CRISPR/Cas-induced targeted mutagenesis with Agrobacterium mediated protein delivery

Schmitz, D.J.

CRISPR/Cas-induced targeted mutagenesis with Agrobacterium mediated protein delivery

Daan Schmitz

Cover image: “TBE salt flats”
Cover & layout design: Daan Schmitz
Printed by : Ridderprint BV

CRISPR/Cas-induced targeted mutagenesis with Agrobacterium mediated protein delivery

Proefschrift

ter verkrijging van de graad van Doctor aan de Universiteit Leiden,
op gezag van de Rector Magnificus prof. mr. C.J.J.M. Stolker,
volgens besluit van het College voor Promoties te verdedigen op
donderdag 20 september 2018 te klokke 15.15 uur

door

Daan Johannes Schmitz
geboren te Haarlem, Nederland
30 maart 1987

Promotiecommissie

Promoter:	Prof. dr. P.J.J. Hooykaas
Co-promoter:	dr. B.S. de Pater
Overige leden:	dr. A.C. Vergunst
	Prof. dr. J. Memelink
	Prof. dr. H.P. Spaink
	Prof. dr. L. Otten

Aan mijn ouders en Amy

Table of contents

Chapter 1

General Introduction

Daan J. Schmitz, Sylvia de Pater, Paul J.J. Hooykaas

General introduction

In the first part of this chapter I will review the developments in the field of genome editing. Specifically the DNA repair mechanisms involved, the tools available and the applications of genome editing in plants. As I have used *Agrobacterium tumefaciens* for delivery, the second part of this chapter will focus on the plant pathogen *Agrobacterium*, describing the mechanisms underlying the transformation of plants by *Agrobacterium* and its biotechnological applications.

Part 1

Genetic engineering and genome editing

Controlled integration of transgenes and controlled genetic replacement of endogenous genes by transgenes based on homologous recombination, also known as gene targeting, is an important tool in biotechnology because it can accelerate the rate of functional gene analysis and guarantees a save introduction of novel traits. This controlled manipulation of genomes is based on the introduction of DNA that contains homology with the chromosomal locus allowing for DNA recombination. In lower eukaryotes such as yeast, integration of DNA occurs predominantly via homologous recombination and therefore the recombinatorial events required for gene-targeting were first studied in the budding yeast *Saccharomyces cerevisiae* [1,2]. In higher eukaryotes this kind of integration also occurs, albeit at low frequencies compared to random integration. The first molecular evidence for gene-targeting in animal cells was provided by the laboratories of Smithies [3] and Capecchi [4,5] in experiments that generated loss-of-function mutations in embryonic mouse stem cells.

In plant cells the first report of gene-targeting appeared in 1988. A DNA repair template transferred to tobacco protoplasts was shown to recombine with an integrated defective copy of a resistance gene leading to integration of a restored gene [6]. In 1989 this was followed by a report showing that delivery of a repair construct via *Agrobacterium*, a bacterium commonly used to transform plants, in the form of a T-DNA could be used for gene targeting leading to the correction of a defective gene at its original locus in the genome [7]. However, to harness the power of homology-directed recombination for gene insertion or gene repair in plants the naturally low HDR frequencies ranging between 10^{-4} and 10^{-6} needed to be enhanced [7–13].

Early experiments showed that DNA-damaging agents stimulated the exchange between sister chromatids [14], but the most compelling evidence that breaks in the DNA enhanced recombination came from studies that showed that a single double-strand break (DSB) in the genome dramatically increases the local frequency of recombination. In these studies a fragment encompassing the recognition site for a specific DNA endonuclease and a defective resistance gene was first inserted into the genome. Recombination with a homologous donor DNA increased several orders of magnitude when the cognate DNA endonuclease was expressed in the recipient cells [15,16].

Initially only naturally occurring site specific nucleases (SSNs) were available for this purpose. During the last decade, however, artificial site specific endonucleases with a customizable DNA recognition and cleavage site were developed that can now be employed for targeted modification of almost any genetic information in the genomes of organisms. These SSNs can also be used for targeted mutagenesis. When SSNs are expressed and used to

induce DSBs, the DNA repair machinery of the cell will seal the break, but when the nuclease is persistently present the DNA will be broken again leading to a cycle of break-repair until imprecise repair leads to a loss of the nuclease restriction site. In this way mutations can be selected at the nuclease target site. If these mutations occur in coding sequences genes can be knocked out making it easier to study the function of the mutated gene. Mutations created in non-coding regions can be used for instance to disrupt binding sites of pathogen produced transcription factors that promote disease [17]. Using SSNs even entire gene clusters can be removed by inducing DNA breaks at opposing ends of the gene cluster [18–20] or for the knockout of multiple (redundant), non-allelic homologous genes at the same time.

Double-strand break repair

The genome is subject to many agents, both exogenous and endogenous leading to DNA damage, that, if unrepaired, may lead to mutation or gross chromosomal rearrangements. Such damage includes base lesions, DNA single-strand breaks (SSBs) and DSBs. Repairing the damage is vital to maintain an organisms genomic integrity and stability and therefore multiple DNA repair pathways have evolved. Exploiting such DNA repair mechanisms of DSBs, and to a lesser extent those involved in SSB repair and base excision repair (BER), underlies genome editing. Two distinct pathways can be used for the repair of DSBs; non-homologous end joining (NHEJ) and homologous recombination (HR). Which repair pathway is used for the repair of DSBs differs greatly between organisms and the cell cycle phase at which the repair of a DSB occurs. Repair via NHEJ operates throughout the cell cycle but predominantly in the G1 phase [21], whereas repair via HR is restricted to the S and G2 phases when sister chromatids are available as a repair template.

NHEJ repairs DSBs by religating the broken ends irrespective of sequence homology and can be precise, but may also result in small deletions or insertions disrupting the genetic information [22]. The first step in NHEJ is the binding of the heterodimeric Ku complex to both ends of a DSB. This complex is composed of a 70kDa and a 80kDa subunit, named Ku70 and Ku80, respectively [23]. After the binding of Ku to a DSB, in mammalian cells DNA-PKcs is recruited to the end of the DNA break [24,25]. DNA-PKcs, however, is not present in plants and fungi. NHEJ is completed by ligation of the DNA ends; this rejoining is carried out by a complex of DNA ligase IV, XRCC4 and the XRCC4-like factor [26]. In plants and yeast orthologues of these C-NHEJ components have been identified [27–31]. Although several bacterial strains have been identified with a NHEJ-like DNA repair mechanism, the repair of DSBs in bacteria is predominately based on repair by HR [32,33].

In the absence of canonical factors involved in Ku-dependent NHEJ, back-up pathways are responsible for residual end joining of DSBs [34–37]. Repair via back-up NHEJ starts with the resection of the ends producing ssDNA ends that can anneal at microhomology regions. The ends may be linked at these microhomology regions, whereafter non-homologous tails are removed. The remaining gaps are filled by a specific DNA polymerase and re-ligated by DNA ligase I [38]. In these back-up pathways microhomologies sequences (5-25 basepairs) are frequently used for the repair of the DSB resulting in deletions [39] and therefore this pathway has also been called microhomology mediated end-joining (MMEJ) [40]. Several factors involved in repair via MMEJ have been identified: PARP1/2, MRN, CtIP, Ligase3, XRCC1 and DNA polymerase θ (Pol θ). The exact molecular mechanism behind MMEJ is poorly understood, but PARP1, the MRN complex and Pol θ have been implicated as key

players [41–43]. Recently, it has been shown that theta mediated end joining (TMEJ) is the dominant pathway for repair in Ku-deficient cells and the occurrence of microhomologies at the break site which is the result of the ability of Pol θ to mediate joining of two resected 3' ends harboring DNA sequence microhomology [42,43].

Homologous recombination is a DNA repair mechanism that uses DNA homology to direct DNA repair. These homologous sequences required for repair by HR are preferentially found on sister chromatids, but other naturally occurring homologous sequences or an artificially introduced repair template may sometimes be used instead. Repair via HR can be divided into several subpathways, classical double-strand break repair (DSBR), synthesis-dependent strand annealing (SDSA), single-strand annealing (SSA) and break-induced replication (BIR) [44]. DSBR was initially described to explain crossover events during meiosis and gene conversion [45]. SDSA is used for mitotic DSB repair [46] and is the predominant repair mechanism for two ended DSBs via HR [47]. In plants evidence has been found for the occurrence of DSBR, SDSA and SSA [48].

HR invariably starts off with the 5' resection of the ends resulting in large 3' ssDNA stretches that can search for complementary sequences. Repair via SDSA starts with a 3' end invading a homologous double-strand forming a D-loop. This is followed by repair synthesis from the 3' end using the newly paired strand as a template. The "X" shaped structure formed at the border between the hetero- and homoduplex of the D-loop is called a Holiday junction. After elongation, in SDSA the invading strand is displaced from the D-loop structure and anneals back to the 3' homologous strand that was formed by resection of the other end of the DSB. Thus in SDSA the donor molecule remains unaltered and gene conversion without loss of sequence information is the final result of the reaction [49]. SDSA is preferably used for HR repair in mitotic cells in the S and G2 phase, when a sister chromatid is available as a template. DSBR is mainly used for the repair of DSB breaks in meiotic cells. In this case both DNA ends invade a homologous chromosome to copy genetic information and this results in a double Holiday junction that may be resolved into a crossover or a non-crossover product.

SSA is a repair mechanism that can be used when two homologous sequences are in close proximity and arranged in a tandem orientation. Repair via SSA starts with resection and the production of 3' single-stranded overhangs. This is followed by immediate pairing of the single stranded complementary sequences found near the break site and trimming of the any remaining 3' ends [50]. SSA is not conservative in contrast to SDSA and DSBR and leads to deletion of one repeat and the intervening sequence. Up to one out of three DSBs in an artificial genomic region with tandemly arranged duplications was repaired via SSA after DSB induction [51].

Site-specific nucleases

To induce targeted DSBs, four different classes of SSNs have been developed: homing endonucleases (HEs), zinc finger nucleases (ZFNs), TAL effector nucleases (TALENs) and the CRISPR/Cas RNA guided endonuclease encoded by the clustered regularly interspaced short palindromic repeats (CRISPR)/CRISPR-associated (Cas) system (Fig. 1) [52].

The first class of SSNs are the HEs which are a class of proteins typically encoded for by introns or inteins and are considered selfish genetic elements, much like transposons [53]. This class of proteins can be divided into eight different subclasses: LAGLIDADG, H-N-H, His-Cys, GIY-YI G, PD(D/E)xK, HJ resolvase-like, EDxHD and Vsr-like [54]. All of these different classes of HEs have large DNA sequence recognition sites varying in length

between 12 to 40 basepairs. Because the DNA binding domain and the cleavage domain of HEs overlaps [55] creating HEs with new DNA binding specificities is challenging. For HEs of the LAGLIDADG class progress has however been made in altering the DNA recognition and cleavage site [56,57]. HEs were the first to be used for targeted mutagenesis in *Arabidopsis* and maize and instrumental in showing that DSB induction enhances targeted mutagenesis and gene targeting in plants (see Table 1).

In ZFNs the DNA binding domain consisting of zinc fingers is fused to the non-specific nuclease domain of the FokI nuclease. The zinc-finger array responsible for DNA binding is created by combining multiple zinc-finger domains, of the C_2H_2 class of zinc-fingers, that were discovered as part of the transcription factor IIIA [58]. This class of zinc-fingers consists of two β-sheets and an α-helix, which fits directly into the major groove of double strand DNA. The side chains from the N-terminal part of this helix contact the edges of the basepairs. Changing the amino acids in the α-helix may alter the affinity for different DNA sequences. Each zinc-finger domain recognizes 3bp of DNA [59]. Zinc-fingers have been created for the recognition of most of the 64 possible triplets [60–62]. By fusing the FokI nuclease domain to the zinc-finger array a zinc-finger nuclease (ZFN) is created. Two ZFNs are usually constructed that bind the DNA at opposing sites of the target sequence, each making a nick in one of the DNA strands leading together to a staggered DSB. A problem that can arise is the formation of a homodimer consisting of two ZFNs for one of the half sites. As a consequence, a DSB may be induced in the genome at a position which was not the initial target, possibly leading to toxicity [63]. Miller et al. [64,65] have designed complementary FokI cleavage domain variants that together function as an obligate heterodimer. The widespread application of ZFNs is however limited because of the limited selectivity conferred by the zinc-finger modules, and the complex context-dependent interactions between neighboring zinc-finger modules [66,67]. Moreover designing ZFNs typically involves multiple rounds of testing without necessarily resulting in a nuclease that performs optimally [66,68–70]. ZFNs have been used for genome editing in *Arabidopsis*, tobacco, soybean and maize (see Table 1).

The third class of artificial restriction enzymes called TAL effector nucleases (TALENs) has been developed as the successor of ZFNs. Transcription activator-like effectors (TALEs) are produced by plant pathogenic bacteria in the genus *Xanthomonas* and are transferred via type III secretion systems to the host, where they function as important virulence factors that act as transcriptional regulators [71–73]. The DNA binding domain of these TALEs contains multiple 30-35 amino acid long repeats that each recognizes a single base pair of DNA. Two hyper variable amino acids found at positions 12 and 13 of these repeats, known as the repeat-variable di-residues (RVDs), determine the base specificity [74,75]. Crystallization of the TALE DNA-binding domain revealed how the TALE wraps around the sense strand of the DNA as a right-handed super helix with each repeat forming a left-handed helix-loop-helix structure [76]. The structure also revealed that the 13th amino acid of the RVD determines the base specificity by interacting with the major groove, and that the 12th amino acid of the RVD stabilizes the loop in the helix-loop-helix structure [76,77]. By fusing the DNA binding domain of TALEs to the nuclease domain of FokI, TALENs are constructed that just like ZFNs can be designed to bind at opposing sites of the target sequences and create a DSB upon dimerization of the nuclease domain [78]. TALENs have been used to modify the genome of the following plants: *Arabidopsis,* barley, *Brachypodium*, cucumber, maize, rice, soybean, tobacco, tomato and wheat (see Table 1).

The newest addition to the family of SSNs is the CRISPR\Cas system derived from a prokaryotic adaptive immune system that cleaves DNA in a sequence dependent manner. In prokaryotes this RNA based defense system provides protection from foreign invading nucleic acids, such as viruses and plasmids [79–81]. The CRISPR/Cas system is made up of a Cas gene operon and CRISPR arrays. Immunity is acquired by integrating short fragments of invading DNA known as spacers between two flanking identical repeats at the proximal region of CRISPR loci. Transcripts of these CRISPR arrays, including the spacers, are processed into small interfering CRISPR RNAs (crRNAs) approximately 40 nucleotides in length [82]. Together with transactivating RNAs (tracrRNAs) these crRNAs form a complex which enables the Cas enzyme to cleave homologous double-stranded DNA sequences, known as protospacers, in previously encountered invading DNA [83,84].

There are three distinct classes of CRISPR/Cas systems from which a type II class system from *Streptococcus pyogenes* has mostly been used for genome editing, because only one protein (Cas9) is sufficient for nuclease activity in combination with the two RNA molecules [85]. A requirement for cleavage is the presence of a conserved protospacer adjacent motif (PAM) downstream of the target DNA, with either a 5'-NGG-3' or 5'-NAG-3' motif [85–87]. The quick development of the CRISPR/Cas system into a tool for genome editing started with the discovery that a 20 nucleotide change in the crRNA was sufficient to reprogram the DNA target specificity [85], and the subsequent development of a single chimeric guide RNA that combines the target specificity of the crRNA with the structural properties of the tracrRNA. The first reports of the CRISPR/Cas system being used for genome editing in different eukaryotes followed shortly thereafter [88–92]. The Cas9 protein has two active sites in a RuvC and a HNH domain, respectively, each of which induces a single-strand break which combined create a DSB. To transform Cas9 into a nickase that induces a single strand break point mutations have been introduced into the RuvC (D10A) and HNH (H840A) domains [85,86,93]. Using a nickase instead of a nuclease strongly decreases the frequency of off target mutations, while still increasing the frequency of HDR [94,95]. The CRISPR/Cas system has been used successfully for genome editing in the following plants: *Arabidopsis*, *Camelina*, common wine grape, tobacco, maize, petunia, orchids, potato, red sage, rice, sweet orange, tomato, sorghum, watermelon wheat and (see Table 2).

Using site-specific nucleases for genome editing in plants

The outcome of repair of a DSB is dependent on which repair pathway is used. In plants and many other higher eukaryotes repair via NHEJ is the predominant pathway used for the repair of DSBs. Repair via HR, but also via NHEJ often results in perfect repair of the lesion, but therefore also restores the target site for the SSNs allowing for the induction of a new DSB. This process can therefore cycle between DSB induction and repair until imperfect repair via NHEJ results in a small deletion or insertion that destroys the recognition site of the site specific nuclease.

When SSNs are used to introduce DSBs into a gene, mutations can be induced that affect gene function. All four classes of SSNs have been used to this end in both model plants as well as in a wide range of crop species (for an overview of see Table 1). High mutation frequencies have been found for instance when poplar leaf discs were transformed with a constitutive CRISPR/Cas construct resulting in 89% of leaf discs transformed showing an albino shoot indicative of mutations created in the phytoene desaturase gene [96]. Maize plants regenerated from calli that constitutively expressed the CRISPR/Cas system were found to have mutations frequencies up to 70-100% [97].

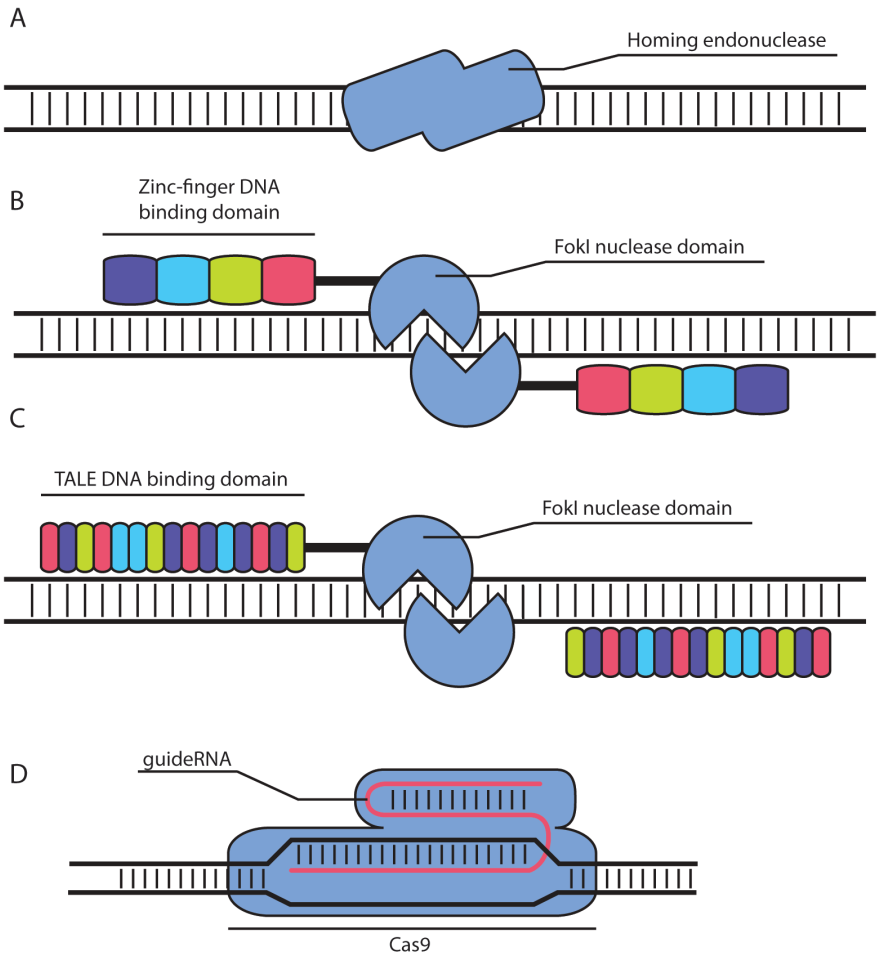

Figure 1. Site-specific nucleases used for genome editing. (**A**). Homing endonuclease with overlapping DNA binding and cleavage domain. (**B**) Zinc-finger nuclease (ZFN) composed of multiple zinc-fingers that form the DNA binding domain fused to the FokI nuclease domain. (**C**) TALEN composed of the DNA binding domain of transcription activator-like effectors (TALEs) fused to the FokI nuclease domain. **(D**) CRISPR/Cas composed of the Cas9 protein and the sgRNA which determines the base specificity with a 20bp sequence that is complementary to the target sequence.

Mutations created in the T0 generation however did not show expected inheritance patterns and are not always inherited by the T1 progeny [98,99]. In *Arabidopsis* chimerism was found to occur in T1 plants obtained after floral dip transformation [100,101]. More complex genetic modifications can be achieved when SSNs are used to induce two DSBs simultaneously, including specific deletions [19,69], inversions [102], duplications [102] and translocations [103]. In plants ZFNs, homing endonucleases and CRISPR/Cas have been used to create targeted chromosomal deletions [20,104–106].

When a DSB is induced and simultaneously an artificial repair template lacking the nuclease restriction site is present, this may be used for HDR, and the specific mutations present in the template introduced into the genome. In plants this has been accomplished in the following species: *Arabidopsis*, tobacco, tomato, soybean, rice, potato, wheat and maize. Similarly by using a repair template containing novel genes these can be inserted in this way at the DSB site (gene targeting). In plants the introduction of DSBs near the site of the desired recombination has been shown to greatly increase the frequency of HDR. The introduction of a SSB by engineered nickases also increases the frequency of HDR, although the increase is less pronounced as with the introduction of a DSB [107,108]. Repair via HDR has been used in plants for targeted modifications of genes in model species as well in several crop species using ZFNs [109–112], TALENs [113,114] and the CRISPR/Cas system [115–120].

Delivery of site-specific nucleases

Although the choice for the delivery system used will often be based on efficiency and feasibility considerations, the resulting expression levels and concentration of the nuclease will also impact the outcome. In plants several methods of delivery have been used for either constitutive or transient expression of SSNs. To obtain plants that constitutively express SSNs varying plant tissues have been transformed using biolistic or *Agrobacterium* based methods (Table 1 and Table 2). These methods, although relatively easy, have some drawbacks. Integration of transgenes may disrupt endogenous genes at the site of transgene integration. The presence of transgenes increases the administrative burden in regulatory processes needed for marketing of genetically modified crops. Furthermore, depending on the position in the genome gene expression levels might be influenced and constitutive expression of the SSN may lead to off target effects. Therefore methods are being developed for the transient expression of SSNs, for instance by controlling them by inducible or cell specific promoters. Alternatively methods for the direct delivery of nuclease mRNA have also been developed [90,121,122], as well for the direct delivery of SSN proteins in human cells and drosophila [88,123,124] and tobacco protoplasts for which direct introduction of I-SceI and TALEN protein was reported [125]. In addition, preassembled CRISPR/Cas complexes of purified Cas9 protein and sgRNA have been transfected into protoplasts of *Arabidopsis*, tobacco, rice and lettuce [126]. Direct protein delivery does however require the isolation and purification of large quantities of SSNs and the isolation of protoplasts. Therefore, a system that eliminates the isolation and purification steps but would still directly deliver the SSN as a protein would have value in both academic and commercial settings. The *Agrobacterium* system may fulfill this promise.

Table 1. Overview of use of homing endonucleases, zing finger nucleaes and TALENs in plants

Site specific nuclease	Transformation method	Mutation type	Reference
Homing endonucleases			
A. thaliana	*Agrobacterium* (Floral dip)	TM / Gene excision	[185]
Z. mays	*Agrobacterium* (Embryo)	TM	[186]
Z. mays	*Agrobacterium* (Embryo)	TM	[187]
Zinc finger nucleases			
A. thaliana	*Agrobacterium* (Floral dip)	TM	[188]
A. thaliana	*Agrobacterium* (Floral dip)	TM	[189]
A. thaliana	*Agrobacterium* (Floral dip)	TM	[190]
A. thaliana	*Agrobacterium* (Floral dip)	TM / GT	[109]
A. thaliana	*Agrobacterium* (Floral dip)	TM / GT	[110]
G. max	*Agrobacterium* (Hairy roots)	TM	[191]
N. tabacum	Protoplast electroporation	TM / GT	[111]
Z. *mays*	*Agrobacterium* (Embryo)	TM / Gene addition	[112]
TALENs			
A. thaliana	*Agrobacterium* (Floral dip)	TM	[192]
A. thaliana	*Agrobacterium* (Floral dip)	TM	[193]
A. thaliana	PEG-protoplast transfection	TM	[194]
B. distachyon	PEG-protoplast transfection	TM	[195]
G. max	*Agrobacterium* (Hairy roots)	TM	[196]
G. max	*Agrobacterium* (Hairy roots)	TM	[197]
H. vulgare	*Agrobacterium* (Embryonic pollen), Particle bombardment	GT	[198]
H. vulgare	*Agrobacterium* (Embryonic pollen)	TM	[199]
H. vulgare	*Agrobacterium* (Embryonic pollen)	TM	[200]
H. vulgare	Particle bombardment	GT	[198]
N. benthamiana	mRNA transfection	TM	[201]
N. benthamiana	PEG-protoplast transfection	GT	[202]
N. tabacum	PEG-protoplast transfection	TM / GT	[113]
O. sativa	*Agrobacterium* (Callus)	TM	[203]
O. sativa	*Agrobacterium* (Callus)	TM	[204]
O. sativa	Particle bombardment	GT	[205]
O. sativa	PEG-protoplast transfection	TM	[120]
O. sativa	PEG-protoplast transfection	TM	[206]
S. lycopersicum	*Agrobacterium* (Seedlings)	TM / GT	[114]
S. tuberosum	PEG-protoplast transfection	TM	[207]
T. aestivum	PEG-protoplast transfection	TM	[208]
Z. mays	*Agrobacterium* (Embryo)	TM	[209]

TM = targeted mutagenesis, GT = gene-targeting, CD = Chromosomal deletions

Table 2. Overview of use CRISPR/Cas in plants

CRISP/Cas9	Transformation method	Mutation type	Reference
A. thaliana	*Agrobacterium* (Floral dip)	TM	[210]
A. thaliana	*Agrobacterium* (Floral dip)	TM	[211]
A. thaliana	*Agrobacterium* (Floral dip)	TM	[212]
A. thaliana	*Agrobacterium* (Floral dip)	TM	[213]
A. thaliana	*Agrobacterium* (Floral dip)	TM	[214]
A. thaliana	*Agrobacterium* (Floral dip)	TM	[215]
A. thaliana	*Agrobacterium* (Floral dip)	TM	[216]
A. thaliana	*Agrobacterium* (Floral dip)	TM	[101]
A. thaliana	*Agrobacterium* (Floral dip)	TM	[217]
A. thaliana	*Agrobacterium* (Floral dip)	TM	[218]
A. thaliana	*Agrobacterium* (Floral dip)	TM	[219]
A. thaliana	*Agrobacterium* (Floral dip)	TM / GT	[116]
A. thaliana	*Agrobacterium* (Floral dip)	TM / GT	[220]
A. thaliana	*Agrobacterium (Floral dip)*	TM (SSA)	[106]
A. thaliana	*Agrobacterium* (Floral dip)	TM (SSA)	[221]
A. thaliana	*Agrobacterium* (Leaf infiltration)	CD	[222]
A. thaliana	*Agrobacterium* (Leaf infiltration)	TM	[223]
A. thaliana	PEG-protoplast transfection	TM	[126]
A. thaliana	PEG-protoplast transfection	TM	[224]
A. thaliana	*PEG-protoplast transfection / Agrobacterium* (Floral dip)	TM / GT	[115]
Camelina sativa	*Agrobacterium* (Floral dip)	TM	[225]
C. lanatus	PEG-protoplast transfection	TM	[226]
C. sinensis	*Agrobacterium* (Leaf infiltration)	TM	[227]
C. sinensis	*Agrobacterium* (Cotelydons)	TM	[228]
C. sativus	*Agrobacterium* (Cotelydons)	TM	[229]
Chrysanthemum morifolium	*Agrobacterium* (Callus)	TM	[230]
Dendrobium officinale	*Agrobacterium* (protocorns)	TM	[231]
G. max	*Agrobacterium* (Hairy roots)	TM	[232]
G. max	*Agrobacterium* (Hairy roots)	TM	[233]
G. max	*Agrobacterium* (Hairy roots)	TM	[196]
G. max	*Agrobacterium* (Hairy roots)	TM	[234]
G. max	PEG-protoplast transfection	TM	[235]
G. max	Particle bombardement	TM / GT	[236]
G. max	*Agrobacterium* (Hairy roots)	TM	[237]
Gossypium hirsutum L.	*Agrobacterium* (Hypocotyl/cotyledonary petiole)	TM	[238]
Gossypium hirsutum L.	*Agrobacterium* (Hypocotyl/cotyledonary petiole)	TM	[239]
Gossypium hirsutum L.	*Agrobacterium* (Hypocotyl/cotyledonary petiole)	TM	[240]
Hordeum vulgare	*Agrobacterium* (Embryo)	CD	[241]
L. sativa	PEG-protoplast transfection	TM	[126]
M. truncatula	*Agrobacterium* (Hairy roots)	TM	[242]
M. truncatula	*Agrobacterium* (Hairy roots)	TM	[237]
Marchantia polyorpha	*Agrobaceterium* (spores)	TM	[243]
N. attenuata	PEG-protoplast transfection	TM	[235]
N. attenuata	PEG-protoplast transfection	TM	[126]
N. benthamiana	*Agrobacterium* (Cas9) / guideRNA (CaLCuV virus)	TM	[244]
N. benthamiana	*Agrobacterium* (Cas9) / guideRNA (TVR)	TM	[245]
N. benthamiana	*Agrobacterium* (Cas9) / guideRNA (TYLCV virus)	TM	[246]
N. benthamiana	*Agrobacterium* (Leaf infiltration)	TM	[247]
N. benthamiana	*Agrobacterium* (Leaf infiltration)	TM	[223]
N. benthamiana	*Agrobacterium* (Leaf infiltration)	TM / CD	[104]
N. benthamiana	*Agrobacterium* (Leaf infiltration)	TM	[248]

N. benthamiana	*Agrobacterium* (Leaf infiltration)	CD	[222]
N. benthamiana	PEG-protoplast transfection	TM	[115]
N. benthamiana	PEG-protoplast transfection	TM	[224]
N. benthamiana	PEG-protoplast transfection	TM	[249]
N. tabacum	*Agrobacterium* (Leaf infiltration)	TM	[250]
N. tabacum	*Agrobacterium* (Leaf disc)	TM	[251]
N. tabacum	PEG-protoplast transfection	TM	[252]
N. tabacum	PEG-protoplast transfection	TM	[253]
N. tabacum	PEG-protoplast transfection	TM / CD	[105]
N. tabacum	PEG-protoplast transfection / *Agrobacterium*(Leaf disc)	TM	[254]
O. sativa	*Agrobacterium* (Callus)	TM	[255]
O. sativa	*Agrobacterium* (Callus)	TM	[250]
O. sativa	*Agrobacterium* (Callus)	TM	[210]
O. sativa	*Agrobacterium* (Callus)	TM	[256]
O. sativa	*Agrobacterium* (Callus)	TM	[257]
O. sativa	*Agrobacterium* (Callus)	TM	[211]
O. sativa	*Agrobacterium* (Callus)	TM	[258]
O. sativa	*Agrobacterium* (Callus)	TM	[259]
O. sativa	*Agrobacterium* (Callus)	TM	[260]
O. sativa	*Agrobacterium* (Callus)	TM	[261]
O. sativa	*Agrobacterium* (Callus)	TM	[262]
O. sativa	*Agrobacterium* (Callus)	TM	[263]
O. sativa	*Agrobacterium* (Callus)	TM	[264]
O. sativa	*Agrobacterium* (Callus)	TM	[265]
O. sativa	*Agrobacterium* (Callus)	TM	[266]
O. sativa	*Agrobacterium* (Callus)	TM	[267]
O. sativa	*Agrobacterium* (Callus)	TM	[98]
O. sativa	*Agrobacterium* (Callus)	TM	[268]
O. sativa	*Agrobacterium* (Callus)	TM	[269]
O. sativa	*Agrobacterium* (Callus)	TM	[270]
O. sativa	*Agrobacterium* (Callus)	TM	[271]
O. sativa	*Agrobacterium* (Callus)	GT	[272]
O. sativa	*Agrobacterium* (Particle bombardment)	TM / GT	[273]
O. sativa	PEG-protoplast transfection	TM	[126]
O. sativa	PEG-protoplast transfection	TM	[274]
O. sativa	PEG-protoplast transfection	GT / GI	[275]
O. sativa	PEG-protoplast transfection / *Agrobacterium*	TM / CD	[20]
O. sativa	PEG-protoplast transfection / *Agrobacterium*	TM / CD	[104]
O. sativa	PEG-protoplast transfection / *Particle bombardement*	TM	[120]
Petunia hybrida	PEG-protoplast transfection ribonucleoproteins	TM	[276]
Petunia hybrida	*Agrobacterium* (Leaves)	TM / CD	[106]
Physcomitrella patens	PEG-protoplast transfection	TM	[277]
Physcomitrella patens	PEG-protoplast transfection	TM	[278]
P. tomentosa	*Agrobacterium* (Leaf disc)	TM	[96]
P. tomentosa	*Agrobacterium* (Leaf disc)	TM	[279]
P. tomentosa	*Agrobacterium* (Leaf disc)	TM	[280]
S. bicolor	*Agrobacterium* (Leaf infiltration)	TM	[223]
Salvia miltiorrhiza	*Agrobacterium* (Hairy roots)	TM	[281]
Scopelophila cataractae	PEG-protoplast transfection	TM	[278]
S. lycopersicum	*Agrobacterium* (Cotyledon segments)	CD	[282]
S. lycopersicum	*Agrobacterium* (Cotyledon segments)	TM	[283]
S. lycopersicum	*Agrobacterium* (Cotyledon segments)	TM	[114]
S. lycopersicum	*Agrobacterium* (Cotyledon segments)	TM	[284]
S. lycopersicum	*Agrobacterium* (Cotyledon segments)	TM	[285]
S. lycopersicum	*Agrobacterium* (Cotyledon segments)	TM	[286]
S. lycopersicum	*A. rhizogenes*	TM	[287]
S. tuberosum	*Agrobacterium* (Callus)	TM / GT	[119]

S. tuberosum	*Agrobacterium* (Stem segments)	TM	[288]
S. tuberosum	PEG-protoplast transfection	TM	[289]
T. aestivum	Particle bombardement	TM	[290]
T. aestivum	PEG-protoplast transfection	TM	[291]
T. aestivum	PEG-protoplast transfection	TM / GT	[120]
T. aestivum	PEG-protoplast transfection	TM	[208]
T. aestivum	PEG-protoplast transfection / Particle bombardement	TM	[292]
T. aestivum	*Agrobacterium* (Leaf infiltration)	TM	[293]
Vitis vinifera	*Agrobacterium* (Embryo)	TM	[294]
Z. mays	*Agrobacterium* (Embryo)	TM	[99]
Z. Mays	*Agrobacterium* (Embryo) / Particle bombardement	TM / GT / GI	[295]
Z. mays	*PEG-protoplast transfection* / Agrobacterium(Embryo)	TM	[296]
Z. mays	*PEG-protoplast transfection* / Agrobacterium(Embryo)	TM	[217]
Z. mays	PEG-protoplast transfection	TM	[297]
Z. mays	PEG-protoplast transfection	TM	[223]
Z. mays	PEG-protoplast transfection ribonucleoproteins	TM	[298]

TM = targeted mutagenesis, GT = gene-targeting, CD = Chromosomal deletions

Part 2

Agrobacterium

Agrobacterium tumefaciens is known as the causative agent of crown gall disease in plants since the beginning of the 20th century. Crown galls arise at the site of infection and are the result of uncontrolled cell division. They can be excised and propagated in vitro without exogenous plant hormones [127]. Crown galls produce opines, amino-acid derivatives, which can be used by *Agrobacterium* as a nitrogen and carbon source and therefore the galls form a favorable niche for the bacterium [128]. The genetic basis for the uncontrolled growth of crown galls is the transfer of a copy of the T-region of the a large Tumor inducing (Ti) plasmid from the bacterium into plant cells. Processing of this T-region in the bacterium results in production of single stranded copies known as T-strands (transfer strands) that are transferred through a virB/virD4 type IV secretion system (T4SS) directly into the host cells. Upon entry into the host cell the T-strand is directed towards the nucleus and eventually may be integrated into the host genome [129]. The genes on this transferred DNA (T-DNA) encode for several proteins involved in the production and activation of plant hormones and the production of opines [130,131].

Induction of the virulence genes

The transformation of plant cells starts with a large cascade of events, the first step of which is the detection of phenolic compounds released by wounded plant cells by the bacterium through the VirA/VirG two-component sensor/regulatory system. VirA is autophosphorylated upon interaction with these phenolic compounds [132–135]. Phosphorylation of VirG by VirA leads to the binding of VirG to the *vir*-boxes of *vir* genes, stimulating their expression [136]. In addition to the VirA/VirG system the chromosomally encoded periplasmic protein ChvE, acts synergistically with the VirA/VirG system to induce *vir* gene expression upon detection of specific sugars in the plant sap [137]. Activation of the *vir* system only occurs in medium with a low pH like plant sap and at moderate temperatures (below 30° degrees). Eight operons designated *VirA*, *B*, *C*, *D*, *E*, *F*, *G* and *H* are largely conserved on the virulence regions of different types of Ti-plasmid. [138–140].

T-DNA processing

The T-region on the Ti-plasmid is flanked by two imperfect direct repeats (borders) that are the only determinants that define the T-DNA, of which only the right border sequence is essential for transfer [141]. Both of these borders are recognized by the virulence proteins VirD1 and the relaxasome VirD2. A single stranded break is introduced by VirD2 which stays covalently attached to the 5'-end of the nick [142,143]. The accessory Vir protein VirD1 enhances the binding and nicking on supercoiled DNA, while VirC1 and VirC2 further increase the efficiency by attaching to the overdrive sequence found close to the right border [144,145]. Probably the T-strand is released by displacement synthesis with VirD2 still covalently attached to the 5' end. This nucleo-protein complex is recognized by the T4SS and transported into the host cell [146].

The type IV secretion system of Agrobacterium

The T4SS of *Agrobacterium* through which the T-strand is transported is a large multi-protein complex that spans the inner membrane, the periplasm and the outer membrane. The genes encoding the T4SS are located on the Ti-plasmid and are encoded by a large *virB* operon, consisting of 11 open reading frames, and the *virD4* gene. Studying the T4SS has led to a division of its components into energy providing subunits (VirB4, VirB11, VirD4), inner membrane channel components located predominantly in the inner membrane (VirB3, VirB6, VirB8), periplasmic and membrane outer components (VirB7, VirB9), connecting component (VirB10), pilus components (VirB2, VirB5) and the VirB1 transglycosylase (extensively reviewed [147]). The T4SS present in *Agrobacterium* also translocates separately the T-strand four other so called effector proteins (VirD5, VirE2, VirE3 and VirF) into the host cell [148–150]. Their translocation via the T4SS is dependent on a C-terminal translocation signal present in all translocated proteins that has a net positive charge with a consensus motif of R-X(7)-R-X-R-X-R-X-X(n)>[149]. Combined with VirD2, VirE2 aids in nuclear targeting of the T-strand [151,152]. The translocated effector protein VirE3 is highly conserved and has two potential NLS by which it binds to importin-α and is transported into the nucleus [153]. Upon mutation of VirE3 hardly any attenuation of virulence is observed, but combined with the inactivation of VirF the role of VirE3 in virulence becomes apparent [154]. Recently, it was shown that VirE3 is a transcription factor which upregulates the VIP1-binding F-box gene thereby removing the requirement for VirF [155]. A recent study focusing on the virulence protein VirD5 suggests that it binds to centromers/kinetochores and there may induce cycle arrest, chromosome missegregation and aneuploidy [156]. The VirF protein is an F-box protein [157] which together with homologues of the yeast protein Skp1 and Cullin forms a Skp-Cullin-F-box protein (SCF) complex [157]. This complex has been reported to target the host protein VIP1 and associated VirE2 for degradation by the 26S proteasome [158], which may be required for the uncoating of the T-DNA to enable integration into the host cell's genome. VirF also targets degradation of the transcription factors VBF3 and VBF4 [159]. Plant species that can be transformed by an *Agrobacterium* strain lacking virF have been shown to induce upon infection the expression of a plant F-box protein, VBF, that can functionally replace VirF [160].

T-DNA transfer and integration

T-DNA transfer across the different T4SS components occurs in a tightly controlled manner in which the T-strand makes sequential contact with VirD4, VirB11, VirB6, VirB8, VirB9 and VirB2 [161]. Although the other components of the T4SS have not been found to interact with the T-strand they are essential [161–163]. The translocated VirE2 protein is thought to be responsible for protecting the T-strand against degradation by host cell proteases/nucleases after the T-strand has been transported to the hosts cytoplasm [164]. In the nucleus the T-strand is converted into a double strand form, which may be expressed for some time and then degraded. Integration into the genome leads to stable transformation and the continued expression that is necessary for tumorigenesis. The integration of the T-DNA occurs at a random position into the genome [165], which can result in mutation and variable expression due to position effects. The molecular mechanism of T-DNA integration was first studied in yeast where integration occurs efficiently via HR when the T-DNA contains yeast homologous sequences [166]. If no homology is present on the T-DNA integration in both

yeast and fungi occurs by NHEJ and is dependent on the C-NHEJ proteins Ku70, Ku80 and DNA ligase 4 [167,168]. Studies on the influence of these proteins on T-DNA integration in plants by different research groups mostly showed that their absence gave no or limited effects [30,169–171]. Disruption of multiple DNA repair pathways simultaneously also did not eliminate transformation [172–174]. Recently, it has been shown that the random integration of the T-strand occurs through a mechanism involving Pol θ [175].

Biotechnological applications of Agrobacterium

After the discovery of the genetic mechanism behind the transformation of plants by *Agrobacterium,* researchers focused on developing *Agrobacterium* into a tool for the introduction of foreign genes into plants. The genes that are naturally located between the borders are not involved in T-DNA transfer and these sequences can be replaced with other sequences of interest. The introduction of these genes of interest into the Ti-plasmid is however difficult due to its large size, low copy number and its inability to replicate in *E. coli*. The *vir* region and the T-DNA can however be separated on different plasmids without impacting *Agrobacterium*'s ability to form tumors [176]. The plasmid harboring the T-DNA, called the binary vector, contains at least one origin of replication that is functional in both *Agrobacterium* and *E. coli* and selectable markers for maintenance in *Agrobacterium* and *E. coli*, hereby allowing the easy cloning of genes on the T-region [176]. The plasmid with the intact *vir* region but lacking the T-region is known as the helper plasmid. The ability to introduce foreign genes into plant cells has made *Agrobacterium* an invaluable tool for plants scientists.

Using *Agrobacterium* plants have been created with enhanced tolerance to abiotic and biotic stress and pest resistance. Furthermore, *Agrobacterium* has been applied to better understand plant biology at a cellular and molecular level for example tagging various proteins with a fluorescent protein to visualize cell compartments [177]. In addition T-DNA insertion libraries have been made in which the T-DNA functions as an insertional mutagen and, by extension, a gene tag [178,179]. More recently SSNs have been introduced into plants, using *Agrobacterium*, for targeted gene disruption and gene targeting in plants (see Table 1). In recent years, *Agrobacterium* has also been used to modify plants for the production of useful proteins, such as edible vaccines and recombinant antibodies [177].

Although *Agrobacterium* only forms tumors on dicotyledonous plants, under laboratory conditions it is also capable of transforming monocotyledonous plants [180], yeast [166] and a wide range of fungi [181]. Because of its ease of use, low cost and precision it has become a preferred vector not only for the genetic modification of plants, but also of yeasts and fungi. As mentioned above *Agrobacterium* not only uses its T4SS for T-DNA transfer but also for the translocation of several effector proteins. It has been found that the translocation signal of these effector proteins can be attached to heterologous proteins such as the Cre recombinase and I-SceI to effectuate their transfer to plant cells [148,149,166,182–184].

Outline of this thesis

In **Chapter 2** we describe the application of CRISPR/Cas in *Agrobacterium*, to cure *Agrobacterium* from the promiscuous plasmid RP4 and from vectors with the replication unit of the octopine Ti plasmid. Furthermore we show that the Cas9 protein fused to a translocation signal recognized by the T4SS does not negate its ability to induce DSBs.

In **Chapter 3** yeast was used as a model organism to show that the Cas9 protein can be delivered through the T4SS of the plant pathogen *Agrobacterium*. The transfer of Cas9 was effectuated by fusion of a T4SS translocation peptide to the Cas9 protein.

In **Chapter 4** a method for targeted mutagenesis in *Nicotiana benthamiana* was developed, that is based on the translocation of the Cas9 protein through the T4SS of *Agrobacterium*. We show that concurrent transfer of Cas9 protein and a T-DNA encoding the sgRNA results in targeted mutations in the infiltrated leaves of *N. benthamiana*.

In **Chapter 5** we describe two novel approaches for the transient expression of the *Agrobacterium* derived isopentenyl transferase that can be used for the selection of transformed plants. The first consisted of the delivery of a T-DNA encoding the ipt gene into Pol-θ-deficient *Arabidopsis* mutants in which only transient expression of the T-DNA occurs but no integration. The second approach involved the direct delivery of the IPT protein through the *Agrobacterium* T4SS into *Arabidopsis.*

References

1. Scherer S, Davis RW. Replacement of chromosome segments with altered DNA sequences constructed in vitro. Proc. Natl. Acad. Sci. USA. 1979;76:4951–5.

2. Orr-Weaver TL, Szostak JW, Rothstein RJ. Yeast transformation: a model system for the study of recombination. Proc. Natl. Acad. Sci. USA. 1981;78:6354–8.

3. Doetschman T, Gregg RG, Maeda N, Hooper ML, Melton DW, Thompson S, et al. Targeted correction of a mutant HPRT gene in mouse embryonic stem cells. Nature. 1987;330:576–8.

4. Thomas KR, Capecchi MR. Site-directed mutagenesis by gene targeting in mouse embryo-derived stem cells. Cell. 1987;51:503–12.

5. Capecchi MR. Altering the genome homologous recombination Science. 1989;244:1288–92.

6. Paszkowski J, Baur M, Bogucki A, Potrykus I. Gene targeting in plants. EMBO J. 1988;7:4021–6.

7. Offringa R, de Groot MJ, Haagsman HJ, Does MP, van den Elzen PJ, Hooykaas PJJ Extrachromosomal homologous recombination and gene targeting in plant cells after Agrobacterium mediated transformation. EMBO J. 1990;9:3077–84.

8. Hrouda M, Paszkowski J. High fidelity extrachromosomal recombination and gene targeting in plants. Mol. Gen. Genet. 1994;243:106–11.

9. Lee KY, Lund P, Lowe K, Dunsmuir P. Homologous recombination in plant cells after Agrobacterium-mediated transformation. Plant Cell. 1990;2:415–25.

10. Halfter U, Morris PC, Willmitzer L. Gene targeting in Arabidopsis thaliana. Mol. Gen. Genet. 1992;231:186–93.

11. Miao ZH, Lam E. Targeted disruption of the TGA3 locus in Arabidopsis thaliana. Plant J. 1995;7:359–65.

12. Risseeuw E, Offringa R, Franke-van Dijk MEI, Hooykaas PJJ. Targeted recombination in plants using Agrobacterium coincides with additional rearrangements at the target locus. Plant J. 1995. 109–19.

13. Hanin M, Volrath S, Bogucki A, Briker M, Ward E, Paszkowski J. Gene targeting in Arabidopsis. Plant J. 2001;28:671–7.

14. Latt S. Sister Chromatid exhange formation. Annu. Rev. Genet. 1977;15:11–55.

15. Rudin N, Sugarman E, Haber JE. Genetic and physical analysis of double-strand break repair and recombination in Saccharomyces cerevisiae. Genetics. 1989;122:519–34.

16. Plessis A, Perrin A, Haber JE, Dujon B. Site-specific recombination determined by I-SceI, a mitochondrial group I intron-encoded endonuclease expressed in the yeast nucleus. Genetics. 1992;130:451–60.

17. Li T, Liu B, Spalding MH, Weeks DP, Yang B. High-efficiency TALEN-based gene editing produces disease-resistant rice. Nat. Biotechnol. 2012;30:390–2.

18. Qi Y, Li X, Zhang Y, Starker CG, Baltes NJ, Zhang F, Targeted deletion and inversion of tandemly arrayed genes in Arabidopsis thaliana using zinc finger nucleases. G3 (Bethesda). 2013;3:1707–15.

19. Xiao A, Wang Z, Hu Y, Wu Y, Luo Z, Yang Z, et al. Chromosomal deletions and inversions mediated by TALENs and CRISPR/Cas in zebrafish. Nucleic Acids Res. 2013;41:1–11.

20. Zhou H, Liu B, Weeks DP, Spalding MH, Yang B. Large chromosomal deletions and heritable small genetic changes induced by CRISPR/Cas9 in rice. Nucleic Acids Res. 2014;42:10903–14.

21. Chapman JR, Taylor MRG, Boulton SJ. Playing the End Game: DNA Double-Strand Break Repair Pathway Choice. Mol. Cell. 2012;47:497–510.

22. Puchta H. The repair of double-strand breaks in plants: mechanisms and consequences for genome evolution. J. Exp. Bot. 2005;56:1–14.

23. Walker JR, Corpina R a, Goldberg J. Structure of the Ku heterodimer bound to DNA and its implications for double-strand break repair. Nature. 2001;412:607–14.

24. Yaneva M, Kowalewski T, Lieber MR. Interaction of DNA-dependent protein kinase with DNA and with Ku: Biochemical and atomic-force microscopy studies. EMBO J. 1997;16:5098–112.

25. Mahaney BL, Meek K, Lees-Miller SP. Repair of ionizing radiation-induced DNA double-strand breaks by non-homologous end-joining. Biochem. J. 2009;417:639–50.

26. Lieber MR. The mechanism of human nonhomologous DNA End joining. J. Biol. Chem. 2008;283:1–5.

27. West CE, Waterworth WM, Story GW, Sunderland P A, Jiang Q, Bray CM. Disruption of the Arabidopsis AtKu80 gene demonstrates an essential role for AtKu80 protein in efficient repair of DNA double-strand breaks in vivo. Plant J. 2002;31:517–28.

28. West CE, Waterworth WM, Jiang Q, Bray CM. Arabidopsis DNA ligase IV is induced by γ-irradiation and interacts with an Arabidopsis homologue of the double strand break repair protein XRCC4. Plant J. 2000;24:67–78.

29. Bundock P, van Attikum H, Hooykaas PJJ. Increased telomere length and hypersensitivity to DNA damaging agents in an Arabidopsis KU70 mutant. Nucleic Acids Res. 2002;30:3395–400.

30. van Attikum H, Bundock P, Overmeer RM, Lee LY, Gelvin SB, Hooykaas PJJ. The Arabidopsis AtLIG4 gene is required for the repair of DNA damage, but not for the integration of Agrobacterium T-DNA. Nucleic Acids Res. 2003;31:4247–55.

31. Tamura K, Adachi Y, Chiba K, Oguchi K, Takahashi H. Identification of Ku70 and Ku80 homologues in Arabidopsis thaliana: Evidence for a role in the repair of DNA double-strand breaks. Plant J. 2002;29:771–81.

32. Shuman S, Glickman MS. Bacterial DNA repair by non-homologous end joining. Nat. Rev. Microbiol. 2007;5:852–61.

33. Pitcher RS, Brissett NC, Doherty AJ. Nonhomologous End-Joining in Bacteria: A Microbial Perspective. Annu. Rev. Microbiol. 2007;61:259–82.

34. Iliakis G. Backup pathways of NHEJ in cells of higher eukaryotes: Cell cycle dependence. Radiother. Oncol. 2009;92:310–5.

35. Mladenov E, Iliakis G. Induction and repair of DNA double strand breaks: The increasing spectrum of non-homologous end joining pathways. Mutat. Res. 2011;711:61–72.

36. Nussenzweig A, Nussenzweig MC. A Backup DNA Repair Pathway Moves to the Forefront. Cell. 2007;131:223–5.

37. Haber JE. Alternative endings. Proc. Natl. Acad. Sci. 2008;105:405–6.

38. Crespan E, Czabany T, Maga G, Hübscher U. Microhomology-mediated DNA strand annealing and elongation by human DNA polymerases λ and β on normal and repetitive DNA sequences. Nucleic Acids Res. 2012;40:5577–90.

39. Kuhfittig-Kulle S, Feldmann E, Odersky A, Kuliczkowska A, Goedecke W, Eggert A, et al. The mutagenic potential of non-homologous end joining in the absence of the NHEJ core factors Ku70/80, DNA-PKcs and XRCC4-LigIV. Mutagenesis. 2007;22:217–33.

40. McVey M, Lee SE. MMEJ repair of double-strand breaks (director's cut): deleted sequences and alternative endings. Trends Genet. 2008;24:529–38.

41. Deriano L, Roth DB. Modernizing the nonhomologous end-joining repertoire: alternative and classical NHEJ share the stage. Annu. Rev. Genet. 2013;47:433–55.

42. Wood RD, Doublié S. DNA polymerase θ (POLQ), double-strand break repair, and cancer. DNA Repair. 2016;44:22–32.

43. Chang HHY, Pannunzio NR, Adachi N, Lieber MR. Non-homologous DNA end joining and alternative pathways to double-strand break repair. Nat. Rev. Mol. Cell Biol. 2017;18:495-506

44. Pardo B, Gómez-González B, Aguilera a. DNA double-strand break repair: How to fix a broken relationship. Cell. Mol. Life Sci. 2009;66:1039–56.

45. Szostak JW, Orr-Weaver TL, Rothstein RJ, Stahl FW. The double-strand-break repair model for recombination. Cell. 1983;33:25–35.

46. San Filippo J, Sung P, Klein H. Mechanism of eukaryotic homologous recombination. Annu. Rev. Biochem. 2008;77:229–57.

47. Llorente B, Smith CE, Symington LS. Break-induced replication: What is it and what is it for? Cell Cycle. 2008;7:859–64.

48. Bray CM, West CE. DNA repair mechanisms in plants: Crucial sensors and effectors for the maintenance of genome integrity. New Phytol. 2005;168:511–28.

49. Puchta H. Repair of genomic double-strand breaks in somatic plant cells by one-sided invasion of homologous sequences. Plant J. 1998;13:331–9.

50. Puchta H, Hohn B. The mechanism of extrachromosomal homologous DNA recombination in plant cells. Mol. Gen. Genet. 1991;230:1–7.

51. Siebert R, Puchta H. Efficient repair of genomic double-strand breaks by homologous recombination between directly repeated sequences in the plant genome. Plant Cell. 2002;14:1121–31.

52. Baltes NJ, Voytas DF. Enabling plant synthetic biology through genome engineering. Trends Biotechnol. 2015;33:120–31.

53. Edgell DR. Selfish DNA: homing endonucleases find a home. Curr. Biol. 2009;19:115–7.

54. Hafez M, Hausner G, Bonen L. Homing endonucleases: DNA scissors on a mission. Genome. 2012;55:553–69.

55. Stoddard BL. Homing endonucleases: from microbial genetic invaders to reagents for targeted DNA modification. Structure. 2011;19:7–15.

56. Smith J, Grizot S, Arnould S, Duclert A, Epinat JC, Chames P, et al. A combinatorial approach to create artificial homing endonucleases cleaving chosen sequences. Nucleic Acids Res. 2006;34:1–12.

57. Chevalier BS, Kortemme T, Chadsey MS, Baker D, Monnat RJ, Stoddard BL. Design, activity, and structure of a highly specific artificial endonuclease. Mol. Cell. 2002;10:895–905.

58. Diakun G, Fairall L, Klug A. EXAFS study of the zinc-binding sites in the protein transcription factor IIIA. Nature. 1986;324:698–9.

59. Pavletich NP, Pabo CO. Zinc finger-DNA recognition: crystal structure of a Zif268-DNA complex at 2.1 A. Science. 1991;252:809–17.

60. Segal DJ, Dreier B, Beerli RR, Barbas CF. Toward controlling gene expression at will: selection and design of zinc finger domains recognizing each of the 5'-GNN-3' DNA target sequences. Proc. Natl. Acad. Sci. USA. 1999;96:2758–63.

61. Dreier B, Beerli RR, Segal DJ, Flippin JD, Barbas CF. Development of zinc finger domains for recognition of the 5'-ANN-3' family of DNA sequences and their use in the construction of artificial transcription factors. J. Biol. Chem. 2001;276:29466–78.

62. Dreier B, Fuller RP, Segal DJ, Lund C V, Blancafort P, Huber A, et al. Development of zinc finger domains for recognition of the 5'-CNN-3' family DNA sequences and their use in the construction of artificial transcription factors. J. Biol. Chem. 2005;280:35588–97.

63. Beumer K, Bhattacharyya G, Bibikova M, Trautman JK, Carroll D. Efficient gene targeting in Drosophila with zinc-finger nucleases. Genetics. 2006;172:2391–403.

64. Miller JC, Holmes MC, Wang J, Guschin DY, Lee Y-L, Rupniewski I, et al. An improved zinc-finger nuclease architecture for highly specific genome editing. Nat. Biotechnol. 2007;25:778–85.

65. Doyon Y, Vo TD, Mendel MC, Greenberg SG, Wang J, Xia DF, et al. Enhancing zinc-finger-nuclease activity with improved obligate heterodimeric architectures. Nat. Methods. 2011;8:74–9.

66. Ramirez CL, Foley JE, Wright D a, Müller-Lerch F, Rahman SH, Cornu TI, et al. Unexpected failure rates for modular assembly of engineered zinc fingers. Nat. Methods. 2008;5:374–5.

67. Händel E-M, Alwin S, Cathomen T. Expanding or restricting the target site repertoire of zinc-finger nucleases: the inter-domain linker as a major determinant of target site selectivity. Mol. Ther. 2009;17:104–11.

68. Carroll D. Genome engineering with zinc-finger nucleases. Genetics. 2011;188:773–82.

69. Kim J-S, Lee HJ, Carroll D. Genome editing with modularly assembled zinc-finger nucleases. Nat. Methods. 2010;7:91–2.

70. Sander JD, Dahlborg EJ, Goodwin MJ, Cade L, Zhang F, Cifuentes D, et al. Selection-free zinc-finger-nuclease engineering by context-dependent assembly (CoDA). Nat. Methods. 2011;8:67–9.

71. Boller T. Innate immunity in plants : an arms race betweeen pattern recognition recepters in plant and effectors in microbial pathogens. Plant-Microbe Interact. 2009;324:742–4.

72. Göhre V, Robatzek S. Breaking the barriers: microbial effector molecules subvert plant immunity. Annu. Rev. Phytopathol. 2008;46:189–215.

73. Kay S, Bonas U. How Xanthomonas type III effectors manipulate the host plant. Curr. Opin. Microbiol. 2009;12:37–43.

74. Boch J, Scholze H, Schornack S, Landgraf A, Hahn S, Kay S, et al. Breaking the code of DNA binding specificity of TAL-type III effectors. Science. 2009;326:1509–12.

75. Moscou M, Bogdanove A. Recognition by TAL Effectors. Science. 2009;326:1501.

76. Deng D, Yan C, Pan X, Mahfouz M, Wang J, Zhu J, et al. Structural basis for sequence-specific recognition of DNA by TAL effectors. Science. 2012;335:720–3.

77. Mak AN-S, Bradley P, Cernadas RA, Bodanove AJ, Stoddard BL. The crystal structure of TAL effector PthXo1 bound to its DNA target. Science. 2012;335:716–9.

78. Christian M, Cermak T, Doyle EL, Schmidt C, Zhang F, Hummel A, et al. Targeting DNA double-strand breaks with TAL effector nucleases. Genetics. 2010;186:757–61.

79. Wiedenheft B, Sternberg SH, Doudna JA. RNA-guided genetic silencing systems in bacteria and archaea. Nature. 2012;482:331–8.

80. Bhaya D, Davison M, Barrangou R. CRISPR-Cas systems in bacteria and archaea: versatile small RNAs for adaptive defense and regulation. Annu. Rev. Genet. 2011;45:273–97.

81. Terns MP, Terns RM. CRISPR-based adaptive immune systems. Curr. Opin. Microbiol. 2011;14:321–7.

82. Deltcheva E, Chylinski K, Sharma CM, Gonzales K. CRISPR RNA maturation by trans -encoded small RNA and host factor RNase III. Nature. 2011;471:602–7.

83. Deltcheva E, Chylinski K, Sharma CM, Gonzales K, Chao Y, Pirzada ZA, et al. CRISPR RNA maturation by trans-encoded small RNA and host factor RNase III. Nature. 2011;471:602–7.

84. Brouns SJJ, Jore MM, Lundgren M, Westra ER, Slijkhuis RJH, Snijders APL, et al. Small CRISPR RNAs guide antiviral defense in prokaryotes. Science. 2008;321:960–4.

85. Jinek M, Chylinski K, Fonfara I, Hauer M, Doudna JA, Charpentier E. A programmable dual-RNA–guided DNA endonuclease in adaptive bacterial immunity. Science. 2012;337:816–22.

86. Gasiunas G, Barrangou R, Horvath P, Siksnys V. Cas9-crRNA ribonucleoprotein complex mediates specific DNA cleavage for adaptive immunity in bacteria. Proc. Natl. Acad. Sci. USA. 2012;109:E2579-

87. Hsu PD, Scott D a, Weinstein J a, Ran FA, Konermann S, Agarwala V, et al. DNA targeting specificity of RNA-guided Cas9 nucleases. Nat. Biotechnol. 2013;31:827–32.

88. Cho SW, Lee J, Carroll D, Kim JS, Lee J. Heritable gene knockout in Caenorhabditis elegans by direct injection of Cas9-sgRNA ribonucleoproteins. Genetics. 2013;195:1177–80.

89. Cong L, Ran FA, Cox D, Lin S, Barretto R, Hsu PD, et al. Multiplex genome engineering using CRISPR/Cas systems. Science. 2013;339:819–23.

90. Hwang WY, Fu Y, Reyon D, Maeder ML, Tsai SQ, Sander JD, et al. Efficient genome editing in zebrafish using a CRISPR-Cas system. Science. Nat. Biotechnol. 2013;31:227–9.

91. Jinek M, East A, Cheng A, Lin S, Ma E, Doudna J. RNA-programmed genome editing in human cells. Elife. 2013;2013:1–9.

92. Dicarlo JE, Norville JE, Mali P, Rios X, Aach J, Church GM. Genome engineering in Saccharomyces cerevisiae using CRISPR-Cas systems. Nucleic Acids Res. 2013;41:4336–43.

93. Sapranauskas R, Gasiunas G, Fremaux C, Barrangou R, Horvath P, Siksnys V. The Streptococcus thermophilus CRISPR/Cas system provides immunity in Escherichia coli. Nucleic Acids Res. 2011;39:9275–82.

94. Sander JD, Joung JK. CRISPR-Cas systems for editing, regulating and targeting genomes. Nat. Biotechnol. 2014;32:347–55.

95. Wolt JD, Wang K, Sashital D, Lawrence-Dill CJ. Achieving Plant CRISPR Targeting that Limits Off-Target Effects. Plant Genome. 2016;9:0.

96. Fan D, Liu T, Li C, Jiao B, Li S, Hou Y, et al. Efficient CRISPR/Cas9-mediated targeted mutagenesis in populus in the first generation. Sci. Rep. 2015;5:12217.

97. Svitashev S, Young JK, Schwartz C, Gao H, Falco SC, Cigan AM. Targeted mutagenesis, precise gene editing, and site-specific gene insertion in maize using Cas9 and guide RNA. Plant Physiol. 2015;169:931–45.

98. Xu R, Li H, Qin R, Li J, Qiu C, Yang Y, et al. Generation of inheritable and " transgene clean " targeted genome-modified rice in later generations using the CRISPR / Cas9 system. Sci. Rep. 2015;9:1–10.

99. Zhang H, Zhang J, Wei P, Zhang B, Gou F, Feng Z, et al. The CRISPR/Cas9 system produces specific and homozygous targeted gene editing in rice in one generation. Plant Biotechnol. J. 2014;12:797–807.

100. Jiang W, Yang B, Weeks DP. Efficient CRISPR/Cas9-mediated gene editing in Arabidopsis thaliana and inheritance of modified genes in the T2 and T3 generations. PLoS One. 2014;9:21–6.

101. Wang Z, Xing H, Dong L, Zhang H, Han C, Wang X, et al. Egg cell-specific promoter-controlled CRISPR / Cas9 efficiently generates homozygous mutants for multiple target genes in Arabidopsis in a single generation. Genome Biol. Genome Biology; 2015;16:1–12.

102. Lee HJ, Kweon J, Kim E, Kim S, Kim J. Targeted chromosomal duplications and inversions in the human genome using zinc finger nucleases. Genome Res. 2012;539–48.

103. Brunet E, Simsek D, Tomishima M, DeKelver R, Choi VM, Gregory P, et al. Chromosomal translocations induced at specified loci in human stem cells. Proc. Natl. Acad. Sci. USA. 2009;106:10620–5.

104. Lowder LG, Zhang D, Baltes NJ, Paul JW, Tang X, Zheng X, et al. A CRISPR/Cas9 toolbox for multiplexed plant genome editing and transcriptional regulation. Plant Physiol. 2015;169:971–85.

105. Mercx S, Tollet J, Magy B, Navarre C, Boutry M. Gene Inactivation by CRISPR-Cas9 in Nicotiana tabacum BY-2 suspension cells. Front. Plant Sci. 2016;7:40.

106. Zhang Z, Mao Y, Ha S, Liu W, Botella JR, Zhu JK. A multiplex CRISPR/Cas9 platform for fast and efficient editing of multiple genes in Arabidopsis. Plant Cell Rep. 2015;1–15.

107. Puchta H, Dujon B, Hohn B. Homologous recombination in plant cells is enhanced by in vivo induction of double strand breaks into DNA by a site-specific endonuclease. Nucleic Acids Res. 1993;21:5034–40.

108. Puchta H, Dujon B, Hohn B. Two different but related mechanisms are used in plants for the repair of genomic double-strand breaks by homologous recombination. Proc. Natl. Acad. Sci. USA. 1996;93:5055–60.

109. de Pater S, Neuteboom LW, Pinas JE, Hooykaas PJJ, van der Zaal BJ. ZFN-induced mutagenesis and gene-targeting in Arabidopsis through Agrobacterium-mediated floral dip transformation. Plant Biotechnol. J. 2009;7:821–35.

110. Pater S De, Pinas JE, Hooykaas PJJ, Zaal BJ Van Der. ZFN-mediated gene targeting of the Arabidopsis protoporphyrinogen oxidase gene through Agrobacterium -mediated floral dip transformation. Plant Cell Physiol. 2013;11:510–5.

111. Townsend J a, Wright D a, Winfrey RJ, Fu F, Maeder ML, Joung JK, et al. High-frequency modification of plant genes using engineered zinc-finger nucleases. Nature. 2009;459:442–5.

112. Shukla VK, Doyon Y, Miller JC, DeKelver RC, Moehle E a, Worden SE, et al. Precise genome modification in the crop species Zea mays using zinc-finger nucleases. Nature. 2009;459:437–41.

113. Zhang Y, Zhang F, Li X, Baller J a., Qi Y, Starker CG, et al. Transcription activator-like effector nucleases enable efficient plant genome engineering. Plant Physiol. 2013;161:20–7.

114. Čermák T, Baltes NJ, Čegan R, Zhang Y, Voytas DF. High-frequency, precise modification of the tomato genome. Genome Biol. 2015;16:232.

115. Li W, Teng F, Li T, Zhou Q. Simultaneous generation and germline transmission of multiple gene mutations in rat using CRISPR-Cas systems. Nat. Biotechnol. 2013;31:684–6.

116. Fauser F, Schiml S, Puchta H. Both CRISPR/Cas-based nucleases and nickases can be used efficiently for genome engineering in Arabidopsis thaliana. Plant J. 2014;79:348–59.

117. Schiml S, Fauser F, Puchta H. The CRISPR/Cas system can be used as nuclease for in planta gene targeting and as paired nickases for directed mutagenesis in Arabidopsis resulting in heritable progeny. Plant J. 2014;80:1139–50.

118. Sun Y, Zhang X, Wu C, He Y, Ma Y, Hou H, et al. Engineering Herbicide Resistant Rice Plants through CRISPR/Cas9-mediated Homologous Recombination of the Acetolactate Synthase. Mol. Plant. 2016;1–4.

119. Butler NM, Atkins PA, Voytas DF, Douches DS. Generation and inheritance of targeted mutations in potato (Solanum tuberosum L.) Using the CRISPR/Cas System. PLoS One. 2015;10:1–12.

120. Shan Q, Wang Y, Li J, Zhang Y, Chen K, Liang Z, et al. Targeted genome modification of crop plants using a CRISPR-Cas system. Nat. Biotechnol. 2013;31:686–8.

121. Niu Y, Shen B, Cui Y, Chen Y, Wang J, Wang L, et al. Generation of gene-modified cynomolgus monkey via Cas9/RNA-mediated gene targeting in one-cell embryos. Cell. 2014;156:836–43.

122. Meng X, Noyes MB, Zhu LJ, Lawson ND, Wolfe SA. Targeted gene inactivation in zebrafish using engineered zinc-finger nucleases. Nat. Biotechnol. 2008;26:695–701.

123. Lee J-S, Kwak S-J, Kim J, Noh HM, Kim J-S, Yu K. RNA-guided genome editing in Drosophila with the purified Cas9 protein. G3 (Bethesda). 2014;4:1291–5.

124. Zuris JA, Thompson DB, Shu Y, Guilinger JP, Bessen JL, Hu JH, et al. Efficient delivery of genome-editing proteins in vitro and in vivo. Nat. Biotechnol. 2015;33:73–80.

125. Luo S, Li J, Stoddard TJ, Baltes NJ, Demorest ZL, Clasen BM, et al. Non-transgenic plant genome editing using purified sequence-specific nucleases. Mol. Plant. ; 2015;8:1425–7.

126. Woo JW, Kim J, Kwon S Il, Corvalán C, Cho SW, Kim H, et al. DNA-free genome editing in plants with preassembled CRISPR-Cas9 ribonucleoproteins. Nat. Biotechnol. 2015;33:1162–4.

127. Braun AC. A physiological basis for autonomous growth of the crown-gall tumor cell. Proc. Natl. Acad. Sci. USA. 1958;44:344–9.

128. Zambryski P, Tempe J, Schell J. Transfer and function of T-DNA genes from Agrobacterium Ti and Ri plasmids in plants. Cell. 1989;56:193–201.

129. Chilton M-D, Drummond MH, Merlo DJ, Sciaky D, Montoya AL, Gordon MP, et al. Stable incorporation of plasmid DNA into higher plant cells: the molecular basis of crown gall tumorigenesis. Cell. 1977;11:263–71.

130. Britton MT, Escobar MA, Dandekar AM. The oncogenes of Agrobacterium tumefaciens and Agrobacterium rhizogenes. In: Tzfira T, Citovsky V, editors. Agrobacterium From Biol. to Biotechnol. New York, NY: Springer New York; 2008. p. 523–63.

131. Dessaux Y, Petit A, Farrand SK, Murphy PJ. Opines and Opine-Like Molecules Involved in Plant-Rhizobiaceae Interactions. In: Spaink HP, Kondorosi A, Hooykaas PJJ, editors. Rhizobiaceae Mol. Biol. Model Plant-Associated Bact. Dordrecht: Springer Netherlands; 1998. p. 173–97.

132. Stachel SE, Messens E, Van Montagu M, Zambryski P. Identification of the signal molecules produced by wounded plant cells that activate T-DNA transfer in Agrobacterium tumefaciens. Nature. 1985;318:624–9.

133. Melchers LS, Regensburg-Tuïnk TJ, Bourret RB, Sedee NJ, Schilperoort RA, Hooykaas PJJ. Membrane topology and functional analysis of the sensory protein VirA of Agrobacterium tumefaciens. EMBO J. 1989;1919–25.

134. Leroux B, Yanofsky MF, Winans SC, Ward JE, Ziegler SF, Nester EW. Characterization of the virA locus of Agrobacterium tumefaciens: a transcriptional regulator and host range determinant. EMBO J. 1987;6:849–56.

135. Melchers LS, Thompson D V., Idler KB, Neuteboom STC, de Maagd RA., Schilperoort RA, et al. Molecular characterization of the virulence gene virA of the Agrobacterium tumefaciens octopine Ti plasmid. Plant Mol. Biol. 1988;11:227–37.

136. Lee YW, Jin S, Sim WS, Nester EW. The sensing of plant signal molecules by Agrobacterium: genetic evidence for direct recognition of phenolic inducers by the VirA protein. Gene. 1996;179:83–8.

137. Cangelosi G a, Ankenbauer RG, Nester EW. Sugars induce the Agrobacterium virulence genes through a periplasmic binding protein and a transmembrane signal protein. Proc. Natl. Acad. Sci. USA. 1990;87:6708–12.

138. Stachel SE, Nester EW. The genetic and transcriptional organization of the vir region of the A6 Ti plasmid of Agrobacterium tumefaciens. EMBO J. 1986;5:1445–54.

139. Hooykaas P, Hofker M, den Dulk-Ras A, Schilperoort RA. A comparison of virulence determinants in an octopine Ti plasmid, a nopaline Ti plasmid, and an Ri plasmid by complementation analysis of Agrobacterium tumefaciens mutants. Plasmid. 1984;11:195–205.

140. Kanemoto RH, Powell a. T, Akiyoshi DE, Regier D a., Kerstetter R a., Nester EW, et al. Nucleotide sequence and analysis of the plant-inducible locus pinF from Agrobacterium tumefaciens. J. Bacteriol. 1989;171:2506–12.

141. van Haaren MJJ, Pronk JT, Schilperoort RA, Hooykaas PJJ. Functional analysis of the Agrobacterium tumefaciens octopine Ti-plasmid left and right T-region border fragments. Plant Mol. Biol. 1987;8:95–104.

142. Ward ER, Barnes WM. VirD2 protein of Agrobacterium tumefaciens very tightly linked to the 5'end of T-strand DNA. Science. 1988;242:927–30.

143. Pansegrau W, Schoumacher F, Hohn B, Lanka E. Site-specific cleavage and joining of single-stranded DNA by VirD2 protein of Agrobacterium tumefaciens Ti plasmids: analogy to bacterial conjugation. Proc. Natl. Acad. Sci. USA. 1993;90:11538–42.

144. Toro N, Datta A, Carmi OA, Young C, Prusti RK, Nester EW. The Agrobacterium tumefaciens virC1 gene product binds to overdrive, a T-DNA transfer enhancer. J. Bacteriol. 1989;171:6845–9.

145. Lu J, den Dulk-Ras A, Hooykaas PJJ, Glover JNM, Dulk-ras A Den, Hooykaas PJJ, et al. Agrobacterium tumefaciens VirC2 enhances T-DNA transfer and virulence through its C-terminal ribbon-helix-helix DNA-binding fold. Proc. Natl. Acad. Sci. USA. 2009;106:9643–8.

146. van Kregten M, Lindhout BI, Hooykaas PJJ, van der Zaal BJ. Agrobacterium-mediated T-DNA transfer and integration by minimal VirD2 consisting of the relaxase domain and a type IV secretion system translocation signal. Mol. Plant. Microbe. Interact. 2009;22:1356–65.

147. Christie PJ, Whitaker N, González-Rivera C. Mechanism and structure of the bacterial type IV secretion systems. Biochim. Biophys. Acta. 2014;1843:1578–91.

148. Vergunst a. C. VirB/D4-dependent protein translocation from Agrobacterium into plant cells. Science. 2000;290:979–82.

149. Vergunst AC, van Lier MCM, den Dulk-Ras A, Stüve T a G, Ouwehand A, Hooykaas PJJ. Positive charge is an important feature of the C-terminal transport signal of the VirB/D4-translocated proteins of Agrobacterium. Proc. Natl. Acad. Sci. USA. 2005;102:832–7.

150. Schrammeijer B. Analysis of Vir protein translocation from Agrobacterium tumefaciens using Saccharomyces cerevisiae as a model: evidence for transport of a novel effector protein VirE3. Nucleic Acids Res. 2003;31:860–8.

151. Zupan JR, Citovsky V, Zambryski P. Agrobacterium VirE2 protein mediates nuclear uptake of single-stranded DNA in plant cells. Proc. Natl. Acad. Sci. USA. 1996;93:2392–7.

152. Ward D V., Zupan JR, Zambryski PC. Agrobacterium VirE2 gets the VIP1 treatment in plant nuclear import. Trends Plant Sci. 2002;7:1–3.

153. Lacroix B, Vaidya M, Tzfira T, Citovsky V. The VirE3 protein of Agrobacterium mimics a host cell function required for plant genetic transformation. Embo J. 2005;24:428–37.

154. García-Rodríguez FM, Schrammeijer B, Hooykaas PJJ, Garcı FM. The Agrobacterium VirE3 effector protein: a potential plant transcriptional activator. Nucleic Acids Res. 2006;34:6496–504.

155. Niu X, Zhou M, Henkel C V., Van Heusden GPH, Hooykaas PJJ. The Agrobacterium tumefaciens virulence protein VirE3 is a transcriptional activator of the F-box gene VBF. Plant J. 2015;84:914–24.

156. Zhang X. Functional analysis of Agrobacterium tumefaciens virulence protein VirD5. Leiden University; 2016.

157. Schrammeijer B, Risseeuw E, Pansegrau W, Regensburg-Tuïnk TJG, Crosby WL, Hooykaas PJJ. Interaction of the virulence protein VirF of agrobacterium tumefaciens with plant homologs of the yeast Skp1 protein. Curr. Biol. 2001;11:258–62.

158. Mycobacterium PET, Biochemistry TD, Tzfira T, Vaidya M, Citovsky V. Involvement of targeted proteolysis in plant genetic transformation by Agrobacterium. Nature. 2004;431:6–11.

159. Garcia-Cano E, Magori S, Sun Q, Ding Z, Lazarowitz SG, Citovsky V. Interaction of arabidopsis trihelix-domain transcription factors VFP3 and VFP5 with agrobacterium virulence protein VirF. PLoS One. 2015;10:1–23.

160. Zaltsman A, Krichevsky A, Kozlovsky S V, Yasmin F, Citovsky V. Plant defense pathways subverted by Agrobacterium for genetic transformation. Plant Signal. Behav. 2010;5:1245–8.

161. Cascales E, Christie PJ. Definition of a bacterial type IV secretion pathway for a DNA substrate. Science. 2004;304:1170–3.

162. Jakubowski SJ, Krishnamoorthy V, Cascales E, Christie PJ. Agrobacterium tumefaciens VirB6 Domains Direct the Ordered Export of a DNA Substrate Through a Type IV Secretion System. Jounral Mol. Biol. 2004;341:961–77.

163. Atmakuri K, Ding Z, Christie PJ. VirE2, a Type IV secretion substrate, interacts with the VirD4 transfer protein a cell poles of Agrobacterium tumefaciens. Changes. 2003;49:1699–713.

164. Abu-Arish A, Frenkiel-Krispin D, Fricke T, Tzfira T, Citovsky V, Wolf SG, et al. Three-dimensional reconstruction of Agrobacterium VirE2 protein with single-stranded DNA. J. Biol. Chem. 2004;279:25359–63.

165. Gelvin SB. Agrobacterium and plant genes involved in T-DNA transfer and integration. Annu. Rev. Plant Biol. 2000; 51:233-256

166. Bundock P, den Dulk-Ras A, Beijersbergen A, Hooykaas PJJ. Trans-kingdom T-DNA transfer from Agrobacterium tumefaciens to Saccharomyces cerevisiae. EMBO J. 1995;14:3206–14.

167. van Attikum H, Bundock P, Hooykaas PJJ. Non-homologous end-joining proteins are required for Agrobacterium T-DNA integration. EMBO J. 2001;20:6550–8.

168. van Attikum H, Hooykaas PJJ. Genetic requirements for the targeted integration of Agrobacterium T-DNA in Saccharomyces cerevisiae. Nucleic Acids Res. 2003;31:826–32.

169. Friesner J, Britt AB. Ku80- and DNA ligase IV-deficient plants are sensitive to ionizing radiation and defective in T-DNA integration. Plant J. 2003;34:427–40.

170. Gallego ME, Bleuyard JY, Daoudal-Cotterell S, Jallut N, White CI. Ku80 plays a role in non-homologous recombination but is not required for T-DNA integration in Arabidopsis. Plant J. 2003;35:557–65.

171. Li JX, Vaidya M, White C, Vainstein A, Citovsky V, Tzfira T. Involvement of KU80 in T-DNA integration in plant cells. Proc. Natl. Acad. Sci. USA. 2005;102:19231–6.

172. Jia Q, Bundock P, Hooykaas PJJ, de Pater S. Agrobacterium tumefaciens T-DNA Integration and Gene Targeting in Arabidopsis thaliana Non-Homologous End-Joining Mutants. J. Bot. 2012;2012:1–13.

173. Mestiri I, Norre F, Gallego ME, White CI. Multiple host-cell recombination pathways act in Agrobacterium-mediated transformation of plant cells. Plant J. 2014;77:511–20.

174. Park SY, Vaghchhipawala Z, Vasudevan B, Lee LY, Shen Y, Singer K, et al. Agrobacterium T-DNA integration into the plant genome can occur without the activity of key non-homologous end-joining proteins. Plant J. 2015;81:934–46.

175. van Kregten M, de Pater S, Romeijn R, van Schendel R, Hooykaas PJJ, Tijsterman M. T-DNA integration in plants results from polymerase-θ-mediated DNA repair. Nat. plants. 2016;2:16164.

176. Hoekema A, Hirsch PR, Hooykaas PJJ, Schilperoort RA. A binary plant vector strategy based on separtion of vir- and T-region of the Agrobacterium tumefaciens Ti-plasmid. Nature. 1983;302:179–80.

177. Krenek P, Samajova O, Luptovciak I, Doskocilova A, Komis G, Samaj J. Transient plant transformation mediated by Agrobacterium tumefaciens: Principles, methods and applications. Biotechnol. Adv. 2015;33:1024–42.

178. Koncz C, Németh K, Rédei GP, Schell J. T-DNA insertional mutagenesis in Arabidopsis. Plant Mol. Biol. 1992;20:963–76.

179. Walden R, Hayashi H, Schell J. T-DNA as a gene tag. Plant J. 1991;1:281–8.

180. Slogteren GH-V, Hooykaas P. Expression of Ti plasmid genes in monocotyledonous plants infected with Agrobacterium tumefaciens. Nature. 1984;311.

181. Michielse CB, Hooykaas PJJ, van den Hondel C a MJJ, Ram AFJ. Agrobacterium-mediated transformation as a tool for functional genomics in fungi. Curr. Genet. 2005;48:1–17.

182. Rolloos M, Hooykaas PJJ, van der Zaal BJ. Enhanced targeted integration mediated by translocated I-SceI during the Agrobacterium mediated transformation of yeast. Sci. Rep. 2015;5:8345.

183. Kregten van M, Boer de P, Pinas JE, Hooykaas PJJ, van der Zaal BJ. Agrobacterium-mediated delivery of a meganuclease into target plant cells. 2011.
184. Kregten van M, Boer de P, Hooykaas PJJ, van der Zaal BJ. Translocation of novel recombinant effector proteins from Agrobacterium tumefaciens to Arabidopsis thaliana. 2011.
185. Antunes MS, Smith JJ, Jantz D, Medford JI. Targeted DNA excision in Arabidopsis by a re-engineered homing endonuclease. BMC Biotechnol. 2012;12:86.
186. Yang M, Djukanovic V, Stagg J, Lenderts B, Bidney D, Carl Falco S, et al. Targeted mutagenesis in the progeny of maize transgenic plants. Plant Mol. Biol. 2009;70:669–79.
187. Gao H, Smith J, Yang M, Jones S, Djukanovic V, Nicholson MG, et al. Heritable targeted mutagenesis in maize using a designed endonuclease. Plant J. 2010;61:176–87.
188. Lloyd A, Plaisier CL, Carroll D, Drews GN. Targeted mutagenesis in Arabidopsis using zinc-finger nucleases. Methods Mol. Biol. 2005;102:2232–7.
189. Zhang F, Maeder ML, Unger-wallace E, Hoshaw JP, Reyon D, Christian M, et al. High frequency targeted mutagenesis in Arabidopsis thaliana using zinc finger nucleases. Proc. Natl. Acad. Sci. 2010;107:12028–33.
190. Osakabe K, Osakabe Y, Toki S. Site-directed mutagenesis in Arabidopsis using custom-designed zinc finger nucleases. Proc. Natl. Acad. Sci. 2010;107:12034–9.
191. Curtin SJ, Zhang F, Sander JD, Haun WJ, Starker C, Baltes NJ, et al. Targeted mutagenesis of duplicated genes in soybean with zinc-finger nucleases. Plant Physiol. 2011;156:466–73.
192. Christian M, Qi Y, Zhang Y, Voytas DF. Targeted mutagenesis of Arabidopsis thaliana using engineered TAL effector nucleases. G3 (Bethesda). 2013;3:1697–705.
193. Forner J, Pfeiffer A, Langenecker T, Manavella P. Germline-Transmitted Genome Editing in Arabidopsis thaliana Using TAL-Effector- Nucleases. PloS One. 2015;1–15.
194. Cermak T, Doyle EL, Christian M, Wang L, Zhang Y, Schmidt C, et al. Efficient design and assembly of custom TALEN and other TAL effector-based constructs for DNA targeting. Nucleic Acids Res. 2011;39:e82.
195. Zhang Y, Shan Q, Wang Y, Chen K, Liang Z, Li J, et al. Rapid and efficient gene modification in rice and brachypodium using TALENs. Mol. Plant. 2013;6:1365–8.
196. Du H, Zeng X, Zhao M, Cui X, Wang Q, Yang H, et al. Efficient targeted mutagenesis in soybean by TALENs and CRISPR/Cas9. J. Biotechnol. 2016;217:90–7.
197. Haun W, Coffman A, Clasen BM, Demorest ZL, Lowy A, Ray E, et al. Improved soybean oil quality by targeted mutagenesis of the fatty acid desaturase 2 gene family. Plant Biotechnol. J. 2014;2023:934–40.
198. Budhagatapalli N, Rutten T, Gurushidze M, Kumlehn J, Hensel G. Targeted modification of gene function exploiting homology-directed repair of TALEN-mediated double strand breaks in barley. G3 (Bethesda). 2015;5:1857–63.
199. Wendt T, Holm PB, Starker CG, Christian M, Voytas DF, Brinch-Pedersen H, et al. TAL effector nucleases induce mutations at a pre-selected location in the genome of primary barley transformants. Plant Mol. Biol. 2013;83:279–85.
200. Gurushidze M, Hensel G, Hiekel S, Schedel S, Valkov V, Kumlehn J. True-Breeding Targeted Gene Knock-Out in Barley Using Designer TALE-Nuclease in Haploid Cells. PLoS One. 2014;9:1–9.
201. Stoddard TJ, Clasen BM, Baltes B, Demorest ZL, Voytas DF, Zhang F, et al. Targeted Mutagenesis in Plant Cells through Transformation of Sequence-Specific Nuclease mRNA. 2016
202. Li J, Stoddard TJ, Demorest ZL, Lavoie P-O, Luo S, Clasen BM, et al. Multiplexed, targeted gene editing in Nicotiana benthamiana for glyco-engineering and monoclonal antibody production. Plant Biotechnol. J. 2015
203. Zhang H, Gou F, Zhang J, Liu W, Li Q, Mao Y, et al. TALEN-mediated targeted mutagenesis produces a large variety of heritable mutations in rice. Plant Biotechnol. J. 2015;1–9.
204. Shan Q, Zhang Y, Chen K, Zhang K, Gao C. Creation of fragrant rice by targeted knockout of the OsBADH2 gene using TALEN technology. Plant Biotechnol. J. 2015;2:791–800.
205. Li Ting, Liu Bo, Chen Chih Ying YB. TALEN-mediated homologous recombination produces site-directed DNA base change nd herbicide-resistant rice. J. Genet. Genomics. 2016;1–9.
206. Chen K, Shan Q, Gao C. An efficient TALEN mutagenesis system in rice. Methods. 2014;69:2–8.
207. Nicolia A, Proux-wéra E, Åhman I, Onkokesung N, Andersson M, Andreasson E, et al. Targeted gene mutation in tetraploid potato through transient TALEN expression in protoplasts. J. Biotechnol. 2015;204:17–24.
208. Wang Y, Cheng X, Shan Q, Zhang Y, Liu J, Gao C, et al. Simultaneous editing of three homoeoalleles in hexaploid bread wheat confers heritable resistance to powdery mildew. Nat. Biotechnol. 2014; 2014;32;947-951.

209. Char SN, Unger-wallace E, Frame B, Briggs SA, Main M, Spalding MH, et al. Heritable site-specific mutagenesis using TALENs in maize. 2015;1002–10.

210. Feng Z, Zhang B, Ding W, Liu X, Yang, Dong-Lei, Wei P, Cao F, et al. Efficient genome editing in plants using a CRISPR/Cas system. Cell Res. 2013;23:1229–32.

211. Ma X, Zhang Q, Zhu Q, Liu W, Chen Y, Qiu R, et al. A robust CRISPR/Cas9 system for convenient, high-efficiency multiplex genome editing in monocot and dicot plants. Mol. Plant. ; 2015;8:1274–84.

212. Mao Y, Zhang H, Xu N, Zhang B, Gou F, Zhu JK. Application of the CRISPR-Cas system for efficient genome engineering in plants. Mol. Plant. 2013;6:2008–11.

213. Mao Y, Zhang Z, Feng Z, Wei P, Zhang H, Botella JR, et al. Development of germ-line-specific CRISPR-Cas9 systems to improve the production of heritable gene modifications in Arabidopsis. Plant Biotechnol. J. 2015;519–32.

214. Ryder P, McHale M, Fort A, Spillane C. Generation of stable nulliplex autopolyploid lines of Arabidopsis thaliana using CRISPR/Cas9 genome editing. Plant Cell Rep. 2017;36:1005–8.

215. Shi Y, Huang J, Sun T, Wang X, Zhu C, Ai Y, et al. The precise regulation of different COR genes by individual CBF transcription factors in Arabidopsis thaliana. J. Integr. Plant Biol. 2017;59:118–33.

216. Tsutsui H, Higashiyama T. PKAMA-ITACHI vectors for highly efficient CRISPR/Cas9-mediated gene knockout in Arabidopsis thaliana. Plant Cell Physiol. 2017;58:46–56.

217. Xing H, Dong L, Wang Z, Zhang H, Han C, Liu B, et al. A CRISPR / Cas9 toolkit for multiplex genome editing in plants. BMC Plant Biol. 2014;14:1–12.

218. Yan L, Wei S, Wu Y, Hu R, Li H, Yang W, et al. High-efficiency genome editing in Arabidopsis using YAO promoter-driven CRISPR/Cas9 system. Mol. Plant. 2015;8:1820–3.

219. Hyun Y, Kim J, Woo S. Site-directed mutagenesis in Arabidopsis thaliana using dividing tissue-targeted RGEN of the CRISPR / Cas system to generate heritable null alleles. Planta. 2015;241:271–84.

220. Schiml S, Fauser F, Puchta H. The CRISPR/Cas system can be used as nuclease for in planta gene targeting and as paired nickases for directed mutagenesis in Arabidopsis resulting in heritable progeny. Plant J. 2014;80:1139–50.

221. Steinert J, Schiml S, Fauser F, Puchta H. Highly efficient heritable plant genome engineering using Cas9 orthologues from Streptococcus thermophilus and Staphylococcus aureus. Plant J. 2015;84:1295–305.

222. Ordon J, Gantner J, Kemna J, Schwalgun L, Reschke M, Streubel J, et al. Generation of chromosomal deletions in dicotyledonous plants employing a user-friendly genome editing toolkit. Plant J. 2017;89:155–68.

223. Jiang W, Zhou H, Bi H, Fromm M, Yang B, Weeks DP. Demonstration of CRISPR/Cas9/sgRNA-mediated targeted gene modification in Arabidopsis, tobacco, sorghum and rice. Nucleic Acids Res. 2013;41:1–12.

224. Li JF, Zhang D, Sheen J. Cas9-based genome editing in arabidopsis and tobacco. Methods Enzymol. 1st ed. 2014;546:459–72.

225. Jiang WZ, Henry IM, Lynagh PG, Comai L, Cahoon EB, Weeks DP. Significant enhancement of fatty acid composition in seeds of the allohexaploid, Camelina sativa, using CRISPR/Cas9 gene editing. Plant Biotechnol. J. 2017;15:648–57.

226. Tian S, Jiang L, Gao Q, Zhang J, Zong M, Zhang H, et al. Efficient CRISPR/Cas9-based gene knockout in watermelon. Plant Cell Rep. 2017;36:399–406.

227. Jia H, Wang N. Targeted genome editing of sweet orange using Cas9/sgRNA. PLoS One. 2014;9:e93806.

228. Jia H, Zhang Y, Orbović V, Xu J, White FF, Jones JB, et al. Genome editing of the disease susceptibility gene CsLOB1 in citrus confers resistance to citrus canker. Plant Biotechnol. J. 2017;15:817–23.

229. Chandrasekaran J, Brumin M, Wolf D, Leibman D, Klap C, Pearlsman M, et al. Development of broad virus resistance in non-transgenic cucumber using CRISPR/Cas9 technology. Mol. Plant Pathol. 2016;1–14.

230. Kishi-Kaboshi M, Aida R, Sasaki K. Generation of gene-edited chrysanthemum morifolium using multicopy transgenes as targets and markers. Plant Cell Physiol. 2017;58:216–26.

231. Kui L, Chen H, Zhang W, He S, Xiong Z, Zhang Y, et al. Building a Genetic Manipulation Tool Box for Orchid Biology: Identification of Constitutive Promoters and Application of CRISPR/Cas9 in the Orchid, Dendrobium officinale. Front. Plant Sci. 2017;7:1–13.

232. Cai Y, Chen L, Liu X, Sun S, Wu C, Jiang B, et al. CRISPR/Cas9-mediated genome editing in soybean hairy roots. PLoS One. 2015;10:1–13.

233. Cai Y, Chen L, Liu X, Guo C, Sun S, Wu C, et al. CRISPR/Cas9-mediated targeted mutagenesis of GmFT2a delays flowering time in soya bean. Plant Biotechnol. J. 2017;1–10.

234. Jacobs TB, Lafayette PR, Schmitz RJ, Parrott WA. Targeted genome modifications in soybean with CRISPR/Cas9. BMC Biotechnol. 2015;15:1–10.

235. Kim H, Kim S-T, Ryu J, Kang B-C, Kim J-S, Kim S-G. CRISPR/Cpf1-mediated DNA-free plant genome editing. Nat. Commun. 2017;8:14406.
236. Li Z, Liu Z, Xing A, Moon BP, Koellhoffer JP, Huang L, et al. Cas9-guide RNA directed genome editing in soybean. Plant Physiol. 2015;169:960–70.
237. Michno J-M, Wang X, Liu J, Curtin SJ, Kono TJY, Stupar RM. CRISPR/Cas mutagenesis of soybean and Medicago truncatula using a new web-tool and a modified Cas9 enzyme. GM Crops Food. 2015;6;243-52.
238. Chen X, Lu X, Shu N, Wang S, Wang J, Wang D, et al. Targeted mutagenesis in cotton (Gossypium hirsutum L.) using the CRISPR/Cas9 system. Sci. Rep. 2017;7:44304.
239. Janga MR, Campbell LM, Rathore KS. CRISPR/Cas9-mediated targeted mutagenesis in upland cotton (Gossypium hirsutum L.). Plant Mol. Biol. 2017;94;349-360.
240. Wang P, Zhang J, Sun L, Ma Y, Xu J, Liang S, et al. High efficient multisites genome editing in allotetraploid cotton (Gossypium hirsutum) using CRISPR/Cas9 system. Plant Biotechnol. J. 2017;1–14.
241. Kapusi E, Corcuera-Gómez M, Melnik S, Stoger E. Heritable Genomic Fragment Deletions and Small Indels in the Putative ENGase Gene Induced by CRISPR/Cas9 in Barley. Front. Plant Sci. 2017;8:540.
242. Meng Y, Hou Y, Wang H, Ji R, Liu B, Wen J, et al. Targeted mutagenesis by CRISPR/Cas9 system in the model legume Medicago truncatula. Plant Cell Rep. 2017;36:371–4.
243. Sugano SS, Shirakawa M, Takagi J, Matsuda Y, Shimada T, Hara-Nishimura I, et al. CRISPR/Cas9-mediated targeted mutagenesis in the liverwort Marchantia polymorpha L. Plant Cell Physiol. 2014;55:475–81.
244. Yin K, Han T, Liu G, Chen T, Wang Y, Yu AYL, et al. A geminivirus-based guide RNA delivery system for CRISPR/Cas9 mediated plant genome editing. Sci. Rep. 2015;5:14926.
245. Ali Z, Abul-Faraj A, Li L, Ghosh N, Piatek M, Mahjoub A, et al. Efficient virus-mediated genome editing in plants using the CRISPR/Cas9 system. Mol. Plant. 2015;8:1288–91.
246. Ali Z, Abulfaraj A, Idris A, Ali S, Tashkandi M, Mahfouz MM. CRISPR/Cas9-mediated viral interference in plants. Genome Biol. Genome Biology; 2015;16:238.
247. Ludman M, Burgyán J, Fátyol K. Crispr/Cas9 mediated inactivation of Argonaute 2 reveals its differential involvement in antiviral responses. Sci. Rep. 2017;7:1010.
248. Nekrasov V, Stakawicz B, Weigel D, Gones JDG, Kamoun S. Targeted mutagenesis in the model plant nicotiana benthamiana using Cas9 RNA-guided endonuclease. Nat. Biotechnol. 2013;31:688–91.
249. Vazquez-Vilar M, Bernabé-Orts JM, Fernandez-Del-Carmen A, Ziarsolo P, Blanca J, Granell A, et al. A modular toolbox for gRNA-Cas9 genome engineering in plants based on the GoldenBraid standard. Plant Methods. BioMed Central; 2016;12:10.
250. Endo A, Masafumi M, Kaya H, Toki S. Efficient targeted mutagenesis of rice and tobacco genomes using Cpf1 from Francisella novicida. Sci. Rep. 2016;6:38169.
251. Xie X, Qin G, Si P, Luo Z, Gao J, Chen X, et al. Analysis of Nicotiana tabacum PIN genes identifies NtPIN4 as a key regulator of axillary bud growth. Physiol. Plant. 2017;160:222–39.
252. Hanania U, Ariel T, Tekoah Y, Fux L, Sheva M, Gubbay Y, et al. Establishment of a tobacco BY2 cell line devoid of plant-specific xylose and fucose as a platform for the production of biotherapeutic proteins. Plant Biotechnol. J. 2017;1–10.
253. Mercx S, Smargiasso N, Chaumont F, De Pauw E, Boutry M, Navarre C. Inactivation of the β(1,2)-xylosyltransferase and the α(1,3)-fucosyltransferase genes in Nicotiana tabacum BY-2 Cells by a Multiplex CRISPR/Cas9 Strategy Results in Glycoproteins without Plant-Specific Glycans. Front. Plant Sci. 2017;8:1–11.
254. Gao J, Wang G, Ma S, Xie X, Wu X, Zhang X, et al. CRISPR/Cas9-mediated targeted mutagenesis in Nicotiana tabacum. Plant Mol. Biol. 2015;87:99–110.
255. Endo M, Mikami M, Toki S. Multigene knockout utilizing off-target mutations of the CRISPR/cas9 system in rice. Plant Cell Physiol. 2015;56:41–7.
256. Hu X, Wang C, Fu Y, Liu Q, Jiao X, Wang K. Expanding the range of CRISPR/Cas9 genome editing in rice. Mol. Plant. 2016;1–3.
257. Ikeda T, Tanaka W, Mikami M, Endo M, Hirano H-Y. Generation of artificial drooping leaf mutants by CRISPR-Cas9 technology in rice. Genes Genet. Syst. 2015;90:231–5.
258. Mikami M, Toki S, Endo M. Comparison of CRISPR / Cas9 expression constructs for efficient targeted mutagenesis in rice. Plant Mol. Biol. 2015;88:561–72.
259. Minkenberg B, Xie K, Yang Y. Discovery of rice essential genes by characterizing a CRISPR-edited mutation of closely related rice MAP kinase genes. Plant J. 2017;89:636–48.

260. Liang G, Zhang H, Lou D, Yu D. Selection of highly efficient sgRNAs for CRISPR/Cas9-based plant genome editing. Sci. Rep. 2016;6:21451.

261. Liu Y, Xu Y, Ling S, Liu S, Yao J. Anther-preferential expressing gene PMR is essential for the mitosis of pollen development in rice. Plant Cell Rep. 2017;36:919–31.

262. Lu Y, Zhu JK. Precise Editing of a Target Base in the Rice Genome Using a Modified CRISPR/Cas9 System. Mol. Plant. 2017;10:523–5.

263. Ma L, Zhang D, Miao Q, Yang J, Xuan Y, Hu Y. Essential role of sugar transporter OsSWEET11 during the early stage of rice grain filling. Plant Cell Physiol. 2017;58:863–73.

264. Sun Y, Jiao G, Liu Z, Zhang X, Li J, Guo X, et al. Generation of High-Amylose Rice through CRISPR/Cas9-Mediated Targeted Mutagenesis of Starch Branching Enzymes. Front. Plant Sci. 2017;8:1–15.

265. Wang M, Liu Y, Zhang C, Liu J, Liu X, Wang L, et al. Gene editing by co-transformation of TALEN and chimeric RNA/DNA oligonucleotides on the rice OsEPSPS gene and the inheritance of mutations. PLoS One. 2015;10:1–16.

266. Wang M, Mao Y, Lu Y, Tao X, Zhu J kang. Multiplex Gene Editing in Rice Using the CRISPR-Cpf1 System. Mol. Plant. 2017;1011–3.

267. Xu R, Li H, Qin R, Wang L, Li L, Wei P, et al. Gene targeting using the Agrobacterium tumefaciens -mediated CRISPR-Cas system in rice. Rice J. 2014;7:2–4.

268. Xu R, Qin R, Li H, Li D, Li L, Wei P, et al. Generation of targeted mutant rice using a CRISPR-Cpf1 system. Plant Biotechnol. J. 2017;15:713–7.

269. Yin X, Biswal AK, Dionora J, Perdigon KM, Balahadia CP, Mazumdar S, et al. CRISPR-Cas9 and CRISPR-Cpf1 mediated targeting of a stomatal developmental gene EPFL9 in rice. Plant Cell Rep. 2017;36:745–57.

270. Yuan J, Chen S, Jiao W, Wang L, Wang L, Ye W, et al. Both maternally and paternally imprinted genes regulate seed development in rice. New Phytol. 2017;

271. Zhou H, He M, Li J, Chen L, Huang Z, Zheng S, et al. Development of Commercial Thermo-sensitive Genic Male Sterile Rice Accelerates Hybrid Rice Breeding Using the CRISPR/Cas9-mediated TMS5 Editing System. Sci. Rep. 2016;6:37395.

272. Wang M, Lu Y, Botella JR, Mao Y, Hua K, Zhu J kang. Gene Targeting by Homology-Directed Repair in Rice Using a Geminivirus-Based CRISPR/Cas9 System. Mol. Plant. 2017;1007–10.

273. Sun Y, Zhang X, Wu C, He Y, Ma Y, Hou H, et al. Engineering herbicide resistant rice plants through CRISPR/Cas9-mediated homologous recombination of the acetolactate synthase. Mol. Plant. 2016;1–4.

274. Xie K, Yang Y. RNA-Guided genome editing in plants using a CRISPR-Cas system. Mol. Plant. © 2013 The Authors. All rights reserved.; 2013;6:1975–83.

275. Li J, Meng X, Zong Y, Chen K, Zhang H, Liu J, et al. Gene replacements and insertions in rice by intron targeting using CRISPR-Cas9. Nat. plants. 2016;2:16139.

276. Subburaj S, Chung SJ, Lee C, Ryu SM, Kim DH, Kim JS, et al. Site-directed mutagenesis in Petunia ?? hybrida protoplast system using direct delivery of purified recombinant Cas9 ribonucleoproteins. Plant Cell Rep. 2016;1–10.

277. Lopez-Obando M, Hoffmann B, Géry C, Guyon-Debast A, Téoulé E, Rameau C, et al. Simple and Efficient Targeting of Multiple Genes Through CRISPR-Cas9 in Physcomitrella patens. G3 (Bethesda). 2016;6:1–27.

278. Nomura T, Sakurai T, Osakabe Y, Osakabe K, Sakakibara H. Efficient and heritable targeted mutagenesis in mosses using the CRISPR/Cas9 system. Plant Cell Physiol. 2016;57:2600–10.

279. Wan S, Li C, Ma X, Luo K. PtrMYB57 contributes to the negative regulation of anthocyanin and proanthocyanidin biosynthesis in poplar. Plant Cell Rep. 2017;36:1–14.

280. Yang L, Zhao X, Ran L, Li C, Fan D, Luo K. PtoMYB156 is involved in negative regulation of phenylpropanoid metabolism and secondary cell wall biosynthesis during wood formation in poplar. Sci. Rep. 2017;7:41209.

281. Li B, Cui G, Shen G, Zhan Z, Huang L, Chen J, et al. Targeted mutagenesis in the medicinal plant Salvia miltiorrhiza. Sci. Rep. 2017;7:43320.

282. Nekrasov V, Wang C, Win J, Lanz C, Weigel D, Kamoun S. Rapid generation of a transgene-free powdery mildew resistant tomato by genome deletion. Sci. Rep. Springer 2017;7:482.

283. Brooks C, Nekrasov V, Lippman ZB, Eck J Van. Efficient hene editing in tomato in the first generation using the Clustered Regularly Interspaced Short Palindromic Repeats / CRISPR-Associated9 System Plant Physiol. 2014;166:1292–7.

284. Ito Y, Nishizawa-Yokoi A, Endo M, Mikami M, Toki S. CRISPR/Cas9-mediated mutagenesis of the RIN locus that regulates tomato fruit ripening. Biochem. Biophys. Res. Commun. ; 2015;467:76–82.

285. Shimatani Z, Kashojiya S, Takayama M, Terada R, Arazoe T, Ishii H, et al. Targeted base editing in rice and tomato using a CRISPR-Cas9 cytidine deaminase fusion. Nat. Biotechnol. 2017;35:441–3.

286. Ueta R, Abe C, Watanabe T, Sugano SS, Ishihara R, Ezura H, et al. Rapid breeding of parthenocarpic tomato plants using CRISPR/Cas9. Sci. Rep. 2017;7:507.

287. Ron M, Kajala K, Pauluzzi G, Wang D, Reynoso M a., Zumstein K, et al. Hairy root transformation using Agrobacterium rhizogenes as a tool for exploring cell type-specific gene expression and function using tomato as a model. Plant Physiol. 2014;166:455–69.

288. Wang S, Zhang S, Wang W, Xiong X. Efficient targeted mutagenesis in potato by the CRISPR / Cas9 system. Plant Cell Rep. 2015;34:1473–6.

289. Andersson M, Turesson H, Nicolia A, F??lt AS, Samuelsson M, Hofvander P. Efficient targeted multiallelic mutagenesis in tetraploid potato (Solanum tuberosum) by transient CRISPR-Cas9 expression in protoplasts. Plant Cell Rep. 2017;36:117–28.

290. Zhang Y, Bai Y, Wu G, Zou S, Chen Y, Gao C, et al. Simultaneous modification of three homoeologs of TaEDR1 by genome editing enhances powdery mildew resistance in wheat. Plant J. 2017;1–11.

291. Gil-Humanes J, Wang Y, Liang Z, Shan Q, Ozuna C V., Sánchez-León S, et al. High-efficiency gene targeting in hexaploid wheat using DNA replicons and CRISPR/Cas9. Plant J. 2017;89:1251–62.

292. Zhang Y, Liang Z, Zong Y, Wang Y, Liu J, Chen K, et al. Efficient and transgene-free genome editing in wheat through transient expression of CRISPR/Cas9 DNA or RNA. Nat. Commun. 2016;7:12617.

293. Upadhyay SK, Kumar J, Alok A, Tuli R. RNA guided genome editing for target gene mutations in wheat. G3. 2013;3:2233–8.

294. Malnoy M, Viola R, Jung M-H, Koo O-J, Kim S, Kim J-S, et al. DNA-Free Genetically Edited Grapevine and Apple Protoplast Using CRISPR/Cas9 Ribonucleoproteins. Front. Plant Sci. 2016;7:1904.

295. Svitashev S, Young JK, Schwartz C, Gao H, Falco SC, Cigan a M. Targeted Mutagenesis, Precise Gene Editing, and Site-Specific Gene Insertion in Maize Using Cas9 and Guide RNA. Plant Physiol. 2015;169:931–45.

296. Zhu J, Song N, Sun S, Yang W, Zhao H, Song W, et al. Efficiency and inheritance of targeted mutagenesis in maize using CRISPR-Cas9. J. Genet. Genomics. 2016;43:25–36.

297. Feng C, Yuan J, Wang R, Liu Y, Birchler J a., Han F. Efficient Targeted Genome Modification in Maize Using CRISPR/Cas9 System. J. Genet. Genomics. 2016;43:37–43.

298. Svitashev S, Schwartz C, Lenderts B, Young JK, Mark Cigan A. Genome editing in maize directed by CRISPR-Cas9 ribonucleoprotein complexes. Nat. Commun. 2016;7:13274.

Chapter 2

CRISPR/Cas mediated curing of RP4 and mini Ti plasmids in *Agrobacterium*

Daan J. Schmitz, Amke den Dulk-Ras, Sylvia de Pater, Paul J.J. Hooykaas

Abstract

The genome editing toolkit has been expanded with the RNA guided Cas9 endonuclease from the type II CRISPR/Cas system from *Streptococcus pyogenus*. Compared to its rapid adaptation in eukaryotes as a genome editing tool, applications of CRISPR/Cas systems in bacteria have progressed relatively slowly. Nonetheless CRISPR/Cas systems have been used to edit bacterial genomes and to control gene expression in a range of prokaryotes. Here we developed an engineered CRISPR/Cas system for use in the plant pathogen *Agrobacterium tumefaciens*. We show that this system can be used to cure *Agrobacterium* from the promiscuous plasmid RP4 and from vectors with the replication unit of the octopine Ti plasmid. Curing of complete Ti plasmids was not successful, possibly because of the presence of a toxin anti-toxin system.

Introduction

CRISPR (clustered regularly interspaced short palindromic repeats) and Cas (CRISPR associated) nucleases are found in many different prokaryotes, where they function as adaptive immune systems that act against bacteriophages and other invading nucleic acids [1]. The prototypical CRISPR/Cas endonuclease from the type II bacterial CRISPR/Cas system found in *Streptococcus pyogenus* can be programmed to create targeted double strand breaks (DSBs) by an engineered single guide RNA (sgRNA) [2]. The Cas9 endonuclease is directed to a desired sequence by specifying a 20 basepair sequence of the sgRNA which directs the Cas9 nuclease to a 20 nucleotide complementary sequence [2,3]. Selection of target sequences is limited by the requirement of a so called protospacer-associated motif (PAM) which is a requisite for Cas9 activity [4,5].

DSBs introduced in the genome with the CRISPR/Cas system can either be repaired via non-homologous end joining (NHEJ) or by homologous recombination (HR) and thus the system can be exploited for targeted mutagenesis or gene replacement. Because most bacteria lack an efficient NHEJ repair mechanism, DSBs are lethal if repair cannot occur by HR. Repair via HR can however be used to replace or correct existing genes by providing an artificial repair template with homology to the target sequence.

DSBs induced by an engineered CRISPR/Cas system from *Streptococcus* have been used for genome editing via HR in *Streptococcus* [6], *E. coli* [6–10], multiple *Streptomyces* species [11–13], *Lactobacillus reuteri* [14] and two different *Clostridium* species [15–17]. Furthermore the CRISPR/Cas system has also been used to selectively repress transcription of target genes using a mutated Cas9 protein without nuclease activity [18–21].

In this chapter we describe development of an engineered CRISPR/Cas system in the gram-negative soil dwelling bacterium *Agrobacterium*. We show that Cas9 is active in *Agrobacterium* and can be used for the curing of the promiscuous plasmid RP4 and small derivatives of the octopine Ti plasmid. Furthermore we show that our version of the Cas9 protein with a C-terminal translocation tag for translocation by the *Agrobacterium* VirB/D4 T4SS is functional and therefore can be used in the following chapters as an effective nuclease.

Results

Design of a Cas9/sgRNA expression system for Agrobacterium

In order to induce DSBs at specific sites in the *Agrobacterium* genome we expressed Cas9 and sgRNAs under the control of the acetosyringone inducible *virF* promoter that has been used previously for the expression of various heterologous proteins [22,23]. The Cas9 protein used in our experiments was modified to contain the C-terminal transport signal that is recognized by the *Agrobacterium* VirB/D4 T4SS secretion channel, which is required for the experiments in Chapters 3 and 4.

Initially a vector was designed for expressing a sgRNA directed against the *sacB* marker (pNCas9FsacB) and introduced into an *Agrobacterium* strain with an pTiAch5 into which a small plasmid conferring kanamycin resistance and containing the *sacB* gene had integrated by HR via a single crossover. This plasmid can be lost again by the reverse HR reaction. This will occur spontaneously at a low frequency, but these deletion events can be selected on sucrose plates. A DSB within the integrated small plasmid can efficiently be repaired by this reverse HR reaction. In order to test CRISPR/Cas activity in *Agrobacterium* we therefore directed the system for induction of DSBs in the *sacB* gene. Efficient DSB induction would therefore lead to loss of kanamycin resistance in the survivors.

Single colonies obtained after electroporation with the Cas9 and sgRNA expression vector were re-streaked on medium with kanamycin and all were sensitive (Fig. 1a) indicating that recombination between the flanking repeats had been induced by the DSB and that the integrated small plasmid had been lost from pTiAch5. To ensure that loss of the plasmid was not the result of spontaneous recombination events that had occurred without DSB induction by the engineered CRISPR/Cas system, a control experiment was performed with a vector expressing the modified NCas9F protein in the absence of the sgRNA. In absence of the sgRNA all of the colonies obtained after electroporation had remained resistant to kanamycin (Fig. 1b).

These results combined indicate that the *Agrobacterium* CRISPR/Cas expression cassette is functional and is capable of efficiently inducing DSBs that promote the recombination of direct repeats flanking the DSB.

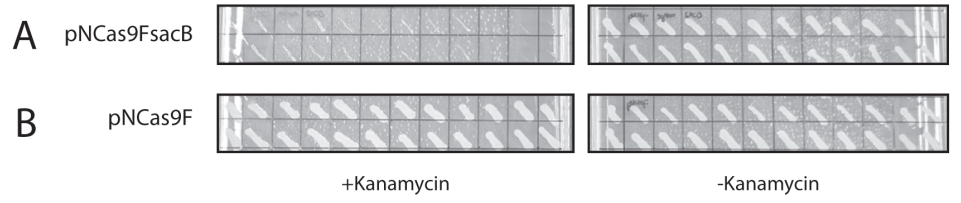

Figure 1. Screening for homologous recombination of repeats surrounding the *SacB* and Km marker present on the modified pTiAch5 identified by the loss of Km marker. LBA4001/pSDM3684 transformed with pNCas9FsacB (**A**) or pNCas9F (**B**) were grown in the absence (control) or presence of kanamycin.

Curing of RP4

As the previous experiments showed that the CRISPR/Cas system was active in *Agrobacterium,* we tested if it could be used to cure plasmids by inducing DSBs. As a candidate for curing we chose the promiscuous plasmid RP4 which is maintained extremely stable under non-selective conditions and has a copy number of about four to seven [24]. The CRISPR/Cas system described in the previous experiments was modified to target a region upstream of the replication initiation gene *trfA* (pNCas9FtrfA) and upstream of the replication origin region (*oriV*) (pNCas9ForiV) both with a protospacer containing a GG motif at the 3' end that enhances DSB break induction [25]. Cells were plated on medium with gentamicin selecting for the presence of pNCas9FtrfA and pNCas9ForiV after electroporation. After *vir* induction these colonies were checked for sensitivity to kanamycin, tetracyclin and carbenicillin as curing of RP4 would result in loss of the *nptII*, *tetA* and *bla* genes found on RP4. Induced *Agrobacterium* cells expressing only the NCas9F protein were all still resistant to kanamycin (Fig. 2b), thereby excluding that the curing of RP4 could occur due to the restrictive induction conditions. However, after expression of the NCas9F protein combined with an sgRNA targeting the region upstream of *trfA* 100% (36/36) of the colonies had become sensitive to kanamycin (Fig. 2a), tetracyclin and carbenicillin. After expression of the NCas9F protein combined with an sgRNA targeting upstream of the oriV about ~69% (25/36) of the tested colonies had become sensitive to kanamycin, tetracyclin and carbenicillin. PCR on the *trfA* locus confirmed the loss of RP4 in these colonies (Fig. 2c). Similar curing efficiencies were obtained after expression of the NCas9F protein and the sgRNA targeting the region upstream of *trfA* under control of the coliphage T5 promoter.

These results showed that DSB induction in the RP4 plasmid is an effective method for the curing of RP4 from *Agrobacterium*.

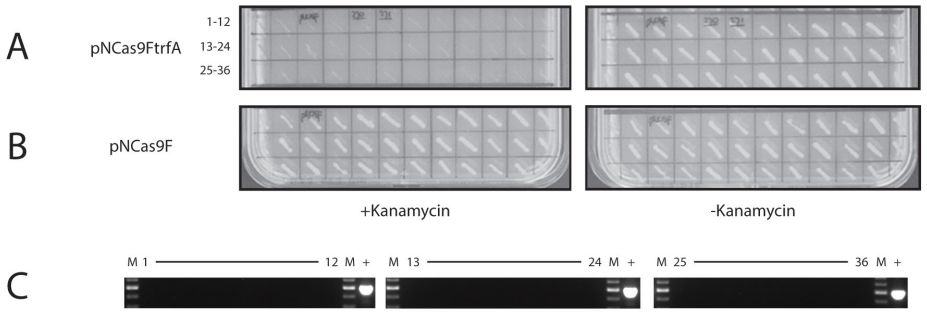

Figure 2. Screening of *Agrobacterium* for the loss of plasmid RP4. Loss is detected by kanamycin sensitivity. LBA1100(RP4) transformed with pNCas9FtrfA (**A**) or pNCas9F (**B**) were grown in the absense (control) or presence of kanamycin. PCR on the *trfA* locus using kanamycin sensitive colonies (**C**). M: DNA marker, +: positive control

Stability of the octopine Ti plasmid

As the previous experiments showed that *Agrobacterium* can be cured of RP4 we performed an experiment to test if *Agrobacterium* can be similarly cured of its octopine Ti plasmid. Because *vir* induction of *Agrobacterium* cells also increases the copy number of the Ti plasmid and its stability [26] an expression system was used in which Cas9 and the sgRNA were brought under the control of the T5 coliphage promoter [27]. With this system curing of RP4 was effective (see previous paragraph). Two new plasmids were created with the T5 expression system targeting the *virD4* (pT5NCasFvirD4) and *agaA* (pT5NCasFagaG) locus on the Ti plasmid and introduced into the *Agrobacterium* strains LBA1010, LBA1100 and LBA288. LBA1010 contains the complete Ti plasmid pTiB6 [28], LBA1100 contains the T-region-deleted helper plasmid [29] and LBA288 does not contain a Ti plasmid [30]. The pT5NCasFvirD4 or pT5NCasFagaG plasmids were introduced by electroporation. Cells were plated on medium with gentamicin selecting for the presence of the plasmids. Selected colonies were analyzed for the presence of the Ti plasmid by PCR. This analysis showed that the Ti plasmid was still present in all the LBA1010 and LBA1100 transformants. Sequence analysis of the segment encompassing the DSB site in the regions of the *virD4* and *agaA* gene, respectively, did not reveal any mutations introduced as the result of error prone repair via a NHEJ mechanism.

The transformation of both pT5NCasFvirD4 and pT5NCasFagaA to LBA1010 resulted in low numbers of transformants. To determine if this was a direct result of DSB induction we introduced a control plasmid that targets the *trfA* locus of RP4 (pT5NCas9FtrfA) and thus introduces no DSBs in the Ti plasmid of LBA1010 and LBA1100. The high numbers of transformants obtained after electroporation with this control plasmid were used to estimate the relative electroporation frequencies of the plasmids targeting the *virD4* and *agaA* locus. These results suggests that only 3% (pT5NCas9FagaA) up to 20% (pT5NCasFvirD4) of the transformants receiving a CRISPR/Cas plasmid inducing a DSB survive.

The construct targeting the *agaA* locus did not cause a strong decrease in the transformation frequency in LBA288 (no Ti plasmid) and LBA1100 (disabled Ti plasmid without the *agaA* locus) (Fig. 3). After transformation of pT5NCas9FvirD4, there was also a decrease in the relative transformation efficiency seen in LBA288 (Ti plasmid less), but presence of *virD4* in LBA1010 and LBA1100 led to a further strong decrease in relative transformation efficiency. These results combined indicate that DSBs induction in the Ti plasmid negatively affects the number of transformants.

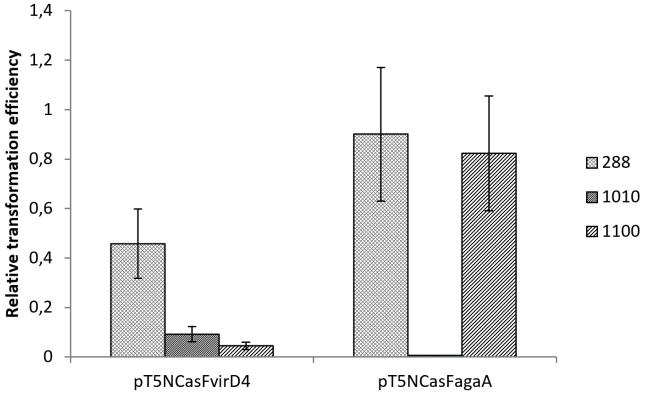

Figure 3. Transformation efficiencies of different *Agrobacterium* strains electroporated with pT5NCasFvirD4 and pT5NcasFagaA, targeting the *virD4* and *agaA* locus respectively. Relative transformation efficiencies are determined by dividing colony count of pT5NCas9FvirD4 and pT5NCas9FagaA, respectively, by the colony count of transformation with pT5NCas9trfa. Error bars indicate the SEM (N=3).

To test if the presence of the Ti plasmid is indeed involved in the reduction of the transformation efficiency a crossing was set up between LBA657, which can transfer its octopine Ti plasmid without requiring induction, and LBA288 derivatives expressing either pT5NCasFvirD4 or pT5NCasFtrfA. The transconjugant frequency decreased ~65 fold with the strain containing pT5NCasFvirD4 in the recipient compared to the transconjugant frequency with pT5NCasFtrfA (Fig. 4). This shows that presence of pT5NCas9virD4 in the recipient can prevent establishment of the Ti plasmid and indicates that a DSB is efficiently formed at the *virD4* locus in the Ti plasmid.

These results combined with the strong decrease in number of transformants seen after electroporation indicate that DSB induction in an established octopine Ti plasmid and its subsequent degradation is lethal to *Agrobacterium*.

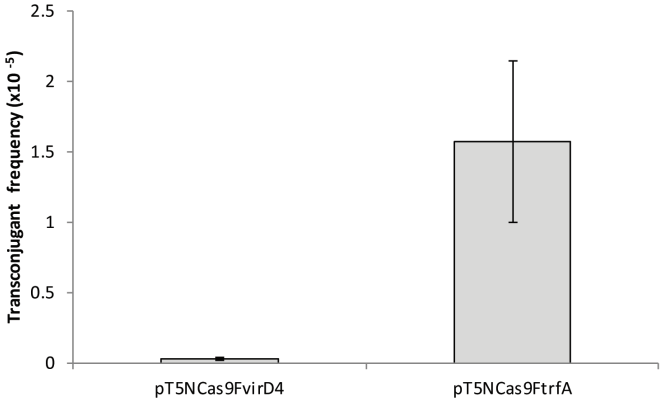

Figure 4. Transconjugant frequency after mating between LBA657 and LBA288 pT5NCas9FvirD4 or LBA288 pT5NCas9F-trfA, respectively. Error bars indicate the SEM (N=3).

Curing of mini Ti plasmids

As curing of the octopine Ti plasmid was not successful and DSB formation in the Ti plasmid led to lethality instead, this suggested the presence of one or more toxin anti-toxin systems on the Ti plasmid preventing its loss from the host. In order to exclude that the Ti *repABC* replicator had special properties preventing its loss, we tested whether small mini Ti plasmids that lack such putative toxin anti-toxin system could be cured by DBS induction in these mini Ti plasmids. Three different *Agrobacterium* strains were used, each carrying a different octopine Ti plasmid derivative all of which have a Tn1 insertion and therefore provide carbenicillin resistance (LBA2803, LBA2811 and LBA2821) [30]. A CRISPR/Cas construct targeting the *bla* gene encoding a β-lactamase present in the Tn1 insertion on these plasmids was tested for its ability to cure the mini octopine Ti plasmid derivatives (pNCas9TN1). The introduction of a vector only encoding the NCas9F protein into LBA2803, LBA2811 and LBA2821 did not result in colonies that had become sensitive to carbenicillin. DSB induction on the small Ti plasmid derivatives with the pNCas9FTN1 vector did however result in high curing efficiencies of 77% (LBA2803) (Fig. 5a), 97% (LBA2811) (Fig. 5b) and 100% (LBA2821) (Fig. 5c) based on sensitivity to carbenicillin. PCR on the *repC* locus confirmed the loss of the mini Ti plasmids in all the carbenicillin sensitive colonies in LBA2811 and LBA2821 expressing Cas9 and the sgRNA. PCR on the LBA2803 colonies expressing NCas9F and the sgRNA that had remained resistant to carbenicillin also showed no PCR product (*repC* locus), indicating that the mini Ti plasmid had been lost also in these colonies and that carbenicillin resistance

may have been maintained by spontaneous mutations or by transposition of Tn1 to another site in the genome. As a control we also introduced pNCas9TN1 in LBA657 (containing a Ti plasmid with an Tn1 insertion) which resulted in a low transformation efficiency but no curing.

These results show that DSB induction in mini octopine Ti plasmids is an effective tool for curing of these mini Ti plasmids. These results also provide a further indication of the presence of toxin anti-toxin systems elsewhere in the Ti plasmid that prevents their curing.

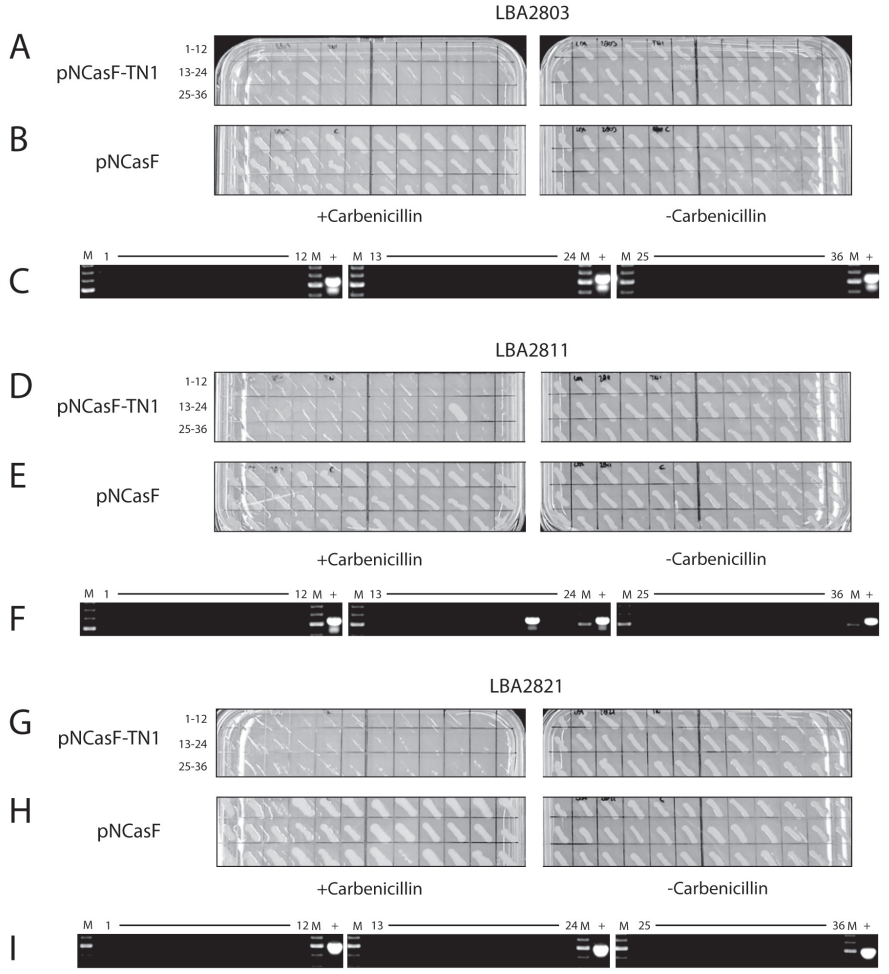

Figure 5. Screen for the loss of mini Ti plasmids. *Agrobacterium* strains expressing both the NCas9F protein and sgRNA targeting the *bla* gene encoding β-lactamase (pNCasF-TN1) were compared to *Agrobacterium* strains only expressing the NCas9F protein. Loss of the mini Ti plasmid results in loss of carbenicillin resistance. LBA2803 (**A**-**C**), LBA2811 (**D**-**F**) and LBA2821 (**G**-**I**) each harboring a different mini Ti plasmid were transformed with pNCasF-TN1 (**A**, **D**, **G**) or pNCasF (**B**, **E**, **H**) and grown in presence or absense (control) of carbenicillin. PCR on the *repC* locus of the mini-Ti plasmids using carbenicillin sensitive colonies of LBA2803, LBA2811 and LBA2821 (**C**, **F**, **I**). M: DNA marker, +: positive control

Discussion

In this study we have shown that the CRISPR/Cas9 based endonuclease system can be expressed in *Agrobacterium* and can be used to induce DBSs. Such DSBs could be used to increase the recombination between flanking direct repeats to excise a previously integrated non-replicating plasmid and can be used therefore to increase the efficiency of obtaining gene deletion events which until now required the screening of vast numbers of colonies. Furthermore the CRISPR/Cas9 system could be used to cure *Agrobacterium* cells of the promiscuous IncPα RP4 plasmid and mini Ti plasmids.

The Cas9 protein used in our experiments has been fused with a hydrophilic C-terminal secretion signal with a net positive charge [31] used previously to translocate several heterologous proteins through the type IV secretion system of *Agrobacterium* [23,31,32]. We show that the addition of this secretion signal does not negate the ability of Cas9 to induce DBSs.

The loss of the integrated non-replicating plasmid through DSB induction occurred at low basal expression levels of Cas9 and the sgRNA without requiring induction of the virF promoter. Curing of RP4 did however require higher expression levels of Cas9 and the sgRNA either through induction of the *virF* promoter or by constitutive expression under control of the T5 promoter. This suggests that with low expression levels of Cas9 and the sgRNA DSB induction does not occur on all copies of RP4 simultaneously and therefore not all copies are lost and cut RP4 molecules can be repaired via HR using still intact copies of RP4 as a template.

DSB induction in the octopine Ti plasmid negatively correlated with the survival of *Agrobacterium* cells harboring the Ti plasmid. Also a very low transconjugant frequency was observed when the Ti plasmid was transferred via mating to a recipient *Agrobacterium* expressing a CRISPR/Cas targeting the Ti plasmid. This shows that Ti plasmids have difficulty in establishing in a host expressing the CRISPR/Cas system which in line with its original biological function in defense. However, removal of an already established Ti plasmid by introduction of CRISPR/Cas turned out to be impossible. This suggests that one or more toxin anti-toxin systems are encoded by the Ti plasmid in which a long-lived toxin is neutralized by a short lived anti-toxin. Evidence for the presence of such a system has already been found for stability of the nopaline pTiC58 plasmid [33]. Attempts to identify the toxin anti-toxin system present on the octopine Ti plasmid using bioinformatics tools for the prediction of toxin anti-toxin systems was however not successful. This could be because the bioinformatics tools used can only identify type II toxin anti-toxin.

In summary, we have created a CRISPR/Cas system that is active in *Agrobacterium* and can be used to promote homologous recombination on DNA repeats and to cure *Agrobacterium* of RP4 and mini octopine Ti plasmids.

Material & Methods

Agrobacterium strains and media

Agrobacterium strains used in this study are listed in Table 1. All *Agrobacterium* strain were grown in LC (LB medium with 5g/l NaCl) at 290C with the appropriate antibiotics at the following concentrations: gentamicin 40 μg/ml; carbenicillin 75 μg/ml; kanamycin 100 μg/ml; tetracyclin 2 μg/ml. Plasmids were electroporated into *Agrobacterium* as described in den Dulk-Ras and Hooykaas [34].

Plasmid construction

The HindIII/NotI fragment with the *virF* promoter and the gene encoding a translation fusion between a nuclear localization signal, the Cre-recombinase and the last 37 amino acids of the *Agrobacterium* virulence protein VirF (pvirFpromoter-NLS::Cre::VirF37c) [32] from pSDM3155 was inserted into the HindIII/NotI sites of pUC18 to create pSDM2131. The Cre-recombinase fragment was removed by digestion of pSDM2131 with SalI and BglII and was replaced by a small linker containing a BglII site (annealed oligos DS061 and DS062). The Cas9 BamHI fragment from plasmid pMJ920 (Addgene plasmid #42234) [35] was inserted into the BglII site of pvirF-NLS::BglII Linker::VirF37c creating a translational fusion between a nuclear localization signal, Cas9 and the C-terminal 37 amino acids of VirF (NLS::Cas9::VirF37c) under the control of the *virF* promoter. The HindIII/NotI fragment with pvirF-NLS::Cas9::VirF37c was inserted into the HindII/NotI sites of pBBR6 (see Table 1) creating plasmid pNCas9F.

For the construction of the sgRNA construct under the control of the *virF* promoter a segment of DNA was synthesized at BaseClear and was provided in the pUC57Kan backbone (pDualsgRNAcassette). This synthesized DNA fragment contains two sgRNA expression cassettes in direct repeat with a BbsI and BsaI restriction site respectively in which annealed oligo's with the protospacer sequence were cloned into the BbsI site. Table 3 shows an overview of the oligo's used to create the different protospacers. The entire fragment was excised with NotI and cloned into the NotI site of pNCasF.

pT5NCasF was created by amplifying the NLS::Cas9::VirF37c fragment from pNCas9F with DS356 and DS359. This fragment was cloned into the HindIII and NotI sites of pBBR6. Into this construct the T5 promoter amplified with DS360 and DS361 was cloned into BamHI and NheI site 5' of the NLS::Cas9::VirF37c resulting in pT5NCas9F. To create the sgRNA expression cassette under control of the T5 promoter and the *virF* 3' terminator the T5 promoter fragment was amplified from pQE-30 with DS349 and DS350 and the structural part of the sgRNA together with the *virF* 3' terminator was amplified from pDualsgRNAcassette with DS351 and DS352. These PCR products were combined and used as template with DS349 and DS352 and the resulting PCR product was cloned into NotI digested pDualsgRNAcassette backbone (pT5sgRNA). To insert the protospacer sequence, annealed and phosphorylated oligo's were inserted into BbsI digested pT5sgRNA. After insertion of the protospacer sequence the T5sgRNA cassettes were cloned into the NotI site of pT5NCas9F.

The pTi Ach5 plasmid with cointegrated pSDM3684 conferring kanamycin resistance and containing the *sacB* gene was created by electroporation of pSDM3684 into wildtype *Agrobacterium* strain Ach5 (LBA4001). Single crossover events of pSDM3684 with pTiAch5 were detected by selection on kanamycin. Plasmid pSDM3684 is a derivative of pIN61 containing fragments homologous to the pTiAch5. As pSDM3684 cannot replicate in *Agrobacterium* maintenance is only possible after cointegration with the Ti plasmid by HR.

Recombination and curing experiments

For the recombination experiments the pNCas9FsacB and pNCas9F plasmids were introduced via electroporation into *Agrobacterium* strain Ach5 with the modified cointegrated Ti-Ach5:: pSDM3684 (described above). Cells were plated on medium with gentamycin selecting for the presence of the plasmids. Single colonies were re-streaked on medium with and without kanamycin (100 μg/ml) to detect loss of the integrated pSDM3684 plasmid. No induction

was required as un-induced levels of expression were sufficient for DSB induction and recombination of direct repeats flanking the DSB.

For the RP4 curing experiments pNCas9FtrfA and pNCas9F were electroporated into LBA1100 (RP4). Colonies were grown overnight at 29°C in LC with appropriate antibiotics. Induction of the *virF* promoter was performed at 28°C at OD600 = 0.25 for 16 hours in induction medium containing 200μM acetosyringone (Sigma-Aldrich Co.) . Dilution series were plated to obtain single colonies which were re-streaked on plates containing tetracycline, carbenicillin or kanamycin and on plates without antibiotics. Loss of RP4 was confirmed by PCR with primers DS347 and DS348 (*trfA* locus).

For Ti plasmid and mini Ti plasmid curing experiments, the pT5NCas9virD4, pT5NCas9FagaA, pT5NCas9FtrfA and pT5NCas9TN1 plasmids were electoporated into LBA288, LBA1010 and LBA1100. Loss of the Ti plasmid was checked via PCR using primers DS317 and DS318 (*repC* locus). Loss of the mini Ti plasmids from LBA2803, LBA2811 and LBA2821 were checked by re-streaking single colonies on medium with and without carbenicillin (75 μg/ml) and was further confirmed by PCR using primers DS317 and DS318 (*repC* locus). The relative transformation efficiencies depicted in Figure 3 were calculated by dividing the colony count after electroporation with pT5NCas9FvirD4 and pT5NCas9FagaA by the colony count after electroporation with pT5NCas9FtrfA.

Table 1. Plasmids used in this study

Plasmid	Marker	Properties	Source
pIN61	Km	sacB vector	Vergunst (unpublished)
pDSM2131	Cb	Cloning vector with NLS-Cre-VirF42n fragment	Vergunst (unpublished)
pDSM3155	Gm	Cre::virFΔ42N	[35]
pSDM3684	Km	Vector containing octopine T-DNA flanks, sacB gene and Km marker	Den Dulk-Ras (unpublished)
pNCas9F	Gm	vir inducible NCas9F expression vector	This study
pQE-30	Cb	Cloning vector for expression under control of the T5 promoter	Qiagen
pDualsgRNAcassette	Km	sgNRA cloning vector, virF promoter	This study
pNCas9FsacB	Gm	vir inducible NCas9F and sgRNA targeting sacB	This study
pT5NCas9	Gm	NCas9F expression vector under control of the T5 promoter	This study
pNCas9FtrfA	Gm	NCas9F and sgRNA targeting upstream of trfA under control of the virF promoter	This study
pNCas9ForiV	Gm	NCas9F and sgRNA targeting upstream of oriV under control of the virF promoter	This study
pT5NCasFT5virD4	Gm	NCas9F and sgRNA targeting virD4 under control of the T5 promoter	This study
pT5sgRNA	Gm	sgNRA cloning vector, T5 promoter	This study
pT5NCasFT5trfA	Gm	NCas9F and sgRNA targeting trfA under control of the T5 promoter	This study
pT5NCasFT5agaG	Gm	NCas9F and sgRNA targeting agaG under control of the T5 promoter	This study
pT5NCasFT5Tn1	Gm	NCas9F and sgRNA targeting Tn1 under control of the T5 promoter	This study
pBBR6	Gm	Derrivative of the broad host-range plasmid pRL662	[22]
pMJ920	Cb	Vector for mammalian expression of Cas9	[35]

Table 2. Overview of *Agrobacterium* strains used in this study

Strain	Chromosomal Background	Properties	Source
LBA288	C58	Cured of Ti plasmid	[36]
LBA657	C58	pTiB6::Tn1	[36]
LBA1010	C58	pTiB6	[28]
LBA1100	C58	pTiB6 disarmed	[29]
LBA2803	C58	mini Ti plasmid	[30]
LBA2811	C58	mini Ti plasmid	[30]
LBA2821	C58	mini Ti plasmid	[30]
LBA1100 RP4	C58	pTiB6 disarmed + RP4	Wittleben (Unpublished)
LBA4001/pSDM3684	Ach5	pTi-Ach5::pSDM3684	den Dulk-Ras, Hooykaas (Unpublished)

Table 3. Overview of primers used in this study

Primer name	Sequence
DS060 Linker	GATCTACTAGTGCTGCACGG
DS061 Linker	GATCCCGTGCAGCACTAGTA
DS256 SacB guideRNA NotI FW	ATATGCGGCCGCAGCTCCTATGATAGTCGATA
DS257 SacB guideRNA NotI RV	ATATGCGGCCGCGGGACCAGCACACTTAGATA
DS270 trfA protospacer	ATTGCTGGACACCAAGGCACCAGG
DS271 trfA protospacer	AAACCCTGGTGCCTTGGTGTCCAG
DS272 OriV protospacer	ATTGTGCCACCCGCGTCGCCGAGG
DS273 OriV protospacer	AAACCCTCGGCGACGCGGGTGGCA
DS317 repCfw	TGCCTAACAAAGGTACAAGG
DS318 repCrv	GTCAAGAACCTGATCAATCGC
DS339 virD4 target fw	ATTGACCTCTGTGCTGGGCGACGG
DS340 virD4 target rv	AAACCCGTCGCCCAGCACAGAGGT
DS345 agaA fw	ATTGCGCGGAGTCTCGGATCGAGG
DS346 agaA rv	AAACCCTCGATCCGAGACTCCGCG
DS349 FW T5 promoter	GATCGCGGCCGCCTCGAGAAATCATAAAAAATTT
DS350 RV T5 promoter	TTCTTCACGAAGACCCCAATCTGTGTGAAATTGTTATCCG
DS351 FW guideRNA + virF 3′	ATTGGGGTCTTCGTGAAGAA
DS352 RV guideRNA + virF 3′	CGATGCGGCCGCGGGACCAGCACACTTAGATA
DS356 RV Cas9F	TCGAGCGGCCGCTCATAGACCGCGCGTTGATC
DS359 NCas9F HindIII BamHI NheI	CTAGAAGCTTGGATCCCTACGGCTAGCATGGATAAAGCG GAATAATTCC
DS360 T5 FW BamHI	CTAGGGATCCCTCGAGAAATCATA
DS361 T5 RV NheI	GGTAGCTAGCAGTTAATTTCTCCT
DS376 TN1 fw1	ATTGTTACTTCTGACAACGATCGG
DS377 TN1 rev1	AAACCCGATCGTTGTCAGAAGTAA

References

1. van der Oost J, Jore MM, Westra ER, Lundgren M, Brouns SJJ. CRISPR-based adaptive and heritable immunity in prokaryotes. Trends Biochem. Sci. 2009;8:401–7.

2. Jinek M, Chylinski K, Fonfara I, Hauer M, Doudna JA, Charpentier E. A programmable dual-RNA–guided DNA endonuclease in adaptice bacterial immunity. Science. 2012;337:816–22.

3. Bhaya D, Davison M, Barrangou R. CRISPR-Cas systems in bacteria and archaea: versatile small RNAs for adaptive defense and regulation. Annu. Rev. Genet. 2011;45:273–97.

4. Deveau H, Barrangou R, Garneau JE, Labonté J, Fremaux C, Boyaval P, et al. Phage response to CRISPR-encoded resistance in Streptococcus thermophilus. J. Bacteriol. 2008;190:1390–400.

5. Mojica FJM, Díez-Villaseñor C, García-Martínez J, Almendros C. Short motif sequences determine the targets of the prokaryotic CRISPR defence system. Microbiology. 2009;155:733–40.

6. Jiang W, Bikard D, Cox D, Zhang F, Marraffini L a. RNA-guided editing of bacterial genomes using CRISPR-Cas systems. Nat. Biotechnol. 2013;31:233–9.

7. Li Y, Lin Z, Huang C, Zhang Y, Wang Z, Tang YJ, et al. Metabolic engineering of Escherichia coli using CRISPR-Cas9 meditated genome editing. Metab. Eng. 2015;31:13–21.

8. Pyne ME, Moo-Young M, Chung DA, Chou CP. Coupling the CRISPR/Cas9 system with lambda red recombineering enables simplified chromosomal gene replacement in Escherichia coli. Appl. Environ. Microbiol. 2015;81:5103–14.

9. Reisch CR, Prather KLJ. The no-SCAR (Scarless Cas9 Assisted Recombineering) system for genome editing in Escherichia coli. Sci. Rep. 2015;5:15096.

10. Ronda C, Pedersen LE, Sommer MOA, Nielsen AT. CRMAGE: CRISPR Optimized MAGE Recombineering. Sci. Rep. 2016;6:19452.

11. Cobb RE, Wang Y, Zhao H. High-Efficiency Multiplex Genome Editing of Streptomyces Species Using an Engineered CRISPR/Cas System. ACS Synth. Biol. 2014;4:723–8.

12. Huang H, Zheng G, Jiang W, Hu H, Lu Y. One-step high-efficiency CRISPR/Cas9-mediated genome editing in Streptomyces. Acta Biochim. Biophys. Sin. 2015;47:231–43.

13. Tong Y, Charusanti P, Zhang L, Weber T, Lee SY. CRISPR-Cas9 Based Engineering of Actinomycetal Genomes. ACS Synth. Biol. 2015;4:1020–9.

14. Oh JH, Van Pijkeren JP. CRISPR-Cas9-assisted recombineering in Lactobacillus reuteri. Nucleic Acids Res. 2014;42:1–11.

15. Wang Y, Zhang ZT, Seo SO, Choi K, Lu T, Jin YS, et al. Markerless chromosomal gene deletion in Clostridium beijerinckii using CRISPR/Cas9 system. J. Biotechnol. 2015;200:1–5.

16. Li Q, Chen J, Minton NP, Zhang Y, Wen Z, Liu J, et al. CRISPR-based genome editing and expression control systems in Clostridium acetobutylicum and Clostridium beijerinckii. Biotechnol. J. 2016;11:961–72.

17. Pyne ME, Bruder MR, Moo-Young M, Chung DA, Chou CP. Harnessing heterologous and endogenous CRISPR-Cas machineries for efficient markerless genome editing in Clostridium. Sci. Rep. 2016;6:25666.

18. Bikard D, Jiang W, Samai P, Hochschild A, Zhang F, Marraffini LA. Programmable repression and activation of bacterial gene expression using an engineered CRISPR-Cas system. Nucleic Acids Res. 2013;41:7429–37.

19. Qi LS, Larson MH, Gilbert LA, Doudna JA, Weissman JS, Arkin AP, et al. Repurposing CRISPR as an RNA-guided platform for sequence-specific control of gene expression. Cell. 2013;152:1173–83.

20. Bruder MR, Pyne ME, Moo-Young M, Chung DA, Chou CP. Extending CRISPR-Cas9 technology from genome editing to transcriptional engineering in the genus Clostridium. Appl. Environ. Microbiol. 2016;82:6109–19.

21. Westbrook AW, Moo-Young M, Chou CP. Development of a CRISPR-Cas9 tool kit for comprehensive engineering of bacillus subtilis. Appl. Environ. Microbiol. 2016;82:4876–95.

22. Vergunst AC, Schrammeijer B, den Dulk-Ras A, Vlaam de CMT, Regensburg-Tuïnk TJ, Hooykaas PJJ. VirB/D4-dependent protein translocation from Agrobacterium into plant cells. Science. 2000;290:979–82.

23. Rolloos M, Hooykaas PJJ, van der Zaal BJ. Enhanced targeted integration mediated by translocated I-SceI during the Agrobacterium mediated transformation of yeast. Sci. Rep. 2015;5:8345.

24. Figurski DH, Helinski DR. Replication of an origin-containing derivative of plasmid RK2 dependent on a plasmid function provided in trans. Proc. Natl. Acad. Sci. USA. 1979;76:1648–52.

25. Farboud B, Meyer BJ. Dramatic enhancement of genome editing by CRISPR/cas9 through improved guide RNA design. Genetics. 2015;199:959–71.

26. Cho H, Winans SC. VirA and VirG activate the Ti plasmid repABC operon, elevating plasmid copy number in response to wound-released chemical signals. Proc. Natl. Acad. Sci. USA. 2005;102:14843–8.

27. Wang Y, Mukhopadhyay A, Howitz VR, Binns AN, Lynn DG. Construction of an efficient expression system for Agrobacterium tumefaciens based on the coliphage T5 promoter. Gene. 2000;242:105–14.

28. Koekman BP, Hooykaas PJJ, Schilperoort RA. A functional map of the replicator region of the octopine Ti plasmid. Plasmid. 1982;7:119–32.

29. Beijersbergen A, Dulk-Ras A, Schilperoort RA, Hooykaas PJJ. Conjugative Transfer by the Virulence System of Agrobacterium tumefaciens. Science. 1992;256:1324–7.

30. Koekman BP, Hooykaas PJ, Schilperoort RA. Localization of the replication control region on the physical map of the octopine Ti plasmid. Plasmid. 1980;4:184–95.

31. Vergunst AC, van Lier MCM, den Dulk-Ras A, Stüve T a G, Ouwehand A, Hooykaas PJJ. Positive charge is an important feature of the C-terminal transport signal of the VirB/D4-translocated proteins of Agrobacterium. Proc. Natl. Acad. Sci. USA. 2005;102:832–7.

32. Vergunst AC, Schrammeijer B, den Dulk-Ras A, Vlaam de CMT, Regensburg-Tuïnk TJ, Hooykaas PJJ. VirB/D4-Dependent Protein Translocation from Agrobacterium into Plant Cells. Science. 2000;290:979–82.

33. Yamamoto S, Kiyokawa K, Tanaka K, Moriguchi K, Suzuki K. Novel toxin-antitoxin system composed of serine protease and AAA-ATPase homologues determines the high level of stability and incompatibility of the tumor-inducing plasmid pTiC58. J. Bacteriol. 2009;191:4656–66.

34. den Dulk-Ras A, Hooykaas PJJ. Electroporation of Agrobacterium Tumefaciens. Methods Mol. Biol. 1995. p. 63–72.

35. Jinek M, East A, Cheng A, Lin S, Ma E, Doudna J. RNA-programmed genome editing in human cells. Elife. 2013;2013:1–9. 36.

36. Hooykaas PJJ, Roobol C, Schilperoort RA. Regulation of the transfer of TI plasmids of Agrobacterium tumefaciens. J. Gen. Microbiol. 1979;110:99–109.

Chapter 3

Targeted mutagenesis in yeast with the Cas9 protein translocated through the type IV secretion system of *Agrobacterium*

Daan J. Schmitz, Sylvia de Pater, Paul J.J. Hooykaas

Abstract

The RNA guided Cas9 endonuclease derived from the type II CRISPR/Cas system from *Streptococcus pyogenes* has been developed as a new potent tool for gene knockout in bacteria, yeast, fungi, animal cells and plants. Usually the system is introduced in cells by transfer of the encoding genes. Here we show that the Cas9 protein can also be delivered in cells through the bacterial type IV secretion system (T4SS) of the plant pathogen *Agrobacterium tumefaciens*. Cas9 transfer was effectuated by fusion of a T4SS translocation peptide to the Cas9 protein. Co-cultivation of yeast with an *Agrobacterium* strain expressing both Cas9 and the sgRNA did not lead to mutations in yeast. However after co-cultivation of a yeast expressing the sgRNA for CAN1 (canavanine) with an *Agrobacterium* strain expressing Cas9 (CAN1) mutations were obtained. Concurrent translocation of a T-DNA together with the Cas9 protein was possible and among T-DNA transformations CAN1 mutants could be identified at low frequency. However, concurrent translocation of a T-DNA encoding the sgRNA and Cas9 did not result in targeted mutations at the CAN1 locus.

Introduction

The RNA guided endonucleases encoded by the Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR) and CRISPR-associated (Cas) systems are of great interest to the biotechnology community because of their use in genome editing. In prokaryotes where the CRISPR/Cas system originates, it functions as an adaptive immune system that acts against bacteriophages and other invading nucleic acids [1]. The Cas9 protein from the type II bacterial CRISPR/Cas system found in *Streptococcus pyogenus*, can be programmed by an engineered single guide RNA (sgRNA) to create targeted double strand breaks (DSBs) at a desired sequence in any host cell [2]. By specifying 20 nucleotides of the sgRNA the Cas9 endonuclease can be directed to 20 basepair complementary target sequences [2,3]. A limitation is that the target sequence must be flanked by a protospacer-associated motif (PAM) that is required for Cas9 activity [4,5]. DSBs introduced in the genome with the CRISPR/Cas system can either be repaired via non-homologous end joining (NHEJ) or by homologous recombination (HR). Repair via the error prone NHEJ pathway can result in small deletions and insertions enabling effective reverse genetics. Repair via HR can be used to replace or correct existing genes by providing an artificial repair template with homology to the target sequence. For targeted mutagenesis of host cells the genes encoding the Cas9 nuclease and the sgRNA are usually introduced in the new host and expressed. High expression levels of these transgenes is however not wanted because it increases the frequency of off-target mutations [6,7]. Here, we developed a system for the direct delivery of the Cas9 protein to yeast through the type IV secretion system of Agrobacterium tumefaciens.

This bacterium is a soil-dwelling organism that is most commonly known for causing crown gall disease in plants by genetically transforming these with a set of oncogenes. It is the preferred vector for plant transformation, but is also frequently used as a vector for the transformation of yeasts [8] and fungi [9,10]. *Agrobacterium* uses a Type IV Secretion System (T4SS), encoded by the *virB* genes and *virD4* gene on its Ti plasmid, for the translocation of the T-DNA [11–13]. Along with the T-DNA, several virulence proteins are translocated independently of the T-DNA into the host cell [14]. Translocation of proteins through the T4SS has been shown to be dependent on a hydrophilic C-terminal secretion signal with a net positive charge [15]. Certain heterologous proteins can be translocated through the T4SS

after fusion to the translocation signal of one of the virulence proteins [14,16,17]. In this way the Cre recombinase and homing endonuclease I-SceI were translocated by *Agrobacterium* into host cells to effect DNA recombination in the genome of target cells [14–16]. Here, we developed a system for the direct delivery of the Cas9 protein to yeast through the type IV secretion system of *Agrobacterium*.

Results

Translocation of the Cas9 endonuclease

To test if the Cas9 protein could be translocated through the T4SS of *Agrobacterium*, an expression plasmid was created encoding a Cas9 fusion protein with a N-terminal nuclear localization signal and the C-terminal 37 amino acid translocation signal of the *Agrobacterium* virulence protein VirF (pNCas9F). The production of the fusion protein (NCas9F) was under control of the acetosyringone inducible *virF* promoter to ensure that production would only occur in the presence of a functional T4SS. The CAN1 gene was used as target locus for double strand break induction by Cas9. As an eukaryotic recipient we used a yeast strain expressing a sgRNA targeting the CAN1 locus from the strong TEF1 promoter [18]. This locus encodes a plasma membrane arginine transporter which mediates the uptake of arginine and its toxic analogue L-canavanine into the cell. Mutation of CAN1 leads to L-canavanine resistance. In our experiments translocation of NCas9F could therefore be detected by an increased number of L-canavanine resistant cells after co-cultivation with *Agrobacterium*.

The CAN1 mutation frequency of the recovered yeast cells was found to be ~2.5 * 10^{-5} after co-cultivation with *Agrobacterium* strain LBA1100 expressing NCas9F compared to ~0.2 * 10^{-5} after control co-cultivation with the same *Agrobacterium* lacking the NCas9F expression plasmid. To show that this about tenfold increase in the frequency of L-canavanine resistant colonies was the result of the combined nuclease activity of the translocated NCas9F and the CAN1 sgRNA, the CAN1 locus of 16 L-canavanine resistant colonies obtained after co-cultivation was amplified by PCR and Sanger sequenced. This revealed that 15 of the 16 L-canavanine resistant colonies had mutations directly upstream of the PAM sequence (Fig. 2a), whereas the CAN1 locus of the 16th colony had a single base pair substitution 300 bp upstream of the PAM sequence. The eight L-canavanine resistant colonies obtained after a co-cultivation with an *Agrobacterium* strain lacking NCas9F had mutations that were not located directly upstream of the PAM sequence but were instead randomly distributed throughout the CAN1 gene and therefore represented spontaneous mutations.

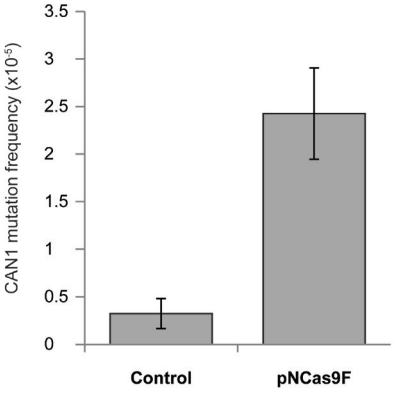

Figure 1. Translocated Cas9 mediated mutagenesis. CAN1 mutation frequency found in yeast after co-cultivation with an *Agrobacterium* strain translocating NCas9F (pNCas9F) and an *Agrobacterium* strain lacking NCas9F (Control). Error bars indicate the SEM (N=3).

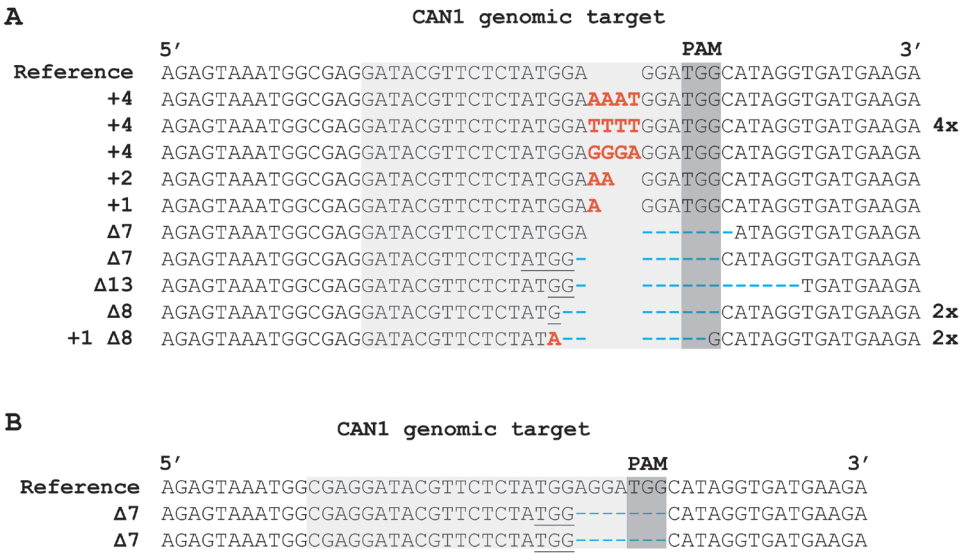

Figure 2. Alignments of the CAN1 target with sequences obtained from L-canavanine resistant colonies after co-cultivation. The PAM sequence is highlighted in dark grey and the recognition sequence of the sgRNA is in light grey. Insertions are marked in red and deletions are marked in blue and microhomologies are underlined. The occurrence of the mutations is noted on the right hand side. (**A**) Mutations found after co-cultivation with an *Agrobacterium* strain delivering the NCas9F protein. (**B**) Mutations found after co-cultivation with an *Agrobacterium* strain delivering both NCas9F and a T-DNA mediating G418 resistance.

Translocation of NCas9F and sgRNA by Agrobacterium

As the previous experiments showed that the NCas9F was functional after translocation we performed an experiment to test the translocation of a NCas9F and a sgRNA complex through the T4SS of *Agrobacterium*. A plasmid was created that puts CAN1 sgRNA production under the control of the *virF* promoter and has the 3' flanking region of *virF* functioning as a terminator sequence (psgRNAvirF5'). The functionality of the *Agrobacterium* sgRNA expression cassette was confirmed with an assay in which DSBs induced by the sgRNA/NCas9F complex stimulated the looping out of a previously plasmid integrated into the genome of *Agrobacterium* (Chapter 2). The frequency of L-canavanine resistant colonies was determined after co-cultivation with several different *Agrobacterium* strains that expressed either the combination of both the sgRNA and NCas9F, neither one or only NCas9F (Fig. 3). A comparison between the different CAN1 mutation frequencies showed the expression of sgRNA in *Agrobacterium* did not enhance the mutation frequency, and that co-cultivations with translocated NCas9F only led to an increased mutation frequency if the sgRNA was expressed in yeast cells. Co-cultivation with the yeast strain expressing the CAN1 sgRNA confirmed that NCas9F protein was translocated from *Agrobacterium* in the presence of sgRNA. These results therefore indicate that sgRNA/NCas9F complexes are not translocated by the T4SS.

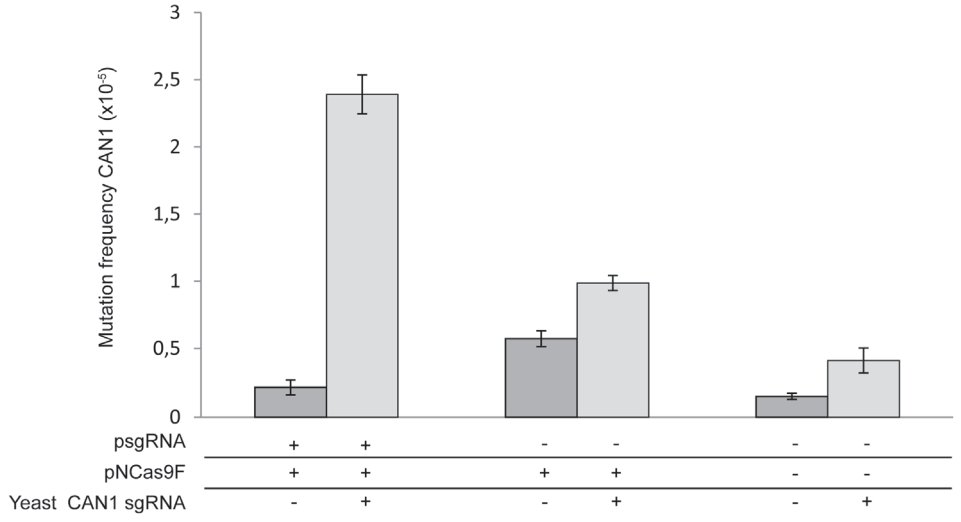

Figure 3. Translocation of sgRNA/NCas9F complex. CAN 1 mutation frequencies after co-cultivation with *Agrobacterium* translocating NCas9F (pNCas9F) in the presence or absence of sgRNA production in (psgRNA) in *Agrobacterium*. Bottom row indicates the presence of the sgRNA targeting the CAN1 in the recipient yeast strain. Error bars indicate the SEM (N=3).

Comparison of mutagenesis by translocated Cas9 protein and Cas9 expressed from a translocated T-DNA

To compare the targeted mutagenesis frequencies between translocated Cas9 and Cas9 from a translocated T-DNA two T-DNA vectors were created both containing an expression cassette in which the NCas9F protein is expressed from the strong TEF1 promoter. The first T-DNA vector pNCas9FPDA1 contains a T-DNA that upon entry of a yeast cell can integrate into the yeast genome at the PDA1 locus via homologous recombination. The second T-DNA vector contains a T-DNA with a 2μ origin of replication that circularizes and replicates in yeast (pNCas9F2μ). Both of these T-DNA vectors carry a dominant KanMX selectable marker which allows selection of transgenic yeast resistant to G418. Immediately after co-cultivation with a recipient yeast expressing a sgRNA directed to the CAN1 locus the mutation frequency was determined by selection on L-canavanine. The results (Fig. 4) obtained indicated that co-cultivation with an *Agrobacterium* strain translocating NCas9F results in a CAN1 mutation frequency that is in the same order of magnitude as the mutation frequency after co-cultivation with either of the strains delivering a T-DNA from which NCas9F is expressed.

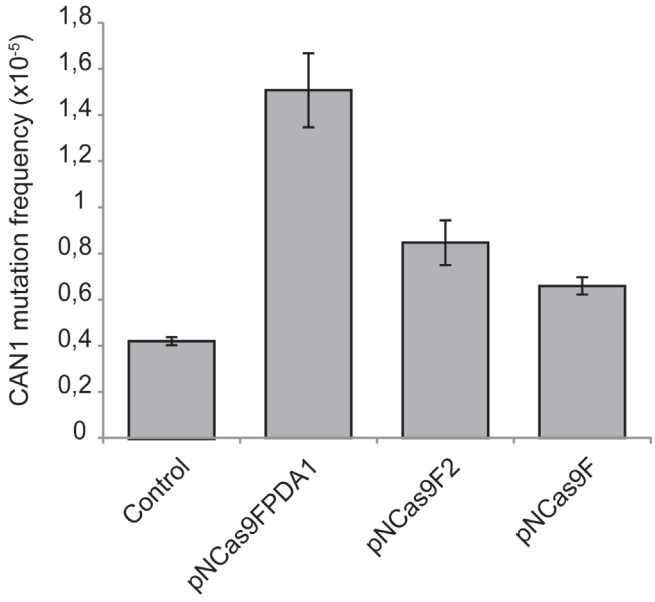

Figure 4. Translocated NCas9F compared to constitutively expressed NCas9F. Mutation frequencies of CAN1 (x10-5) after co-cultivation of a recipient yeast expressing the CAN1 sgRNA with *Agrobacterium* donor strains with integrative T-DNA vector expressing NCas9F (pNCas9FPDA1), circularizing T-DNA vector expressing NCas9F (pNCas9F2µ) or protein translocation vector with NCas9F (pNCas9F). Error bars indicate the SEM (N=3).

Recovery frequency of mutants at non-selectable loci

To determine at which frequency yeast colonies with mutations at non-selectable loci could be recovered by our methodology, an *Agrobacterium* strain was created that not only translocates the NCas9F protein but also a T-DNA conferring G418 resistance (pSDM8002). This T-DNA was used to determine the frequency at which mutated yeast colonies could be recovered for non-selectable loci.

A co-cultivation was conducted with this *Agrobacterium* strain and a yeast strain expressing the sgRNA targeting the CAN1 locus. After co-cultivation yeast transformants were selected on G418 containing medium, and the G418 resistant colonies were subsequently tested for L-canavanine resistance. Out of a total of 1890 transformants, two colonies from independent co-cultivations were found that were also L-canavanine resistant. Sequencing of the CAN1 locus revealed that both colonies had the same 7 basepair deletion upstream of the PAM (Fig. 2b) that was previously found already in the first experiment. These results indicate that after NCas9 translocation targeted mutations can be recovered in yeast at non-selectable loci at a frequency of about 1:1000 T-DNA transformants.

Translocation of both Cas9 and a T-DNA expressing the sgRNA

As we could not obtain evidence for translocation of the complete NCas9F complex from *Agrobacterium* into yeast, we subsequently studied whether we could obtain targeted mutagenesis in yeast by translocation of the NCas9F protein combined with the translocation of a T-DNA coding for the sgRNA could induce targeted mutations. For this purpose two sgRNA expression vectors were created, an integrative T-DNA vector (psgRNAPDA1) and a

circularizing T-DNA vector (psgRNA2µ) both conferring to G418 resistance. Co-cultivations were performed with *Agrobacterium* strains expressing NCas9F in combination with either of the sgRNA encoding T-DNA vectors. These co-cultivations did however not increase the CAN1 mutation frequency (Fig. 5). Transfer of both T-DNA's was confirmed by selection on G418. The transfer of the NCas9F protein in the presence of either T-DNA vector was confirmed when the co-cultivations were performed with a recipient yeast strain expressing the CAN1 sgRNA (Fig. 5). Because previous experiments showed that the concurrent transfer of the NCas9F protein and a T-DNA occurs at low frequencies, the same co-cultivations were performed at a much larger scale and a total of $2*10^9$ recovered yeast cells was plated on dual selection medium containing both L-canavanine and G418. No colonies were however found that were L-canavanine resistant due to mutations directly upstream of the PAM. These results combined indicate that concurrent transfer of the NCas9F protein and a T-DNA expressing the sgRNA did not induce targeted mutations.

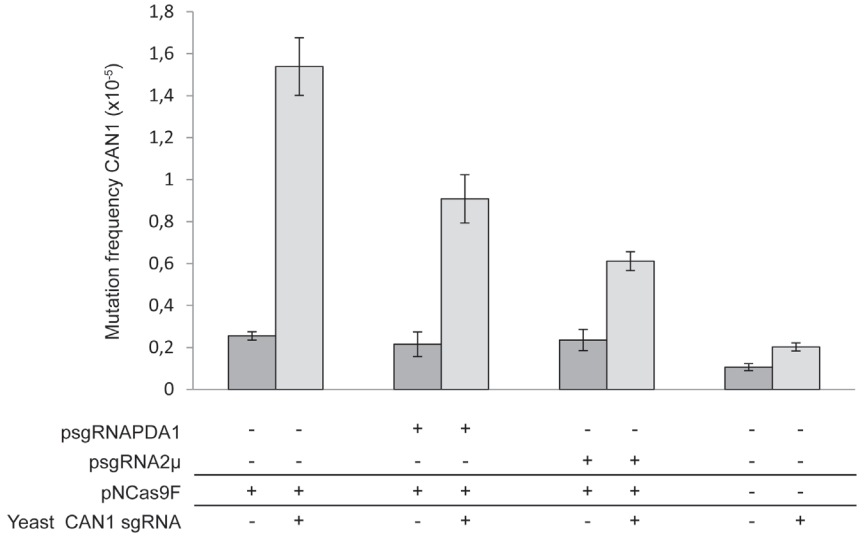

Figure 5. Mutation frequency after co-cultivation with integrative T-DNA expressing the CAN1 sgRNA (psgRNAPDA1), circularizing T-DNA expressing the CAN1 sgRNA (psgRNA2µ) and translocated NCas9F (pNCAS9F). CAN1 mutation frequencies after co-cultivation with yeast strain not expressing the sgRNA (dark grey) and a yeast strain expressing the sgRNA (light grey). Bottom row indicates the presence of the sgRNA targeting the CAN1 in the recipient yeast strain. Error bars indicate the SEM (N=3).

Discussion

In this study we have shown that it is possible to translocate Cas9 of the type II bacterial CRISPR system from *Agrobacterium* to yeast. After translocation, Cas9 was capable of forming a complex with a sgRNA and create targeted DSBs that after imperfect repair via NHEJ resulted in targeted mutations. Several of the mutations found within the CAN1 gene were probably the result of repair using small existing 4bp microhomologies (Fig. 2a and Fig. 2b). This may be one of the preferred ways to repair the DSB in the absence of a repair template [19–21].

The mutation frequency which we found with translocated NCas9F is about a tenfold lower than the previously reported mutation frequencies obtained when Cas9 was expressed directly from a plasmid in yeast [18]. This could be due to limiting levels of translocated NCas9F, the short presence of NCas9F in the host after translocation or because not every yeast cell in the co-cultivation mixture is susceptible to *Agrobacterium* mediated transfer of NCas9F, or could be the consequence of a combination of these factors.

Mutations induced by the translocation of the sgRNA/NCas9F complex from *Agrobacterium* to yeast could not be detected in the experiments performed. Mutations were however found if an *Agrobacterium* strain expressing both NCas9F and the sgRNA was co-cultivated with a yeast strain expressing the sgRNA, confirming the translocation of NCas9F in the presence of the sgRNA. Furthermore we showed that the sgRNA/NCas9F complex forms a functional complex in *Agrobacterium* (Chapter 2). Therefore we speculate that the large size of the NCas9F/sgRNA complex (100 Å x 100 Å x 50 Å) [22] and other large proteins probably requires (partial) unfolding to be transferred complex through the T4SS pore which is only 10 Å at its narrowest point [23]. Such unfolding would disrupt the interaction between NCas9F and the sgRNA.

The mutation frequencies found with translocated NCas9F protein and Cas9 expressed from a T-DNA were in the same order of magnitude. This showed that protein translocation of Ncas9F occurs at a similar frequency as T-DNA transfer. Transient presence of NCas9F could have a positive effect on the frequency of off-target mutations by the limiting availability of the nuclease [6,7].

Combined transfer of NCas9F and a T-DNA encoding the CAN1 sgRNA did not result in targeted mutations after co-cultivation, although transfer of both NCas9F and the T-DNA could be detected individually. Combined transfer of NCas9F and a T-DNA encoding the sgRNA is however capable of inducing targeted mutations in plants after ten days (unpublished data). Therefore we speculate that the lack of targeted mutations is caused by insufficient levels of sgRNA, or because sufficient levels of sgRNA requires the double strand formation of the T-DNA expressing the sgRNA at which point the translocation of NCas9F has potentially already stopped.

In summary the experimental data discussed above show that it is possible to use *Agrobacterium* for the translocation of the Cas9 protein into yeast cells to create targeted mutations in yeasts. Because under laboratory conditions *Agrobacterium* is capable of transforming various other hosts including filamentous fungi, our methodology may be applicable more broadly.

Material and methods

Yeast strains and media

Derivatives of the *Saccharomyces cerevisiae* strain YPH499 (MATa ura3-52 lys2-801 ade2-101 trp1-Δ63 his3-Δ200 leu2-Δ1) were used in all experiments. YPH499 was grown on YEPD before transformation and then propagated on SD medium minus the auxotrophic compound(s) complemented by the plasmid(s). Transformation of YPH499 was done using the standard LiAc method [24].

Agrobacterium strains

The *Agrobacterium* strain LBA1100 (C58 containing pTiB6Δ (ΔT-DNA, Δocc, Δtra), Rif, Spc) was used for all experiments [25]. *Agrobacterium* was grown in LB (5 g/l NaCl) medium containing the appropriate antibiotics at the following concentrations: gentamicin 40 μg/ml, rifampicin 10 μg/ml, spectinomycin 250 μg/ml, kanamycin 100 μg/ml. Plasmids were electroporated to LBA1100 as described in den Dulk-Ras and Hooykaas (1995).

Plasmid construction

For the construction of pNCas9F see Chapter 2. The pSDM8002 backbone was created by cloning the 2μ replicon EcoRI fragment from pSLF178k [28] into the EcoRI site of pSDM8000 [29]. To create the T-DNA vectors that expres NCas9F, the NLS::Cas9F::VirF fragment was amplified via PCR with DS122 and DS123. This PCR fragment was digested with XbaI, and cloned into the XbaI and blunted (Klenow) XhoI site of p414-TEF1p-Cas9-CYC1t (Addgene #43802, [18]). From the resulting backbone the NLS::Cas9::VirF including the 5' TEF1 promoter and the 3' CYC1 terminator was amplified by PCR with DS124 and DS125. This PCR fragment was cloned into the blunted XbaI site (Klenow) of pSDM8001 [30] to create pNCas9FFPDA1 and the blunted XhoI site of pSDM8002 to create pNCas9F2μ. The pSDM8002 backbone was created by cloning the 2μ replicon EcoRI fragment from pSLF178k [28] into EcoRI site of pSDM8000 [29]. The T-DNA sgRNA expression vectors were created by amplifying the CAN1 sgRNA expression cassette including the SNR52 promoter and the SUP4 terminator from p426-SNR52p-gRNA.CAN1.Y-SUP4t (Addgene #43803, [18]) via PCR with either DS114 and DS115 or DS119 and DS120 and inserted into the XbaI site of pSDM8001 and the XhoI site of pSDM8002, respectively.

To create the *Agrobacterium* sgRNA expression vector, the *virF* 3' region was amplified using DS097 and DS099 adding a XbaI, StuI and SalI site and two flanking BglII sites. The resulting PCR fragment was cloned into the BglII site pOPHIS Borderless [31]. The *virF* 5' flanking region was amplified using DS100 and DS101 and inserted into the XbaI and StuI sites of pOPHIS with the *virF* 3' flanking region. The complementary oligo's DS104 and DS105 were annealed and inserted into the BsbI sites of pEN_Chimera [32]. From pEN_Chimera the entire CAN1 sgRNA cassete was amplified by PCR with DS102 and DS103 and cloned into the StuI and SalI sites of pOPHIS with the *virF* 5' and *virF* 3' flanking regions creating psgRNA.

Co-cultivations of Agrobacterium with yeast

Co-cultivations of *Agrobacterium* with the YPH499 yeast strain were carried out as previously described with the following minor modifications [33]. *Agrobacterium* was grown overnight at 29°C in LB (5 g/l NaCl) with appropriate antibiotics. Induction of the *virF* promoter was performed at 28°C at OD_{600} = 0.25 for 6 hours in induction medium [30] containing 200 μM acetosyringone (Sigma-Aldrich Co.). The yeast strain with the p426-SNR52p-gRNA.CAN1.Y-SUP4t plasmid [18] was grown overnight in MY minus the auxotrophic growth compound (uracil) complemented by the plasmid and then diluted 10 times in YEPD and cultured for 6 hours. Yeast (10^7 cells) and *Agrobacterium* ($2*10^8$ cells) were mixed and spotted on nitrocellulose filters on induction medium plates [8] containing 200μM acetosyringone (Sigma-Aldrich Co.) followed by an incubation at 21°C for 7 days.

Determination of the CAN1 mutation frequency and T-DNA transfer frequency

Yeast was recovered from the nitrocellulose filter and plated on minimal yeast medium containing L-canavanine sulfate (60 μg mL^{-1}, Santa Cruz Biotechnology Inc.) to select for mutations of the CAN1 locus and cefotaxime (200 μg mL^{-1}, FORMEDIUM™) to stop the growth of *Agrobacterium*. Total yeast cell numbers were determined by plating serial dilutions of yeast on YEPD plates containing cefotaxime (200 μg mL^{-1}, FORMEDIUM™). The CAN1 mutation frequency was determined by dividing the number of L-canavanine resistant colonies by total colony count based on the serial dilutions on YEPD. To determine the T-DNA transfer frequency yeast recovered from the nitrocellulose filter was plated on YEPD containing G418 (200 μg mL^{-1}, FORMEDIUM™) and cefotaxime (200 μg mL^{-1}, FORMEDIUM™). The transfer frequency was determined by dividing the number of G418 resistant colonies by the total colony count based on the serial dilutions on complete yeast medium.

Mutation analysis of the CAN1 locus

Yeast was grown overnight in YEPD at 30°C. Genomic DNA was isolated from 2 ml cultures using the yeastar^{tm} genomic DNA kit from Zymoclean (protocol 1). The CAN1 locus was amplified from L-canavanine resistant colonies by PCR with primers DS086 and DS088 and the PCR fragment was cloned into pJET2.1 (CloneJET PCR Cloning Kit, Thermo Fischer Inc.) before sequencing (Macrogen Europe Inc.).

Table 1. Overview of primers used in this study

Primer name	Sequence
DS060	GATCTACTAGTGCTGCACGG
DS061	GATCCCGTGCAGCACTAGTA
DS072	AGTCAGATCTGAAAAACATCAAAAAAAACCAG
DS073	GCTAAGATCTCGTTTCAGGAAAGTTTCGGAGGAG
DS086	CTTCAGACTTCTTAACTCCTGT
DS088	TGAGGGTGAGAATGCGAAATG
DS097	AGGCCTCCTTCCGTCGACCTATCCGTGCTGTTCGTCAC
DS098	AGATCTTCTAGACCTTCCAGGCCTCCTTCCGTCGACCT
DS099	AGATCTGGGACCAGCACACTTAGATA
DS100	TCTAGAAGCTCCTATGATAGTCGATA
DS101	AGGCCTATCGCTCCTGTGCTTTTGAA
DS102	AGGCCTATTGGGGTCTTCGAGAAGAC
DS103	GTCGACTAATGCCAACTTTGTACAAG
DS104	ATTGGATACGTTCTCTATGGAGGA
DS105	AAACTCCTCCATAGAGAACGTATC
DS114	CTCGAGTCTTTGAAAAGATAATGTATG
DS115	CTCGAGAGACATAAAAAACAAAAAAAG
DS119	ACTAGTTCTTTGAAAAGATAATGTATG
DS120	ACTAGTAGACATAAAAAACAAAAAAAG
DS122	CTATTCTAGAATGGATAAAGCGGAATTAAT
DS123	ATCGGAATTCTCATAGACCGCGCGTTGATC
DS124	CCCGGGAGCTCCGGATGCAAGGGTTC
DS125	CCCGGGGGTACCGGCCGCAAATTAAAG
DS126	ATGGATAAAGCGGAATTAATTCC
DS127	TCATAGACCGCGCGTTGATCG

Table 2. Overview of plasmids used in this study

Name	Description	Vector type, organism
pCas9F	pvirFpromoter:NLS::Cas9::VirF37C	Expression vector, Agrobacterium
pCas9FPDA1	rightborder:PDA1:KANMX:tef:NLS::Cas9::VirF37C:cyc1t:PDA1:leftborder	Binary vector, yeast
pCas9F2μ	rightborder:KANMX:tef:NLS::Cas9::VirF37C:cyc1t:2μ:leftborder	Binary vector, yeast
psgRNAPDA1	rightborder:PDA1:KANMX:SNR52:gRNACan1:SUP43':PDA1:leftborder	Binary vector, yeast
psgRNA2μ	rightborder:KANMX:SNR52:gRNACan1:SUP43':2μ:leftborder	Binary vector, yeast
pSDM8001	rightborder:PDA1:KANMX:PDA1:leftborder	Binary vector, yeast
pSDM8002	rightborder:KANMX:2μ:leftborder	Binary vector, yeast
psgRNA	pvirFpromoter:SNR52:gRNACan1:SUP4:3'virF	Expression vector, Agrobacterium
p426-SNR52p-gRNA.CAN1.Y-SUP4t	SNR52:gRNACan1:SUP4	Expression vector, yeast

Acknowledgements

This work was supported by TTI-Green Genetics (project number TTI GG CORE 4CC057RP). The authors would like to thank George Church for providing the plasmid p426-SNR52p-gRNA.CAN1.Y-SUP4t (Addgene plasmid # 43803).

References

1. Oost J Van Der, Jore MM, Westra ER, Lundgren M, Brouns SJJ. CRISPR-based adaptive and heritable immunity in prokaryotes. Trends Biochem Sci. 2009;401–7.

2. Jinek M, Chylinski K, Fonfara I, Hauer M, Doudna JA, Charpentier E. A Programmable Dual-RNA–Guided DNA Endonuclease in Adaptive Bacterial Immunity. Science. 2012;337:816–22.

3. Bhaya D, Davison M, Barrangou R. CRISPR-Cas Systems in Bacteria and Archaea: Versatile Small RNAs for Adaptive Defense and Regulation. Annu. Rev. Genet. 2011;45:273–97.

4. Deveau H, Barrangou R, Garneau JE, Labonté J, Fremaux C, Boyaval P, et al. Phage response to CRISPR-encoded resistance in Streptococcus thermophilus. J. Bacteriol. 2008;190:1390–400.

5. Mojica FJM, Díez-Villaseñor C, García-Martínez J, Almendros C. Short motif sequences determine the targets of the prokaryotic CRISPR defence system. Microbiology. 2009;155:733–40.

6. Hsu PD, Scott DA, Weinstein JA, Ran FA, Konermann S, Agarwala V, et al. DNA targeting specificity of RNA-guided Cas9 nucleases. Nat. Biotechnol. 2013;31:827–32.

7. Pattanayak V, Lin S, Guiliniger JP, Ma E, Doudna JA, Liu DR. High-throughput profiling of off-target DNA cleavage reveals RNA-programmed Cas9 nuclease specificity. Nat. Biotechnol. 2013;31:839–43.

8. Bundock P, den Dulk-Ras A Beijersbergen A, Hooykaas PJJ. Trans-kingdom T-DNA transfer from Agrobacterium tumefaciens to Saccharomyces cerevisiae. EMBO J. 1995;14:3206–14.

9. de Groot MJ, Bundock P, Hooykaas PJJ, Beijersbergen GM. Agrobacterium tumefaciens-mediated transformation of filamentous fungi. Nat. Biotechnol. 1998;16:839–42.

10. Michielse CB, Hooykaas PJJ, van den Hondel CA MJJ, Ram AFJ. Agrobacterium-mediated transformation as a tool for functional genomics in fungi. Curr. Genet. 2005;48:1–17.

11. Christie PJ, Atmakuri K, Krishnamoorthy V, Jakubowski S, Cascales E. Biogenesis, architecture, and function of bacterial type IV secretion systems. Annu. Rev. Microbiol. 2005;59:451–85.

12. Jiang W, Zhou H, Bi H, Fromm M, Yang B, Weeks DP. Demonstration of CRISPR/Cas9/sgRNA-mediated targeted gene modification in Arabidopsis, tobacco, sorghum and rice. Nucleic Acids Res. 2013;41:1–12.

13. Alvarez-Martinez CE, Christie PJ. Biological diversity of prokaryotic type IV secretion systems. Microbiol. Mol. Biol. Rev. 2009;73:775–808.

14. Vergunst AC, Schrammeijer B, den Dulk-Ras A, Vlaam de CMT, Regensburg-Tuïnk TJ, Hooykaas PJJ. VirB/D4-Dependent Protein Translocation from Agrobacterium into Plant Cells. Science. 2000;290:979–82.

15. Vergunst AC, van Lier MCM, den Dulk-Ras A, Stüve TG, Ouwehand A, Hooykaas PJJ. Positive charge is an important feature of the C-terminal transport signal of the VirB/D4-translocated proteins of Agrobacterium. Proc. Natl. Acad. Sci. 2005;102:832–7.

16. Rolloos M, Hooykaas PJJ, van der Zaal BJ. Enhanced targeted integration mediated by translocated I-SceI during the Agrobacterium mediated transformation of yeast. Sci. Rep. 2015;5:8345.

17. Schrammeijer B, den Dulk-Ras A, Vergunst A, Jurado Jácome E, Hooykaas PJJ. Analysis of Vir protein translocation from Agrobacterium tumefaciens using Saccharomyces cerevisiae as a model: evidence for transport of a novel effector protein VirE3. Nucleic Acids Res. 2003;31:860–8.

18. Dicarlo JE, Norville JE, Mali P, Rios X, Aach J, Church GM. Genome engineering in Saccharomyces cerevisiae using CRISPR-Cas systems. Nucleic Acids Res. 2013;41:4336–43.

19. Yu X, Gabriel A. Ku-dependent and Ku-independent end-joining pathways lead to chromosomal rearrangements during double-strand break repair in Saccharomyces cerevisiae. Genetics. 2003;163:843–56.

20. Ma J-L, Kim EM, Haber JE, Lee SE. Yeast Mre11 and Rad1 Proteins Define a Ku-Independent Mechanism To Repair Double-Strand Breaks Lacking Overlapping End Sequences. Mol. Cell. Biol. 2003;23:8820–8.

21. Boulton SJ, Jackson SP. Identification of a Saccharomyces cerevisiae Ku80 homologue: roles in DNA double strand break rejoining and in telomeric maintenance. Nucleic Acids Res. 1996;24:4639–48.

22. Jinek M, Jiang F, Taylor DW, Sternberg SH, Kaya E, Ma E, et al. Structures of Cas9 endonucleases reveal RNA-mediated conformational activation. Science. 2014;343:1247997.

23. Rivera-Calzada A, Fronzes R, Savva CG, Chandran V, Lian PW, Laeremans T, et al. Structure of a bacterial type IV secretion core complex at subnanometre resolution. EMBO J. 2013;32:1195–204.

24. Schiestl RH, Gietz RD. High efficiency transformation of intact yeast cells using single stranded nucleic acids as a carrier. Curr. Genet. 1989;16:339–46.

25. Beijersbergen A, Dulk-Ras AD, Schilperoort RA, Hooykaas PJJ. Conjugative Transfer by the Virulence System of Agrobacterium tumefaciens. Science. 1992;256:1324–7.

26. den Dulk-Ras A, Hooykaas PJJ. Electroporation of Agrobacterium Tumefaciens. Methods Mol. Biol. 1995. p. 63–72.

27. Jinek M, East A, Cheng A, Lin S, Ma E, Doudna J. RNA-programmed genome editing in human cells. Elife. 2013;2013:1–9.

28. Forsburg SL, Guarente L. Mutational Analysis of Upstream Activation Sequence 2 of the CYCI Gene of Saccharomyces cerevisiae : a HAP2-HAP3- Responsive Site UAS1 UAS2. Mol Cell Biol. 1988;8:647–54.

29. van Attikum H, Bundock P, Hooykaas PJ. Non-homologous end-joining proteins are required for Agrobacterium T-DNA integration. EMBO J. 2001;20:6550–8.

30. van Attikum H, Hooykaas PJJ. Genetic requirements for the targeted integration of Agrobacterium T-DNA in Saccharomyces cerevisiae. Nucleic Acids Res. 2003;31:826–32.

31. Rolloos M, Dohmen MHC, Hooykaas PJJ, Zaal BJ Van Der. Involvement of Rad52 in T-DNA circle formation during Agrobacterium tumefaciens-mediated transformation of Saccharomyces cerevisiae. Mol Microbiol. 2014;91:1240–51.

32. Fauser F, Schiml S, Puchta H. Both CRISPR/Cas-based nucleases and nickases can be used efficiently for genome engineering in Arabidopsis thaliana. Plant J. 2014;79:348–59.

33. Bundock P, Van Attikum H, Dulk-Ras A Den, Hooykaas PJJ. Insertional mutagenesis in yeasts using T-DNA from Agrobacterium tumefaciens. Yeast. 2002;19:529–36.

Chapter 4

CRISPR/Cas-induced mutagenesis in Nicotiana benthamiana through Cas9 protein translocation via the type IV secretion system of *Agrobacterium*

Daan J. Schmitz, Sylvia de Pater, Paul J.J. Hooykaas

Abstract

The RNA guided endonuclease based on the CRISPR/Cas system of *Streptococcus pyogenus* is a potent new tool for genome engineering in plants. Delivery of this system is usually done by transforming plants with *Agrobacterium tumefaciens* which transfers a T-DNA encoding the required genes through its Type IV Secretion System (T4SS). Instead, here we report the delivery of the Cas9 protein of the CRISPR/Cas system of *Streptococcus pyogenus* through the T4SS of *Agrobacterium*, simultaneously with transfer of a T-DNA encoding the sgRNA directed against the phytoene desaturase gene. Transfer of Cas9 was accomplished by fusion of a T4SS translocation peptide to the Cas9 protein. Infiltration of *Nicotiana benthamiana* leaves with an *Agrobacterium* suspension resulted in targeted mutations at the phytoene desaturase locus. Deep amplicon sequencing showed that translocation of NCas9F through the T4SS resulted in deletion and insertion frequencies that are eighteen and twelve fold lower, respectively, than after delivery of a T-DNA encoding both Cas9 and the sgRNA. Mutations induced by translocated NCas9F were found upstream of the PAM and were distributed in the same manner as mutations induced with Cas9 expressed from a T-DNA.

Introduction

Genome editing in plants can be achieved by the induction of site-specific double stranded breaks (DSBs) in the genome with several classes of nucleases. Meganucleases, Zinc-Finger Nucleases (ZFNs), Transcription Activator-Like Effector Nucleases (TALENs) and more recently the CRISPR/Cas system (for: clustered regularly interspaced palindromic repeats / CRISPR-associated) have become important tools for the targeted genetic modification of plant genomes. DSBs induced by these nucleases can be repaired through non-homologous end joining (NHEJ) or homologous recombination (HR). Repair via error-prone NHEJ can result in small insertions and deletions enabling reverse genetics. By providing an artificial repair template with homology to the target sequence, repair via HR can be used to replace or correct existing genes.

The CRISPR/Cas system, originating from prokaryotes, functions as an adaptive immune system that acts against invading bacteriophages and other invading nucleic acids [1]. The prototypical CRISPR/Cas endonuclease from the type II bacterial CRISPR/Cas system found in *Streptococcus pyogenus* can be programmed to create targeted double strand breaks (DSBs). The tracrRNA and crRNAs can be engineered into a single guide RNA (sgRNA) [2]. The Cas9 endonuclease can be directed to a specific target sequence by the incorporation of a 20 nucleotide sequence in the sgRNA that is complementary to the target sequence [2,3]. This 20 base pair (bp) target sequence has to be flanked by a 3 bp protospacer-associated motif that is required for Cas9 activity [4,5].

For targeted mutagenesis in plant cells, the genes encoding the RNA guided endonuclease system are usually introduced using *Agrobacterium tumefaciens*. Translocation of the T-DNA occurs via its Type IV Secretion System (T4SS), encoded by the *virB* genes and *virD4* gene on its Ti plasmid [6,7]. Independently of the T-DNA, several virulence proteins are transported alongside the T-DNA into the host cell [8]. A hydrophilic secretion signal with a net positive charge is responsible for the translocation of proteins through the T4SS [9]. Several heterologous proteins, fused to this secretion signal, have been translocated through the T4SS [8–11]. In this way the Cre recombinase and homing endonuclease I-SceI were translocated by *Agrobacterium* into host cells to effect DNA recombination in the genome of target cells [8–10].

In this study we show that delivery of the Cas9 protein through the T4SS of *Agrobacterium* can be used for targeted mutagenesis in *Nicotiana benthamiana*.

Results

Engineering the Cas9 expression constructs

In order to be able to test the translocation of the Cas9 protein of the type II class of CRISPR/Cas systems from bacteria to *N. benthamiana* cells a translational fusion between Cas9 and the last 37 amino acids of the *Agrobacterium* virulence protein VirF was created. These 37 amino acids are known to function as a recognition signal for the T4SS of *Agrobacterium* and can be used to target proteins for translocation [8–10]. Target recognition of the CRISPR/Cas system requires an RNA guide. In Chapter 3, it was shown that the sgRNA, when expressed in *Agrobacterium*, was not translocated via the T4SS together with the Cas9 protein to yeast cells. Therefore, a T-DNA vector was created encoding a sgRNA targeting the phytoene desaturase (*PDS*) gene, for expression of the sgRNA in the host cells. The protospacer of this sgRNA contains an GG motif at the 3' end of the target sequence, which has been shown to increase the frequency of targeted mutagenesis [12]. To compare the targeted mutagenesis frequency of the translocated NCas9F protein to constitutively expressed Cas9, a T-DNA vector was used expressing Cas9 under control of the ubiquitin promoter and the sgRNA under control of the small nuclear RNA promoter U6-26 [13].

Targeted mutagenesis in N. benthamiana leaves with translocated Cas9

To test the functionality of the translocated Cas9 in *N. benthamiana*, leaves were infiltrated with an *Agrobacterium* strain expressing the NCas9F protein and a T-DNA encoding the PDS sgRNA or a T-DNA expressing both Cas9 and the sgRNA. Ten days post infiltration, the tissue was harvested and the DNA was extracted. To easily detect mutations induced by the NCas9F nuclease we used the restriction enzyme loss method [14], as the target sequence of the sgRNA overlaps with a DdeI restriction site. To enrich for DNA molecules carrying mutations the genomic DNA of the infiltrated leaves was pre-digested with DdeI. A 490 bp fragment was then amplified by PCR with primers flanking the target sequence, and the resulting PCR products were again digested with DdeI. Restriction digestion resistant bands were cloned and analyzed by sequencing. Translocation of a T-DNA encoding both Cas9 and the sgRNA resulted in restriction digestion resistant PCR products in all of the nine infiltrated leaves (Fig. 1a). The resistent PCR products were isolated and cloned into a high-copy vector transformed into *E. coli* and individual plasmid clones were sequenced. Deletions ranging from 1 to 11 bp, single nucleotide insertions and small deletions combined with insertions were found (Fig. 1b). Translocation of the NCas9F protein together with a T-DNA encoding the PDS sgRNA resulted in restriction digestion resistant PCR products in seven out of nine infiltrated leaves (Fig. 1a). The sequences of the PCR products were very similar to those obtained by T-DNA encoded Cas9, including deletions ranging from 1 to 9 bp, single nucleotide insertions and combinations (Fig. 1c). The results showed that translocated NCas9F protein together with a sgRNA encoded by a T-DNA can bring about targeted mutagenesis in cells of *N. benthamiana* leaves.

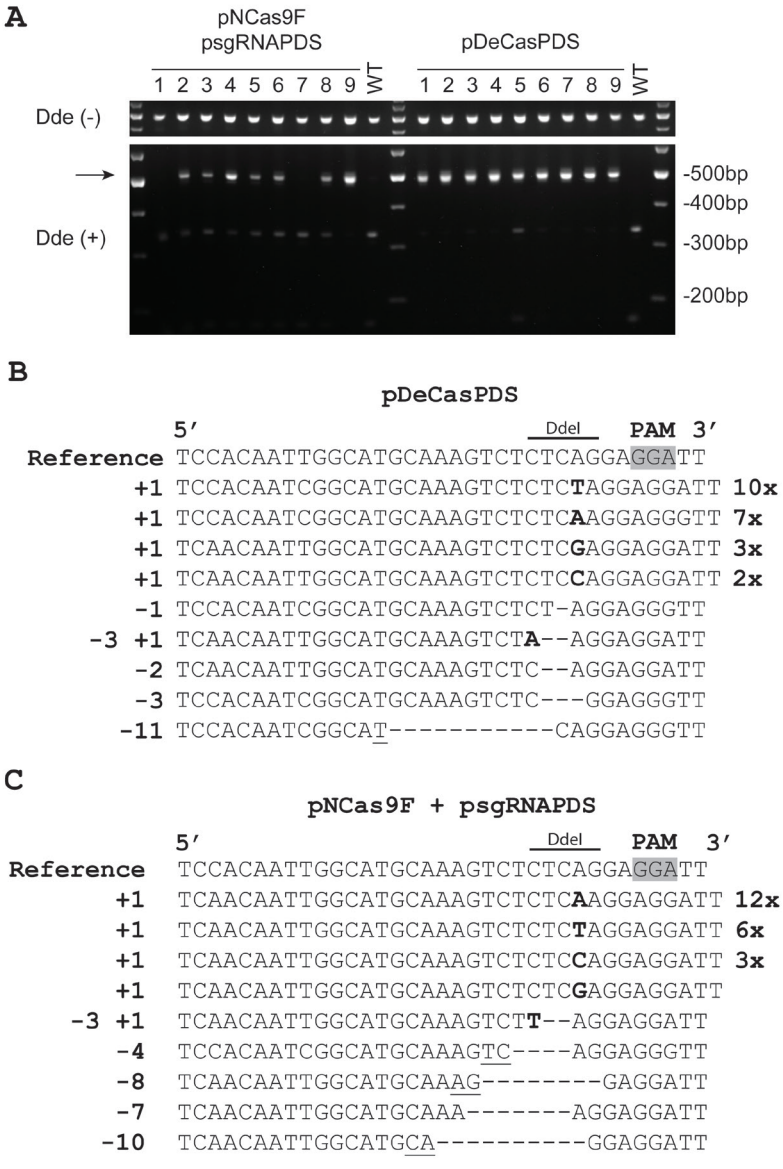

Figure 1. CRISPR/Cas9 endonuclease-induced mutagenesis. The *PDS* target site was amplified using genomic DNA that was pre-digested with DdeI. The resulting 490bp PCR products were digested with DdeI. (**A**) DdeI resistant 490bp bands (arrow) obtained from tissue co-transformed by NCas9F protein and T-DNA encoding the sgRNA and by pDeCasPDS containing genes encoding Cas9 and sgRNA. (**B**) Sequence analysis of mutations obtained after NCas9F protein translocation and T-DNA encoding the sgRNA. (**C**) Sequence analysis of mutations obtained after transfer of T-DNA encoding Cas9 and sgRNA. The DdeI target site and PAM are indicated. Insertions are shown in bold, deletions with dashes and microhomologies are underlined. Length of insertions and deletions are shown at the left. Numbers of multiple clones with the same sequence are indicated at the right.

Efficiency of mutagenesis by translocated Cas9 protein in comparison with Cas9 expressed from a T-DNA

In order to detect mutations, enrichment by restriction enzyme digestion was applied in the experiments described in the previous paragraph. In order to compare the efficiency of targeted mutagenesis after NCas9F protein delivery with that after T-DNA transfer directly, amplicon deep sequencing was performed. Amplicons were generated by PCR from DNA isolated from nine independently infiltrated leaves, 10 days after inoculation. These PCR amplicons covering the protospacer and PAM sequences were sequenced using the illumina HiSeq 2500 platform. Figure 2 depicts the results as relative numbers of mutated reads per position. Most deletions detected were found 3 to 13 bp upstream of the PAM sequence irrespective of how Cas9 was delivered (Fig. 2a). Deletions in cells expressing Cas9 and sgRNA from a single T-DNA were likewise found 3 to 13 bp upstream of the PAM, but at a higher frequency than in cells into which the NCas9F protein had been translocated (Fig. 2a). Almost all insertions found with both translocated NCas9F and Cas9 and sgRNA expressed from a single T-DNA were located 3 to 4 bp upstream of the PAM (Fig. 2b). Such single bp insertions were not detected in sequenced amplicons obtained from non-infiltrated *N. benthamiana* tissue. The percentage of sequences with a deletion obtained from tissue with translocated NCas9F was found to be 0.24%, but the percentage seen after expression of Cas9 and sgRNA expressed from a single T-DNA was found to be eighteen fold higher at 4.31%. Similarly, the insertion frequency with translocated NCas9F was found to be 1.55% and 12 fold higher at 19.8% after expression of Cas9 and sgRNA from a single T-DNA. Further analysis revealed that around ~80% of these single bp insertions are adenosine or thymidine insertions.

These results show that mutations created with translocated NCas9F protein have a similar distribution of both deletions and insertions as those seen after expression from a single T-DNA, albeit at 12-18 fold lower frequencies.

Regeneration of plants from mutated leaves

As previous experiments showed that translocated NCas9F was effective in creating targeted mutations in the leaf cells of *N. benthamiana* we tested if plants could be regenerated from such cells. Leaf discs were excised from *N. benthamiana* leaves infiltrated with an *Agrobacterium* strain expressing the NCas9F protein and a T-DNA encoding the sgRNA and were placed on selective medium containing kanamycin to select for the presence of the T-DNA. This resulted in the formation of 150 resistant calli from which shoots were regenerated on shoot-induction medium. Total DNA was isolated from these regenerated shoots from which the target locus was amplified by PCR. The amplified bands were treated with the restriction enzyme DdeI as loss of the restriction site was taken as indicator for the presence of a mutation. Undigested bands were not found, which indicated that none of the regenerated shoots had homozygous or heterozygous mutation at the desaturase target locus.

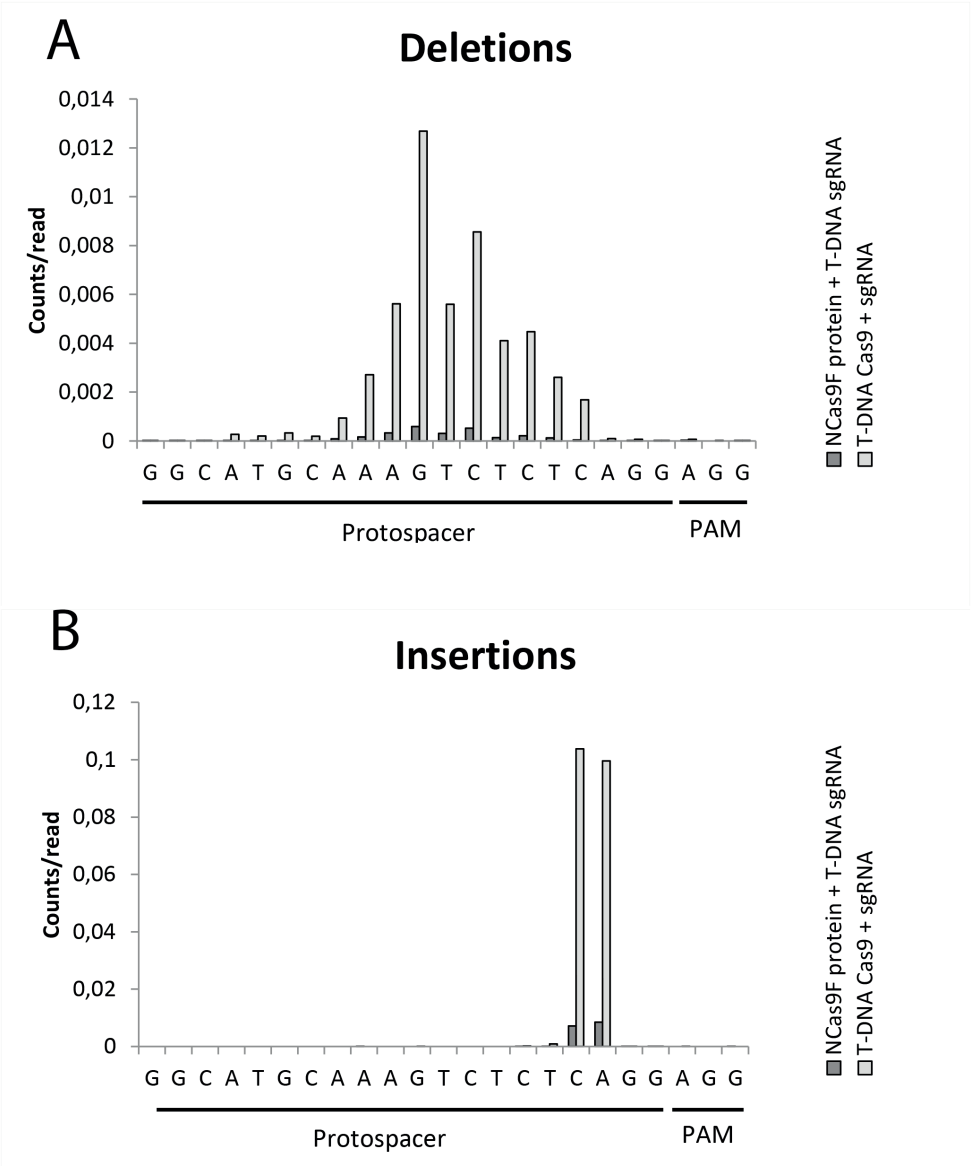

Figure 2. Deletion frequency (**A**) and insertion frequency (**B**) for each nucleotide of the *PDS* target. Deep sequencing analysis of the target sequence obtained from tissue infiltrated with an *Agrobacterium* transferring NCas9F protein combined with a T-DNA expressing the sgRNA and tissue infiltrated with *Agrobacterium* transferring a T-DNA expressing both Cas9 and sgRNA. Relative numbers of mutated reads are depicted as a fraction of the total number of reads numbers per position.

Discussion

In this study we have shown that the NCas9F protein of the type II bacterial CRISPR system can be translocated from *Agrobacterium* to *N. benthamiana* and combined with a T-DNA expressing a sgRNA induced targeted DSBs that after imperfect repair via NHEJ resulted in targeted mutations.

Mutations induced by translocated NCas9F were distributed similarly as mutations induced with Cas9 expressed from a T-DNA. The deletion and insertion frequencies were however about eighteen and twelve fold lower, respectively, with translocated NCas9F protein than with Cas9 expressed from a T-DNA. These lower frequencies could be due to limiting levels of translocated NCas9F, the short presence of NCas9F in the host after translocation, lower activity of NCas9F protein compared to Cas9 protein, or because transfer of T-DNA and NCas9F does not always occur simultaneously.

High levels of small insertions several basepair upstream of the PAM were detected in our experiments using amplicon sequencing. Using the restriction enzyme loss method, the most common mutations found were also small insertions although the frequency was slightly lower than in our experiments using amplicon sequencing. Other reports about targeted mutagenesis in *N. benthamiana* with CRISPR/Cas using the enzyme loss method to detect mutations do not mention these high levels of small insertions [15–18]. High levels of small insertions were however reported in *Arabidopsis* after DSB induction with CRISPR/Cas using next-generation sequencing [13]. This suggests that the outcome of repair of DSBs induced with CRISPR/Cas in plants results in a low frequency of small deletions and a high frequency of single bp insertions upstream of the PAM. These single bp insertions mainly consist of adenosine or thymidine insertions which is consistent with the 'A-rule' which states that polymerases are known to preferentially incorporate deoxyadenosine-monophosphate (dAMP) when template base coding is not available [19].

We attempted to obtain *N. benthamiana* shoots with targeted mutations induced by translocated NCas9F. A total of a 150 shoots were regenerated from leaf discs taken from infiltrated leaves and analyzed for targeted mutations. With a combined insertion and deletion frequency of 1.79% a few mutated shoots were expected. However we did not find a single line with targeted mutations. This suggests that either we were very unlucky (chance less than 7%) or that the frequency is in fact lower, for instance because cells with mutations mostly have both alleles mutated, which would reduce the percentage of mutated cells to 0.9%.

In summary the experimental data discussed above show that it is possible to use *Agrobacterium* for the translocation of the Cas9 protein into plant cells to create targeted mutations, but that the frequency is low. Therefore, it will be laborious to identify and regenerate the mutated cells into complete mutated plants.

Material & Methods

Plasmid construction

For the construction of NCas9F see materials & methods of Chapter 2.

To create the sgRNA expression vectors oligo's DS196/DS197 were phosphorylated, annealed and subsequently cloned into the BbsI site of pEn-Chimera [13]. A correct clone was used as entry vector for a Gateway reaction with destination vector pDe-CAS9 [13] resulting in pDECas9PDS and destination vector pMDC100 [20] resulting in psgRNAPDS.

Leaf infiltration

Seeds of *N. benthamiana* were germinated and grown in controlled climate chambers at 24°C with a 16 h light /8 h dark photoperiod with 75 % humidity for a period of three weeks before infiltration. *Agrobacterium* strain AGL1 with pNCasF and psgRNAPDS or pDECasPDS were grown overnight at 29°C shaking at 180 RPM in a 10 ml culture of LB (5 g/l NaCl) supplemented with appropriate antibiotics. The next day *Agrobacterium* cells were re-suspended in induction medium [21] to an OD600 of 1.2 and were then kept at room temperature for 3 hours without shaking. The *Agrobacterium* cells were then introduced into the two youngest leaves of 3 weeks old *N. benthamiana* plants using needleless 1 ml syringes.

Footprint detection

Leaf discs were harvested 4 days post infiltration from which genomic DNA was isolated using CTAB DNA extraction [22]. Genomic DNA (500ng) was pre-digested with DdeI (Thermo Scientific Inc.) before the target sequence was amplified using primers DS192 and DS193. Amplified products were digested again with DdeI (Thermo Scientific Inc.) after which resistant bands were cloned into pJET1.2 (Thermo Scientific Inc.). Individual clones were sent for Sanger sequencing (Macrogen Inc.).

Shoot regeneration

Infiltrated leaves were surface sterilized by immersion into 70% ethanol for 1 minute followed by immersion into 1% fresh sodium hypochlorite with 0.1% Tween 20 for 30 minutes. From these leaves discs (with a diameter of 1 cm) were excised and put on selection medium (1X Murashige and Skoog basal salt mixture, 1x Gamborg's B5 vitamins, 3% Sucrose, 0,59 g/l MES, 1.0 mg/l BAP, 0.1 mg/l NAA, 0.4% Gelrite pH 5.7, 100 µg/ml kanamycin and 320 µg/ml timentin). Leaf discs were transferred onto fresh medium every 10 days until shoots appeared. Shoots were transferred to 1/2 MS medium (100 µg/ml kanamycin and 320 µg/ml timentin).

Amplicon deep sequencing

Leaves of three weeks old *N. benthamiana* plants were infiltrated and tissue was collected 10-days post infiltration. DNA was isolated using CTAB DNA isolation; 100ng of this DNA was used for amplification with the primers listed in Table 1 with Illumina adapter sequences using Phusion® High-Fidelity DNA Polymerase. Amplicons were purified from gel using the Zymoclean™ Gel DNA Recovery Kit. Paired-end sequencing was performed by BaseClear B.V. on the Illumina HiSeq 2500 platform. Data analysis was performed using the integrative genome viewer [23,24] and the Galaxy webserver [25–27]. Deletion and insertion frequencies were determined by dividing the number of reads containing insertions or deletions by the

total number of reads. Number of reads containing insertions or deletions was determined using a custom script that sorted and counted unique reads.

Table 1. Overview of primers used in this study

Primer	Sequence
DS190 FW Nb target GGNGG	ATTGGGCATGCAAAGTCTCTCAGG
DS191 RV Nb target GGNGG	AAACCCTGAGAGACTTTGCATGCC
DS192 FW GGNGG Restric	GGTGTGCCTGATAGGGTGAC
DS193 RV GGNGG Restric	GGCAAACACAAAAGCATCTCC
DS217 Cas9-F FW 1	TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGAAAGGCCACCTTTTGACTC AATATG
DS218 Cas9-F RV 1	GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAGAAAAATCTGCAGGAGAAA CATGG
DS219 Cas9-F FW 2	TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGAACGGCCACCTTTTGACTC AATATG
DS220 Cas9-F RV 2	GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAGAACAATCTGCAGGAGAAA CATGG
DS221 Cas9-F FW 3	TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGAAGGGCCACCTTTTGACTC AATATG
DS222 Cas9-F RV 3	GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAGAAGAATCTGCAGGAGAAA CATGG
DS223 Cas9-F FW 4	TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGAATGGCCACCTTTTGACTCA ATATG
DS224 Cas9-F RV 4	GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAGAATAATCTGCAGGAGAAA CATGG
DS225 Cas9-F FW 5	TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGACAGGCCACCTTTTGACTCA ATATG
DS226 Cas9-F RV 5	GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAGACAAATCTGCAGGAGAAA CATGG
DS227 Cas9-F FW 6	TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGACCGGCCACCTTTTGACTCA ATATG
DS228 Cas9-F RV 6	GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAGACCAATCTGCAGGAGAAA CATGG
DS229 Cas9-F FW 7	TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGACGGGCCACCTTTTGACTC AATATG
DS230 Cas9-F RV 7	GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAGACGAATCTGCAGGAGAAA CATGG
DS231 Cas9-F FW 8	TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGACTGGCCACCTTTTGACTCA ATATG
DS232 Cas9-F RV 8	GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAGACTAATCTGCAGGAGAAA CATGG
DS233 Cas9-F FW 9	TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGAGAGGCCACCTTTTGACTC AATATG
DS234 Cas9-F RV 9	GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAGAGAAATCTGCAGGAGAAA CATGG
DS253 WT FW	TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGCATGGCCACCTTTTGACTCA ATATG
DS254 WT RV	GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAGCATAATCTGCAGGAGAAAC ATGG
DS235 T-DNA Cas9 FW 1	TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGAGCGGCCACCTTTTGACTCA ATATG

DS236 T-DNA Cas9 RV 1	GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAGAGCAATCTGCAGGAGAAA CATGG
DS237 T-DNA Cas9 FW 2	TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGAGGGGCCACCTTTTGACTCA ATATG
DS238 T-DNA Cas9 RV 2	GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAGAGGAATCTGCAGGAGAAA CATGG
DS239 T-DNA Cas9 FW 3	TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGAGTGGCCACCTTTTGACTCA ATATG
DS240 T-DNA Cas9 RV 3	GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAGAGTAATCTGCAGGAGAAA CATGG
DS241 T-DNA Cas9 FW 4	TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGATAGGCCACCTTTTGACTCA ATATG
DS242 T-DNA Cas9 RV 4	GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAGATAAATCTGCAGGAGAAA CATGG
DS243 T-DNA Cas9 FW 5	TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGATCGGCCACCTTTTGACTCA ATATG
DS244 T-DNA Cas9 RV 5	GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAGATCAATCTGCAGGAGAAA CATGG
DS245 T-DNA Cas9 FW 6	TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGATGGGCCACCTTTTGACTCA ATATG
DS246 T-DNA Cas9 RV 6	GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAGATGAATCTGCAGGAGAAA CATGG
DS247 T-DNA Cas9 FW 7	TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGATTGGCCACCTTTTGACTCA ATATG
DS248 T-DNA Cas9 RV 7	GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAGATTAATCTGCAGGAGAAA CATGG
DS249 T-DNA Cas9 FW 8	TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGCAAGGCCACCTTTTGACTC AATATG
DS250 T-DNA Cas9 RV 8	GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAGCAAAATCTGCAGGAGAAA CATGG
DS251 T-DNA Cas9 FW 9	TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGCACGGCCACCTTTTGACTCA ATATG
DS252 T-DNA Cas9 RV 9	GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAGCACAATCTGCAGGAGAAA CATGG

Acknowledgements

This work was supported by TTI-Green Genetics (project number TTI GG CORE 4CC057RP).

References

1. van der Oost J, Jore MM, Westra ER, Lundgren M, Brouns SJJ. CRISPR-based adaptive and heritable immunity in prokaryotes. Trends Biochem. Sci. 2009;8:401–7.

2. Jinek M, Chylinski K, Fonfara I, Hauer M, Doudna JA, Charpentier E. A programmable dual-RNA–guided DNA endonuclease in adaptice bacterial immunity. Science. 2012;337:816–22.

3. Bhaya D, Davison M, Barrangou R. CRISPR-Cas systems in bacteria and archaea: versatile small RNAs for adaptive defense and regulation. Annu. Rev. Genet. 2011;45:273–97.

4. Deveau H, Barrangou R, Garneau JE, Labonté J, Fremaux C, Boyaval P, et al. Phage response to CRISPR-encoded resistance in Streptococcus thermophilus. J. Bacteriol. 2008;190:1390–400.

5. Mojica FJM, Díez-Villaseñor C, García-Martínez J, Almendros C. Short motif sequences determine the targets of the prokaryotic CRISPR defence system. Microbiology. 2009;155:733–40.

6. Christie PJ, Cascales E. Structural and dynamic properties of bacterial type IV secretion systems (review). Mol. Membr. Biol. 2005;22:51–61.

7. Alvarez-Martinez CE, Christie PJ. Biological diversity of prokaryotic type IV secretion systems. Microbiol. Mol. Biol. Rev. 2009;73:775–808.

8. Vergunst AC, Schrammeijer B, den Dulk-Ras A, Vlaam de CMT, Regensburg-Tuïnk TJ, Hooykaas PJJ. VirB/D4-Dependent Protein Translocation from Agrobacterium into Plant Cells. Science. 2000;290:979–82.

9. Vergunst AC, van Lier MCM, den Dulk-Ras A, Stüve T a G, Ouwehand A, Hooykaas PJJ. Positive charge is an important feature of the C-terminal transport signal of the VirB/D4-translocated proteins of Agrobacterium. Proc. Natl. Acad. Sci. USA. 2005;102:832–7.

10. Rolloos M, Hooykaas PJJ, van der Zaal BJ. Enhanced targeted integration mediated by translocated I-SceI during the Agrobacterium mediated transformation of yeast. Sci. Rep. 2015;5:8345.

11. Schrammeijer B, den Dulk-Ras A, Vergunst A, Jurado Jácome E, Hooykaas PJJ. Analysis of Vir protein translocation from Agrobacterium tumefaciens using Saccharomyces cerevisiae as a model: evidence for transport of a novel effector protein VirE3. Nucleic Acids Res. 2003;31:860–8.

12. Farboud B, Meyer BJ. Dramatic enhancement of genome editing by CRISPR/cas9 through improved guide RNA design. Genetics. 2015;199:959–71.

13. Fauser F, Schiml S, Puchta H. Both CRISPR/Cas-based nucleases and nickases can be used efficiently for genome engineering in Arabidopsis thaliana. Plant J. 2014;79:348–59.

14. Voytas DF. Plant genome engineering with sequence-specific nucleases. Annu. Rev. Plant Biol. 2013;64:327–50.

15. Lowder LG, Zhang D, Baltes NJ, Paul JW, Tang X, Zheng X, et al. A CRISPR/Cas9 Toolbox for Multiplexed Plant Genome Editing and Transcriptional Regulation. Plant Physiol. 2015;169:971–85.

16. Ali Z, Abul-Faraj A, Li L, Ghosh N, Piatek M, Mahjoub A, et al. Efficient virus-mediated genome editing in plants using the CRISPR/Cas9 system. Mol. Plant. 2015;8:1288–91.

17. Yin K, Han T, Liu G, Chen T, Wang Y, Yu AYL, et al. A geminivirus-based guide RNA delivery system for CRISPR/Cas9 mediated plant genome editing. Sci. Rep. 2015;5:14926.

18. Nekrasov V, Stakawicz B, Weigel D, Gones JDG, Kamoun S. Targeted mutagenesis in the model plant nicotiana benthamiana using Cas9 RNA-guided endonuclease. Nat. Biotechnol. 2013;31:688–91.

19. Strauss BS. The "A" rule revisited: Polymerases as determinants of mutational specificity. DNA Repair (Amst). 2002;1:125–35.

20. Curtis MD, Grossniklaus U. A Gateway cloning vector set for high-throughput functional analysis of genes in planta. Breakthr. Technol. 2003;133:462–9.

21. Bundock P, den Dulk-Ras A, Beijersbergen A, Hooykaas PJ. Trans-kingdom T-DNA transfer from Agrobacterium tumefaciens to Saccharomyces cerevisiae. EMBO J. 1995;14:3206–14.

22. de Pater S, Neuteboom LW, Pinas JE, Hooykaas PJJ, van der Zaal BJ. ZFN-induced mutagenesis and gene-targeting in Arabidopsis through Agrobacterium-mediated floral dip transformation. Plant Biotechnol. J. 2009;7:821–35.

23. Thorvaldsdóttir H, Robinson JT, Mesirov JP. Integrative Genomics Viewer (IGV): High-performance genomics data visualization and exploration. Brief. Bioinform. 2013;14:178–92.

24. Robinson JT, Thorvaldsdóttir H, Winckler W, Guttman M, Lander ES, Getz G, et al. Integrative genomics viewer. Nat. Biotechnol. 2011;29:24–6.

25. Goecks J, Nekrutenko A, Taylor J. Galaxy: a comprehensive approach for supporting accessible, reproducible, and transparent computational research in the life sciences. Genome Biol. 2010;11:R86.

26. Blankenberg D, Kuster G Von, Coraor N, Ananda G, Lazarus R, Mangan M, et al. Galaxy: A web-based genome analysis tool for experimentalists. Curr. Protoc. Mol. Biol. 2010;1–21.

27. Giardine B, Riemer C, Hardison RC, Burhans R, Elnitski L, Shah P, et al. Galaxy: A platform for interactive large-scale genome analysis. Genome Res. 2005;15:1451–5.

Chapter 5

Transient expression of the isopentenyl transferase for (non)transgenic shoot induction in *Arabidopsis thaliana*

Daan J. Schmitz, Amke den Dulk-Ras, Sylvia de Pater, Paul J.J. Hooykaas

Abstract

Plant transformation systems use a selectable marker gene which is co-delivered with the gene of interest for efficient selection of transformation events among the large numbers of non-transformed plants cells. Throughout the years several marker genes have been developed usually based on conditional dominant genes many of which are antibiotic resistance genes. An alternative non antibiotic marker gene is the isopentenyl transferase gene (*ipt*), found on the Ti-plasmid of *Agrobacterium tumefaciens*, which increases cytokinine levels stimulating organogenesis in many cultured plant tissues and which is widely used to regenerate transgenic plants from cultured cells after transformation. Constitutive expression of *ipt* however results in loss of apical dominance and an inability to form roots and therefore its removal after selection is essential to produce normal plants. Instead of integrating and subsequently removing the *ipt* gene we have tested whether the transient expression of IPT can be used for the selection of transformed plants. The first approach consisted of the delivery of a T-DNA encoding the *ipt* gene into the Pol-θ-deficient *Arabidopsis* integration mutant in which only transient expression of the T-DNA occurs but no integration. The second approach involved the direct delivery of the IPT protein through the *Agrobacterium* VirB/D4 T4SS into *Arabidopsis*. We show that the combined transfer of the IPT protein with a T-DNA encoding a CRISPR/Cas system can be used to obtain mutated shoots. Furthermore if the transfer of a T-DNA is combined with the transfer of the IPT protein, T-DNA transformants can be identified based on shoot induction without requiring selection for the T-DNA.

Introduction

For efficient selection of transformation events among the large numbers of non-transformed plants cells, plant transformation systems use a selectable marker gene which is co-delivered with the gene of interest. Throughout the years several marker genes have been developed usually based on conditional dominant genes. By selection for these genes transgenic plants can be obtained eventually. These selection systems do however have several shortcomings: (1) the presence of marker genes prevents usage of the same marker in a next round of transformation; (2) integration of the marker genes limits the usage of these plants due to regulatory concerns; (3) concerns have been raised specifically on the release of antibiotic resistance genes.

An alternative non antibiotic selection gene which been used successfully is the isopentenyl transferase gene (*ipt*) found on the Ti-plasmid of *Agrobacterium tumefaciens*. This isopentenyl transferase catalyzes the condensation of isopentenyl pyrophosphate, a precursor of several cytokinines [1,2]. Increased levels of these cytokinines have been shown to induce cell proliferation and shoot formation in several plants species [3–6]. The *ipt* gene derived from *Agrobacterium* has been used as a visible marker for identifying transgenic plants, which are bushy due to enhanced cytokinin levels. Unfortunately plants expressing *ipt* lose apical dominance and are unable to form roots and therefore removal of the *ipt* gene is required to obtain normal plants. For the removal of the *ipt* gene from transgenic cells two different approaches have been developed based on site specific recombination [7–13] or on transposition by the maize transposable element Ac [14]. Naturally the *ipt* gene is introduced into plant cells by *Agrobacterium* as part of the T-DNA. This bacterium has a type IV secretion system (T4SS), encoded by the *virB* genes and *virD4* gene on its Ti plasmid, through which translocation of the T-DNA occurs [15,16]. Several virulence protein

are transported independently alongside the T-DNA into the host cell [17]. Recognition and translocation of proteins through the *Agrobacterium* VirB/D4 T4SS is dependent on a hydrophilic secretion signal with a net positive charge in the C-terminal part of the proteins [18]. Several heterologous proteins, fused to this secretion signal, have been translocated through the *Agrobacterium* VirB/D4 T4SS [17–20]. In this way the homing endonuclease I-SceI and the Cre recombinase were translocated into host cells to effect DNA recombination in the genome of target cells [17–19].

In this study we have developed two different methods for the transient expression of IPT that can be used for the selection of transformed plants. The first consists of delivery of a T-DNA encoding the *ipt* gene in the Pol-θ-deficient *Arabidopsis* mutant in which only transient expression of the T-DNA occurs but no integration [21]. The second involves the direct delivery of the IPT protein into *Arabidopsis* through the *Agrobacterium* VirB/D4 T4SS.

Results

Shoot formation via transient expression of the ipt gene in Pol-θ-deficient Arabidoposis

To test if *Arabidopsis* shoots can be recovered after the transient expression of the *ipt* gene we created a T-DNA vector containing the *ipt* gene from the *Agrobacterium* octopine Ti plasmid (pSDM3679). This T-DNA was introduced into an *Agrobacterium* strain already containing a T-DNA vector that provides resistance to the herbicide phosphinothricin (PPT), but also encoded a CRISPR/Cas system targeting the protoporphyrinogen oxidase (*PPO*) locus (pSDM3905). Both of these T-DNA vectors could stably replicate together in *Agrobacterium* as their replication units were compatible: incP for pSDM3679 and pVS1 for pSDM3905. The resulting strain LBA3718 was used to transform the roots of wild type *Arabidopsis* and the roots of the T-DNA integration resistant Pol-θ-deficient mutant (*teb-5*). This second T-DNA, pSDM3905, was added to test if the transient expression of a CRISPR/Cas system was effective in inducing targeted mutations (discussed in the next paragraph). Although no T-DNA integration occurs in *teb-5* roots, genes on a transferred T-DNA are still expressed transiently [21].

After cocultivation for 3 days *Arabidopsis* wild type roots were transferred to hormone free medium containing PPT to select for the presence of the T-DNAs from pSDM3905. After three weeks such roots developed dark green callus tissue. Shoot formation was observed after six weeks with ~6.5% (47/720) of the calli (Fig. 1d). As this was not seen after cocultivation with strains lacking the *ipt* gene, this probably reflects the temporary or ongoing transient expression of the *ipt* gene or continuing activity of the encoded IPT protein. Cocultivated roots placed on shoot induction medium developed dark callus tissue with shoots on 14.6% of the calli (Fig. 1f). Similar cocultivations were done with the *teb-5* mutant. On hormone free medium containing PPT, dark green calli were formed from which shoots appeared after 6 weeks. Shoot formation was observed with ~1.2% (10/864) of the calli, a ~fivefold reduction compared to wild type roots. With *teb-5* roots dark green callus tissue always developed shoot tissue whereas with wild type roots dark green callus tissue did not always develop shoot tissue (Fig. 1a; Fig. 1d). Five of the ten shoots obtained after transformation of *teb-5* roots that were transferred to hormone free medium exhibited normal growth and root formation suggesting that they were not stably transformed with the *ipt* gene. This suggests that transient expression of the *ipt* gene from a T-DNA in *teb-5* roots is effective in inducing shoot formation. Neither non transformed *teb-5* roots nor wild type roots developed dark green calli and shoots (Fig. 1b; Fig 1e).

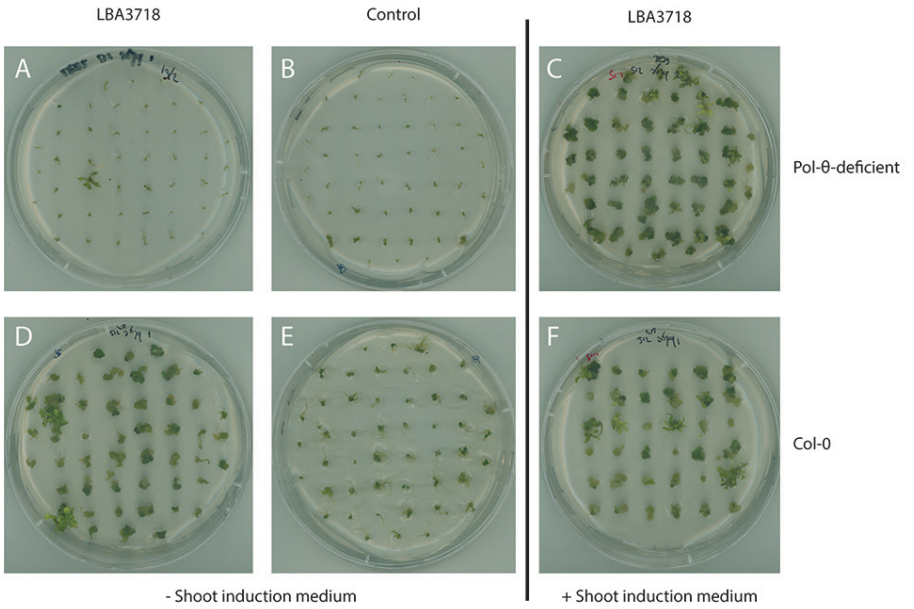

Figure 1. Shoot formation after cocultivation of wild-type (**D**, **F**) and *teb-5* roots (**A**,**C**) with *Agrobacterium* strain LBA3718 that transfers a T-DNA encoding the *ipt* gene (pSDM3679) and a T-DNA encoding a CRISPR/Cas system targeting the *PPO* locus (pSDM3905). (**A**,**D**) Roots cocultivated with LBA3718 placed on hormone free medium containing PPT. (**B**,**E**) Non cocultivated roots on hormone free medium containing PPT. (**C**,**F**) Roots cocultivated with LBA3718 placed on shoot induction medium containing PPT.

Targeted mutagenesis in shoots recovered after transient expression of the ipt gene and CRISPR/Cas

In the previous experiments we were able to recover several shoots on selection medium after the cocultivation of *teb-5* roots with LBA3718. To test if (transient) presence of the T-DNA expressing CRISPR/Cas9 had resulted in targeted mutations the recovered shoots were analyzed for the presence of footprints at the *PPO* locus. A 950 basepair (bp) fragment was amplified by PCR from the *PPO* locus with primers flanking the target sequence from genomic DNA isolated from a single leaf. This PCR product was digested with FauI, as the target site of the sgRNA overlaps with a FauI site, and restriction digestion resistant bands were cloned and analyzed by sequencing. Three shoots (3/10) that were analyzed showed FauI resistant PCR products (Fig. 2a). Sequencing of the PCR products showed that mutations were present consisting of small deletions several bp upstream of the PAM (Fig. 2c). Several of the targeted mutations we detected after the co-transfer of the T-DNA expressing *ipt* and the T-DNA expressing the CRISPR/Cas system using the restriction enzyme site loss method [22] did not contain a mutated FauI site and therefore are probably the result of incomplete digestion of the PCR product. The three shoots with targeted mutations (2, 6, and 10) showed a bushy phenotype consistent with constitutive expression of the *ipt* gene suggesting that the T-DNA encoding *ipt* might still be present (Fig. 2d). Therefore all shoots were analyzed for the presence of both T-DNA's using PCR. The pSDM3905 T-DNA was still detected in shoot number 2, 5, 6 and 10 and the T-DNA encoding the *ipt* gene (pSDM3679) was detected in shoot number 2, 6 and 10 (Fig. 2b).

These results indicated that non-integrated T-DNA can remain present for a long period of time or that T-DNA's can integrate by a process independent of Pol θ in a low percentage of the cells.

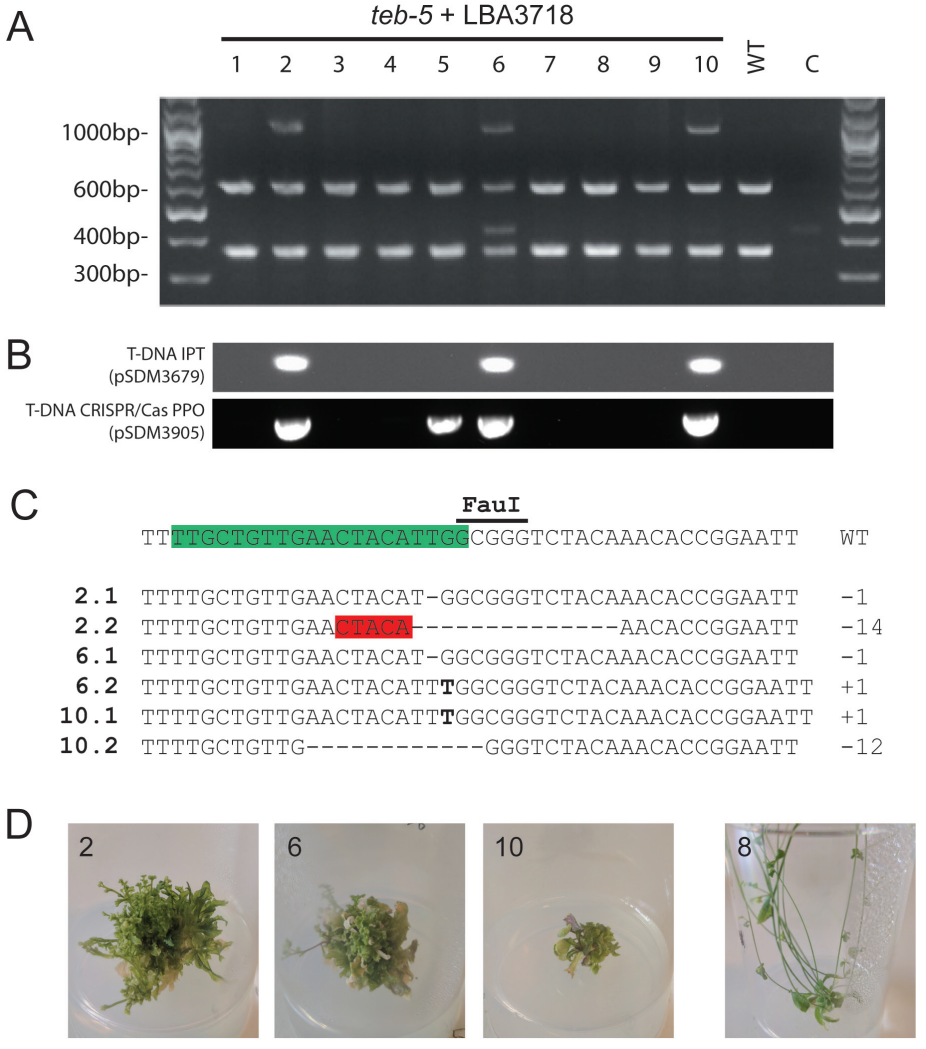

Figure 2. CRISPR/Cas induced mutagenesis in *teb-5* roots. (**A**) The *PPO* target site was amplified using genomic DNA of 10 *teb-5* shoots transformed by LBA3718 and a wild type leaf and the resulting PCR products were digested with FauI and separated on agarose gel. A control PCR sample without template (**C**) was included. (**B**) Detection of the presence of the T-DNA from pSDM3679 and pSDM3905 via PCR. (**C**) Sequence analysis of mutations in shoot number 2, 6 and 10. The sgRNA is in green, the restriction site is underlined, deletions are shown by dashes, insertions in bold and microhomology in red. Numbers on the right are length of deletions (-) and insertions (+). (**D**) Shoot number 2, 6, and 10 showing a bushy phenotype. Shoot number 8 is an example of a shoot showing a normal phenotype.

Shoot induction after translocation of the IPT protein

The previous experiments suggested that transient expression of *ipt* from a non-integrated T-DNA is effective in inducing shoot formation in the roots of *teb-5*. To eliminate the requirement for this mutant line we tested if the IPT protein can be translocated through the *Agrobacterium* VirB/D4 type IV secretion system and is capable of inducing shoot formation after translocation.

For the translocation of the IPT protein an expression plasmid was created encoding the isopentenyl transferase fused to the C-terminal 37 amino acid translocation signal of the *Agrobacterium* virulence protein VirF. This translocation signal has previously been used for the translocation of several heterologous proteins [17–19]. The production of the fusion protein (IPTF) was under the control of the acetosyringone inducible *virF* promoter to ensure that production would occur concomitantly with formation of a functional VirB/D4 type IV channel.

To assay for the translocation of the IPTF protein roots had been co-cultivated with *Agrobacterium* expressing the IPTF protein (LBA3720) after which the root segments were placed on hormone free medium. After six weeks shoot formation was observed on the plates with root fragments that had been co-cultivated with *Agrobacterium* strain expressing the IPTF protein (Fig. 3a; Fig. 3b). On plates with root fragments that were co-cultivated with an *Agrobacterium* strain not expressing IPTF no shoot induction was observed (Fig. 3c).

Because translocated IPTF successfully initiated shoot formation we tested if shoot formation could be used to visually identify T-DNA transformants if the translocation of a T-DNA is combined with the transfer of IPTF. A binary vector (pBIN19) was introduced into the *Agrobacterium* expressing the IPTF protein. The resulting strain (LBA3721) was used to transform the roots of wild type *Arabidopsis*. After co-cultivation for three days roots were transferred to hormone free medium. After five weeks shoot formation was observed on several root fragments. These shoots were analyzed for the presence of the *nptII* gene present on the T-DNA by PCR. Shoot number 1, 3 and 5 contained the T-DNA showing that T-DNA transformants can be selected for using shoot formation induced by transferred IPTF as a visual selection marker (Fig. 3d).

These results show that the IPTF protein is effectively translocated to *Arabidopsis* roots at sufficient levels to induce shoot formation in *Arabidopsis* root fragments and this induction of shoot formation can be used to identify T-DNA transformants.

Combined transfer of the IPT protein and a T-DNA encoding the CRISPR/Cas system

As the previous experiments showed that translocated IPTF is effective in inducing shoot formation we added the binary vector encoding the CRISPR/Cas system targeting the *PPO* locus (pSDM3905) to the *Agrobacterium* strain expressing the IPTF protein resulting in *Agrobacterium* strain LBA3719. Roots were co-cultivated with LBA3719 for three days after which roots were placed on hormone free medium containing PPT to select for the T-DNA encoding the CRISPR/Cas system targeting the *PPO* locus. Shoot formation was observed on one root fragment (1/1392). DNA was isolated from three individual leaves of this shoot. To easily detect sgRNA-guided mutations induced by Cas9 we used the same restriction enzyme loss method described above. The target locus was amplified via PCR with primers flanking the target sequence using three different parts of the shoot and the resulting PCR products were digested with FauI. Resistant bands were observed in all three samples (Fig. 4a).

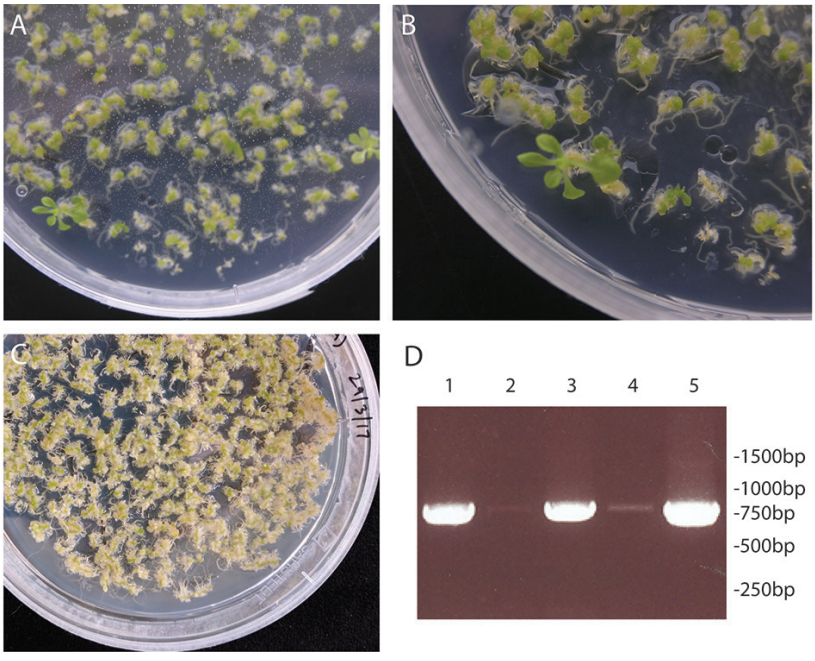

Figure 3. Shoot regeneration after translocation of the IPTF protein. (**A**) Roots after co-cultivation with *Agrobacterium* expressing IPTF. (**B**) Close up of shoot after co-cultivation of roots with *Agrobacterium* expressing IPTF (**C**). Non co-cultivated roots on hormone free medium (**D**). PCR on the *nptII* locus in shoots that were recovered after co-cultivation with *Agrobacterium* containing pBIN19 and expressing IPTF.

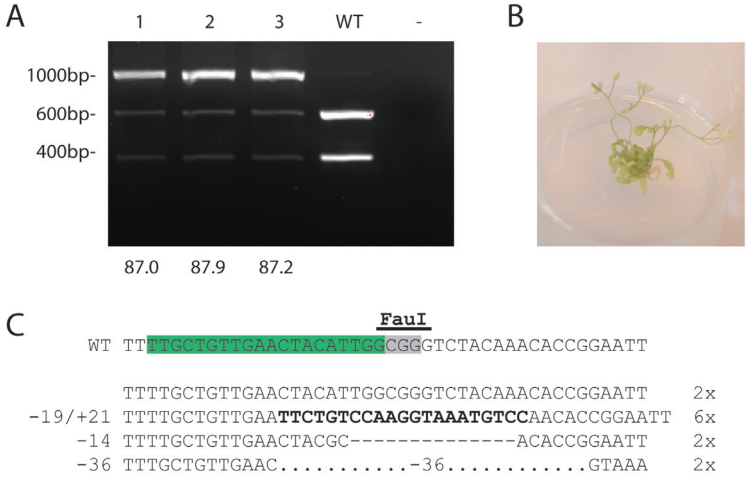

Figure 4. CRISPR/Cas induced mutagenesis in a wild type shoot after co-cultivation with LBA3719. (**A**) The *PPO* target site was amplified using genomic DNA from three individual leaves from a shoot recovered after co-transfer of IPTF and the T-DNA from pSDM3905 (1-3) and a wild type leaf (WT). A control sample without template (-) was included. The resulting PCR products were digested with FauI and separated on an agarose gel. (**B**) Shoot regenerated after co-transfer of IPTF and the T-DNA from pSDM3905. (**C**) Sequence analysis of mutations detected. The sgRNA is in green, restriction site is underlined, deletions are shown by dashes, insertions in red and templated insertion is in bold. Numbers on the left are length of deletions (-) and insertions (+). Numbers of multiple clones with the same sequence are shown on the right.

Using the relative band intensities the number of mutations was estimated at around ~87% in each of the three samples (Fig. 4a). The resistant PCR products were cloned into a high-copy vector, transformed to *E. coli* and individual clones were sequenced. Analysis of these sequences showed that the plant contains three different kinds of mutations; a 16 bp deletion, a 36 bp deletion and a templated insertion (Fig. 4c). Because of the high mutation frequency in the essential *PPO* gene stunted growth was observed in the recovered plant (Fig. 4a).

These results combined show that the translocation of the IPTF protein combined with the transfer of a T-DNA encoding a CRISPR/Cas system resulted in a shoot with targeted mutations.

Discussion

In this study we have shown that the *ipt* gene can be transiently expressed from a T-DNA in the Pol- θ-deficient *teb-5* mutant to induce shoot formation. Furthermore we showed that the IPT protein can be translocated through the *Agrobacterium* VirB/D4 T4SS and is effective in inducing shoot formation. If the translocation of the IPTF protein is combined with the transfer of a T-DNA encoding a CRISPR/Cas system it is possible to recover plants with mutations. We also showed that T-DNA transformants could be identified using shoot formation induced by translocated IPTF as a visual identification method.

The T-DNA encoding the *ipt* gene and the T-DNA encoding the CRISPR/Cas system which were used to transform *teb-5* roots could still be detected via PCR in 40% (4/10) and 30% (3/10) of the obtained shoots, respectively. This suggests that non-integrated T-DNA persists in the plant cells for a prolonged period up to five weeks. We can however not exclude that the T-DNA was still detected due to incomplete removal of all *Agrobacterium* or that T-DNA integration still occurs in *teb-5* plants via an alternative integration pathway.
When selection of shoot formation by IPTF protein transfer was done after cocultivation with an *Agrobacterium* strain containing a T-DNA encoding a CRISPR/Cas system, a shoot was obtained with a high frequency of targeted mutations that were evenly distributed throughout the plant. This suggests that these mutations occurred early in the development of the shoot. If the frequency of shoot formation after co-transfer can be improved, it will allow for the easy recovery of plants with a high frequency of targeted mutations.

In summary we developed two systems for which we employ shoot regeneration by IPT activity to identify transformants, which grow with a normal phenotype, because they do not contain an integrated *ipt* gene as in previous methods.

Material & Methods

Agrobacterium strains and media

Agrobacterium strains and plasmids used in this study are listed in Table 1 and Table 2, respectively. All *Agrobacterium* strain were grown in LB (5 g/l NaCl) with the appropriate antibiotics: gentamicin (40 μg/ml); carbenicillin (75 μg/ml); kanamycin (100 μg/ml); spectinomycin (250 μg/ml). Plasmids were electroporated into AGL1 as described in den Dulk-Ras and Hooykaas (1995).

Plasmid construction

To create the expression plasmid for expression of the IPT protein fused to the 37 last amino acids of the *Agrobacterium* virulence protein VirF (IPTF) in *Agrobacterium*, the *ipt* gene was amplified by PCR from LBA1 with primers IPT1/IPT2 and was inserted into the EcoRV and SalI sites of pSDM3190. From this modified pSDM3190 vector a 1900bp HindIII/XbaI fragment was cut and inserted into the HindIII and XbaI sites of pBBR6, creating pSDM3678. For construction of the binary vector for the expression of *ipt* in plant cells, the *ipt* gene was amplified by PCR from LBA1 with primers IPT3 and IPT4 and inserted into the XbaI and XhoI sites of pART7-YFP-HAII. The NotI fragment (p35S::ipt::t35S) from this vector was cut and inserted into the NotI site of pBluescript creating pBSK-p35S-IPT-t35S. From pBSK-p35S-IPT-t35S a HindIII and SacI fragment (p35S::ipt::t35S) was cut and inserted into the HindIII and SacI sites of pBIN19 creating pSDM3679.

Protein translocation and plant transformation experiments

Root transformations were performed as described previously [17,24,25], using *Agrobacterium* strain AGL1. Briefly, seedlings from wild type *Arabidopsis* (ecotype Col-0) and the *teb-5* mutant (Pol-θ-deficient line, ecotype Col-0, [26]) were grown for 10 days. Roots were removed from seedlings and precultured on callus induction medium [25], followed by a three day co-cultivation period with *Agrobacterium*. After co-cultivation roots were transferred to B5 medium [27] containing vancomycin (100 μg/ml) and timentin (100 μg/ml) to kill remaining *Agrobacterium*. The selection for the T-DNA of pSDM3905 was done by adding PPT (30 μg/ml) to the medium. After three weeks calli were transferred to medium without PPT.

Detection of the T-DNAs

The presence of pSDM3905 and pSDM3679 T-DNAs was performed with PCR using primers pair SP558/SP559 (detecting Cas9) and DS585/DS589 (detecting *ipt*), respectively. The presence of the pBIN19 T-DNA was detected using primers nosNPTII1/nosNPTII2 (detecting *nptII*).

Detection of mutations

DNA was isolated from a single leaf using CTAB DNA extraction [28]. The target sequence was amplified using primers SP392 and SP538. Amplified products were digested with FauI (New England Biolabs) after which resistant bands were cloned into pJET1.2 (Thermo Scientific Inc.). Individual clones were sent for Sanger sequencing (Macrogen Inc.).

Table 1. Overview of plasmids used in this study

Plasmid	Marker	Origin of replication	Properties	Source
pART7-YFP-HAII	Sp	ColE1, incP	Modified pART7 backbone	Galvan Ampudia (unpublished)
pBBR6	Gm	pBBR	Derivative of the broad host-range plasmid pRL662	[17]
pBIN19	Km	incP	Plant binary vector	[29]
pBSK-p35S-IPT-t35S	Cb	ColE1	pBluescript containing 35S promoter, ipt and 35S terminator	This study
pSDM3190	Cb	ColE1	pUC21 pvirFpromoter:NLS::Cre::VirFΔ42N	den Dulk (unpublished)
pSDM3678	Gm	pBBR	IPTF expression vector0	This study
pSDM3679	Km	incP	Binary vector for the expression of the ipt gene	This study
pSDM3905	Sp	ColE1, pVS1	T-DNA vector with CRISPR/Cas system targeting the PPO locus	[30]

Sp = spectinomycin, Gm = gentamicin, Cb = carbenicillin, Km = kanamycin

Table 2. Overview of strains used in this study

Strain	Genomic background	Plasmids	Source
AGL1	AGL1	-	[31]
LBA3718	AGL1	pSDM3679 + pSDM3905	This study
LBA3719	AGL1	pSDM3578 + pSDM3905	This study
LBA3720	AGL1	pSDM3679 + pBIN19	This study
LBA3721	AGL1	pSDM3679	This study

Table 3. Overview of primers used in this study

Primer name	Sequence
IPT1	agcgatatcATGGACCTGCATCTAATTTTCG
IPT2	acggtcgactATACATTCCGAACGGATGAC
IPT3	ccgctcgagCAGTTTGTATTCAATATACTGC
IPT4	gctctagaATACATTCCGAACGGATGAC
SP392	CACTTTGACAGATTAGGTAG
SP538	CTAAGGCTACACCAGCGACG
DS558	GGAACTAACTCTGTGGGATG
DS559	CACACCTGAAGCGTTGATAG
NOSNPTII1	AAGCCTGAACTCACCGCGAC
NOSNPTII2	CCGGCACCAACCGAACATAC

Acknowledgements

This work was supported by TTI-Green Genetics (project number TTI GG CORE 4CC057RP).

References

1. Akiyoshi DE, Klee H, Amasino RM, Nester EW, Gordon MP. T-DNA of Agrobacterium tumefaciens encodes an enzyme of cytokinin biosynthesis. Proc. Natl. Acad. Sci. 1984;81:5994–8.

2. Barry GF, Rogers SG, Fraley RT, Brand L. Identification of a cloned cytokinin biosynthetic gene. Proc. Natl. Acad. Sci. 1984;81:4776–80.

3. Ooms G, Karp A, Roberts J. From tumour to tuber; tumour cell characteristics and chromosome numbers of crown gall-derived tetraploid potato plants (Solanum tuberosum cv. "Maris Bard"). Theor. Appl. Genet. 1983;66:169–72.

4. Smigocki a C, Owens LD. Cytokinin-to-auxin ratios and morphology of shoots and tissues transformed by a chimeric isopentenyl transferase gene. Plant Physiol. 1989;91:808–11.

5. Smigocki AC, Owens LD. Cytokinin gene fused with a strong promoter enhances shoot organogenesis and zeatin levels in transformed plant cells. Proc. Natl. Acad. Sci. 1988;85:5131–5.

6. Medford JI, Horgan R, El-Sawi Z, Klee HJ. Alterations of endogenous cytokinins in transgenic plants using a chimeric isopentenyl transferase gene. Plant Cell. 1989;1:403–13.

7. Ebinuma H, Sugita K, Endo S, Matsunaga E, Yamada K. Elimination of marker genes from transgenic plants using MAT vector systems. In: Peña L, editor. Transgenic Plants Methods Protoc. Totowa, NJ: Humana Press; 2004. p. 237–53.

8. López-Noguera S, Petri C, Burgos L. Combining a regeneration-promoting ipt gene and site-specific recombination allows a more efficient apricot transformation and the elimination of marker genes. Plant Cell Rep. 2009;28:1781–90.

9. Scaramelli L, Balestrazzi A, Bonadei M, Piano E, Carbonera D, Confalonieri M. Production of transgenic barrel medic (Medicago truncatula Gaernt.) using the ipt-type MAT vector system and impairment of recombinase-mediated excision events. Plant Cell Rep. 2009;28:197–211.

10. Khan RS, Nakamura I, Mii M. Production and selection of marker-free transgenic plants of petunia hybrida using site-specific recombination. Biol. Plant. 2010;54:265–71.

11. Khan RS, Ntui VO, Chin DP, Nakamura I, Mii M. Production of marker-free disease-resistant potato using isopentenyl transferase gene as a positive selection marker. Plant Cell Rep. 2011;30:587–97.

12. Zou X, Peng A, Xu L, Liu X, Lei T, Yao L, et al. Efficient auto-excision of a selectable marker gene from transgenic citrus by combining the Cre/loxP system and ipt selection. Plant Cell Rep. 2013;32:1601–13.

13. Zheng Y, Pan Y, Li J, Zhou Y, Pan Y, Ding Y, et al. Visible marker excision via heat-inducible Cre/LoxP system and ipt selection in tobacco. Vitr. Cell. Dev. Biol. - Plant. In Vitro Cellular & Developmental Biology - Plant; 2016;52:492–9.

14. Ebinuma H, Sugita K, Matsunaga E, Yamakado M. Selection of marker-free transgenic plants using the isopentenyl transferase gene. Proc. Natl. Acad. Sci. 1997;94:2117–21.

15. Christie PJ, Cascales E. Structural and dynamic properties of bacterial type IV secretion systems (review). Mol. Membr. Biol. 2005;22:51–61.

16. Alvarez-Martinez CE, Christie PJ. Biological diversity of prokaryotic type IV secretion systems. Microbiol. Mol. Biol. Rev. 2009;73:775–808.

17. Vergunst AC, Schrammeijer B, den Dulk-Ras A, Vlaam de CMT, Regensburg-Tuïnk TJ, Hooykaas PJJ. VirB/D4-dependent protein translocation from Agrobacterium into plant cells. Science. 2000;290:979–82.

18. Vergunst AC, van Lier MCM, den Dulk-Ras A, Stüve T a G, Ouwehand A, Hooykaas PJJ. Positive charge is an important feature of the C-terminal transport signal of the VirB/D4-translocated proteins of Agrobacterium. Proc. Natl. Acad. Sci. 2005;102:832–7.

19. Rolloos M, Hooykaas PJJ, van der Zaal BJ. Enhanced targeted integration mediated by translocated I-SceI during the Agrobacterium mediated transformation of yeast. Sci. Rep. 2015;5:8345.

20. Schrammeijer B, den Dulk-Ras A, Vergunst A, Jurado Jácome E, Hooykaas PJJ. Analysis of Vir protein translocation from Agrobacterium tumefaciens using Saccharomyces cerevisiae as a model: evidence for transport of a novel effector protein VirE3. Nucleic Acids Res. 2003;31:860–8.

21. van Kregten M, de Pater S, Romeijn R, van Schendel R, Hooykaas PJJ, Tijsterman M. T-DNA integration in plants results from polymerase-θ-mediated DNA repair. Nat. Plants. 2016;2:16164.

22. Voytas DF. Plant genome engineering with sequence-specific nucleases. Annu. Rev. Plant Biol. 2013;64:327–50.

23. den Dulk-Ras A, Hooykaas PJJ. Electroporation of Agrobacterium Tumefaciens. Methods Mol. Biol. 1995. p. 63–72.

24. Vergunst AC, Lier MCM, Dulk-ras A, Hooykaas PJJ. Recognition of the Agrobacterium tumefaciens VirE2 translocation signal by the VirB/D4 transport system does not require VirE1. Plant Physiol. 2003;133:978–88.

25. Vergunst AC, de Waal EC, Hooykaas PJJ. Root Transformation by Agrobacterium tumefaciens. In: Martinez-Zapater JM, Salinas J, editors. Arab. Protoc. Totowa, NJ: Humana Press; 1998. p. 227–44.

26. Inagaki S, Suzuki T, Ohto M, Urawa H, Horiuchi T, Nakamura K, et al. Arabidopsis TEBICHI , with helicase and DNA polymerase domains , is required for regulated cell division and differentiation in meristems. The Plant Cell. 2006;18:879–92.

27. Gamborg OL, Miller RA, Ojima K. Nutrient requirements of suspension cultures of soybean root cells. Exp. Cell Res. 1968;50:151–8.

28. de Pater S, Neuteboom LW, Pinas JE, Hooykaas PJJ, van der Zaal BJ. ZFN-induced mutagenesis and gene-targeting in Arabidopsis through Agrobacterium-mediated floral dip transformation. Plant Biotechnol. J. 2009;7:821–35.

29. Bevan M. Binary Agrobacterium vectors for plant transformation. Nucleic Acids Res. 1984;12:8711–21.

30. Shen H, Strunks GD, Klemann BJPM, Hooykaas PJJ, de Pater S. CRISPR/Cas9-induced double-strand break repair in Arabidopsis nonhomologous end-joining mutants. G3 Genes|Genomes|Genetics. 2017;7:193–202.

31. Lazo GR, Stein PA, Ludwig RA. A DNA Transformation-Competent Arabidopsis Genomic Library in Agrobacterium. Nat Biotech. 1991;9:963–7.

Chapter 6

Nederlandse samenvatting

CRISPR/Cas-geïnduceerde gerichte mutagenese met *Agrobacterium* gemedieerde eiwitoverdracht

Het kunnen aanbrengen van mutaties in het genoom van planten is om (in ieder geval) twee redenen relevant. In de eerste plaats vormt het de basis van veel fundamenteel plantonderzoek. Daarnaast wordt het gebruikt voor het creëren van nieuwe plantenrassen die (bijvoorbeeld) resistent zijn tegen insecten of herbiciden. Vroeger werden er een technieken gebruikt die resulteerden in mutaties op willekeurige plekken in het genoom. Daardoor kostte het veel moeite om planten te identificeren met de juiste mutatie. Dat probleem doet zich niet voor bij gericht aangebrachte mutaties. Om die reden is het aanbrengen van gerichte mutaties voor zowel het toegepast plantonderzoek als voor het creëren van nieuwe plantenrassen van grote toegevoegde waarde.

Er zijn verschillende manieren waarop gerichte mutaties kunnen worden aangebracht. Eén van die manieren is door middel van het aanbrengen van dubbelstrengsbreuken (DSB). Een DSB in het DNA kan door twee verschillende DNA-reparatietrajecten gereperareerd worden, te weten het homologe recombinatietraject en het Niet-Homologe End-Joining traject (NHEJ). Het homologe recombinatietraject kan gebruik maken van het zusterchromatide als sjabloon voor nauwkeurige reparatie, maar kan ook gebruik maken van een artificieel geïntroduceerd reparatiesjabloon waarmee genen kunnen worden vervangen, toegevoegd of aangepast. Het NHEJ maakt geen gebruik van een sjabloon, maar stimuleert de directe ligatie van de DSB-uiteinden, wat kan resulteren in kleine inserties en deleties op de plaats van herstel, zogenaamde mutaties.

Voor het aanbrengen van gerichte dubbelstrengsbreuken kan gebruik worden gemaakt van plaats-specifieke nucleases (PSN-en). Tot vijftien jaar geleden waren er alleen natuurlijk voorkomende PSN-en beschikbaar waarvan de DNA herkenningsplaats lastig was aan te passen. In de laatste twee decennia zijn echter verschillende artificiële PSN-en ontwikkeld waarvan de DNA herkenningsplaats vrij eenvoudig kan worden aangepast. Deze PSN-en zijn eiwit of eiwit/RNA-complexen die een specifieke DNA-sequentie herkennen en daar een DSB induceren.

Om deze PSN-en te kunnen gebruiken, moeten deze worden geïntroduceerd in doelwitcellen. In dit proefschrift hebben we getest of de bacterie *Agrobacterium tumefaciens* hiervoor als vector zou kunnen worden gebruikt. Vanwege het efficiënte mechanisme van DNA-overdracht is de bodembacterie *Agrobacterium* een veelgebruikte vector geworden voor het genetisch modificeren van planten. Tijdens de modificatie met *Agrobacterium*wordt het T-DNA afkomstig van het tumor-inducerende plasmide naar de kern van de gastheercel getransporteerd, waarna het wordt ingebouwd in het plantengenoom. De genen aanwezig op het T-DNA zijn niet essentieel voor de overdracht van het T-DNA naar de gastheercel en kunnen daarom worden vervangen door andere DNA-sequenties. Tegelijkertijd met het transport van het T-DNA worden ook enkele virulentie eiwitten overgebracht naar de gastheercel. Zowel het T-DNA als deze eiwitten worden via het Type IV secretie systeem (T4SS) van *Agrobacterium* overgebracht. Naast het transformeren van planten kan *Agrobacterium* onder laboratorium omstandigheden ook worden gebruikt voor de genetische modificatie van onder andere gisten en schimmels.

Het in dit proefschrift beschreven onderzoek richt zich op het gebruik van het CRISPR/Cas systeem voor mutagenese en op het gebruik van *Agrobacterium* als vector voor de overdracht van verschillende eiwitten.

In **Hoofdstuk 1** wordt een overzicht gepresenteerd van de PSN-en die beschikbaar zijn voor het aanbrengen van gerichte DSB-en en van de wijze waarop DSB-en worden gerepareerd. Daarnaast bevat dit hoofdstuk een beschrijving van enerzijds het mechanisme waarop *Agrobacterium* het T-DNA en virulentie eiwitten overbrengt en anderzijds van de biotechnologische toepassingen van A. tumefaciens.

Hoofdstuk 2 beschrijft de aanpassing van het CRISPR/Cas systeem voor gebruik in *Agrobacterium*. De DSB-en aangebracht met het CRISPR/Cas systeem leiden tot effectieve recombinatie van de flankerende "repeats" rond de negatieve selectie marker *sacB* die eerder is geïntegeerd via homologe recombinatie met een enkele cross-over. DBS-inductie op het stabiele RP4 plasmide resulteerde in het verlies van het RP4 plasmide. DSB-inductie op het octopine Ti plasmide van *Agrobacterium* leidde niet tot verlies van het Ti plasmide, maar wel tot een sterke afname in het aantal transformanten na expressie van het CRISPR/Cas systeem. Dat duidt erop dat DSB-inductie op het Ti plasmide en de daarop volgende degradatie lethaal is voor *Agrobacterium*.

Hoofdstuk 3 beschrijft hoe het Cas9 eiwit van het CRISPS/Cas systeem, gefuseerd met het translocatie signaal van het virulentie eiwit VirF, kon worden overgebracht van *Agrobacterium* naar *Saccharomyces cerevisiae*. Na overdracht zorgde het Cas9 eiwit, in combinatie met het in *S. cerevisiae* tot expressie gebrachte sgRNA, voor gerichte mutaties (inserties/deleties) in het CAN1 locus. Co-cultivatie van een *Agrobacterium* stam die zowel het sgRNA als het heterologe Cas9 eiwit tot expressie brengt, leidde niet tot mutaties op het CAN1 locus. Dit duidt erop dat het eiwit/RNA-complex van het Cas9 eiwit en sgRNA niet wordt getransporteerd door het T4SS.

Hoofdstuk 4 beschrijft de toepassing van het eerder ontwikkelde CRISPR/Cas systeem met het heterologe Cas9 eiwit dat door het T4SS systeem kan worden getransporteerd met als doel het aanbrengen van gerichte mutaties in *Nicotiana benthamiana*. Simultaan transport van het heterologe Cas9 eiwit met een T-DNA dat codeert voor het sgRNA gericht tegen het fytoeen desaturase gen, resulteerde in mutaties in het fytoeen desaturase gen. Amplicon sequencing liet zien dat mutaties die op deze wijze waren aangebracht, plaatsvonden met een lagere frequentie dan mutaties die waren aangebracht door de introductie van een T-DNA dat codeert voor zowel het Cas9 eiwit als het sgRNA.

Hoofdstuk 5 beschrijft tot slot de ontwikkeling van een systeem dat kan worden ingezet voor de selectie van transformanten door gebruik te maken van het isopentenyl transferase gen (*ipt*). De overdracht van het *ipt* gen naar een Pol-θ deficiënte *Arabidopsis thaliana* mutant resulteerde in transiënte expressie van het *ipt* gen en scheut formatie, maar resulteerde niet in integratie van het T-DNA. Ook direct transport van het IPT eiwit door het T4SS systeem van *Agrobacterium* naar *Arabidopsis* resulteerde in scheut formatie. Gecombineerd transport van het IPT eiwit samen met een T-DNA dat codeert voor het CRISPR/Cas systeem, dat DSB-en induceert in het polyphenol oxidase gen, resulteerde in regeneratie van scheuten met gerichte mutaties zonder te selecteren op de aanwezigheid van het T-DNA.

Chapter 7

Curriculum vitae

Curriculum vitae

Daan Schmitz was born on the 30st of March, 1987, in the city of Haarlem, the Netherlands. He grew up in Overveen and received his high school diploma from het Eerste Christelijk Lyceum located in the city of Haarlem in 2005. In September 2006, he started his BSc studies of Biology at Leiden University, receiving his BSc diploma in December 2009. In September 2009 he continued his scientific education by entering the 'Molecular and Cellular Biosciences' MSc program at Leiden University. After he received his MSc degree in August 2011, he started his PhD research in the group of Molecular and Developmental Genetics of the Institute Biology Leiden under the supervision of Dr. Sylvia de Pater and Prof. Dr. Paul J.J. Hooykaas, the results of which are described in this thesis. Daan's project was part of a research programme entitled "*Agrobacterium*-mediated protein therapy: a non-GMO approach for creating crop diversity" funded by ALW-NWO and TTI Green Genetics. Currently, Daan is employed as researcher at the department of 'Onderzoek, Informatie en Statistiek' of the municipality of Amsterdam.