Check for updates

# A comprehensive review on *Gossypium hirsutum* resistance against cotton leaf curl virus

Sahar Nadeem[1†], Syed Riaz Ahmed[1,2†], Tahira Luqman[1],
Daniel K. Y. Tan[3], Zahra Maryum[1], Khalid Pervaiz Akhtar[1],
Sana Muhy Ud Din Khan[1], Muhammad Sayyam Tariq[1],
Nazar Muhammad[4], Muhammad Kashif Riaz Khan[1,5]* and
Yongming Liu[6]*

[1]Nuclear Institute for Agriculture and Biology College, Pakistan Institute of Engineering and Applied
Sciences (NIAB-C, PIEAS), Faisalabad, Pakistan, [2]Pakistan Agriculture Research Council (PARC),
Horticulture Research Institute Khuzdar Baghbana, Khuzdar, Pakistan, [3]School of Life and Environmental
Sciences, Plant Breeding Institute, Sydney Institute of Agriculture, Faculty of Science, The University of
Sydney, Sydney, NSW, Australia, [4]Agriculture and Cooperative Department, Quetta, Pakistan, [5]Plant
Breeding and Genetics Division, Cotton Group, Nuclear Institute for Agriculture and Biology, Faisalabad,
Pakistan, [6]National Nanfan Research Institute (Sanya), Chinese Academy of Agricultural Sciences,
Sanya, China

Cotton (*Gossypium hirsutum* L.) is a significant fiber crop. Being a major
contributor to the textile industry requires continuous care and attention.
Cotton is subjected to various biotic and abiotic constraints. Among these,
biotic factors including cotton leaf curl virus (CLCuV) are dominant. CLCuV is a
notorious disease of cotton and is acquired, carried, and transmitted by the
whitefly (Bemisia tabaci). A cotton plant affected with CLCuV may show a wide
range of symptoms such as yellowing of leaves, thickening of veins, upward or
downward curling, formation of enations, and stunted growth. Though there
are many efforts to protect the crop from CLCuV, long-term results are not yet
obtained as CLCuV strains are capable of mutating and overcoming plant
resistance. However, systemic-induced resistance using a gene-based
approach remained effective until new virulent strains of CLCuV (like
Cotton Leaf Curl Burewala Virus and others) came into existence. Disease
control by biological means and the development of CLCuV-resistant cotton
varieties are in progress. In this review, we first discussed in detail the evolution
of cotton and CLCuV strains, the transmission mechanism of CLCuV, the
genetic architecture of CLCuV vectors, and the use of pathogen and
nonpathogen-based approaches to control CLCuD. Next, we delineate the
uses of cutting-edge technologies like genome editing (with a special focus
on CRISPR-Cas), next-generation technologies, and their application in
cotton genomics and speed breeding to develop CLCuD resistant cotton
germplasm in a short time. Finally, we delve into the current obstacles related
to cotton genome editing and explore forthcoming pathways for enhancing
precision in genome editing through the utilization of advanced genome
editing technologies. These endeavors aim to enhance cotton's resilience
against CLCuD.

KEYWORDS

Cotton (*G. hirsutum* L.), CLCuD, genome editing, next-generation technologies,
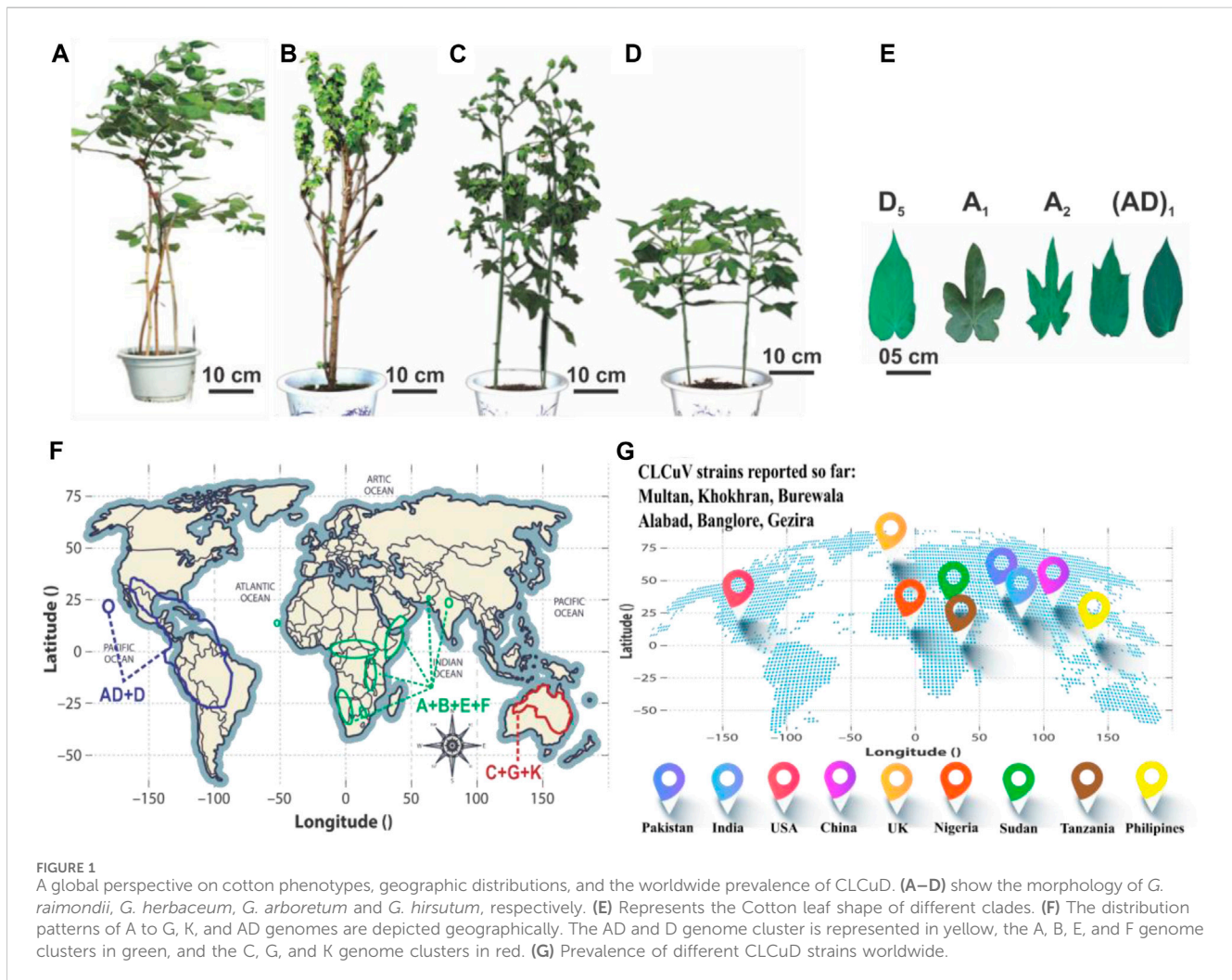speed breeding

GRAPHICAL ABSTRACT
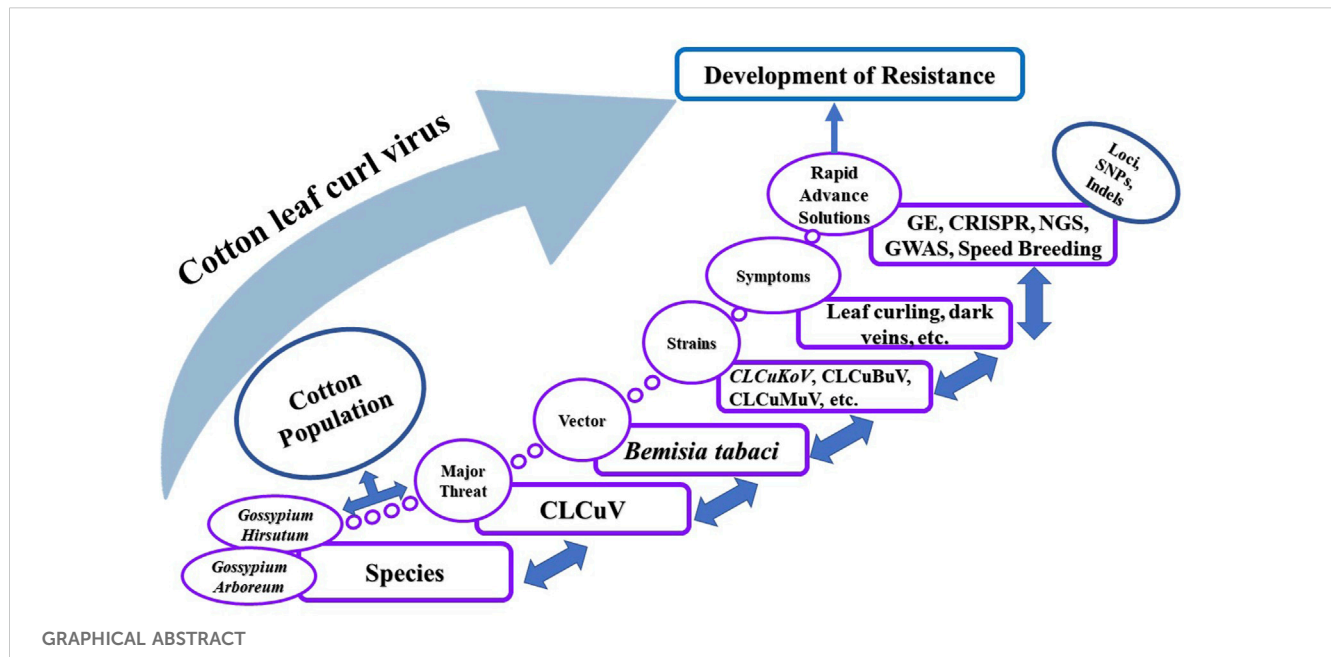
1 Introduction

Cotton (*Gossypium hirsutum* L.) stands as a high-quality fiber-producing plant that contributes greatly to the world's textile industry by generating an annual economic income of 600 billion US dollars (Ashraf et al., 2018). It serves as a model system for investigating plant polyploidization, cell wall biogenesis, and cell elongation (Lu et al., 2018). The genus *Gossypium* encompasses seven tetraploid species (2n = 4x = 52) and 45 diploid species (2n = 2x = 26), showcasing remarkable morphological diversity (Lu et al., 2018). This variation includes a spectrum of plant structures, spanning from untamed perennial trees and shrubs to cultivated annual herbaceous plants, accompanied by unique fiber traits and diverse leaf shapes. Cotton has captivated the interest of agricultural researchers, evolutionary biologists, and taxonomists for its remarkable genomic diversity and widespread dispersion. This diversity has given rise to the evolution of eight distinct diploid cotton groups, designated as A-, B-, C-, D-, E-, F-, G-, and K-genomes, along with an AD-genome clade. The classification of the genus *Gossypium* into three primary lineages delineated mainly by geographical and morphological factors, including the Australian clade (C-, G-, and K-genomes), the African-Asian clade (A-, B-, E-, and F-genomes), and the New World clade (D- and AD-genomes) (Figure 1) (Ahmed, 1999; Huang et al., 2021). Among these, *G. hirsutum* [(AD)$_1$], commonly called upland cotton, currently dominates the global cotton industry by contributing approximately 95% of the natural lint fiber essential for textiles.

The textile industries rely on good quality cotton lint fiber to fulfill market demands. Each clade has a separate fiber lint quality and yield. Therefore, the fiber lint is of great importance in cotton domestication and quality among cotton breeders. The progression of fiber development can be categorized into four successive phases: initiation, elongation, secondary cell wall (SCW) biosynthesis, and maturation, categorized based on the days post-anthesis (DPA)

(Prasad et al., 2022). Cotton fibers can be additionally classified into two forms: adherent fuzz fibers, which commence development around 5 to 10 DPA and reach a final length of about 5 mm, and spinnable lint fibers, originating before flowering and growing to a final length of roughly 3 cm.

Pakistan is the fourth-largest producer of cotton following India, China, and the United States of America (Shuli et al., 2018). Unfortunately, cotton production is continuously being challenged by various biotic and abiotic stresses, among which CLCuD is most prevalent. CLCuV has emerged as a prominent threat to cotton, causing major losses of up to 80%–87% in Pakistan and North India (Mansoor et al., 2003; Varma and Malathi, 2003; Sattar et al., 2013; Monga and Sain, 2021; Chauhan et al., 2023). The importance of cotton in the textile industry is undeniable, therefore, requires continuous attention to its improvement. The virus is a major cause of low productivity in cotton and hence, efforts are being made to induce resistance in cotton crop against the disease (Ali et al., 2015). The story of the emergence of CLCuV as a threat is not new. For the first time, it was reported in Nigeria in 1912. Later on, it was documented in Sudan in 1924, followed by occurrences in Tanzania in 1926, and later in the Philippines in 1959. In Pakistan, CLCuD was discovered in 1967 on scattered cotton plants of the *G. hirsutum* species in the Multan district (Hussain and Ali, 1975). However, the disease was not considered a serious threat until the 1990s, when it appeared as an epidemic and disastrously brought a monetary loss of $5 US billion, throughout the world (Briddon and Markham, 1994; Leke et al., 2015)

CLCuV is acquired, carried, and transmitted via whitefly (*Bemisia tabaci*) (Nogia et al., 2014). Once the vector acquires the virus, it is retained by the carrier whitefly throughout its life. Plants infected with CLCuD show a broad range of symptoms ranging from stunted growth, yellowing of leaves, and appearance of dark green thickened veins on leaves followed by upward or downward curling (Heigwer et al., 2016). In case of high

FIGURE 1
A global perspective on cotton phenotypes, geographic distributions, and the worldwide prevalence of CLCuD. **(A–D)** show the morphology of *G. raimondii*, *G. herbaceum*, *G. arboretum* and *G. hirsutum*, respectively. **(E)** Represents the Cotton leaf shape of different clades. **(F)** The distribution patterns of A to G, K, and AD genomes are depicted geographically. The AD and D genome cluster is represented in yellow, the A, B, E, and F genome clusters in green, and the C, G, and K genome clusters in red. **(G)** Prevalence of different CLCuD strains worldwide.

disease severity, curled leaves may even develop small outgrowths (also called enations) on their lower surface. Moreover, affected leaves become brittle and leaf size is substantially reduced (Mansoor et al., 1993; Harrison et al., 1997; Qazi et al., 2007). At the seedling stage, CLCuV affects the cotton yield byreducing the cotton boll's number and size, flowering time, and posing undesirable effects on seed and fiber quality, and final yield (Ali et al., 2013; Ji et al., 2015; Heigwer et al., 2016). It is challenging to fight CLCuV because of the prevalence of active viral strains and greater recombination rates of CLCuV complex. A case study by Qadir et al. (2019), revealed that the recombination potential of CLCuV is significantly higher. A greater recombination rate was observed in all genes (*IR, V1, V2, C1, C4,* and *C5*) associated with cotton leaf curl Kokhran Virus (CLCuKoV). Furthermore, CLCuKoV also donated in the *C2, C3* regions of cotton leaf curl Multan virus (CLCuMuV). Entirely, these observations clearly show the uniqueness of Indian CLCuMuV isolates representing the contribution of CLCuKoV in all the genes (Qadir et al., 2019). Higher recombination rate of CLCuV leads to multiple strain infection which makes viral control difficult. The presence of various alternative hosts and favorable environmental conditions has further complicated the disease control. Many efforts were made during the 1990s to produce CLCuV resistant varieties, yet the outbreak of another CLCuV

strain, namely, Cotton Leaf Curl Burewala (CLCuBuV) broke the resistance and the varieties fell prey to it (Sarwar et al., 2022).

Note: Cotton in Asia, experienced two significant epidemics within this time frame: the 'Multan epidemic,' spanning from 1988 to 1999, followed by a period of relative calm. However, in 2002, the 'Burewala epidemic' emerged in the cotton fields of the Indo-Pak subcontinent and persisted until 112 2013–2014.

The availability of various alternative hosts like hibiscus, okra, etc., and different farming practices could provide a reservoir of novel sources of new virulent CLCuV strains (Ali et al., 2022). Systemic-induced resistance using a gene-based approach remained effective until a new virulent strain of CLCuV came into existence named Cotton Leaf Curl Burewala Virus (CLCuBV). The damaging effects of the virus could only be overcome by proper management strategies. Many efforts have been used to protect the cotton crop from CLCuD. Disease control by biological means and development of CLCuV resistant cotton varieties are in progress. Use of pathogen and non-pathogenbased approaches and advanced molecular practices can be an effective way to fight viral causative agents (Hasan et al., 2019). We present an insightful overview encompassing the comprehension of the cotton genome, unraveling the genetic architecture of CLCuV, understanding cotton's defense mechanisms against CLCuD, and exploring the

applications of genome editing techniques for disease management, with a focus on cotton. Furthermore, we delineate the broader applications of other innovative technologies such as Next-Generation technologies, functional genomics, and speed breeding in cotton biological research and enhancement. Lastly, we discuss the ongoing challenges related to cotton genome editing and deliberate on the future opportunities involving various strategies and advancements in genome editing tools to foster the sustainability of cotton production.

2 Navigating CLCuV vectors: unveiling transmission mechanisms

Over the past three decades, insects from the genus *Bemisia* have become the major devastating pests of agriculture and horticulture (Ribeiro et al., 2003). The most significant and widespread among these are the whiteflies that are associated with *Geminiviruses* or more specifically to *Begomoviruses*. Their transmission may be directly associated with the silver leaf type of the whitefly. *Bemisia tabaci* is an indiscriminate feeder, facilitating rapid and proficient spread of CLCuV from affected plant hosts to neighbouring crops (Brown and Czosnek, 2002). Among all damaging pests of cotton, whitefly is responsible for about 50% yield loss and is a major restraint of boll formation (Mansoor et al., 2003). Intensive escalation in the whitefly population is because of the immense propagation of diseases by viruses through whiteflies.

Begomoviruses exist in either mono-partite or bi-partite forms. Bi-partite *Begomoviruses* contain both DNA A and DNA B, associated with New World pathogens. On the other hand, mono-partite *Begomoviruses* only contain a circular single-strand DNA-A and two satellite molecules viz., α and β satellite, classified as an essential component for disease severity in Old World pathogens (Hasan et al., 2019). An example of a mono-partite particle is CLCuV (Liu et al., 1998; Mansoor et al., 1999). PCR amplification of CLCuD-infected cotton indicated the presence of CLCuV (Fiallo-Olivé and Navas-Castillo, 2023). CLCuV possesses a circular single-stranded deoxyribonucleic acid (ssDNA) molecule, tightly encapsulated inside a geminate particle. Previous research has shown a wide range of hosts for mono-partite and bi-partite *Begomoviruses* (Yazdani-Khameneh et al., 2016). As the most destructive pathogens, CLCuVs are widely spread in Central and South-East Asia. The viruses carry distinct recombinations to break down transgenic resistance in cotton cultivars (Devendran et al., 2022).

2.1 Geminiviruses

Geminivirus is a group of pathogenic viruses transmitted by insects. The family *Geminiviridae* is responsible for causing several diseases in plants around the globe (Arif et al., 2022). *Geminiviruses* contribute to the largest group of plant viruses. Zerbini and colleagues (2017), categorized *Geminiviruses* into nine genera, utilizing criteria such as genome arrangement, transmission vectors, and host plants. This classification comprises *Turncurtovirus*, *Topocuvirus*, *Mastrevirus*, *Grablovirus*, *Eragovirus*, *Curtovirus*, *Capulavirus*, *Begomovirus*, and *Becurtovirus* along with a subset of species that remain unclassified (Zerbini et al., 2017). Recently, this family has been increased from nine recognized genera to fourteen genera (Materatski

et al., 2021). *Begomoviruses* is the largest group of *Geminiviruses* with around 520 accepted viral species ((Uniyal et al., 2019; Ouattara et al., 2022; Fiallo-Olivé and Navas-Castillo, 2023).

CLCuV *Geminiviruses* are responsible for the transmission of CLCuV through whiteflies. These viruses are taken up by whiteflies when they feed on infected plants and are then transmitted to healthy plants during subsequent feeding. This transmission mechanism makes *Geminiviruses* key contributors to the spread and persistence of CLCuV in cotton crops, leading to significant agricultural and economic impacts (Jain et al., 2023). Significant crop losses (such as high plains viral disease of wheat and corn cause by wheat mosaic virus and maize red stipe virus, sorghum mosaic virus and sugarcane mosaic virus cause by potyvirus, and rice stripe virus cause and transmitted by *Laodelphax striatellus*) have been reported due to various disease infestations leading to viral diseases spread worldwide (Lu et al., 2021; Tatineni and Hein, 2021; Xu, et al., 2021). Major reasons that resulted in disease epidemics include the recombination of many different *Geminiviruses* (i.e., the exchange of genetic materials among at least two *Geminiviruses* to become more virulent) that co-infect the same crop, expansion, and development of agriculture in newer areas, or transport of infectious plant materials to other regions. Additionally, plant co-infections can be spread by vectors that pick up the pathogens from plants infected with multiple diseases, either from different plants one after another, or by different vectors each carrying a different disease. Moreover, the migration of carrier insect vectors from one location to another also brings epidemic explosion (Gray and Banerjee, 1999).
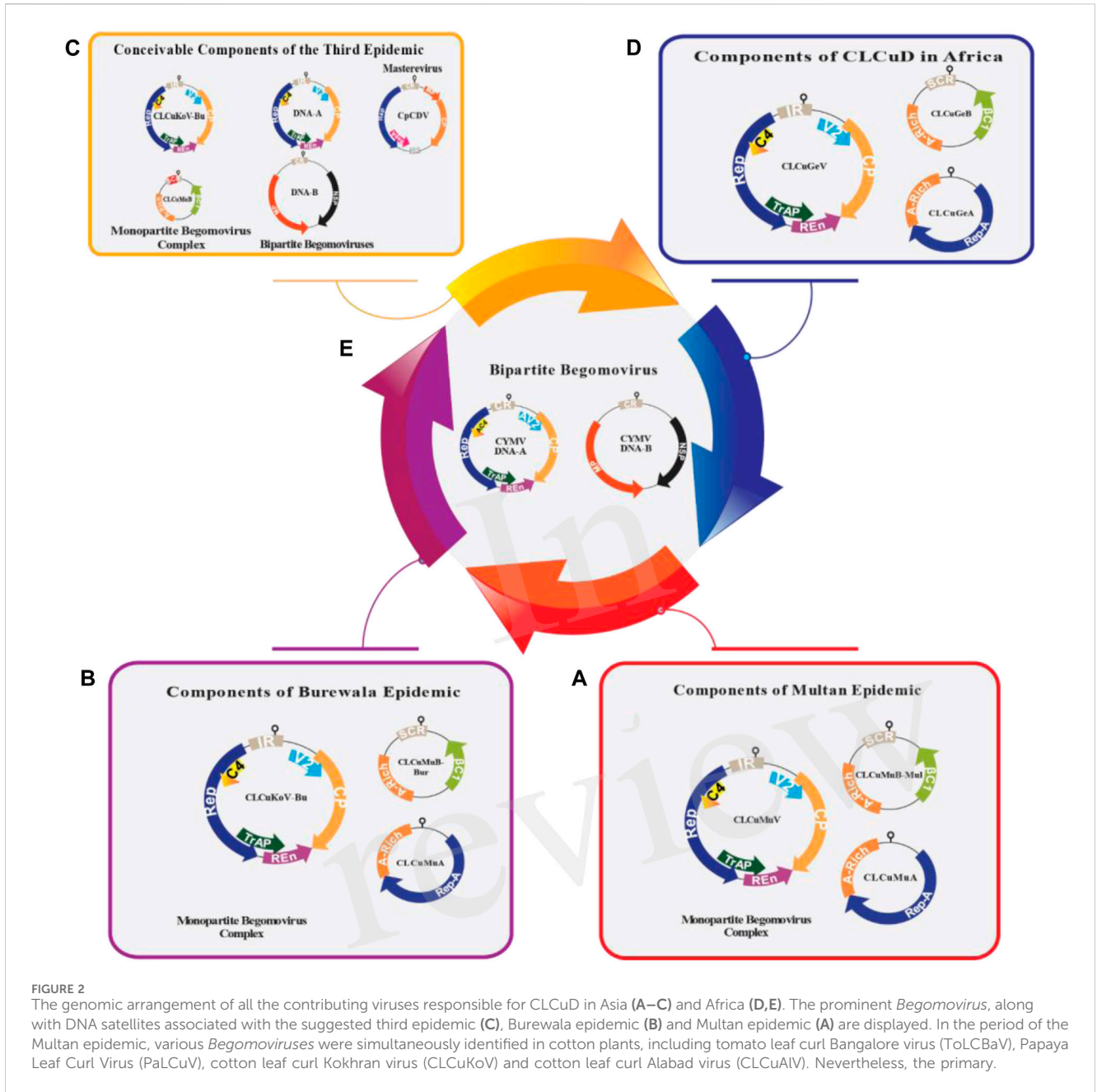
2.2 Begomoviruses

The genus *Begomovirus* is a serious threat not only to cotton but also to other crops such as tomato and cassava (Donnelly and Gilligan, 2023; Fortes et al., 2023). *Begomovirus* behaves the same way as *Geminiviruses* in transmitting the virus. The complex interaction between *Begomoviruses*, whiteflies, and cotton plants highlights the significance of understanding and managing these factors to mitigate the impact of CLCUV on cotton crops. Bi-partite begomoviruses are comprised of DNA-A and DNA-B, each approximately 2,600 nucleotides long (Fondong, 2013). The mono-partite particles contain single DNA-A molecule of about 2,800 nucleotides (Fauquet and Stanley, 2003). *Betasatellites* encompass circular, single-stranded DNA molecules comprising around 1,350 nucleotides, responsible for disease inducing symptoms (Zhou et al., 2003; Saunders et al., 2004). *Alphasatellites* are also circular single-stranded DNA made up of approximately 1,375 nucleotides and associated with a rolling-circle replication (RCR) initiator protein.

3 Decoding CLCuV genome: unveiling the genetic architecture

3.1 DNA A

CLCuV is an example of a mono-partite particle that contains single-stranded DNA (Briddon et al., 2000). DNA-A of mono-partite particles encodes for viral DNA replication, insect transmission and control gene expression (Noueiry et al., 1994). DNA-A has six open

FIGURE 2
The genomic arrangement of all the contributing viruses responsible for CLCuD in Asia **(A–C)** and Africa **(D,E)**. The prominent *Begomovirus*, along with DNA satellites associated with the suggested third epidemic **(C)**, Burewala epidemic **(B)** and Multan epidemic **(A)** are displayed. In the period of the Multan epidemic, various *Begomoviruses* were simultaneously identified in cotton plants, including tomato leaf curl Bangalore virus (ToLCBaV), Papaya Leaf Curl Virus (PaLCuV), cotton leaf curl Kokhran virus (CLCuKoV) and cotton leaf curl Alabad virus (CLCuAlV). Nevertheless, the primary.

reading frames (ORFs) that encode various proteins (Yaqoob et al., 2020). In the virion sense, there are two ORFs, namely, AV2 (encodes the movement of a protein; also called precoat) and AV1 (encodes the coat protein), while in the complementary sense, four ORFs [AC4 (determines the expression of symptoms), AC3 (encodes replication enhancer protein), AC2 (encodes a transcription activator protein), and AC1 (encodes the inititation protein called Rep)] are found (Marwal et al., 2014). The genomes of mono-partite viruses and the DNA-A component of bi-partite *Begomoviruses* contain the instructions for the coat protein (CP) and V2 (also known as AV2) in the orientation corresponding to the virion sense (Figure 2). The coat protein is involved in virus proliferation, vector transmission, disease virulence and encapsidation (Singh et al., 2021). AV2 is associated with cell-to-cell movement and pathogenicity (Arif et al., 2022). The proteins linked to

replication (AC1/C1 or Rep protein), transcriptional activation (AC2/ C2 or TrAP protein), replication enhancement (AC3/C3 or REn protein), and the AC4/C4 protein are encoded in the orientation corresponding to the complementary sense (Fondong, 2013; Zhou, 2013). AC1 serves viral replication and gene expression whereas AC2 is responsible for transcription activators of rightward ORFs and suppressors of post-transcriptional gene silencing (PTGS) (Dilip, 2016; Sharma and Prasad, 2020). AC3 is involved in viral replication and symptom development and AC4 in gene silencing (Faiz and Abhinav, 2021; Roumagnac et al., 2021). DNA-B consists of two ORFs, BC1 in the complementary sense and BV1 in the virion sense. BV1 plays a role in nuclear trafficking and BC1 encodes movement protein MP, involved in inter- and intracellular movement (Diamos et al., 2019; Yadav et al., 2021).

3.2 DNA B

The genome component DNA-B (absent in CLCuV but present in bi-partite *Begomoviruses*) encode the movement protein (MP) in the complementary-sense orientation and nuclear shuttle protein (NSP) in the virion sense (Sahu et al., 2014). DNA-B has two open reading frames and encodes proteins associated with inter and intracellular movements (Mubarik et al., 2021). An intergenic intronic region exists between ORFs of the virion and complementary-sense regions that contain a hair-pin loop-like structure and cis-acting regulatory elements essential for gene expression. The loop-like structurecontains a nona-nucleotide conserved sequence 'TAATATTAC' and small repetitive sequences 'iterons' that provide sequence-specific binding sites for Rep (C_1/AC_1) protein. Both iterons and hairpin loop form the origin of replication (*ori*) for initiation of viral DNA replication (Zhou, 2013).

3.3 DNA α-satellite and β-satellite

Alpha and *betasatellites* are assumed to function in encapsulation, insect mediated transmission and cellular movement inside the infective host plants Gnanasekaran and Chakraborty, 2018). Additionally, betasatellites impact symptom development and boost the pathogenicity of helper viruses by elevating viral DNA levels in host plants and inhibiting the plant's antiviral defense mechanisms (Gnanasekaran et al., 2019). Satellite molecules share a conserved loop like structure with DNA-A. Genome size of these molecules are about 1.4 kb (Azeem et al., 2022). Both satellites contain conserved structures which share broad characteristics. Furthermore, the molecules comprise an "*Ori*" region separated by a non-nucleotide sequence "NANTATTAC" at the apex of the hairpin structure (Rosario et al., 2012). Though, *betasatellites* contain a nonanucleotide motif "TAATATTAC", observed in a large number of *Begomoviruses* whereas *alpha satellites* contain "TAGTATTAC" motif, seen in various members of *Nanoviridae* family. *Alphasatellite* is capable of replicating itself autonomously. Both *alpha* and *betasatellites* are rich in adenine residues which function in proficient encapsulation and aid in systemic movement of the virus. Both molecules encode a single protein. The α-satellite encodes a Replication protein (RP) associated with rolling circle replication inside the host cells. On the contrary, helper *Begomovirus* is required for *betasatellite* replication. The *β-satellite* also encodes βC1 protein which is multi-functional and significant for inducing pathogenicity (Zhou, 2013).

Begomovirus responsible was for Cotton Leaf Curl Multan Virus (CLCuMuV). In the case of the Burewala epidemic, a recombinant cotton leaf curl Multan betasatellite-Burewala strain (CLCuMuB^Bur) and a truncated transcriptional activator protein (TrAP) were observed, incorporating approximately 98 nucleotides from the SCR region acquired from Tomato leaf curl betasatellite (ToLCB) (B). The third epidemic in Asia presents several possible scenarios: i) it could involve Cotton leaf curl Kokhran virus-Burewala (CLCuKoV-Bu), which retains an intact TrAP, along with Cotton leaf curl Multan betasatellite (CLCuMuB), containing a ~24 nt stretch from tomato leaf curl betasatellite (ToLCB), ii)

Alternatively, it might include tomato leaf curl New Delhi virus (ToLCNDV) (comprising both DNA-A and DNA-B) and tomato leaf curl virus (ToLCV) (with DNA-A), possibly in conjunction with ToLCNDV (DNA-B), iii) Another potential contributor could be Mastrevirus chickpea chlorotic dwarf virus (CpCDV) (C). Concurrently, the CLCuD situation in Africa comprises: (i) a mono-partite *Begomovirus* complex (D), a recently identified bipartite *Begomovirus*, cymbidium mosaic virus (CYMV), with DNA-A and DNA-B (E). The DNA-A of mono-partite *Begomoviruses* encodes two open reading frames (ORFs) in the virion-sense orientation, specifically the V2 protein and the coat protein. In the complementary-sense orientation, there are four ORFs: the replication-associated protein (Rep), TrAP, replication enhancer protein (REn), and C4 protein (with functions yet to be fully understood). Bi-partite *Begomovirus* DNA-A resembles the mono-partite genome but includes a second genomic component, DNA-B, which encodes two ORFs: the movement protein (MP) and the nuclear shuttle protein (NSP), in both complementary-sense orientations and virion sense. The alphasatellite encodes a single ORF (Rep-A) for autonomous replication, while the betasatellite encodes its single ORF, βC1, in the complementary-sense orientation. βC1 plays a role in assisting the helper *Begomovirus* in various functions.

4 Defense mechanisms of cotton against CLCuV

Besides viruses, several other organisms including insects such as cotton aphids, *Hemipteran* sucking insects, cotton bollworms, and Lepidopteran chewing caterpillars are commonly observed in cotton fields. Cotton bollworm predation triggers the activation of genes related to the gibberellic acid (GA) (*GhMPK11*, *SLR1*), ethylene (*GhWRKY70D13*, *GhERF91*), and jasmonic acid (JA) (*GhWRKY70D13*, *JAZ3*) pathways, while simultaneously suppressing genes associated with the salicylic acid pathway within cotton plants (Kumar et al., 2016; Wang et al., 2016; Xia et al., 2018; Xiong et al., 2020; Kundu et al., 2023; Nurimanguli et al., 2023). JA acts as a regulator in facilitating insect defense, working in conjunction with its corresponding target (SPL9) and miR156 (Arora and Chaudhary, 2021). The response to JA is swift in young plants but progressively slows as plants age, displaying an age-related pattern inversely linked to SPL9 group protein levels. Moreover, not all but some of the herbivorous insects have also developed intricate strategies, such as the discharge of effector molecules into the host and the synthesis of diverse detoxification enzymes, to overcome the resistance mechanisms of their host plants. The defense mechanisms of cotton against CLCuD are multifaceted and include both pathogen-derived and non-pathogen-derived strategies. The most significant defense mechanism is the presence of resistance (R) genes in plants (Table 1). These genes provide specific resistance against particular viruses by inducing cell death around infected plant cells, thereby preventing the movement of the virus. Cotton has two associated R genes, *R1CLCuDhir* and *R2CLCuDhi*, which confer resistance against CLCuD (Khan et al., 2015). Furthermore, the exclusive protein βC1, encoded by the satellite, has been demonstrated to hinder the ubiquitination process of the host

TABLE 1 List of genes associated with CLCUD and other biotic and abiotic factors, directly and indirectly, involved in cotton infections.

Sr. No	Gene ID/ Accession no.	Gene symbol	Gene description	Chromosomal position/locus tag	Gene function	References
1	30144432	AV1	Coat protein	BMU10_gp3	Virus proliferation, vector transmission, disease virulence and encapsidation	Singh et al. (2021)
2	30144433	AV2	Pre-coat protein	BMU10_gp2	Involved in Pathogenicity and cell to cell movement	Sodha et al. (2022)
3	30144436	AC1	Replication associated protein	BMU10_gp6	Viral replication and gene expression	Sharma and Prasad. (2020)
4	30144435	AC2	Transcription activator protein	BMU10_gp5	Responsible for transcription activators of rightward ORFs. Suppressor of post-transcriptional gene silencing (PTGS)	Dilip (2016)
5	30144434	AC3	Replication enhancer protein	BMU10_gp4	Viral replication and symptom development	Roumagnac et al. (2021)
6	30144437	AC4	Regulatory protein	BMU10_gp7	Gene silencing	Faiz and Abhinav. (2021)
7	80557347	βC1	BetaC1 protein	QKT41_gp1	Symptom determinant, interact with other helper viruses, affect number of plant species including cotton	Iqbal et al. (2012)
8	988208	ac5	AC5 protein	ToLCNDVsAgp4	Viral replication	Amudha et al. (2010)
9	107958459	WRKY4 0	Probable WRKY transcription factor 40	Chromosome A05/ ERO13_A05G119 500v2	Regulate cotton defense for controlling whitefly infestation	Li et al. (2016)
10	107924381	MPK	Mitogen-activated protein kinase kinase 3	Chromosome A11	Silencing of MPK leads to enhanced whitefly susceptibility	Li et al. (2016)
11	988213	BV1	Nuclear shuttle protein	ToLCNDVsBgp1	Role in nuclear trafficking	Diamos et al. (2019)
12	2767401	BC1	BC1 protein	TYLCKV_sBgp2	Encodes movement protein MP, involved in inter- and intracellular movement	Yadav et al. (2021)
13	—	R1CLCuDhir	Resistance gene	—	Provides resistance to cotton against CLCuD	Hagan et al. (2019)
14	—	R2CLCuDhir	Resistance gene	—	Provides resistance to cotton against CLCuD	Hagan et al. (2019)
15	—	SCLCuDhir	suppressor gene	—	Suppressor of resistance to CLCuD	Hagan et al. (2019)
16	107934102	LOC107 934102	Probable zinc metallopeptida se EGY3, chloroplastic	ERO13_A01G009 420v2/ Chromosome A01	Chloroplast development. Copes with stress	Naqvi et al. (2017)
17	107928973	LOC107 928973	Probable zinc metallopeptida se EGY3, chloroplastic	ERO13_D01G009 000v2/ Chromosome D01	Development and stress response	(Naqvi et al. 2017)
18	105786625	LOC105 786625	Probable zinc metallopeptida se EGY3, chloroplastic	B456_002G01100 0/ chromosome 1	Probable membrane-associated metalloprotease that may be involved in chloroplast development.	Marino and Funk. (2012)
19	107909622	LOC107 909622	Chaperone protein dnaJ 16	ERO13_D02G166 200v2/ chromosome D02	Pathogen defense, vital in antiviral defense	Verchot (2012), Park and Seo (2015)
20	105802570	LOC105 802570	Transcription factor UNE10	B456_007G08830 0/ chromosome 11	Antiviral defense	Wang et al. (2015)
21	107911889	LOC107 911889	Transcription factor UNE10- like	ERO13_A11G078 200v2/ chromosome A11	Antiviral defense	Wang et al. (2015)
22	107904753	LOC107 904753	Protein REVEILLE 7	ERO13_D12G273 200v2/ chromosome D12	Plant growth contributes to response against pathogens	Wang et al. (2015), Zhong et al. (2016)

TABLE 1 (*Continued*) List of genes associated with CLCUD and other biotic and abiotic factors, directly and indirectly, involved in cotton infections.

Sr. No	Gene ID/ Accession no.	Gene symbol	Gene description	Chromosomal position/locus tag	Gene function	References
23	107902644	LOC107 902644	Transcription factor RADIALIS	ERO13_D05G196 766v2/ Chromosome D05	Plant defense response	Katiyar et al. (2012)
24	107951700	LOC107 951700	Putative disease resistance protein RGA3	ERO13_D11G336 100v2/ chromosome D11	R-gene-mediated resistance and disease response	Chen et al. (2015)
25	107916180	LOC107 916180	Bidirectional sugar transporter SWEET17	ERO13_D10G171 600v2/ chromosome D10	Confer's tolerance against abiotic stress, pathogenesis-related protein	Chandran (2015)
26	105792917	LOC105 792917	Protein ECERIFERU M 16	B456_004G28270 0/ chromosome 8	Role in biotic and abiotic stresses	Bourdenx et al. (2011)
27	107928876	LOC107 928876	Aquaporin PIP2-7-like	ERO13_D01G001 900v2/ chromosome D01	Critical in plant immunity, biotic and abiotic stresses	Matsumoto et al. (2009), Meng et al. (2016), Tian et al. (2016)
28	107927933	LOC107 927933	Potassium channel AKT2/3	ERO13_A03G193 300v2/ chromosome A03	Plant development, stress responses, antiviral defense	(Wang and Wu 2010, Zhou et al. 2014)
29	107937358	LOC107 937358	Boron transporter 4	ERO13_A06G187 400v2/ chromosome A06	Involved in R-gene-mediated viral defense	Lim et al. (2014)
30	107949718	LOC107 949718	Vignain	ERO13_D03G102 700v2/ chromosome D03	Plant immunity, pathogenesis, and plant defense	Figueiredo et al. (2014)
31	SAMS	GhSAMS	S-adenosylmethionine synthase	Chromosome AD1	Drought tolerance	Sun et al. (2022)
32	TR104989_c0_g1	LSM14-like protein B	Zf-C2H2 transcription factors	—	CLCuV resistance	Farooq et al. (2022)
33	AT5G11270.1	Gh_D08G2462	A homeodomain transcription factor	Chromosome _D8	Potential candidate gene for drought tolerance	Guo et al. (2022)
34	AT2G29560.1	Gh_A13G1963	A putative Phosphoenolpyruvate enolase	Chromosome A13	Potential candidate gene for drought tolerance	(Guo et al. 2022)
35	—	GhCFIm25	Pre-mRNA cleavage factor Im 25 kDa subunit	—	A potential candidate gene for *Aphis gossypii* tolerance	Yang et al. (2023)
36	—	GhGUX5	Involved in several physiological processes		A candidate gene associated with resistance to *Verticillium wilt*	Zhang et al. (2023)

plant (Gnanasekaran et al., 2019). It achieves this by interacting with the SKP1/CUL1/F- box complex, disrupting plant hormone signaling (Zhang Y. et al., 2019). This disruption ultimately leads to the buildup of the virus and the clustering of symptoms in the host plants. Moreover, βC1 also disrupts the plant's autophagy pathway by obstructing the normal function of autophagy-related protein 3 and cytosolic glyceraldehyde-3-phosphate dehydrogenase within the host cells (Ismayil et al., 2020).

4.1 Pathogen derived resistance strategies

RNA interference (RNAi) is another potent defense mechanism in plants. This mechanism involves the degradation of viral mRNA, effectively silencing the genes responsible for viral replication. For example, after the initial cloning of the first effector glucose oxidase from *Helicoverpa zea* (Musser et al., 2002), many other effectors like HARP1 (Chen et al., 2019), and cytochrome P450 monooxygenases (TAO et al., 2012), were discovered. In addition to phytochemicals, phytohormones (SA and JA) also have the capability to activate insect P450 genes when insects invade (Li et al., 2002). Both

alterations in the structure of P450 proteins and increases in the expression of P450 genes have been linked to the ability of insects to tolerate gossypol and insecticides such as fenvalerate and deltamethrin (Joußen et al., 2012). Upon introducing a double-stranded RNA (dsRNA) construct that targeted the gossypol-inducible P450 gene *CYP6AE14* (referred to as *dsCYP6AE14*) into *Arabidopsis* plants, there was a significant reduction in the expression of *CYP6AE14*, resulting in a simultaneous decrease in the weight gain of larvae (Mao et al., 2007). Similarly, employing dsRNA constructs to target genes like *dsFAR* (fatty acyl-CoA reductase), *dsJHBP* (JH-binding protein), *dsJHAMT* (juvenile hormone acid methyltransferase) and *dsNDUFV2* (encoding a mitochondrial complex I subunit) in the sap-sucking pest *Adelphocoris suturalis* of the *Miridae* family and *Helicoverpa armigera* yielded positive outcomes when transformed into cotton plants for pest control (Huang et al., 2021). Additionally, plants produce natural defense chemicals like nitric oxide (NO), salicylic acid ($C_7H_6O_3$), and reactive oxygen species (ROS) to combat pathogens (Ibrahim et al., 2023). Short interfering RNAs (siRNAs) also play a role in plant defense against double-stranded RNA, although some viruses have developed proteins to counteract

this response. In the event of injury, plants restrict viral movement through plasmodesmata. These findings now urge cotton breeders to use RNAi extensively in controlling CLCuD.

The primary approach to controlling CLCuD involves managing the whitefly population through insecticides. However, this method is not entirely reliable, as delayed application could lead to transmission of the infectious strain. Moreover, insecticide use has adverse environmental effects and can be toxic to various organisms. Therefore, conventional methods of protecting cotton crops from insects are not preferable. Biotechnological and molecular breeding approaches present the most promising option for inducing resistance in cotton against viruses (Saakre et al., 2023). Pathogen-derived resistance can be further categorized into two types: resistance with protein expression and resistance without protein expression. Proteins play a crucial role in plant defense against biotic stresses, with research indicating that proteins act as the primary defense against CLCuD in cotton (Chauhan et al., 2023). Viral replication, regulated by Rep proteins, can be disrupted to inhibit the synthesis of viral components. However, while efforts have been made to develop transgenic cotton varieties against CLCuD, these varieties are still undergoing testing. The second approach, without protein expression, involves inducing resistance through RNAi. RNAi is a sequence-specific defense mechanism that leads to the degradation of viral mRNA transcripts. In plants, RNAi can be triggered by introducing a hairpin loop structure with a sequence homologous to the virus, thus stimulating the silencing signal (Karthik et al., 2023). This approach could also aid in developing insect-resistant genes against whitefly-transmitted viruses. RNAi-based resistance has been successful against Bean Golden Mosaic Virus in beans in Brazil (Souza et al., 2018). Another approach involves coat protein (CP) expression from other viruses, such as the Tobacco Mosaic Virus (TMV), to induce resistance. For example; Rasool et al. (2021), in their study, used synthetic genes to explore pathogen-derived resistance (PDR) for managing CLCuKoV-Bu and its associated betasatellite linked to CLCuD. Synthetic coat protein (CP_{syn}) and replication-associated protein (Rep_{syn}) genes from CLCuKoV-Bu were designed and inserted into *Nicotiana benthamiana* plants via *Agrobacterium*-mediated transformation. Transgenic plants carrying CP_{syn} and Rep_{syn} genes showed milder symptoms and lower virus concentrations when exposed to CLCuKoV-Bu or CLCuKoV-Bu-CLCuMuB inoculums. These findings suggest that CP_{syn} and Rep_{syn} genes have the potential for conferring resistance against CLCuKoV-Bu and related betasatellites and the study indicates the development of a PDR mechanism for controlling CLCuD. However, the exact mechanisms behind these strategies remain only partially understood (Rasool et al., 2021).

4.2 Non-pathogen derived resistance strategies

Non-pathogen-derived resistance strategies have been explored in cotton against CLCuD. The Need for non-pathogen-derived resistance strategies emerged when the applications of biotechnological approaches were found successful in other crops like wheat (Jastrzębska et al., 2020), rice (Montesinos et al., 2017), and maize (Gould and Northcote, 1986), but with a low success ratio

in cotton (Khan et al., 2015). These strategies include enhancing natural defenses like trichomes (which are hair-like structures, developed at the aerial plant surfaces involved in plant defense) and increasing the concentration of waxes and inorganic salts, which can protect cotton plants against whiteflies, CLCuD, and other pests. For example; Suthar et al., 2021, studied the tolerance response of interspecific derivatives (resulting from a cross between *Gossypium hirsutum* L. × *G. armourianum Kearney*) against whitefly. Their results revealed that the concentration of trichome at early developing stages was high and developed strong tolerance against whitefly. However, with age progression, a significant decline was observed in trichome concentration in mature leaves (Suthar et al., 2021). Majid et al. (2020) conducted an experiment to study the effect of epicuticular wax against CLCuD in different varieties of cotton crop. Their results revealed a strong significant negative correlation between leaf epicuticular wax and CLCuD infection, indicating that the higher the concentration of wax the lower will be the CLCuD infection (Majid et al., 2020).

5 Revolutionizing genetics: genome editing and beyond

5.1 CRISPR-Cas mediated genome editing for CLCuD

Genome editing (GE) technologies have brought about a transformative shift in the realm of plant research and hold immense promise for enhancing crop attributes. The Clustered Regularly Interspaced Short Palindromic Repeat (CRISPR)-CRISPR-associated protein (CRISPR-Cas) system, renowned for its adaptability, simplicity, and cost-effectiveness in achieving precise modifications to DNA sequences such as gene knockout or knockdown, allele/gene *in vivo* replacement, and single base substitution, has emerged as the dominant force in GE over recent years (Khan et al., 2022; Atiq et al., 2023). Upon creating double-strand DNA breaks (DSBs) by CRISPR-Cas, subsequent repair occurs via either the error-prone non-homologous end joining (NHEJ) pathway or the more precise homology-directed repair (HDR), or a combination of both pathways (Liu et al., 2022). NHEJ primarily facilitates DSB repair and typically induces random indels (also called insertions and deletions) at the rejoining point of the chromosomes. This propensity for indels has led to the predominant use of NHEJ in current plant GE endeavors for inducing gene knockouts and mutations. In the context of precise genome editing, HDR can be invoked through the presence of a DNA or RNA donor repair template (DRT) containing homologous sequences encircling the DSB. This permits more precise gene insertions or replacements (Tang et al., 2023). Alternatively, two other strategies also exist that still offer precise GE even without DRTs and DSBs. The first is called base editing (BE), a powerful technique facilitating gene replacement through HDR, enabling the precise substitution of a single base with a designated single nucleotide polymorphism (SNP) (Sony et al., 2023). This technique facilitates the precise alteration of plant genomes in a programmed manner. The second strategy, known as prime editing (PE), greatly extends the scope of precise genome editing by allowing all types of minor indels and base substitutions (Karlson et al., 2021).

Near in the future, the widespread adoption of various CRISPR toolsets will significantly hasten progress in both crop genetic improvement and biological research, particularly for those plant species with available reference genome sequences.

Until now, most of the research on genome editing in crop plants, such as cotton, has relied on NHEJ to create targeted loss-of-function mutations at particular gene locations. The applications of CRISPR-Cas9 have been widely employed to enhance both crop yield and quality such as resilience against abiotic and biotic factors including CLCuD. For instance; Binyameen et al. (2021), used multiplexed CRISPR-based GE system to develop resistance against CLCuD in cotton. They obtained the targeted CLCuD sequences from the National Center for Biotechnology Information (NCBI) and designed three guided RNA (gRNA) molecules to target three different sites in the genome. They successfully obtained 60%–70% resistance against CLCuD at the seedling stage compared with control plants (Binyameen et al., 2021). The βC1 and Rep genes were responsible for replication of CLCuV within the host genome. Using a newly developed multiplexed GE approach was more efficient and effective to knock down or knock-out these genes to develop resistance against CLCuD (Khan et al., 2020). Improving cotton quality and meeting the needs of the cotton industry is an important breeding objective for cotton breeders. The aim is to develop a cotton variety that not only showcases improved quality and resistance against CLCuD but also achieves increased production without sacrificing yield.

5.2 Base editing and its applications against CLCuD

BE offers a mechanism to achieve exact substitutions of one base pair for another at a designated genomic position, without necessitating a DRT and without instigating DSBs (Basu et al., 2023). This technique operates independently of the HDR pathway. Recently existing base editors, encompassing adenine base editors (ABEs) and cytosine base editors (CBEs), have been engineered by fusing the inactive Cas9 (dCas9) or Cas9 nickase (D10A) (nCas9) with adenine deaminase or cytidine deaminase, respectively (Komor et al., 2016; Gaudelli et al., 2017). In this manner, single-point mutations such as the conversion of A*T to G*C or C*G to T*A in crop plants can be generated using ABEs or CBEs (Bharat et al., 2020; Ali et al., 2023; Khan et al., 2023). So far, base editing has been effectively used to target several important traits in various crops like maize, cotton, *Arabidopsis thaliana*, rice and others. For example,; Qin et al. (2020), developed the novel *G. hirsutum*-Base Editor 3 (GhBE3) to induce single base mutation in the allotetraploid genome of G. *hirsutum*. To test the efficiency of this system, two genes (*GhPEBP* and *GhCLA*) were targeted at three different sites in the genome. The efficiency of this base editing system was found highly effective and ranged from 26.67% to 57.78% (Qin et al., 2020). In another study, Wang et al. (2022), created diverse ABE vectors utilizing modified adenosine deaminase (TadA) proteins linked with two Cas9 variants named nCas9 and dCas9. These vectors demonstrated proficient A to G editing, achieving efficiencies up to 64% at targeted locations in the complex allotetraploid cotton genome. The results revealed

GhABE7.10n as the most effective, particularly altering the A5 position (considering PAM as positions 21–23). RNA off-target and DNA assessments using *GhABE7.10d* and *GhABE7.10n* edited cotton plants revealed minimal RNA off-target mutations and no DNA off-target changes. Additionally, a novel base editor, *GhABE7.10d*Cpf1, was introduced, recognizing a PAM rich in thymine (T), leading to targeted A-to- G substitutions and a single amino acid modification in the cotton phosphatidyl ethanolamine-binding protein (*GhPEBP*). This alteration resulted in a condensed cotton plant structure that is ideal for mechanized harvesting in contemporary cotton farming (Wang et al., 2022). Recently, Mubarik et al. (2021) demonstrated a successful CRISPR-Cas9 system application to combat CLCuD by targeting overlapping genes of prevalent CLCuVs with three guide RNAs (gRNAs). Utilizing a multiplex CRISPR-Cas9 construct targeting six CLCuV genes proved more effective in inhibiting virus proliferation compared to individual targeting. Targeting CLCuV at multiple sites simultaneously resulted in superior interference and inefficient recovery of altered virus molecules. When tested in cotton plants, the multiplex construct significantly reduced virus accumulation in leaves. These findings highlight the potential of CRISPR-Cas9 for engineering virus resistance in crops, especially against mixed virus infections, through multiplex genome editing.

5.3 Prime editing and its applications against CLCuD

Prime editing, also known as PE, is a revolutionary genome-editing approach that allows precise modifications to the genetic code. Unlike traditional methods that rely on donor DNA templates or DSBs, prime editing utilizes a fusion protein complex to introduce specific edits at desired gene loci in plant cells (Hillary and Ceasar, 2022). Prime editing involves the conjugation of a reverse transcriptase Moloney murine leukemia virus reverse transcriptase (M-MLV-RT) with a catalytically impaired Cas9 (nCas9) to form a protein complex. A guide RNA for prime editing (pegRNA) is employed to designate the desired target location and encode the intended genetic information. Within the prime editing system, the protein complex attaches to the targeted DNA sequence and creates a nick in the non-target strand (Li et al., 2021). Prime editing has gone through several iterations, with each version aiming to improve product purity and editing efficiency. The initial version, PE1, was developed by fusing the C-terminus of nCas9 with M-MLV-RT. This fusion protein showed promise but lacked optimal efficacy. To improve PE efficacy, PE2 was developed by replacing the original M-MLV-RT with an engineered version containing six mutations: *L603W*, *T330P*, *W313F*, *T306K*, *D200N*, and *H9Y*. Building upon the success of PE2, PE3 was created by introducing another nicking sgRNA at various distances from the nicks induced by the pegRNA. This additional cleavage in the non-edited strand further increased the efficiency of the editing process (Anzalone et al., 2019).

PE shows significant potential for introducing precise genome modifications in cotton crop to enhance agronomic traits and develop resistance against CLCuD. In cotton, CLCuD tolerant plants were developed using prime editing (Cao et al., 2020). The multiple gRNA was cloned to target C1 and Rep genes

simultaneously. The resulted cottons plants did not display any CLCuD symptoms after 30 days of transformation and the results were further confirmed through PCR by extracting the RNA at different growth stages (Shafique et al., 2022). Shortly, prime editing represents a significant advancement in genome-editing technology, offering improved precision and efficiency compared to traditional methods (Lin et al., 2020). The development of various prime editing systems, from PE1 to PE3b, has enabled researchers to achieve precise modifications in both human cells and crop plants. As the technology continues to evolve, prime editing may become a powerful tool in the hands of cotton breeders for creating sustainable and resilient cotton varieties to meet the challenges of the future (Li et al., 2020).

6 Beyond genome editing

6.1 Next-generation sequencing technologies

Cotton has been extensively studied in the field of genomics to understand its genetic makeup and improve its agronomic traits. In the past, classical methods such as microarray analysis were used to study the genetics of plants, but they were costly, low-yielding, slow and labor-intensive (Hu et al., 2021). However, in the last two decades, high-throughput sequencing methods, collectively called Next-Generation Sequencing (NGS), have revolutionized plant genetics research, including cotton. NGS methods offer higher throughput, faster turnaround times, and reduced costs compared to traditional sequencing methods (Salk et al., 2018).

6.2 First generation technology

The first generation of sequencing technology, which emerged in the early 2000s, included chemical analysis and enzymatic sequencing methods. These methods relied on chain termination using $2',3'$- dideoxynucleotides (ddNTPs) (Medžiūnė et al., 2022). First-generation sequencers, such as the ABI Prism, were developed to automate the sequencing process and could analyze 4 to 96 samples per single run (Kchouk et al., 2017). Although this technology produced high-quality sequencing, it had limitations in terms of laboriousness, slow methodology, and limited data output. It was primarily used for small genomes and large sequencing projects like the Human Genome Project (Collins et al., 2003). Mansoor et al. (1999) identified and characterized a new circular single-stranded (ss) DNA found in naturally infected cotton plants. Purified geminate particles from infected plants not only contain a component resembling DNA A but also a smaller, unrelated ssDNA termed DNA 1. DNA 1 was cloned from the double-stranded replicative form of viral DNA isolated from infected cotton plants. Blot hybridization using specific probes for CLCuV DNA or DNA 1 demonstrated the co-infection of naturally infected cotton plants in various geographical locations by both DNAs. DNA 1 was detected in viruliferous Bemisia tabaci and in laboratory-infected tobacco plants using B. *tabaci*, indicating its transmission by whiteflies. Sequence analysis revealed that DNA 1 is approximately half the size of CLCuV DNA, with no homology,

suggesting it is not a defective *Geminivirus* component. DNA 1 shares some homology with a genomic component of Nanoviridae, a family of DNA viruses typically transmitted by aphids or planthoppers. DNA 1 encodes a homologue of the nanovirus replication-associated protein (Rep) and can autonomously replicate in tobacco. The findings imply that a nanovirus-like DNA has acquired whitefly transmissibility through its association with a *Geminivirus*. This association may contribute to a mutually dependent relationship between two distinct DNA virus families with a similar replication strategy, potentially influencing the development of cotton leaf curl disease. Similarly, Sharma et al. (2005) used polymerase chain reaction (PCR) technique to detect viral DNA in samples stored for up to 3 days after collection, making it widely applicable for field diagnosis. In their study, the complete nucleotide sequence of the coat protein (CP) gene component of the Indian isolate of cotton leaf curl *Geminivirus* (CLCuV-HS2) was determined using CP-specific primers in PCR amplification. The infected cotton plants were obtained from fields in Haryana, India. Amino acid sequence comparison of the putative CP with other mono and bipartite *Geminiviruses* revealed a maximum identity of 97.3% with Pakistan cotton leaf curl virus (CLCuV-62). Additionally, a nuclear localization signal close to the N-terminal of the CP gene was identified. Moreover, Briddon et al. (2000) obtained full-length clones of Cotton Leaf Curl Virus (CLCuV), equivalent to the DNA A component of bipartite *Begomoviruses*. These CLCuV clones were found to be systemically infectious in both Nicotiana benthamiana and cotton plants. Interestingly, the infected plants did not display the typical symptoms associated with cotton leaf curl disease; instead, they exhibited mild leaf curling, yellowing, and some stunting. Despite efforts to identify a second genomic component, none was found. These results suggest that CLCuV alone may not be the sole cause of cotton leaf curl disease. However, the transmission of the disease by Bemisia tabaci, a whitefly, implies that the *Begomovirus* CLCuV might still play a role in disease transmission, even if it is not the primary causative agent.

6.3 Second generation technology

The advent of second-generation sequencing technology marked a significant milestone in the field of genomics. Unlike first-generation methods, second-generation or short-read sequencing techniques utilized PCR to amplify DNA fragments before sequencing. These techniques employed hybridization or synthesis-based methods to sequence millions of DNA fragments in parallel, enabling high-throughput sequencing (Adewale, 2020). Platforms like 454 Life Sciences Roche, SOLiD (Thermo-Fisher Scientific), Solexa (later acquired by Illumina), and Ion Torrent revolutionized the sequencing landscape. Illumina, the extensively employed technology for genotyping investigations in cotton, provides diverse choices for transcriptome and genome sequencing, delivering outputs spanning from gigabases to terabases (Meera Krishna et al., 2019). Patel (2016), detailed the *de novo* assembly process and characterization of transcriptome and genome data for an Asia I population. The research aimed to establish a genetic framework for insecticide discovery. Using the Roche 454 sequencing platform, the authors assembled a

comprehensive transcriptome of adult females from the Asia I population, generating 0.864 million reads and producing 29,418 contigs through CLC Genomics. Of these, 8,563 contigs were assigned putative functions. A draft genome of 828 Mbp for the Asia I population was constructed via *de novo* assembly of 206 million 250 bp paired-end reads from the Illumina MiSeq 2,500 platform using Platanus. This genome comprised 41,981 protein-coding genes (PCGs) and included 990 non-coding RNAs and repetitive elements (45.66%). The study examined 741 full-length genes, revealing larger intron sizes and repetitive elements contributing to the larger Asia I genome size, which explained challenges in genome assembly. Additionally, the research obtained one mitogenome and three endosymbiont genomes (Portiera, Wolbachia, and Arsenophonus) from the same sequence library, alongside the B. *tabaci* Asia I genome. The development of a genetic framework involved using *Drosophila* essential genes as a reference to identify orthologs in other insects, validated with ChEMBL targets. The presented 'omics' data serves as a comprehensive sequence resource for Asia I populations, illustrating the workflow for obtaining genetic information on hosts and their endosymbionts.

6.4 Third generation technology

Third-generation sequencing techniques, often referred to as single molecule sequencing (SMS), represent the latest advancement in sequencing technology. These methods do not require PCR amplification and can sequence single DNA molecules directly. The ability to sequence DNA molecules without amplification provide advantages such as reduced bias, improved accuracy, and the ability to detect modified bases. PacBio's Single Molecule Real-Time (SMRT) sequencing (Ardui et al., 2018), and Oxford Nanopore Technologies' nanopore sequencing (Weirather et al., 2017) are two prominent examples of third-generation sequencing platforms. These technologies have opened up new avenues for studying viral diversity, metagenomics, and epigenomics. Vij et al. (2022) presented a groundbreaking study on the introgression and mapping of Cotton Leaf Curl Disease (CLCuD) resistance from a 'synthetic cotton polyploid' to upland cotton. They developed a backcross population (synthetic polyploid/Gossypium hirsutum Acc. PIL 43/G. hirsutum Acc. PIL 43) to investigate the inheritance and mapping of CLCuD resistance. The study observed the dominance of CLCuD resistance over susceptibility, revealing the presence of two dominant genes conferring resistance to CLCuD. Molecular analysis using genotyping-by-sequencing identified chromosomes A01 and D07 as carrying one CLCuD resistance gene each. This research marks the first instance of successfully transferring and mapping CLCuD resistance from a synthetic polyploid to upland cotton. Furthermore, Hussain et al. (2023) conducted a sequencing study on the Mac7 accession, a source of Gossypium hirsutum known for its resistance against various biotic stresses. Through alignment with the Gossypium hirsutum 'TM-1′' genome (AD1), they identified 4.7 million SNPs and 1.2 million InDels in the Mac7 genome. Gene ontology and metabolic pathway enrichment analyses indicated the involvement of these SNPs and InDels in processes such as nucleotide binding, secondary metabolite synthesis, and plant-pathogen interaction

pathways. Additionally, RNA-seq data from different tissues and qPCR expression profiling under Cotton Leaf Curl Disease (CLCuD) revealed the roles of individual genes in resistant and susceptible accessions. Notably, the expression of differential NLR genes was higher in resistant plants compared to susceptible ones. The resequencing results offer foundational data for identifying DNA resistance markers, potentially aiding marker-assisted breeding for the development of Mac7-derived resistance lines. Ahmed et al. (2021), utilized CIDER-seq (Circular DNA Enrichment Sequencing), a recently developed PCR-free virus enrichment sequencing method, in conjunction with Sanger sequencing to investigate the genetic diversity of the Cotton leaf curl disease (CLCuD) complex. The study identified a highly recombinant strain of Cotton leaf curl Multan virus and a recently evolved strain of Cotton leaf curl Multan betasatellite, both prevalent in major cotton-growing regions. Additionally, multiple species of alphasatellites were identified, including Mesta yellow vein mosaic alphasatellite (MeYVMA) observed for the first time in cotton. The study also employed real-time quantitative PCR to determine the relative abundance of the virus and associated satellites. This research represents the first study to characterize the CLCuD complex during its third epidemic.

7 Next-generation sequencing applications in cotton genomics

7.1 Genome sequencing and assembly

NGS technologies have displayed a significant role in sequencing and assembling the cotton genome. The availability of the cotton sequenced genome has facilitated the discoveries of genes responsible for important cotton traits, like disease resistance (CLCuD) and fiber quality. Both diploid and tetraploid cotton genomes have been sequenced using NGS technologies, providing valuable insights into the genetic basis of cotton traits. For example,; Hussain et al. (2023), sequenced

G. hirsutum accessions named Mac7 (known to possess resistance against several biotic stresses) to identified the genetic bases associated with resistance against CLCuD. A total of 1.2 and 4.7 million InDels and SNPs were identified and their roles in resistance against CLCuD were accessed using metabolic pathway enrichment and gene ontology. Furthermore, RNA-seq data collected from different tissues at different growth stages identified individual genes associated with resistance and susceptibility in Mac7 against CLCuD (Hussain et al., 2023). Virus is famous for changing its strain rapidly and it became difficult to separate the newly developed strain from the previous one.

Genome sequencing technologies have now made it simple to separate the new strain one from the previous one and also detect the genes/alleles associated with the CLCuD. For example,; Kumar et al. (2010) collected unknown leaf sample affected with CLCuD from two different sites of Rajasthan and DNA-A components of both samples were characterized and amplified. DNA-A of one sample consist of 2,759 nucleotides and the second sample consist of 2,751 nucleotides with both having total of six open reading frames (ORF's). Upon sequence characterization, the first sample

showed sequence similarity with CLCuV strain reported in Pakistan and the second sample showed sequence similarity with the local Rajasthan CLCuV strain. These sequence findings suggest a potential movement of certain strains of the CLCuV from Pakistan to India, or the concurrent existence of diverse isolates in similar geographical conditions. Whereas CLCuV-SG01 shared the highest nucleotide sequence similarity with CLCuV Rajasthan (Abohar), the V1 ORF (encoding coat protein) of SG01 displayed the highest nucleotide identity (100%) with CLCuV Multan (Bhatinda) and Abohar virus, albeit with variations in the AC1 region. The entire nucleotide sequence of SG01 exhibited an 86% similarity to CLCuV Multan virus. A similarity search unveiled significant disparities in AC1 and AV1 regions concerning DNA-A, indicative of a recombination-driven evolutionary history. Computational analysis through the Recombination Detection Program (RDP) supports this recombination hypothesis, suggesting the occurrence of recombination with other *Begomoviruses* within the AC1-ORF of CLCuV-SG01 and V1 ORF, as well as the AC1-ORF of CLCuV-SG02, alongside the intergenic noncoding region (Kumar et al., 2010).

7.2 Transcriptome analysis and gene expression profiling

RNA sequencing (RNA-seq) has become a popular approach for studying gene expression in cotton. By sequencing the transcriptome, researchers can identify differentially expressed genes under various conditions, such as stress, development, and tissue-specific expression (Berkowitz et al., 2021). RNA-seq has enabled the discovery of non-coding RNAs, alternative splicing events and novel transcripts in cotton (Qin et al., 2022; Singh et al., 2023). The commonly grown cotton species, *Gossypium hirsutum* is easily affected by CLCuD whereas; the diploid species *G. arboreum* naturally has some resistance (Hussain et al., 2023). However, the mechanism behind how CLCuD affects the genes expression profile of *G. arboreum* and how the disease interacts with it is unclear. Therefore, Naqvi et al. (2017) utilized RNA-Seq technique to look at how genes behave in *G. arboreum* when it is dealing with CLCuD. They infected *G. arboreum* plants with CLCuD by 600 grafting parts from infected *G. hirsutum* plants and studied the genes in both healthy and infected *G. arboreum* plants using RNA-seq with an Illumina HiSeq 2,500 machine. A total of 1,062 differentially expressed genes (DEGs) were identified of which 17 significant DEGs having possible role in disease resistance were studied for their expression profiles using qPCR. They also found genes that are related to fighting the disease and protecting the plant. Using the RNA-Seq data, they made a network of genes that work together. Within this network, they identified 50 hub genes, many of which help transport things inside the plant. These genes might be important for how *G. arboreum* defends itself against CLCuD. This study is a starting point for understanding how the virus and the plant interact and finding the key genes that help *G. arboreum* resist CLCuD (Naqvi et al., 2017). Furthermore, transcriptome analysis and gene expression profiling offer insights into how the plant reacts when CLCuD is transmitted through whiteflies in *G. hirsutum*. Naqvi et al. (2019) utilized RNA-Seq approach to delve into the variations in gene activity

within the susceptible *G. hirsutum* variety when it is infected by whiteflies carrying the CLCuD. To compare, they subjected both uninfected and CLCuD-infected plants to RNA-Seq analysis using the Illumina HiSeq 2,500 platform. They identified 468 differentially expressed genes (DEGs), consisting of 248 downregulated and 220 upregulated DEGs that play roles in responding to diseases and defending against pathogens. Ten of these genes were selected for further analysis using RT-qPCR on two susceptible cultivars, MNH 786 and Karishma. The consistent gene expression patterns across both cultivars affirmed the reliability of their transcriptome data, suggesting the applicability of their study's insights into the broader context of host susceptibility to CLCuD. They proceeded to conduct weighted gene co-expression network analysis, revealing six modules within the gene interaction networks. This analysis also pinpointed highly co-expressed genes, with 55 hub genes demonstrating co-expression with at least 50 other genes. Notably, many of these hub genes were found to be downregulated and involved in cellular processes. The diminished expression of these co-expressed genes indicated their potential role in favoring the virus and amplifying the plant's vulnerability to CLCuD. They also discuss the potential mechanisms underpinning the establishment of disease susceptibility. Altogether, their study presents a comprehensive analysis of gene activity changes in *G. hirsutum* when infected by whitefly transmitted CLCuD. This research contributes to a better understanding of the combined impact of whiteflies and the virus on their host, offering insights into significant *G. hirsutum* genes that play a role in its susceptibility to CLCuD (Naqvi et al., 2019). Transcriptome analysis and gene expression profiling have also been successfully applied to crops other than cotton, such as wheat (Lee et al., 2022; Xu et al., 2022), rice (Thomas et al., 2019; Kong et al., 2020), maize (Wang et al., 2019; Zhang et al., 2020), tomato (Wen et al., 2019; Riccini et al., 2021) and many others.

7.3 Identification of genetic variation

The creation of genetic diversity has always been the prime target of plant breeders to improve crop growth, development, and yield. Plant breeders use different techniques from hybridization to induce mutations (chemical or physical) to create genetic diversity in plants (Wang et al., 2020). In the case of CLCuD, inducing mutation via chemical or physical means to develop resistance against CLCuD in susceptible cotton genotypes is another modern breeding technique (Rehman et al., 2023). However, after inducing tolerance against CLCuD via mutation in susceptible line, detecting the mutated gene or single nucleotide polymorphisms (SNPs) associated with resistance is another objective of breeders which can only be detected after whole genome sequencing (Chen et al., 2022). NGS technologies have revolutionized the identification and characterization of genetic variation in cotton. By sequencing multiple cotton genotypes, researchers can detect SNPs, insertions, deletions, and structural variations. This information is crucial for marker-assisted selection (MAS) or marker-assisted breeding (MAB), genetic mapping, and understanding the complex traits genetic architecture in cotton (Kushanov et al., 2021). Molecular markers have shown important role in agriculture for cultivar and agronomical trait improvement such as detection of resistance genes

in plants. Numerous QTLs have been developed based on different experimental designs and genome mapping has been greatly improved and has provided a number of QTLs for major and minor crops including cotton (Zargar et al., 2015). Molecular markers linked with QTLs have been identified for various diseases in different crops. MAS has been successfully used in the development of resistant varieties. In cotton, QTLs related to disease are described as simple inherited traits. However, molecular breeding for disease resistance/tolerance is more specific, that is why a lot of work is needed to be done for locating QTLs linked with CLCuD resistance in cotton (Javed et al., 2019). In the recent years, genome map construction has gained an immense significance in MAS. MAS has been acquired as a molecular tool for the selection of desirable phenotypes using DNA markers such as RFLP, RAPD and AFLP on the basis of agriculturally important traits and the genes linked with tolerance against several abiotic and biotic factors (Jones et al., 1997; Khan et al., 2023). Use of MAS in plant breeding has opened new doors in crop improvement (Winter and Kahl, 1995).

Although, CLCuD has always been an important topic of discussion for botanists, yet exact opinion is still needed to be developed on the inheritance and spread of CLCuD (Mahmood et al., 2003). A study conducted by Rahman et al. was based on the collection of minor genes associated with CLCuV disease resistant material, selected by recurrent method (Rahman et al., 2005). Resistance is controlled by major genes, which are dominant and may be affected by the evolution of pathogens (Iqbal et al., 2003).

7.4 Epigenomics and DNA methylation analysis

Epigenetic modifications, such as DNA methylation, play a crucial role in regulating gene expression. NGS technologies have enabled the study of DNA methylation patterns in cotton, providing insights into epigenetic regulation of important agronomic traits. DNA methylation analysis can help understand the impact of environmental factors on gene expression and phenotype (He et al., 2022). DNA methylation in plants take place at cytosine residues in the context of CHH, CHG and CG (where H can be T, C or A) sequences (Feng et al., 2020). DNA methylation patterns can be heritable and can influence gene expression and phenotype. In cotton, DNA methylation has been shown to play a role in various developmental processes, stress responses, and disease resistance. To discover DNA methylation alterations during cotton domestication and evolution, researchers have developed single-base resolution methylomes from different cotton species and varieties. They found significant differences in DNA methylation patterns among diploid and tetraploid cottons, as well as between wild and cultivated cottons (Song et al., 2017). These DNA methylation changes were found to be associated with gene expression changes and may have contributed to the domestication and adaptation of cotton plants. Moreover, transposable elements (TEs) are usually linked with DNA methylation and can contribute to complexity of genome. In diploid cotton species, there are differences in TE content and DNA methylation patterns between different species (Song et al., 2017). The presence of TEs in genic regions can affect gene expression and evolution. The study of DNA methylation divergence among diploid

cotton species provides deep information in understanding the role of DNA methylation in genome evolution and speciation. DNA methylation can play a role in disease resistance in plants (Tirnaz and Batley, 2019). In cotton, DNA methylation changes have been associated with resistance to CLCuD. The betasatellite, a key component of the CLCuD complex, is involved in manipulating host cellular functions and suppressing gene silencing. DNA methylation changes in response to betasatellite infection may contribute to disease resistance by regulating gene expression and host defense responses. qPCR analysis has been used to evaluate the CLCuD-associated *Begomoviruses* titers and their impact on cotton yield. qPCR is a powerful method for quantifying the levels of virus and satellite DNA in infected plants. The results of qPCR analysis have unveiled a significant negative correlation between betasatellite titer and seed cotton yield (SCY) (Iqbal et al., 2023). This finding suggests that betasatellite titer is a critical factor in determining cotton yield and productivity.

7.5 Functional genomics

Functional genomics has emerged as a crucial scientific discipline in recent years, revolutionizing our understanding of plant biology and genetics improvement (Van Bel et al., 2022). In the case of cotton, functional genomics offers the potential to systematically exploit genetic resources, improve resistance against biotic stresses, enhance yield, and optimize germplasm for cultivation (Chen et al., 2007; Wei et al., 2022). However, several challenges persist in cotton functional genomics, including the long growth cycle, low transformation efficiency, large genome size and the lack of efficient molecular tools. Nonetheless, considerable advancements have been achieved in the development of genetically modified cotton varieties, particularly against resistance to CLCuD and other insects (Shafique et al., 2022). The sequencing of genomes plays a pivotal role in functional genomics. In 2007, the Cotton Genome Consortium strategized to sequence cotton genomes, prioritizing less-complicated diploid genomes that could be directly applied to tetraploid cotton (Wang et al., 2012). Consequently, the draft genome sequence of *G. raimondii*, a diploid species, was released in 2012, providing valuable insights into the larger 'A' diploid and 'AD' tetraploid cotton genomes (Li et al., 2014). Subsequently, the genome of *G. arboreum*, a supposed donor species for the A chromosome group in tetraploid cotton, was also sequenced (Watson et al., 2018). However, the precise species directing the development of tetraploid cotton currently remains unknown, highlighting the need for further genome sequencing to understand the biology and evolutionary history of CLCuD. Despite the availability of reference genome sequences for tetraploid and diploid cotton species, some discrepancies exist, potentially due to assembly errors. Therefore, efforts should be directed towards improving the quality control standards for genome assemblies and re-sequencing for accurate comparison and analysis of different cotton species. Vij et al. (2022) used one of the functional genomics approach (genotyping by sequencing) to detect the genes associated with the CLCuD in both upland cotton and backcross population (synthetic cotton polyploidy). They identified two potential genes on chromosomes D07 and

A01 that confer resistance to CLCuD (Vij et al., 2022). Similarly; Farooq et al. (2022), compared the CLCuD transmission behavior of different whitefly species (MEAM1 and Asia II 7) with different time frames using transcriptomics. Sequencing using Illumina technology unveiled, after 6 and 12 h of acquiring CLCuD, different sets of genes were differentially expressed in viruliferous (VF) whiteflies (MEAM 1 and Asia II 7) compared to aviruliferous (AVF) whiteflies. The main categories of these DEGs were associated with transcription factors (TFs) of HTH-1 class in Asia II 7 and zf-C2H2 class in MEAM1. Interestingly, MEAM1 species showed more transcriptional changes and increased immune-related responses after CLCuD infection compared with Asia II 7. Both species had specific genes linked to antiviral innate immunity, lysosome function, autophagy/phagosome pathways, spliceosome, detoxification, signaling pathways, carbohydrate metabolism, transport, and protein digestion. Validation through RT-qPCR and RNA-seq showed consistent expression in 23 out of 28 selected genes. These findings provide insights into how *Begomoviruses* interact with whiteflies, affecting their transmission dynamics, and offer new targets for managing insect-transmitted plant viruses in a sustainable way (Farooq et al., 2022).

To facilitate the comprehensive study of cotton genomes, several functional genomics databases have been developed for the cotton research community. These databases provide access to crucial information such as genome sequences, map positions, protein expression and mRNA data, allelic variation and metabolic pathways. CottonGen (https://www.cottongen.org), is a curated web based intellectual database that offers convenient access to cotton genetic and genomic data. It includes germplasm resources, genes, trait loci, markers, genetic maps, unigenes from ESTs and annotated whole-genome sequences. Another significant database, Cotton Functional Genomic Database (CottonFGD; https://cottonfgd.org), provides access to transcriptome data, functional annotations, genome sequences, and genome re-sequencing data for sequenced *Gossypium* genomes. Additionally, databases like the Database for Co-expression Networks with Function Modules (ccNET; http://structuralbiology.cau.edu.cn/gossypium/), Join Genome Institute (JGI; http://jgi.doe.gov), Comparative Evolutionary Genomics of Cotton (http://cottonevolution.info/), Platform of Functional Genomics Analysis in *Gossypium raimondii* (GraP; http://structuralbiology.cau.edu.cn/GraP/about.html), Cotton Genome Database (CottonDB; http://www.cottondb.org), and Cotton Genome Resource Database (CGRD; http://cgrd.hzau.edu.cn/index.php) offer valuable resources for in-depth analysis and exploration of cotton functional genomics.

7.6 Speed breeding

In the world of plant breeding, the traditional methods of crossbreeding and selection can be 757 time-consuming and labor-intensive. However, with the advent of speed breeding, plant breeders now have a powerful tool at their disposal to accelerate the generation of new plant varieties. Speed breeding involves manipulating environmental factors to promote rapid plant growth and shorter generation times and development of tolerance against diseases like CLCuD (Jähne et al., 2020). Plant breeding has long been a slow and meticulous process, often taking years or even decades to develop new crop varieties with desirable traits. This time-consuming nature of traditional breeding techniques has posed challenges in addressing urgent agricultural needs, such as developing crops with high nutritional profile, higher yield, and resilience to biotic and abiotic factors. This is where speed breeding comes into play (Ghosh et al., 2018).

Speed breeding offers a way to shorten the breeding cycle by creating optimal growth conditions for plants. By manipulating factors such as light exposure, temperature, and humidity, breeders can accelerate plant growth and achieve multiple generations within a single year. This rapid generation turnover opens up new possibilities for crop improvement and allows breeders to respond more effectively to emerging challenges in agriculture (Chiurugwi et al., 2019). To combat CLCuD and develop resistant cotton varieties, speed breeding has emerged as a promising technique. By accelerating the breeding cycle, breeders can quickly screen large populations of cotton plants for disease resistance and select the best-performing individuals for further breeding. This rapid selection process allows breeders to develop new cotton varieties with improved resistance to CLCuD in a shorter timeframe. So far, no research has reported the application of speed breeding against CLCuD in cotton. In addition to cotton, speed breeding has been used successfully in other cereal crops such as wheat and rice to improve disease resistance and final grain yield (Alahmad et al., 2018; Chimmili et al., 2022; Schoen et al., 2023). Given the successful use of speed breeding in other crops, cotton breeders and molecular biologists are now working on its use to improve biotic and abiotic stress tolerance in cotton.

One of the critical environmental factors in speed breeding is lighting. Plants have well-established light-sensitive receptor pathways that respond to specific wavelengths of light, triggering growth and development processes (Hickey et al., 2019). In speed breeding, breeders use specialized LED lamps to provide the necessary light spectrum for optimal plant growth and development. LED lamps offer several advantages over traditional lighting sources, such as fluorescent lamps. LED lamps can be precisely tuned to emit specific wavelengths of light; allowing breeders to customize the light spectrum based on the crop and desired traits (Gavhane et al., 2023). By manipulating the blue and redlight wavelengths, breeders can induce early flowering in plants, effectively shortening the breeding cycle. Furthermore; Nexsel, a leading provider of speed breeding solutions, has been at the forefront of developing optimized lighting protocols for various crops, including cotton. With their specialized grow cabinets and a tailored lighting solution, Nexsel offers a comprehensive package for speed breeding in cotton. The Nexsel Speed Breeding Growth Chamber provides precise control over critical parameters such as temperature, humidity, CO_2 levels, lighting spectrum, light intensity, and photoperiod. This level of control allows breeders to create the ideal growth conditions for cotton plants, maximizing their potential for rapid growth and development. Nexsel's lighting solutions for speed breeding, such as the All-in-One Grow Light and the 3-in-1 Grow Light, offer customizable spectra and intensity options to meet the specific needs of cotton plants at different growth stages (Jois, 2021; Gavhane et al., 2023).

8 Current challenges and future directions

In recent years, advanced biotechnological and molecular breeding techniques have shown promise in the field of cotton research. Precision genome editing *via* multiplex genome editing and HDR have emerged as potential strategies for controlling CLCuD and improving cotton varieties. Additionally, the *de novo* domestication of wild species and the use of synthetic biology approaches offer new avenues for developing resistance against CLCuD and enhancing cotton productivity.

8.1 Precision genome editing via HDR in cotton

Precision genome editing through HDR holds great potential for accelerating cotton improvement and developing resistance against CLCuD. HDR allows for the precise replacement of existing alleles, offering a more targeted and efficient approach compared to traditional breeding methods. However, the application of HDR for particular allele or gene replacement in cotton remains challenging, and as of now, no stable lines with precise editing have been successfully obtained. To enhance HDR efficiency in cotton, several strategies used in other plant species and mammalian cells can be evaluated. One such strategy is the use of a modular RNA aptamer-streptavidin system to enrich the availability of donor repair templates (DRTs), which has shown significant improvements in HDR efficiency in human cells (Carlson-Stevermer et al., 2017).

Another approach involves increasing the temporal and spatial colocalization of DRTs and the Cas9 protein, which can be achieved through the use of SNAP-tag technology. Covalently tethering single-stranded donor oligonucleotides (ssODNs) to the Cas9/guide RNA ribonucleoprotein complex has also been shown to increase HDR efficiency by bringing the DRTs closer to the DNA double-strand break (DSB) site (Savic et al., 2018). Furthermore, utilizing tissue-specific promoters, such as the egg cell- and early embryo-specific DD45 gene promoter, can increase the efficiency of HDR-mediated genome editing. This is particularly relevant in cotton, as HDR take place with more frequency in egg cells (Miki et al., 2018; Wolter et al., 2018). Additionally, the development of a tandem repeat-HDR (TR-HDR) strategy (Lu et al., 2020), which involves the random insertion of a modified DNA fragment, shows promise for seamless in-locus tagging and fragment replacement in other crops and could potentially be applied to cotton. While these techniques show potential, further research is needed to optimize and adapt them specifically for cotton. Continued efforts to improve HDR efficiency in cotton will pave the way for precise gene/allele replacement and the development of novel germplasm with resistance to CLCuD.

8.2 Multiplex genome editing in cotton

Many agriculturally important traits, including yield per plant, are governed by a complex 827 genetic network or multiple genes. In order to understand the functions of these genes more thoroughly, expedite gene discovery, and facilitate breeding efforts, a proficient multiplex genome editing platform is essential. Multiplex genome editing enables the knockout of individual genes, combined genes, or specific gene regions, providing the means to dissect traits and support marker-assisted breeding. In cotton, the formulation of a proficient multiplex editing strategy is still in its early stages. However, insights from other plant species, such as wheat (Kumar et al., 2019) and *Arabidopsis* (Luo et al., 2021), can guide the optimization and systemic comparison of different vector components in cotton. For example, the use of robust constitutive promoters for the expression of Cas9 protein or guide RNA (gRNA), along with additional nuclear localization signals (NLSs), can contribute to the development of a more effective multiplex editing strategy.

Recent advancements in multiplex genome editing in wheat using a polycistronic tRNA strategy offer potential insights for cotton research. In an elite wheat variety, the simultaneous editing of multiple genes at up to 15 genomic loci was achieved by producing multiple mature single chimeric guide RNAs (sgRNAs) simultaneously. This approach utilized an expression cassette controlled by a rice Actin promoter and terminated by a Poly A sequence to enhance the transcripts stability (Luo et al., 2021). By adapting similar strategies to cotton, it may be possible to achieve efficient multiplex editing and accelerate gene discovery and breeding efforts. Expanding the horizons of multiplex editing in cotton can also involve the use of multiple Cas nucleases with different PAM requirements, like Cas12a, SaCas9, SpRY, SpG, xCas9 and SpCas9-NG (Zhang et al., 2019a; Jordan et al., 2021; Li et al., 2021). Additionally, multiplexing systems can be employed to target non-coding RNA regions and several other genetic regulatory elements, enabling the evaluation of their roles and further promoting cotton improvement against CLCuD. The development of a flexible and efficient CRISPR/Cas multiplexing system in cotton will highly facilitate fundamental biological research and translational breeding processes.

8.3 *De novo* domestication of wild species and site-directed artificial evolution of agriculturally important genes

Conventional breeding practices usually result in low stress resistance and loss of genetic diversity, posing challenges to crop stability and food security. To address these issues, the *de novo* domestication of wild species and site-directed artificial evolution of agriculturally important genes have emerged as potential approaches (Hickey et al., 2019). Through the comprehensive application of synthetic biology, genome editing, genomics and other cutting-edge technologies, wild species or landraces can be rapidly domesticated without disturbing the important traits (Eshed and Lippman, 2019). For instance, the introduction of traits of interest into stress-resilient wild tomato genotypes *via* multiplex genome editing resulted in offspring that displayed domesticated phenotypes while retaining salt tolerance and parental disease resistance (Li et al., 2018; Zsögön et al., 2018).

A rational approach for developing new crops varieties involves the *de novo* domestication of wild allotetraploid rice (*Oryza alta*, CCDD) has also been proposed. By editing several traits in *O. alta*

and successfully producing allotetraploid rice lines with targeted improvements in agronomically important genes, researchers have demonstrated the potential for *de novo* domestication in cotton relatives or landraces (Yu et al., 2021). Additionally, base editing techniques could permit the artificial evolution of agriculturally significant genes in existing crop varieties, leading to the development of new germplasm with high genetic diversity. While these approaches have not been documented in cotton to date, the lessons learned from other crop species can inform future research and applications. The exploration of wild cotton species and related plant genomes, along with the manipulation of diverse gene modules, will provide cotton breeders with noval genes and modules for crop improvement. Comparative genomic analysis and in-depth studies of *Gossypium* species will contribute to the development of novel germplasm and the improvement of cotton varieties against CLCuD.

8.4 Development of genotype-independent and in planta delivery strategies

CRISPR-based genome editing techniques rely on the delivery of foreign DNA into plant cells to enable the function and expression of the CRISPR/Cas reagents. In cotton, this has traditionally been achieved through *Agrobacterium*-mediated transformation or biolistic transformation (Ribeiro et al., 2003; Rajasekaran, 2012). However, these methods are usually limited by low plant regeneration efficiency and the limited number of transformable genotypes, particularly in modern cotton varieties. To overcome these limitations, the development of genotype-independent and *in planta* delivery strategies is essential. A genotype-independent approach has been implemented in maize by introducing the Bbm expression cassette into embryos under the control of the maize auxin-inducible promoter (Lowe et al., 2018). This approach allows for the direct germination of embryos into plants without a callus phase, enabling more uniform and efficient transformation. Synthetic biology approaches, which involve the design and construction of new biological parts and systems, also offer potential solutions for improving cotton transformation. By utilizing synthetic biology principles, it may be possible to redesign the cotton genome to enhance transformation efficiency and broaden the range of transformable genotypes.

9 Concluding remarks

The future of controlling and developing resistance in cotton against CLCuD lies in the 893 application of advanced biotechnological and molecular breeding techniques. Cotton employs a range of defense mechanisms against CLCuD, including resistance genes, RNA interference, and natural defense chemicals. While controlling whitefly populations through insecticides is the primary method, biotechnological and molecular breeding approaches offer the most promising solutions for inducing resistance. NGS technologies have transformed the field of cotton genomics, enabling researchers to unravel the complexities of the cotton genome and understand the genetic basis of important traits. These technologies have opened up new avenues for studying gene expression, genetic variation, epigenomics, metagenomics, and functional genomics in cotton. With advancements in sequencing technologies and bioinformatics tools, the future of cotton genomics looks promising, paving the way for the development of improved cotton varieties with enhanced productivity, fiber quality, and disease resistance. Moreover, DNA methylation also plays a crucial role in cotton evolution, domestication, and disease resistance. Understanding the epigenetic regulation of CLCuD and the role of DNA methylation in cotton can provide valuable insights into the development of disease-resistant cotton varieties. The qPCR analysis of virus and satellite DNA levels can be a useful tool for assessing cotton yield and selecting resistant varieties for cultivation. Future research should focus on unraveling the complex interactions between DNA methylation, gene expression, disease resistance and genome editing in cotton to develop sustainable strategies fo disease management in cotton cultivation. Precision genome editing *via* HDR holds promise for achieving specific allele/gene replacement and developing resistance against CLCuD. Similarly, multiplex genome editing offers opportunities for MAS, gene discovery and trait dissection. The newly domesticated wild species and site-directed artificial evolution of agriculturally important genes provide avenues for enhancing genetic diversity and developing novel germplasm. Additionally, the formulation of genotype-independent and *in planta* delivery strategies, along with the utilization of synthetic biology principles, can overcome the limitations of traditional transformation methods. Furthermore, speed breeding has also revolutionized the field of plant breeding, offering breeders a powerful tool to speed-up the development of new crop varieties. In the case of cotton, speed breeding holds immense promise for combating Cotton Leaf Curl Disease and addressing other agronomic challenges. By leveraging the advantages of specialized lighting solutions and controlled growth environments, breeders can expedite the breeding process, enhance disease resistance, improve fiber quality, increase yield, and develop climate-resilient cotton varieties. With companies like Nexsel leading the way in providing cutting-edge speed breeding solutions, the future of cotton breeding looks brighter than ever.

As these techniques continue to be optimized and adapted specifically for cotton, they have the potential to revolutionize cotton improvement and pave the way for sustainable cotton production. By developing cotton varieties with enhanced resistance against CLCuD, the devastating impacts of this disease can be mitigated, leading to increased yields and improved livelihoods for cotton farmers. The future of cotton research is bright, with biotechnological advancements driving innovation and unlocking the full potential of this important fiber crop.

Author contributions

SN: Conceptualization, Investigation, Writing–original draft. SR: Conceptualization, Writing–original draft, Writing–review and editing. TL: Writing–original draft. DT: Conceptualization, Formal Analysis, Supervision, Writing–review and editing. ZM: Investigation, Writing–original draft. KA: Formal Analysis,

Funding

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.

References

Adewale, B. A. (2020). Will long-read sequencing technologies replace short-read sequencing technologies in the next 10 years? *Afr. J. laboratory Med.* 9 (1), 1–5. doi:10.4102/ajlm.v9i1.1340

Aguiar, M. S., Quintela, E. D., Melo, C. L. P., Hungria, M., Faria, L. C., Wendland, A., et al. (2018). Agronomic performance and yield stability of the RNA interference-based bean golden mosaic virus-resistant common bean. *Crop Sci.* 58 (2), 579–591. doi:10.2135/cropsci2017.06.0355

Ahmed, N., Amin, I., Zaidi, S. S. E. A., Rahman, S. U., Farooq, M., Fauquet, C. M., et al. (2021). Circular DNA enrichment sequencing reveals the viral/satellites genetic diversity associated with the third epidemic of cotton leaf curl disease. *Biol. Methods Protoc.* 6 (1), bpab005. doi:10.1093/biomethods/bpab005

Ahmed, Z. (1999). Prospects and bottlenecks of cotton crop in Pakistan. *Pak Cotton Grow.* 3, 6–7.

Alahmad, S., Dinglasan, E., Leung, K. M., Riaz, A., Derbal, N., Voss-Fels, K. P., et al. (2018). Speed breeding for multiple quantitative traits in durum wheat. *Plant methods* 14 (1), 36–15. doi:10.1186/s13007-018-0302-y

Ali, A., Ahmed, Z., and Guo, Z. (2022). *Cotton growing in Pakistan, Bangladesh and Myanmar. Pest management in cotton: a global perspective*. Wallingford UK: CABI, 53–79.

Ali, A., Zafar, M. M., Farooq, Z., Ahmed, S. R., Ijaz, A., Anwar, Z., et al. (2023). Breakthrough in CRISPR/Cas system: current and future directions and challenges. *Biotechnol. J.* 18, 2200642. doi:10.1002/biot.202200642

Ali, H., Hussain, G., Hussain, S., Shahzad, A., Ahmad, S., Javeed, H., et al. (2015). Early sowing reduces cotton leaf curl virus occurrence and improves cotton productivity. *Cercetari Agronomice Moldova* 47 (4), 71–81. doi:10.1515/cerce-2015-0006

Ali, I., Amin, I., Briddon, R. W., and Mansoor, S. (2013). Artificial microRNA-mediated resistance against the monopartite begomovirus Cotton leaf curl Burewala virus. *Virology J.* 10 (1), 231–238. doi:10.1186/1743-422X-10-231

Amudha, J., Balasubramani, G., Malathi, V., Monga, D., Bansal, K., and Kranthi, K. (2010). Cotton transgenics with antisense AC1 gene for resistance against cotton leaf curl virus. *Electron. J. Plant Breed.* 1 (4), 360–369.

Anzalone, A. V., Randolph, P. B., Davis, J. R., Sousa, A. A., Koblan, L. W., Levy, J. M., et al. (2019). Search-and-replace genome editing without double-strand breaks or donor DNA. *Nature* 576 (7785), 149–157. doi:10.1038/s41586-019-1711-4

Ardui, S., Ameur, A., Vermeesch, J. R., and Hestand, M. S. (2018). Single molecule real-time (SMRT) sequencing comes of age: applications and utilities for medical diagnostics. *Nucleic acids Res.* 46 (5), 2159–2168. doi:10.1093/nar/gky066

Arif, M., Farooq, S., Alasmari, A., Alshehri, M. A., Hashem, M., Alamri, S., et al. (2022). Molecular study of geminiviruses: complex biology, host-vector 970interactions, and increasing diversity. *J. King Saud University-Science* 34 (4), 102051. doi:10.1016/j.jksus.2022.102051

Arora, S., and Chaudhary, B. (2021). Global expression dynamics and miRNA evolution profile govern floral/fiber architecture in the modern cotton (Gossypium). *Planta* 254 (3), 62. doi:10.1007/s00425-021-03711-3

Ashraf, J., Zuo, D., Wang, Q., Malik, W., Zhang, Y., Abid, M. A., et al. (2018). Recent insights into cotton functional genomics: progress and future perspectives. *Plant Biotechnol. J.* 16 (3), 699–713. doi:10.1111/pbi.12856

Atiq, M., Talib, M. Z., Rajput, N. A., Sahi, S. T., Khan, M. A., Usman, M., et al. (2023). NEW fangled tactics towards cotton leaf curl virus disease A review. *J. Nat. Fibers* 20 (2), 2217364. doi:10.1080/15440478.2023.2217364

Azeem, H., Perveen, R., Tahir, M. N., Umar, U.-u.-d., Ölmez, F., and Ali, A. (2022). Prevalence, transmission and molecular characterization of Cotton leaf curl Multan virus infecting hollyhock plants in Pakistan. *Mol. Biol. Rep.* 49 (6), 5635–5644. doi:10.1007/s11033-022-07557-0

Basu, U., Riaz Ahmed, S., Bhat, B. A., Anwar, Z., Ali, A., Ijaz, A., et al. (2023). A CRISPR way for accelerating cereal crop improvement Progress and challenges. *Front. Genet.* 13, 866976. doi:10.3389/fgene.2022.866976

Berkowitz, O., Xu, Y., Liew, L. C., Wang, Y., Zhu, Y., Hurgobin, B., et al. (2021). RNA-seq analysis of laser microdissected *Arabidopsis thaliana* leaf epidermis, mesophyll and vasculature defines tissue-specific transcriptional responses to multiple stress treatments. *Plant J.* 107 (3), 938–955. doi:10.1111/tpj.15314

Bharat, S. S., Li, S., Li, J., Yan, L., and Xia, L. (2020). Base editing in plants: current status and challenges. *Crop J.* 8 (3), 384–395. doi:10.1016/j.cj.2019.10.002

Binyameen, B., Khan, Z., Khan, S. H., Ahmad, A., Munawar, N., Mubarik, M. S., et al. (2021). Using multiplexed CRISPR/Cas9 for suppression of cotton leaf curl virus. *Int. J. Mol. Sci.* 22 (22), 12543. doi:10.3390/ijms222212543

Bourdenx, B., Bernard, A., Domergue, F., Pascal, S., Léger, A., Roby, D., et al. (2011). Overexpression of Arabidopsis ECERIFERUM1promotes wax very-long-chain alkane biosynthesis and influences plant response to biotic and abiotic stresses. *Plant physiol.* 156 (1), 29–45. doi:10.1104/pp.111.172320

Briddon, R., and Markham, P. (1994). Universal primers for the PCR amplification of dicot-lnfecting geminiviruses. *Mol. Biotechnol.* 1 (2), 202–205.

Briddon, R. W., Mansoor, S., Bedford, I. D., Pinner, M. S., and Markham, P. G. (2000). Clones of cotton leaf curl geminivirus induce symptoms atypical of cotton leaf curl disease. *Virus Genes* 20 (1), 19–26. doi:10.1023/a:1008151921937

Brown, J., and Czosnek, H. (2002). Whitefly transmission of plant viruses-III whitefly biology in relation to transmission. *Adv. Botanical Res.* 36, 71–74.

Cao, Y., Zhou, H., Zhou, X., and Li, F. (2020). Control of plant viruses by CRISPR/Cas system-mediated adaptive immunity. *Front. Microbiol.* 11, 593700. doi:10.3389/fmicb.2020.593700

Carlson-Stevermer, J., Abdeen, A. A., Kohlenberg, L., Goedland, M., Molugu, K., Lou, M., et al. (2017). Assembly of CRISPR ribonucleoproteins with biotinylated oligonucleotides via an RNA aptamer for precise gene editing. *Nat. Commun.* 8 (1), 1711. doi:10.1038/s41467-017-01875-9

Chandran, D. (2015). Co-option of developmentally regulated plant SWEET transporters for pathogen nutrition and abiotic stress tolerance. *IUBMB life* 67 (7), 461–471. doi:10.1002/iub.1394

Chauhan, P., Mehta, N., Chauhan, R., Kumar, A., Singh, H., Lal, M. K., et al. (2023). Utilization of primary and secondary biochemical compounds in cotton as diagnostic markers for measuring resistance to cotton leaf curl virus. *Front. Plant Sci.* 14, 1185337. doi:10.3389/fpls.2023.1185337

Chen, C.-Y., Liu, Y.-Q., Song, W.-M., Chen, D.-Y., Chen, F.-Y., Chen, X.-Y., et al. (2019). An effector from cotton bollworm oral secretion impairs host plant defense signaling. *Proc. Natl. Acad. Sci.* 116 (28), 14331–14338. doi:10.1073/pnas.1905471116

Chen, J.-Y., Huang, J.-Q., Li, N.-Y., Ma, X.-F., Wang, J.-L., Liu, C., et al. (2015). Genome-wide analysis of the gene families of resistance gene analogues in cotton and their response to Verticillium wilt. *BMC plant Biol.* 15, 1–15. doi:10.1186/s12870-015-0508-3

Chen, Y., Gao, Y., Chen, P., Zhou, J., Zhang, C., Song, Z., et al. (2022). Genome-wide association study reveals novel quantitative trait loci and candidate genes of lint

percentage in upland cotton based on the CottonSNP80K array. *Theor. Appl. Genet.* 135 (7), 2279–2295. doi:10.1007/s00122-022-04111-1

Chen, Z. J., Scheffler, B. E., Dennis, E., Triplett, B. A., Zhang, T., Guo, W., et al. (2007). Toward sequencing cotton (Gossypium) genomes. *Plant physiol.* 145 (4), 1303–1310. doi:10.1104/pp.107.107672

Chimmili, S. R., Kanneboina, S., Hanjagi, P. S., Ps, B., Sakhare, A. S., P, S., et al. (2022). "Integrating advanced molecular, genomic, and speed breeding methods for genetic improvement of stress tolerance in rice," in *Next-generation plant breeding approaches for stress resilience in cereal crops* (Singapore: Springer Nature Singapore), 263–283.

Chiurugwi, T., Kemp, S., Powell, W., and Hickey, L. T. (2019). Speed breeding orphan crops. *Theor. Appl. Genet.* 132, 607–616. doi:10.1007/s00122-018-3202-7

Collins, F. S., Morgan, M., and Patrinos, A. (2003). The Human Genome Project: lessons from large-scale biology. *Science* 300 (5617), 286–290. doi:10.1126/science.1084564

Devendran, R., Kumar, M., Ghosh, D., Yogindran, S., Karim, M. J., and Chakraborty, S. (2022). Capsicum-infecting begomoviruses as global pathogens: host–virus interplay, pathogenesis, and management. *Trends Microbiol.* 30 (2), 170–184. doi:10.1016/j.tim.2021.05.007

Diamos, A. G., Crawford, J. M., and Mason, H. S. (2019). Fine-tuning expression of begomoviral movement and nuclear shuttle proteins confers cell-to-cell movement to mastreviral replicons in Nicotiana benthamiana leaves. *J. General Virology* 100 (6), 1038–1051. doi:10.1099/jgv.0.001275

Dilip, M. (2016). Cotton leaf curl virus disease: an overview. *Agric. Res. J.* 53 (4), 466–474. doi:10.5958/2395-146x.2016.00093.4

Donnelly, R., and Gilligan, C. A. (2023). A new method for the analysis of access period experiments, illustrated with whitefly-borne cassava mosaic begomovirus. *PLOS Comput. Biol.* 19 (8), e1011291. doi:10.1371/journal.pcbi.1011291

Eshed, Y., and Lippman, Z. B. (2019). Revolutions in agriculture chart a course for targeted breeding of old and new crops. *Science* 366 (6466), eaax0025. doi:10.1126/science.aax0025

Faiz, H., and Abhinav, K. (2021). Development of immunity against cotton leaf curl virus using antisense ac4 gene. *Res. J. Biotechnol.* 16, 9.

Farooq, T., Lin, Q., She, X., Chen, T., Tang, Y., and He, Z. (2022). Comparative transcriptome profiling reveals a network of differentially expressed genes in Asia II 7 and MEAM1 whitefly cryptic species in response to early infection of Cotton leaf curl Multan virus. *Front. Microbiol.* 13, 1004513. doi:10.3389/fmicb.2022.1004513

Fauquet, C. M., and Stanley, J. (2003). Geminivirus classification and nomenclature: progress and problems. *Ann. Appl. Biol.* 142 (2), 165–189. doi:10.1111/j.1744-7348.2003.tb00241.x

Feng, S., Zhong, Z., Wang, M., and Jacobsen, S. E. (2020). Efficient and accurate determination of genome-wide DNA methylation patterns in *Arabidopsis thaliana* with enzymatic methyl sequencing. *Epigenetics chromatin* 13 (1), 42–17. doi:10.1186/s13072-020-00361-9

Fiallo-Olivé, E., and Navas-Castillo, J. (2023). Begomoviruses: what is the secret (s) of their success? *Trends Plant Sci.* 28, 715–727. doi:10.1016/j.tplants.2023.01.012

Figueiredo, A., Monteiro, F., and Sebastiana, M. (2014). Subtilisin-like proteases in plant–pathogen recognition and immune priming: a perspective. *Front. plant Sci.* 5, 739. doi:10.3389/fpls.2014.00739

Fondong, V. N. (2013). Geminivirus protein structure and function. *Mol. plant Pathol.* 14 (6), 635–649. doi:10.1111/mpp.12032

Fortes, I. M., Pérez-Padilla, V., Romero-Rodríguez, B., Fernández-Muñoz, R., Moyano, C., Castillo, A. G., et al. (2023). Begomovirus tomato leaf curl New Delhi virus is seedborne but not seed transmitted in melon. *Plant Dis.* 107 (2), 473–479. doi:10.1094/PDIS-09-21-1930-RE

Gaudelli, N. M., Komor, A. C., Rees, H. A., Packer, M. S., Badran, A. H., Bryson, D. I., et al. (2017). Programmable base editing of A●T to G●C in genomic DNA without DNA cleavage. *Nature* 551 (7681), 464–471. doi:10.1038/nature24644

Gavhane, K. P., Hasan, M., Singh, D. K., Kumar, S. N., Sahoo, R. N., and Alam, W. (2023). Determination of optimal daily light integral (DLI) for indoor cultivation of iceberg lettuce in an indigenous vertical hydroponic system. *Sci. Rep.* 13 (1), 10923. doi:10.1038/s41598-023-36997-2

Ghosh, S., Watson, A., Gonzalez-Navarro, O. E., Ramirez-Gonzalez, R. H., Yanes, L., Mendoza-Suárez, M., et al. (2018). Speed breeding in growth chambers and glasshouses for crop breeding and model plant research. *Nat. Protoc.* 13 (12), 2944–2963. doi:10.1038/s41596-018-0072-z

Gnanasekaran, P., and Chakraborty, S. (2018). Biology of viral satellites and their role in pathogenesis. *Curr. Opin. virology* 33, 96–105. doi:10.1016/j.coviro.2018.08.002

Gnanasekaran, P., KishoreKumar, R., Bhattacharyya, D., Vinoth Kumar, R., and Chakraborty, S. (2019). Multifaceted role of geminivirus associated betasatellite in pathogenesis. *Mol. Plant Pathol.* 20 (7), 1019–1033. doi:10.1111/mpp.12800

Gould, J., and Northcote, D. (1986). Cell-cell recognition of host surfaces by pathogens the adsorption of maize (Zea mays) root mucilage by surfaces of pathogenic fungi. *Biochem. J.* 233 (2), 395–405. doi:10.1042/bj2330395

Gray, S. M., and Banerjee, N. (1999). Mechanisms of arthropod transmission of plant and animal viruses. *Microbiol. Mol. Biol. Rev.* 63 (1), 128–148. doi:10.1128/MMBR.63.1.128-148.1999

Guo, X., Wang, Y., Hou, Y., Zhou, Z., Sun, R., Qin, T., et al. (2022). Genome-wide dissection of the genetic basis for drought tolerance in Gossypium hirsutum L. races. *Front. Plant Sci.* 13, 876095. doi:10.3389/fpls.2022.876095

Hagan, A. K., Kloepper, J. W., Lawrence, K. S., Guertal, E. A., Smith, R. H., Shaw, J. N., et al. (2019). *Auburn university crops*. Cotton Research Report.

Harrison, B., Liu, Y., Khalid, S., Hameed, S., Otim-Nape, G., and Robinson, D. (1997). Detection and relationships of cotton leaf curl virus and allied whitefly-transmitted geminiviruses occurring in Pakistan. *Ann. Appl. Biol.* 130 (1), 61–75. doi:10.1111/j.1744-7348.1997.tb05783.x

Hasan, I., Rasul, S., Malik, T. H., Qureshi, M. K., Aslam, K., Shabir, G., et al. (2019). Present status of cotton leaf curl virus disease (CLCUVD): a major threat to cottonproduction. *Int. J. Cotton Res. Technol.* 1 (1), 1–13. doi:10.33865/ijcrt.001.01.0240

He, S., Zhang, Y., Wang, J., Wang, Y., Ji, F., Sun, L., et al. (2022). H3K4me2, H4K5ac and DNA methylation function in short-and long-term heat stress responses through affecting the expression of the stress-related genes in G. hirsutum. *Environ. Exp. Bot.* 194, 104699. doi:10.1016/j.envexpbot.2021.104699

Heigwer, F., Zhan, T., Breinig, M., Winter, J., Brügemann, D., Leible, S., et al. (2016). CRISPR library designer (CLD): software for multispecies design of single guide RNA libraries. *Genome Biol.* 17 (1), 55–10. doi:10.1186/s13059-016-0915-2

Hickey, L. T., Hafeez, A. N., Robinson, H., Jackson, S. A., Leal-Bertioli, S. C., Tester, M., et al. (2019). Breeding crops to feed 10 billion. *Nat. Biotechnol.* 37 (7), 744–754. doi:10.1038/s41587-019-0152-9

Hillary, V. E., and Ceasar, S. A. (2022). Prime editing in plants and mammalian cells: mechanism, achievements, limitations, and future prospects. *BioEssays* 44 (9), 2200032. doi:10.1002/bies.202200032

Hu, T., Chitnis, N., Monos, D., and Dinh, A. (2021). Next-generation sequencing technologies: an overview. *Hum. Immunol.* 82 (11), 801–811. doi:10.1016/j.humimm.2021.02.012

Huang, G., Huang, J.-Q., Chen, X.-Y., and Zhu, Y.-X. (2021). Recent advances and future perspectives in cotton research. *Annu. Rev. plant Biol.* 72, 437–462. doi:10.1146/annurev-arplant-080720-113241

Hussain, A., Farooq, M., Naqvi, R. Z., Aslam, M. Q., Siddiqui, H. A., Amin, I., et al. (2023). Whole-genome resequencing deciphers new insight into genetic diversity and signatures of resistance in cultivated cotton Gossypium hirsutum. *Mol. Biotechnol.* 65 (1), 34–51. doi:10.1007/s12033-022-00527-8

Hussain, T., and Ali, M. (1975). *Review of cotton diseases of Pakistan*. Pakistan cottons.

Ibrahim, Y., Abdelhamed, M. F., and Ibrahim, F. M. (2023). Impact of Foliar Application of Some Growth stimulants on the vegetative growth and the Essential oil characters of Ocimum basilicum var. thyrsiflorum (L.). *Egypt. J. Chem.* 66 (4), 435–447. doi:10.21608/EJCHEM.2022.173276.7163

Iqbal, M., Chang, M. A., and Mahmood, A. (2003). *Inheritance of response to cotton leaf curl virus (CLCuV) infection in cotton*. Pakistan: Asian Journal of Plant Sciences. (Pakistan).

Iqbal, Z., Sattar, M. N., Kvarnheden, A., Mansoor, S., and Briddon, R. W. (2012). Effects of the mutation of selected genes of Cotton leaf curl Kokhran virus on infectivity, symptoms and the maintenance of Cotton leaf curl Multan betasatellite. *Virus Res.* 169 (1), 107–116. doi:10.1016/j.virusres.2012.07.016

Iqbal, Z., Shafiq, M., Ali, S., Mahmood, M. A., Siddiqui, H. A., Amin, I., et al. (2023). qPCR assay as a tool for examining cotton resistance to the virus complex causing CLCuD: yield loss inversely correlates with betasatellite, not virus, DNA titer. *Plants* 12 (14), 2645. doi:10.3390/plants12142645

Ismayil, A., Yang, M., Haxim, Y., Wang, Y., Li, J., Han, L., et al. (2020). Cotton leaf curl multan virus βC1 protein induces autophagy by disrupting the interaction of autophagy-related protein 3 with glyceraldehyde-3-phosphate dehydrogenases. *Plant Cell* 32 (4), 1124–1135. doi:10.1105/tpc.19.00759

Jähne, F., Hahn, V., Würschum, T., and Leiser, W. L. (2020). Speed breeding short-day crops by LED-controlled light schemes. *Theor. Appl. Genet.* 133 (8), 2335–2342. doi:10.1007/s00122-020-03601-4

Jain, H., Chahal, S., Singh, I., Sain, S. K., and Siwach, P. (2023). The rising threat of geminiviruses: molecular insights into the disease mechanism and mitigation strategies. *Mol. Biol. Rep.* 50 (4), 3835–3848. doi:10.1007/s11033-023-08266-y

Jastrzębska, M., Wachowska, U., and Kostrzewska, M. K. (2020). Pathogenic and non-pathogenic fungal communities in wheat grain as influenced by recycled phosphorus fertilizers: a case study. *Agriculture* 10 (6), 239. doi:10.3390/agriculture10060239

Javed, M., Hussain, S., and Baber, M. (2019). Role of QTL mapping to circumscribe various diseases in different crops with special emphasis on cotton. *J. Genet. Mol. Biol.* 3 (1). doi:10.35841/genetics-molecular-biology.3.23-33

Ji, X., Zhang, H., Zhang, Y., Wang, Y., and Gao, C. (2015). Establishing a CRISPR–Cas-like immune system conferring DNA virus resistance in plants. *Nat. Plants* 1 (10), 15144–4. doi:10.1038/nplants.2015.144

Jois, S. (2021). Morphological characteristics of hydroponically grown lettuce (Lactuca sativa) treated with Pranic Agriculture. *Egypt. J. Agric. Res.* 99 (4), 391–396. doi:10.21608/EJAR.2021.98969.1159

Jones, N., Ougham, H., and Thomas, H. (1997). Markers and mapping: we are all geneticists now. *New Phytologist* 137 (1), 165–177. doi:10.1046/j.1469-8137.1997.00826.x

Jordan, W. T., Currie, S., and Schmitz, R. J. (2021). Multiplex genome editing in *Arabidopsis thaliana* using Mb3Cas12a. *Plant Direct* 5 (9), e344. doi:10.1002/pld3.344

Joußen, N., Agnolet, S., Lorenz, S., Schöne, S. E., Ellinger, R., Schneider, B., et al. (2012). Resistance of Australian Helicoverpa armigera to fenvalerate is due to the chimeric P450 enzyme CYP337B3. *Proc. Natl. Acad. Sci.* 109 (38), 15206–15211. doi:10.1073/pnas.1202047109

Karlson, C. K. S., Mohd-Noor, S. N., Nolte, N., and Tan, B. C. (2021). CRISPR/dCas9-based systems: mechanisms and applications in plant sciences. *Plants* 10 (10), 2055. doi:10.3390/plants10102055

Karthik, K., Hada, A., Bajpai, A., Patil, B. L., Paraselli, B., Rao, U., et al. (2023). A novel tasi RNA-based micro RNA-induced gene silencing strategy to tackle multiple pests and pathogens in cotton (Gossypium hirsutum L). *Planta* 257 (1), 20. doi:10.1007/s00425-022-04055-2

Katiyar, A., Smita, S., Lenka, S. K., Rajwanshi, R., Chinnusamy, V., and Bansal, K. C. (2012). Genome-wide classification and expression analysis of MYB transcription factor families in rice and Arabidopsis. *BMC genomics* 13, 1–19. doi:10.1186/1471-2164-13-544

Kchouk, M., Gibrat, J.-F., and Elloumi, M. (2017). Generations of sequencing technologies: from first to next generation. *Biol. Med.* 9 (3). doi:10.4172/0974-8369.1000395

Khan, M. A. U., Shahid, A. A., Rao, A. Q., Shahid, N., Latif, A., ud Din, S., et al. (2015). Defense strategies of cotton against whitefly transmitted CLCuV and Begomoviruses. *Adv. Life Sci.* 2 (2), 58–66.

Khan, M. K. R., Ditta, A., Wang, B., Fang, L., Anwar, Z., Ijaz, A., et al. (2023). "The intervention of multi-omics approaches for developing abiotic stress resistance in cotton crop under climate change," in *Sustainable agriculture in the era of the OMICs revolution* (Springer), 37–82.

Khan, S., Mahmood, M., Rahman, S., Rizvi, F., and Ahmad, A. (2020). Evaluation of the CRISPR/Cas9 system for the development of resistance against Cotton leaf curl virus in model plants. *Plant Prot. Sci.* 56 (3), 154–162. doi:10.17221/105/2019-pps

Khan, Z. A., Kumar, R., and Dasgupta, I. (2022). CRISPR/Cas-mediated resistance against viruses in plants. *Int. J. Mol. Sci.* 23 (4), 2303. doi:10.3390/ijms23042303

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

Kong, W., Zhang, C., Qiang, Y., Zhong, H., Zhao, G., and Li, Y. (2020). Integrated RNA-seqanalysis and Meta-QTLs mapping provide insights into cold stress response in rice seedling roots. *Int. J. Mol. Sci.* 21 (13), 4615. doi:10.3390/ijms21134615

Kumar, R., Kaur, A., Pandey, A., Mamrutha, H., and Singh, G. (2019). CRISPR-based genome editing in wheat: a comprehensive review and future prospects. *Mol. Biol. Rep.* 46 (3), 3557–3569.

Kumar, S., Kanakachari, M., Gurusamy, D., Kumar, K., Narayanasamy, P., Kethireddy Venkata, P., et al. (2016). Genome-wide transcriptomic and proteomic analyses of bollworm-infested developing cotton bolls revealed the genes and pathways involved in the insect pest defence mechanism. *Plant Biotechnol. J.* 14 (6), 1438–1455. doi:10.1111/pbi.12508

Kundu, P., Bera, P., Mishra, S., and Vadassery, J. (2023). *Regulatory role of phytohormones in the interaction of plants with insect herbivores*. Plant Hormones in Crop Improvement, Elsevier, 41–64.

Kushanov, F. N., Turaev, O. S., Ernazarova, D. K., Gapparov, B. M., Oripova, B. B., Kudratova, M. K., et al. (2021). Genetic diversity, QTL mapping, and marker-assisted selection technology in cotton (Gossypium spp.). *Front. plant Sci.* 12, 779386. doi:10.3389/fpls.2021.779386

Lee, M. H., Kim, K.-M., Sang, W.-G., Kang, C.-S., and Choi, C. (2022). Comparison of gene expression changes in three wheat varieties with different susceptibilities to heat stress using RNA-seq analysis. *Int. J. Mol. Sci.* 23 (18), 10734. doi:10.3390/ijms231810734

Leke, W. N., Mignouna, D. B., Brown, J. K., and Kvarnheden, A. (2015). Begomovirus disease complex: emerging threat to vegetable production systems of West and Central Africa. *Agric. Food Secur.* 4 (1), 1–14. doi:10.1186/s40066-014-0020-2

Li, F., Fan, G., Wang, K., Sun, F., Yuan, Y., Song, G., et al. (2014). Genome sequence of the cultivated cotton Gossypium arboreum. *Nat. Genet.* 46 (6), 567–572. doi:10.1038/ng.2987

Li, J., Li, H., Chen, J., Yan, L., and Xia, L. (2020). Toward precision genome editing in crop plants. *Mol. Plant* 13 (6), 811–813. doi:10.1016/j.molp.2020.04.008

Li, J., Zhu, L., Hull, J. J., Liang, S., Daniell, H., Jin, S., et al. (2016). Transcriptome analysis reveals a comprehensive insect resistance response mechanism in cotton to infestation by the phloem feeding insect Bemisia tabaci (whitefly). *Plant Biotechnol. J.* 14 (10), 1956–1975. doi:10.1111/pbi.12554

Li, S., Zhang, C., Li, J., Yan, L., Wang, N., and Xia, L. (2021). Present and future prospects for wheat improvement through genome editing and advanced technologies. *Plant Commun.* 2 (4), 100211. doi:10.1016/j.xplc.2021.100211

Li, T., Yang, X., Yu, Y., Si, X., Zhai, X., Zhang, H., et al. (2018). Domestication of wild tomato is accelerated by genome editing. *Nat. Biotechnol.* 36 (12), 1160–1163. doi:10.1038/nbt.4273

Li, X., Schuler, M. A., and Berenbaum, M. R. (2002). Jasmonate and salicylate induce expression of herbivore cytochrome P450 genes. *Nature* 419 (6908), 712–715. doi:10.1038/nature01003

Lim, H.-S., Nam, J., Seo, E.-Y., Nam, M., Vaira, A. M., Bae, H., et al. (2014). The coat protein of Alternanthera mosaic virus is the elicitor of a temperature-sensitive systemic necrosis in Nicotiana benthamiana, and interacts with a host boron transporter protein. *Virology* 452, 264–278. doi:10.1016/j.virol.2014.01.021

Lin, Q., Zong, Y., Xue, C., Wang, S., Jin, S., Zhu, Z., et al. (2020). Prime genome editing in rice and wheat. *Nat. Biotechnol.* 38 (5), 582–585. doi:10.1038/s41587-020-0455-x

Liu, S.-C., Feng, Y.-L., Sun, X.-N., Chen, R.-D., Liu, Q., Xiao, J.-J., et al. (2022). Target residence of Cas9-sgRNA influences DNA double-strand break repair pathway choices in CRISPR/Cas9 genome editing. *Genome Biol.* 23 (1), 165. doi:10.1186/s13059-022-02736-5

Liu, Y., Robinson, D. J., and Harrison, B. D. (1998). Defective forms of cotton leaf curl virus DNA-A that have different combinations of sequence deletion, duplication, inversion and rearrangement. *J. General Virology* 79 (6), 1501–1508. doi:10.1099/0022-1317-79-6-1501

Lowe, K., La Rota, M., Hoerster, G., Hastings, C., Wang, N., Chamberlin, M., et al. (2018). Rapid genotype "independent" Zea mays L.(maize) transformation via direct somatic embryogenesis. *Vitro Cell. Dev. Biology-Plant* 54, 240–252. doi:10.1007/s11627-018-9905-2

Lu, G., Wang, Z., Xu, F., Pan, Y. B., Grisham, M. P., and Xu, L. (2021). Sugarcane mosaic disease: characteristics, identification and control. *Microorganisms* 9 (9), 1984. doi:10.3390/microorganisms9091984

Lu, H., Cui, X., Liu, Z., Liu, Y., Wang, X., Zhou, Z., et al. (2018). Discovery and annotation of a novel transposable element family in Gossypium. *BMC plant Biol.* 18, 307–310. doi:10.1186/s12870-018-1519-7

Lu, Y., Tian, Y., Shen, R., Yao, Q., Wang, M., Chen, M., et al. (2020). Targeted, efficient sequence insertion and replacement in rice. *Nat. Biotechnol.* 38 (12), 1402–1407. doi:10.1038/s41587-020-0581-5

Luo, J., Li, S., Xu, J., Yan, L., Ma, Y., and Xia, L. (2021). Pyramiding favorable alleles in an elite wheat variety in one generation by CRISPR-Cas9-mediated multiplex gene editing. *Mol. Plant* 14 (6), 847–850. doi:10.1016/j.molp.2021.03.024

Mahmood, T., Arshad, M., Gill, M. I., Mahmood, H. T., Tahir, M., and Hussain, S. (2003). Burewala strain of cotton leaf curl virus: a threat to CLCuV cotton resistant varieties. *Asian J. Plant Sci.* 2, 968–970. doi:10.3923/ajps.2003.968.970

Majid, M. U., Sher, Z., Rashid, B., Ali, Q., Sarwar, M. B., Hassan, S., et al. (2020). Role of leaf epicuticular wax load and composition against whitefly population and cotton leaf curl virus in different cotton varieties. *Cytol. Genet.* 54, 472–486. doi:10.3103/s009545272005014x

Mansoor, S., Bedford, I., Pinner, M., Stanley, J., and Markham, P. (1993). A whitefly-transmitted geminivirus associated with cotton leaf curl disease in Pakistan. *Pak. BOTANICAL SOC DEPT Bot. UNIV KARACHI, 32 KARACHI, Pak.* 25, 105–107.

Mansoor, S., Briddon, R. W., Zafar, Y., and Stanley, J. (2003). Geminivirus disease complexes: an emerging threat. *Trends plant Sci.* 8 (3), 128–134. doi:10.1016/S1360-1385(03)00007-4

Mansoor, S., Khan, S. H., Bashir, A., Saeed, M., Zafar, Y., Malik, K. A., et al. (1999). Identification of a novel circular single-stranded DNA associated with cotton leaf curl disease in Pakistan. *Virology* 259 (1), 190–199. doi:10.1006/viro.1999.9766

Mao, Y.-B., Cai, W.-J., Wang, J.-W., Hong, G.-J., Tao, X.-Y., Wang, L.-J., et al. (2007). Silencing a cotton bollworm P450 monooxygenase gene by plant-mediated RNAi impairs larval tolerance of gossypol. *Nat. Biotechnol.* 25 (11), 1307–1313. doi:10.1038/nbt1352

Marino, G., and Funk, C. (2012). Matrix metalloproteinases in plants: a brief overview. *Physiol. Plant.* 145 (1), 196–202. doi:10.1111/j.1399-3054.2011.01544.x

Marwal, A., Sahu, A. K., and Gaur, R. K. (2014). "Transmission and host interaction of Geminivirus in weeds," in *Plant virus–host interaction* (Academic Press), 143–161.

Materatski, P., Jones, S., Patanita, M., Campos, M. D., Dias, A. B., Félix, M. D. R., et al. (2021). A bipartite geminivirus with a highly divergent genomic organization identified in olive trees may represent a novel evolutionary direction in the family Geminiviridae. *Viruses* 13 (10), 2035. doi:10.3390/v13102035

Matsumoto, T., Lian, H.-L., Su, W.-A., Tanaka, D., Liu, C. w., Iwasaki, I., et al. (2009). Role of the aquaporin PIP1 subfamily in the chilling tolerance of rice. *Plant Cell physiology* 50 (2), 216–229. doi:10.1093/pcp/pcn190

Medžiūnė, J., Kapustina, Ž., Žeimytė, S., Jakubovska, J., Sindikevičienė, R., Čikotienė, I., et al. (2022). Advanced preparation of fragment libraries enabled by oligonucleotide-modified 2′, 3′-dideoxynucleotides. *Commun. Chem.* 5 (1), 34. doi:10.1038/s42004-022-00649-9

Meera Krishna, B., Khan, M. A., and Khan, S. T. (2019). Next-generation sequencing(NGS) platforms: an exciting era of genome sequence analysis. *Microb. Genomics Sustain. Agroecosyst.* 2, 89–109. doi:10.1007/978-981-32-9860-6_6

Meng, D., Walsh, M., and Fricke, W. (2016). Rapid changes in root hydraulic conductivity and aquaporin expression in rice (Oryza sativa L.) in response to shoot removal– xylem tension as a possible signal. *Ann. Bot.* 118 (4), 809–819. doi:10.1093/aob/mcw150

Miki, D., Zhang, W., Zeng, W., Feng, Z., and Zhu, J.-K. (2018). CRISPR/Cas9-mediated gene targeting in Arabidopsis using sequential transformation. *Nat. Commun.* 9 (1), 1967. doi:10.1038/s41467-018-04416-0

Monga, D., and Sain, S. K. (2021). Incidence and severity of cotton leaf curl virus disease on different BG II hybrids and its effect on the yield and quality of cotton crop. *J. Environ. Biol.* 42 (1), 90–98. doi:10.22438/jeb/42/1/mrn-1296

Montesinos, L., Bundó, M., Badosa, E., San Segundo, B., Coca, M., and Montesinos, E. (2017). Production of BP178, a derivative of the synthetic antibacterial peptide BP100, in the rice seed endosperm. *BMC plant Biol.* 17 (1), 63–14. doi:10.1186/s12870-017-1011-9

Mubarik, M. S., Wang, X., Khan, S. H., Ahmad, A., Khan, Z., Amjid, M. W., et al. (2021). Engineering broad-spectrum resistance to cotton leaf curl disease by CRISPR-Cas9 based multiplex editing in plants. *GM Crops Food* 12 (2), 647–658. doi:10.1080/21645698.2021.1938488

Musser, R. O., Hum-Musser, S. M., Eichenseer, H., Peiffer, M., Ervin, G., Murphy, J. B., et al. (2002). Herbivory: caterpillar saliva beats plant defences. *Nature* 416 (6881), 599–600. doi:10.1038/416599a

Naqvi, R. Z., Zaidi, S. S.-e.-A., Akhtar, K. P., Strickler, S., Woldemariam, M., Mishra, B., et al. (2017). Transcriptomics reveals multiple resistance mechanisms against cotton leaf curl disease in a naturally immune cotton species, Gossypium arboreum. *Sci. Rep.* 7 (1), 15880. doi:10.1038/s41598-017-15963-9

Naqvi, R. Z., Zaidi, S. S.-e.-A., Mukhtar, M. S., Amin, I., Mishra, B., Strickler, S., et al. (2019). Transcriptomic analysis of cultivated cotton Gossypium hirsutum provides insights into host responses upon whitefly-mediated transmission of cotton leaf curl disease. *PloS one* 14 (2), e0210011. doi:10.1371/journal.pone.0210011

Nogia, V. K., Singh, V., and Meghwal, R. R. (2014). Effect of Cotton leaf curl virus infected plants on the biology of the whitefly, Bemisia tabaci (Hemiptera: aleyrodidae): vector–virus mutualism. *Phytoparasitica* 42 (5), 619–625. doi:10.1007/s12600-014-0402-9

Noueiry, A. O., Lucas, W. J., and Gilbertson, R. L. (1994). Two proteins of a plant DNA virus coordinate nuclear and plasmodesmal transport. *Cell* 76 (5), 925–932. doi:10.1016/0092-8674(94)90366-2

Nurimanguli, A. I. N. I., Wu, Y. L., Pan, Z. Y., An, Q. S., Shui, G. L., Shao, P. X., et al. (2023). Cotton ethylene response factor GhERF91 involved in the defense against Verticillium dahliae. *J. Integr. Agric.* doi:10.1016/j.jia.2023.07.022

Ouattara, A., Tiendrébéogo, F., Becker, N., Urbino, C., Thébaud, G., Hoareau, M., et al. (2022). Synergy between an emerging monopartite begomovirus and a DNA-B component. *Sci. Rep.* 12 (1), 695. doi:10.1038/s41598-021-03957-7

Park, C.-J., and Seo, Y.-S. (2015). Heat shock proteins: a review of the molecular chaperones for plant immunity. *plant pathology J.* 31 (4), 323–333. doi:10.5423/PPJ.RW.08.2015.0150

Patel, M. V. (2016). *De novo assembly of transcriptome and genome data from the Asia I mtCOI genetic clade of Bemisia tabaci. Doctoral dissertation. University of Greenwich.*

Prasad, P., Khatoon, U., Verma, R. K., Sawant, S. V., and Bag, S. K. (2022). Data mining of transcriptional biomarkers at different cotton fiber developmental stages. *Funct. Integr. Genomics* 22 (5), 989–1002. doi:10.1007/s10142-022-00878-0

Qadir, R., Khan, Z. A., Monga, D., and Khan, J. A. (2019). Diversity and recombination analysis of Cotton leaf curl Multan virus: a highly emerging begomovirus in northern India. *BMC genomics* 20 (1), 1–13. doi:10.1186/s12864-019-5640-2

Qazi, J., Amin, I., Mansoor, S., Iqbal, M. J., and Briddon, R. W. (2007). Contribution of the satellite encoded gene betaC1 to cotton leaf curl disease symptoms. *Virus Res.* 128 (1-), 135–139. doi:10.1016/j.virusres.2007.04.002

Qin, L., Li, J., Wang, Q., Xu, Z., Sun, L., Alariqi, M., et al. (2020). 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

Rahman, M., Hussain, D., Malik, T., and Zafar, Y. (2005). Genetics of resistance to cotton leaf curl disease in Gossypium hirsutum. *Plant pathol.* 54 (6), 764–772. doi:10.1111/j.1365-3059.2005.01280.x

Rajasekaran, K. (2012). *Biolistic transformation of cotton embryogenic cell suspension cultures.* Transgenic Cotton: Methods and Protocols, Springer, 59–70.

Rasool, G., Yousaf, S., Ammara, U. e., Iqbal, A., Saeed, M., Amin, I., et al. (2021). Transgenic expression of synthetic coat protein and synthetic replication associated protein produces mild symptoms and reduce begomovirus-betasatellite accumulation in Nicotiana benthamiana. *Front. Agron.* 3, 676820. doi:10.3389/fagro.2021.676820

Rehman, A., Almas, H. I., Qayyum, A., Li, H., Peng, Z., Qin, G., et al. (2023). *Mutation breeding in cotton. Biotechnologies and genetics in plant mutation breeding.* Apple Academic Press, 23–51.

Ribeiro, S., Ambrozevicius, L., Avila, A., Bezerra, I., Calegario, R., Fernandes, J., et al. (2003). Distribution and genetic diversity of tomato-infecting begomoviruses in Brazil. *Archives virology* 148 (2), 281–295. doi:10.1007/s00705-002-0917-0

Riccini, A., Picarella, M., De Angelis, F., and Mazzucato, A. (2021). Bulk RNA-Seq analysis to dissect the regulation of stigma position in tomato. *Plant Mol. Biol.* 105, 263–285. doi:10.1007/s11103-020-01086-9

Rosario, K., Duffy, S., and Breitbart, M. (2012). A field guide to eukaryotic circular single-stranded DNA viruses: insights gained from metagenomics. *Archives virology* 157 (10), 1851–1871. doi:10.1007/s00705-012-1391-y

Roumagnac, P., Lett, J.-M., Fiallo-Olivé, E., Navas-Castillo, J., Zerbini, F. M., Martin, D. P., et al. (2021). Establishment of five new genera in the family Geminiviridae: citlodavirus, maldovirus, mulcrilevirus, opunvirus, and topilevirus. *Archives virology* 167, 695–710. doi:10.1007/s00705-021-05309-2

Saakre, M., Jaiswal, S., Rathinam, M., Raman, K. V., Tilgam, J., Paul, K., et al. (2023). Host-delivered RNA interference for durable pest resistance in plants: advanced methods, challenges, and applications. *Mol. Biotechnol.*, 1–20. doi:10.1007/s12033-023-00833-9

Sahu, P. P., Sharma, N., Puranik, S., Muthamilarasan, M., and Prasad, M. (2014). Involvement of host regulatory pathways during geminivirus infection: a novel platform for generating durable resistance. *Funct. Integr. genomics* 14 (1), 47–58. doi:10.1007/s10142-013-0346-z

Salk, J. J., Schmitt, M. W., and Loeb, L. A. (2018). Enhancing the accuracy of next-generation sequencing for detecting rare and subclonal mutations. *Nat. Rev. Genet.* 19 (5), 269–285. doi:10.1038/nrg.2017.117

Sarwar, G., Sharif, I., Younas, A., Nazir, A., and Farooq, J. (2022). Historical work to combat with cotton leaf curl disease (clcud) in Pakistan. *J. Agric. Res.* 60 (3), 03681157.

Sattar, M. N., Kvarnheden, A., Saeed, M., and Briddon, R. W. (2013). Cotton leaf curl disease–an emerging threat to cotton production worldwide. *J. General Virology* 94 (4), 695–710. doi:10.1099/vir.0.049627-0

Saunders, K., Norman, A., Gucciardo, S., and Stanley, J. (2004). The DNA beta satellite component associated with ageratum yellow vein disease encodes an essential pathogenicity protein (betaC1). *Virology* 324 (1), 37–47. doi:10.1016/j.virol.2004.03.018

Savic, N., Ringnalda, F. C., Lindsay, H., Berk, C., Bargsten, K., Li, Y., et al. (2018). Covalent linkage of the DNA repair template to the CRISPR-Cas9 nuclease enhances homology-directed repair. *elife* 7, e33761. doi:10.7554/eLife.33761

Schoen, A., Wallace, S., Holbert, M. F., Brown-Guidera, G., Harrison, S., Murphy, P., et al. (2023). Reducing the generation time in winter wheat cultivars using speed breeding. *Crop Sci.* 63 (4), 2079–2090. doi:10.1002/csc2.20989

Shafique, T., Shafique, J., Shafique, A., Shafique, S., Shahid, M., Uddin, N., et al. (2022). Crispr/cas9 in gossypium hirsutum (cotton) coker 312 for Clcud cotton leaf curl virus disease resistance mediated by Agrobacterium. *J. Microbiol. Mol. Genet.* 3 (3), 81–96.

Sharma, N., and Prasad, M. (2020). Silencing AC1 of Tomato leaf curl virus using artificial microRNA confers resistance to leaf curl disease in transgenic tomato. *Plant Cell Rep.* 39, 1565–1579. doi:10.1007/s00299-020-02584-2

Sharma, P., Rishi, N., and Malathi, V. G. (2005). Molecular cloning of coat protein gene of an Indian cotton leaf curl virus (CLCuV-HS2) isolate and its phylogenetic relationship with others members of Geminiviridae. *Virus Genes* 30, 85–91. doi:10.1007/s11262-004-4585-x

Shuli, F., Jarwar, A. H., Wang, X., Wang, L., and Ma, Q. (2018). Overview of the cotton in Pakistan and its future prospects. *Pak. J. Agric. Res.* 31 (4), 396. doi:10.17582/journal.pjar/2018/31.4.396.407

Singh, A., Vivek, A., Gupta, K., Sharma, S., and Kumar, S. (2023). Long non-coding RNA and microRNA landscape of two major domesticated cotton species. *Comput. Struct. Biotechnol. J.* 21, 3032–3044.

Singh, I., Kaur, R., Kumar, A., Singh, S., and Sharma, A. (2021). Differential expression of gut protein genes and population density of Arsenophonus contributes to sex-biased transmission of Bemisia tabaci vectored Cotton leaf curl virus. *Plos one* 16 (11), e0259374. doi:10.1371/journal.pone.0259374

Sodha, D., Verma, S. K., Chhokar, V., and Paul, D. (2022). Cotton lead curl viral disease in American cotton (G. Hirsutum): genetic basis of resistance and role of genetic engineering tools in combating CLCUD. *Scientist* 1 (3), 5122–5137.

Song, Q., Zhang, T., Stelly, D. M., and Chen, Z. J. (2017). Epigenomic and functional analyses reveal roles of epialleles in the loss of photoperiod sensitivity during domestication of allotetraploid cottons. *Genome Biol.* 18, 99–14. doi:10.1186/s13059-017-1229-8

Sony, S. K., Kaul, T., Motelb, K. F. A., Thangaraj, A., Bharti, J., Kaul, R., et al. (2023). CRISPR/Cas9-mediated homology donor repair base editing confers glyphosate resistance to rice (Oryza sativa L.). *Front. Plant Sci.* 14, 1122926. doi:10.3389/fpls.2023.1122926

Sun, F., Ma, J., Wang, P., and Yang, Y. (2022). Genome-wide identification of the SAMS gene family in upland cotton (Gossypium hirsutum L.) and expression analysis in drought stress treatments. *Genes* 13 (5), 860. doi:10.3390/genes13050860

Suthar, T., Gupta, N., Pathak, D., Sharma, S., and Rathore, P. (2021). *Morpho-anatomical characterization of interspecific derivatives of Gossypium hirsutum L.× G. armourianum Kearney cross for whitefly tolerance*. Springer: Phytoparasitica, 1–19.

Tang, Y., Zhang, Z., Yang, Z., and Wu, J. (2023). CRISPR/Cas9 and Agrobacterium tumefaciens virulence proteins synergistically increase efficiency of precise genome editing via homology directed repair in plants. *J. Exp. Bot.*, erad096.

Tao, X. Y., Xue, X. Y., Huang, Y. P., Chen, X. Y., and Mao, Y. B. (2012). Gossypol-enhanced P450 gene pool contributes to cotton bollworm tolerance to a pyrethroid insecticide. *Mol. Ecol.* 21 (17), 4371–4385. doi:10.1111/j.1365-294X.2012.05548.x

Tatineni, S., and Hein, G. L. (2021). High Plains wheat mosaic virus: an enigmatic disease of wheat and corn causing the High Plains disease. *Mol. Plant Pathol.* 22 (10), 1167–1179.

Thomas, J., Kim, H. R., Rahmatallah, Y., Wiggins, G., Yang, Q., Singh, R., et al. (2019). RNA-seq reveals differentially expressed genes in rice (Oryza sativa) roots during interactions with plant-growth promoting bacteria, Azospirillum brasilense. *PLoS One* 14 (5), e0217309. doi:10.1371/journal.pone.0217309

Tian, S., Wang, X., Li, P., Wang, H., Ji, H., Xie, J., et al. (2016). Plant aquaporin AtPIP1; 4 links apoplastic H2O2 induction to disease immunity pathways. *Plant physiol.* 171 (3), 1635–1650. doi:10.1104/pp.15.01237

Tirnaz, S., and Batley, J. (2019). DNA methylation: toward crop disease resistance improvement. *Trends plant Sci.* 24 (12), 1137–1150. doi:10.1016/j.tplants.2019.08.007

Uniyal, A. P., Yadav, S. K., and Kumar, V. (2019). The CRISPR–Cas9, genome editing approach: a promising tool for drafting defense strategy against begomoviruses including cotton leaf curl viruses. *J. Plant Biochem. Biotechnol.* 28 (2), 121–132. doi:10.1007/s13562-019-00491-6

Van Bel, M., Silvestri, F., Weitz, E. M., Kreft, L., Botzki, A., Coppens, F., et al. (2022). PLAZA 5.0: extending the scope and power of comparative and functional genomics in plants. *Nucleic Acids Res.* 50 (D1), D1468–D1474. doi:10.1093/nar/gkab1024

Varma, A., and Malathi, V. (2003). Emerging geminivirus problems: a serious threat to crop production. *Ann. Appl. Biol.* 142 (2), 145–164. doi:10.1111/j.1744-7348.2003.tb00240.x

Verchot, J. (2012). Cellular chaperones and folding enzymes are vital contributors to membrane bound replication and movement complexes during plant RNA virus infection. *Front. plant Sci.* 3, 275. doi:10.3389/fpls.2012.00275

Vij, S., Pathak, D., Rathore, P., Kumar, H., Sekhon, P. S., Bhatia, D., et al. (2022). Molecular mapping of CLCuD resistance introgressed from synthetic cotton polyploid in upland cotton. *J. Genet.* 101 (1), 25. doi:10.1007/s12041-022-01365-y

Wang, F., Wang, C., Yan, Y., Jia, H., and Guo, X. (2016). Overexpression of cotton GhMPK11 decreases disease resistance through the gibberellin signaling pathway in transgenic Nicotiana benthamiana. *Front. plant Sci.* 7, 689. doi:10.3389/fpls.2016.00689

Wang, G., Xu, Z., Wang, F., Huang, Y., Xin, Y., Liang, S., et al. (2022). Development of an efficient and precise adenine base editor (ABE) with expanded target range in allotetraploid cotton (Gossypium hirsutum). *BMC Biol.* 20 (1), 45–15. doi:10.1186/s12915-022-01232-3

Wang, J., Hu, Z., Zhao, T., Yang, Y., Chen, T., Yang, M., et al. (2015). Genome-wide analysis of bHLH transcription factor and involvement in the infection by yellow leaf curl virus in tomato (Solanum lycopersicum). *BMC genomics* 16, 39–14. doi:10.1186/s12864-015-1249-2

Wang, K., Wang, Z., Li, F., Ye, W., Wang, J., Song, G., et al. (2012). The draft genome of a diploid cotton Gossypium raimondii. *Nat. Genet.* 44 (10), 1098–1103. doi:10.1038/ng.2371

Wang, W., Niu, S., Dai, Y., Wang, M., Li, Y., Yang, W., et al. (2019). The Zea mays mutants opaque2 and opaque 16 disclose lysine change in waxy maize as revealed by RNA-Seq. *Sci. Rep.* 9 (1), 12265. doi:10.1038/s41598-019-48478-6

Wang, W., Xu, J., Fang, H., Li, Z., and Li, M. (2020). Advances and challenges in medicinal plant breeding. *Plant Sci.* 298, 110573. doi:10.1016/j.plantsci.2020.110573

Wang, Y., and Wu, W.-H. (2010). Plant sensing and signaling in response to K+-deficiency. *Mol. plant* 3 (2), 280–287. doi:10.1093/mp/ssq006

Watson, A., Ghosh, S., Williams, M. J., Cuddy, W. S., Simmonds, J., Rey, M.-D., et al. (2018). Speed breeding is a powerful tool to accelerate crop research and breeding. *Nat. plants* 4 (1), 23–29. doi:10.1038/s41477-017-0083-8

Wei, Y., Liu, Y., Ali, A. M., Xiao, R., Liang, C., Meng, Z., et al. (2022). Rich variant phenotype of Gossypium hirsutum L. saturated mutant library provides resources for cotton functional genomics and breeding. *Industrial Crops Prod.* 186, 115232. doi:10.1016/j.indcrop.2022.115232

Weirather, J. L., de Cesare, M., Wang, Y., Piazza, P., Sebastiano, V., Wang, X.-J., et al. (2017). *Comprehensive comparison of pacific biosciences and Oxford nanopore technologies and their applications to transcriptome analysis*. 1000Research 6.

Wen, J., Jiang, F., Weng, Y., Sun, M., Shi, X., Zhou, Y., et al. (2019). Identification of heat-tolerance QTLs and high-temperature stress-responsive genes through

conventional QTL mapping, QTL-seq and RNA-seq in tomato. *BMC plant Biol.* 19, 1–17. doi:10.1186/s12870-019-2008-3

Winter, P., and Kahl, G. (1995). Molecular marker technologies for plant improvement. *World J. Microbiol. Biotechnol.* 11, 438–448. doi:10.1007/BF00364619

Wolter, F., Klemm, J., and Puchta, H. (2018). Efficient in planta gene targeting in Arabidopsis using egg cell-specific expression of the Cas9 nuclease of Staphylococcus aureus. *plant J.* 94 (4), 735–746. doi:10.1111/tpj.13893

Xia, X. C., Hu, Q. Q., Li, W., Chen, Y., Han, L. H., Tao, M., et al. (2018). Cotton (Gossypium hirsutum) JAZ3 and SLR1 function in jasmonate and gibberellin mediated epidermal cell differentiation and elongation. *Plant Cell, Tissue Organ Cult. (PCTOC)* 133, 249–262. doi:10.1007/s11240-018-1378-9

Xiong, X. P., Sun, S. C., Zhang, X. Y., Li, Y. J., Liu, F., Zhu, Q. H., et al. (2020). GhWRKY70D13 regulates resistance to Verticillium dahliae in cotton through the ethylene and jasmonic acid signaling pathways. *Front. Plant Sci.* 11, 69. doi:10.3389/fpls.2020.00069

Xu, F., Chen, S., Zhou, S., Yue, C., Yang, X., Zhang, X., et al. (2022). Genome-wide association, RNA-seq and iTRAQ analyses identify candidate genes controlling radicle length of wheat. *Front. Plant Sci.* 13, 939544. doi:10.3389/fpls.2022.939544

Xu, Y., Fu, S., Tao, X., and Zhou, X. (2021). Rice stripe virus: exploring molecular weapons in the arsenal of a negative-sense RNA virus. *Annu. Rev. Phytopathology* 59, 351–371. doi:10.1146/annurev-phyto-020620-113020

Yadav, N., Yadav, D. K., Yadav, S., and Khurana, S. P. (2021). *Viral movement-cellular protein interaction. Plant Virus-Host Interaction*. Elsevier, 59–109.

Yang, J., Zhang, H., Chen, H., Sun, Z., Ke, H., Wang, G., et al. (2023). Genome-wide association study reveals novel SNPs and genes in Gossypium hirsutum underlying Aphis gossypii resistance. *Theor. Appl. Genet.* 136 (8), 171. doi:10.1007/s00122-023-04415-w

Yaqoob, S., Fatima, N., Khan, S., Ali, Q., Hafeez, M., and Malik, A. (2020). Begomoviruses and betasatellites associated with CLCuD. *Biol. Clin. Sci. Res. J.* 2020 (1). doi:10.54112/bcsrj.v2020i1.2

Yazdani-Khameneh, S., Aboutorabi, S., Shoori, M., Aghazadeh, A., Jahanshahi, P., Golnaraghi, A., et al. (2016). Natural occurrence of tomato leaf curl New Delhi virus in Iranian cucurbit crops. *plant pathology J.* 32 (3), 201–208. doi:10.5423/PPJ.OA.10.2015.0210

Yu, H., Lin, T., Meng, X., Du, H., Zhang, J., Liu, G., et al. (2021). A route to de novo domestication of wild allotetraploid rice. *Cell* 184 (5), 1156–1170. doi:10.1016/j.cell.2021.01.013

Zargar, S. M., Raatz, B., Sonah, H., Bhat, J. A., Dar, Z. A., Agrawal, G. K., et al. (2015). Recent advances in molecular marker techniques: insight into QTL mapping, GWAS and genomic selection in plants. *J. crop Sci. Biotechnol.* 18 (5), 293–308. doi:10.1007/s12892-015-0037-5

Zerbini, F. M., Briddon, R. W., Idris, A., Martin, D. P., Moriones, E., Navas-Castillo, J., et al. (2017). ICTV virus taxonomy profile: Geminiviridae. *J. general virology* 98 (2), 131–133. doi:10.1099/jgv.0.000738

Zhang, H., Zhang, J., Xu, Q., Wang, D., Di, H., Huang, J., et al. (2020). Identification of candidate tolerance genes to low-temperature during maize germination by GWAS and RNA-seq approaches. *BMC plant Biol.* 20, 1–17. doi:10.1186/s12870-020-02543-9

Zhang, L., Yu, Y., Zhang, M., Rong, K., Wu, Y., Zhang, M., et al. (2023). Genome-wide identification of xylan glucuronosyltransferase family in cotton and function characterization of GhGUX5 in regulating Verticillium wilt resistance. *Int. J. Biol. Macromol.* 245, 124795. doi:10.1016/j.ijbiomac.2023.124795

Zhang, S., Tian, Z., Li, H., Guo, Y., Zhang, Y., Roberts, J. A., et al. (2019a). Genome-wide analysis and characterization of F-box gene family in Gossypium hirsutum L. *BMC genomics* 20, 1–16. doi:10.1186/s12864-019-6280-2

Zhang, Y., Malzahn, A. A., Sretenovic, S., and Qi, Y. (2019b). The emerging and uncultivated potential of CRISPR technology in plant science. *Nat. Plants* 5 (8), 778–794. doi:10.1038/s41477-019-0461-5

Zhong, Y., Cheng, C., Jiang, B., Jiang, N., Zhang, Y., Hu, M., et al. (2016). Digital gene expression analysis of Ponkan Mandarin (Citrus reticulata Blanco) in response to Asia citrus psyllid-vectored Huanglongbing infection. *Int. J. Mol. Sci.* 17 (7), 1063. doi:10.3390/ijms17071063

Zhou, L., He, H., Liu, R., Han, Q., Shou, H., and Liu, B. (2014). Overexpression of GmAKT2potassium channel enhances resistance to soybean mosaic virus. *BMC plant Biol.* 14 (1), 1–11. doi:10.1186/1471-2229-14-154

Zhou, X. (2013). Advances in understanding begomovirus satellites. *Annu. Rev. phytopathology* 51, 357–381. doi:10.1146/annurev-phyto-082712-102234

Zhou, X., Xie, Y., Tao, X., Zhang, Z., Li, Z., and Fauquet, C. M. (2003). Characterization of DNAbeta associated with begomoviruses in China and evidence for co-evolution with their cognate viral DNA-A. *J. General virology* 84 (1), 237–247. doi:10.1099/vir.0.18608-0

Zsögön, A., Čermák, T., Naves, E. R., Notini, M. M., Edel, K. H., Weinl, S., et al. (2018). De novo domestication of wild tomato using genome editing. *Nat. Biotechnol.* 36 (12), 1211–1216. doi:10.1038/nbt.4272