Engineering Molecular Immunity against Plant Viruses

Syed Shan-e-Ali Zaidi\textsuperscript{1,2}, Manal Tashkandi\textsuperscript{1}, and Magdy M. Mahfouz\textsuperscript{1,*}

\textsuperscript{1}Laboratory for Genome Engineering, Division of Biological Sciences, 4700 King Abdullah University of Science and Technology, Thuwal 23955-6900, Saudi Arabia

\textsuperscript{2}National Institute for Biotechnology and Genetic Engineering (NIBGE), Faisalabad, Pakistan

*Correspondence: M.M. Mahfouz
magdy.mahfouz@kaust.edu.sa

Abstract

Genomic engineering has been used to precisely alter eukaryotic genomes at the single-base level for targeted gene editing, replacement, fusion, and mutagenesis, and plant viruses such as Tobacco rattle virus have been developed into efficient vectors for delivering genome-engineering reagents. In addition to altering the host genome, these methods can target pathogens to engineer molecular immunity. Indeed, recent studies have shown that clustered regularly interspaced short palindromic repeats (CRISPR)/CRISPR-associated 9 (Cas9) systems that target the genomes of DNA viruses can interfere with viral activity and limit viral symptoms \textit{in planta}, demonstrating the utility of this system for engineering molecular immunity in plants. CRISPR/Cas9 can efficiently target single and multiple viral infections and confer plant immunity. Here, we discuss the use of site-specific nucleases to engineer molecular immunity against DNA and RNA viruses in plants. We also explore how to address the potential challenges encountered when producing plants with engineered resistance to single and mixed viral infections.
Keywords

Site-specific nuclease, genome engineering, CRISPR/Cas9, Geminiviruses, tobacco rattle virus

1. Introduction

The ability to alter DNA sequences in a site-specific manner and to generate gene mutants and protein variants can facilitate the interrogation of gene function and understanding of functional biology. The term “genome engineering” refers to the precise, site-specific manipulation of the genome at the single-base level. Genome engineering uses site-specific nucleases (SSNs) to generate site-specific double-stranded breaks (DSBs) in the genome. DSBs trigger the cellular repair machinery and two major DNA repair pathways, the imprecise non-homologous end joining (NHEJ) pathway and the precise homology-directed repair (HDR) pathway, can repair DSBs (Figure 1). NHEJ repair is suitable for site-specific mutagenesis and for subsequent functional knockout of genes. Therefore, genome engineering via NHEJ is ideal for interrogating gene functions. By contrast, HDR uses a DNA template to provide the information needed by the HDR repair machinery to copy the sequence across the break. HDR constitutes a precise form of repair and is ideal for generating gene and protein variants based on a DNA template that is supplied with the nuclease that produces the DSBs. Several genome engineering platforms have been developed and repeatedly improved, but limitations to their efficiency and ease of use remain, especially in plants. One limitation is the choice of vector for delivering genome-engineering reagents into the host cell.
Work in a variety of systems has used viruses to deliver genome-engineering reagents into target cells. For example, in mammalian systems, the effective delivery of genome engineering reagents has been achieved using viruses, including retroviruses, adenoviruses, and adeno-associated viruses. These viral vectors can be delivered into virtually any human cell type and function as templates not only for expression of the designer site-specific nucleases (SSNs), but also for targeted chromosomal integration of foreign DNA.

Recent studies have developed DNA and RNA viruses for genome engineering in plants. In addition to using viruses as delivery vectors, genome engineering reagents can now be used to target invading DNA viruses for degradation and to establish plant molecular immunity against single and multiple viral infections. For example, Ali et al. engineered tobacco plants with clustered regularly interspaced short palindromic repeats (CRISPR)/CRISPR-associated (Cas9) targeting the Tomato yellow leaf curl virus (TYLCV), establishing virus interference and significantly limiting viral symptoms. Similarly, Baltes et al. and Ji et al. established virus interference in tobacco against Bean yellow dwarf virus (BeYDV) and in Arabidopsis against Beet severe curly top virus (BSCTV). Here, we discuss the current research on the use of genome engineering to develop virus interference strategies, and address the potential challenges of engineering plants with effective, durable virus resistance.

2. The use of SSNs for genome engineering

SSNs, including zinc finger nucleases (ZFNs), transcription activator-like effector nucleases (TALENs), and the recently developed CRISPR/Cas9 system, have been used to produce targeted mutations and edit the genomes of various eukaryotic species (Figure 1).
ZNMs and TALENs have been used for targeted editing of plant genomes\textsuperscript{27-29}, but the specificity of these SSNs\textsuperscript{30} comes from protein–DNA binding and therefore ZFNs and TALENs require protein engineering for each user-selected target sequence\textsuperscript{31}. Engineering ZFNs and TALENs is time-consuming and resource-intensive and must be performed by skilled molecular biologists\textsuperscript{32}. Several repeat assembly protocols have been developed for TALEN engineering, but the need for two TALEN monomers to bind to the sense and antisense strands simultaneously complicates the broad utility and applicability of this technique in plants\textsuperscript{4}. These limitations have been largely overcome in the recently developed CRISPR/Cas9 system\textsuperscript{34-36}. CRISPR/Cas9 engineering is highly predictable and much easier to perform than the other techniques, with greater reproducibility and affordability\textsuperscript{37}.

In nature, the CRISPR/Cas9 system provides adaptive molecular immunity to bacterial and archaeal species to confer resistance against invading DNA species, including conjugative plasmids and phages\textsuperscript{38}. The CRISPR/Cas9 locus in the bacterial genome is typically composed of Cas genes and the repeat and spacer sequences downstream of a leader sequence\textsuperscript{39,40}. CRISPR/Cas9-based immunity is achieved via three distinct phases: foreign DNA is acquired from the plasmid or phage (acquisition phase), the CRISPR locus is transcribed and CRISPR RNA processing and maturation take place (transcription and maturation phase), and the paired trans-activating crRNAs (tracrRNAs) and CRISPR RNAs (crRNAs) are used as guides for the Cas9 endonuclease to target or interfere with subsequent invasions (interference phase)\textsuperscript{39,41,42}. Therefore, the CRISPR/Cas9 system acquires pieces of DNA from an invading phage or plasmid and uses these sequences as molecular records to guide Cas9 endonuclease to bind to the phage or plasmid DNA in subsequent invasions, binding in a region preceding the protospacer associated motif (PAM) sequence\textsuperscript{43}. The PAM, a trinucleotide sequence that is important for
Cas9 activity, is used to distinguish self from non-self DNA\textsuperscript{44}. The PAM sequence of
\textit{Streptococcus pyogenes} Cas9 is NGG (with N being any nucleotide) and that of \textit{S. thermophilus}
Cas9 is NGGNG\textsuperscript{43,45}.

Notably, the tracrRNAs and crRNAs required for Cas9 specificity can be fused to
generate a chimeric single-guide RNA molecule (sgRNA) that can direct Cas9 to a specific
sequence preceding the PAM trinucleotide sequence\textsuperscript{46}. Since engineering via the CRISPR/Cas9
system involves changing only the 20-nucleotide sequence, the CRISPR/Cas9 system is the
system of choice for genome engineering applications\textsuperscript{26}.

The CRISPR/Cas9 system is based on Watson-Crick base pairing between the sgRNA
and the DNA of the target genome\textsuperscript{47}. Therefore, the CRISPR/Cas9 system supersedes and
outperforms other systems due to its ease of use, reproducibility, and affordability\textsuperscript{26,48}. The
CRISPR/Cas9 system has recently been used for various genome-engineering applications across
eukaryotic species, including plants\textsuperscript{49-51}. For example, the CRISPR/Cas9 system was used to edit
the wheat and maize genomes for targeted trait development\textsuperscript{52-54}. Various CRISPR/Cas9-based
genome engineering modalities have been used involving different Cas9 variants\textsuperscript{55} or
catalytically inactive Cas9 (dCas9) to modulate gene expression patterns and to subsequently
interrogate gene functions\textsuperscript{56,57}.

Editing of genes and genomes usually involves site-specific mutagenesis, including DNA
addition, deletion, or creation of gene fusions. For any genome-engineering platform or
modality, there are always concerns about specificity, delivery of reagents, and HDR efficiency
in the target cells. Therefore, ongoing, active research aims to improve the specificity, delivery,
and editing efficiency in a target cell or species to maximize the usefulness of the genome
engineering modality. Very recently, a novel class of DNA-guided SSNs involving FEN:FokI has been reported; these are expected to be suitable for various engineering purposes.

3. Viruses for delivery of genome engineering reagents

As mentioned in the Introduction, many studies in mammalian systems have successfully used viruses to deliver genome engineering reagents. For plants, the TRV RNA virus has great potential for use as a vector for genome engineering. Important features of the TRV-based delivery system include the following: 1) TRV can systemically infect many plant species, both naturally and under laboratory conditions; 2) TRV can easily introduced into plants via Agrobacterium and delivered into the growing points of the plant; 3) the small genome size of TRV facilitates cloning, multiplexing, library construction, and agroinfection; and 4) the RNA genome of this virus does not integrate into the plant genome. A TRV-based expression system has been developed for transient delivery of ZFNs into a variety of tissues and cells of intact plants, representing a non-transgenic approach for ZFN delivery and mutant plant production. TRV has a bipartite genome, with two RNA components: RNA1 (or TRV1) and RNA2 (or TRV2). In the above-mentioned system, RNA2 was replaced with RNA for the Zif268:FokI ZFN, and targeted genome modifications were recovered in somatic tobacco (Nicotiana tabacum) and petunia (Petunia hybrida) cells at an integrated reporter gene. TRV systemically infected its hosts, and ZFN-mediated, precise mutagenesis was clearly observed in newly developed infected tissues, as measured by the activation of a mutated reporter transgene in tobacco and petunia plants.
If TRV could infect germ cells, this would improve its utility as a vector for genome engineering, making it possible to harvest mutant seeds from infected plants. Indeed, the ability of TRV to move to developing buds and regenerating tissues has enabled researchers to recover mutated plants. Sequence analysis and examination of transmission of the mutations to the next generation confirmed the stability of the ZFN-induced genetic changes. Because TRV can infect a wide range of plant species, it provides a viable alternative to the production of ZFN-mediated mutants while avoiding the use of classical plant transformation methods.

Recently, we developed a TRV system that delivers single and multiple sgRNAs for targeted editing of the host genome and for targeting invading DNA viruses (Figure 3). In this system, TRV delivers single or multiple sgRNAs for genomic targets to modify the genomes of host plants, including tobacco and Arabidopsis. To determine whether the newly developed TRV could produce and deliver sgRNA molecules systemically and to evaluate the persistence of the genomic modification in growing tissues, the engineered TRV was delivered by agroinfiltration and reconstituted in the leaves of N. benthamiana expressing Cas9. Interestingly, TRV-mediated delivery of sgRNA was not limited to the original infiltrated plants; this precise targeting can also be transmitted to the next generation (Figure 2). This TRV-mediated sgRNA delivery system stands out because it bypasses the requirement for plant transformation and tissue culture. This technique opens up new possibilities for producing plants with desired traits using CRISPR/Cas9 without the need for laborious, time-consuming intermediate steps.

One potential limit of RNA viruses is that these viruses are limited in the amount of heterologous DNA they can carry (approximately 2–3 kb); excess cargo results in the loss of systemic movement or recombinational loss of the cargo DNA. This limits the utility of
these viruses for delivering TALENs or Cas9 expression cassettes, which typically exceed the
virus’s cargo capacity. However, for relatively small molecules such as ZFNs, meganucleases,
and sgRNAs, TRV was proven to be an effective delivery agent, achieving mutagenesis in
somatic and germline cells of several plant species, including tobacco 59, N. benthamiana 12,60,
and petunia 59. In addition to RNA plant viruses, certain DNA viruses have been used to deliver
genome engineering reagents, including SSNs and DNA templates, for efficient genome editing;
readers are directed to the following articles on the development of DNA viruses as vehicles for
genome engineering: 13,63-66.

4. Using SSNs to develop virus resistance

As noted above, several studies have reported engineering of SSNs to target viral
genomes for degradation. Successful attempts used ZFNs to target TYLCV in Arabidopsis 67,68;
resistance developed in transgenic plants expressing ZFNs targeting the TYLCV genome. Initial
studies aimed at targeting plant viruses via genome engineering have focused on using ZFN to
target begomovirus Rep binding sites/iterons, the iterated nucleotide sequences within the
begomovirus genome that provides binding site for virus replication associated protein (Rep).
ZFNs developed for iterons of Beet severe curly top virus (BSCTV) and Tomato yellow leaf curl
virus (TYLCV) effectively targeted the respective viruses 68,69. Transgenic Arabidopsis plants
expressing ZFNs that target iterons of BSCTV demonstrated complete resistance to this virus 69.
Chen et al. (2014) demonstrated that ZFN can efficiently target BSCTV and TYLCV but argued
that this viral sequence-specific strategy may not be effective in the field, where mixed viral
infections are common 67. As an alternative, targeting the conserved regions of viruses would
confer relatively durable resistance. Three conserved regions in Rep of monopartite begomoviruses were identified and tested for broad-spectrum resistance. Among these three targets, one targeting a 25-base pair sequence functioned well against both *Tomato yellow leaf curl China virus* (TYLCCNV) and *Tobacco curly shoot virus* (TbCSV)\(^{67}\).

Artificial TALE (ATALE), an artificial DNA-binding protein that has a highly conserved central region with tandem repeat units of 34 amino acids, also has been used for designing broad-spectrum resistance to begomoviruses\(^{70}\). Two highly conserved targets, AC1 and a nonanucleotide sequence, were selected and targeted using an artificial TALE (ATALE). The DNA binding efficiency of the ATALE was confirmed, and transgenic *N. benthamiana* plants were developed. The authors tested two begomoviruses, *Tobacco curly shoot virus* (TbCSV) and *Tomato leaf curl Yunnan virus* (TLCYnV), as well as one begomovirus with its associated betasatellite, *Tomato yellow leaf curl China virus* (TYLCCNV) with its cognate Tomato yellow leaf curl China betasatellite (TYLCCNB). The resulting transgenic plants demonstrated partial resistance, with delayed symptom development and reduced viral DNA accumulation compared to wild type\(^{70}\).

5. **CRISPR/Cas9-mediated virus interference**

Since the native CRISPR/Cas9 system can confer resistance against invading phages in bacterial and archaeal species, we investigated whether this system could be transferred to plant species to confer molecular immunity against plant DNA viruses just as ZFNs and TALENs were shown to do in the studies discussed above. Among plant DNA viruses, geminiviruses infect several economically important crops such as cotton\(^{71,72}\), tomato\(^{73}\), potato\(^{74}\), and several
legumes\textsuperscript{75,76}, as well as ornamentals\textsuperscript{77,78} and weeds\textsuperscript{79-81}. In the past two decades, molecular
biologists and breeders have worked to develop geminivirus-resistant crops\textsuperscript{82,83}. Recently,
several studies have demonstrated that engineering plant species, including tobacco and
\textit{Arabidopsis}, with CRISPR/Cas9 machinery targeting geminiviruses can confer molecular
immunity\textsuperscript{14,16,17}.

We recently showed that engineering plants to express the CRISPR/Cas9 machinery
targeting the TYLCV genome led to interference and degradation of the viral genome, resulting
in a significant reduction in viral symptoms\textsuperscript{14}. In Ali \textit{et al.}, (2015) we showed that \textit{N. benthamiana} plants expressing the CRISPR/Cas9 machinery exhibited resistance against
TYLCV, \textit{Beet curly top virus} (BCTV), and \textit{Merremia mosaic virus} (MeMV)\textsuperscript{14}. Baltes \textit{et al.} and
Ji \textit{et al.} demonstrated virus interference activities in \textit{N. benthamiana} against \textit{Bean yellow dwarf virus} (BeYDV) and \textit{Beet severe curly top virus} (BSCTV), respectively\textsuperscript{16,17}. Ji \textit{et al.} correlated
Cas9 expression with the levels of virus suppression, indicating the need to use a single
background transgenic line with optimum expression of Cas9 and sgRNA for practical
applications\textsuperscript{17}. Baltes \textit{et al.} showed that one sgRNA targeting the BeYDV genome could confer
plant resistance without cleavage activity, suggesting that catalytically inactive Cas9 (dCas9) can
be used to mediate virus interference, thereby eliminating concerns about off-target activities in
the plant genome\textsuperscript{16}.

In a follow-up study, we demonstrated that this technology can be used to target and
cleave \textit{Cotton leaf curl Kokhran virus} (CLCuKoV) and that by targeting the conserved
nonanucleotide sequence, multiple begomoviruses can be targeted simultaneously (CLCuKoV, TYLCV, TYLCSV, MeMV, BCTV-Worland, and BCTV-Logan), conferring broad-spectrum
geminivirus resistance\textsuperscript{15}. In these studies, \textit{N. benthamiana} plants expressing the CRISPR/Cas9
system displayed considerably reduced viral titers, which abolished or significantly reduced
disease symptoms.

We subsequently investigated the differences in efficiency of virus interference when
targeting coding versus non-coding sequences. Since invading viruses and host plants exist in
an arms race where pathogens rapidly evolve mechanisms to evade plant defenses, we
systematically analyzed the ability of geminiviruses to evade the CRISPR/Cas9 machinery. We
determined the ability of multiple geminiviruses to evade this machinery by targeting coding and
non-coding sequences, finding that targeting coding sequences led to the generation of viral
variants capable of evading the CRISPR/Cas9 machinery. Since these viral variants can have
variable replication efficiencies, variants with more proficient replication should predominate
over those with lower replication rates under natural field conditions, where mixed infections
predominate.

Interestingly, targeting non-coding intergenic sequences led to high levels of virus
interference, with no detectable viral escapes from the CRISPR/Cas9 machinery, thereby
providing an effective strategy for conferring potentially durable resistance. Given the
considerable heterogeneity and recombination ability of geminiviruses and the ability for
replication and systemic movement in CRISPR/Cas9-induced variants, strategies could be
devised in which the CRISPR/Cas9 system targets the non-coding region to inhibit or prevent
viral replication by mutating the essential parts of geminiviruses. Alternatively, targeting
multiple regions of the virus genome simultaneously would abrogate the ability of the virus to
use the NHEJ repair system, and these cleaved molecules would ultimately be degraded.
6. Developing resistance against RNA viruses

Plant viruses that store and transmit their genetic information in form of RNA are called RNA plant viruses. Several RNA viruses cause devastating yield losses in economically important crops like tomato, potato, cucurbits, etc. Genome engineering strategies aimed at manipulating host factors in the plant genome have been used to develop resistance against RNA viruses. For example, the CRISPR/Cas9 system is used to modify plant genes that affect plant responses to viral infection, thus generating virus resistance. The translation initiation-like factors eIF4E and eIF(iso)4E are directly involved in the infection process of RNA viruses. Arabidopsis plants with mutations in these translation initiation factors exhibit resistance to Turnip mosaic virus (TuMV). Therefore, these targets were mutated using the CRISPR/Cas9 system to develop virus-resistant plants. The utility of CRISPR/Cas9 technology for generating novel genetic resistance to the potyvirus TuMV was demonstrated in Arabidopsis via deletion of the host factor, eIF(iso)4E, which is strictly required for viral survival. Virus resistance was also developed in cucumber (Cucumis sativus) by disrupting eIF4E, resulting in non-transgenic, heterozygous eIF4E mutant plants. These non-transgenic plants exhibited partial resistance to an ipomovirus (Cucumber vein yellowing virus) and two potyviruses (Zucchini yellow mosaic virus and Papaya ring spot mosaic virus-W). Translation initiation factor genes are prime candidates for host genes that can be targeted, but any host gene encoding a factor that the virus requires is a potential target for modification. This form of recessive resistance could be exploited with the aid of the CRISPR/Cas9 system to create novel resistance alleles in crop plants to protect them against problematic viruses that use host translation initiation factors.
The CRISPR/Cas9 system has recently emerged as a groundbreaking genome-engineering platform. This system has been used to improve several traits in plants, e.g., engineering biotic stress tolerance (against invading viruses, fungus, bacteria, and insects), abiotic stress tolerance (against drought, heat, and cold stress), and yield improvement (e.g., by increasing grain weight in maize). Moreover, all of these works can be performed using a “DNA-free”, non-genetically modified organisms (GMO) crop production system. In addition to crop improvement, the CRISPR/Cas9 system can be used to examine the plant genome, e.g., by creating knockout libraries or by enhancing plant genome imaging (Figure 4). The CRISPR/Cas9 system has been used for targeted genome editing in diverse plant species, from model plants such as N. benthamiana and Arabidopsis to economically important plants such as wheat, maize, sorghum, and rice. The specificity of this system is determined by the 20-nucleotide sequence of the sgRNA, allowing for unprecedented, simple genome engineering. Furthermore, the CRISPR/Cas9 system can be used to simultaneously edit multiple genomic targets.

Viruses have long been used for genome engineering, mainly as delivery vehicles for mammalian cells. Plant viruses can potentially be powerful tools for improving the efficiency of genome editing. Recent studies have demonstrated the utility and applicability of using viruses for genome editing via CRISPR/Cas9 genome engineering platforms. In these studies, to generate plants with heritable modifications, it was necessary to generate transgenic lines that stably expressed Cas9 and gRNA; progeny with targeted modifications were then recovered in subsequent generations. Since producing transgenic plants is time-consuming, efficient delivery
methods are needed to expedite and maximize the utility of this technology for trait discovery
and development. The recently developed TRV-mediated CRISPR/Cas9 delivery system can
help overcome this hurdle. The use of heterozygous Cas9-overexpressing plants with this simple,
versatile genome-editing platform allows plants free of foreign DNA to be generated, which
might help overcome the regulatory hurdles that impede the commercialization of genetically
engineered plants. TRV has multiple advantages as a delivery vehicle for sgRNA, including its
small RNA genome, which is amenable to library construction and does not integrate into the
host genome. In addition, TRV can infect germline cells and has a wide host range under both
natural and laboratory conditions. However, the limitations in the carrying capacity of TRV
and other RNA viruses prevent their use beyond the expression of relatively small SSNs and
sgRNAs, and they are unable to efficiently deliver DNA repair templates. Perhaps geminiviruses
can be used to overcome the limitations of RNA viruses due to their larger carrying capacity and
their ability to produce a DNA replicon capable of acting as a repair template for gene targeting
97. In an effort to facilitate efficient genome engineering, various viruses have recently been
developed to deliver genome-engineering reagents into plants.

Several genome-engineering platforms have been used to engineer plants with molecular
immunity against invading viruses. However, the CRISPR/Cas9 system has shown promise for
providing immunity against economically important, widespread geminiviruses such as TYLCV.
Moreover, targeting the non-coding regions of viruses provides significant interference activity
and limits the generation of viral variants that escape the CRISPR/Cas9 machinery and are
capable of replication and systemic movement. These developments are exciting but have thus
far been limited to laboratory demonstrations. Field trials in multiple locations would be helpful
for further evaluating the actual success of this technology under natural conditions. Moreover,
previous studies have mainly focused on targeting DNA viruses due to the robust activity of the CRISPR/Cas9 system for cleaving DNA substrates. Novel variants of the CRISPR/Cas system were recently shown to be capable of targeting RNA genomes. These systems are based on the concept that the Cas9 binds with high affinity to single-stranded RNA (ssRNA) targets matching the Cas9-associated gRNA sequence when the PAM is presented in trans as a separate DNA oligonucleotide. Furthermore, PAM-presenting oligonucleotides (PAMmers) stimulate site-specific endonucleolytic cleavage of ssRNA targets; using specially designed PAMmers, Cas9 can be specifically directed to bind or cut RNA targets. Such systems could be applied to target the RNA genomes of single and multiple viruses and to confer resistance against RNA viruses.

One of the salient features of the CRISPR/Cas9 system is that it can be used in a ‘DNA-Free’ crop production system to develop non-transgenic plants with desired traits. This is even more practical when endogenous host genes are targeted, since the progeny plants in the following segregating generations can be selected for the desired phenotypes caused by the edited traits, but lacking the Cas9 and sgRNA transgenes. This system has direct application in crop production and variety development. An alternate approach to avoid the incorporation of foreign genes into the host genome can be the introduction of CRISPR/Cas9 reagents directly into the cell, for example by using a Cas9/sgRNA ribonucleoprotein complex. In this way the final product would be a plant, no different from the one produced as a result of classical plant breeding approaches. The DNA-free mushroom and maize varieties, that have been developed using similar CRISPR/Cas9 mediated genome editing approach have been deemed by the US Department of Agriculture to not be subject to their regulatory system. This decision may pave the way for these and other plants free of foreign DNAs to enter the marketplace.

Overall these strategies can be used to devise a variety development system that might bypass
GMO regulations (Figure 4). However, the actual regulatory status of most genome-edited crops is still under debate.

References


Figure legends

Figure 1. Genome engineering platforms and their utilization.

The zinc finger nuclease (ZFN) system is composed of a zinc finger domain and a FokI endonuclease domain. The trinucleotide binding sites within each zinc finger domain dictate the specificity of DNA binding, while coupling of the nuclease domains (derived from the type II restriction endonuclease FokI) of two ZFNs produces a double stranded break (DSB) at the designated target site. Transcription activator-like effector nucleases (TALENs) also contain two domains, a DNA binding domain and a FokI endonuclease domain. But in this case the DNA binding specificity is dictated by the TALE array that contains 33–34 nucleotide conserved residue repetitive motifs. These motifs contain repeated variable di-residues (RVDs) at positions twelve and thirteen that specify binding to either an A, G, C or T nucleotide in the DNA target site. Similar to ZNF, the dimerized FokI endonuclease produces a DSB at the target site. The clustered regularly interspaced short palindromic repeat (CRISPR)/CRISPR associated 9 (Cas9) system consists of Cas9 endonuclease and a single guide RNA (sgRNA). In case of CRISPR/Cas9, the DNA binding specificity is dictated by specific hybridization to the sgRNA where the two Cas9 nuclease domains produce a DSB three base pairs upstream of the protospacer adjacent motif (PAM). The repair of the DSB is either by the error-prone non-homologous end joining (NHEJ) system or by homology-directed repair (HDR). During NHEJ, random insertions or deletions can occur, which ultimately may lead to gene disruption. By contrast, HDR uses a donor DNA template that can be used to direct gene replacement/addition or gene sequence modification.
Figure 2. Molecular basis of CRISPR/Cas9-mediated genome engineering using the TRV system.

The engineered TRV genome is transcribed inside the plant cell’s nucleus, producing a single guide RNA (sgRNA) under the control of a user-defined promoter. The target plant must be genetically engineered with a Cas9 gene to constitutively produce the Cas9 endonuclease. Cas9 binds the TRV-encoded sgRNA and cleaves the plant genome at the gRNA-directed site (red arrow), three base pairs upstream of the PAM sequence. The resulting double-stranded break (DSB) is repaired by the cell’s error-prone non-homologous end joining (NHEJ) DNA repair system. As a result, certain insertions or deletions (indels) are produced, ultimately disrupting the target gene (in most cases). This engineered mutation is a permanent part of the plant’s genome and any resulting altered phenotype is transmissible to the next generation.

Figure 3. Stepwise methodology for TRV-mediated genome engineering.

In the TRV2 genome, a 20-nucleotide target sequence preceding the PAM sequence is cloned into the sgRNA gene that is under the control of an appropriate promoter. *Agrobacterium* cultures carrying DNA segments encoding the engineered TRV2 genome (conferring user-selected sequence specificity) and the TRV1 genome are co-infiltrated into the leaves of plants overexpressing Cas9 (Cas9-OE) via agroinfection. For example in figure the sgRNA (cloned in TRV2) is designed to target a malfunctioning gene makes plant leaves yellow. After agroinfection, the plants are analyzed for the presence of the targeted modification that restors the green leaf phenotype. Plant leaf discs carrying modified genomes can be regenerated to
recover mutant plants, or the seed progeny can be screened for the presence of the modification in the next generation, thereby bypassing the need for tissue culture.

Figure 4. Applications of CRISPR/Cas9 in plants

CRISPR/Cas9 can be (and has been) used to improve several plant traits, e.g., engineering biotic stress tolerance (resistance against invading viruses, bacteria, and insects), abiotic stress tolerance (resistance against drought, heat, and cold stress) and yield improvement (e.g., by increasing grain weight in wheat, kernel size in corn, and the number of bolls in cotton).

Moreover, all of this work can be performed in a “DNA-free”, non-GMO crop production system. In addition to crop improvement, the CRISPR/Cas9 system can be used to greatly accelerate fundamental plant research, e.g., by constructing extensive gene knockout libraries or by producing plants with key gene replacements or modifications.

Figure 1:
Figure 2:
Figure 3:
Cas9 overexpression plant with target genes

TRV, with gRNA for target gene/s, is agro infiltrated; gene editing starts (green leaf)

Gene targeting confirmation

Systemic movement of TRV helps editing whole plant (green leaves)

T0 seeds

Tissue culture

Tissue culture-independent genome engineering

Tissue culture-dependent genome engineering
Figure 4: