INTERNATIONAL COTTON RESEARCHERS ASSOCIATION

COTTON

INNOVATIONS

Volume 3 .Issue7
September 2023

ICRA
INTERNATIONAL COTTON RESEARCHERS ASSOCIATION
www.icra-cotton.org

Content

ICRA launched a new website 1 — 1

EST-SSRs revealed the Pakistani cotton varieties' genetic bottleneck
Fatima Javeria, Muhammad Arslan, Saba Yaseen, Zulqurnain Khan and Shoaib Ur Rehman — 2

Targeting gene regulatory sequences for inhibition of cell-to-cell movement of plant DNA virus
Fatima Javeria, Muhammad Arslan, Saba Yaseen, Zulqurnain Khan and Shoaib Ur Rehman — 14

Permanent Suppression of SPY Gene via CRISPR/Cas9 Technology for Enhanced Drought Tolerance in Cotton
Saba Yaseen and Zulqurnain Khan — 30

The Cotton Innovations Newsletter is published twelve times annually. Contributed articles, pictures, cartoons, and feedback are welcome at any time. Please send contributions to the General Editors (see below). The editors reserve the right to edit. The deadline for contributions is one month before the publication date.

Editorial Board

Published by ICRA Secretariat, Pakistan Central Cotton Committee, Multan-Pakistan
http://icra-cotton.org
The newsletter is also available at URL: https://icra-cotton.org/cotton-innovation/

ISSN 2788-6611

ICRA launched a new website

The contract of ICRA website with the previous service provider expired in February 2023. The Executive Committee of ICRA decided to get a new website for ICRA as there were some limitations in the previous website. The group e-mail functionality stopped working as the number of ICRA members exceeded 1500. This was a major limitation and the ICRA Secretariat had to send e-mails to members manually and from multiple e-mail addresses. Moreover, new members could just sign up with the existing email credentials that permitted, many fake registrations.

The new website has a better look with added functionalities and probably more secure as detailed information is required from the members at the time of registration. The website can be accessed at "www.icra-cotton.org". Previously registered members are required to register themselves again at the new website. For new registration, click the website address, while you are at homepage, click the sign up icon, at top right corner, for the new registration, a register form opens. Fill in the all required information and make sure to fill as much detail as possible in all fields, then tick your area of expertise from the given options. If you do not find your subject area, click other. There are different options in member type field, please select the correct one. If you are a researcher select member, fellow membership is for ??, student member for students, associate membership for a farmer, corporate membership for persons from industry and sponsor/partner membership for organizations who contribute financially in organizing events of ICRA or other activities.

The new website has an option for email subscriptions for members and non-members. If a person does not opt to be a member of ICRA but wants to receive ICRA newsletter or updates about events, he/she can do it by subscribing through his/her email address. The members and subscribers will receive updates of events and the latest issues of newsletter when uploaded on the website "https://icra-cotton.org/cotton-innovation/". The website will also provide multiple language support when fully active. Although the website is functional but different functions are being improved so occasional halts in functions may be expected.

EST-SSRs revealed the Pakistani cotton varieties' genetic bottleneck

Fatima Javeria[1], Muhammad Arslan[1], Saba Yaseen[1], Zulqurnain Khan[1] and Shoaib Ur Rehman[1*]

[1]Institute of Plant Breeding and Biotechnology, MNS-University of Agriculture-Multan, Pakistan

[*]Corresponding author: shoaib.rehman@mnsuam.edu.pk

Abstract

Cotton, belongs to the *Malvaceae* family and the genus *Gossypium,* is an important source of fibre, edible oil, and seedcake. It is well-known for its adaptability in a variety of settings. Notably, cotton improvement is often based on the creation of new types via innovative combinations, which is heavily dependent on genetic diversity. The goal of our research was to look at the genetic diversity of several cotton cultivars. The 10 EST-SSRs in twenty cotton cultivars was examined. The results indicated a striking lack of genetic variation among the selected types. The PIC was 3.75 for the cultivars NAU1048, NAU3920, and NAU4871. The dendrogram revealed the greatest degree of similarity between chosen Bt and non-Bt cultivars. This genetic bottleneck might be alleviated by using measures such as interspecific hybridization and the introduction of foreign genetic lines.

Keywords: EST-SSRs, Cotton, Genetic bottleneck

INTRODUCTION

Cotton, a fibre crop of global recognition, falls under the genus Gossypium and is a member of the Malvaceae family. It has been observed to possess adaptability to various environmental conditions (Seelanan et al., 1997; Zhang et al., 2008; Ranga et al., 2020; Abdellatif et al., 2012; Parekh et al., 2016). The cultivation, processing, and marketing of this particular crop hold considerable economic importance in numerous countries, serving as a primary source of income for a significant number of individuals (Fayet and Vermeulen, 2014; Malik and Ahsan, 2016). Moreover, the economic ramifications of this phenomenon

transcend national borders, exerting a significant influence on the economic equilibrium of multiple countries (MacDonald, 2012; Devadoss and Sabala, 2020).

Researchers are placing significant emphasis on improving the productivity of this crucial agricultural crop. Improvement techniques encompass a range of strategies pertaining to breeding and biotechnological interventions, which leverage genetics information and genetic diversity. The utilisation of cotton can be traced back to ancient civilizations, as evidenced by the discovery of the earliest cotton fabric at the Mohenjo-daro archaeological site situated along the Indus River in Pakistan, approximately in 3200 BC (Lee, 1984; Lee and Fang, 2015; Akhtar, 2020). Currently, cotton production has achieved global expansion, with Asia emerging as the leading contributor, accounting for approximately 70% of the world's cotton output. Subsequently, America follows with a contribution of around 20%, while Africa accounts for approximately 6%. In contrast, Europe's contribution remains relatively minimal, constituting less than 2% of the global cotton production (Kahriz et al., 2019; Raper et al., 2019; Amanet et al., 2019; Basal et al., 2019).

According to the Food and Agriculture Organisation (FAO), there is a projected decline of 3% in global cotton production for the 2020-21 period, in comparison to the preceding year (FAO, 2020-21). The decline mentioned is a matter of concern, given the projected 102% increase in global cotton demand by 2030, as reported by Xiao et al. (2009). The mismatch between supply and demand emphasises the need to increase cotton output, with a focus on using genetic variety.

It is crucial to protect genetic information and natural resources. Utilising a variety of molecular markers to map cotton genetics may help to provide a more thorough knowledge of genetic diversity and make it easier to find genetic variants that can be used to improve cotton harvests (Chen et al., 2007). The use of genetic resources, in accordance with Abdurakhmonov et al. (2012), may result in the creation of efficient ways to deal with a variety of problems, including insect resistance, tolerance to abiotic stress, improvement of plant health, and enhancement of fibre quality. However, the frequent use of similar genotypes in certain places may lead to genetic erosion, highlighting the need of protecting and using a variety of genetic resources.

Several molecular markers, including Simple Sequence Repeats (SSR), Amplified Fragment Length Polymorphism (AFLP), Random Amplified Polymorphic DNA (RAPD), and Expressional Sequence Tags (EST), have been used to evaluate genetic diversity in different crop genotypes, particularly within the Gossypium genus (Tatineni et al., 1996; Xu et al., 2001; Lu and Myers

Breeders may use these DNA markers to determine the genetic foundation of better qualities for cotton improvement since they are resistant to environmental variation and free from epistatic or pleiotropic effects (Saha et al., 2003). This makes these markers a valuable tool for cotton

germplasm resource exploitation and genetic studies (Liu et al., 2000).

MATERIAL AND METHODS

The Cotton Research Station in Multan provided the twenty different types of cotton. Immature leaf samples from plants were carefully identified, placed in zip-lock bags, and stored at -80°C in a freezer to avoid any deterioration. The following wet laboratory tests were carried out at Multan's MNS-University of Agriculture's Central Laboratory.

Table 1: The following is a compilation of chosen cotton cultivars, along with their respective places of origin

Sr. No.	Variety	Origin of Release	Sr. No.	Variety	Origin of Release
1	BH-184	CRS, Bahawalpur	11	NIAB-Kiran	NIAB, Faisalabad
2	BT-121	India	12	NIAB-824	NIAB, Faisalabad
3	CIM-554	CCRI, Multan	13	NIAB-878	NIAB, Faisalabad
4	CIM-602	CCRI, Multan	14	NIAB-2008	NIAB, Faisalabad
5	CIM-608	CCRI, Multan	15	NIAB-878	NIAB, Faisalabad
6	CIM-612	CCRI, Multan	16	OHO-608	CRI, Multan
7	CYTO-179	CCRI, Multan	17	RH-647	CRS, Khanpur
8	FD-512	CRS, Faisalabad	18	RH-662	CRS, Khanpur
9	FH-Lalazar	CRS, Faisalabad	19	SITARA-008	AzizGroup, Pakistan
10	FH-Noor	CRS, Faisalabad	20	VH-738	CRS, Vehari

DNA extraction and marker selection

DNA from leaves of selected 20 cotton varieties was isolated utilizing the method established by Doyle and Doyle (1987), and the integrity of the extracted DNA was evaluated using 1% agarose gel electrophoresis. The concentration of the DNA samples was then quantified with a spectrophotometer (KAIAO-K5800C). To conduct this study, we employed 10 EST-SSR (NAU-series) markers (Table 2) (Han et al., 2006; Guo et al., 2007). These primers were carefully chosen from the cotton marker database (https://www.cottongen.org/) , ensuring broad genomic representation of the cotton species. The primers were appropriately diluted according to the provider's instructions, and the conditions for the PCR reactions were fine-tuned for optimal amplification outcomes

Table 2: Investigation of NAU Series EST-SSRs: Chromosomal Location and Optimal Annealing Temperatures

Sr. No.	Primer Name	Chr. No.	Primer Sequence	Annealing Temperature
1.	NAU 1048	7	F: GGCCATATTATTGCAGAACC R: ACAGCCTTGAGTTGAGCTTT	57 °C

Sr. No.	Primer Name	Chr. No.	Primer Sequence	Annealing Temperature
2.	NAU 2869	10 and 20	F: GCAAACCAACACCAATACAC R: TCTCCCTTTCAGCATTTAGG	57 °C
3.	NAU 3009	04	F: CCTCCACTTTCAGATGTCCT R: GCCAGCTCAGGATCTATGTT	57 °C
4.	NAU 3201	08 and 24	F: CGGGCCTAACTCCACTAATA R: GATAGGCAAAAGGCAGTGTC	57 °C
5.	NAU 3306	25	F: ACAACCCAAGAGGACAAAGA R: ATAACCACAGCGACCACTTT	57 °C
6.	NAU 3598	14	F: GCTTAAGATCGAATAATCAC R: TGATTGTGCTTTATCGGGTA	57 °C
7.	NAU 3897	12	F: CTCCAATTGGGTCATCATTC R: GTACTCTTCAATCGGCCTTT	57 °C
8.	NAU 3920	26	F: GGGGCAATTTAGGAGTTGTA R: CGTCAGTTTTACACCCCATT	57 °C
9.	NAU 4871	18	F: CTGCTTCTAATGGCCGTAAT R: AAGCCTGGAAAAAGAACCTT	57 °C
10.	NAU 5005	19	F: AAGGTAGGAAGCAATGCAAC R: AAAACATGTAGGAACGAGCA	57 °C

PCR Profiling

The polymerase chain reaction (PCR) procedure involves the utilisation of various essential components to optimise reagents. These components include diluted DNA template, buffer solution, deoxyribonucleotide triphosphates (dNTPs), diluted primers, magnesium chloride ($MgCl_2$), and polymerase enzymes. The PCR amplification protocol typically begins with an initial denaturation step, where the temperature is set to 95°C and maintained for a period of 5 minutes. Following this, a sequence of 35 iterations is conducted, with each iteration comprising of denaturation at a temperature of 95°C for a duration of 45 seconds, annealing at 57°C for a duration of 30 seconds, and extension at a temperature of 72°C for a duration of 45 seconds. The amplification process is finalized by a concluding extension step at a temperature of 72°C, which persists for a duration of 10 minutes.

Bands scoring

The PCR products are separated by employing a 2% agarose gel, which are subjected to electrophoresis at a voltage of 90V for duration of 35 minutes. The utilisation of this procedure facilitates the observation and isolation of alleles on the gel (Wang et al., 2009; Adawy, 2007). The subsequent scoring is determined by the presence (1) or absence (0) of bands, with only unique bands being taken into account. The software known as "Power Marker V3.25" is employed for the purpose of conducting diversity analysis. The similarity index is visualised by constructing a phylogenetic tree using the UPGMA clustering method.

RESULTS

This study aimed to assess the genetic diversity of 20 cotton varieties by employing 10 NAU series markers. The

markers exhibited a moderate degree of genetic diversity. The range of the Polymorphism Information Content (PIC) was observed to be between 0.31924 and 0.3750, with an average value of 0.3593. The Genotype Diversity index (H') exhibited a range of values from 0.2500 to 0.8500, with an average value of 0.605. The genetic similarity coefficient observed among the germplasm sources under investigation ranged from 0.398 to 0.5000, with an average value of 0.471. These findings suggest a moderate level of genetic diversity among the studied sources.

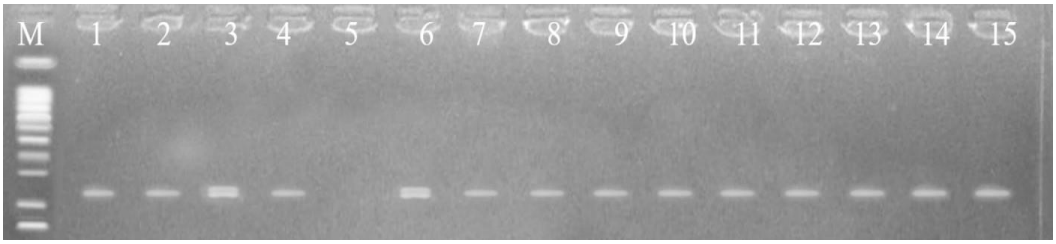

Figure 1: The primer NAU1048 was used to amplify the product from a variety of samples including CIM-473, BT-121, FH-Lalazar, Sitara-008, CIM-554, RH-647, CYTO-179, NIAB-878, NIAB-Kiran, RH-662, BH-184, NIAB- 2008, CIM-608, CIM-612, and FD-512. These samples are represented in lanes 1 to 15 respectively on a 2% agarose gel. The reference used was a 100bp ladder, labeled as 'M'.

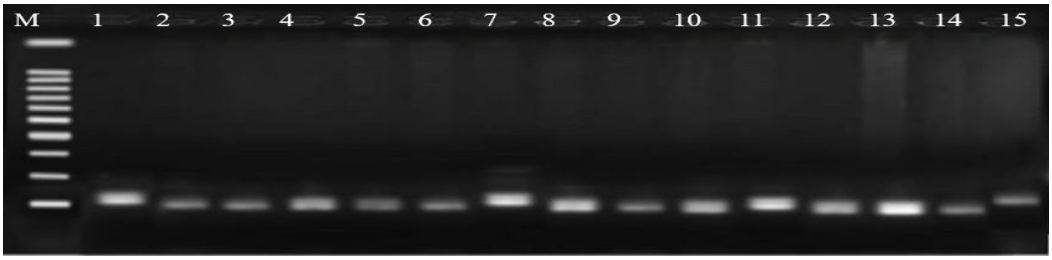

Figure 2: Using the primer NAU3201, the product was amplified from a range of varieties, namely FH-Noor, CIM-602, VH-738, OHO-608, NIAB-824, FD-512, CIM-612, NIAB-Kiran, FH-Lalazar, Sitara-008, RH-662, BH-184, CYTO-179, CIM-473, and BT-121. These amplified products, corresponding to lanes 1 through 15 respectively, were run on a 2% agarose gel. A 100bp ladder, labeled as 'M', was used for reference.

Table 3: Analysis of Major Allele Frequency, Gene Diversity, Heterozygosity, and PIC Using NAU Markers

Sr. No.	Marker	Major Allele Frequency	Gene Diversity	Heterozygosity	PIC
1.	NAU1048	0.5	0.5	0.8	0.375
2.	NAU2869	0.725	0.39875	0.55	0.3192
3.	NAU3201	0.725	0.39875	0.45	0.3192
4.	NAU3306	0.575	0.48875	0.65	0.3692

5.	NAU3598	0.575	0.48875	0.25	0.3694
6.	NAU3897	0.525	0.49875	0.85	0.3744
7.	NAU3920	0.5	0.5	0.7	0.375
8.	NAU3009	0.525	0.49875	0.85	0.3744
9.	NAU4871	0.5	0.5	0.5	0.375
10.	NAU5005	0.675	0.43875	0.45	0.3425
11.	**Mean**	**0.5825**	**0.471125**	**0.605**	**0.3593**

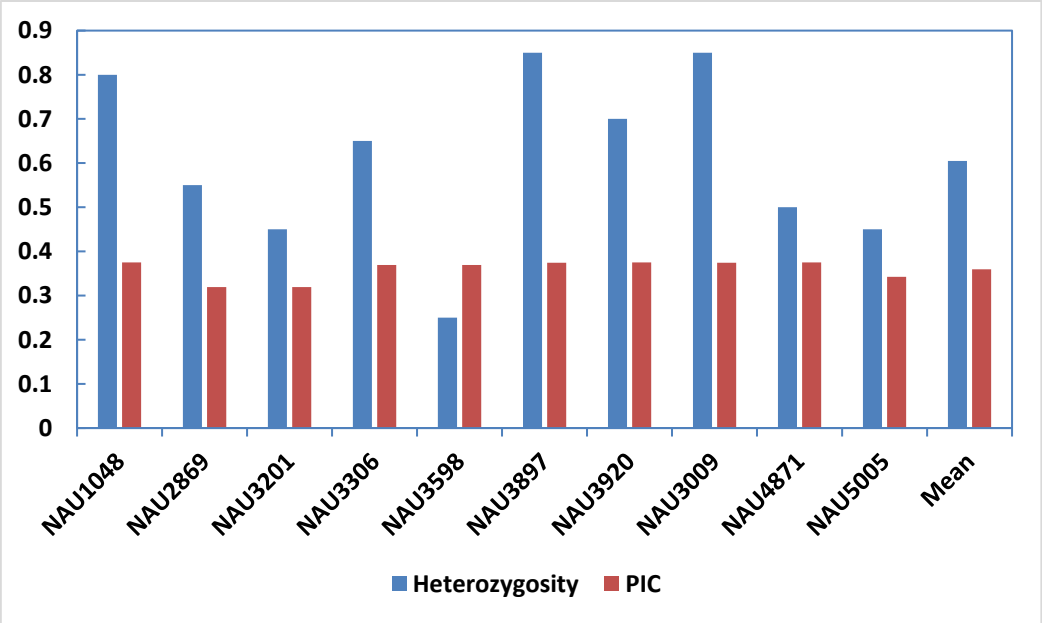

Figure 3: A graphical representation of the PIC and heterozygosity values of EST-SSR markers.

Phylogenetic tree

The genetic similarity analysis conducted on the different varieties led to the creation of a phylogenetic bootstrap neighbor-joining tree, using the Nei 1973 similarity index. The genetic relationships among the 20 genotypes examined in the study were assessed utilising the provided tree diagram. The research found three main groups that represented the genotypes' phylogenetic similarity, demonstrating their genetic relatedness.

The genotypes RH-647 and FH-Lalazar were discovered to be closely linked in cluster one and to have a same genetic ancestor. This clustering shows that RH-647 and FH-Lalazar have a great deal of phylogenetic similarities. NIAB-878, Sitara-008, VH-738, FH-Noor, NIAB-824, CIM-473, NIAB-2008, BH-184, CIM-554, CIM-612, and RH-662 made up Cluster 2. These genotypes were put together in the phylogenetic tree because they showed genetic relatedness.

CIM-602, OHO-608, NIAB-Kiran, BT-121, CYTO-179, and FD-512 made up the third cluster. In the phylogenetic tree, these genotypes formed a separate cluster, demonstrating their shared genetic traits. Overall, the phylogenetic tree showed

various clusters that showed the analysed genotypes' genetic relatedness to one another, with unique groups depending on their ancestral ties.

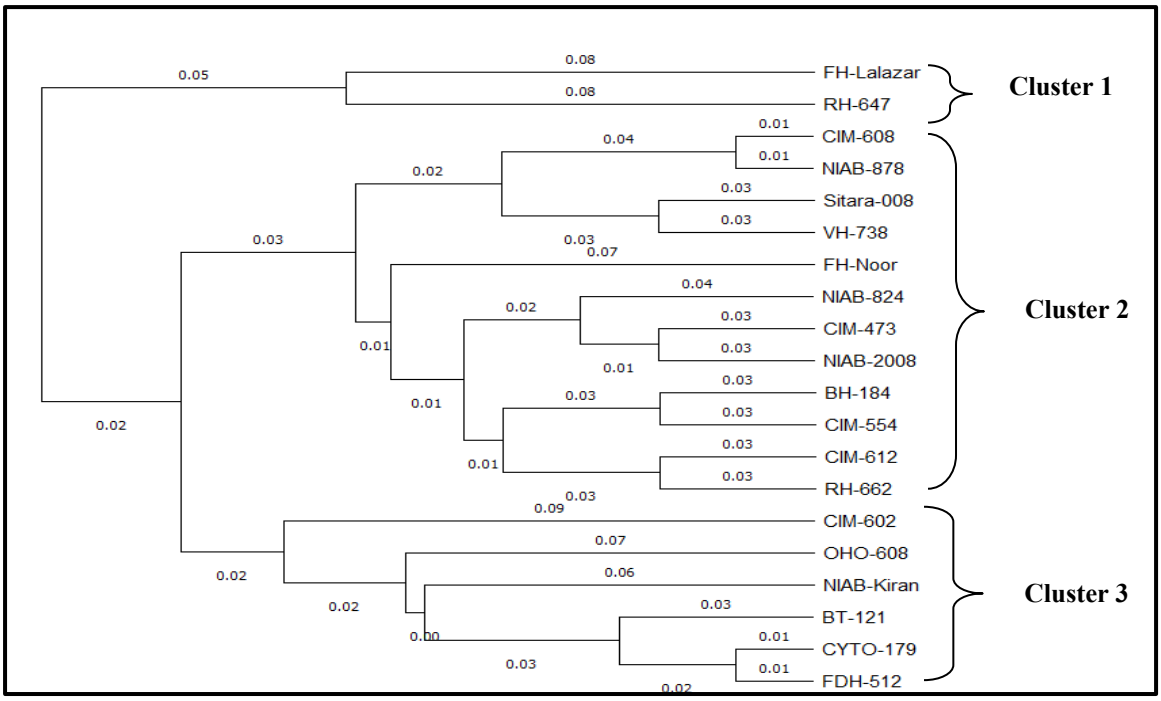

Figure 4: Diversity Analysis of 20 Selected Cotton Varieties Using UPGMA Dendrogram.

DISCUSSION

As a significant revenue commodity, cotton is essential to meeting domestic demand for fibre and oil. It is an essential resource for the textile industry. However, genetic constraints and limited molecular bases pervasive in commonly cultivated varieties impede the genetic improvement of cotton cultivars. Consequently, genetic diversity studies provide a superior alternative to conventional breeding programmes. These studies assist in overcoming limitations and identifying enhancement opportunities for cotton varieties.

It has been observed at the molecular level that many extensively used cotton varieties share a similar genetic background, highlighting the need for genetic diversity research. Singh et al. (2020) highlighted the importance of molecular markers in estimating genetic diversity for plant development and conservation. Among these markers, SSRs are highly preferred due to their high polymorphism, codominant nature, and ability to detect allelic variability across diverse germplasms (Kesawat and

Kumar, 2009; Vieira et al., 2016). The multi-allelic nature of SSRs allows the detection of comparative allelic variability, making them valuable tools for genetic fingerprinting in various plant species (Kirungu et al., 2018).

Assessing the level of polymorphism in EST-SSR markers is commonly achieved using the PIC values (Botstein et al., 1980; Gupta and Varshney, 2000). PIC values reflect the ability of a primer to assess multiple heterozygous alleles. Higher PIC values indicate greater potential for molecular diversity. According to Botstein et al. (1980), SSR markers with PIC values above 0.75 are considered highly informative, while values below 0.25 suggest low or natural information content. In the present study, the analysis of 10 NAU series markers in 20 cotton varieties revealed a low degree of genetic diversity. The PIC values ranged from 0.31924 to 0.3750, with an average of 0.3593. The range of the Diversity genotype index (H') was observed to be between 0.2500 and 0.8500, with a calculated mean value of 0.605. According to Singh et al. (2020), the genetic similarity coefficient (SSR) among the germplasm sources varied between 0.398 and 0.5000, with an average value of 0.471. These findings suggest that there is limited genetic diversity among the germplasm sources examined.

The genetic relatedness of the genotypes was investigated by constructing a phylogenetic bootstrap neighbor-joining tree using the Nei (1973) similarity index. The 20 genotypes were classified into distinct clusters based on their phylogenetic similarities. Cluster one was comprised of two entities, namely RH-647 and FH-Lalazar, which exhibited a notable level of resemblance. Cluster two included CIM-608, NIAB-878, Sitara-008, VH-738, FH-Noor, NIAB-824, CIM-473, NIAB-2008, BH-184, CIM-554, CIM-612, and RH-662, suggesting shared origins and genetic makeup.

The third cluster consisted of CIM-602, OHO-608, NIAB-Kiran, BT-121, CYTO-179, and FD-512 (Singh et al., 2020). The observed clustering pattern aligns with the findings of a study conducted by Aslam et al. (2020), which examined the genetic diversity of different cotton varieties through the utilisation of diverse microsatellites. The study revealed limited diversity among the chosen varieties. FH-Lalazar was observed to be clustered distinctly in both studies, suggesting its unique characteristics and diverse genome (Aslam et al., 2020; Singh et al., 2020).

Bertini et al. (2006) conducted an experiment using SSR markers to assess the genetic diversity of 53 cotton varieties. With an average of 2.1 loci per primer and PIC values ranging from 0.2 to 0.6, the study revealed a relatively low level of diversity among the cotton varieties examined. Using 320 microsatellites, Lacape et al. (2007) analysed the diversity of 47 tetraploid cotton varieties from various regions of the world. The research demonstrated a genetic limitation among cotton cultivars. Lacape et al., 2007; Bertini et al., 2006; Lacape et al., 2007; similar results were observed in the present investigation, indicating limited

genetic diversity among the cotton varieties used in Pakistan.

Other investigations using different molecular markers by Kantartzi et al. (2009), Obaid-ur-Rehman et al. (2009), Kalivas et al. (2011), JIA et al. (2014), Abdellatif et al. (2012), Saleem et al. (2020), and Shoukat et al. (2020) consistently revealed minimal genetic diversity in cotton cultivars. These findings support the idea that widening the genetic basis of cotton germplasm is critical for generating varieties with superior attributes because it introduces new genetic variants and facilitates genetic recombination (Kantartzi et al., 2009; Saleem et al., 2020; Shoukat et al., 2020).

Finally, the genetic variety of cotton types is critical for their advancement. The current research discovered a lack of genetic variety among regularly used cotton cultivars, which impedes genetic advancement. To extend the genetic base and improve the traits and adaptability of cotton cultivars, measures such as adding new sources of genetic diversity, inter-specific hybridization, and mutation breeding should be addressed (Singh et al., 2020). Expanding the genetic basis is critical for guaranteeing cotton cultivation's sustainability and resilience in the face of developing difficulties (Singh et al., 2020).

CONCLUSION

The EST-SSR analysis revealed a lack of genetic diversity among the regularly used cotton cultivars studied in this work, posing obstacles to genetic improvement initiatives. Alternative tactics like as introduction from other sources, inter-specific hybridization, and mutation breeding may be used to add new sources of variety. These techniques seek to extend the genetic base and encourage the development of unique genetic variants that may contribute to cotton genetic improvement.

REFERENCES

Abdellatif, K.F., Y.A. Khidr, Y.M. El-Mansy, M.M. El-Lawendey and Y.A. Soliman. 2012. Molecular diversity of Egyptian cotton (Gossypium barbadense L.) and its relation to varietal development. Journal of Crop Science and Biotechnology. 15:93-99.

Abdurakhmonov, I.Y., A. Abdukarimov, A.E. Pepper, A.A. Abdullaev, F. Kushanov, Z.Y. John, J.N. Jenkins, K. Urmonov, M. Ulloa and R.J. Kohel. 2012. Genetic diversity in Gossypium genus.

Adawy, S.S. 2007. An evaluation of the utility of simple sequence repeat loci (SSR), expressed sequence tags (ESTs) and expressed sequence tag microsatellites (EST-SSR) as molecular markers in cotton. Journal of Applied Sciences Research. 3:1581-1588.

Amanet, K., E.O. Chiamaka, G. Willie, M.M. Quansah, H.U. Farid, R. Akram and W. Nasim. 2019. cotton Production in Africa. Cotton Production.359.

Aslam, S., H. Iqbal, M. Hussain and M. Baber. 2020. Molecular diversity of

Pakistani upland cotton (Gossypium hirsutum L.) varieties based on microsatellite markers. Pakistan Journal of Biochemistry and Biotechnology. 1.

Basal, H., E. Karademir, H.K. Goren, V. Sezener, M.N. Dogan, I. Gencsoylu and O. Erdogan. 2019. Cotton production in Turkey and Europe. Cotton production.297-321.

Bertini, C.H., I. Schuster, T. Sediyama, E.G.D. Barros and M.A. Moreira. 2006. Characterization and genetic diversity analysis of cotton cultivars using microsatellites. Genet. Mol. Biol. 29:321-329.

Botstein, D., R.L. White, M. Skolnick and R.W. Davis. 1980. Construction of a genetic linkage map in man using restriction fragment length polymorphisms. Am. J. Hum. Genet. 32:314.

Chen, Z.J., B.E. Scheffler, E. Dennis, B.A. Triplett, T. Zhang, W. Guo, X. Chen, D.M. Stelly, P.D. Rabinowicz and C.D. Town. 2007. Toward sequencing cotton (Gossypium) genomes. Plant Physiol. 145:1303-1310.

Doyle, J.J. and J.L. Doyle 1987. A rapid DNA isolation procedure for small quantities of fresh leaf tissue.

Guo, W., C. Cai, C. Wang, Z. Han, X. Song, K. Wang, X. Niu, C. Wang, K. Lu and B. Shi. 2007. A microsatellite-based, gene-rich linkage map reveals genome structure, function and evolution in Gossypium. Genetics. 176:527-541.

Gupta, P.K. and R. Varshney. 2000. The development and use of microsatellite markers for genetic analysis and plant breeding with emphasis on bread wheat. Euphytica. 113:163-185.

Han, Z., C. Wang, X. Song, W. Guo, J. Gou, C. Li, X. Chen and T. Zhang. 2006. Characteristics, development and mapping of Gossypium hirsutum derived EST-SSRs in allotetraploid cotton. Theor. Appl. Genet. 112:430-439.

Iqbal, M., O. Reddy, K. El-Zik and A. Pepper. 2001. A genetic bottleneck in the'evolution under domestication'of upland cotton Gossypium hirsutum L. examined using DNA fingerprinting. Theor. Appl. Genet. 103:547-554.

Jia, Y.-H., J.-L. Sun, X.-W. Wang, Z.-L. Zhou, Z.-E. Pan, S.-P. He, B.-Y. Pang, L.-R. Wang and X.-M. Du. 2014. Molecular diversity and association analysis of drought and salt tolerance in Gossypium hirsutum L. germplasm. Journal of Integrative Agriculture. 13:1845-1853.

Kahriz, M.P., P.P. Kahriz and K.M. Khawar. 2019. cotton Production in central Asia. Jabran K, Singh Chauhan B (eds) Cotton Production. Wiley, New York.323-339.

Kalivas, A., F. Xanthopoulos, O. Kehagia and A. Tsaftaris. 2011. Agronomic characterization, genetic diversity and association analysis of cotton cultivars using simple sequence repeat molecular markers. Genet Mol Res. 10:208-217.

Kantartzi, S.K., M. Ulloa, E. Sacks and J.M. Stewart. 2009. Assessing genetic diversity in Gossypium arboreum L. cultivars using genomic and EST-derived microsatellites. Genetica. 136:141-147.

Kesawat, M.S. and B.D. Kumar. 2009. Molecular markers: it's application in crop improvement. Journal of Crop Science and Biotechnology. 12:169-181.

Kirungu, J.N., Y. Deng, X. Cai, R.O. Magwanga, Z. Zhou, X. Wang, Y. Wang, Z. Zhang, K. Wang and F. Liu. 2018. Simple sequence repeat (SSR) genetic linkage map of D genome diploid cotton derived from an interspecific cross between Gossypium davidsonii and Gossypium klotzschianum. International journal of molecular sciences. 19:204.

Lacape, J.-M., D. Dessauw, M. Rajab, J.-L. Noyer and B. Hau. 2007. Microsatellite diversity in tetraploid Gossypium germplasm: assembling a highly informative genotyping set of cotton SSRs. Mol. Breed. 19:45-58.

Liu, S., R. Cantrell, J. Mccarty Jr and J.M. Stewart. 2000. Simple sequence repeat–based assessment of genetic diversity in cotton race stock accessions. Crop Sci. 40:1459-1469.

Obaid-Ur-Rehman, S.H.K. and M. Sajjad. 2009. Genetic Diversity Among Cotton (Gossypium hirsutum L.) Genotypes Existing in Pakistan. American-Eurasian Journal of Sustainable Agriculture. 3:816-823.

Park, Y.-H., M.S. Alabady, M. Ulloa, B. Sickler, T.A. Wilkins, J. Yu, D.M. Stelly, R.J. Kohel, O.M. El-Shihy and R.G. Cantrell. 2005. Genetic mapping of new cotton fiber loci using EST-derived microsatellites in an interspecific recombinant inbred line cotton population. Mol. Genet. Genomics. 274:428-441.

Raper, T.B., C. Pilon, V. Singh, J. Snider, S. Stewart and S. Byrd. 2019. Cotton production in the United States of America: An overview. Cotton production.217.

Saha, S., M. Karaca, J.N. Jenkins, A.E. Zipf, O.U.K. Reddy and R.V. Kantety. 2003. Simple sequence repeats as useful resources to study transcribed genes of cotton. Euphytica. 130:355-364.

Saleem, M.A., M.W. Amjid, M.Q. Ahmad, H. Riaz, S.F. Arshad and Z.U. Zia. 2020. EST-SSR based analysis revealed narrow genetic base of in-use cotton varieties of pakistan. Pak. J. Bot. 52:1667-1672.

Shoukat, S., M. Anwar, S. Ali, S. Begum, M. Aqeel, M. Zia, S. Shah and G. Ali. 2020. Estimation of morphological and molecular diversity of seventy-two advanced pakistani cotton genotypes using simple sequence repeats. Appl. Ecol. Environ. Res. 18:2043-2056.

Singh, R.B., M.D. Mahenderakar, A.K. Jugran, R.K. Singh and R.K. Srivastava. 2020. Assessing genetic diversity and population structure of sugarcane cultivars, progenitor species and genera using microsatellite (SSR) markers. Gene. 753:144800.

Tanksley, S.D. and S.R. Mccouch. 1997. Seed banks and molecular maps: unlocking genetic potential from the wild. Science. 277:1063-1066.

Vieira, M.L.C., L. Santini, A.L. Diniz and C.D.F. Munhoz. 2016. Microsatellite markers: what they mean and why they are so useful. Genet. Mol. Biol. 39:312-328.

Wang, X., T.A. Rinehart, P.A. Wadl, J.M. Spiers, D. Hadziabdic, M.T. Windham and R.N. Trigiano. 2009. A new electrophoresis technique to separate microsatellite alleles. African Journal of Biotechnology. 8.

Xiao, J., K. Wu, D.D. Fang, D.M. Stelly, J. Yu and R.G. Cantrell. 2009. New SSR markers for use in cotton (Gossypium spp.) improvement.

Targeting gene regulatory sequences for inhibition of cell-to-cell movement of plant DNA virus

Maria Siddique [1] , Sarmad Frogh Arshad [2] , Hafaz Nazar Faried [3] and Zulqurnain Khan [1]

1 Institute of Plant Breeding and Biotechnology, MNS-University of Agriculture-Multan, Pakistan

2 Faculty of Veterinary and Animal Science, MNS-University of Agriculture-Multan, Pakistan

3 Department of Horticultural Sciences, MNS-University of Agriculture-Multan, Pakistan

Corresponding author: zulqurnain.khan@mnsuam.edu.pk

Abstract

Plant DNA viruses are the most devastating group of viruses that are responsible to cause the gigantic loss worldwide. Geminiviridae is known to be the largest class of DNA viruses which are further classified into the seven other groups (genera) of viruses. These viruses have a wide host range and are responsible to cause a significant damage to sustainable agriculture sector. Begomoviruses are a class of ssDNA viruses that may cause a huge loss to Pakistan's economy. Cotton leaf curl viral disease (CLCuD) is reported the biggest threat to agricultural sector because it can decrease the yield up to 40-50%. In past, many conventional and non-conventional approaches were used to control the plant DNA viruses, but they were unable to control the epidemic cause by these viruses. Some silencing approaches, including RNAi and RSS (RNA silencing suppressors), have been used for developing CLCuD-resistant plant varieties by suppressing or inhibiting the activity of viral genes. Genome editing techniques have been used to get rid of these viruses by developing resistance in plants. CRISPR/Cas9 is the most efficient tool that has been used for developing the virus resistant plants. The objective of the study was using CRISPR/Cas9 for the inhibition of cell-to-cell movement of plant DNA viruses. CP, IR and Rep regions are highly targeted regions that can be used for inhibition of the movement of begomoviruses. These regions are responsible for the replication and movement of viruses and also these sites are highly conserved so, targeting these regions may give a favourable result. The DNA sequence of the plant virus was retrieved from the NCBI to select the target site. The gRNA was designed by using the CRISPRdirect online tool. Colony PCR was used for the confirmation of the clones. *Agrobacterium*-mediated transformation method was used for the co-infiltration and transformation of viruses into plants. The confirmation for the inhibition of movement of virus was determined by using the PCR and RT-qPCR. The results of the study were promising for suppression of viral movement in the plants suggesting it a potential strategy for developing virus resistance in plants. Translation of this study will result as an effective approach for the development of transgenic plants against DNA viruses.

Key words: ssDNA viruses, coat protein, movement inhibition

1. Introduction

Plant viruses are responsible to cause the gigantic loss all around the world because viruses are the major threat to sustainable agricultural sector (Islam *et al.*, 2018a). Among them, the Geminiviridae (Genus Begomovirus) having 320 or more species (Zerbini *et al.*, 2017) which are known to be the most destructive and diverse group of viruses (Islam and Wu, 2017). Begomovirus is responsible to cause huge loss in the Pakistan economy i.e., recently Cotton leaf curl viral disease (CLCuD) is reported as a biggest threat due to its multiple epidemic (Sattar *et al.*, 2013). Hence, to mitigate these losses breeders are trying to develop the virus resistant plants and protect the food crops by developing the improved resistance plant varieties (Islam *et al.*, 2018b). The mechanism of virus replication and transmission can give multiple target sites which are used for developing the resistance in food crops but the rapid evolution found in virus genome make it difficult. Plant viruses use different type of mechanism for their replication and transmission, some transmitted from one plant to another with the help of an intermediate vector (Hogenhout *et al.*, 2008) and some feed on the host plant and transmit virus from infected plant to healthy plant. The major hosts that are responsible for the virus transmission are whiteflies, nematodes, grass hoppers, beetles and few other insects (Whitfield *et al.*, 2015). Geminiviridae is known as the largest family of ssDNA viruses that may infect the variety of plants and its genome is ~2.5–3.1kb in length. These viruses are categorized by their encapsidate circular, quasi-icosahedral joined at missing vertex and the capsid range in size is almost 18×30 nm (Stanley, 1985). On the bases of their genomic organization and host range, the geminiviruses are further classified into the seven genera such as Eragrovirus, Topocuvirus, Turncurtovirus, Curtovirus, Becurtovirus, and Begomovirus (Varsani *et al.*, 2014). They all affect the crop plants but the major contribution is from the begomoviruses because the begomoviruses are infecting the large variety of economically important crop plants (Zaidi *et al.*, 2016). Begomoviruses having a monopartite or bipartite genome, while in case of monopartite they having a single component in their genome but in in case of bipartite they have the two components DNA-A and DNA-B in their genome. The large diversity of viruses have the bipartite genome (DNA-A or DNA-B), each of which is about the 2.5 to 2.8kb in size (Briddon *et al.*, 2010). In case of bipartite genome both components (DNA-a or DNA-B) are playing almost an equal role in infecting the host but DNA-B is totally dependent on DNA-A for replication in a host genome (Hou and Gilbertson, 1996).

Viruses having bipartite genome, DNA-A plays role in replication of genome inside the nucleus by two step procedures by conversion of single stranded DNA (ssDNA) to double stranded DNA (dsDNA) through rolling circle duplication (Gutierrez and pathology, 2002) and encapsidation process while the DNA-B is responsible for the symptoms development in the host. Both types of genome either monopartite and bipartite having same size, excepting a common region of short sequence

TAATATTAC found in intergenic region (Arguello-Astorga *et al.*, 2004). This is stem loop structure having inverted repeats and a polynucleotide loop of 6-12 bp which is responsible for the binding of replication initiator protein during replication. There's need to understand the movement of viruses from inoculated to non-inoculated leaves to understand the mechanism of infection. Viruses are transported by two ways one is the shielding of genome with a shell produced by the assembly of capsid protein (CP) subunits, and other one is through ribonucleoprotein (RNP) complexes, in which the viral genome is linked with viral and cellular proteins. The CP protein plays a main role in long-distance transport of different viral species. Yet, little is known about viral transport mechanisms in the vascular system because of difficulty of thoroughly understanding of study of deeply buried tissue. Furthermore, depending on the host, collecting phloem sap to identify viral phloem partners may be difficult or impossible. The efficiency of cell-to-cell mobility influences viral long-distance transmission, these two linked processes can be difficult to discern at times. As a result, identifying viral and host variables that are specifically necessary for virus long-distance transmission is sometimes misunderstood and remains a difficulty. To acquire control of these deadly viruses, the researchers utilized a variety of techniques.

Now the biotechnological tools for genome editing along with advance breeding techniques such as sequence-specific nucleases (SSNs) such as mega-nucleases (MNs), zinc finger nucleases (ZFNs), Transcription Activator-like Effector Nucleases (TALENs), and clustered regularly interspaced short palindrome repeats (CRISPR)/CRISPR-associated protein 9 (Cas9) has helped to achieve the targeted modification of genome of many crops. All these genomes editing techniques have several advantages over the conventional technique (Das *et al.*, 2019). ZFNs and (TALENs) were the first-generation tools of genome editing technology that has been used for targeted editing of plant genomes (Boch *et al.*, 2009; Kim *et al.*, 1996). All of these genome editing tools have a common sequence-specific nucleases (SSNs), often known as "designer nucleases." They produce double stranded breaks (DSBs), hence cause targeted genome modification, which is later repaired by non-homologous end joining (NHEJ) or homology-directed repair (HDR) and this repair mechanism leads to an insertion or deletion (indel) of few nucleotides at the double stranded break (Wright *et al.*, 2016). Both of the ZFNs and TALENs are made up of chimeric proteins.

The CRISPR/Cas system is an adaptive immune system that can originate from the prokaryotes and used to protect the organism from the foreign DNA invasion e.g., RNA-guided DNA nuclease are used to cut the nucleic acid of viruses in a sequence specific manner. Recent studies indicate that the CRISPR/Cas system have successfully used to develop the geminivirus resistance in plants including Arabidopsis and tobacco. For developing the resistance CRISPR/Cas9 was used to target the genomic DNA of virus at the replication stage and exploit the viral genome. Single guided RNA (sgRNA) is designed that is used to target the Rep, CP

and IR region of the viral genome and significantly reduced the disease symptoms of ssDNA viruses (geminivirus) (Mehta *et al.*, 2018). In case of dsDNA viruses CRISPR/Cas9 has also been used to develop the resistance e.g., Cas9 machinery is used in Arabidopsis to inhibit the virulence of Cauliflower mosaic virus (CaMV) and African cassava mosaic virus (ACMV). Recent studies indicate that the use of CRISPR/Cas in plants have been responsible to develop the novel mutation in viruses that can't be controlled by Cas9 and due to this conserved mutation in viruses the transgenic plants have been failed to develop the resistance against geminiviruses. To avoid the induction of resistant viruses the researches need to focus on the application of CRISPR/Cas9 while developing the virus resistance plants (Mehta *et al.*, 2019). CRISPR/Cas modification system have been used to activate the targeted nuclease activity of the Cas associated proteins that further used to resist virus infection. Different variants of Cas14 protein have been used in comparative sequence analysis which have been used to cleave the genome of ssDNA viruses and develop the resistant plants. Cas14a have been tightly linked the IR, CP and Rep domain of the Single stranded viruses and develop modified plants. Despite these challenges CRISPR/Cas have been successfully used to target the highly conserved region of plant viruses which have been used to inhibit the movement of these viruses by developing the transgene plants (Khan *et al.*, 2019). Although, few other varients of CRISPR has also been used successfully for developing resistance against viruses in plants by inhibiting their replicating ability. Whereas, all the above mentioned technologies are used for increasing the productivity of economically important crop plants.

Table 1.1. Some successful examples of controlling viruses via genome editing tools

Genome targeting platform	**Virus**	**Genus**	**DNA Target**
CRISPR/Cas9	Merremia-mosaic virus, Cotton leaf curl Kokhran-virus, TYLCV, BCTV	Begomovirus	Viral DNA
TALE	TYLCCNV, TbCSV, Tomato leaf curl Yunnan virus	Geminiviridae	Viral DNA
ZFN	TYLCCNV, TbCSV	Geminiviridae	Viral DNA
AZP	BSCTV, TYLCV	Geminiviridae	Viral DNA

2. Material and Methods

2.1. Selecting the target sequence and designing the gRNA

The DNA sequence of different plant DNA viruses such Cotton Leaf Curl Khokhran Virus (CLCuKV) was obtained from the NCBI. CRISPR-direct software was used for identifying the potential target site of respective virus. The common region of the DNA sequence of respective virus gene was used for recognizing the potential target site for dCas9. Select the target sequence with respect to PAM sequence, design the forward, and reverse primer according to selected sequence (Khan *et al*., 2019).

2.2. Cloning of dCas9 expression vector

The cloning of gRNA was done in three different steps:

2.2.1. Primer hybridization

For primer hybridization, 25 µL to 30 µL of primer was mixed in a PCR tube (200 µL). Master mix was added in the 200 µL PCR tube and placed the tube inside the thermocycler. Started the thermocycler, under following conditions the reaction was performed.

Table 2.1. PCR reaction Profile

Denaturation	**Annealing**	**Extension**	**Time**
95 °C (initial)	70°C	40°C	3 min
90°C	60°C	30°C	1 min
80°C	50°C	20°C	1 min
	(30-35)	**Cycles**	
Hold		4°C	∞

2.2.2. Restriction digestion of Cas9 backbone vector

Bsal and the gel DNA was used for restriction digestion of backbone vector by using the thermo scientific GeneJET extraction kit. Moreover, T4 ligase was used for the ligation reaction and the reaction was run at 22 ºC for 240 mints.

2.2.3. Ligation into the backbone vector

E. coli DH5α competent cells were used for the transformation purpose. The ligation products were transformed into these *E. coli* competent cells. Confirmation of cloning of Cas9 vector the colony PCR was performed. The colony PCR was conducted by using the specific primers:

U6-26-F:
5'TGTCCCAGGATTAGAATGATTAGGC3'

dT4-R:
5'AAACGTAATATTAAACGGATGGCC3'

Miniprep was performed for isolating PCR confirmed clones Plasmid DNA through the GneJET kit (Thermo Scientific kit). Before transformation into *Agrobacterium* (GV3101) the Plasmid DNA were sequenced.

2.3. Restriction digestion and colony PCR-based confirmation

Blue and white bacterial colonies were screened out on Xgal/IPTG. After picking the white colonies further PCR was performed for conformation purpose. Moreover, after picking the white colonies, mixed these colonies in 20µL reaction mixture. While following to the supplier's guidelines the restriction endonuclease *BsaI* and their respective buffer was used for the plasmid digestion and PCR product accomplishment.

2.4. Infectious colonies and their growth

After reviving the culture, picked up the single colony on media plates. The media plates had the culture that contained the agrobacterium mediated desired construct. Take 50ml autoclaved falcon tubes, add 40ml of LB media in it and then take a single colony of desired construct and add it in LB media contains falcon tube. After that add the 40µL Rifampicin and Kanamycin (antibiotics) into the falcon tube and tightly covered the falcon tube led or seal it properly with parafilm. Then placed that primary culture in a shaker at 28 ºC for 2 days in dark. After the 2 days in dark, the culture was grown and successfully used for the transformation.

2.5. Optimum growth condition for plant

The seeds of *N. benthamian* will be taken as a plant material. The seeds were spread on petri plates containing sand and placed these plates in the growth room for 10-12 days. Another process was also used for the initial growth of the tobacco plants. In this process, the *N. benthamiana* plants were firstly grow on the MS media contain petri plates. Ten seeds were sow on per petri plate seal the petri plates with parafilm and place the plates in a growth room under proper growth condition. This experiment was conducted in completely randomized

design (CRD) in triplicate. After 7 to 10 days, nursery was growing and ready to transfer it into the pots. Later, for maintaining the triplicates the plants were shifted as 10 pots per 10 plants containing peat moss.

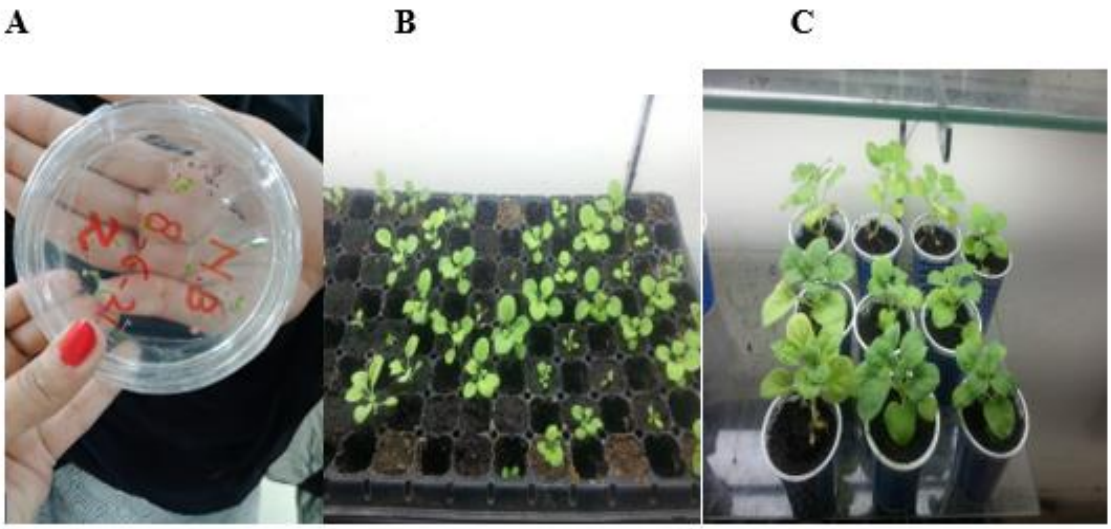

Figure 2.1. Diagram shows the development of nursery. A. showing the growing of *N. benthamiana* plants on media under control conditions. B. shows after 15 to 20 days plants are shifted into trays. C. shows plants which are ready for infiltration.

2.6. Disease resistance assay

In *N. benthamiana,* through the process of Agroinfiltration the infectious clones' infectivity was checked. *Agrobacterium* containing the viruses was used to infiltrate the two weeks old plants. Inoculation culture of *Agrobacterium* was maintained to an optical density of 0.6 at 600 nm after incubating at 28°C for 48 hours with Rifampicin (25µg/ml) and Kanamycin (50µg/ml). *Agrobacterium* cells was spun for15 min at 20°C and resuspend in10mM $MgCl_2$ containing acetosyringone 150ug/ml. Infiltration medium was left on bench for 3 hours. Infiltration was done into the fully expanded young leaves. After syringe infiltration, the plants were kept under suitable growth condition. After 9 to 15 days plants the symptoms were observed in the infiltrated plants. Khan *et al.*, 2019 clearly mentioned that after 14 to 17 days leaf samples were taken of post inoculation (dpi) for the extraction of DNA and then PCR/qPCR were performed. For DNA extraction, plant sample was extracted from the diseased plants or the plants that having viral symptoms after agroinfiltration.

2.7. Accumulation of virus

For checking the virus accumulation, the leaf sample was collected after 14 to 20 days of the post inoculation for the extraction of DNA to conduct the PCR/qPCR.

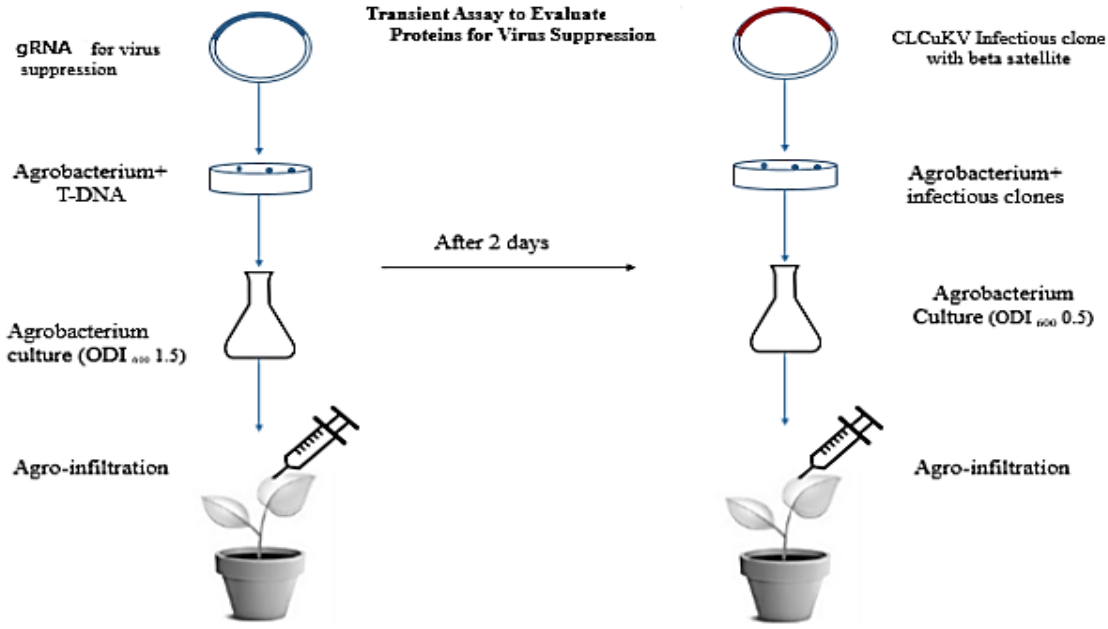

Figure 2.2. Strategy of Agroinfiltration in *N. Benthamian* (Khan, 2017)

2.8. Plasmid Restriction digestion

Initially, buffer, plasmid with enzymes and ddH_2O was added in an Eppendorf tube. After mixing all the reagents in the tube placed the tube into the water at 37°C for 20 minutes. Digestion was completed and then prepared 1% agarose gel for running the sample on the gel.

Table 2.4. Reaction of restriction digestion

Chemical reagents	Quantity
ddH_2O	**15 µl**
Plasmid	**2 µl**
Buffer	**2 µl**
BamH1	**1 µl = 20 µl (total)**

3. Results

3.1. Vector designing

pHSE401 Cas9 expression vector was used for the cloning of gRNA by confirming

through colony PCR. Cloning of gRNA into the Cas9 expression vector was confirmed. The vector was cloned with promoter and terminator sequences.

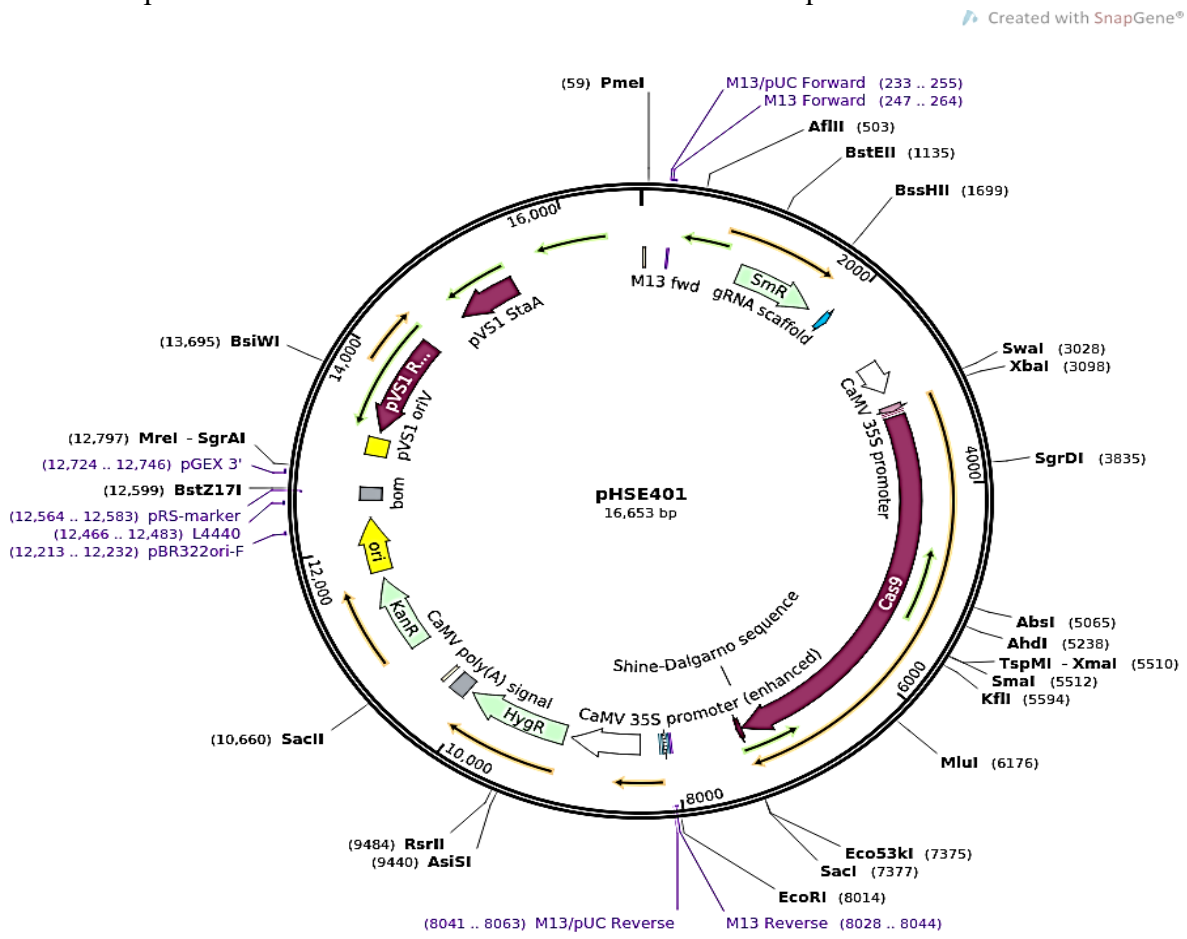

Figure 3.1. Map of the pHSE401 plant expression vector with gRNA sequence by Vector Builders (https://en.vectorbuilder.com).

3.2. gRNA cloning

By using the designed oligos, gRNA was cloned inside the dCAS9 plant expression vector pHSE401. Cloning of gRNA was confirmed through colony PCR into the expression vector by using the same primers. DNA amplification of about 280bp stands was shown by positive clones. For checking the orientation of the cloned guide RNA inside the vector by using different primer plasmid was sequenced. Sequencing results shown below in fig 4.3.

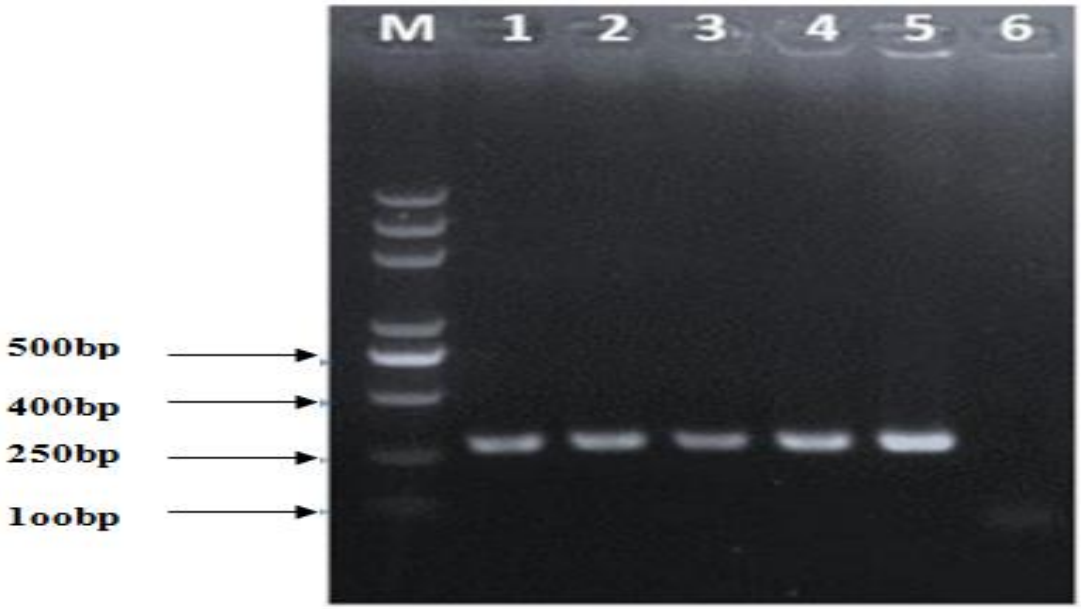

Figure 3.2. gRNA clones' confirmation through colony PCR. For confirming he 20nt gRNA sequence in Cas9 vector the process of electrophoresis have been performed. In 2% agarose gel the amplified product have been run. Negative clones are present in 6 columns while the positive clones (required) are shown in line 1 to 5.

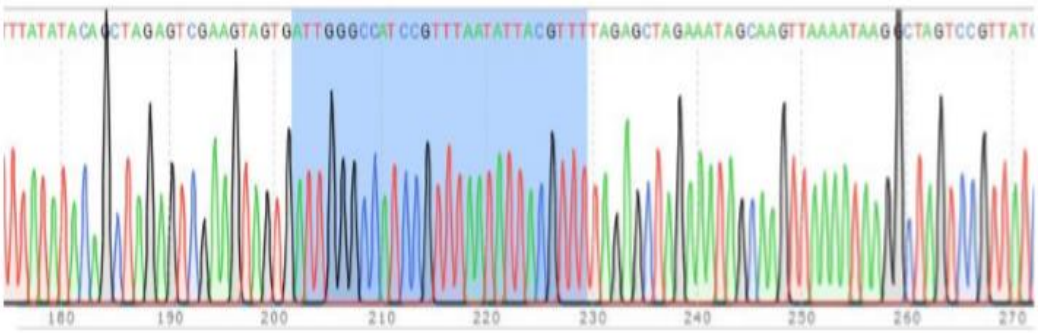

Figure 3.3. The sequencing results of gRNA by using CLUSTRLW.

3.3. Plasmid restriction digestion

BamH1 was used for the restriction digestion of plasmid and after that miniprep was performed for checking the purity of plasmid. Digestion was done successfully through this enzyme and 2000bp band was observed.

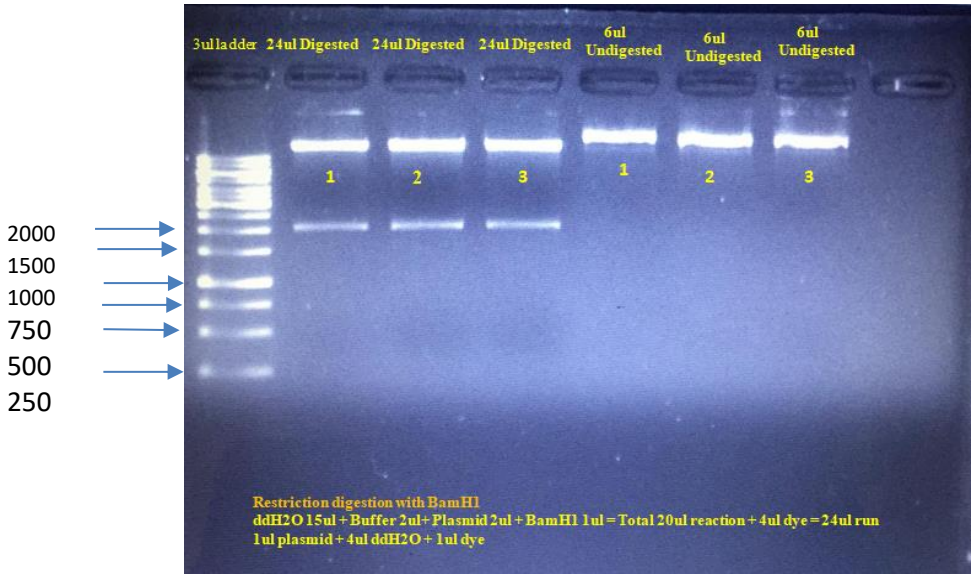

Figure 3.4. Results shows the restriction digestion of plasmid with BamH1. The digested band size is 2000bp.

3.4. Virus infectivity assay in *N. benthamiana*

Symptoms were observed in the plants after the infiltration with CLCuKoV infectious clones. Disease symptoms were appeared in those control plants that were infiltrated with the Cas9 gRNA. The most commonly observed symptoms are vein thickening, upward and downward curling of leaves and stunted growth. Infectious clones infiltrated leaves were taken at 14 dpi to detect the presence of virus through PCR and compared with symptom development.

Figure 3.5. Show the virus infectivity assay of CLCuKV infectious clones in tobacco. At 14dpi the symptoms have been develop. B: showing control plants (no infiltration) A or C. infiltrated plants (infectious clones) D. Diseased plant (inward/outward curling).

3.5. Virus accumulation through qPCR

The attenuated and delayed symptoms of virus infectious clones have been shown after the infiltration of the plants with Cas9 gRNA. By performing qPCR, it was confirmed. Under the transient expression of Cas9 the virus accumulation was determined in *N. benthamiana*. Due to Cas9 the disease symptoms (stunted grow, curling of leaves, thickening of veins) was appeared. The symptoms might be mild or delayed due to the low virus titer (50-80% decreased). Moreover, as compared with the control (no infiltration) the reduction of virus titer in infiltrated leaves (Cas9 construct) was observed.

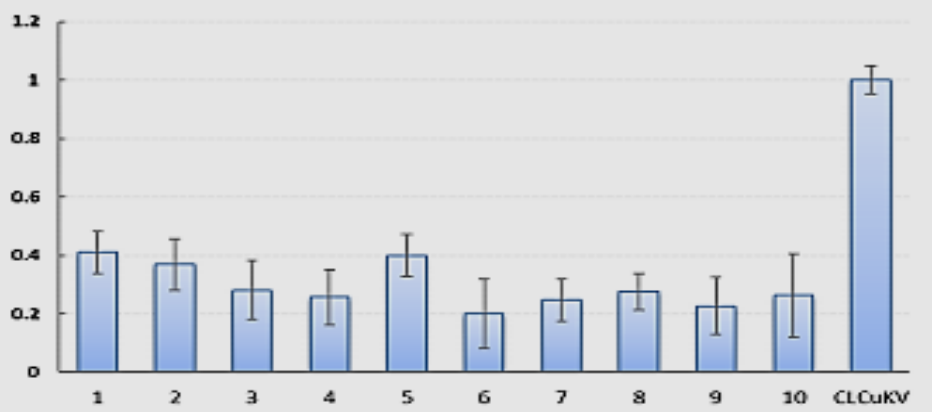

Figure 3.6. Outcomes of virus accumulation through qPCR. As a reference CLCuKV is used. The height of graph bar shows the accumulation of virus in plant sample. While comparing with the reference the virus accumulation is minimum from 1 to 10 in sample 0.2 to 0.4.

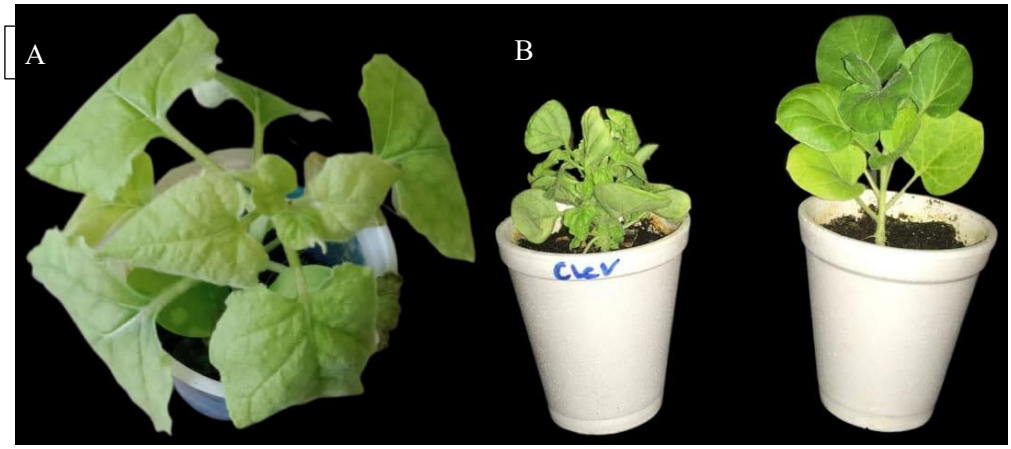

Figure 3.7. Virus infectivity assay. A and B shows the virus infectivity assay of CLCuKV infectious clones in tobacco

3.6. Delay in symptoms development

Within 10 to 15 days the virus symptom started appearing with equivalent frequency. At 12 to 14dpi the co-infiltrated plants (CLCuKV) able to show the symptoms. Furthermore, as compared to the upper leaves the lower leaves can show the mild or delayed symptoms. Instaed of control plants the co-infiltred plants start recovering from infection at the level of maturity.

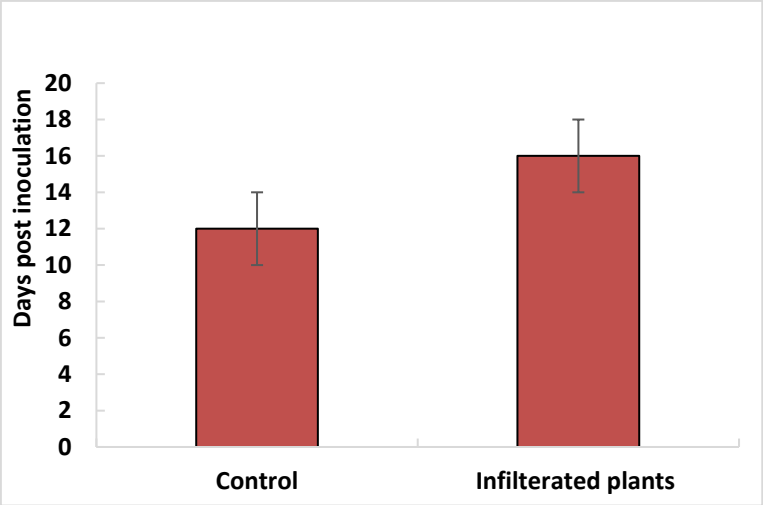

Figure 3.8. Delay in symptoms development. Control plants were infiltrated with CLCuKV. dCas9 plants were co-infiltrated with CLCuKV and dCas9

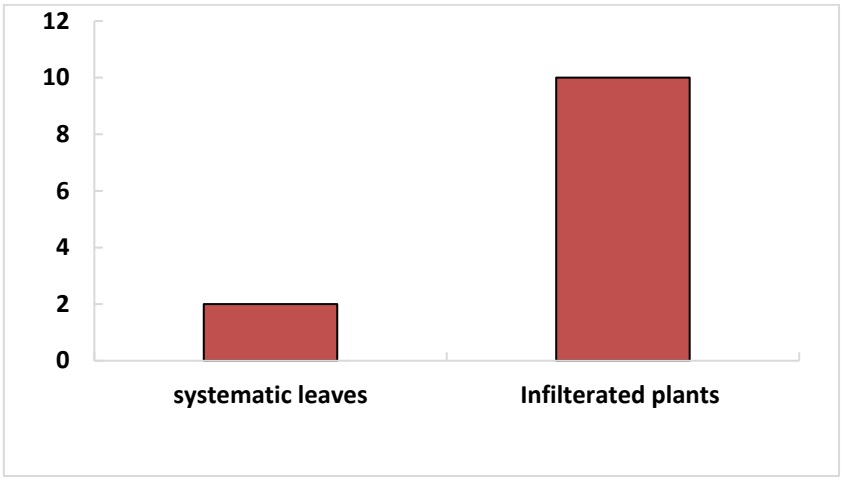

Figure 3.9. qPCR graph results confirming presence and accumulation of CLCuKV in upper young leaves**.**

Summary

The plant DNA viruses are the major threating factor for the agricultural sector in all around the world. Begomoviruses are one of the biggest class of ssDNA viruses, having the wide host rang from simpler to complex organisms. CLCuD is reported to be most devastating group of viruses that can cause the huge loss to Pakistan economy. Conventional approaches are not as much efficient so the advance genome editing technologies are used for the development of resistant varieties. For inhibiting the movement of ssDNA viruses in plants there is need to stop the replication of these viruses inside the plant cell by targeting the IR, CP and Rep region of begomoviruses. The genome of begomoviruses have many overlapping regions, just target one the over lapping region by the gRNA and inhibit its movement. After retrieving the DNA sequence of these viruses from NCBI, targeted region was selected and gRNA was designed by using the CRISPR-direct tool. Moreover, for the confirmation of these clones' colony PCR was performed and then the *Agrobacterium*-mediated transformation was done by following the process of co-infiltration. When the construct was successfully transferred inside the plant the confirmation for the inhibition of virus movement have done through the qPCR. So, we concluded that this study will be used for developing the resistance against the ssDNA/dsDNA viruses and prevent the huge crop loss.

Reference

Islam, W., M. Adnan, M. Tayyab, M. Hussain and S.U. Islam. 2018a. Phyto-metabolites; an impregnable shield against plant viruses. Nat. Prod. Commun. 13(1):19345-19363.

Zerbini, F.M., R.W. Briddon, A. Idris, D.P. Martin, E. Moriones, J. Navas-Castillo, R. Rivera-Bustamante, P. Roumagnac, A. Varsani and

Consortium. 2017. ICTV virus taxonomy profile: Geminiviridae. J. Gen. Virol. 98(2):131-146.

Sattar, M.N., A. Kvarnheden, M. Saeed and R.W. Briddon. 2013. Cotton leaf curl disease–an emerging threat to cotton production worldwide. J. Gen. Virol. 94(4):695-710.

Islam, W., W. Lin, S.U. Islam, M. Arif, X. Li, Y. Yang, X. Ding, Z. Du and Z. Wu. 2018b. Genetic diversity of begomoviruses in Pakistan captured through a vector based survey. Microb. Pathog. 118(56):91-97.

Hogenhout, S.A., E.-D. Ammar, A.E. Whitfield and M.G. Redinbaugh. 2008. Insect vector interactions with persistently transmitted viruses. J. Annu. Rev. Phytopathol. 46(4):327-359.

Whitfield, A.E., B.W. Falk and D. Rotenberg. 2015. Insect vector-mediated transmission of plant viruses. Virology. 479(27):278-289.

Stanley, J. 1985. The molecular biology of geminiviruses. Adv. Virus Res. 30(3):139-177.

Varsani, A., J. Navas-Castillo, E. Moriones, C. Hernández-Zepeda, A. Idris, J.K. Brown, F.M. Zerbini and D.P. Martin. 2014. Establishment of three new genera in the family Geminiviridae: Becurtovirus, Eragrovirus and Turncurtovirus. Arch. Virol. 159(8):2193-2203.

Zaidi, S.S.-E.-A., I. Amin, Z. Iqbal, K. Pervaiz Akhtar, B.E. Scheffler and S. Mansoor. 2016a. Sesbania bispinosa, a new host of a begomovirus-betasatellite complex in Pakistan. Can. J. Plant Pathol. 38(1):107-111.

Briddon, R.W., B.L. Patil, B. Bagewadi, M.S. Nawaz-Ul-Rehman and C.M. Fauquet. 2010. Distinct evolutionary histories of the DNA-A and DNA-B components of bipartite begomoviruses. BMC Evol. Biol. 10(1):1-17.

Hou, Y.-M. and R.L. Gilbertson. 1996. Increased pathogenicity in a pseudorecombinant bipartite geminivirus correlates with intermolecular recombination. J. Virol. 70(8):5430-5436.

Gutierrez, C. 2002. Strategies for geminivirus DNA replication and cell cycle interference. Physiol. Mol. Plant Pathol. 60(5):219-230.

Arguello-Astorga, G., L. Lopez-Ochoa, L.-J. Kong, B.M. Orozco, S.B. Settlage and L. Hanley-Bowdoin. 2004. A novel motif in geminivirus replication proteins interacts with the plant retinoblastoma-related protein. J. Virol. 78(9):4817-4826.

Das, A., N. Sharma and M. Prasad. 2019. CRISPR/Cas9: a novel weapon in the arsenal to combat plant diseases. Front. Plant Sci. 9(2):2008-2022.

Boch, J., H. Scholze, S. Schornack, A. Landgraf, S. Hahn, S. Kay, T. Lahaye, A. Nickstadt and U. Bonas. 2009. Breaking the code of DNA binding specificity of TAL-type III effectors. Sci. 326(5959):1509-1512.

Kim, Y.-G., J. Cha and S. Chandrasegaran. 1996. Hybrid restriction enzymes: zinc finger fusions to Fok I cleavage

domain. Proc. Natl. Acad. Sci. 93(3):1156-1160.

Wright, A.V., J.K. Nunez and J.A. Doudna. 2016. Biology and applications of CRISPR systems: harnessing nature's toolbox for genome engineering. Cell. 164(1-2):29-44.

Mehta, D., A. Stürchler, R.B. Anjanappa, S.S.-E.-A. Zaidi, M. Hirsch-Hoffmann, W. Gruissem and H. Vanderschuren. 2019. Linking CRISPR-Cas9 interference in cassava to the evolution of editing-resistant geminiviruses. Genome Biol. 20(1):1-10.

Mehta, D., A. Stürchler, M. Hirsch-Hoffmann, W. Gruissem and H. Vanderschuren. 2018. CRISPR-Cas9 interference in cassava linked to the evolution of editing-resistant geminiviruses. bioRxiv. 24(11):314542-314558.

Khan, Z., S.H. Khan, A. Ahmad, S. Aslam, M.S. Mubarik and S. Khan. 2019b. CRISPR/dCas9-mediated inhibition of replication of begomoviruses. Int. J. Agric. Biol. 21(4):711-718.

Khan, M.Z., S. Haider, S. Mansoor and I. Amin. 2019a. Targeting plant ssDNA viruses with engineered miniature CRISPR-Cas14a. Trends Biotechnol. 37(8):800-804.

Khan, Z. 2017. Use Of Transcription activator like effectors to suppress cotton leaf curl virus. Ph.D. diss., Univ. Agric., Faisalabad, Pakistan.

Permanent Suppression of *SPY* Gene via CRISPR/Cas9 Technology for Enhanced Drought Tolerance in Cotton

Saba Yaseen and Zulqurnain Khan*

[1]*Institute of Plant Breeding and Biotechnology (IPBB), MNS University of Agriculture, Old Shujaabad Road, Multan, Pakistan.*

*Correspondence: zulqurnain.khan@mnsuam.edu.pk

Abstract

Climate change exerts significant pressures on agricultural productivity, threatening global food security with increasingly severe weather conditions. These stresses, particularly drought, adversely affect crop yield and quality. This study explores a two-fold strategy for developing drought-resistant cotton, starting with transient gene silencing using Virus-Induced Gene Silencing (VIGS) and following with permanent gene editing through CRISPR/Cas9. We initially targeted the Spindly (SPY) gene, known to negatively regulate drought stress response. The SPY homolog in cotton was successfully silenced using VIGS, leading to a measured reduction in gene expression from 7-14 days post infiltration, and an observed improvement in drought tolerance. Upon validating the role of the SPY gene, CRISPR/Cas9 technology was employed for stable, long-term gene silencing. By designing guide RNAs (gRNAs) to direct the Cas9 nuclease to the SPY gene, we induced targeted double-strand breaks, leading to error-prone repair and subsequent gene knockout. The successful permanent suppression of the SPY gene conferred improved drought tolerance in the progeny of transformed plants, showing promise for the development of climate-resilient cotton cultivars. While our study illuminates the potential of combining VIGS and CRISPR/Cas9 in crop improvement, it also underscores the necessity of comprehensive risk assessment and regulatory oversight in genetic engineering.

Keywords: CRISPR/Cas9, Abiotic stress, climate resilient breeding, Food security, functional genomics, VIGS

Introduction and Background

Researchers have resorted to the synergistic combination of functional genomics, Virus-Induced Gene Silencing (VIGS), and CRISPR/Cas9 technologies to improve drought tolerance in cotton. These cutting-edge techniques give significant tools for deciphering the molecular pathways behind stress responses and engineering more resilient crops (1). Scientists want to uncover essential genes involved in stress tolerance, confirm their function, and generate genetically improved cotton cultivars that can endure drought conditions by combining their capacities.

By thoroughly studying the roles and interconnections of genes within an organism's genome, functional genomics serves as a cornerstone in the research of gene function. Researchers obtain insights into gene expression patterns and protein functions related to stress reactions using approaches such as transcriptomics and proteomics (2). This plethora of information enables the identification of candidate genes that may play critical roles in stress tolerance in cotton.

Functional genomics, on the other hand, does not give a direct proof of gene function or its influence on stress tolerance. Additional methods are necessary to confirm the functional importance of candidate genes and assess their involvement in stress responses. This is when VIGS shines as a useful method (3). VIGS enables researchers to examine changes in stress tolerance and analyse the functional repercussions of gene silencing by temporarily reducing the expression of particular target genes using viral vectors. It allows for fast functional confirmation, reducing the list of genes for subsequent analysis.

Although VIGS provides useful insights, its benefits are fleeting. Researchers use the astonishing CRISPR/Cas9 technology to make permanent changes to the DNA. Scientists may create particular mutations or knockouts in target genes with astonishing accuracy and permanence with this gene-editing technology (4). CRISPR/Cas9 enables direct examination of the effects of gene changes on stress tolerance, allowing for a greater understanding of gene function and its influence on plant physiology.

The combination of functional genomics, VIGS, and CRISPR/Cas9 technologies provides a complete approach to elucidating the genetic basis of stress tolerance in cotton. VIGS provides for transitory gene silencing and functional validation (5), whereas CRISPR/Cas9 technology allows for permanent gene editing and validation. Researchers may bridge the gap between gene discovery, functional validation, and genetic engineering by combining these techniques (6). This comprehensive approach improves our knowledge of gene function and its influence on stress tolerance, paving the way for the generation of drought-tolerant cotton types.

Workflow

Virus-induced gene silencing (VIGS) offers an influential strategy in plant science for scrutinizing gene functions by

facilitating specific gene silencing using viral vectors (7). VIGS enables in-depth exploration of the implications of gene silencing on plant growth, physiology, and responses to various stressors. For altering gene expression, Tobacco rattle virus (TRV)-based VIGS is used efficiently across several plant species including *Nicotiana spp*., *Arabidopsis thaliana*, *G. arboreum*, *Petunia hybrida*, *Solanum lycopersicum* (8). Owing to its mild infection symptoms, extensive cellular infectivity, migration to rapidly growing meristematic cells, and effective spread to new plant tissues, TRV is considered a superior method for VIGS (9).

In the current study, the Spindly (*SPY*) gene, a protein superfamily member known to negatively regulate drought stress, was analyzed. An Insilco study was conducted to examine the structure and function of the *SPY* gene. The complete coding sequence of the gene was obtained from Cotton FGD (10). The Pfam tool was utilized to conduct a domain analysis of the protein, which helps understand gene structure patterns and motifs critical for protein-protein interaction. The *SPY* gene's structural analysis was accomplished through the gene structure display server 2.0 (11).

The vector used was 16437 bp in size and contained a 2.7 kb *SPY* gene, controlled by the 35S terminator and promoter, with Kanamycin serving as the selection marker (12). This binary vector comprises left (LB) and right borders (RB), with the *SPY* gene designed within the T-DNA region. The Agrobacterium was prepared for agro-infiltration, followed by mobilization of the vector into Agrobacterium (13).

Inoculation was conducted on two-leaf-stage cotton plants. Leaves were pricked with a needle, and the Agrobacterium culture containing the TRV binary vector (TRV1+TRV2) with the cloned *SPY* gene was inoculated to induce drought stress tolerance. The success of gene silencing was subsequently evaluated by analyzing the expression levels of endogenous genes using RT-PCR with RNA extracted from control and silenced cotton plants (14). The gene was found to be effectively silenced (Fig:2).

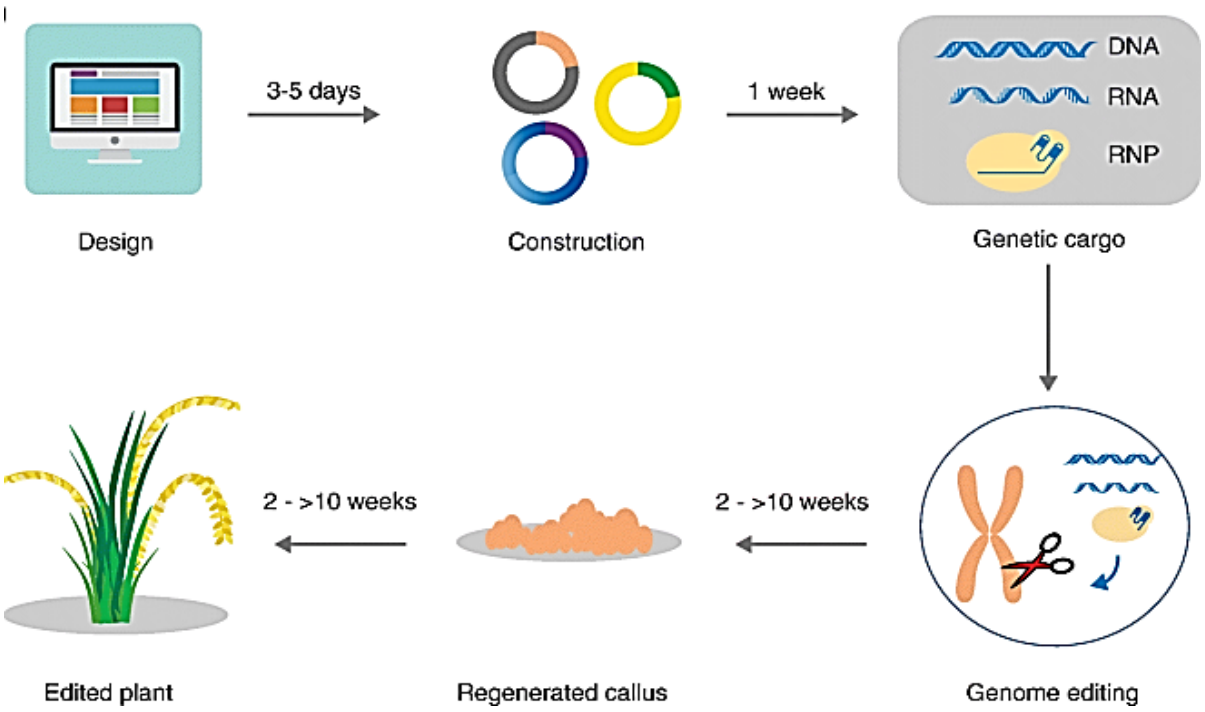

Fig:2 CRISPR Workflow for developing gene edited plants

Post-VIGS treatment, the impacts of gene silencing were assessed by evaluating phenotypic changes and gene expression alterations in silenced plants. This entailed measuring plant growth, development, or response to specific stressors, along with quantifying expression levels of the target gene and related genes (5).

To achieve permanent suppression of the *SPY* gene in cotton, CRISPR/Cas9 technology can be applied after the successful VIGS experiment. This would ensure a more stable, long-term alteration. The first step is the design of guide RNAs (gRNAs) for targeting the Cas9 endonuclease to the *SPY* gene (6). Following the synthesis and validation of gRNAs, they are assembled with the Cas9 nuclease into a binary vector, akin to the one used for VIGS. This CRISPR/Cas9 complex is then introduced into cotton plant cells using Agrobacterium-mediated transformation. This leads to double-strand breaks at the target *SPY* gene location, which the cell's DNA repair machinery fixes by introducing small indels that disrupt the gene's reading frame, leading to permanent gene silencing (11).

Post-transformation, the cotton plants are allowed to grow, and their progeny are screened for the desired mutation. Phenotypic analysis is performed once the mutation is confirmed through sequencing or techniques like the T7E1 mismatch cleavage assay, to verify that permanent silencing of the *SPY* gene confers improved

drought tolerance, as observed with transient VIGS silencing (7).

While VIGS provides a swift and temporary impact on mutation, leading to rapid suppression of the target gene, it can also serve as a precursor for permanent transformation. VIGS has proven to be a valuable tool in functional genomics and plant research, enabling rapid and efficient investigation of gene function and decipherings complex gene networks (8). The use of CRISPR/Cas9 for permanent gene silencing can be a valuable adjunct to VIGS, potentially leading to the development of stable, drought-resistant cotton varieties. However, it is critical to ensure thorough risk assessment and regulatory oversight with CRISPR/Cas9-mediated gene editing to prevent adverse effects on the environment or biodiversity.

Conclusion:

The combination of functional genomics, virus-induced gene silencing (VIGS), and CRISPR/Cas9 technology offers a strong and thorough strategy for figuring out the molecular processes behind cotton's stress tolerance. Candidate gene identification is based on functional genomics, whereas quick and temporary functional validation is made possible by VIGS and precise and long-lasting gene editing is made possible by CRISPR/Cas9. Combining these methods will enable scientists to produce cotton cultivars with higher drought resistance by developing a deeper knowledge of gene activity and how it affects stress tolerance.

Researchers may find out important genes involved in stress responses, confirm their function, and evaluate how much they contribute to stress tolerance thanks to the cooperative application of various tools. Exploring gene networks and signalling pathways and identifying possible relationships and hierarchies within them is made easier by this integrated approach. Comparative studies also make it possible to pinpoint the genetic factors that contribute to phenotypic variability and to create focused breeding plans for stress tolerance.

By creating crops that are better able to tolerate environmental challenges like drought, researchers may pave the road for sustainable agriculture by using the capabilities of functional genomics, VIGS, and CRISPR/Cas9 technologies. Global food security will be significantly impacted by the capacity to design crops with increased stress tolerance since it will increase agricultural yield and resistance to changing climatic circumstances. Researchers will learn more about the complex systems operating stress responses as they continue to dive further into functional genomics and gene editing technologies. We can make great progress towards creating robust crop types that can survive in difficult situations by integrating information from these domains. By combining these methods, we may get closer to a day when agriculture will be better able to feed the expanding world population and lessen the consequences of climate change.

References

Anu, K., Limiya Joseph, and C. Bindu Roy. "Tobacco Rattle Virus Vector-Based

Virus Induced Gene Silencing-Optimization of A Functional Genomics Tool for Hevea brasiliensis using Reverse Genetics Approach." (2021).Biswas, D., S. C. Saha and A. Dey. 2021. CRISPR-Cas genome-editing tool in plant abiotic stress-tolerance. Plant Gene. 26: 100286.

Bhattacharjee, Bipasha, and Vipin Hallan. "Geminivirus-derived vectors as tools for functional genomics." *Frontiers in Microbiology* 13 (2022).

Cunha, Beatriz Ribeiro da, Christian Andreasen, Jesper Rasmussen, Jon Nielsen, Christian Ritz, and Jens Carl Streibig. "Assessing herbicide symptoms by using a logarithmic field sprayer." *Pest management science* 75, no. 4 (2019): 1166-1171.Jain, S. M. 2010. Mutagenesis in crop improvement under the climate change. Rom Biotechnol Lett. 15(2): 88-106.

Li, Guan, Rong Liu, Rongfang Xu, Rajeev K. Varshney, Hanfeng Ding, Mengwei Li, Xin Yan et al. "Development of an Agrobacterium-mediated CRISPR/Cas9 system in pea (Pisum sativum L.)." *The Crop Journal* 11, no. 1 (2023): 132-139.

Manchur, Christopher. "The improvement and application of topically applied double stranded RNAs to control Sclerotinia sclerotiorum and Hyaloperonospora arabidopsidis." Master's thesis, 2022.

Shi, Yao, Wenlin Li, Yuliang Zhou, Xiaolan Liao, and Li Shi. "Contribution of multiple overexpressed carboxylesterase genes to indoxacarb resistance in Spodoptera litura." *Pest Management Science* 78, no. 5 (2022): 1903-1914.

Samantara, Kajal, Abhishek Bohra, Sourav Ranjan Mohapatra, Riry Prihatini, Flora Asibe, Lokendra Singh, Vincent P. Reyes et al. "Breeding more crops in less time: A perspective on speed breeding." *Biology* 11, no. 2 (2022): 275.

Yue, Ning, Zhihao Jiang, Xuan Zhang, Zhenggang Li, Xueting Wang, Zhiyan Wen, Zongyu Gao et al. "Palmitoylation of γb protein directs a dynamic switch between Barley stripe mosaic virus replication and movement." *The EMBO Journal* 41, no. 13 (2022): e110060.

Yadav, U. P., J. F. Evers, M. A. Shaikh, and B. G. Ayre. 2022. Cotton phloem loads from the apoplast using a single member of its nine-member sucrose transporter gene family. J. Exp. Bot. 73(3): 848-859.

Yang, Q.Q., X.X. Zhao, D.Wang, P.J. Zhang, X.N. Hu, S. Wei, and X.P. Yu. 2022. A reverse transcription-cross-priming amplification method with lateral flow dipstick assay for the rapid detection of Bean pod

mottle virus. Scientific Reports.12(1): 1-8.

Yang, S.J., S. A. Carter, A. B. Cole, N.H. Cheng, and R. S. Nelson. 2004. A natural variant of a host RNA-dependent RNA polymerase is associated with increased susceptibility to viruses by Nicotiana benthamiana. Proc. Natl. Acad. Sci. U.S.A. 101(16): 6297-6302.

Zhang, Y., and N. Zhang. 2018. Imaging technologies for plant high-throughput phenotyping: a review. Front Agric. Sci. Eng. 5(4): 406-419.

Zhao, M., L. Ma, N. Song, J. Cheng, Z. Zhao, and J. Wu. 2022. The regulation of Alternaria alternata resistance by LRR-RK4 through ERF109, defensin19, and phytoalexin scopoletin in Nicotiana attenuata. Plant Sci. 323: 111414.

Zheng, B., J. Liu, A. Gao, X. Chen, L. Gao, L. Liao, and Y. Han. 2022. Epigenetic reprogramming of H3K27me3 and DNA methylation during leaf-to-callus transition in peach. Hort. Res. 5.