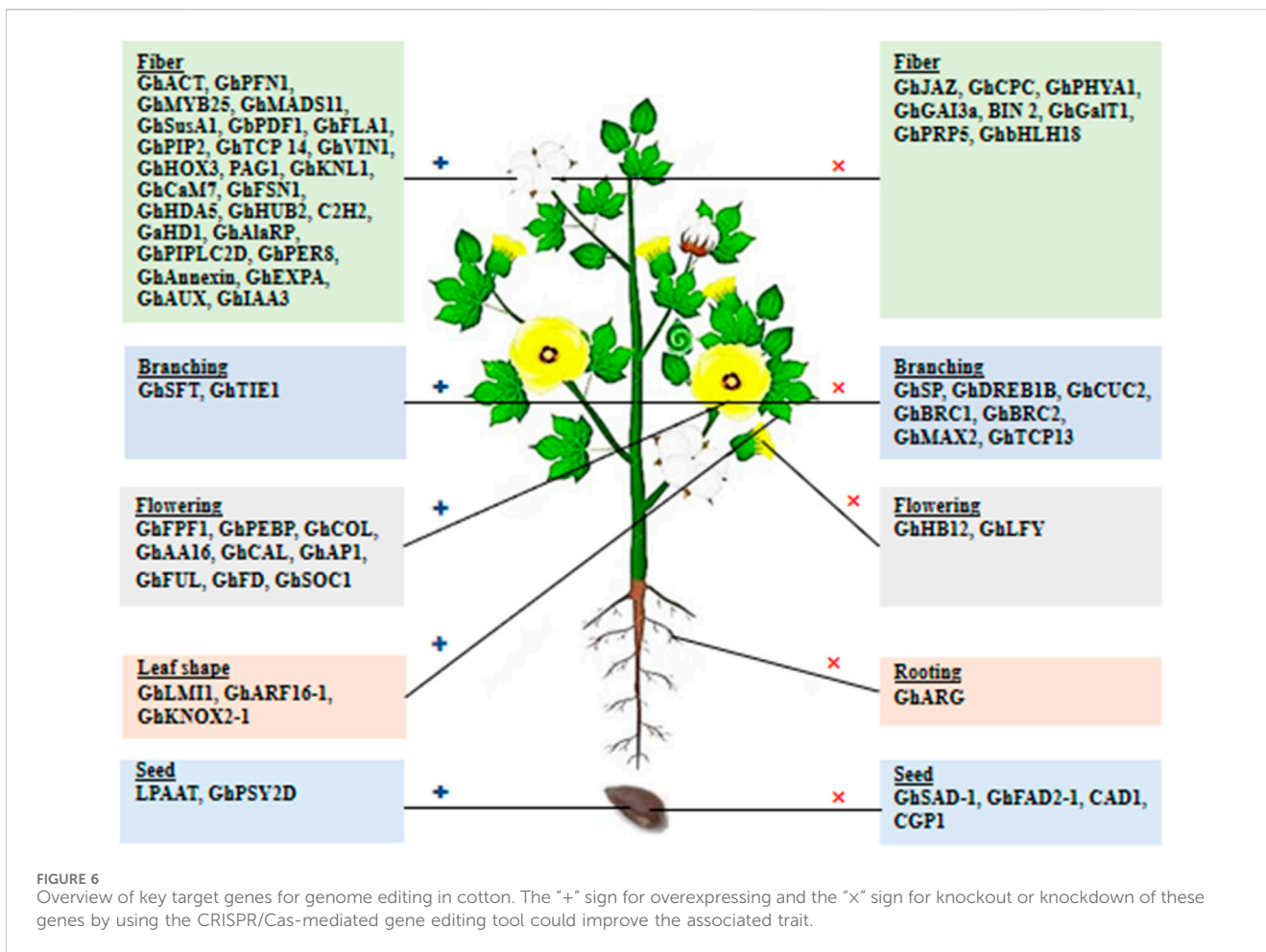


FIGURE 6 Overview of key target genes for genome editing in cotton. The "+" sign for overexpressing and the "x" sign for knockout or knockdown of these genes by using the CRISPR/Cas-mediated gene editing tool could improve the associated trait.

MERISTEM IDENTITY1 (*GhLMI1*)-like, AUXIN RESPONSE FACTOR (*GhARF16-1*), and *GhKNOX2-1* genes using CRISPRa tools can improve leaf shape in cotton. This enhancement leads to increased light interception in the leaves, ultimately improving crop productivity. Several reported transcription factor genes involved in branching regulation represent potential targets for CRISPR/Cas-based genome editing to enhance crop management and yield in cotton (Table 3; Figure 6). By using the CRISPRa tool, the expression of *GhSFT*, *GhSP*, and *GhTIE1* genes can be increased, promoting indeterminate and determinate growth by modulating monopodial and sympodial branching patterns. Additionally, CRISPR/Cas9 or Cas12a-mediated knockout of *GhDREB1B*, *GhCUC2*, *GhBRC1*, *GhBRC2*, *GhMAX2*, and *GhTCP13* genes could induce monopodial and sympodial branching in cotton.

Several reported transcription factor genes involved in flower organ development and flowering time in cotton represent potential targets for editing through CRISPR/Cas tools to induce early maturity (Table 3; Figure 6). The overexpression of genes such as *GhFPP1*, *GhPEBP*, *GhCOL*, *GhAA16*, *GhCAL*, *GhAPI*, *GhFUL*, *GhFD*, and *GhSOC1* using the CRISPRa tool, along with the knockout of *GhHB12* and *GhLFY* genes through CRISPR/Cas9 or Cas12a-mediated approaches, can effectively promote the early transition from generative to reproductive growth (early flowering) in cotton. Through CRISPR/Cas-based knockout of negative regulator genes and overexpression of positive regulator

genes associated with output traits, significant improvements in the phenotype of cotton plants can be achieved, facilitating desired trait enhancements.

## 10 Future prospective

Despite conventional breeding approaches yielding improved traits in the cotton crop, their lengthy breeding cycles, limited fidelity in hybridization, high level of heterozygosity, and infrequent occurrence of desirable mutations require substantial resources for the development of new cotton varieties. Transgenic approaches have been utilized for cotton improvement due to their ability to overcome compatibility barriers between cotton species. However, these approaches face challenges such as public acceptance, high costs, and time-consuming processes due to strict biosafety regulations. The emergence of CRISPR/Cas tools now offers a promising solution, enabling researchers to generate DNA-free edited crop plants that may be more readily accepted by the public compared to transgenic methods. The simplicity, versatility, and robustness of CRISPR/Cas tools have addressed many drawbacks associated with genome editing, leading to renewed approaches for enhancing cotton crop improvement in terms of precision, efficiency, and time savings. These tools facilitate gene

knockout, knock-in, replacement, base-editing, fine-tuning of gene regulation, and epigenetic modifications. However, further research is required to fully understand the potential of CRISPR tools in cotton. By leveraging CRISPR/Cas technology, it becomes possible to enhance agronomic traits, significantly improving total cotton productivity and quality. To tackle the complexity of agronomic traits, multiplex genome editing allows for the efficient stacking of multiple desirable traits in cotton plants. Ultimately, CRISPR/Cas-based genome editing in cotton has the potential to increase food, feed, fuel, and fiber production sustainably, thereby contributing to global food security in the future.

## Author contributions

MSS: Writing—original draft. SHK: Writing—review and editing. AA: Writing—review and editing. IAR: Writing—review and editing. ZAN: Writing—review and editing. AIK: Writing—review and editing.

## References

- Abdurakhmonov, I. Y., Ayubov, M. S., Ubaydullaeva, K. A., Buriev, Z. T., Shermatov, S. E., Ruziboev, H. S., et al. (2016). RNA interference for functional genomics and improvement of cotton (*Gossypium* sp). *Front. Plant Sci.* 7, 202. doi:10.3389/fpls.2016.00202
- Abdurakhmonov, I. Y., Buriev, Z. T., Saha, S., Jenkins, J. N., Abdurakimov, A., and Pepper, A. E. (2014). Phytochrome RNAi enhances major fiber quality and agronomic traits of the cotton *Gossypium hirsutum* L. *Nat. Commun.* 5 (1), 3062. doi:10.1038/ncomms4062
- Adli, M. (2018). The CRISPR tool kit for genome editing and beyond. *Nat. Commun.* 9 (1), 1911. doi:10.1038/s41467-018-04252-2
- Ahmad, S., and Hasanuzzaman, M. (2020). *Cotton production and uses: agronomy, crop protection, and postharvest technologies*. Springer.
- Allard, R. W. (2002). *Principles of plant breeding, 2nd edn*, by R. W. ALLARD. ix+254 pp. Chichester: John Wiley and Sons, Inc. (1999). £48.50 (hardback). ISBN 0 471 02309 4. *J. Agric. Sci.* 135, 331–334. doi:10.1017/s002185969921814x
- Álvarez, I., and Wendel, J. F. (2006). Cryptic interspecific introgression and genetic differentiation within *Gossypium aridum* (malvaceae) and its relatives. *Evolution* 60 (3), 505–517. doi:10.1111/j.0014-3820.2006.tb01132.x
- Andres, R. J., Coneva, V., Frank, M., Tuttle, J. R., Samayoa, L. F., Han, S., et al. (2016). Modifications to a Late Meristem IDENTITY1 gene are responsible for the major leaf shapes of Upland cotton (*Gossypium hirsutum* L.). *PNAS* 114 (1), E57–E66. doi:10.1073/pnas.1613593114
- Arora, N. K. (2019). Impact of climate change on agriculture production and its sustainable solutions. *Environ. Sustain* 2 (2), 95–96. doi:10.1007/s42398-019-00078-w
- Ashraf, S., Ahmad, A., Khan, S. H., Jamil, A., Sadia, B., and Brown, J. K. (2023). LbCas12a mediated suppression of Cotton leaf curl Multan virus. *Front. Plant Sci.* 14, 1233295. doi:10.3389/fpls.2023.1233295
- Aslam, S., Khan, S. H., Ahmad, A., Walawage, S. L., and Dandekar, A. M. (2022). Founder transformants of cotton (*Gossypium hirsutum* L.) obtained through the introduction of DS-Red, Rec, Rep, and CRISPR/Cas9 expressing constructs for developing baselines of recombinase mediated gene stacking. *PLoS One* 17 (2), e0263219. doi:10.1371/journal.pone.0263219
- Aslam, S., Khan, S. H., Ahmed, A., and Dandekar, A. M. (2020). The tale of cotton plant: from wild type to domestication, leading to its improvement by genetic transformation. *Am. J. Mol. Biol.* 10 (02), 91–127. doi:10.4236/ajmb.2020.102008
- Bassam, N. E. (2001). Renewable energy for rural communities. *Renew. Energy* 24 (3–4), 401–408. doi:10.1016/s0960-1481(01)00022-2
- Billon, P., Bryant, E. E., Joseph, S. A., Nambiar, T. S., Hayward, S. B., Rothstein, R., et al. (2017). CRISPR-mediated base editing enables efficient disruption of eukaryotic genes through induction of STOP codons. *Mol. Cell* 67 (6), 1068–1079. doi:10.1016/j.molcel.2017.08.008
- Cai, D., Liu, H., Sang, N., and Huang, X. (2017). Identification and characterization of CONSTANS-like (COL) gene family in upland cotton (*Gossypium hirsutum* L.). *PLoS ONE* 12 (6), e0179038. doi:10.1371/journal.pone.0179038
- Cao, Y., Huang, H., Yu, Y., Dai, H., Hao, H., Zhang, H., et al. (2021). A modified actin (Gly65Val substitution) expressed in cotton disrupts polymerization of actin filaments leading to the phenotype of ligon lintless-1 (Li1) mutant. *Int. J. Mol. Sci.* 22, 3000. doi:10.3390/ijms22063000
- Carroll, D. (2014). Genome engineering with targetable nucleases. *Annu. Rev. Biochem.* 83 (1), 409–439. doi:10.1146/annurev-biochem-060713-035418
- Chakravarthy, V. S. K., Reddy, T. P., Reddy, V. D., and Rao, K. V. (2012). Current status of genetic engineering in cotton (*Gossypium hirsutum* L.): an assessment. *Crit. Rev. Biotechnol.* 34 (2), 144–160. doi:10.3109/07388551.2012.743502
- Chaudry, R. M., and Guitchounts, A. (2022). *Cotton facts*. Washington, DC: International Cotton Advisory Committee.
- Chen, X., Lu, X., Shu, N., Wang, S., Wang, J., Wang, D., et al. (2017). Targeted mutagenesis in cotton (*Gossypium hirsutum* L.) using the CRISPR/Cas9 system. *Sci. Rep.* 7 (1), 44304. doi:10.1038/srep44304
- Chen, Y., Fu, M., Li, H., Wang, L., Liu, R., Liu, Z., et al. (2020). High-oleic acid content, nontransgenic allotetraploid cotton (*Gossypium hirsutum* L.) generated by knockout of *GhFAD2* genes with CRISPR/Cas9 system. *Plant Biotechnol. J.* 19 (3), 424–426. doi:10.1111/pbi.13507
- Cheng, S., Chen, P., Su, Z., Ma, L., Hao, P., Zhang, J., et al. (2020). High-resolution temporal dynamic transcriptome landscape reveals a GhCAL-mediated flowering regulatory pathway in cotton (*Gossypium hirsutum* L.). *Plant Biotechnol. J.* 19 (1), 153–166. doi:10.1111/pbi.13449
- Cheng, X., Wang, H., Wei, H., Gu, L., Hao, P., Sun, H., et al. (2021). The MADS transcription factor *GhAPI.7* coordinates the flowering regulatory pathway in upland cotton (*Gossypium hirsutum* L.). *Gene* 769, 145235. doi:10.1016/j.gene.2020.145235
- Cheng, Y., Lu, L., Yang, Z., Wu, Z., Qin, W., Yu, D., et al. (2016). GhCam7-like, a calcium sensor gene, influences cotton fiber elongation and biomass production. *Plant Physiol. Biochem.* 109, 128–136. doi:10.1016/j.plaphy.2016.09.009
- C H Knight Limited (2024). *Using cotton byproducts in beef cattle diets*. Georgia, United States: UGA Cooperative Extension. Available at: <https://extension.uga.edu/>.
- Christian, M., Čermák, T., Doyle, E. L., Schmidt, C. H., Zhang, F., Hummel, A. W., et al. (2010). Targeting DNA double-strand breaks with TAL effector nucleases. *Genetics* 186 (2), 757–761. doi:10.1534/genetics.110.120717
- Cody, W. B., and Scholthof, K. G. (2019). Plant virus vectors 3.0: transitioning into synthetic genomics. *Annu. Rev. Phytopathol.* 57 (1), 211–230. doi:10.1146/annurev-phyto-082718-100301
- Cong, L., Ran, F. A., Cox, D., Lin, S., Barretto, R. P. J., Habib, N., et al. (2013). Multiplex genome engineering using CRISPR/CAS systems. *Science* 339 (6121), 819–823. doi:10.1126/science.1231143
- Cotton Incorporated (2020). Cotton fiber development and processing an illustrated overview. Available at: <https://www.cottoninc.com/quality-products/textile-resources/cotton-fiber-development-and-processing/>.

## Funding

The author(s) declare that no financial support was received for the research, authorship, and/or publication of this article.

## Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

## 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.

- Curtin, S. J., Voytas, D. F., and Stupar, R. M. (2012). Genome engineering of crops with designer nucleases. *Plant Genome* 5 (2), 42–50. doi:10.3835/plantgenome2012.06.0008
- Deng, F., Tu, L., Tan, J., Yang, L., Nie, Y., and Zhang, X. (2011). *GBPDF1* is involved in cotton fiber initiation via the core CIS-Element HDZIP2ATATHB2. *Plant Physiol.* 158 (2), 890–904. doi:10.1104/pp.111.186742
- D'Halluin, K., Vanderstraeten, C., Van Hulle, J., Rosolowska, J., Van Den Brande, L., Pennewaert, A., et al. (2013). Targeted molecular trait stacking in cotton through targeted double-strand break induction. *Plant Biotechnol. J.* 11 (8), 933–941. doi:10.1111/pbi.12085
- Diao, Y., Zhan, J., Zhao, Y., Liu, L., Li, P., Wei, X., et al. (2019). *GHTIE1* regulates branching through modulating the transcriptional activity of TCPs in cotton and arabidopsis. *Front. Plant Sci.* 10, 1348. doi:10.3389/fpls.2019.01348
- Dilnur, T., Peng, Z., Pan, Z., Palanga, K. K., Jia, Y., Gong, W., et al. (2019). Association analysis of salt tolerance in asiatic cotton (*Gossypium arboreum*) with SNP markers. *Int. J. Mol. Sci.* 20 (9), 2168. doi:10.3390/ijms20092168
- Ding, M., Cao, Y., He, S., Sun, J., Hirata, D., Zhang, H., et al. (2020). *GaHDI1*, a candidate gene for the *Gossypium arboreum* SMA-4 mutant, promotes trichome and fiber initiation by cellular H<sub>2</sub>O<sub>2</sub> and Ca<sup>2+</sup> signals. *Plant Mol. Biol.* 103 (4–5), 409–423. doi:10.1007/s11103-020-01000-3
- Dutt, Y., Wang, X. D., Zhu, Y., and Li, Y. Y. (2004). Breeding for high yield and fibre quality in coloured cotton. *Plant Breed.* 123 (2), 145–151. doi:10.1046/j.1439-0523.2003.00938.x
- Economic Survey of Pakistan (2017). *Ministry of food, agriculture and livestock, economic advisor wing, islamabad, Pakistan.*
- Ellison, E. E., Nagalakshmi, U., Gamo, M. E., Huang, P. J., Dinesh-Kumar, S. P., and Voytas, D. F. (2020). Multiplexed heritable gene editing using RNA viruses and mobile single guide RNAs. *Nat. Plants* 6 (6), 620–624. doi:10.1038/s41477-020-0670-y
- Fang, D. D., Naoumkina, M., Thyssen, G. N., Bechere, E., Li, P., and Florane, C. B. (2020). An EMS-induced mutation in a tetratricopeptide repeat-like superfamily protein gene (Ghir\_A12G008870) on chromosome A12 is responsible for the liy short fiber phenotype in cotton. *Theor. Appl. Genet.* 133, 271–282. doi:10.1007/s00122-019-03456-4
- Feng, H., Li, X., Chen, H., Deng, J., Zhang, C., Liu, J., et al. (2018). *GhHUB2*, a ubiquitin ligase, is involved in cotton fiber development via the ubiquitin–26S proteasome pathway. *J. Exp. Bot.* 69 (21), 5059–5075. doi:10.1093/jxb/ery269
- Frisch, M., and Melchinger, A. E. (2005). Selection theory for marker-assisted backcrossing. *Genetics* 170 (2), 909–917. doi:10.1534/genetics.104.035451
- Gaba, Y., Pareek, A., and Singla-Pareek, S. L. (2021). Raising climate-resilient crops: journey from the conventional breeding to new breeding approaches. *Curr. genomics* 22 (6), 450–467. doi:10.2174/1389202922666210928151247
- Gadelha, I. C. N., Fonseca, N. B. S., Oloris, S. C. S., Melo, M. M., and Soto-Blanco, B. (2014). Gossypol toxicity from cottonseed products. *Sci. World J.* 2014, 231635. doi:10.1155/2014/231635
- Gao, J., Shen, L., Yuan, J., Zheng, H., Su, Q., Yang, W., et al. (2019b). Functional analysis of *GhCHS*, *GhANR* and *GhLAR* in colored fiber formation of *Gossypium hirsutum* L. *BMC Plant Biol.* 19 (1), 455. doi:10.1186/s12870-019-2065-7
- Gao, J., Shi, Y., Wang, W., Wang, Y. H., Yang, H., Shi, Q. H., et al. (2021). Genome sequencing identified novel mechanisms underlying virescent mutation in upland cotton *Gossypium hirsutum*. *BMC Genomics* 22, 498. doi:10.1186/s12864-021-07810-z
- Gao, W., Long, L., Tian, X., Xu, F., Liu, J., Singh, P. K., et al. (2017). Genome editing in cotton with the CRISPR/Cas9 system. *Front. Plant Sci.* 8, 1364. doi:10.3389/fpls.2017.01364
- Gao, W., Xu, F., Long, L., Li, Y., Zhang, J., Chong, L., et al. (2020). The gland localized CGP1 controls gland pigmentation and gossypol accumulation in cotton. *Plant Biotechnol. J.* 18 (7), 1573–1584. doi:10.1111/pbi.13323
- Gao, Z., Sun, W., Wang, J., Zhao, C., and Zuo, K. (2019a). *GhbHLH18* negatively regulates fiber strength and length by enhancing lignin biosynthesis in cotton fibers. *Plant Sci.* 286, 7–16. doi:10.1016/j.plantsci.2019.05.020
- Ge, X., Xu, J., Yang, Z., Yang, X., Wang, Y., Chen, Y., et al. (2023). Efficient genotype-independent cotton genetic transformation and genome editing. *J. Integr. Plant Biol.* 65 (4), 907–917. doi:10.1111/jipb.13427
- Gilbert, L. A., Larson, M. H., Morsut, L., Liu, Z., Brar, G. A., Torres, S. E., et al. (2013). CRISPR-Mediated modular RNA-Guided regulation of transcription in eukaryotes. *Cell* 154 (2), 442–451. doi:10.1016/j.cell.2013.06.044
- GITNEX (2024). Global textile industry statistics. Available at: <https://gitnux.org/global-textile-industry/>.
- Gong, S., Huang, G., Sun, X., Qin, L., Yang, L., Zhou, L., et al. (2014). Cotton KNL1, encoding a class II KNOX transcription factor, is involved in regulation of fibre development. *J. Exp. Bot.* 65 (15), 4133–4147. doi:10.1093/jxb/eru182
- Guo, H., Zhang, L., Guo, H., Cui, X., Fan, Y., Li, T., et al. (2024). Single-cell transcriptome atlas reveals somatic cell embryogenic differentiation features during regeneration. *Plant Physiol.* 195 (2), 1414–1431. doi:10.1093/plphys/kiad107
- Guo, W., Guo, D., Li, F., Shang, S., Li, T., Tang, Y. A., et al. (2023). Efficient genome editing in cotton using the virus-mediated CRISPR/Cas9 and grafting system. *Plant Cell Rep.* 42 (11), 1833–1836. doi:10.1007/s00299-023-03061-2
- Hajiahmadi, Z., Movahedi, A., Wei, H., Li, D., Orooji, Y., Ruan, H., et al. (2019). Strategies to increase On-Target and reduce Off-Target effects of the CRISPR/CAS9 system in plants. *Int. J. Mol. Sci.* 20 (15), 3719. doi:10.3390/ijms20153719
- Hannon, G. J. (2002). RNA interference. *Nature* 418, 244–251. doi:10.1038/418244a
- He, P., Zhang, Y., Li, H., Fu, X., Shang, H., Zou, C., et al. (2021). *GhARF16-1* modulates leaf development by transcriptionally regulating the *GhKNOX2-1* gene in cotton. *Plant Biotechnol. J.* 19 (3), 548–562. doi:10.1111/pbi.13484
- He, X., Wang, T., Xu, Z., Liu, N., Wang, L., Hu, Q., et al. (2018). The cotton HD-Zip transcription factor *GhHB12* regulates flowering time and plant architecture via the GhmiR157-GhSPL pathway. *Commun. Biol.* 1 (1), 229. doi:10.1038/s42003-018-0234-0
- Henikoff, S., and Comai, L. (2003). Single-Nucleotide mutations for plant functional genomics. *Annu. Rev. Plant Biol.* 54 (1), 375–401. doi:10.1146/annurev.arplant.54.031902.135009
- Hernandez, E. (2016). *Cottonseed: overview*. Elsevier eBooks, 242–246.
- Hu, H., He, X., Tu, L., Zhu, L., Zhu, S., Ge, Z., et al. (2016). *GhJAZ2* negatively regulates cotton fiber initiation by interacting with the R2R3-MYB transcription factor GhMYB25-like. *Plant J* 88 (6), 921–935. doi:10.1111/tpj.13273
- Hu, J., Li, S., Li, Z., Li, H., Song, W., Zhao, H., et al. (2019). A barley stripe mosaic virus-based guide RNA delivery system for targeted mutagenesis in wheat and maize. *Mol. Plant Pathol.* 20 (10), 1463–1474. doi:10.1111/mpp.12849
- Huang, G., Gong, S., Xu, W., Li, W., Li, P., Zhang, C., et al. (2013). A Fasciclin-Like arabinogalactan protein, *GhFLA1*, is involved in fiber initiation and elongation of cotton. *Plant Physiol.* 161 (3), 1278–1290. doi:10.1104/pp.112.203760
- International Feed Industry Federation (2024). Global feed statistics. Available at: <https://ifif.org/>.
- Jacobs, T. B., Zhang, N., Patel-Tupper, D., and Martin, G. B. (2017). Generation of a collection of mutant tomato lines using pooled CRISPR libraries. *Plant Physiol.* 174 (4), 2023–2037. doi:10.1104/pp.17.00489
- James, C. (2011). *Global status of commercialized biotech/GM crops*. NY: ISAAA. Brief No. 43.
- Ji, G., Liang, C., Cai, Y., Pan, Z., Meng, Z., Li, Y., et al. (2020). A copy number variant at the HPDA-D12 locus confers compact plant architecture in cotton. *New Phytol.* 229 (4), 2091–2103. doi:10.1111/nph.17059
- Jiang, Y., Guo, W., Zhu, H., Ruan, Y., and Zhang, T. (2012). Overexpression of *GhSusA1* increases plant biomass and improves cotton fiber yield and quality. *Plant Biotechnol. J.* 10 (3), 301–312. doi:10.1111/j.1467-7652.2011.00662.x
- Kang, M., Zuo, Z., Yin, Z., and Gu, J. (2022). Molecular mechanism of D1135E-Induced Discriminated CRISPR-CAS9 PAM recognition. *J. Chem. Inf. Model.* 62 (12), 3057–3066. doi:10.1021/acs.jcim.1c01562
- Khan, Z., Khan, S. H., Mubarik, M. S., Sadia, B., and Ahmad, A. (2016). Use of TALEs and TALEN technology for genetic improvement of plants. *Plant Mol. Biol.* 35 (1), 1–19. doi:10.1007/s11105-016-0997-8
- Komor, A. C., Kim, Y. B., Packer, M. S., Zuris, J. A., and Liu, D. R. (2016). Programmable editing of a target base in genomic DNA without double-stranded DNA cleavage. *Nature* 533 (7603), 420–424. doi:10.1038/nature17946
- Kumar, M., Tomar, M., Punia, S., Grasso, S., Arruti, F., Choudhary, J., et al. (2021). Cottonseed: a sustainable contributor to global protein requirements. *Trends Food Sci. Technol.* 111, 100–113. doi:10.1016/j.tifs.2021.02.058
- Kumar, V., Singh, B., Singh, S. K., Mohan, K., Singh, S., Sable, A., et al. (2018). Role of *GhHDA5* in H3K9 deacetylation and fiber initiation in *Gossypium hirsutum*. *Plant J* 95 (6), 1069–1083. doi:10.1111/tpj.14011
- Lange, M., Yellina, A. L., Orashakova, S., and Becker, A. (2013). Virus-induced gene silencing (VIGS) in plants: an overview of target species and the virus-derived vector systems. *Methods Mol. Biol.* 975, 1–14. doi:10.1007/978-1-62703-278-0\_1
- Lei, J., Dai, P., Li, J., Yang, M. L., Li, X., Zhang, W., et al. (2020). Tissue-specific CRISPR/Cas9 system of cotton pollen with GhPLIMP2b and GhMYB24 promoters. *J. Plant Biol.* 64 (1), 13–21. doi:10.1007/s12374-020-09272-4
- Lei, J., Li, Y., Dai, P., Liu, C., Zhao, Y., You, Y., et al. (2022). Efficient virus-mediated genome editing in cotton using the CRISPR/Cas9 system. *Front. Plant Sci.* 13, 1032799. doi:10.3389/fpls.2022.1032799
- Li, B., Fu, C., Zhou, J., Hui, F., Wang, Q., Wang, F., et al. (2022). Highly efficient genome editing using Geminivirus-Based CRISPR/CAS9 system in cotton plant. *Cells* 11 (18), 2902. doi:10.3390/cells11182902
- Li, B., Liang, S., Alariqi, M., Wang, F., Wang, G., Wang, Q., et al. (2020). The application of temperature sensitivity CRISPR/LbCpf1 (LbCas12a) mediated genome editing in allotetraploid cotton (*G. hirsutum*) and creation of nontransgenic, gossypol-free cotton. *Plant Biotechnol. J.* 19 (2), 221–223. doi:10.1111/pbi.13470
- Li, B., Rui, H., Li, Y., Wang, Q., Alariqi, M., Qin, L., et al. (2019b). Robust CRISPR/Cpf1 (Cas12a)-mediated genome editing in allotetraploid cotton (*Gossypium hirsutum*). *Plant Biotechnol. J.* 17 (10), 1862–1864. doi:10.1111/pbi.13147

- Li, C., Ünver, T., and Zhang, B. (2017). A high-efficiency CRISPR/Cas9 system for targeted mutagenesis in Cotton (*Gossypium hirsutum* L.). *Sci. Rep.* 7 (1), 43902. doi:10.1038/srep43902
- Li, D., Ruan, X., Zhang, J., Yu, W., Wang, X., and Li, X. (2013). Cotton plasma membrane intrinsic protein 2s (PIP2s) selectively interact to regulate their water channel activities and are required for fibre development. *New Phytol.* 199 (3), 695–707. doi:10.1111/nph.12309
- Li, J., Manghwar, H., Liang, S., Wang, P., Wang, G., Sheng, H., et al. (2019a). Whole genome sequencing reveals rare off-target mutations and considerable inherent genetic or/and somaclonal variations in CRISPR/Cas9-edited cotton plants. *Plant Biotechnol. J.* 17 (5), 858–868. doi:10.1111/pbi.13020
- Li, P., Liu, Q., Wei, Y., Xing, C., Xu, Z., Ding, F., et al. (2024a). Transcriptional landscape of cotton roots in response to salt stress at single-cell resolution. *Plant Commun.* 5 (2), 100740. doi:10.1016/j.xplc.2023.100740
- Li, X., Shahzad, K., Guo, L., Qi, T., Zhang, X., Wang, H., et al. (2018). Using yield quantitative trait locus targeted SSR markers to study the relationship between genetic distance and yield heterosis in upland cotton (*Gossypium hirsutum*). *Plant Breed.* 138 (1), 105–113. doi:10.1111/pbr.12668
- Li, X. B., Fan, X. P., Wang, X. L., Cai, L., and Yang, W. (2005). The cotton ACTIN1 gene is functionally expressed in fibers and participates in fiber elongation. *Plant Cell* 17 (3), 859–875. doi:10.1105/tpc.104.029629
- Li, Y., Hua, N., Zhang, Z., Wu, Y., Jiang, J., Su, S., et al. (2011). A cotton gene encoding novel MADS-box protein is preferentially expressed in fibers and functions in cell elongation. *Acta Biochim. Biophys. Sin.* 43 (8), 607–617. doi:10.1093/abbs/gmr055
- Li, Y., Ma, H., Wu, Y., Ma, Y., Yang, J., et al. (2024b). Single-cell transcriptome atlas and regulatory dynamics in developing cotton anthers. *Adv. Sci.* 11 (3), e2304017. doi:10.1002/advs.202304017
- Lin, J., Chen, L., Wu, W., Guo, X., Yu, C., Xu, M., et al. (2023). Single-cell RNA sequencing reveals a hierarchical transcriptional regulatory network of terpenoid biosynthesis in cotton secretory glandular cells. *Mol. Plant* 16 (12), 1990–2003. doi:10.1016/j.molp.2023.10.008
- Liu, B., Zhu, Y., and Zhang, T. (2015). The R3-MYB gene GHCP2 negatively regulates cotton fiber elongation. *PLoS ONE* 10 (2), e0116272. doi:10.1371/journal.pone.0116272
- Liu, Q., Singh, S., and Green, A. (2002). High-stearic and high-oleic cottonseed oils produced by HairPin RNA-mediated post-transcriptional gene silencing. *Plant Physiol.* 129 (4), 1732–1743. doi:10.1104/pp.001933
- Long, L., Xu, F., Wang, C., Zhao, X., Yuan, M., Song, C. P., et al. (2023). Single-cell transcriptome atlas identified novel regulators for pigment gland morphogenesis in cotton. *Plant Biotechnol. J.* 21 (6), 1100–1102. doi:10.1111/pbi.14035
- Lu, H., and Myers, G. O. (2002). Genetic relationships and discrimination of ten influential Upland cotton varieties using RAPD markers. *Theor. Appl. Genet.* 105 (2), 325–331. doi:10.1007/s00122-002-0947-8
- Luo, S., Jin, L., Stoddard, T. J., Baltes, N. J., Demorest, Z. L., Clasen, B. M., et al. (2015). Non-transgenic plant genome editing using purified Sequence-Specific nucleases. *Mol. Plant* 8 (9), 1425–1427. doi:10.1016/j.molp.2015.05.012
- Luterbacher, C. E. S., and Luterbacher, J. S. (2015). Break it down! How scientists are making fuel out of plants. *Front. Young Minds* 3. doi:10.3389/frym.2015.00010
- Lv, Y., Zhao, G., Xie, Y., Owusu, A. G., Wu, Y., and Gao, J. (2023). Transcriptome and metabolome profiling unveil pigment formation variations in Brown cotton lines (*Gossypium hirsutum* L.). *Int. J. Mol. Sci.* 24 (6), 5249. doi:10.3390/ijms24065249
- Ma, D., Hu, Y., Yang, C., Liu, B., Fang, L., Wan, Q., et al. (2016). Genetic basis for glandular trichome formation in cotton. *Nat. Commun.* 7 (1), 10456. doi:10.1038/ncomms10456
- Ma, X., Zhang, X., Liu, H., and Li, Z. (2020). Highly efficient DNA-free plant genome editing using virally delivered CRISPR-Cas9. *Nat. Plants* 6 (7), 773–779. doi:10.1038/s41477-020-0704-5
- Machado, A. F., Wu, Y., Yang, Y., Llewellyn, D., and Dennis, E. S. (2009). The MYB transcription factor GhMYB25 regulates early fibre and trichome development. *Plant J* 59 (1), 52–62. doi:10.1111/j.1365-313x.2009.03847.x
- May, D. C., Paldi, K., and Altpeter, F. (2023). Targeted mutagenesis with sequence-specific nucleases for accelerated improvement of polyploid crops: progress, challenges, and prospects. *Plant Genome* 16 (2), e20298. doi:10.1002/tpg2.20298
- McGarry, R. C., Prewitt, S., Culpepper, S., Eshed, Y., Lifschitz, E., and Ayre, B. G. (2016). Monopodial and sympodial branching architecture in cotton is differentially regulated by the *Gossypium hirsutum* SINGLE FLOWER TRUSS and SELF-PRUNING orthologs. *New Phytol.* 212 (1), 244–258. doi:10.1111/nph.14037
- Meng, X., Yu, H., Zhang, Y., Zhuang, F., Song, X., Gao, S., et al. (2017). Construction of a Genome-Wide mutant library in Rice using CRISPR/CAS9. *Mol. Plant* 10 (9), 1238–1241. doi:10.1016/j.molp.2017.06.006
- Mikhailova, A. S., Strygina, K., and Khlestkina, E. K. (2019). The genes determining synthesis of pigments in cotton. *Commun. Biol.* 64 (2), 133–145. doi:10.21638/spbu03.2019.205
- Muthusamy, A., and Jayabalan, N. (2011). *In vitro* induction of mutation in cotton (*Gossypium hirsutum* L.) and isolation of mutants with improved yield and fiber characters. *Acta Physiol. Plant* 33 (5), 1793–1801. doi:10.1007/s11738-011-0718-8
- Naoumkina, M., Thyssen, G. N., Fang, D. D., Bechere, E., Li, P., and Florane, C. B. (2021). Mapping-by-sequencing the locus of EMS-induced mutation responsible for tufted-fuzzless seed phenotype in cotton. *Mol. Genet. Genomics* 296, 1041–1049. doi:10.1007/s00438-021-01802-0
- National Cotton Council of America (2024). *World of cotton*. Washington, DC: National Cotton Council of America. Available at: <https://www.cotton.org/econ/world/>.
- Papikian, A., Liu, W., Gallego-Bartolomé, J., and Jacobsen, S. E. (2019). Site-specific manipulation of *Arabidopsis* loci using CRISPR-Cas9 SunTag systems. *Nat. Commun.* 10 (1), 729. doi:10.1038/s41467-019-08736-7
- Parihar, A., Shiwani, S., Mondal, S., Singh, P. K., and Singh, R. L. (2022). Introduction, scope, and applications of biotechnology and genomics for sustainable agricultural production. *Plant Genomics Sustain. Agric.*, 1–14. doi:10.1007/978-981-16-6974-3\_1
- Pathirana, R. (2011). “Plant mutation breeding in agriculture,” in *Cab reviews: perspectives in agriculture, veterinary science, nutrition and natural resources*, 1–20. doi:10.1079/pavsnr20116032
- 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. doi:10.1016/j.tibtech.2020.09.001
- Pingali, P. (2012). Green Revolution: impacts, limits, and the path ahead. *PNAS* 109 (31), 12302–12308. doi:10.1073/pnas.0912953109
- Podevin, N., Devos, Y., Davies, H. V., and Nielsen, K. M. (2012). Transgenic or not? No simple answer. *EMBO Rep.* 13 (12), 1057–1061. doi:10.1038/embor.2012.168
- Puchta, H., and Höhn, B. (2010). Breaking news: plants mutate right on target. *PNAS* 107 (26), 11657–11658. doi:10.1073/pnas.1006364107
- Qanmber, G., Lu, L., Liu, Z., Yu, D., Zhou, K., Huo, P., et al. (2019). Genome-wide identification of *GhAAI* genes reveals that *GhAAI66* triggers a phase transition to induce early flowering. *J. Exp. Bot.* 70 (18), 4721–4736. doi:10.1093/jxb/erz239
- Qi, L. S., Larson, M. H., Gilbert, L. A., Doudna, J. A., Weissman, J. S., Arkin, A. P., et al. (2013). Repurposing CRISPR as an RNA-Guided platform for Sequence-Specific control of gene expression. *Cell* 152 (5), 1173–1183. doi:10.1016/j.cell.2013.02.022
- Qin, L., Chen, Y., Zeng, W., Li, Y., Gao, L., Li, D., et al. (2017). The cotton  $\beta$ -galactosyltransferase 1 (*GalT1*) that galactosylates arabinogalactan proteins participates in controlling fiber development. *Plant J* 89 (5), 957–971. doi:10.1111/tpj.13434
- Qin, L., Li, J., Wang, Q., Xu, Z., Liang, S., Alariqi, M., et al. (2019). High-efficient and precise base editing of C●G to T●A in the allotetraploid cotton (*Gossypium hirsutum*) genome using a modified CRISPR/Cas9 system. *Plant Biotechnol. J.* 18 (1), 45–56. doi:10.1111/pbi.13168
- Qin, Y., Sun, M., Li, W., Xu, M., Shao, L., Liu, Y., et al. (2022). Single-cell RNA-seq reveals fate determination control of an individual fibre cell initiation in cotton (*Gossypium hirsutum*). *Plant Biotechnol. J.* 20 (12), 2372–2388. doi:10.1111/pbi.13918
- Ramadan, M., Alariqi, M., Ma, Y., Li, Y., Liu, Z., Zhang, R., et al. (2021). Efficient CRISPR/Cas9 mediated Pooled-sgRNAs assembly accelerates targeting multiple genes related to male sterility in cotton. *Plant Methods* 17 (1), 16. doi:10.1186/s13007-021-00712-x
- Rathore, K. S., Pandeya, D., Campbell, L. M., Wedegaertner, T., Puckhaber, L. S., Stipanovic, R. D., et al. (2020). Ultra-Low Gossypol cottonseed: selective gene silencing opens up a vast resource of Plant-Based protein to improve human nutrition. *Crit. Rev. Plant Sci.* 39 (1), 1–29. doi:10.1080/07352689.2020.1724433
- Salih, H., Odongo, M. R., Gong, W., He, S., and Du, X. (2019). Genome-wide analysis of cotton C2H2-zinc finger transcription factor family and their expression analysis during fiber development. *BMC Plant Biol.* 19 (1), 400. doi:10.1186/s12870-019-2003-8
- Scheffler, J. A. (2016). Evaluating protective terpenoid aldehyde compounds in cotton (*Gossypium hirsutum* L.) roots. *Am. J. Plant Sci.* 07 (07), 1086–1097. doi:10.4236/ajps.2016.77104
- Shan, C., Shangguan, X. X., Zhao, B., Zhang, X. F., Chao, L., Yang, C., et al. (2014). Control of cotton fibre elongation by a homeodomain transcription factor *GhHOX3*. *Nat. Commun.* 5 (1), 5519. doi:10.1038/ncomms6519
- Si, Z., Liu, H., Zhu, J., Chen, J., Wang, Q., Fang, L., et al. (2018). Mutation of SELF-PRUNING homologs in cotton promotes short-branching plant architecture. *J. Exp. Bot.* 69, 2543–2553. doi:10.1093/jxb/ery093
- Silva, G. (2018). *Feeding the world in 2050 and beyond*. Michigan, United States: Michigan State University Extension. Available at: <https://www.canr.msu.edu/news/feeding-the-world-in-2050-and-beyond-part-1#>.
- Smith, W. C. (1999). “Production statistics,” in *Cotton origin, history, technology, and production*. Editors C. W. Smith and J. T. Cothorn (New York: Wiley), 435–449.
- Soomro, S. R., Soomro, S. N., Syed, S., Hassan, S., and Tabassum, B. (2023). “Cotton stalks: potential biofuel recourses for sustainable environment,” in *Biotechnology and omics approaches for bioenergy crops*, 203–234. doi:10.1007/978-981-99-4954-0\_10
- Statista (2022a). Cottonseed oil: market value worldwide 2021 and 2028. Available at: <https://www.statista.com/statistics/999312/cottonseed-oil-market-value-worldwide/>.
- Statista (2022b). Global cotton production 2021/2022, by country. Available at: <https://www.statista.com/statistics/263055/cotton-production-worldwide-by-top-countries/>.
- Sun, J., Xiong, X., Zhu, Q., and Li, Y. (2019). Transcriptome sequencing and metabolome analysis reveal genes involved in pigmentation of Green-Colored cotton fibers. *Int. J. Mol. Sci.* 20 (19), 4838. doi:10.3390/ijms20194838

- Sun, Q., Xie, Y., Li, H., Liu, J., Geng, R., Wang, P., et al. (2022). Cotton GhBRC1 regulates branching, flowering, and growth by integrating multiple hormone pathways. *Crop J.* 10 (1), 75–87. doi:10.1016/j.cj.2021.01.007
- Sun, Y., and Allen, R. D. (2005). Functional analysis of the BIN2 genes of cotton. *Mol. Genet.* 74 (1), 51–59. doi:10.1007/s00438-005-1122-0
- Sun, Y., Han, Y., Sheng, K., Yang, P., Cao, Y., Li, H., et al. (2023). Single-cell transcriptomic analysis reveals the developmental trajectory and transcriptional regulatory networks of pigment glands in *Gossypium bickii*. *Mol. Plant* 16 (4), 694–708. doi:10.1016/j.molp.2023.02.005
- Sunilkumar, G., Campbell, L. M., Puckhaber, L. S., Stipanovic, R. D., and Rathore, K. S. (2006). Engineering cottonseed for use in human nutrition by tissue-specific reduction of toxic gossypol. *PNAS* 103 (48), 18054–18059. doi:10.1073/pnas.0605389103
- United Nations (2017). *World population projected to reach 9.8 billion in 2050, and 11.2 billion in 2100*. New York, NY: United Nations. Available at: <https://www.un.org/en/desa/world-population-projected-reach-98-billion-2050-and-112-billion-2100>.
- Urnov, F. D., Rebar, E. J., Holmes, M. C., Zhang, H. S., and Gregory, P. D. (2010). Genome editing with engineered zinc finger nucleases. *Nat. Rev. Genet.* 11 (9), 636–646. doi:10.1038/nrg2842
- Walford, S., Wu, Y., Llewellyn, D., and Dennis, E. S. (2011). *GhMYB25-like*: a key factor in early cotton fibre development. *Plant J.* 65 (5), 785–797. doi:10.1111/j.1365-313x.2010.04464.x
- Walton, R. T., Christie, K. A., Whittaker, M. N., and Kleinstiver, B. P. (2020). Unconstrained genome targeting with near-PAMless engineered CRISPR-Cas9 variants. *Science* 368 (6488), 290–296. doi:10.1126/science.aba8853
- Wang, C., Liu, J., Xie, X., Wang, J., Ma, Q., Chen, P., et al. (2023a). *GhAPI-D3* positively regulates flowering time and early maturity with no yield and fiber quality penalties in upland cotton. *J. Integr. Plant Biol.* 65 (4), 985–1002. doi:10.1111/jipb.13409
- Wang, D., Hu, X., Ye, H., Wang, Y., Yang, Q., Liang, X., et al. (2023b). Cell-specific clock-controlled gene expression program regulates rhythmic fiber cell growth in cotton. *Genome Biol.* 24 (1), 49. doi:10.1186/s13059-023-02886-0
- Wang, G., Wang, F., Xu, Z., Wang, Y., Zhang, C., Zhou, Y., et al. (2024a). Precise fine-tuning of *GhTFL1* by base editing tools defines ideal cotton plant architecture. *Genome Biol.* 25 (1), 59. doi:10.1186/s13059-024-03189-8
- Wang, H., Yu, Y., Chen, Z., and Xia, G. (2005). Functional characterization of *Gossypium hirsutum* profilin 1 gene (*GhPFN1*) in tobacco suspension cells. Characterization of *in vivo* functions of a cotton profilin gene. *Planta* 222 (4), 594–603. doi:10.1007/s00425-005-0005-2
- Wang, J., Wang, H., Zhao, P., Han, L., Jiao, G., Zheng, Y., et al. (2010). Overexpression of a profilin (*GhPFN2*) promotes the progression of developmental phases in cotton fibers. *Plant Cell Physiol.* 51 (8), 1276–1290. doi:10.1093/pcp/pcq086
- Wang, L., Cook, A., Patrick, J. W., Chen, X., and Ruan, Y. (2014b). Silencing the vacuolar invertase gene *GhVIN1* blocks cotton fiber initiation from the ovule epidermis, probably by suppressing a cohort of regulatory genes via sugar signaling. *Plant J* 78 (4), 686–696. doi:10.1111/tpj.12512
- Wang, M., Zhao, P., Cheng, H., Han, L., Wu, X., Gao, P., et al. (2013). The cotton transcription factor TCP14 functions in Auxin-Mediated Epidermal cell differentiation and elongation. *Plant Physiol.* 162 (3), 1669–1680. doi:10.1104/pp.113.215673
- Wang, N., Ma, J., Pei, W., Wu, M., Li, H., Li, X., et al. (2017a). A genome-wide analysis of the lysophosphatidate acyltransferase (LPAAT) gene family in cotton: organization, expression, sequence variation, and association with seed oil content and fiber quality. *BMC Genomics* 18 (1), 218. doi:10.1186/s12864-017-3594-9
- Wang, W., Baiming, C., Yu, X., Chen, Q., Zheng, Y., Xia, Y., et al. (2012). Functional analysis of cotton DELLA-like genes that are differentially regulated during fiber development. *Plant Mol. Biol. Report.* 30 (4), 1014–1024. doi:10.1007/s11105-012-0412-z
- Wang, W., Yang, X., Zeng, J., Liang, A., Liu, D., et al. (2024b). Construction of ethyl methane sulfonate mutant library in *G. arboreum* and rapid identification of mutant genes via repeated re-sequencing. *Ind. Crops Prod.* 213, 118373. doi:10.1016/j.indcrop.2024.118373
- Wang, X., Fan, S., Song, M., Pang, C., Wei, H., Yu, J., et al. (2014a). Upland cotton gene GhFPF1 confers promotion of flowering time and shade-avoidance responses in *Arabidopsis thaliana*. *PLOS ONE* 9 (3), e91869. doi:10.1371/journal.pone.0091869
- Wang, Y., Meng, Z., Liang, C., Meng, Z., Wang, Y., Sun, G., et al. (2017b). Increased lateral root formation by CRISPR/Cas9-mediated editing of arginase genes in cotton. *Sci. China Life Sci.* 60 (5), 524–527. doi:10.1007/s11427-017-9031-y
- Weinthal, D., Tovkach, A., Zeevi, V., and Tzfira, T. (2010). Genome editing in plant cells by zinc finger nucleases. *Trends Plant Sci.* 15 (6), 308–321. doi:10.1016/j.tplants.2010.03.001
- Wen, X., Chen, Z., Yang, Z., Wang, M., Jin, S., et al. (2023). A comprehensive overview of cotton genomics, biotechnology and molecular biological studies. *Sci. China Life Sci.* 66 (10), 2214–2256. doi:10.1007/s11427-022-2278-0
- Wendel, J. F., and Grover, C. E. (2015). “Taxonomy and evolution of the cotton genus, *Gossypium*,” in *Agronomy*, 25–44. doi:10.2134/agronmonogr57.2013.0020
- Wilkins, T. A., Rajasekaran, K., and Anderson, D. (2000). Cotton biotechnology. *Crit. Rev. Plant Sci.* 19 (6), 511–550. doi:10.1080/07352680091139286
- Workman, D. (2022). “Cotton exports by country,” World’s top exports. Available at: <https://www.worldstopexports.com/cotton-exports-by-country/>.
- Xu, W., Zhang, D., Wu, Y., Qin, L., Huang, G., Li, J., et al. (2013). Cotton PRP5 gene encoding a proline-rich protein is involved in fiber development. *Plant Mol. Biol.* 82 (4–5), 353–365. doi:10.1007/s11103-013-0066-8
- Yan, Q., Wang, Y., Li, Q., Zhang, Z., Ding, H., Zhang, Y., et al. (2018). Up-regulation of *GhTT2-3A* in cotton fibres during secondary wall thickening results in brown fibres with improved quality. *Plant Biotechnol. J.* 16 (10), 1735–1747. doi:10.1111/pbi.12910
- Yang, Z., Zhang, C., Yang, X., Liu, K., Wu, Z., Zhang, X., et al. (2014). PAG1, a cotton brassinosteroid catabolism gene, modulates fiber elongation. *New Phytol.* 203 (2), 437–448. doi:10.1111/nph.12824
- Yao, D., Wang, Y., Li, Q., Ouyang, X., Li, Y., Wang, C., et al. (2018). Specific upregulation of a cotton phytoene synthase gene produces golden cottonseeds with enhanced provitamin A. *Sci. Rep.* 8 (1), 1348. doi:10.1038/s41598-018-19866-1
- Yin, X., and Ding, Y. (2021). Advanced technology in agronomy to secure food, fiber, feed, and fuel supply and maintain environmental sustainability. *Technol. Agron.* 1 (1), 1–2. doi:10.48130/tia-2021-0001
- Zetsche, B., Gootenberg, J. S., Abudayyeh, O. O., Slaymaker, I. M., Makarova, K. S., Essletzbichler, P., et al. (2015). CPF1 is a single RNA-Guided endonuclease of a Class 2 CRISPR-CAS system. *Cell* 163 (3), 759–771. doi:10.1016/j.cell.2015.09.038
- Zhan, J., Chu, Y., Wang, Y., Diao, Y., Zhao, Y., Liu, L., et al. (2021). The miR164-GhCUC2-GhBRC1 module regulates plant architecture through abscisic acid in cotton. *Plant Biotechnol. J.* 19 (9), 1839–1851. doi:10.1111/pbi.13599
- Zhang, J., Huang, G., Zou, D., Yan, J., Li, Y., Huang, S., et al. (2017a). The cotton (*Gossypium hirsutum*) NAC transcription factor (FSN1) as a positive regulator participates in controlling secondary cell wall biosynthesis and modification of fibers. *New Phytol.* 217 (2), 625–640. doi:10.1111/nph.14864
- Zhang, M., Han, L., Wang, W., Wu, S., Jiao, G., Su, L., et al. (2017b). Overexpression of *GhFIM2* propels cotton fiber development by enhancing actin bundle formation. *J. Integr. Plant Biol.* 59 (8), 531–534. doi:10.1111/jipb.12552
- Zhang, X., Wang, C., Pang, C., Wei, H., Wang, H., Song, M., et al. (2016). Characterization and functional analysis of PEBP family genes in upland cotton (*Gossypium hirsutum* L.). *PLOS ONE* 11 (8), e0161080. doi:10.1371/journal.pone.0161080
- Zhang, Y., Malzahn, A., Sretenovic, S., and Qi, Y. (2019). The emerging and uncultivated potential of CRISPR technology in plant science. *Nat. Plants* 5 (8), 778–794. doi:10.1038/s41477-019-0461-5
- Zhang, Z., Ge, X., Luo, X., Wang, P., Fan, Q., Hu, G., et al. (2018). Simultaneous editing of two copies of *gh14-3-3d* confers enhanced transgene-clean plant defense against *Verticillium dahliae* in allotetraploid upland cotton. *Front. Plant Sci.* 9, 842. doi:10.3389/fpls.2018.00842
- Zhao, L., Fang, J., Xing, J., Liu, W., Peng, P., Long, H., et al. (2017). Identification and functional analysis of two cotton orthologs of MAX2 which control shoot lateral branching. *Plant Mol. Biol. Report.* 35 (5), 480–490. doi:10.1007/s11105-017-1040-4
- Zhu, L., Liu, D., Shāng, H., Li, H., Yu, J., and Xiao, G. (2021a). *GhPIPLC2D* promotes cotton fiber elongation by enhancing ethylene biosynthesis. *iScience* 24 (3), 102199. doi:10.1016/j.isci.2021.102199
- Zhu, S., Li, Y., Zhang, X., Liu, F., Xue, F., Zhang, Y., et al. (2021b). *GhAlaRP*, a cotton alanine rich protein gene, involves in fiber elongation process. *Crop J.* 9 (2), 313–324. doi:10.1016/j.cj.2020.08.007
- Zhu, S., Yu, X., Li, Y., Sun, Y., Zhu, Q., and Sun, J. (2018). Highly efficient targeted gene editing in upland cotton using the CRISPR/CAS9 system. *Int. J. Mol. Sci.* 19 (10), 3000. doi:10.3390/ijms19103000
- Zhu, X., Xu, Z., Wang, G., Cong, Y., Yu, L., Jia, R., et al. (2023). Single-cell resolution analysis reveals the preparation for reprogramming the fate of stem cell niche in cotton lateral meristem. *Genome Biol.* 24 (1), 194. doi:10.1186/s13059-023-03032-6