



THE UNIVERSITY *of* EDINBURGH

Edinburgh Research Explorer

Engineering of CRISPR/Cas9-mediated potyvirus resistance in transgene-free Arabidopsis plants

Citation for published version:

Pyott, D, Sheehan, E & Molnar, A 2016, 'Engineering of CRISPR/Cas9-mediated potyvirus resistance in transgene-free Arabidopsis plants' *Molecular Plant Pathology*. DOI: 10.1111/mpp.12417

Digital Object Identifier (DOI):

[10.1111/mpp.12417](https://doi.org/10.1111/mpp.12417)

Link:

[Link to publication record in Edinburgh Research Explorer](#)

Document Version:

Peer reviewed version

Published In:

Molecular Plant Pathology

General rights

Copyright for the publications made accessible via the Edinburgh Research Explorer is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy

The University of Edinburgh has made every reasonable effort to ensure that Edinburgh Research Explorer content complies with UK legislation. If you believe that the public display of this file breaches copyright please contact openaccess@ed.ac.uk providing details, and we will remove access to the work immediately and investigate your claim.



Title

Engineering of CRISPR/Cas9-mediated *potyvirus* resistance in transgene-free *Arabidopsis* plants

Authors: Douglas E. Pyott, Emma Sheehan and Attila Molnar*

*Corresponding author, Tel: +44 131 6505335, Fax: +44 131 6508650

attila.molnar@ed.ac.uk

Author affiliation: Institute of Molecular Plant Sciences, University of Edinburgh, EH9 3JR, UK

Running head: CRISPR/Cas9-induced potyvirus resistance

Key words: CRISPR/Cas9, potyvirus, eIF(iso)4E, TuMV, virus resistance, transgene-free

Word counts: 6570

Abstract

Members of the eukaryotic translation initiation factor (eIF) gene family including *eIF4E*, and its paralogue, *eIF(iso)4E*, have previously been identified as recessive resistance alleles against various *Potyruses* in a range of different hosts. However, identification and introgression of these alleles into important crop species is often limited. In this study, we utilise CRISPR/Cas9 technology to introduce sequence-specific deleterious point mutations at the *eIF(iso)4E* locus in *Arabidopsis thaliana* to successfully engineer complete resistance to Turnip Mosaic Virus (TuMV), a major pathogen in field grown vegetable crops. By segregating the induced mutation from the *CRISPR/Cas9* transgene, we outline a framework for producing heritable, homozygous mutations in transgene-free T₂ generation in self-pollinating species. Analysis of dry weights and flowering times for four independent T₃ lines revealed no differences to wild type plants under standard growth conditions, suggesting that homozygous mutations in *eIF(iso)4E* does not affect plant vigour. Thus the established CRISPR/Cas9 technology provides a new approach for generating *Potyvirus* resistance alleles in important crops without the use of persistent transgenes.

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process which may lead to differences between this version and the Version of Record. Please cite this article as an 'Accepted Article', doi: 10.1111/mpp.12417

This article is protected by copyright. All rights reserved.

Introduction

Plant viruses are ubiquitous in natural environments and can severely limit plant growth and fertility. Globally, viruses are a significant economic burden to both well-developed and under-developed agriculture due to absolute yield losses in the field, and decreased marketability of harvested crops. The *Potyvirus* family contains a greater number of virus species than any other plant virus family (Gibbs & Oshima, 2010), and certain species within this family (notably its type member, PVY) are particularly damaging to economically important crops (Karasev & Gray, 2013). *Potyviruses* exist as flexuous, rod shaped virions comprised of a positive sense, single stranded RNA (+ssRNA) genome coated by a virally encoded coat protein (Jagdish *et al.*, 1991). Upon entry into plant cells, the +ssRNA genome is uncoated and translated into a single polypeptide which subsequently generates the range of *potyviral* proteins by autocatalysis (Carrington *et al.*, 1989). Translation of the *potyviral* RNA is largely dependent on host translation factors.

In eukaryotes, translation of mRNAs is orchestrated by multi-component translation complexes composed of eukaryotic initiation factors (eIFs), which recruit ribosomes to the 5' UTR. eIF4E and eIF4G associate to form an eIF4F core complex. eIF4G acts as a scaffold protein which associates with the DEAD box RNA helicase eIF4A, and the polyA binding protein (PABP) which unwind and circularise the mRNA, respectively. eIF4E associates with the 5' m⁷GpppN cap structure which is crucial for mRNA circularisation and anchoring of the translation complex to the 5' UTR. In higher plants, gene duplication has resulted in a second eIF4F complex called eIF(iso)4F which is composed of eIF(iso)4E and eIF(iso)4G (Browning, 1996). While eIF4F and eIF(iso)4F are usually formed by cognate pairing of their respective subunits (eIF4E/eIF4G and eIF(iso)4E/eIF(iso)4G) (Bush *et al.*, 2009) functional redundancy exists such that single mutations in one complex can be compensated for by activity of the other (Duprat *et al.*, 2002).

The *potyviral* 'VPg' (viral protein genome-linked) protein is involved in the usurpation of host translation complexes to aid viral translation. This is partly achieved by VPg interacting with both the 5' UTR of the viral genome and host eIFs (such as eIF4E or eIF(iso)4E). In contrast to cellular mRNAs, which can utilise both eIF4E and eIF(iso)4E, different *potyviral* VPgs have evolved to bind specifically to one or other isoform. This exclusive binding was first demonstrated for the VPg of TuMV in a yeast two hybrid screen which showed VPg interaction with *Arabidopsis* eIF(iso)4E but not eIF4E (Wittmann *et al.*, 1997). Later it was shown that this specific requirement for eIF(iso)4E could result in TuMV resistance if both alleles of *eIF(iso)4E* are knocked out (Duprat *et al.*, 2002; Lellis *et al.*, 2002; Sato *et al.*, 2005). More broadly, this exclusivity of different *potyviral* VPg for their cognate eIF partners has formed the basis for natural, recessive resistance to a wide range of *Potyviruses* in various crops. For example; pepper *pvr2* (Ruffel *et al.*, 2002), lettuce *mo1* (Nicaise *et al.*, 2003), pea *sbm1* (Gao *et al.*, 2004), tomato *pot1* (Ruffel *et al.*, 2005), barley *rym4* (Kanyuka *et al.*, 2005;

Stein *et al.*, 2005), melon *nsv* (Nieto *et al.*, 2006) bean *bc-3* (Naderpour *et al.*, 2010), and potato *eva1* (Duan *et al.*, 2012) have all been mapped to either *eIF4E* or *eIF(iso)4E* homologues.

The field of site-specific genome editing was recently revolutionised by the discovery and characterisation of a programmable, RNA guided DNA endonuclease from *Streptococcus pyrogenes* called Cas9 which operates together with CRISPR (Clustered Regularly Interspaced Palindromic Repeats) loci in the bacterial genome as a form of defence against invading plasmids or DNA viruses, in bacteria (Jinek *et al.*, 2012). In its natural context, Cas9 associates with a so-called crRNA (CRISPR RNA) derived from the invading DNA, and a tracrRNA (trans-acting CRISPR RNA), which associate with the Cas9 protein. Cas9 is guided to the invading DNA by the crRNA via Watson-Crick base pairing, and it destroys the invading sequences by introducing double stranded breaks (DSBs) at a specific site. A breakthrough in the application of CRISPR/Cas9 came from the discovery that Cas9 can be programmed to introduce DSBs in eukaryotic cells guided by an artificial sgRNA (single guide RNA), which fuses components of the crRNA and tracrRNA (Jinek *et al.*, 2012). The only constraint on the design of the sgRNA is that the 20nt region of complementarity between the sgRNA and the target DNA must be immediately upstream of an NGG sequence (where 'N' is any base) known as a PAM (pro-sacer adjacent motif). Introduction of sgRNA-programmed Cas9 into eukaryotic cells leads to Cas9-induced DSBs in the target DNA specifically 3nt upstream of the PAM (Gasiunas *et al.*, 2012; Jinek *et al.*, 2012). As these DSBs are primarily repaired by the error-prone non-homologous end joining (NHEJ) DNA repair pathway, Cas9-induced DSBs will often result in short insertions/deletions (indels) at the site of DNA cleavage. As such, CRISPR/Cas9 technology has opened up a facile means for introducing site-specific mutations in eukaryotic genomes. Since its watershed publication, CRISPR/Cas9 technology has been adopted in various model organisms including zebrafish (Hwang *et al.*, 2013), mouse (Wang *et al.*, 2013), rat (Li *et al.*, 2013), *Arabidopsis thaliana* (Feng *et al.*, 2013), and *Nicotiana benthamiana* (Nekrasov *et al.*, 2013). Importantly, CRISPR/Cas9 has recently been used to modify genomes of several crop plants (Jiang *et al.*, 2013; Shan *et al.*, 2013; Brooks *et al.*, 2014; Liang *et al.*, 2014; Shan *et al.*, 2014; Zhou *et al.*, 2014; Cai *et al.*, 2015; Ito *et al.*, 2015; Jacobs *et al.*, 2015; Zhu *et al.*, 2016), though most of these studies have been a proof of concept and there are few examples to date where the technology has been used for crop improvement (Ito *et al.*, 2015).

As loss of function mutations in components of the eIF4F translation complex have repeatedly been associated with stable resistance to several *Potyvirus*s, we aimed to generate virus resistant plants by novel mutation at the *eIF(iso)4E* locus in *Arabidopsis thaliana* using CRISPR/Cas9 technology. Our rationale for inducing such mutations by CRISPR/Cas9 genome editing was to showcase the concept for generating virus resistance, which can be applied directly to important crops in the future. As CRISPR/Cas9 has been shown to be a viable technology for site-specific genome editing in several plant species we

believe this work will pave the way as a strategy for reverse engineering *potyviral* resistance in a wide variety of crops.

Results

Site-specific mutation of *eIF(iso)4E* by transgenic expression of a sgRNA guided Cas9

We chose to target *eIF(iso)4E* in *Arabidopsis* (At5G35620) for CRISPR/Cas9 mutagenesis as mutations caused by ethyl methanesulfonate (EMS) (Lellis *et al.*, 2002; Sato *et al.*, 2005) and transposon insertion (Duprat *et al.*, 2002) at this locus have previously been shown to result in complete resistance to several *Potyvirus*es including Turnip Mosaic Virus (TuMV) (Duprat *et al.*, 2002), Lettuce Mosaic Virus (LMV) (Duprat *et al.*, 2002), and Tobacco Etch Virus (TEV) (Lellis *et al.*, 2002). We decided to target the 5' region of the open reading frame (ORF) as mutations here would increase the likelihood of creating non-functional proteins by causing a coding frameshift or early stop codons. Specifically, we designed an sgRNA complementary to bases +15 – +35 relative to the translation start site of the gene (Figure 1). This particular region was selected because it allowed the end of the target region of the sgRNA (immediately upstream of the PAM) to end in a guanine dinucleotide (GG), which has been reported to increase the efficacy of Cas9-induced mutations at the correct target site (Farboud & Meyer, 2015). Additionally, when we queried this sgRNA for possible off-targeting, using the online search tools CRISPR-Plant and CCTop (Stemmer *et al.*, 2015), we found no off-target loci. For this we used search criteria which only qualified off-targets with fewer than 5 mismatching bases to the sgRNA and no more than 2 mismatching bases in the seed region (12 consecutive nucleotides upstream of the PAM), as these parameters have been experimentally validated as the minimum requirements for directing Cas9 to cleave DNA targets (Sternberg *et al.*, 2013; Hsu *et al.*, 2013). We cloned this sgRNA into a binary vector (pDe-CAS9) (Fauser *et al.*, 2014), which allowed for the tandem expression of a plant codon optimised Cas9 and the sgRNA (driven by *PcUbi4-2* and *AtU6-26* promoters, respectively) along with the BASTA resistance gene. This construct was introduced into *Arabidopsis* (Col-0 accession) plants by *Agrobacterium*-mediated transformation (floral dipping).

Transgenic T₁ seeds were BASTA selected, and 6 healthy looking plants were chosen to test for genome editing at the *eIF(iso)4E* locus. For this we adopted an assay, which allows edited sequences to be detected on a gel after incubation with T7 endonuclease (hereafter T7). Briefly, primers spanning the target site (Figure 1) were used to generate an *eIF(iso)4E* amplicon, which was subsequently denatured and re-annealed before incubation with T7. Bulges caused by mismatching bases are recognised by T7, which results in cleavage of the re-annealed amplicons. Hence, samples with a mixture of wild type and mutant DNA will yield cleavage products which can be resolved on a gel, whereas homogenous samples of either wild type or fully mutated DNA will be resistant to T7 activity; yielding a full, un-

cleaved target amplicon. A possible drawback of this assay is that samples with identical lesions in all DNA copies will be indistinguishable from wild type, and hence such samples would be missed in a screen for detecting mutations. However, we reasoned that this would be a highly unlikely scenario in the T₁ generation as Cas9 expression could only have been present in one of the parental gametes and as such, genome editing at this stage would most likely result in chimeric plants containing a mixture of wild type and mutant sequences. We performed the T7 assay on *eIF(iso)4E* amplicons from 6 BASTA selected T₁ transformants alongside a wild type, non-transformed control and a T₁ plant transformed with a Cas9 vector lacking a sgRNA sequence (Figure 2). T7 cleavage products were absent from both negative controls but were clearly visible in candidate plants numbered 1,3, and 6 (Figure 2C). As the signal for the T7 cleavage product was strongest in sample number 1, we selected this line for further analysis.

Segregation of the induced mutation from the transgene in the T₂ generation

We generated T₂ seeds by allowing the selected T₁ plant (line number 1) to self-pollinate. We reasoned that it would be possible to segregate stable, uniform *eIF(iso)4E* mutations from the transgene at this generation. To establish the number of integration sites for the transgene in this line, we sprayed approximately 200 of these T₂ seedlings from the T₁ line number 1 with BASTA. The BASTA resistant:susceptible ratio was approximately 3:1, indicative of a single integration site of the transgene. In parallel, with a separate batch of segregating T₂ seedlings, we utilised a multiplex PCR screen to identify T₂ plants, which had lost the *Cas9/sgRNA* expressing transgene. The general strategy to segregate the transgene from the induced mutation is outlined in Figure 3. We tested 144 plants in this way and recovered 55 non-transgenic plants (Figure 4). Next, we tested this population of transgene-free plants for *eIF(iso)4E* mutations. As we hoped to identify homozygous mutants, the T7 assay was inappropriate for this analysis and we instead opted to directly sequence *eIF(iso)4E* amplicons for each of the 55 lines by Sanger sequencing. By carefully checking the quality of the base calls at the predicted mutation site we were able to distinguish homozygous from heterozygous mutations, as the quality score for the latter would show a sharp fall at the mutation site due to mixed signal base calling (Supplemental Figure 1, bottom panel). Of the 55 non-transgenic plants, 39 (70.9%) harboured mutations in *eIF(iso)4E*, and 4 of these mutations were found to be homozygous (Figure 5A, Supplemental Figure 1). Interestingly the majority of all mutations were single nucleotide indels, with the exception of one heterozygous mutant, which had a 57bp deletion, spanning the predicted Cas9 cleavage site (Supplemental Figure 2). This over-representation of single nucleotide indels is consistent with previous reports for Cas9-induced editing in plants (Nekrasov *et al.*, 2013). Importantly, all of the indels aligned perfectly with the expected Cas9 cleavage site, 3bp upstream of the PAM (Supplemental Figure 1). Furthermore, mutations of interest were confirmed by re-sequencing the *eIF(iso)4E*

amplicons from the opposite direction to rule out the unlikely possibility of false positives due to sequencing errors. Each of the homozygous mutants have early stop codons in their predicted amino acid sequences (Figure 5B), which we reasoned would create complete functional knock-outs by severe truncation of the *eIF(iso)4E* protein. These 4 homozygous mutants (named #44, #65, #68 and #98) were self-pollinated to produce T₃ populations for each of the different *eIF(iso)4E* point mutations. A non-transgenic T₂ plant with homozygous wild type *eIF(iso)4E* alleles (#105) was also selected to produce a wild type T₃ population.

The induced mutation in *eIF(iso)4E* confers complete resistance to TuMV

The 4 separate T₃ populations (#44, #65, #68 and #98) harbouring each of the induced *eIF(iso)4E* mutations were grown alongside the T₃ wild type control (#105) and a previously published transposon insertion *eIF(iso)4E* mutant, which is known to be resistant to TuMV infection (Duprat *et al.*, 2002). After 4 weeks of growth, 40 plants from each of the 6 different genotypes (#44, #65, #68, #98, #105 and the transposon insertion mutant) were rub-inoculated with a GFP expressing TuMV clone (hereafter TuMV-GFP) (Garcia-Ruiz *et al.*, 2010). At 7 and 14 days post inoculation (dpi) TuMV infection was assessed by monitoring the expression of GFP in inoculated and systemic leaves (Figure 6A). GFP expression (indicative of TuMV-GFP infection) was clearly visible at 14dpi in 37/40 (92.5%) of the wild type plants (#105) but not in any of the CRISPR/Cas9-induced mutants (#44, #65, #68 and #98) (Supplemental Figure 3) nor the previously reported transposon insertion mutant. To ascertain that the lack of GFP signal in the *eIF(iso)4E* mutants was due to complete TuMV-GFP resistance, and not just lower viral titres, inoculated and systemic leaves for each genotype were analysed by RT-PCR amplifying the coat protein coding region of TuMV-GFP. TuMV-specific amplicons were clearly detected in both inoculated and systemic leaves from 3 wild type plants, but no TuMV-- amplicons were detected in any samples from any of the *eIF(iso)4E* mutants (Figure 6B). To gain a more quantitative insight into the viral titres in the inoculated plants, qRT-PCR was performed for the inoculated and systemic leaves at 7dpi (Figure 6C). A serial dilution of purified TuMV-GFP RNA was included in the qRT-PCR to construct a standard curve (Supplemental Figure 4) from which we could interpolate the titre (in pg) of TuMV-GFP in the 7dpi samples. In agreement with the RT-PCR results (Figure 6B) we were able to detect TuMV-GFP in all of the wild type samples but not in systemic leaves of the *eIF(iso)4E* mutants. Interestingly, we measured a very low titre of TuMV-GFP in the inoculated leaves of the *eIF(iso)4E* mutants. We suspect that this TuMV-GFP signal may be due to residual infectious sap from the rub-inoculation, though it is also possible that a low level of TuMV replication occurs in the inoculated leaves of *eIF(iso)4E* mutants. To stringently test the TuMV resistance of the *eIF(iso)4E* mutants, 20 systemic leaves for each genotype were collected at 28dpi and pooled to make sap for back-inoculating *Nicotiana benthamiana* plants, which are highly susceptible to TuMV and hence would reveal the

presence of very low TuMV-GFP titres present in the sap. 5 days after the back-inoculations, TuMV-GFP infection was clearly visible in both local and systemic leaves inoculated with sap taken from wild type *Arabidopsis* plants. In contrast, TuMV-GFP was not detected at all in any of the back-inoculations using sap prepared from the *eIF(iso)4E* mutants (Figure 7). This indicates that the ablation of the *eIF(iso)4E* protein by the induced single nucleotide mutations in lines #44, #65, #68 and #98 render the plants completely resistant to TuMV-GFP in a similar way to the previously published transposon mutant (Duprat *et al.*, 2002).

The *eIF(iso)4E* mutants show no growth defects compared to wild type plants, when grown under standard growth conditions.

We noted that the *eIF(iso)4E* mutants looked indistinguishable from the wild type plants as they grew. To more accurately assess the growth vigour of the *eIF(iso)4E* mutants in comparison to wild type plants, we measured the dry mass and flowering times for non-infected populations of each of the 4 homozygous *eIF(iso)4E* mutant lines (#44, #65, #68, #98) and a wild type line (#105) in the T₃ generation. For both the dry weight and flowering time experiments seeds of each genotype were grown in a randomised block design. 10 plants from each genotype were grown together in a tray, with 6 replicate trays (resulting in 60 plants per genotype, spread across 6 different trays). The relative position of the different genotypes within any tray was randomised. These measures were taken to avoid systematic biases of the growth environment affecting one genotype more than another and hence confounding the analysis. After 4 weeks of growth, just prior to the onset of flowering, 30 plants from each set were randomly sampled to estimate the dry weights of each genotype. The remaining plants were scored for flowering emergence by counting the number of days from germination to the first appearance of a floral bolt. From these experiments, we can conclude that there are no statistically significant differences in the growth of *eIF(iso)4E* mutants and wild type plants, with respect to total dry mass ($F_{4,70} = 1.372$, $p=0.252$) or flowering time ($F_{4,119} = 1.597$, $p=0.180$) (Figure 8, Supplemental table 2 and 3). From this, we surmise that this strategy of site-specific disruption of *eIF(iso)4E* will be useful for generating virus resistant crops without concomitant constraints on plant growth.

Discussion

In the context of a rapidly growing global population, food security is one of the major challenges facing our generation. It is estimated that food production will need to at least double in the period between 2005 and 2050 (Tilman *et al.*, 2011) to meet demand. As viral infections in crops are estimated to result in approximately 10-15% global yield reductions each year (Regenmortel & Mahy, 2009), mitigation of these losses by improved viral

resistance is a worthwhile strategy for meeting global yield targets. Arguably, utilising genetic resistance in crops is the most sustainable approach for controlling virus infections – other methods such as pesticides to control insect vectors or manual inspection and removal of infected plants are costly, laborious, and often ineffective. There are several approaches for introducing resistance genes (or R genes) into crops. Firstly, classical breeding can be used to identify and introgress R genes from related species to the crop of interest. Often wild relatives of the domesticated crop are used for these breeding programmes and successive back-crosses are required to segregate out additional non-beneficial (or even deleterious) alleles from the hybrid. Hence breeding programmes can be very costly and laborious, and take many years to achieve the desired elite cultivar. Breeding programmes can be additionally limited by constraints such as hybrid incompatibility or outbreeding depression.

Biotechnological approaches can also be adopted to generate R genes in crops. A popular platform for introducing and selecting beneficial mutations in crops is TILLING (Targeting Induced Local Lesions IN Genomes). This relies on non-site-specific mutation of the genome of interest, usually by EMS treatment (McCallum *et al.*, 2000). The induced mutations can be subsequently identified and screened for positive attributes, such as viral resistance (Piron *et al.*, 2010). This strategy overcomes some of the pitfalls of classical breeding as it generates novel genetic diversity within the species of interest. However, this approach is also costly, time consuming, and slow to produce elite cultivars. Furthermore, off-target mutations due to the random nature of the mutagenesis can result in unintended deleterious effects in the new cultivars. A second biotechnological approach to generate viral resistance is the stable integration of transgenes into crops. This transgenic approach has been utilised in a variety of ways including; overexpression of dominant R genes identified in other species (Oldroyd *et al.*, 1998), overexpression of non-coding sequences to direct RNA silencing of specific viruses (Asad *et al.*, 2003 ; Poogin *et al.*, 2003 ; Yang *et al.*, 2004 ; Bonfim *et al.*, 2007), overexpression of null alleles to sequester functional host susceptibility factors in a dominant negative manner (Cavatorta *et al.*, 2011), and stable expression of a *Cas9/sgRNA* cassette to target DNA viruses directly for CRISPR/Cas9-mediated cleavage (Ali *et al.*, 2015). While these methods can be effective for introducing viral resistance into crops, transformation of plants with transgenes can be costly and laborious and limited to only certain species or varieties with well-established transformation protocols. Moreover, many governments are opposed to the use of transgenic products, which seriously undermines the application of these technologies for improving crop yields in the near future. More recently, site-specific genome editing technologies have offered new ways to improve crops. One attractive feature of these technologies is that, once the desired genome alterations have been made, the transgenes can be crossed out from the improved variety, thus circumventing public and political concerns around the use of persistent transgenes in crops. Early genome editing technologies include TALENs (Transcription Activator-Like Effector Nucleases) and ZFNs (Zinc

Finger Nucleases) (Gaj *et al.*, 2013). Both TALENs and ZFNs combine DNA nucleases (such as FokI) with a DNA binding protein to induce DNA DSBs at specific site. A major limitation of these technologies is that tailoring the DNA binding proteins to target a sequence of interest can be costly and time-consuming. Furthermore, it is only possible to engineer DNA binding proteins for certain DNA target sequences. The advent of CRISPR/Cas9 technology has revolutionised the field of genome editing. Crucially, the fact that the Cas9 nuclease is guided by RNA rather than protein overcomes the major limitations of TALEN and ZFN technologies. RNA-based guiding is cheaper and easier to engineer and the range of possible target sequences is greatly expanded, requiring only the commonly occurring NGG PAM sequence.

In this study we show-case the utility of CRISPR/Cas9 technology for generating novel genetic resistance to TuMV in *Arabidopsis* by deletion of a host factor (*eIF(iso)4E*) which is strictly required for viral survival. We hope that this work will pave the way for a similar strategy to be adopted in important crop species and thus provide an alternative, novel strategy for the introduction of R genes. Credence for this approach is given by the fact that many natural sources of *Potyvirus* resistance rely on the same principle: loss of function mutations in host eIFs (Ruffel *et al.*, 2002; Nicaise *et al.*, 2003; Gao *et al.*, 2004; Ruffel *et al.*, 2005; Kanyuka *et al.*, 2005; Stein *et al.*, 2005; Nieto *et al.*, 2006; Naderpour *et al.*, 2010; Duan *et al.*, 2012). Therefore, we believe that it would be hard to justify objections to commercial application of such a strategy, as the final genome-edited product is essentially no different to varieties carrying mutant alleles arisen from 'natural' methods of mutagenesis. It is noteworthy that the engineered viral resistance reported here is the result of a single nucleotide point mutation arisen from the plant's own natural DNA damage repair mechanism, namely NHEJ. Moreover, while our approach was to use transgenic delivery of the *CRISPR/Cas9* cassette, we have shown that it is feasible to segregate out the transgene from the induced mutation at the target *eIF(iso)4E* locus at an early stage to produce stable, heritable point mutations without a persistent transgene. The fact that CRISPR/Cas9 induces sequence specific mutations, and that there were no detected off-targets in this study, means that we can be confident that the genetic differences between the wild type and mutant plants are solely at the *eIF(iso)4E* locus. This results in a better system for investigating the effects of the mutation on plant growth as previous studies using EMS (Lellis *et al.*, 2002; Sato *et al.*, 2005) or transposon insertion (Duprat *et al.*, 2002) mutagenesis are likely to be confounded by multiple off-target genomic mutations. Our analysis did not reveal any significant differences in the growth and development of the *eIF(iso)4E* mutants compared to wild type plants. It is still possible that under certain growth environments, particularly stress conditions, the *eIF(iso)4E* mutants may grow differently to wild type plants, though this is entirely speculative and requires further investigation. The durability of this engineered resistance also remains to be tested. It is assumed that recessive resistance arising from the loss of a host factor required by the virus will be more durable than dominant R genes, due to lower selective pressures on the

virus to evolve counter defence strategies (Ronde *et al.*, 2014). However, resistance breaking has been previously reported for recessive *eIF(iso)4E* resistance to TuMV, possibly due to VPg polymorphisms acting via an *eIF(iso)4E* independent pathway (Gallois *et al.*, 2010). It remains to be seen whether this resistance breaking will also occur in the CRISPR/Cas9-induced *eIF(iso)4E* mutants in this study. We hypothesise that our approach will result in a more durable resistance due to the complete absence of *eIF(iso)4E* protein. We hope that the coming years will provide more detailed analysis to this end, and will eventually lead to the introduction of this technique for a variety of marketable crops. Indeed, while preparing this manuscript, a similar approach was adopted to generate viral resistance in cucumber by CRISPR/Cas9-induced mutation of *eIF4E* (Chandrasekaran *et al.*, accepted article). Hence utilisation of CRISPR/Cas9 technology may provide an efficient and publicly acceptable method for crop improvement in the future.

Materials and Methods

Plant growth conditions

For the growth and development experiments, plants were grown in 6cmx6cmx8cm (length x width x depth) pots on Levington F2+S professional growth compost. Plants were grown in growth rooms at 21°C with 16:8 hour light:dark cycles under cool, white fluorescent bulbs at approximately 100 $\mu\text{Molesm}^{-2}\text{s}^{-1}$ light intensity. T₁ seed were grown under the same growth conditions in 50cmx25cmx6cm (length x width x depth) trays filled with Levington F2+S professional growth compost. For all other experiments, plants were grown in 4cmx4cmx5cm (length x width x depth) pots on Levington F2+S professional growth compost. Plants were grown in SANYO/Panasonic growth chambers set to 21°C with 16:8 hour light:dark cycles with side illumination from cool, white fluorescent bulbs at approximately 200 $\mu\text{Molesm}^{-2}\text{s}^{-1}$ light intensity. All seeds were stratified for 48 hours in darkness at 4°C prior to planting.

Guide RNA design and cloning

sgRNA was designed by CRISPR Design (<http://crispr.mit.edu/>) and DNA2.0 CRISPR gRNA design tool (<https://www.dna20.com/eCommerce/cas9/input>). The corresponding oligos Iso_Fw and Iso_Rv (Supplemental Table 1) were annealed and cloned into BbsI digested entry vector pEn-Chimera (Fauser *et al.*, 2014). After sequence verification, the sgRNA expression cassette was recombined into the destination vector pDe-Cas9 by Gateway cloning (LR reaction) according to the manufacturer's instructions (Life Technologies) resulting in pDe-Cas9-sgAtelF(iso)4E.

Plant transformation

Agrobacterium tumefaciens (strain AGL1) cells were transformed with the pDE-CAS9 and pDE-CAS9-sgAtelF(iso)4E plasmids by electroporation and transformed cultures were grown overnight in LB liquid medium at 28°C with 230rpm shaking to a final OD₆₀₀ of 1.0. Cells were pelleted by centrifugation at 4000rpm for 20 minutes. Pelleted cells were washed then resuspended in liquid LB and the OD₆₀₀ was adjusted to 0.8. Silwet was added to the culture at a final concentration of 0.1% (v/v) and Acetosyringone was added to a final concentration of 150µM. 6 week old flowering *Arabidopsis* (Col-0 accession) plants were dipped in the *Agrobacterium* suspension for approximately 10s. This process was repeated 7 days later.

BASTA selection

T₁ seeds were collected and scattered evenly on compost filled 70cmx30cmx6cm (length x width x depth) trays. 7 days after germination, seedlings were sprayed with a 120mg/l solution of BASTA by applying a fine mist of the herbicide across the entire area of the tray, using a hand held spray bottle. The BASTA spray treatment was repeated at 14 and 21 days after germination. Plants were grown in a temperature controlled glasshouse at 21°C with 16:8 light:dark cycles and approximately 300µMolesm⁻²s⁻¹ light intensity provided by halogen light bulbs and daylight. BASTA resistant plants were transplanted into 6cmx6cmx8cm (length x width x depth) pots to produce T₂ seed.

PCR conditions

For preparing *eIF(iso)4E* PCR amplicons for the T7 assay and sanger sequencing, Q5® High fidelity DNA polymerase (NEB #M0491S) was used with the following reaction conditions: 95°C for 120s, (95°C for 15s, 57°C for 15s, 72°C for 30s) x35, 72°C for 300s. All other PCRs were performed using Taq DNA polymerase (NEB #M0267L) with the following reaction conditions: 95°C for 120s, (95°C for 20s, T_a°C for 30s, 72°C for 30s) x35, 72°C for 300s.

Where 'T_a' is the primer-pair-specific annealing temperature listed in Supplemental table 1.

T7 endonuclease assay

For the denaturing/annealing reaction the total Q5 PCR product (10µl) was mixed with 1.5µl 10X NEB buffer 2 (NEB #B7002S) and 1.5µl water. This mixture was incubated at 95°C for 10 minutes, then ramped from 95°C-85°C at a rate of -2°C s⁻¹, then from 85°C-25°C at a rate of -0.3°C s⁻¹. T7 endonuclease I (NEB #M0302S) was diluted to a concentration of 2U µl⁻¹ (in NEB buffer 2 #B7002S) and 2µl of this was added to the denaturing/annealing reaction product, then incubated at 37°C for 1 hour. The total T7 digestion product (15µl) was loaded onto a 2% agarose gel and separated at 100V.

Sanger sequencing

eIF(iso)4E amplicons were generated as described above. The PCR products were purified using a MinElute PCR Purification kit (Qiagen). The cleaned products were sequenced

directly with either eIF(iso)4E_Fw or eIF(iso)4E_Rv primers (Supplemental table 1) using BigDye (Applied Biosystems) according to the manufacturer's protocols.

Viral inoculations

Infectious TuMV-GFP sap was prepared by syringe infiltrating 4 week old *Nicotiana benthamiana* plants with *Agrobacterium tumefaciens* (strain AGL1) cells containing the pCB-TuMV-GFP plasmid (Garcia-Ruiz *et al.*, 2010). *Agrobacterium* cells were grown as described above and resuspended in infiltration medium containing 10mM MES (pH 5.6), 10mM MgCl₂, and 150μM Acetosyringone. The cells were incubated at room temperature without shaking for 1 hour before syringe infiltration into the abaxial surface of 3 fully expanded leaves. 14 days after infiltration, systemic leaves showing viral symptoms were harvested and homogenised with a sterile mortar and pestle. The homogenate was diluted 1:5 (w/v) in 1mM sodium phosphate buffer (pH7) and frozen in 0.5ml aliquots at -80°C.

Arabidopsis/N.benthamiana leaves were rub-inoculated on the adaxial surface with 10μl of the 1:5 diluted viral sap aliquots, using aluminium oxide powder as an abrasive. The third and fourth oldest rosette leaves of 4 week old *Arabidopsis* plants were used for rub-inoculations. A second and third round of inoculations was performed 3 and 7 days after the initial inoculation using the fourth/fifth and sixth/seventh oldest leaves, respectively.

Viral GFP imaging and RT-PCR/qRT-PCR

Viral expression of GFP was monitored using hand-held UV lamps (UVP B-100AP Lamp 100W 365nm) and imaged using a Canon powershot G16 digital camera.

Total nucleic acids (TNA) were purified from plant tissue by phenol/chloroform extraction as previously described (White & Kaper, 1989). 3μg of TNA was DNase treated using Turbo DNase kit (Ambion) according to the manufacturer's protocol. 1μg of DNase treated RNA was used for cDNA synthesis using SuperscriptII reverse transcriptase (Invitrogen) according to the manufacturer's protocol. 1μl of cDNA was used for the PCR reactions with TuMV and Ef1a-specific primers as described above.

For qRT-PCR, 2ul of cDNA was used in a 10μl reaction with LightCycler® 480 SYBR Green I Master mix, according to the manufacturer's protocol. TuMV-specific primers (Supplemental table 1) were used at a final concentration of 1μM. qRT-PCR was performed in a LightCycler® 480 machine with the following amplification reaction conditions: 95°C for 300s, (95°C for 10s, 60°C for 10s, 72°C for 10s) x50.

Dry weight measurements

4 week old *Arabidopsis* plants were harvested for dry weight measurements. Whole plants were prepared by carefully removing the plants and soil from the pots and soaking the plant roots in water to wash the soil from the root system. 2 plants per genotype were pooled and placed in pre-weighed waxed paper cups, giving 15 replicates of 2 pooled plants per

genotype. The plants were dried in a Binder drying and heating chamber (Model E 28) set to 100°C for 48 hours. The difference in weight between the empty paper cups and the paper cups with dried plants was used to estimate the dry mass of the 2 pooled plants.

Statistical analysis

R software was used to perform the one way ANOVAs.

Acknowledgements

DEP is supported by the BBSRC EastBio DTP award BB/J01446X/1 and the James Hutton Institute studentship. AM is a Chancellor's Fellow at the University of Edinburgh. The authors have no conflict of interest.

We would like to thank James Carrington for providing the pCB-TUMV-GFP infectious clone, and Christophe Robaglia for providing the *eIF(iso)4E* transposon insertion seeds. We thank Karen Halliday, Lesley Torrance, and Graham Cowan for discussions.

Figure legends

Figure 1: Schematic of the *eIF(iso)4E* locus targeted for editing by CRISPR/Cas9. Primers flanking the target site are shown by half arrows over the 5' untranslated region (UTR) and open reading frame (ORF) shaded grey and blue, respectively. Scale bar depicts the approximate positions, in base pairs, relative to the translation start site (+1). The enlarged area in red dashes shows the position and sequence of the sgRNA, with the protospacer adjacent motif (PAM) underlined. The position of the sgRNA relative to the +1 of the ORF is indicated by an arrow. The Cas9 cleavage site is indicated by an arrow head.

Figure 2: PCR/T7 endonuclease products for six independently transformed plants (T_1 generation) containing a Cas9 transgene with a sgRNA targeting the *eIF(iso)4E* locus. Transformation with a *Cas9* transgene with no sgRNA (Cas9 only) and a non-transformed wild type (WT) plant were used for controls. L denotes a 100bp DNA ladder. **A** 1% agarose gel showing multiplex PCR products confirming the presence/absence of the *Cas9* transgene, using the constitutively expressed house-keeping gene *EF1a* as a loading control. **B** 1 % agarose gel showing PCR amplicons spanning the putative mutation site at the *eIF(iso)4E* locus. **C** 2% agarose gel showing *eIF(iso)4E* cleavage products after a self annealing reaction of the PCR amplicon, and subsequent digestion with T7 endonuclease. Presence of the 378bp cleavage product in samples 1, 3, and 6 is indicative of Cas9-induced mutation at the *eIF(iso)4E* locus (the corresponding 196bp cleavage product is not visible on the gel).

Figure 3: Schematic of project work-flow. *Arabidopsis* Col-0 plants were transformed by floral dipping with the *Cas9/sgRNA* recombinant binary vector containing a BASTA resistance gene [pDe-Cas9-*sgAtelF(iso)4E*]. After self-pollination, the seeds were collected and germinated in soil. Plants carrying the transgenic construct (red circle) were selected in the T₁ generation by spraying with BASTA. The transgenic plants were then tested for *eIF(iso)4E* mutation (M) to identify plants with an active CRISPR/Cas9 nuclease using a T7 assay. One transgenic T₁ plant with clear signs of *eIF(iso)4E* editing was used to produce the T₂ generation. PCR was used to identify T₂ generation plants which had lost the transgene by Mendelian segregation. The non-transgenic T₂ plants were then screened for *eIF(iso)4E* mutations by Sanger sequencing (please note: the mutations depicted as 'M' in the diagram are not identical, as the mutations in the T₁ generation occurred in somatic cells so were not heritable, and different mutations were recovered in the T₂ generation due to independent editing events in the germline of T₁. For simplicity 'M' was used to depict all CRISPR/Cas9-induced mutations). Non-transgenic T₂ plants, which were homozygous for either the mutated or wild type *eIF(iso)4e* alleles, were used to produce T₃ populations, which were then tested for viral resistance.

Figure 4: Representative 1% agarose gel for the selection of T₂ candidates lacking the *Cas9* transgene. Multiplex PCR was used to confirm the presence/absence of the *Cas9* transgene, using the constitutively expressed housekeeping gene *EF1a* as a loading control. L denotes a 100bp DNA ladder. A *Cas9* transformant (T₁ generation) and a non-transformed wild type plant were used as positive and negative controls for *Cas9* amplification, respectively. Samples #41 - #49 are a representative selection of T₂ progeny from T₁ plant number 1 (as shown in Figure 2). Candidates #44 and #45 represent two out of a total of 55 candidates lacking the *Cas9* transgene, which were selected by this method.

Figure 5: Summary of CRISPR/Cas9-induced *eIF(iso)4E* mutations. **A** DNA sequence alignments for the four homozygous *eIF(iso)4E* mutants (#44, #65, #68, #98) identified in the T₂ generation, along with a wild type (WT) control. Lines #65, #68, and #98 exhibit single nucleotide insertions while line #44 has a single nucleotide deletion. **B** Predicted amino acid sequence alignments for the four homozygous mutants and the wild type consensus. Each of the mutant alleles code for severely truncated and disrupted proteins.

Figure 6: **A** Representative photos of TuMV-GFP infected plants, imaged under UV light, 7 days post infection. A transposon-induced *eIF(iso)4E* mutant (Tn) was used as a resistant control. An enlarged image of inoculated (left) and systemic (right) leaves are shown below each rosette. **B** RT-PCR to detect the presence of TuMV-GFP in inoculated (I) and systemic (S) leaves for each genotype. Three separate plants were analysed per genotype. The first two lanes for each genotype correspond to the leaves imaged in panel A. Amplicons of the TuMV coat protein region (537bp) and the housekeeping gene *EF1a* (418bp) were PCR amplified separately from the same cDNA, then mixed and run together on a 2% agarose gel. L denotes a 100bp DNA ladder. TuMV-specific amplicons are clearly visible in each of

the wild type (WT) samples but completely absent from any of the *eIF(iso)4E* mutant samples. **C** qRT-PCR to detect the mean absolute viral titre (in pg) for the samples shown in panels **A** and **B**. qRT-PCR reactions were performed with cDNA from a healthy plant (H) and water (W) as negative controls (NC). Error bars show SEM of 3 biological replicates.

Figure 7: Back-inoculations of *Nicotiana benthamiana* plants using sap from TuMV-GFP inoculated *Arabidopsis*. Sap was prepared by pooling 20 systemic leaves from TuMV-GFP inoculated *Arabidopsis* plants as shown in Supplemental Figure 3. Labels above each quadrant refer to the genotype [Wild type (WT), #44, #65, #68, #98 and Transposon (Tn)] of the inoculated *Arabidopsis* used to make sap. Each quadrant shows an inoculated leaf (I) and systemic tissue (S) for two replicate plants, imaged under UV light. A 1cm scale bar is shown for each image.

Figure 8: Box plots of dry weights (**A**) and flowering times (**B**) for the CRISPR/Cas9-edited *eIF(iso)4E* mutants (lines #44,#65,#68 and #98) alongside wild type (WT) plant (#105).

References

- Ali, Z., Abulfaraj, A., Idris, A., Ali, S., Tashkandi, M., & Mahfouz, M. M. (2015). CRISPR / Cas9-mediated viral interference in plants. *Genome Biology*, 1–11. doi:10.1186/s13059-015-0799-6
- Asad, S., Haris, W. A. A., Bashir, A., Zafar, Y., Malik, K. A., Malik, N. N., & Lichtenstein, C. P. (2003). Transgenic tobacco expressing geminiviral RNAs are resistant to the serious viral pathogen causing cotton leaf curl disease. *Arch Virol* 2341–2352. doi:10.1007/s00705-003-0179-5
- Bonfim, K., Faria, J. C., Nogueira, E. O. P. L., Mendes, É. A., Aragão, F. J. L., Recursos, E., ... Brasília, U. De. (2007). RNAi-Mediated Resistance to Bean golden mosaic virus in Genetically Engineered Common Bean (*Phaseolus vulgaris*), *Molecular Plant Microbe Interactions* 20(6), 717–726.
- Brooks, C., Nekrasov, V., Lippman, Z. B., & Eck, J. Van. (2014). Efficient Gene Editing in Tomato in the First Generation Using the Clustered Regularly Interspaced Short Palindromic Repeats / CRISPR-Associated9 System 1. *Plant Physiol* 166(November), 1292–1297. doi:10.1104/pp.114.247577
- Browning, K.S. (1996) The plant translational apparatus. *Plant Molecular Biology*, 32, 107-144.
- Bush, M. S., Hutchins, A. P., Jones, A. M. E., Naldrett, M. J., Jarmolowski, A., Lloyd, C. W., & Doonan, J. H. (2009). Selective recruitment of proteins to 5' cap complexes during the growth cycle in *Arabidopsis*. *The Plant Journal* 400–412. doi:10.1111/j.1365-313X.2009.03882.x

Cai, Y., Chen, L., Liu, X., Sun, S., Wu, C., & Jiang, B. (2015). CRISPR / Cas9-Mediated Genome Editing in Soybean Hairy Roots. *PloS one*, 1–13. doi:10.1371/journal.pone.0136064

Carrington, J. C., Freed, D. D., & Sanders, T. C. (1989). Autocatalytic Processing of the Potyvirus Helper Component Proteinase in *Escherichia coli* and In Vitro Cl-I. *Journal of Virology*, 63(10), 4459–4463.

Cavatorta, J., Perez, K. W., Gray, S. M., Van Eck, J., Yeam, I., & Jahn, M. (2011). Engineering virus resistance using a modified potato gene. *Plant Biotechnology Journal*, 9(9), 1014–21. doi:10.1111/j.1467-7652.2011.00622.x

Chandrasekaran, J., Brumin, M., Wolf, D., Klap, C., Sherman, A., & Arazi, T. (Accepted article) Development of broad virus resistance in non-transgenic cucumber using CRISPR / Cas9 technology. *Molecular Plant Pathology*.

Duan, H., Richael, C., & Rommens, C. M. (2012). Overexpression of the wild potato eIF4E-1 variant Eva1 elicits Potato virus Y resistance in plants silenced for native eIF4E-1. *Transgenic Research*, 21(5), 929–38. doi:10.1007/s11248-011-9576-9

Duprat, A., Caranta, C., Revers, F., Menand, B., Browning, K. S., & Robaglia, C. (2002). The Arabidopsis eukaryotic initiation factor (iso)4E is dispensable for plant growth but required for susceptibility to potyviruses. *The Plant Journal*, 32(6), 927–934. doi:10.1046/j.1365-313X.2002.01481.x

Farboud, B. & Meyer, B. (2015). Dramatic Enhancement of Genome Editing by CRISPR / Cas9 Through Improved Guide RNA Design, *Genetics* 199(4), 959–971. doi:10.1534/genetics.115.175166

Fausser, F., Schiml, S., & Puchta, H. (2014). Both CRISPR / Cas-based nucleases and nickases can be used efficiently for genome engineering in *Arabidopsis thaliana*. *The Plant Journal*, 348–359. doi:10.1111/tpj.12554

Feng, Z., Zhang, B., Ding, W., Liu, X., Yang, D.-L., Wei, P., ... Zhu, J.-K. (2013). Efficient genome editing in plants using a CRISPR/Cas system. *Cell Research*, 1–4. doi:10.1038/cr.2013.114

Gaj, T., Gersbach, C. A., & Barbas, C. F. (2013). ZFN, TALEN, and CRISPR / Cas-based methods for genome engineering. *Trends in Biotechnology*, 31(7), 397–405. doi:10.1016/j.tibtech.2013.04.004

Garcia-Ruiz, H., Takeda, A., Chapman, E. J., Sullivan, C. M., Fahlgren, N., Bremel, K. J., & Carrington, J. C. (2010). Arabidopsis RNA-dependent RNA polymerases and dicer-like proteins in antiviral defense and small interfering RNA biogenesis during Turnip Mosaic Virus infection. *The Plant Cell*, 22(2), 481–96. doi:10.1105/tpc.109.073056

Gasiunas, G., Barrangou, R., Horvath, P., & Siksnys, V. (2012). Cas9 – crRNA ribonucleoprotein complex mediates specific DNA cleavage for adaptive immunity in bacteria. *Proc. Natl. Acad. Sci. USA*, 109(39), 2579–2586. doi:10.1073/pnas.1208507109

Gallois, J.-L., Charron, C., Sánchez, F., Pagny, G., Houvenaghel, M.-C., Moretti, A., ... German-Retana, S. (2010). Single amino acid changes in the turnip mosaic virus viral genome-linked protein (VPg) confer virulence towards *Arabidopsis thaliana* mutants knocked out for eukaryotic initiation factors eIF(iso)4E and eIF(iso)4G. *The Journal of General Virology*, *91*(Pt 1), 288–93. doi:10.1099/vir.0.015321-0

Gao, Z., Johansen, E., Eyers, S., Thomas, C. L., Ellis, T. H. N., Maule, A. J., ... Nr, N. (2004). The potyvirus recessive resistance gene, *sbm1*, identifies a novel role for translation initiation factor eIF4E in cell-to-cell trafficking. *The Plant Journal*, *4*, 376–385. doi:10.1111/j.1365-313X.2004.02215.x

Gibbs, A., & Ohshima, K. (2010). Potyviruses and the Digital Revolution. *Annual Review of Phytopathology*, *48*, 205–23 doi:10.1146/annurev-phyto-073009

Hsu, P. D., Scott, D. A., Weinstein, J. A., Ran, F. A., Konermann, S., Agarwala, V., ... Zhang, F. (2013). DNA targeting specificity of RNA-guided Cas9 nucleases. *Nature Biotechnology*, *31*(9). doi:10.1038/nbt.2647

Hwang, W. Y., Fu, Y., Reyon, D., Maeder, M. L., Tsai, S. Q., Sander, J. D., ... Joung, J. K. (2013). Efficient genome editing in zebrafish using a CRISPR-Cas system. *Nature Biotechnology*, *31*(3), 227–229. doi:10.1038/nbt.2501

Ito, Y., Nishizawa-yokoi, A., Endo, M., & Mikami, M. (2015). Biochemical and Biophysical Research Communications CRISPR / Cas9-mediated mutagenesis of the RIN locus that regulates tomato fruit ripening. *Biochemical and Biophysical Research Communications*, *467*(1), 76–82. doi:10.1016/j.bbrc.2015.09.117

Jacobs, T. B., Lafayette, P. R., Schmitz, R. J., & Parrott, W. A. (2015). Targeted genome modifications in soybean with CRISPR/Cas9. *BMC Biotechnology*, *15*:16. doi:10.1186/s12896-015-0131-2

Jagadish, M. N., Ward, C. W., Gough, K. H., Tulloch, P. A., Whittaker, L. A., & Shukla, D. D. (1991). Expression of potyvirus coat protein in *Escherichia coli* and yeast and its assembly into virus-like particles. *Journal of General Virology*, *72*, 1543–1550.

Jiang, W., Zhou, H., Bi, H., Fromm, M., Yang, B., & Weeks, D. P. (2013). Demonstration of CRISPR/Cas9/sgRNA-mediated targeted gene modification in *Arabidopsis*, tobacco, sorghum and rice. *Nucleic Acids Research*, 1–12. doi:10.1093/nar/gkt780

Jinek, M., Chylinski, K., Fonfara, I., Hauer, M., Doudna, J. A., & Charpentier, E. (2012). A Programmable Dual-RNA – Guided DNA Endonuclease in Adaptive Bacterial Immunity. *Science*, *337*(August), 816–822.

Kanyuka, K., Druka, A., Caldwell, D.G., Tymon, A., McCallum, N., Waugh, R & Adams, M.J. (2005). Evidence that the recessive bymovirus resistance locus *rym 4* in barley corresponds to the eukaryotic translation initiation factor 4E. *Molecular Plant Pathology*, *6*, 449–458. doi:10.1111/J.1364-3703.2005.00294.X

Karasev, A. V., & Gray, S. M. (2013). Continuous and Emerging Challenges of Potato virus Y in Potato. *Annual Review of Phytopathology*, *51*, 571–86. doi:10.1146/annurev-phyto-082712-102332

Lellis, A. D., Kasschau, K. D., Whitham, S. A., & Carrington, J. C. (2002). Loss-of-Susceptibility Mutants of *Arabidopsis thaliana* Reveal an Essential Role for eIF(iso)4E during Potyvirus Infection. *Current Biology*, *12*(02), 1046–1051.

Li, W., Teng, F., Li, T. & Zhou, Q. (2013) Simultaneous generation and germline transmission of multiple gene mutations in rat using CRISPR-Cas systems. *Nature Biotechnology* *31*, 684-686. doi:10.1038/nbt.2652

Liang, Z., Zhang, K., Chen, K., & Gao, C. (2014). Targeted Mutagenesis in *Zea mays* Using TALENs and the CRISPR/Cas System. *Journal of Genetics and Genomics*, *41*(2), 63–68. doi:10.1016/j.jgg.2013.12.001

Mccallum, C. M., Comai, L., Greene, E. A., & Henikoff, S. (2000). Targeted screening for induced mutations. *Nature Biotechnology*. *18*, 455–457.

Naderpour, M., Lund, O. L. E. S., Larsen, R., & Johansen, E. (2010). Potyviral resistance derived from cultivars of *Phaseolus vulgaris* carrying bc-3 is associated with the homozygotic presence of a mutated eIF4E allele. *Molecular Plant Pathology*, *11*, 255–263. doi:10.1111/J.1364-3703.2009.00602.X

Nekrasov V., Staskawicz B., Weigel D., Jones JD., Kamoun S. (2013) Targeted mutagenesis in the model plant *Nicotiana benthamiana* using Cas9 RNA-guided endonuclease. *Nature Biotechnology*, *31*:691–693 doi:10.1038/nbt.2655

Nicaise, V., German-retana, S., Sanjua, R., Dubrana, M., Mazier, M., Maisonneuve, B., Candresse, T., ... Legall, O. (2003). The Eukaryotic Translation Initiation Factor 4E Controls Lettuce Susceptibility to the Potyvirus Lettuce mosaic virus. *Plant Physiol* *1*, *132*, 1272–1282. doi:10.1104/pp.102.017855.recessive

Nieto, C., Morales, M., Orjeda, G., Clepet, C., Monfort, A., Sturbois, B., ... Aranda, M. A. (2006). An eIF4E allele confers resistance to an uncapped and non-polyadenylated RNA virus in melon. *The Plant Journal*, *45*2–462. doi:10.1111/j.1365-313X.2006.02885.x

Oldroyd, G. & Staskawicz, B. (1998). Genetically engineered broad-spectrum disease resistance, *Proc. Natl. Acad. Sci. USA*, *95*, 10300–10305.

Piron, F., Nicolai, M., Minoia, S., Piednoir, E., Moretti, A., Salgues, A., ... Bendahmane, A. (2010). An induced mutation in tomato eIF4E leads to immunity to two potyviruses. *PLoS One*, *5*(6), e11313. doi:10.1371/journal.pone.0011313

Pooggin, M. M., Shivaprasad, P. V., Veluthambi, K., and Hohn, T. (2003). RNAi targeting of DNA virus in plants. *Nature Biotechnology*, (*21*), 131-132

Regenmortel, M. & Mahy, B. (2009) *Desk Encyclopedia of Plant and Fungal Virology*. 426-430

Ronde, D. De, Butterbach, P., & Kormelink, R. (2014). Dominant resistance against plant viruses. *Frontiers in Plant Science*, 5(June), 1–17. doi:10.3389/fpls.2014.00307

Ruffel, S., Dussault, M., Palloix, A., Moury, B., Bendahmane, A., Robaglia, C., & Caranta, C. (2002). A natural recessive resistance gene against potato virus Y in pepper corresponds to the eukaryotic initiation factor. *The Plant Journal*, 1067–1075.

Ruffel, S., Gallois, J. L., Lesage, M. L., & Caranta, C. (2005). The recessive potyvirus resistance gene pot-1 is the tomato orthologue of the pepper pvr2-eIF4E gene. *Molecular Genetics and Genomics : MGG*, 274(4), 346–53. doi:10.1007/s00438-005-0003-x

Sato, M., Nakahara, K., Yoshii, M., Ishikawa, M., & Uyeda, I. (2005). Selective involvement of members of the eukaryotic initiation factor 4E family in the infection of *Arabidopsis thaliana* by potyviruses. *FEBS Letters*, 579, 1167–1171. doi:10.1016/j.febslet.2004.12.086

Shan, Q., Wang, Y., Li, J., Zhang, Y., Chen, K., Liang, Z., Zhang, K., Liu, J., Xi, J., Qui, J. & Gao, C. (2013) Targeted genome modification of crop plants using CRISPR-Cas system. *Nature Biotechnology*. 31, 686-688. doi:10.1038/nbt.2650

Shan, Q., Wang, Y., Li, J., & Gao, C. (2014). Genome editing in rice and wheat using the CRISPR / Cas system. *Nature Protocols*, 9(10), 2395–2410. doi:10.1038/nprot.2014.157

Stein, N., Perovic, D., Kumlehn, J., Pellio, B., Stracke, S., Streng, S., ... Graner, A. (2005). The eukaryotic translation initiation factor 4E confers multiallelic recessive Bymovirus resistance in *Hordeum vulgare* (L .). *The Plant Journal*, 912–922. doi:10.1111/j.1365-313X.2005.02424.x

Stemmer, M., Thumberger, T., Keyer, S., Wittbrodt, J., & Mateo, J. L. (2015). CCTop : An Intuitive , Flexible and Reliable CRISPR / Cas9 Target Prediction Tool. *PloS one*, 9, 1–11. doi:10.1371/journal.pone.0124633

Sternberg, S. H., Lafrance, B., Kaplan, M., Doudna, J. A., & Cas, C. (2015). Conformational control of DNA target cleavage by CRISPR-Cas9. *Nature*, 527(7576), 110–113. doi:10.1038/nature15544

Tilman, D., Balzer, C., Hill, J., & Befort, B. L. (2011). Global food demand and the sustainable intensification of agriculture. *Proc. Natl. Acad. Sci. USA*, 108(50). doi:10.1073/pnas.1116437108

Wang, H., Yang, H., Shivalila, C. S., Dawlaty, M. M., Cheng, A. W., & Zhang, F. (2013). Resource One-Step Generation of Mice Carrying Mutations in Multiple Genes by CRISPR / Cas-Mediated Genome Engineering. *Cell*, 153(4), 910–918. doi:10.1016/j.cell.2013.04.025

White, J. L., & Kaper, J. M. (1989). A simple method for detection of viral satellite RNAs in small plant tissue samples. *Journal of Virological Methods*, 23, 83-94

Wittmann, S., Chatel, H., & Fortin, M. G. (1997). Interaction of the Viral Protein Genome Linked of Turnip Mosaic Potyvirus with the Translational Eukaryotic Initiation Factor (iso) 4E of *Arabidopsis thaliana* Using the Yeast Two-Hybrid System. *Virology*, 92(234), 84–92.

Yang, Y., Sherwood, T. A., Patte, C. P., Hiebert, E., & Polston, J. E. (2004). Use of Tomato yellow leaf curl virus (TYLCV) Rep Gene Sequences to Engineer TYLCV Resistance in Tomato. *Phytopathology*, 94(5), 490-496.

Zhou, H., Liu, B., Weeks, D. P., Spalding, M. H., & Yang, B. (2014). Large chromosomal deletions and heritable small genetic changes induced by CRISPR / Cas9 in rice. *Nucleic Acids Research*, 1–12. doi:10.1093/nar/gku806

Zhu, J., Song, N., Sun, S., Yang, W., Zhao, H., Song, W., & Lai, J. (2016). Efficiency and Inheritance of Targeted Mutagenesis in Maize Using CRISPR-Cas9. *Journal of Genetics and Genomics*, 43(1), 25–36. doi:10.1016/j.jgg.2015.10.006

Accepted Article

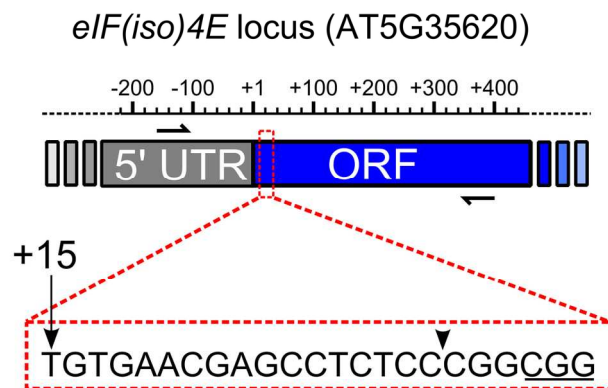


Figure 1 : Schematic of the *eIF(iso)4E* locus targeted for editing by CRISPR/Cas9. Primers flanking the target site are shown by half arrows over the 5' untranslated region (UTR) and open reading frame (ORF) shaded grey and blue, respectively. Scale bar depicts the approximate positions, in base pairs, relative to the translation start site (+1). The enlarged area in red dashes shows the position and sequence of the sgRNA, with the protospacer adjacent motif (PAM) underlined. The position of the sgRNA relative to the +1 of the ORF is indicated by an arrow. The Cas9 cleavage site is indicated by an arrow head.

165x116mm (300 x 300 DPI)

Accept

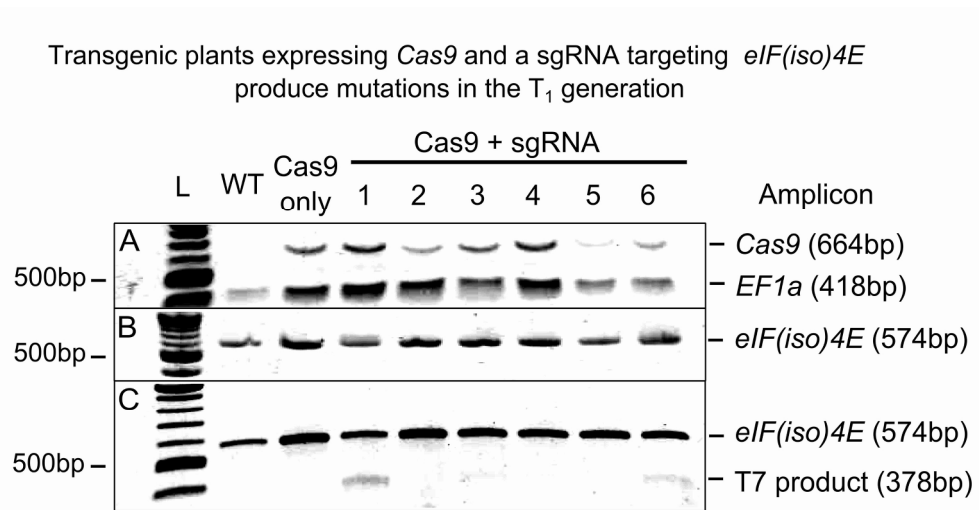


Figure 2 : PCR/T7 endonuclease products for six independently transformed plants (T₁ generation) containing a *Cas9* transgene with a sgRNA targetting the *eIF(iso)4E* locus. Transformation with a *Cas9* transgene with no sgRNA (*Cas9* only) and a non-transformed, wild type (WT) plant were used for controls. L denotes a 100bp DNA ladder. **A** 1% agarose gel showing multiplex PCR products confirming the presence/absence of the *Cas9* transgene, using the constitutively expressed house-keeping gene *EF1a* as a loading control. **B** 1% agarose gel showing PCR amplicons spanning the putative mutation site at the *eIF(iso)4E* locus. **C** 2% agarose gel showing *eIF(iso)4E* cleavage products after a self annealing reaction of the PCR amplicon, and subsequent digestion with T7 endonuclease. Presence of the 378bp cleavage product in samples 1, 3, and 6 is indicative of *Cas9* induced mutation at the *eIF(iso)4E* locus (the corresponding 196bp cleavage product is not visible on the gel).

213x150mm (300 x 300 DPI)

Accept

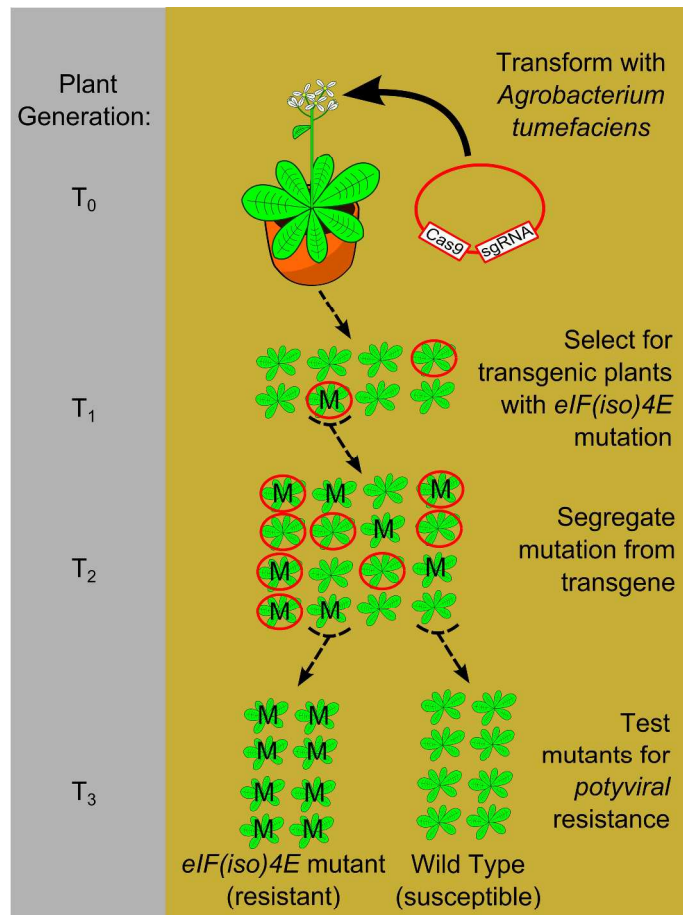


Figure 3: Schematic of project work-flow. Plants carrying the transgenic construct were selected in the T_1 generation by spraying with BASTA. The transgenic plants were then tested for *eIF(iso)4E* mutation (M) to identify plants with an active CRISPR/Cas9 nuclease using a T7 assay. One transgenic T_1 plant with clear signs of *eIF(iso)4E* editing was used to produce the T_2 generation. PCR was used to identify T_2 generation plants which had lost the transgene by Mendelian segregation. The non-transgenic T_2 plants were then screened for *eIF(iso)4E* mutations by Sanger sequencing (please note: the mutations depicted as 'M' in the diagram are not identical, as the mutations in the T_1 generation occurred in somatic cells so were not heritable, and different mutations were recovered in the T_2 generation due to independent editing events in the germline of T_1 gametes. For simplicity 'M' was used to depict all CRISPR/Cas9 induced mutations). Non-transgenic T_2 plants, which were homozygous for either the mutated or wild type *eIF(iso)4E* alleles, were used to produce T_3 populations, which were then tested for viral resistance.

369x647mm (300 x 300 DPI)

AC

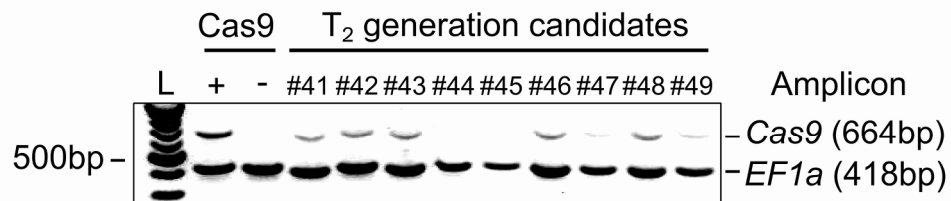
PCR selection of T₂ generation plants lacking the Cas9 transgene

Figure 4 : Representative 1% agarose gel for the selection of T₂ candidates lacking the Cas9 transgene. Multiplex PCR was used to confirm the presence/absence of the Cas9 transgene, using the constitutively expressed house-keeping gene *EF1a* as a loading control. L denotes a 100bp DNA ladder. A Cas9 transformant (T₁ generation) and a non-transformed wild type plant were used as positive and negative controls for Cas9 amplification, respectively. Samples #41 - #49 are a representative selection of T₂ progeny from T₁ plant number 1 (as shown in Figure 2) Candidates #44 and #45 represent two out of a total of 55 candidates lacking the Cas9 transgene, which were selected by this method.

209x143mm (300 x 300 DPI)

Accepte

Mutation summary for homozygous *eIF(iso)4E* mutants recovered in the T₂ generation

A

Sample	Nucleotide Sequence	Mutation
WT	TGTGAACGAGCCTCTCC-CGGCGG	
#44	TGTGAACGAGCCTCTCC--GGCGG	Deletion (-C)
#65	TGTGAACGAGCCTCTCCACGGCGG	Insertion (+A)
#68	TGTGAACGAGCCTCTCCTCGGCGG	Insertion (+T)
#98	TGTGAACGAGCCTCTCCCCGGCGG	Insertion (+C)

B

Sample	Amino Acid Sequence
WT	1 MATDDVNEPLPAAELPATEAEKQPHK.LERKMSFVFDNQSKKGAAWGASLRKAYTFDTV
#44	1 MATDDVNEPLRRRRNYRQRRRNHTSSKESGVSGSITNQRKA.PPGELLFVKPILSTP
#65	1 MATDDVNEPLGGGGITGDRGGETTQARKKVEFLVR*~
#68	1 MATDDVNEPLGGGGITGDRGGETTQARKKVEFLVR*~
#98	1 MATDDVNEPLGGGGITGDRGGETTQARKKVEFLVR*~
WT	60 EDFWGLHETIFQTSKLTANAIEIHLFKAGVEPKMEDPECANGKNTNVVTANRKEALDKGW
#44	59 SKIFGDCTRLY.....FRLAN*~
#65	38 ~~~~~~
#68	38 ~~~~~~
#98	38 ~~~~~~
WT	120 LETLMALIGEQFDEADEICGVVASVRPQSKQDKLSLWTRTKSNEAVLMGIGKKNKEILDV
#44	75 ~~~~~~
#65	38 ~~~~~~
#68	38 ~~~~~~
#98	38 ~~~~~~
WT	180 TDKITFNNHDSRRSRFTV*
#44	75 ~~~~~~
#65	38 ~~~~~~
#68	38 ~~~~~~
#98	38 ~~~~~~

Figure 5: Summary of CRISPR induced *eIF(iso)4E* mutations. **A** DNA sequence alignments for the four homozygous *eIF(iso)4E* mutants (#44, #65, #68, #98) identified in the T₂ generation, along with a wild type (WT) control. Lines #65, #68, and #98 exhibit single nucleotide insertions while line #44 has a single nucleotide deletion. **B** Predicted amino acid sequence alignments for the four homozygous mutants and the wild type consensus. Each of the mutant alleles code for severely truncated and disrupted proteins.

302x419mm (300 x 300 DPI)

AC

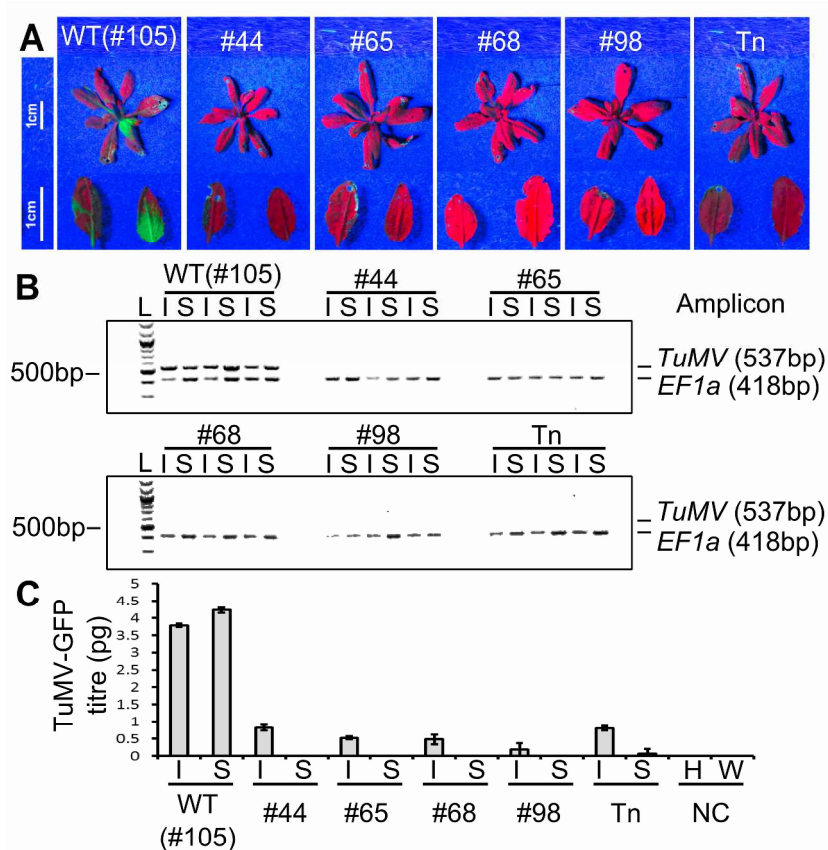


Figure 6 : **A** Representative photos of TuMV-GFP infected plants, imaged under UV light, 7 days post infection. A transposon induced *eIF(iso)4E* mutant (Tn) was used as a resistant control. An enlarged image of inoculated (left) and systemic (right) leaves are shown below each rosette. **B** RT-PCR to detect the presence of TuMV-GFP in inoculated (I) and systemic (S) leaves for each genotype. Three separate plants were analysed per genotype. The first two lanes for each genotype correspond to the leaves imaged in panel A. Amplicons of the TuMV coat protein region (537bp) and the housekeeping gene *EF1a* (418bp) were PCR amplified separately from the same cDNA, then mixed and run together on a 2% agarose gel. L denotes a 100bp DNA ladder TuMV specific amplicons are clearly visible in each of the wild type (WT) samples but completely absent from any of the *eIF(iso)4E* mutant samples. **C** qRT-PCR to detect the mean absolute viral titre (in pg) for the samples shown in panels A and B. qRT-PCR reactions were performed with cDNA from a healthy plant (H) and water (W) as negative controls (NC). Error bars show SEM of 3 biological replicates.

307x448mm (300 x 300 DPI)

AC

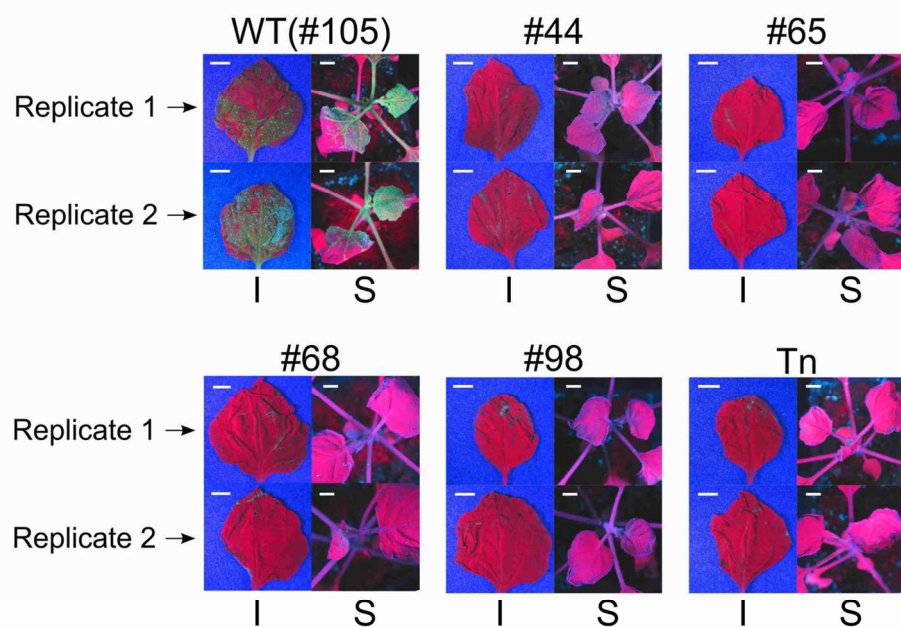


Figure 7: Back-inoculations of *Nicotiana benthamiana* plants using sap from TuMV-GFP inoculated *Arabidopsis*. Sap was prepared by pooling 20 systemic leaves from TuMV-GFP inoculated *Arabidopsis* plants as shown in Supplemental Figure 3. Labels above each quadrant refer to the genotype [Wild type (WT), #44, #65, #68, #98 and Transposon (Tn)] of the inoculated *Arabidopsis* used to make sap. Each quadrant shows an inoculated leaf (I) and systemic tissue (S) for two replicate plants, imaged under UV light. A 1cm scale bar is shown for each image.

127x103mm (300 x 300 DPI)

Accepted

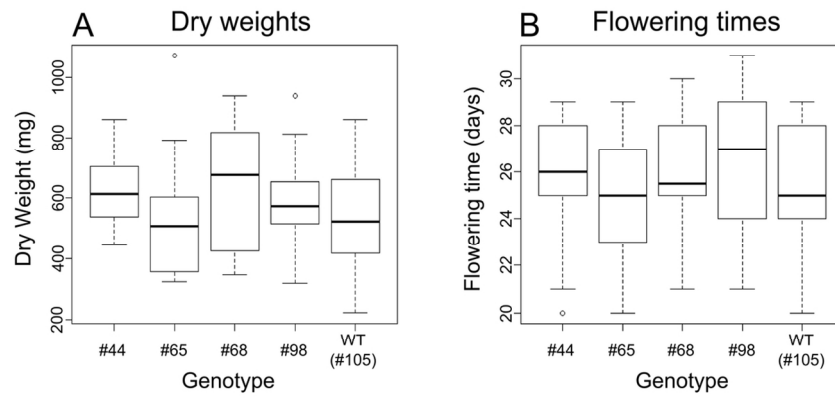


Figure 8 : Box plots of dry weights (**A**) and flowering times (**B**) for the CRISPR/Cas9 edited *eIF(iso)4E* mutants (lines #44,#65,#68 and #98) alongside wild type (WT) plant (#105).

119x67mm (300 x 300 DPI)

Accepted