Evaluation of the CRISPR/Cas9 system for the development of resistance against Cotton leaf curl virus in model plants

Sehrish Khan1*, Muhammad Shahid Mahmood1, Sajjad ur Rahman1, Farzana Rizvi2, Aftab Ahmad3*

1Institute of Microbiology, University of Agriculture, Faisalabad, Pakistan
2Department of Pathology, Faculty of Veterinary Science, University of Agriculture, Faisalabad, Pakistan
3Department of Biochemistry/U.S.-Pakistan Center for Advanced Studies in Agriculture and Food Security, University of Agriculture Faisalabad, Faisalabad, Pakistan
*Corresponding authors: aftab.ahmad@uaf.edu.pk, sana88k@gmail.com


Abstract: Over the last decade, the Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR) CRISPR/Cas9 system has been used by biologists in various fields. In plant biology, the technology is being utilised to manufacture transgenic plants resistant to different diseases. In Pakistan, the Cotton leaf curl virus (CLCuV) (a begomovirus) affects cotton plants causing significant loss to the economy of this agriculturally based country. In the present study, we use the CRISPR/Cas9 system in Nicotiana benthamiana Domin (a model plant) to develop resistance against CLCuV. An interesting facet of the study was the comparison of two constructs (pHSE401 and pKSE401) with regards to their efficacy in the virus inhibition. The pKSE401 vector contained a Cas9 nuclease and two guide RNAs (gRNAs), one targeting the Replication associated protein (Rep) gene and the other targeted the βC1 gene of the Betasatellite. The vector pHSE401 had only one sgRNA that targeted the (Rep) gene. Both genes that are intended to be targeted play important roles in the replication of CLCuV. Plants infiltrated with pKSE401 exhibited a delay in the development of the symptoms of the disease and showed lower virus titres. Our proposed multiplexing approach gave efficient results of the resistance in the model plants, and the results in this communication may be extended to the CRISPR/Cas9 based editing of cotton plants.

Keywords: Nicotiana benthamiana; sgRNA; Begomovirus; vector; viruses

In recent times, genome editing in plants has been subject to vast improvements with several traits of interest being introduced into different crops by site-specific modifications of the genome (Sovová et al. 2017). For a comprehensive review about important molecular tools being used for gene editing in plants, see the review (Mohanta et al. 2017). One of these tools, and perhaps the most well-known, is the CRISPR/Cas9 system; a prokaryotic adaptive immune system that is found in 90% of Archaea and 50% of bacteria (Ishino et al. 2018). In comparison to previous gene editing techniques, the CRISPR/Cas9 system is more efficient and can edit multiple genes (multiplexing) of interest simultaneously (Bortesi & Fischer 2015). The CRISPR/Cas system consists of a short synthetic gRNA sequence of 20 nucleotides

Supported by the USAid under United States Pakistan Center for Advanced Studies in Agriculture and Food Security, (USPCAS-AFS), University of Agriculture, Faisalabad, Pakistan.
that bind to the target DNA. The second component is the Cas9 nuclease which contains two domains: the RuVc and HNH domain, each domain cuts a single DNA strand. The mechanism of action of Cas9 includes the cleavage of 3–4 bases after the protoscaler adjacent motif (PAM, 5’ NGG) (Jinek et al. 2012).

To perform a CRISPR project in any molecular biology laboratory, the following steps have to be followed, (i) identify the PAM sequence in the target gene, (ii) design and synthesise the sgRNA, (iii) clone the sgRNA as a construct, (iv) introduce the construct into the target cell followed by screening and validation. The ease of the CRISPR mediated genome editing enables laboratories with only basic facilities to perform gene editing in various organisms. The technology has been adopted by scientists around the globe and is being used as a resource against diseases in plants.

The genus begomovirus (family Geminiviridae) consists of small ssDNA viruses, which possess either one (monopartite, DNA-A) or two (bipartite, DNA-A and DNA-B) genomic molecules of 2 800 nucleotides each (Czosnek et al. 2017). Each component has its own open reading frame (ORF) in a bidirectional fashion (Zaidi et al. 2017). The DNA-A typically has six ORFs with four in complementary sense orientation (AC1/C1-AC4/C4) and two (A1V1/V1 and A2V2/V2) in virion sense orientation. The C1 gene encodes the replication associated protein (Rep), while the C2 gene encodes the replication enhancer (REN), C3 encodes the transcriptional activator protein (TraP) and C4 encodes the C4 protein (Rojas et al. 2005). The DNA-B has two ORF’s, the BC1 gene encodes the movement protein (MP) on the complementary sense strand, and BV1 encodes the nuclear shuttle protein on the virus sense strand (Ha et al. 2008). Conserved sequence called the common region (CR) is present in both the DNA-A and DNA-B. The CR has a hairpin structure with a conserved nonanucleotide (TAATATT/AC). Iterons, which are 5–7 nucleotide long, are present at 5’ of the hairpin and form binding sites for the Rep protein.

Old world begomoviruses are associated with pathogenic determinant betasatellites and alphasatellites (self-replicating); the betasatellites encode the βC1 protein (considered important in the spread of the disease) (Chatterji et al. 2000). The vector of begomoviruses is the whitefly (Bemisia tabaci), which transmits the virus in a persistent circular non-propagative manner (Wei et al. 2017). The entry of virions into the vector occurs during feeding on the phloem sap of the infected plants. The virions then traverse the gut via a transcytotic pathway. From the gut, the virions proceed to the haemolymph and finally reach the salivary glands where the transmission process occurs (Czosnek et al. 2017). One of the most notorious begomoviruses, the Cotton leaf curl virus (CLCuV), is the causative agent of the cotton leaf curl disease (CLCuD). This disease is a major biotic restraint of cotton plants and is prevalent in Pakistan (Mansoor et al. 1999), India (Kirthi et al. 2004) and some African nations (Tahir et al. 2011). Recent reports indicate the presence of the virus in South-eastern China and the Philippines (Masood & Briddon 2018).

Major disease symptoms in plants are swelling of the veins with the upward and downward curling of the leaves (Briddon & Markham 2000). In the case of severe infections, leaf enation occurs which eventually leads to the retardation of the growth of the cotton plants. The severity of the infection depends upon the variety of the cotton and the age of the infected plants (Sattar et al. 2013). The occurrence of the disease has catastrophic consequences on the yield. According to an estimate, the virus can cause a staggering loss of yield in the range of 10–70% (Rahman et al. 2017). In Pakistan, the first CLCuV epidemic took place from 1991–1992, in which the outbreak was attributed to several monopartite viruses (Saeed et al. 2015). In India, the first cases of CLCuD were reported in 1993. Recently, there have been reports on the emergence of several variants of CLCuV in the Indian subcontinent with chances of potential outbreaks (Qadir et al. 2019).

During the last decade, much effort has been undertaken to develop resistance in cotton plants against CLCuV. In 2017, Ahmad et al. (2017) used a small interfering RNA (siRNA) construct to target the AC1 gene of the Cotton leaf curl Kokhran virus-Burewala (CLCuKoV-Bu) and the βC1 gene of the Cotton leaf curl Multan betasatellite (CLCuMB). The results of the assay showed that the plants with a transient expression did not exhibit any characteristic leaf curl symptoms, the qPCR results confirmed a significant reduction in the accumulation of the begomoviral-betasatellite in the plants. In a recent study, a CRISPR/Cas9 multiplexing approach was used to target two essential regions of the single stranded DNA genome of the Cotton leaf curl Multan virus (CLCuMvV) (Yin et al. 2019). The transient expression resulted in resistance to CLCuMvV.

In the present study, we planned the transient expression of two vectors (pHSE401 and pKSE401)
containing the sgRNAs and the Cas9 nuclease. The notion was to transform both vectors in Nicotiana benthamiana (the model plant) and to evaluate the virus interference by analysis of the signs and symptoms of the disease and virus titre.

MATERIAL AND METHODS

Analysis of the viral genome. The whole genome sequence of the Cotton leaf curl Kokhran virus (CLCKV) (Accession# NC_004583.1) was retrieved from the National Center for Biological Information (NCBI) in January 2018 (the sequence is given in the ESM- Electronic supplementary material). The two genes to be targeted were the Replication associated protein (Rep) gene (ESM) and the βC1 gene of the betasatellite (ESM).

Construction of pHSE401 and pKSE401. For pHSE401, the CRISPRdirect (Naito et al. 2015), software was used to analyse the DNA sequences of the target site (Rep gene). The manual selection of a 20 bp target site was performed in the noncoding region, according to the previously defined criteria (Mali et al. 2013). The forward and reverse DNA oligos encoding the gRNA were designed manually and synthesised. The sgRNA encoding DNA oligos were cloned into the pHSE401 expression vector containing a promoter and gRNA scaffold. In a small polymerase chain reaction (PCR) tube, the sgRNA encoding oligos and double-distilled water were added in a given amount. The conditions of the PCR reactions were as follows, 95 + 90 + 80 + 70 + 60 + 50 + 40 + 30 + 20 + 4 °C ∞ (1 min each). Furthermore, the mixture was incubated in a thermocycler to anneal the oligos for 5 min at 95 °C and at 25 °C for 20 minutes. For the visual analysis, the annealed oligos were run on 2.5% agarose gel by gel electrophoresis. The annealed oligos were stored at –20 °C.

The digested pHSE401 vector and the annealed oligos were ligated. Then, the ligation reaction mixture was incubated at 25 °C for 2–3 hours. The ligation mixture was transformed into an Escherichia coli strain (DH5α) according to the above-mentioned protocol. The transformed colonies were selected on Luria Bertani (LB) plates with the antibiotic kanamycin. A similar procedure was used for the construction for pKSE401, with the addition that the plasmid contained a maize codon optimised Cas9 gene with two gRNAs (1 and 2), with both gRNAs under the U6-26 promoter.

Confirmation of sgRNA. The colony PCR was used for the confirmation of the transformation into the DH5α cells. The primers for the two constructs are shown in Table 1.

Transformation to Agrobacterium tumefaciens strain gv 3101. Agrobacterium tumefaciens gv 3101 stock was prepared. The transformation of the gv 3101 cells with the plasmid was performed by electroporation, and the suspected transformed colonies were grown in the LB agar with kanamycin overnight at 28 °C. The positive clones were screened by PCR for the confirmation of the vector transformation.

Plant materials. Seeds of N. benthamiana were initially placed on a Murashige and Skoog MS media (MS salt 4.43 g, sucrose 30%, agar 12 g) in petri plates. After a duration of seven days, the plants were shifted to pots having soil (one plant each pot, 2.5” × 3”, with 250 g of soil). A total of 10 pots/10 plants were kept for each replicate and were watered regularly. The conditions of the growth room were a 16 h light period and an 8 h dark period at 25 °C.

Analysis of the expression levels of the sgRNAs and Cas9 by quantitative PCR (qPCR). The inoculation culture of A. tumefaciens was grown to an optical density of 0.5 at 600 nm after growing at 28 °C for 48 h with 25 μg/mL of rifampicin and 50 μg/mL of kanamycin. The Agrobacterium cells were spun at 5 000 × g for 15 min at 20 °C and re-suspended in 10 mM MgCl2 containing 150 μg of acetosyringone per mL, after which the infiltration medium was kept at room temperature for 3 hours. The agro-infiltration of the culture and the infiltration medium was undertaken using a 5 mL syringe into the plant (N. benthamiana) leaves. The total RNA was extracted from the plant leaves using TRIzol

Table 1. Primers used for the two constructs

<table>
<thead>
<tr>
<th>Vector</th>
<th>Primer</th>
</tr>
</thead>
</table>
| pHSE401 | U626F: TGTCCCAGGATTAGAATGATTAGGC  
|         | dT4-R: AAAACGTATTATAAAACGGAATGCC |
| pKSE401 | gRNA1-F: AGTGGATTGATCGTTTAATTTACCG  
|         | gRNA2-F: TGATTGGAAATGCTCTTTGACGTGTAAC  
|         | Cas9-Ver-R: T2TGTAATCAATGTCGTGGTCCTTGTAATC |
The cDNA was synthesised with 1 μg of RNA using the HiScript Q RT SuperMix for qPCR (+gDNA wiper) kit (Vazyme, China) according to the manufacturer’s instructions. The expression levels of the sgRNAs and Cas9 were detected by qPCR using the Applied Biosystems 7500 Fast Dx Real-Time system (Thermofisher, USA) according to the manufacturer’s protocol. The pentatricopeptide repeat containing protein (PPR) gene (Accession No. GO602734 (At1g62930) of *N. benthamiana* was used as an internal control with the forward and reverse primers F-GAGGGTCCATT TGAGTGAC/R-AGGCTGATGGTTGAATCTGG. The concentration of each DNA sample was adjusted to 20 μg/μL using a NanoDrop 2000 (Thermoscientific, USA).

**Virus infectivity assay.** The infectious clones of the *Cotton leaf curl Kokhran virus* (CLCuKV) and *Cotton leaf curl Multan* (CLCuVMB) were agro-infiltrated in the *N. benthamiana* plants via *Agrobacterium tumefaciens* (Leuzinger et al. 2013). This was done 2 days post agro-infiltration of the constructs in the plants (Khan et al. 2018; Mubarik et al. 2019). A total of 40 plants were divided into five groups (A–E). Group A and B were given pHSE401 and pKSE401 without the virus. Group C was kept as a positive control and the plants were infected with both viruses (CLCuKV/CLCuMB) without any sgRNA/Cas9. Group D contained both infectious clones and pHSE401, while Group E contained the infectious clones and pKSE401. The plant leaf samples were taken at 14 days post inoculation (dpi) for the DNA extraction to perform the PCR and qPCR. The same primers were used for the PCR and qPCR as given previously.

**Statistical analysis.** For the statistical analysis of the readings of the qPCR, the standard error mean (SEM) was calculated.

**RESULTS**

**Cloning of the plasmid into *E. coli* (DH5α).** The map of the Cas9 expression vector pHSE401 with the cloned gRNA is shown in Figure 1. The colony PCR was used to confirm the cloning of the sgRNA. For pHSE 401, the primers U-26-F and dt4 were used. The positive clones gave a band size of 0.8 kb (Figure 2) indicating the successful cloning of the vector into the *E. coli* cells. Furthermore, the sequencing of the plasmid using the specific primers revealed information about the presence and orientation of the gRNA in the vector. For the PKSE-401 vector (map in Figure 3), the primers
gRNA-1 F, gRNA-2F and Cas9-verR2 were used. The positive clones gave a band of 1.4 kb shown in Figure 4. The sequencing of the plasmid using specific primers revealed information about the presence and orientation of the gRNA in the vector.

Expression analysis of gRNA/Cas9 in the *N. benthamiana* leaves. The expression levels of the sgRNAs/Cas9 in the leaves of the plants were evaluated at 3 days post-inoculation using the qPCR. The expression of the three sgRNAs and Cas9 are shown in Figures 5 and 6.

Viral infectivity assay. Groups A and B (only the constructs with no virus) had no signs or symptoms of the disease, while the control group C (given the infectious clones only) exhibited signs and symptoms of the disease at 8–15 days. The plants of group D

![Figure 4](image)

Figure 4. The confirmation of the gRNA clones of the pKSE401 vector by the colony PCR. The cloning of the 20-nucleotide gRNA was confirmed in the vector. The PCR product was run on a 1.2% agarose gel. M is the ladder used, lanes 1–4 contain the positive clones of 1.4 kb in size.

![Figure 5](image)

Figure 5. (A) The qPCR results of the expression level of the sgRNA of pHSE401 in the plant leaves at 3 days post-inoculation and (B) the qPCR results of the expression level of the Cas9 of pHSE401 in the plant leaves at 3 days post inoculation.

(A) The relative expression of ten samples (A1–A10) is shown in the figure bars relative to the control. The error bars indicate the SEM. The highest expression level of the sgRNA1 was observed in sample A6 which was 1.3 relative to the control (0.1). For the sgRNA2, the highest expression value was found for sample A17, which was 1.5 relative to the control (0.1). (B) The relative expression of ten samples (A1–A10) are shown in the figure bars relative to the control. The error bars indicate the standard mean error. The expression level of the Cas9 was highest for sample A8, which was 2.0 relative to the control.

![Figure 6](image)

Figure 6. (A) The qPCR results of the expression level of the sgRNA1 and sgRNA2 of pKSE401 in the plant leaves at 3 days post inoculation and (B) the qPCR results of the expression level of the Cas9 of pKSE401 in the plant leaves at 3 days post inoculation.
developed symptoms of the disease with a delay of 2–4 days, while a 3–6-day delay of the symptoms was seen in group E. Overall, the expression of the sgRNAs/Cas9 suppressed the viral proliferation. However, the group given the pKSE401 construct (two sgRNAs) showed better results when compared to pHSE401 (Figure 7).

All the experiments were performed in three replications and are shown in Table 2.

**Determination of the virus accumulation by qPCR.** The accumulation of the virus under the transient expression of both pHSE401 and pKSE401 was analysed in *N. benthamiana* at 15 days post inoculation. The results of the qPCR indicated a low virus titre in the range of 0.3–0.6 compared to that of the control (a value of 1) for pHSE401 as shown in Figure 8, and an even lower virus titre range (0.2–0.4) for pKSE401 as shown in Figure 9. The co-infiltrated plants with both vectors and CLCuKV/CLCuMB showed a decrease of 40–80% in the virus titre.

---

**Figure 7.** (A) The plants with only pHSE 401 are shown without the inoculation of the infectious clones, no signs and symptoms of disease were observed. (B) This group, the plants were infiltrated with only pKSE401 without any virus, consequently, no signs and symptoms of disease were observed. (C) Group C was the positive control, which contained only the infectious clones of CLCuKV/CLCuMV without any of the constructs. The usual signs and symptoms of CLCuD were observed at 8–15 days post inoculation. (D) The plants of Group D were infiltrated with pHSE401 and the infectious clones of both viruses. A delay of 2–4 days in the occurrence of the infection was observed, in addition, the symptoms of the disease were also mild, and the plants recovered from the disease, 30 days post inoculation. (E) The plants of Group E were infiltrated with pKSE401 and the infectious clones of the viruses. A delay of 3–6 days in the development of symptoms was observed. The plants recovered from infection by the 30th day of the initial inoculation.

---

**Figure 8.** The results of the virus accumulation analysed by the qPCR in the ten plants with pHSE401 and the virus as a reference. On the 15th day post infection, the low accumulation of the virus was observed in the plants of the range 0.3–0.6 in comparison to the reference.

**Figure 9.** The results of the virus accumulation analysed by the qPCR in the ten plants with pKSE401 and the virus as reference. The low accumulation of the virus was observed in the plants of the range 0.2–0.4 in comparison to the reference.
DISCUSSION

Cotton holds a pivotal position in the global economy, as it accounts for 40% of the international fibre market (Long et al. 2018). Pakistan is the 4th largest producer of cotton, and its economy gets a substantial boost by its export (2nd largest exporter in the world). Despite the importance of cotton, there have been few functional studies on its genes (Zhang et al. 2016). Recent advancements in the field of genomics, proteomics and transcriptomics have paved the way for biotechnological tools that could improve cotton plants in various ways. One of these improvements is the genetic modification of crops to make them resistant to different diseases. The Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR)/Cas9 system has emerged as a powerful and precise gene editing tool in eukaryotes (Jiang et al. 2014). There has been a dramatic increase in the number of CRISPR publications in PubMed over the past 12 years, and it was projected that the publication count will exceed 5000 by the end of 2018 (Adli 2018). The use of CRISPR/Cas9 in gene editing in cotton has been hindered by issues with the genetic transformation process, making it imperative to maximise its efficiency (Long et al. 2018).

The CRISPR/Cas9 system has proven to be more efficient than Zinc Finger Nucleases (ZFNs) and transcription activator-like effector nucleases (TALENs). The cost-effectiveness and ease of design are certain benefits of the system, however, a higher frequency of off-targets have been reported (Fu et al. 2013). The use of CRISPR/Cas9 in suppressing the DNA plant viruses has gained fame (Ji et al. 2017), however, reports have indicated virus resistance towards the insertions and deletion mediated by the CRISPR/Cas9 system and escape after the double stranded breaks (DSB) (Ali et al. 2016).

In similar studies on CLCuKV, RNA interference (RNAi) technologies have been used to disrupt the replication of the virus, but off-target activity limited the scope of the molecular tool (Carthew & Sontheimer 2009). In another study by (Chen et al. 2014), ZFN’s were used to target the conserved motifs in the Rep gene. The cleavage of the target site was observed, which resulted in the inhibition of the viral replication.

In this study, the Rep gene and βC1 gene of the betasatellite of the CLCV were targeted to generate transgenic plants resistant to the virus. For this purpose, we designed two constructs, pHSE 401 (One sgRNA and Cas9) to target only the Rep gene, while the pKSE401 construct contained two sgRNAs and Cas9, one sgRNA targeting the Rep gene and the other targeting the βC1 gene. The expression of the sgRNAs and Cas9 in the transformed leaves was measured using qPCR (Figures 4–6). The expression quantification results were similar to those of Ji et al. (2015) who observed similar patterns. The efficient transformation of the CRISPR/Cas9 constructs are pivotal in the effective integration of sgRNA and Cas9 into the target plants. So far, Agrobacterium mediated transformation has yielded promising results and the same approach was used in the present communication. We successfully transferred both vectors into the model plants, and confirmed the success by measurement of the expression levels of both the sgRNA and Cas9.

In our study, we found a delayed onset of the disease symptoms in the transgenic plants. In addition, we found low virus titres in the range of 0.3–0.6 for the plants with pHSE401 and an even lower titre of the plants with pKSE401. This shows that our construct with two sgRNAs targeting the rep gene and βC1 gene could provide efficient resistance in plants against CLCV. The multiplex approach with CRISPR/Cas9 in plants has been proposed previously by (Xing et al. 2014).

A limitation of this study was the inability to perform an analysis of the type of mutation induced by both constructs. We relied only on the measurement of the expression levels of the sgRNA and Cas9 by qPCR followed by the analysis of the virus titre in
the different groups. Further efforts will focus on sequencing to unveil details of the type of mutations (Insertion/deletion) that were induced in the model plants. Overall, the results of this research lay the foundation for using the multiplexing approach for the CRISPR based editing of cotton plants.

In Pakistan, the CRISPR/Cas9 technology is still in its nativity with progress being made for the uplift of this gene editing technology. Being an agriculturally-based country, a major portion of the Gross Domestic Product (GDP) is derived from this sector. So, any loss to crops directly influences the progress of the economy. With diseases and severe climate changes on the rise, the development of disease and stress resistance plants is the need of the hour. The development of transgenic plants with disease resistance characteristics will certainly be a move in the right direction. In addition, plants resistant to different environment stresses is also being researched. The vector of CLCuV can also be targeted using different gene drive systems (already discussed in the Review of Literature section) developed by the CRISPR/Cas9 system. The modified vectors may be released into the wild and the populations of vectors may flourish that do not possess the ability to carry the virus. As the present communication was performed in-vitro, not much can be said about the effect of the technique on plants in the field. Future studies may be performed to check the efficiency of the present research by inoculating cotton plants in the field. However, our multiplexing approach proved to be effective in the model plants by targeting the two genes. Further multiplexing approaches may be made to target more genes in the genome of CLCuV to converse the resistance in cotton plants. In this regard, future field trials with transformed cotton plants is highly recommended.

**CONCLUSION**

The CRISPR/Cas9 system can provide resistance against different diseases. One possible outcome is shown in this manuscript. However, any future testing of the mutants will be required to verify the success of the technology. The coming years will be interesting, as more molecular and bioinformatic tools will become available to analyse the success of CRISPR/Cas9 in various plant species.

**Acknowledgement:** We acknowledge Dr. Hassan Zafar and Prof. Dr. Milton H Saier, University of California, San Diego, USA for their technical advice.

**REFERENCES**


https://doi.org/10.17221/105/2019-PPS

Original Paper


